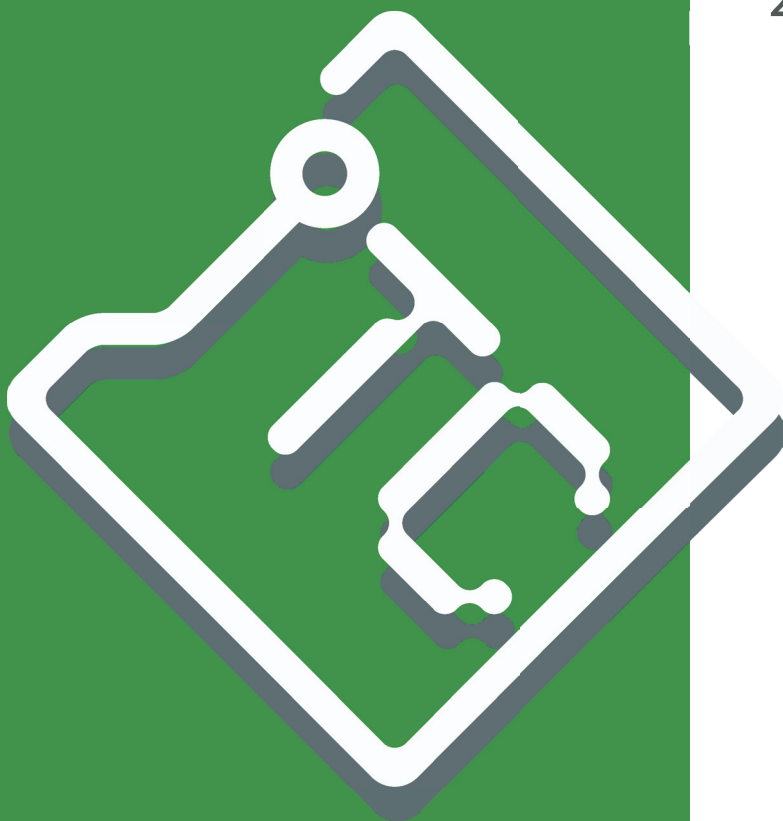


# 年報 第 11 号

2009 年度

Annual Report No.11

2009



東京大学情報基盤センター

Information Technology Center  
the University of Tokyo

# 目 次

巻頭言	1
<b>PART1 概 況</b>	<b>3</b>
組 織	5
組織図	5
職員数	5
教職員	6
平成 21 年度中の人事異動	11
東京大学情報基盤センター運営委員会委員名簿	13
予 算	14
収入・支出	14
外部資金	14
補助金等	16
平成 21 年度科学研究費採択状況	16
平成 21 年度受託研究費受入状況	18
平成 21 年度共同研究費受入状況	19
平成 21 年度政府系委託費受入状況	20
平成 21 年度政府系補助金受入状況	20
<b>PART2 センター活動報告</b>	<b>21</b>
学際大規模情報基盤共同利用・共同研究拠点	23
PKI プロジェクト	29
外部機関との共同研究（ローレンスバークレイ国立研究所）	31
CO <sub>2</sub> 排出量削減への取り組み	32
<b>PART3 各研究部門 研究活動報告</b>	<b>37</b>
情報メディア教育研究部門	39
情報メディア教育研究部門概要	41
柴山 悦哉	44
プログラミング言語に関する研究	
田中 哲朗	50
ゲームプログラミングに関する研究	
関谷 貴之	57
教育支援システムの開発と運用—講義データベース—	
丸山 一貴	63
Wiki とペンタブレットにおけるユーザインタフェースの研究	

<b>学術情報研究部門</b> .....	69
学術情報研究部門概要 .....	71
中川 裕志.....	78
統計的機械学習とそのテキスト処理への応用	
二宮 崇.....	84
自然言語処理のための基礎理論と応用—オンライン学習と英文読解支援—	
吉田 稔.....	92
Web 文書の分析・活用システムの研究—人名曖昧性解消システムの性能向上—	
清田 陽司.....	98
これからの電子図書館を支える知識ナビゲーション技術の研究	
清水 伸幸.....	104
知識発見と機械翻訳の自然言語処理要素技術—機械学習テンプレートライブラリー—	
<b>ネットワーク研究部門</b> .....	111
ネットワーク研究部門概要.....	113
若原 恭.....	118
ネットワークの機能の高度化と設計・運用・保守・管理の合理化—アドホックネットワーク、車車間・路車間通信、P2P ネットワーク—	
中山 雅哉.....	124
広域分散環境の高度基盤技術に関する研究	
小川 剛史.....	126
実空間における状況認識およびデータ送受信に関する研究	
関谷 勇司.....	132
連邦型クラウド構築のための技術研究ならびに広域分散サービス分析のための計測手法の研究	
中村 文隆.....	142
P2P ネットワーク計測 ネットワークシミュレーションの信頼性 音声処理言語	
加藤 朗.....	145
ネットワークの運用とセキュリティに関する研究	
<b>スーパーコンピューティング研究部門</b> .....	149
スーパーコンピューティング研究部門概要 .....	151
中島 研吾.....	165
大規模並列シミュレーション手法に関する研究、教育	
金田 康正.....	174
スーパーコンピューターに基づく大規模数値計算に関する研究	
佐藤 周行.....	191
言語処理系とサービス体系におけるセキュリティ保証の研究 Optimization Verifying Compiler, Document Carrying Authorization, and LoA of Services	
田浦 健次朗 .....	197
データ集約的な分散計算のための枠組み	
黒田 久泰.....	204
シームレス高生産・高性能プログラミング環境—高効率・高可搬性ライブラリの研究開発—	

松葉 浩也	207
クラスタからクラスタ外部へのネットワーク出力に関する研究	
大島 聡史	213
GPU を用いた高速数値計算と GPU プログラミング環境に関する研究	
掘 敦史	218
高効率かつ透過なファイルステージングシステム	
片桐 孝洋	226
ソフトウェア自動チューニングおよび高性能数値計算ライブラリの研究と HPC 教育	
吉廣 保	236
先端的大規模計算利用サービス	
渡辺 宙志	238
マルチスケールな現象の直接計算—大規模分子動力学法による気泡生成現象の研究—	
鴨志田 良和	244
大規模分散計算環境を有効活用するソフトウェアの研究	
藤田 肇	252
ブロードキャスト型単一 IP アドレスクラスタにおける柔軟なリクエスト分散機構	

## PART4 各部門 教育・サービス活動報告 259

<b>情報メディア教育部門</b>	261
情報メディア教育部門概要	263
教育用計算機システム運用報告	264
学内組織向けメールサーバ (MAILHOSTING) 運用報告	274
WEB PARK サービス運用報告	277
遠隔講義支援サービス運用報告	280
CFIVE 運用報告	285
<b>図書館電子化部門</b>	291
図書館電子化部門概要	293
デジタルコンテンツサービス	294
学術情報リテラシー支援	296
図書館関係システム運用・管理	304
<b>ネットワーク部門</b>	307
ネットワーク部門概要	309
東京大学情報ネットワークシステム (UTnet3) の運用管理	312
セキュリティ対応	322
東京大学情報システム緊急対応チーム (UT-CERT) との連携	325
学内ソフトウェアライセンス	327
<b>スーパーコンピューティング部門</b>	239
スーパーコンピューティング部門概要	331
スーパーコンピューティング業務	334
講習会	339
シンポジウム・研究会	342



研究プロジェクト .....	345
文部科学省 研究開発施設共用等促進費補助金（先端研究施設共用促進事業）	
先端的大規模計算利用サービス .....	353
HA8000 クラスタシステムの企業利用支援 .....	360
スーパーコンピュータ利用による研究成果報告（2009年） .....	362
<b>PART5 その他</b> .....	<b>373</b>
平成 21 年度委員会委員等 .....	375
平成 21 年度講習会・セミナー .....	378
平成 21 年度報道関係一覧 .....	381

## 巻 頭 言

東京大学情報基盤センター長 石川 裕

2009年度の東京大学情報基盤センター年報をお届けします。前任者の米澤明憲センター長、前々任者の岡部洋一センター長のリーダーシップの下、教員ならびに事務スタッフが建設的に取り組み、センターのアクティビティのさらなる活性化が図られてきました。2009年度は、国立大学法人化後の第1期中期目標・中期計画の最終年度にあたり、ここにお届けする年報は、第1期中期計画における最終報告となります。

当センターは、情報メディア教育部門、図書館電子化部門、ネットワーク部門、スーパーコンピューティング部門の4部門から構成される研究・教育・サービスの3つの役割を有する組織です。今までの年報では、各部門毎に、これら3つのミッションのご報告をしまいましたが、部門横断的な取り組みも活発となりました。そこで、今回の年報では、部門横断的取り組みの報告の章を設けました。これには、学際大規模情報基盤共同利用・共同研究拠点、PKIプロジェクト、外部機関との共同研究、CO<sub>2</sub>排出量削減への取り組み、などがあります。

2010年度から、第2期中期目標・中期計画が開始されます。また、当センターでは、学際大規模情報基盤共同利用・共同研究拠点として、2010年4月から正式に活動を開始しました。本年報が、単なる報告ではなく、我々のアクティビティの紹介を通して、新たなる共同利用・共同研究につながることを期待しております。皆さまからのご意見やご提案をお待ちしております。

最後に、本年報をまとめるにあたり、年報編集委員長の田中哲朗先生をはじめとする年報編集委員の皆様のご尽力に感謝申し上げます。

2010年5月吉日



PART 1

# 概 況

組 織

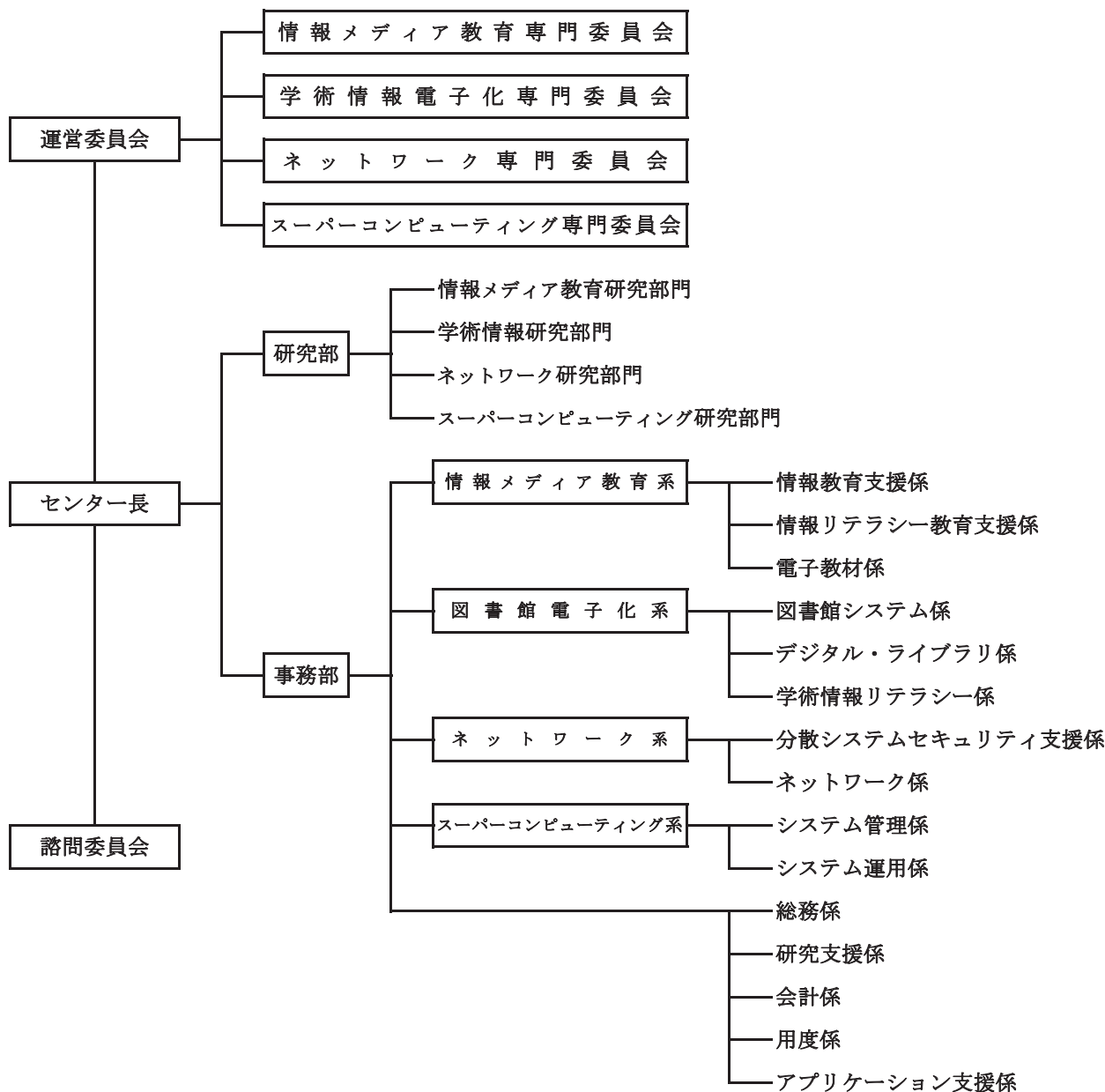
予 算

補助金等



# 組 織

組織図



職員数

H22. 3. 31

	専任	兼務	特任	合計
センター長		1		1
教授	5	1	2	8
准教授	3	2	3	8
講師	3		1	4
助教	7		3	10
研究員				0
事務職員	23			23
技術職員	26			26
事務補佐員	8			8
技術補佐員	4			4
派遣職員	4			4
合計	83	4	9	96

## 教職員

平成 22 年 3 月 31 日現在

情報基盤センター長 米 澤 明 憲  
(情報理工学系研究科教授)  
秘書(事務補佐員) 渡 邊 かがり

## 研究部

### 情報メディア教育研究部門

教 授	柴 山 悦 哉
准教授	田 中 哲 朗
助 教	関 谷 貴 之
助 教	丸 山 一 貴

### 学術情報研究部門

教 授	中 川 裕 志
講 師	二 宮 崇
助 教	吉 田 稔
助 教	清 田 陽 司
特任助教	清 水 伸 幸
事務補佐員	吉 富 美 樹
技術補佐員	松 岡 敦 美

### ネットワーク研究部門

教 授	若 原 恭
准教授	中 山 雅 哉
講 師	小 川 剛 史
講 師	関 谷 勇 司
助 教	中 村 文 隆
特任教授	加 藤 朗
事務補佐員	川 崎 しのぶ
事務補佐員	吉 澤 文 代

### スーパーコンピューティング研究部門

教 授	中 島 研 吾
教 授	金 田 康 正
教 授(兼務)	石 川 裕
准教授	佐 藤 周 行

准教授（兼務）	田 浦 健次朗
准教授（兼務）	黒 田 久 泰
助 教	松 葉 浩 也
助 教	大 島 聡 史
特任教授	堀 敦 史
特任准教授	片 桐 孝 洋
特任准教授	吉 廣 保
特任准教授	伊 藤 祥 司
特任講師	渡 辺 宙 志
特任助教	鴨志田 良 和
特任助教	藤 田 肇
事務補佐員	亀 田 文美代
事務補佐員	高 橋 弘 美

## PKI プロジェクト

事務補佐員	伊 東 雅 美
-------	---------

## 事務部

事務長	大 星 敏 明
副事務長	大 塚 浩 一
(総務担当)	
副事務長	早 野 裕 士
(業務担当)	
専門員	風 間 勉
(財務・研究協力担当)	
専門員	本 多 玄
主 査	西 澤 明 生
(研究支援担当)	
主 査	平 野 光 敏
(スーパーコンピューティング担当)	
専門職員	柏 芳 美
(eラーニング支援担当)	
専門職員	小 村 昇
(会計担当)	

## 総務係

係 長	米 山 浩
主 任	宮 地 亨



組 織

一般職員 柳 川 圭 介  
派遣職員 齊 藤 幸 子

研究支援係

係 長 上 杉 将 史  
主 任 草 開 泰 之

会計係

専門職員 小 村 昇  
(会計担当)  
主 任 木 崎 信 一

用度係

係 長 大 橋 公一郎  
主 任 岡 戸 彰 二  
主 任 山 口 拳 史  
派遣職員 栢 沙代子

アプリケーション支援係

係 長 岩 崎 直 樹  
技術職員 永 岡 陽 香

[情報メディア教育系] チーフ 増 田 均

専門職員 柏 芳 美  
(eラーニング支援担当)

情報教育支援係

係 長 増 田 均  
主 任 山 本 和 男  
技術職員 秋 田 英 範  
技術補佐員 宮 北 美 保  
派遣職員 後 閑 直 子  
派遣職員 五十嵐 崇

情報リテラシー教育支援係

係 長 松 岡 喜美代  
主 任 佐々木 馨  
主 任 田 川 善 教

## 電子教材係

係 長	有 賀	浩
主 任	坂 井	朱 美
技術職員	岩 藤	健 弘
技術補佐員	加 藤	康 一

## [図書館電子化系] チーフ 本 多 玄

## 図書館システム係

係 長 (兼)	本 多	玄
一般職員	小 松	陽 一
一般職員	田 口	忠 祐

## デジタル・ライブラリ係

係 長	吉 田	幸 苗
一般職員	小野澤	さわ子

## 学術情報リテラシー係

係 長	森 田	美由紀
一般職員	濱 田	智恵子
一般職員 (臨時)	笹 原	可 恵
事務補佐員	白 鳥	めぐみ

## [ネットワーク系] チーフ 早 野 裕 士

専門職員	中 山	仁 史
分散システムセキュリティ支援係		
係 長	椿 山	惣一郎
技術職員	小 藺	隆 弘
技術補佐員	阿 部	大

## ネットワーク係

係 長	下 田	哲 郎
技術職員	友 西	大
技術職員	松 本	浩 一

## [スーパーコンピューティング系] チーフ 平 野 光 敏

専門職員	丹 下	藤 夫
専門職員	佐 島	浩 之

組 織

システム管理係

係 長	宮 寄 洋
技術職員	佐 藤 孝 明
一般職員	諏 佐 賢 司

システム運用係

係 長	石 崎 勉
技術職員	安 部 達 巳

## 平成21年度中の人事異動

(転入・昇任・配置換等)

21. 4. 1	早野裕士	副事務長(業務担当) / 主査(情報基盤整備担当) から
21. 4. 1	本多 玄	専門員(図書館システム係長兼務) / 図書館情報係長から
21. 4. 1	西澤明生	研究支援担当主査(研究支援係長兼務) / 共同利用担当主査(共同利用係長兼務) から
21. 4. 1	柏 芳美	専門職員(eラーニング支援担当) / 電子教材係長から
21. 4. 1	有賀 浩	電子教材係長 / システム運用係長から
21. 4. 1	吉田幸苗	デジタル・ライブラリ係長 / 横浜国立大学図書館・情報部情報企画課情報企画係長から
21. 4. 1	下田哲郎	ネットワーク係長 / 国立情報学研究所学術基盤推進部学術ネットワーク課SINET運用チーム係長から
21. 4. 1	石崎 勉	システム運用係長 / 文部科学省研究振興局情報課研修生から
21. 4. 1	中山仁史	専門職員 / キャンパスネットワーク係長から
21. 4. 1	草開泰之	研究支援係主任 / 共同利用係主任から
21. 4. 1	三上 匠	用度係主任 / 用度係から
21. 4. 1	永岡陽香	アプリケーション支援係 / 新規
21. 4. 1	秋田英範	情報教育支援係 / 新規
21. 4. 1	小松陽一	図書館システム係 / 図書館情報係から
21. 4. 1	田口忠祐	図書館システム係 / 附属図書館総務課企画渉外係(文部科学省研究振興局情報課研修生) から
21. 4. 1	友西 大	ネットワーク係 / キャンパスネットワーク係から
21. 4. 1	松本浩一	ネットワーク係 / キャンパスネットワーク係から
21. 4. 1	安部達巳	システム運用係 / 情報リテラシー教育支援係から
21. 4. 1	白鳥めぐみ	学術情報リテラシー係事務補佐員 / 新規
21. 6. 1	風間 勉	専門員(財務・研究協力担当) / 東京国立博物館総務部経理課室長(経理・契約担当) から
21. 7. 1	西澤明生	研究支援担当主査 / 研究支援担当主査(研究支援係長兼務) から
21. 7. 1	小村 昇	専門職員(会計担当) / 薬学部会計チーム係長から
21. 7. 1	上杉将史	研究支援係長 / 国立科学博物館経営管理部財務課係長から
21. 7. 1	宮地 亨	総務係主任 / 工学系・情報理工学系等事務部総務グループ人事・給与チーム主任から
21. 7. 1	山口 挙史	用度係主任 / 附属病院管理課経理チーム主任から
21. 7. 1	柳川圭介	総務係 / 人事・労務系人事企画グループ人事業務推進チームから

組 織

21.10.1	大島 聡 史	スーパーコンピューティング研究部門助教／新規
21.10.1	笹原 可 恵	学術情報リテラシー係／新規
22.1.1	松岡 敦 美	学術情報研究部門技術補佐員／新規
22.2.1	伊藤 祥 司	スーパーコンピューティング研究部門特任准教授／新規 (任期 22.2.1～27.3.31)

(転出・退職等)

21.4.1	大日方 一 男	副事務長／情報系情報化推進グループ副課長へ
21.4.1	鴨 邦 宏	デジタル・ライブラリ係長／国立情報学研究所学術基盤推進部学術コンテンツ課コンテンツチーム係長へ
21.4.1	佐島 浩 之	システム運用係主任／専門職員（文部科学省研究振興局情報課研修生）へ
21.4.1	坂 牧 一 博	図書館情報係／地震研究所庶務チームへ
21.5.1	中 村 誠	ネットワーク研究部門助教／情報システム本部講師へ
21.6.30	吉 田 和 弘	学術情報研究部門特任研究員／退職
21.7.1	野 呂 清 隆	会計係長／調達・経理系調達グループ旅費チーム係長へ
21.7.1	灰 塚 毅 弘	総務係主任／教養学部等事務部総務課専門職員へ
21.7.1	入 江 健 司	用度係主任／東京農工大学府中地区会計チーム係長へ
21.7.1	三 上 匠	用度係主任／工学系・情報理工学系等事務部財務グループ調達チーム主任へ
21.7.1	佐々木 一 孝	総務係／人事・労務系人事企画グループ人事業務推進チームへ
21.7.1	赤 津 愛 美	デジタル・ライブラリ係／筑波大学附属図書館情報管理課雑誌受入係へ
22.2.28	麦 谷 真 弓	総務係事務補佐員／任期満了
22.3.31	米 澤 明 憲	センター長／任期満了
22.3.31	松 葉 浩 也	スーパーコンピューティング研究部門助教／退職
22.3.31	白 鳥 めぐみ	学術情報リテラシー係事務補佐員／任期満了

## 東京大学情報基盤センター運営委員会委員名簿

任期：平成21年4月1日～平成23年3月31日

平成21年4月1日現在

氏 名	所 属 ・ 職 名	適 用
米 澤 明 憲	情報基盤センター長	規則第3条第1号
金 田 康 正	情報基盤センター・教授	規則第3条第2号
中 川 裕 志	情報基盤センター・教授	規則第3条第2号
若 原 恭	情報基盤センター・教授	規則第3条第2号
柴 山 悦 哉	情報基盤センター・教授	規則第3条第2号
中 島 研 吾	情報基盤センター・教授	規則第3条第2号
太 田 勝 造	大学院法学政治学研究科・教授	規則第3条第3号
相 田 仁	大学院工学系研究科・教授	規則第3条第3号
堀 浩 一	大学院工学系研究科・教授	規則第3条第3号
近 山 隆	大学院工学系研究科・教授	規則第3条第3号
佐 藤 隆 夫	大学院人文社会系研究科・教授	規則第3条第3号
星 野 真 弘	大学院理学系研究科・教授	規則第3条第3号
新 井 富 雄	大学院経済学研究科・教授	規則第3条第3号
山 口 和 紀	大学院総合文化研究科・教授	規則第3条第3号
山 口 泰	大学院総合文化研究科・教授	規則第3条第3号
森 下 真 一	大学院新領域創成科学研究科・教授	規則第3条第3号
石 川 裕	大学院情報理工学系研究科・教授	規則第3条第3号
喜連川 優	生産技術研究所・教授	規則第3条第3号
横 山 伊 徳	史料編纂所・教授	規則第3条第3号
古 田 元 夫	附属図書館長	規則第3条第4号

オブザーバー 鶴岡 賀雄 総長補佐（大学院人文社会系研究科・教授）

## 予 算

## 収入・支出

## 平成21年度決算額

## 収入

区 分	決算額 (千円)
受託研究費等収入	148,894
自己収入	152,575
計	301,469

## 支出

区 分	決算額 (千円)
人 件 費	543,471
物 件 費	4,532,391
計	5,075,862

## 外部資金

## 1) 科学研究費補助金(平成21年度)

内 訳	受入件数	受入額 (千円)
情報メディア教育研究部門	1	15,400
学術情報研究部門	5	28,480
ネットワーク研究部門		
スーパーコンピューティング研究部門	7	18,880
計	13	62,760

## 2) 受託研究(平成21年度)

内 訳	受入件数	受入額 (千円)
情報メディア教育研究部門		
学術情報研究部門		
ネットワーク研究部門	2	7,210
スーパーコンピューティング研究部門	3	35,710
計	5	42,920

## 3) 共同研究(平成21年度)

内 訳	受入件数	受入額 (千円)
情報メディア教育研究部門		
学術情報研究部門	1	3, 420
ネットワーク研究部門	1	1, 540
スーパーコンピューティング研究部門	3	2, 315
計	5	7, 275

## 4) 政府系委託費(平成21年度)

内 訳	受入件数	受入額 (千円)
情報メディア教育研究部門		
学術情報研究部門		
ネットワーク研究部門		
スーパーコンピューティング研究部門	1	52, 037
計	1	52, 037

## 5) 政府系補助金(科学研究費補助金を除く。)(平成21年度)

内 訳	受入件数	受入額 (千円)
情報メディア教育研究部門		
学術情報研究部門	1	17, 762
ネットワーク研究部門		
スーパーコンピューティング研究部門	1	28, 900
計	2	46, 662



## 補助金等

### 平成 21 年度 科学研究費採択状況

研究代表者	教授 柴山 悦哉
研究種目	特定領域研究
研究期間	平成 18～22 年度
研究課題	情報爆発に対応する高度にスケーラブルでセキュアなソフトウェア構成・更新方式
研究費	15,400,000 円（平成 21 年度）
研究代表者	教授 中川 裕志
研究種目	特定領域研究
研究期間	平成 21～22 年度
研究課題	Web 上のテキスト情報の信頼性と有益性の評価システムに関する研究
研究費	5,000,000 円（平成 21 年度）
研究代表者	教授 中川 裕志
研究種目	基盤研究（A）
研究期間	平成 21～23 年度
研究課題	機械学習によるロングテール現象の解決方法に関する研究
研究費	17,550,000 円（平成 21 年度）
研究代表者	准教授 佐藤 周行
研究種目	基盤研究（B）
研究期間	平成 20～22 年度
研究課題	階層的なアルゴリズム選択機構を有する自動チューニング方式の研究
研究費	3,770,000 円（平成 21 年度）
研究代表者	特任准教授 片桐 孝洋
研究種目	基盤研究（B）
研究期間	平成 21～23 年度
研究課題	メニーコア・超並列時代に向けた自動チューニング記述言語の方式開発
研究費	9,880,000 円（平成 21 年度）

研究代表者	教授 金田 康正
研究種目	基盤研究 (C)
研究期間	平成 21～23 年度
研究課題	科学技術計算に有用な高速 4 倍長演算に関する研究
研究費	1,820,000 円 (平成 21 年度)
研究代表者	准教授 佐藤 周行
研究種目	挑戦的萌芽研究
研究期間	平成 20～21 年度
研究課題	ワークフローの認証と証明による PKI の応用拡大の研究
研究費	1,200,000 円 (平成 21 年度)
研究代表者	講師 二宮 崇
研究種目	若手研究 (A)
研究期間	平成 19～21 年度
研究課題	主辞駆動句構造文法のための統計同期文法による機械翻訳
研究費	3,640,000 円 (平成 21 年度)
研究代表者	助教 清田 陽司
研究種目	若手研究 (B)
研究期間	平成 20～21 年度
研究課題	情報探索支援を目的としたハイブリッド情報資源オントロジー
研究費	1,690,000 円 (平成 21 年度)
研究代表者	特任講師 渡辺 宙志
研究種目	若手研究 (B)
研究期間	平成 19～22 年度
研究課題	非ハミルトンダイナミクスにおける統計力学諸概念の力学からの基礎付け
研究費	1,040,000 円 (平成 21 年度)
研究代表者	日本学術振興会特別研究員 佐藤 一誠
研究種目	特別研究員奨励費
研究期間	平成 20～22 年度
研究課題	確率的生成モデルにおけるノンパラメトリックベイズ学習と自然言語処理への応用 (受入研究者: 中川裕志)
研究費	600,000 円 (平成 21 年度)
研究分担者	教授 中島 研吾 (研究代表者: 電気通信大学・今村俊幸准教授)

補助金等

研究種目	基盤研究 (B)
研究期間	平成 21 年度
研究課題	マルチコア複合環境を指向した適応型自動チューニング技術 (分担事項等: 適応型自動チューニング機能+前処理付反復法 ソルバの開発)
研究費	585,000 円 (平成 21 年度)
研究分担者	特任准教授 片桐 孝洋 (研究代表者: 電気通信大学・今村俊幸准教授)
研究種目	基盤研究 (B)
研究期間	平成 21 年度
研究課題	マルチコア複合環境を指向した適応型自動チューニング技術 (分担事項等: 固有地ソルバのチューニング・適応型チューニ ング方式の研究)
研究費	585,000 円 (平成 21 年度)

平成 21 年度 受託研究費受入状況

研究代表者	教授 石川 裕
相手機関名	(独) 科学技術振興機構
研究期間	平成 21 年度
研究課題	並列・分散型組込みシステムのためのディペンダブルシングルシ ステムイメージ OS
研究費	32,630,000 円 (平成 21 年度)
研究代表者	教授 中島 研吾
相手機関名	(独) 科学技術振興機構
研究期間	平成 21 年度
研究課題	ペタスケール大規模並列シミュレーション用ミドルウェア構築
研究費	2,080,000 円 (平成 21 年度)
研究分担者	准教授 中山 雅哉 (研究代表者: 本学工学系研究科・藤野陽三教授)
相手機関名	(独) 科学技術振興機構
研究期間	平成 21 年度
研究課題	都市基盤広域センシングに関する手法・データ伝送と処理・統 合リスク評価そしてフィールド適応に関する研究
研究費	4,810,000 円 (平成 21 年度)

研究代表者 准教授 佐藤 周行  
 相手機関名 大学共同利用機関法人情報・システム研究機構  
 研究期間 平成 21 年度  
 研究課題 最先端学術情報基盤の構築に関する研究開発と調査  
 研究費 1,000,000 円 (平成 21 年度)

研究代表者 講師 関谷 勇司 他  
 相手機関名 (独) 日本学術振興会 (フランス CNRS)  
 研究期間 平成 21 年度  
 研究課題 One Lab システムにおける要素技術と拡張に関する研究  
 研究費 2,400,000 円 (平成 21 年度)

## 平成 21 年度 共同研究費受入状況

研究代表者 教授 石川 裕  
 相手機関名 (独) 国立環境研究所地球環境研究センター  
 研究期間 平成 21 年度  
 研究課題 大量ジョブの効率的な処理方式に関する研究  
 研究費 1,600,000 円 (平成 21 年度)

研究代表者 教授 石川 裕  
 相手機関名 NEC ソフト (株)  
 研究期間 平成 21 年度  
 研究課題 大規模 HPC 向け ASP・SaaS サービスを適用したモデルについて  
 事業化に向けた実現可能性の研究  
 研究費 715,000 円 (平成 21 年度)

研究代表者 教授 中川 裕志  
 相手機関名 ソニー (株) システム技術研究所  
 研究期間 平成 21 年度  
 研究課題 自然言語処理による知識抽出技術の研究  
 研究費 3,420,000 円 (平成 21 年度)

研究代表者 助教 中村 文隆  
 相手機関名 日本電信電話 (株) サービスインテグレーション基盤  
 研究期間 平成 21 年度  
 研究課題 P2P ネットワークにおけるユーザ行動の研究  
 研究費 1,540,000 円 (平成 21 年度)

## 補助金等

研究代表者	教授 中島 研吾
相手機関名	大成建設（株）技術センター
研究期間	平成 21 年度
研究課題	地下多相流体挙動解析の並列計算手法に関する技術開発
研究費	0 円（平成 21 年度）

## 平成 21 年度 政府系委託費受入状況

研究代表者	教授 石川 裕
相手機関名	文部科学省
研究期間	平成 21 年度
研究課題	シームレス高生産・高性能プログラミング環境
研究費	52,037,000 円（平成 21 年度）

## 平成 21 年度 政府系補助金受入状況

研究代表者	教授 石川 裕
相手機関名	文部科学省
研究期間	平成 21 年度
研究課題	先端的大規模計算利用サービス
研究費	28,900,000 円（平成 21 年度）

研究分担者	教授 中川 裕志 (研究代表者：本学情報理工学系研究科・辻井潤一教授)
相手機関名	文部科学省
研究期間	平成 21 年度
研究課題	日中・中日言語処理技術の開発研究
研究費	17,762,334 円（平成 21 年度）

PART 2

# センター活動報告

学際大規模情報基盤共同利用・共同研究拠点

PKIプロジェクト

外部機関との共同研究

CO<sub>2</sub>排出量削減への取り組み



# 学際大規模情報基盤共同利用・共同研究拠点

## 研究支援係

### 1 共同利用・共同研究拠点 “ Joint Usage / Research Center ”

本節では、まず、文部科学省が説明している「共同利用・共同研究拠点の概要」を引用し、拠点制度の実施背景について説明する。次いで、拠点制度の性質と形態について説明し、平成 20、21 年度における認定状況について説明する。

『我が国の学術研究の発展には、個々の大学の枠を越えて大型の研究設備や大量の資料・データ等を全国の研究者が共同で利用したり、共同研究を行う「共同利用・共同研究」のシステムが大きく貢献してきました。』

共同利用・共同研究は、従来、国立大学の全国共同利用型の附置研究所等を中心に推進されてきましたが、我が国全体の学術研究の更なる発展のためには、国公私立大学を問わず大学の研究ポテンシャルを活用し、研究者が共同で研究を行う体制を整備することが重要です。

このため、文部科学省では、科学技術・学術審議会学術分科会研究環境基盤部会の報告を踏まえ、平成 20 年 7 月に、学校教育法施行規則を改正し、国公私立大学を通じたシステムとして、新たに文部科学大臣による共同利用・共同研究拠点の認定制度を設けました。

本制度の実施により、広範な研究分野にわたり、共同利用・共同研究拠点が形成されるなど、我が国の学術研究の基盤強化と新たな学術研究の展開が期待されます。』

これまでの国立大学の全国共同利用型附置研究所等は一分野につき一拠点とすることが原則であったが、今後国公私立大学を通じた共同利用・共同研究拠点では、研究分野によっては、一定の役割分担の下で複数の拠点を設けて相互に連携を図るなど、柔軟な形態による当該研究分野の特性や研究者コミュニティの議論を踏まえた最適な拠点構成の方向性が謳われている。このような拠点のタイプとしては、単一機関内に設置される「単独型」と、複数の拠点により構成される「ネットワーク型」とがある。後者は、「特定の国公私立大学の研究所等が中心となって、他の研究組織とネットワークを形成する形態」の拠点である。

これまでに認定された拠点数は、平成 20 年度が単独型 7 件、平成 21 年度が単独型 70 件、ネットワーク型 3 件である。

### 2 学際大規模情報基盤共同利用・共同研究拠点(情報基盤拠点)について

本章のタイトルでもある「学際大規模情報基盤共同利用・共同研究拠点」が正式名称であり、略称として「情報基盤拠点」が用いられる。英語表記の正式タイトルは、“Joint Usage / Research Center for Interdisciplinary Large-Scale Information Infrastructures”である。また、本拠点正式発足に向けては、当事業の特長を示す研究分野の英語名から“JHPCN : Japan High Performance Computing and Networking plus Large-scale Data Analyzing and Information Systems”というキーワードも用意した。

#### 2.1 本拠点の目的

本拠点は、北海道大学、東北大学、東京大学、東京工業大学、名古屋大学、京都大学、大阪大学、九州大学にそれぞれ附属するスーパーコンピュータを持つ8つの共同利用施設から構成され、



東京大学情報基盤センターがその中核機関として機能する「ネットワーク型」の共同利用・共同研究拠点である。

我が国の学術情報基盤全体の構築と次世代スーパーコンピュータとの連携という観点から、従来の大型計算機センターから発展してきた 7 つのスーパーコンピュータセンターと東京工業大学のそれが、緊密なネットワークを形成して共同利用・共同研究の拠点を構築することは、国家的観点から大変に意義深い。

本ネットワーク型拠点の目的は、超大規模計算機と大容量のストレージおよびネットワークなどの情報基盤を用いて、地球環境、エネルギー、物質材料、ゲノム情報、Web データ、学術情報、センサーネットワークからの時系列データ、映像データ、プログラム解析、その他情報処理一般の分野における、これまでに解決や解明が極めて困難とされてきた、いわゆるグランドチャレンジ的な問題について、学際的な共同利用・共同研究を実施することにより、我が国の学術・研究基盤の更なる高度化と恒常的な発展に資することにある。

## 2.2 拠点の全体計画の概要

本ネットワーク型拠点が共通的に対象とする研究分野は、地球環境、エネルギー、物質材料などの超大規模数値計算系応用分野、ゲノム情報や学術情報コンテンツに対する高度なアクセスや検索などの超大規模データ処理系応用分野、超大規模データを共有するため等の超大容量ネットワーク技術分野、および、これらの技術分野を統合した大規模情報システム関連研究分野を想定する。共同利用・共同研究拠点の重要事項は、8 大学内外の教員・研究者から構成され、中核機関に置かれる運営委員会において審議される。拠点が実施する共同研究の公募課題の設定や実施課題の採択等の重要事項は、運営委員会が審議する。

## 2.3 拠点の目指す役割

本ネットワーク型組織は、我が国の学際大規模情報基盤の共同利用・共同研究の拠点として、地球環境、エネルギー、物質材料、ゲノム情報、学術情報、プログラム解析等の分野を中心に、グランドチャレンジ的な問題に取り組む研究を、計算科学・工学および計算機科学に基づき、学際的に遂行・支援する中心的な役割を果たす。さらに、神戸に開設される次世代スーパーコンピュータ拠点に対して、本拠点は戦略性をもつ支援体制や共同研究体制を整備する役割を担う。また、次世代スーパーコンピュータが想定していない超大規模データ処理系応用分野においては、スーパーコンピュータを利用した学際共同研究を推進するという大きな役割を果たす。

## 2.4 拠点形成の必要性

超大規模数値計算系応用分野、超大規模データ処理系応用分野、超大容量ネットワーク技術分野、および、大規模情報システム関連研究分野における技術や知見は、科学技術の飛躍的発展に不可欠であり、我が国の持続的繁栄を支える基盤となる。8大学の施設から構成されるネットワーク型の拠点では、大規模計算資源を集中的に連携させ、かつ運営委員会を通して、これらの分野の共同利用・共同研究を戦略的に遂行していく。

また、ネットワーク型拠点が可能にする緊密な連携とシナジー効果により、それぞれの施設が実施可能な特徴的な研究課題が他の多くの施設にも共有されるだけでなく、他の施設が持つ研究ポテンシャルを援用することによって問題解決が容易になり、一つの施設では解決不能と思われていた問題を解決・解明する。シナジー効果は研究能力の側面にとどまらず、ネットワークを構成する各大学センターのもつ設備を、ネットワーク内のルールに基づき、極めて柔軟性の高い形で集約・結合し、互いの共同利用・共同研究の資することが可能になるので、この点においても従来解決が困難とされていた諸問題への対処が可能になる。

このような拠点形成の必要性は、関係する研究者コミュニティからも明言されており、本拠点の申請に当たっては、以下の学会から推薦をいただいた。

日本計算工学会、日本応用数理学会、情報処理学会、電子情報通信学会、人工知能学会、言語処理学会

さらに、拠点形成の必要性の要望について、拠点全体および拠点を構成する各大学に対しては、上記学会のほか、以下の研究者コミュニティからも要望書が寄せられた。

統計関連学会連合、 計算機学術利用北海道地区協議会、 航空宇宙学会、 日本機械学会 計算力学部門、 気象庁予報部数値予報課、 理化学研究所 情報基盤センター、 北陸先端大学 先端融合領域研究院、 産業技術総合研究所 計算科学部門、 情報・システム研究機構 国立情報学研究所 学術基盤推進部、 情報・システム研究機構 国立情報学研究所 リサーチグリッド研究開発センター、 宇宙航空研究開発機構 情報・計算工学センター、 京都高度技術研究所、 愛媛大学 総合情報メディアセンター、 日本バーチャリアリティー学会、 大型計算機利用大阪地区(第6地区)協議会、 日本機械学会 計算力学部門、 分子科学会、 西日本地区大学情報関連センター長会議、 財団法人九州先端科学技術研究所、 財団法人福岡県産業・科学技術振興財団

## 2.5 全国的な学術研究の発展への寄与

本拠点(8 大学)のもつ施設は、国内では唯一の巨大な計算基盤であり、これが上記の研究分野の大半の研究者によって共同利用・共同研究として利用されることになる。このため、本ネットワーク型拠点の共同利用・共同研究による成果は、わが国における上述グランドチャレンジ的研究の成果全体のなかで非常に大きな部分を占めることになり、国内のみならず海外からも注目されることが期待できる。また本拠点では、このような研究分野を中心にしながら、ネットワークを構成する各センターがもつ特徴を生かすことにより、研究分野に広がりを持たせ、さらにスーパーコンピュータの資源や研究者を量的に集積・確保することにより、質・量双方の観点から当該コミュニティの中心的存在となり、総じて本ネットワーク型拠点が全国的な学術研究の発展へ大きく貢献できると考えている。

## 3 平成 21 年度公募型共同研究(試行)について

平成 22 年 4 月からの本格的な活動開始に先駆け、試行的に公募型共同研究を実施した。

### 3.1 試行時の運用

試行は本運用に向けた準備と調整の意味合いも含めて実施されたが、負担金は徴収していない。

試行においても、研究課題募集要項、研究課題申込書と課題採択後に提出していただく誓約書を用意し、本運用に向けてのひな型を作成した。また、応募手続き、課題審査のプロセス自体も本運用の原型となった。試行における利用成果については、研究成果報告書を提出していただいている。

● 試行時のホームページ URL(平成 22 年度内には、正式運用の URL 配下へ移動の予定):

<http://www.itc.u-tokyo.ac.jp/kyoten/>

### 3.2 試行に関する主なスケジュール

平成 21 年 9 月 25 日	公募開始(試行用 Web 公開, 各種 ML にて案内)
平成 21 年 10 月 27 日	課題申込み〆切(当初 10/20 を延期)
平成 21 年 11 月 6 日	共同研究実施
～平成 22 年 3 月 15 日	
平成 22 年 5 月 14 日	成果報告書提出

### 3.3 試行運用の採択課題

課題審査のプロセスを経て、最終的に下記の 23 課題を採択した。試行では申請課題全て採択となった。

研究課題名	共同研究分野	研究課題責任者(所属)	共同研究大学
Wikipedia トピックを利用したブログ・ニュースからの知識マイニング	データ処理系	宇津呂武仁(筑波大・システム情報工学研究科)	東大
ハイパフォーマンス計算力学	数値計算系	牛島 省(京大・学術情報メディアセンター)	京大
クラスタ計算機での数値予報モデルの高速化に関する研究	数値計算系	佐藤清富(気象庁・予報部数値予報課)	東工大
言語間差異を活用した Web 情報資源へのアクセスシステムに関する研究	データ処理系	増田英孝(電機大・未来科学部)	東大
高エネルギー原子衝突現象のための大規模シミュレーション基盤	数値計算系	青木学聡(京大・工学研究科)	京大
1ミリ径を有する任意屈折率分布極大口径プラスチック光ファイバの有限要素法による全モード解析に関する研究	数値計算系	江口真史(千歳科学技術大・光システム学科)	北大
宇宙物理学的応用のための GPGPU 流体コードの開発	数値計算系	中村卓史(京大・理学部)	東工大
次世代ジオスペースシミュレーション拠点の構築	数値計算系	荻野瀧樹(名古屋大・太陽地球環境研究所)	東北大, 東大, 東工大, 名大, 京大, 阪大, 九大
GPU を用いた疎結合スピングラス模型に関する計算の加速化	数値計算系	高橋久尚(東工大・総合理工学研究科)	東工大
津波発生時の短周期波(周期: 数分～数 10 分)が係留船体運動へ及ぼす影響評価について	数値計算系	久保雅義(大島商船高専)	北大
等間隔直交格子法を基礎とした大規模並列流体解析ツールの開発	数値計算系	新井紀夫(農工大・工学府)	東北大
学術グリッド基盤の構築・運用技術に関する研究	情報システム	合田憲人(国立情報学研究所リサーチグリッド研究開発センター)	北大, 東北大, 東大, 東工大, 名大, 京大, 阪大, 九大
GPGPU の地震ハザード予測シミュレーションへの適応性評価	数値計算系	青井 真(防災研・防災システム研究センター)	東工大
超並列プラズマシミュレーションの研究	数値計算系	中島 浩(京大・学術情報メディアセンター)	京大
格子ガス法超並列計算手法の実用規模問題への適用可能性評価	数値計算系	松岡 浩(理研・次世代スパコン開発実施本部)	東北大, 阪大
大規模データベースレプリケーションの検証実験	データ処理系	小島 功(産総研・情報技術研究部門)	阪大
大規模並列計算における陰的時間積分法を使用した MHD 非線形コードの高速化	数値計算系	佐藤雅彦(核融合研・シミュレーション科学研究部)	東大
ベクトル型スーパーコンピュータ群のシームレスで効率的な運用技術に関する検討	情報システム	渡邊國彦(海洋研・地球シミュレータセンター)	東北大, 阪大

不確実性を扱う大規模有限要素解析のためのメタコンピューティング	数値計算系	奥田洋司(東大・人工物工学研究センター)	東大
超多自由度流動現象の大規模データ可視化解析システムの開発	数値計算系	石原 卓(名大・計算科学連携教育研究センター)	名大
データセントリック手法による人間行動処理の研究	データ処理系	間瀬健二(名大・情報科学研究科)	名大
アカデミッククラウド環境におけるソーシャルコンピューティングアーキテクチャの構築	情報システム	松尾啓志(名工大・情報基盤センター)	名大
グリッドデータファームによる大規模分散ストレージの構築と宇宙天気クラウド (OneSpaceNet) でのシミュレーション技術に関する研究	情報システム	村田健史(電磁波計測研究センター・宇宙環境計測グループ)	東北大, 名大, 阪大

## 4 拠点発足に向けての活動

拠点申請に向けての取り組みは、全国共同利用情報基盤センター長懇談会をとおり平成 20 年度から行なわれてきた。平成 21 年 6 月 25 日付けで学際大規模情報基盤共同利用・共同研究拠点(ネットワーク型)が認定されて以後、本拠点正式発足に向けて以下の活動が行なわれた。尚、正式な拠点発足は平成 22 年 4 月 1 日である。

### 4.1 センター長懇談会

本拠点の正式発足前の状況では、上述のセンター長懇談会にて、8拠点のセンター長および帯同の教員と事務系職員により、試行と拠点発足に向けた準備が進められた。また、一連の懇談会は、文部科学省研究振興局情報課学術基盤整備室職員も同席の上、開催された。

平成 21 年 7 月 2 日(於:東京大学), および、平成 21 年 9 月 4 日(於:東京大学)

主な議題: 試行に向けた課題募集要項と課題申込書の内容について検討、  
拠点の組織及び運営に関する協定書について検討、  
拠点運営委員会規則と課題審査委員会規則について検討

平成 21 年 11 月 27 日(於:メルパルク名古屋)

主な議題: 本拠点の体制(協定書、規則、委員名簿)検討、試行の状況報告

平成 22 年 2 月 16 日(於:東京大学)

主な議題: 拠点運営委員会規則と運営委員会名簿について承認、  
課題審査委員会規則(案)と委員会名簿について検討、  
本運用に向けた課題募集要項と課題申込書の内容について検討

### 4.2 各種規則等の整備

- ・「東京大学情報基盤センター規則」(東京大学役員会議決、東大規則第 119 号)にて、共同利用・共同研究拠点の形成と中核拠点としての役割について制定し、拠点運営委員会についても制定された。
- ・「学際大規模情報基盤共同利用・共同研究拠点運営委員会規則」(平成 22 年 1 月 21 日東京大学役員会議決、東大規則第 64 号)にて、拠点運営委員会の組織及び運営について制定され、共同研究課題審査委員会について制定された。
- ・「学際大規模情報基盤共同利用・共同研究拠点の組織及び運営に関する協定書」(平成 22 年 1 月 28 日東京大学役員会承認、2 月 16 日発行)にて、構成 8 拠点による構成拠点の基本事項について協定締結された。

上記に加え、平成 22 年度の本拠点発足に向けて、下記の各種規則(案)等も整備した。

- ・「学際大規模情報基盤共同利用・共同研究拠点共同研究課題審査委員会規則(案)」



- ・「平成 22 年度 学際大規模情報基盤共同利用・共同研究拠点 公募型共同研究 課題募集要項(予告)」
- ・「平成 22 年度 学際大規模情報基盤共同利用・共同研究拠点 公募型共同研究 課題申込書(案)」

### 4.3 拠点発足に関する取り組み

主な準備活動は「センター長懇談会」での検討と公募型共同研究(試行)の実施であるが、その他、以下の活動を実施した。

平成 22 年 3 月 5 日

文部科学省「革新的ハイパフォーマンス・コンピューティング・インフラ(HPCI)フォーラム」にて構成8拠点(および、筑波大学、JAEA、JAXA、JAMSTEC)のポスター出展協力  
主催:文部科学省  
会場:文部科学省 第二講堂

平成 22 年 3 月 9 日

本拠点の正式ホームページを公開の上、平成 22 年度 課題募集(予告)も公開し、各種マーケティングリスト等によるアナウンスも開始された。

- 正式ホームページ URL:

<http://jhpcn-kyoten.itc.u-tokyo.ac.jp/>

平成 22 年 3 月 12 日

情報処理学会創立 50 周年記念(第 72 回)全国大会ポストイベント  
「計算科学技術と次世代スーパーコンピューティング基盤」  
～革新的ハイパフォーマンス・コンピューティング・インフラの構築～  
にて構成8拠点(および、筑波大学、JAEA、JAXA、JAMSTEC)のポスター出展協力  
主催:(社)情報処理学会  
共催:文部科学省, (独)理化学研究所  
会場:東京大学 安田講堂

平成 22 年 3 月末日までに

「共同利用・共同研究拠点 平成 22 年度実施計画書」を構成8拠点各々から文部科学省へ提出

# PKI プロジェクト

佐藤周行

## プロジェクト概要

PKI プロジェクトは、CA（認証局）を安定的に運用するための運用技術を、発足以来重要項目として検討してきたが、いくつかの試行錯誤を経て、UT-CA としてキャンパス PKI のための CA を実際に構築し、運用技術を蓄積するために、インソースでの運用でのプロトタイプを構築した。2010 年 3 月現在、UT-CA の実験協力部局として情報学環、情報基盤センター・本部合同を得て実験開始以来 3 年 5 ヶ月の間安定して運用できた。このことで当プロジェクトが提唱している分散 RA の運用モデルの有効性が明らかになった。

NII が主導している CSI UPKI プロジェクトが開始した「オープンドメイン証明書発行実験」に参加し、パブリックサーバ証明書発行のための学内の体制を構築した。ここにおいても分散 RA 方式を取っている。2010 年 3 月末で参加部局は 8 部局を数え、サーバ証明書を発行した部局は 21、運用開始からの審査枚数は 314 枚に及ぶ。さらに EV 証明書の運用を開始した。

さらに、NII が主導する認証フェデレーションへに参加すべく、認証環境の整備に努めている。

## 運用報告

### 1 UT-CA

[http://www.pki.itc.u-tokyo.ac.jp/UT\\_CA-top.html](http://www.pki.itc.u-tokyo.ac.jp/UT_CA-top.html)

現在では、UT-CA の実験協力部局として情報学環、情報基盤センター・本部合同を得て実験開始以来 3 年 5 ヶ月の間安定して運用することができた。このことで当プロジェクトが提唱している分散 RA の運用モデルの有効性が明らかになった。

### 2 サーバ証明書登録局運用

<http://www.pki.itc.u-tokyo.ac.jp/serpj/>

NII が主導している CSI UPKI プロジェクトが開始した「オープンドメイン証明書発行実験」に参加し、パブリックサーバ証明書発行のための学内の体制を構築した。ここにおいても分散 RA 方式を取っている。2010 年 3 月末で参加部局は 8 部局を数え、サーバ証明書を発行した部局は 21 部署、314 件に及ぶ。

東大は、学内体制として TRA を立ち上げ、CP/CPS を公開した上で可能ならば発行にともなう審査を部局に委譲するモデルを取っている。これは従来からあった UTnet が安定して運用されていることに大きく依存するモデルであった。この運用モデルは成功し、参加部局の広がりが見られるようになった。

## EV 証明書の運用開始

2009 年末より、EV 証明書の発行権利を購入し、学内にサービスすることとなった。TRA は申請を代行する立場にあるが、申請に当たって、従来の審査を援用することで信頼を得るようにした。

## サービス統計

2009 年 4 月–2010 年 3 月

参加部局	発行枚数
TRA 直接	37
情報基盤センター	27
新領域創成科学研究科	2
史料編纂所	8
本部事務局	20
生産技術研究所	19
理学系研究科	15
医学部附属病院	7
数理科学研究科	5

## 3 SSL-VPN 試行サービス

図書館のもつ電子ジャーナルを ECC の運用管理する ID 体系でサービスしている。理工系にとどまらず、広い分野でアクセスがなされている。2008 年度からはサービス可能な電子ジャーナルの数が大幅に増え、主だったところがサービス対象になり、利用が増大している。運用に対して、PKI プロジェクトは主にセキュリティの観点から関与している。

## 4 認証フェデレーション

Shibboleth を利用した SSO への対応を検討中である。

## 5 CSI UPKI プロジェクトへの係わり

NII が主導している CSI UPKI プロジェクトが開始した「オープンドメイン証明書発行実験」に参加し、パブリックサーバ証明書発行のために構築した学内の体制を維持した。

さらに NII の主導する認証フェデレーションにテストサイトで参加している。本参加に際しての条件を検討中である。

## 外部機関との共同研究(ローレンスバークレイ国立研究所)

東京大学情報基盤センターとローレンスバークレイ国立研究所(Lawrence Berkeley National Laboratory<sup>1</sup>)は2009年9月4日にハイパフォーマンスコンピューティング(HPC)及び計算科学・工学分野における基礎的、学術的な共同研究の促進を目的とした研究交流協定覚書(Memorandum of Understanding)を取り交わしました<sup>2</sup>。ローレンスバークレイ国立研究所はアメリカ合衆国エネルギー省(U.S. Department of Energy)付属の国立研究所であり、カリフォルニア大学(University of California)がその運営にあたっています。特に原子核物理学の分野では大きな業績をあげており、過去に11名のノーベル賞受賞者(物理学賞, 化学賞)を輩出しています。スーパーコンピュータを利用した計算科学分野(非軍事目的)においては世界のトップクラスの研究が行われており、計算機科学, 科学・工学, 応用数学各分野の専門家の緊密な協力のもとに研究開発が進められています。

最先端の計算機システムは大規模化, アーキテクチャの複雑化により, その性能を引き出すことは一層困難となっています。当情報基盤センターでは次世代, 次次世代スーパーコンピュータに備えて, 安定で性能の高いアプリケーションの開発支援環境の整備が急務と認識しており, アプリケーション, 数値アルゴリズム, 計算機システムの研究者が一体となった研究開発を進めています。ローレンスバークレイ国立研究所とは同様の価値観を共有しており, 個別の分野では研究協力をこれまで実施してきました。本交流協定によって, 以下に示すような多分野にわたった体系的な研究協力が可能となり, 大規模計算向けの新規アルゴリズムの開発, 新しい計算機システムの提案につながるとともに, 両センターの有する計算機資源の有効活用にも資するものと期待されます。



図1：協定覚書に署名する  
米澤明憲センター長（当時）

- 科学技術計算のための数値計算ライブラリ, 自動チューニング技術
- マルチフィジックスシミュレーションのためのカプラー
- 科学技術計算ベンチマーク
- ペタ/エクサスケール計算機向け並列プログラミングモデル, 実行環境

<sup>1</sup> <http://www.lbl.gov/>

<sup>2</sup> <http://www.itc.u-tokyo.ac.jp/news/n090924/>



## CO<sub>2</sub> 排出量削減への取り組み

### 1 概要

2010年4月1日より、東京都では全国に先駆けて「CO<sub>2</sub>削減義務化」がスタートした。対象は、電気などのエネルギー消費量が原油換算で年間1500kl以上の事業所である。本学においても、サステイナブルキャンパスプロジェクト室<sup>1</sup>(TSCP)を中心に全学的な取り組みが行われており、2010年3月9日付けで総長より各部局に対して、サステイナブル・キャンパス実現に向けた節電・節エネルギーの実行についての要請が行われている。

TSCPによる具体的な取り組みは以下の通りである：

アクションプラン	取り組み期間	目標(2006年度比)
TSCP-2012	2008年度～2012年度	2012年度末には、非実験系のCO <sub>2</sub> 排出量を15%削減
TSCP-2030	2013年度～2030年度	2030年度末には、CO <sub>2</sub> 排出総量を50%削減

情報基盤センターは、定格消費電力の総量が3MWを超えるスーパーコンピュータシステムを始めとして、大量のエネルギーを消費する部局であり、TSCPを始めとする大学当局との緊密な連携のもと、長期的見通しに基づいた組織的取り組みが必要との認識で、CO<sub>2</sub>削減に向けた取り組みを概略以下の通り推進した。

- ・ スーパーコンピュータ

スーパーコンピュータシステムに関する技術の研究開発動向を、特に消費電力量の観点から調査し、その結果に基づき、利用者ニーズに基づく性能と消費電力量の最適なバランスを達成できるよう次期調達システムの仕様策定を進めた。

- ・ その他

非実験系の消費電力量を削減するため、空調設備の設定温度の調整や不在時の照明の消灯等、主に費用対効果が高くかつ比較的取り組みやすいソフト面での施策を盛り込んだ実施計画を策定し、これに基づいて取り組みに着手した。

<sup>1</sup> <http://www.tscp.u-tokyo.ac.jp/>

## 2 スーパーコンピュータに関する取り組み

### 2.1 背景

スーパーコンピュータの大規模化とともに、電力消費量が増加し、電力料金という経済的な問題の他、CO<sub>2</sub> 排出量の増加という環境への影響も生じている。特にスーパーコンピュータの最も重要な用途の一つが温暖化など気候予測シミュレーションであることもあり、世界的に業界全体として消費電力量削減の取り組みが始まりつつある。

スーパーコンピュータの性能を比較する「TOP 500 List<sup>2</sup>」というサイトがあるが、2007 年からは消費電力あたりの性能を比較する「Green 500 List<sup>3</sup>」というカテゴリーも登場している。

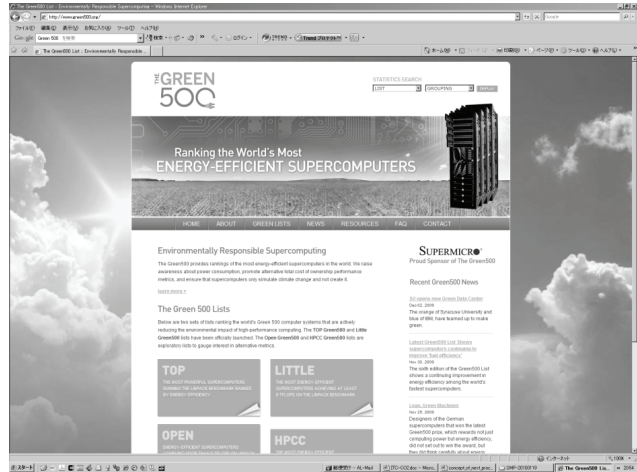


図 1 : Green500 Top Page  
(<http://www.green500.org>)

### 2.2 取り組みの内容

現在、情報基盤センターには表 1 に示す 2 つのスーパーコンピュータシステムを有している。いずれも稼働率は 90% を超えており、多くの利用者によって様々な研究開発用途に使用されている。

このうち、ベクトル並列型スーパーコンピュータシステム(Hitachi SR11000/J2)は 2010 年度末にリプレース時期を迎える。次期マシンは柏キャンパスに 2011 年 3 月末竣工予定の新研究棟に設置され、2011 年度中に稼働開始の予定であり、調達に向けて詳細仕様の策定を進めた。この策定では、性能もさることながら、環境に配慮したエネルギー効率の高いシステムを導入することを重要と考えており、利用者のニーズも踏まえてバランスの良いシステムを導入する方針とした。

消費電力あたりの性能を高めるための技術革新は、スーパーコンピュータを構成するプロセッサ、メモリ、ネットワーク、ファイルシステムなど様々な分野で日々進められおり、できるだけ最新のシステムを導入することが、消費電力を削減し、環境に配慮することにつながる。特に、情報基盤センターのスーパーコンピュータのリプレース周期は 6 年であるため、あらゆる観点から技術開発の動向、上記のバランスを考慮して、最適な解を得ることが重要と考えている。

現在の Hitachi SR11000/J2 のシステム全体の消費電力(冷却設備除く定格消費電力より算出)あたりの性能は 26MFLOPS/W (MFLOPS: Million (10<sup>6</sup>) of Floating-Point Operations per Second, 1 秒間に 10<sup>6</sup> 回の浮動小数点演算)であるが、次期システムでは最高では 500MFLOPS/W を超えるもの

<sup>2</sup> <http://www.top500.org/>

<sup>3</sup> <http://www.green500.org/>

と予想されている。2018年頃に登場すると言われているピーク性能 1ExaFLOPS ( $10^{18}$ FLOPS) のシステムを定格消費電力 20MW 程度で運用するためには、この効率を 100 倍、すなわち 50,000MFLOPS/W 程度まで高める必要がある。そのために様々な技術開発が行われている。

スーパーコンピュータの大規模化によって、電力消費量だけでなく設置面積も増加している。この影響をできるだけ抑えるために、様々な部品の小型化及び高集積化が進められている。高集積化によって、単位面積あたりの電力消費量(発熱量)が著しく増加するため、最新のスーパーコンピュータシステムは効率の良い水冷方式を採用する場合がほとんどであり、水冷に依存するシステムがしばらくは主流となるものと考えられる。前節で述べたような 100 倍の効率化が達成できれば、空冷が主流となる可能性も否定はできないが、少なくとも 2011 年、2014 年に予定されているリプレースにおいて、限られた面積で利用者のニーズに応えるためには、水冷方式によるシステムが中心となる。柏キャンパスで建造が開始されつつある新研究棟もそのような傾向を考慮して設計を実施した。

ピーク性能 10PFLOPS ( $10^{16}$ FLOPS) クラスの次世代スーパーコンピュータ(神戸)は 2012 年 6 月のフル稼働を目指している。2011 年度中に稼働する、東大情報基盤センターの次期システムは、学際大規模情報基盤共同利用・共同研究拠点を構成する 8 大学などととも、文部科学省の提唱する「革新的ハイパフォーマンス・コンピューティング・インフラ(HPCI)」の一翼を担うものと期待され、次世代スーパーコンピュータの性能等を勘案して、ピーク性能 1PFLOPS ( $10^{15}$ FLOPS) 以上が要求されている。500MFLOPS/W とすると、電力消費量(計算機本体のみ)は 2MW となり、実際はスーパーコンピュータシステム全体として 3MW 程度になる可能性もある。一時的にはこのように全体の電力消費量が増加する場合もあるが、2014 年以降はむしろ全体の処理性能は増加しても効率改善により、電力消費量は減少する可能性が強い。したがって、情報基盤センターにおいては TSCP-2012 の達成は困難であるが、長期的に考えた TSCP-2030 には十分に貢献できると考えられる(2006 年の日単位の実効消費電力(合計値)の平均値は 1.517MW である)。

10 年単位の様々な技術動向を考慮して、スーパーコンピュータシステムの選定、建物も含めた運用環境の整備を実施していく必要がある。スーパーコンピュータは様々な分野の技術が結集した複雑なシステムであり、学内の多様な分野の専門家によって、消費電力を削減し、環境にやさしいシステムを実現するための継続的な検討を進めていくことは大きな意義がある。

表 1:情報基盤センター現有スーパーコンピュータシステム

	ベクトル並列型スーパー コンピュータシステム	超並列型スーパー コンピュータシステム	
型番	Hitachi SR11000/J2	Hitachi HA8000 クラスタシ ステム (T2K オープンスパ コン (東大))	
Architecture	Power 5+	Quad-core AMD Opteron	
システム定格消費電力 (うち計算機本体)	1.345 MW (0.716 MW)	1.734 MW (1.0512 MW)	
日単位平均実効消費電力 (うち計算機本体)	0.988 MW (0.462 MW)	0.792 MW (0.534 MW)	
日単位ピーク実効消費電力 (うち計算機本体)	1.135 MW (0.515 MW)	1.001 MW (0.675 MW)	
冷却方式	水冷+空冷	空冷	
ノ ー ド	コア数	16	16
	ピーク性能 (GFLOPS)	147.2	147.2
	メモリ容量 (GB)	128 (コア当たり 8GB)	32 (一部 128)
	BYTE/FLOP (物理性能)	0.924	0.272
全 体	ノード数	128	952
	コア数	2,048	15,232
	ピーク性能 (TFLOPS)	18.84	140.13
	メモリ容量 (TB)	16.38	32.00
稼動開始時期	2005年3月 (2007年アップグレード)	2008年6月	
リプレース時期	2011年4月以降	2014年4月以降	

### 3 その他の取り組み

情報基盤センターでは、CO<sub>2</sub> 排出量削減を効果的に進めるため TSCP 推進委員会を設置し、当面消費電力量の削減を目指して積極的に取り組んでいく方針を決定した。特に、本学の TSCP 室のアクションプランと歩調を合わせることにしたが、加えてセンター独自の施策も考案して、それらを総合した具体的な施策を表1の通り整理し、可能なものから実施を開始した。更に、2010 年度には、これらの施策による効果を定量的に評価できる仕組みの導入についても検討し実施に移していく計画とした。

表1 消費電力量削減に向けた実施施策  
\*財務面の理由から、実施は 2010 年度の予定

照明スイッチへの「節電ラベル」貼り付けによる利用者の節電意識向上
廊下・通路など共用部分における蛍光灯の点灯間引き
利用頻度が低い廊下等の蛍光灯の原則消灯
洗浄便座の保温温度「低」設定への変更
空調設備の冬季設定温度 20℃（夏季設定温度 28℃）の徹底
室不在時における照明消灯の徹底
パソコン・プリンター等 OA 機器のスリープ設定の徹底
パソコン・プリンター等 OA 機器の長時間不使用時及び帰宅時の電源遮断
洗浄便座の非使用時の便座蓋閉じの徹底
両面印刷や 2 アップ機能による印刷用紙枚数の適正化・不要なカラー印刷の削減
女子トイレ擬音装置設置による節水
冷蔵庫の省エネ型機種への更新*
天吊型空調機へのファンの取り付けによる空調効率の向上*
廊下・トイレ等の照明器具の人感式スイッチへの取替え*
窓への遮蔽フィルムの貼り付けによる空調効率の向上*
蛍光灯上部への反射フィルム貼り付けによる照明効率の向上*

## PART 3

各研究部門

# 研究活動報告

情報メディア教育研究部門

学術情報研究部門

ネットワーク研究部門

スーパーコンピューティング研究部門



# 情報メディア教育研究部門

## 情報メディア教育研究部門概要

柴山 悦哉

プログラミング言語に関する研究

田中 哲朗

ゲームプログラミングに関する研究

関谷 貴之

教育支援システムの開発と運用

—講義データベース—

丸山 一貴

Wikiとペンタブレットにおけるユーザインタフェースの研究





# 情報メディア教育研究部門

## 概要

部門長 柴山悦哉

情報メディア教育研究部門は、情報メディア教育部門の研究部門であり、今年度は、教授1名、准教授1名、助教2名が在籍して以下にあげるような研究を行なった。

**実運用システムの設計・改善等を直接の目的とした研究:** 情報メディア教育部門が業務として運用している教育用計算機システムなどは、教育機関が管理運用するエンドユーザ向けの情報基盤としては、我が国の中でも有数の規模や複雑度を有する。そのため、既存のノウハウだけに頼っているだけでは、安定的かつ効率的な運用は不可能であり、研究として解決すべき課題も多い。また、この規模と複雑さのシステムの実運用から得られた知見やノウハウには、他の組織にとって有用なものも多く、これらを系統的な形にまとめて公開することが研究活動として意味を持つこともある。今年度は、情報メディア教育部門が運用する一部のサブシステムに関係した次のような研究成果の発表を行なった。これらは、業務部門を研究部門がサポートした成果である。

- 学習管理システム CFIVE の運用 (関谷)
- 学内向けメールサービスの運用 (丸山、関谷)

**情報システムの構成や開発に関連した研究:** 情報メディア教育部門が業務として運用しているシステムを直接的に対象とするものではなく、より一般的に、情報基盤を開発・運用するための技術に関連した研究も行なわれている。今年度は以下のような研究成果の発表を行なった。

- 動的言語処理系の高速化 (柴山)
- イベント駆動オブジェクト指向プログラムの理解支援 (柴山)

**情報メディアの使い勝手の向上に関連した研究:** 情報メディア教育部門では、3万人を超えるエンドユーザを対象に情報基盤を提供している。システムのユーザビリティの向上や情報メディアに容易にアクセスするための支援は重要な課題であり、これに関連した技術の研究開発も行なわれている。今年度は次のような研究成果の発表を行なった。

- 講義データベース (関谷)
- RDB との連携による Wiki データ入力の強化 (丸山)
- ペンタブレットの空中動作による3次元データ入力 (丸山)
- 手書きを用いた動画上の非同期コミュニケーション (丸山)

**その他の研究:** 今年度は、コンピュータゲームに関する次のような研究開発成果を発表/公開した。

- オープンソース将棋プログラムの作成 (田中)
- Retrograde Analysis によるボードゲームの完全解析 (田中)

このうちコンピュータ将棋に関しては、他研究室などと共同で開発している GPS 将棋が、第19回世界コンピュータ選手権および第5回コンピュータ将棋世界最強決定戦において優勝し、さらに人間のトップアマとの公開対局でも圧勝した。

## 情報メディア教育研究部門 成果要覧

### 招待講演／招待論文

[招待 1] 田中哲朗: 査読者はどこを見る? - 査読を通るゲーム研究論文を書くために, 第 14 回ゲームプログラミングワークショップ, 2009 年 11 月

### 受賞関連

[受賞 1] 金子知適, 田中哲朗, 山口和紀, 川合慧: 平成 20 年度論文賞, 情報処理学会, 2009 年 5 月

### 著書／編集

[著書 1] 笈捷彦, 山口利恵, 近山隆, 石畑清, 松崎公紀, 柴山悦哉, 田中哲朗: 目指せ! プログラミング世界 — 大学対抗プログラミングコンテスト ICPC への挑戦 —, 近代科学社, 2009.

### 査読付論文

[査読付 1] Salikh Zakirov, Shigeru Chiba, Etsuya Shibayama: How to Select Superinstructions for Ruby, *IPSJ Transactions on Programming*, Vol. 3, No. 2, pp. 1-8, March 2010.

[査読付 2] Izuru Kume, Etsuya Shibayama: Feature Interactions in Object-Oriented Effect Systems from a Viewpoint of Program Comprehension, *Proceedings of 10th International Conference on Feature Interactions in Telecommunications and Software Systems*, pp. 123-130, June 2009.

[査読付 3] Izuru Kume, Etsuya Shibayama: A Conceptual Model for Comprehension of Object-oriented Interactive Systems, *Proceedings of International Conference on Software Engineering and Knowledge Engineering*, pp. 376-381, July 2009.

[査読付 4] T. Sekiya, Y. Matsuda, and K. Yamaguchi: Analysis of Curriculum Structure Based on LDA, ICEIT'09, Proceedings of International Conference on Education and Information Technology, pp.561-566, San Francisco, CA, USA, 2009.

[査読付 5] 山本 三雄, 関谷 貴之, 山口 和紀: 情報教育における理論と技術に関する研究の調査. 情報処理学会論文誌 (ジャーナル), Vol.50, No.10, pp. 2449 - 2461, 2009.

[査読付 6] 梅林靖弘, 丸山一貴, 寺田実: ペンタブレットの空中動作を利用したインタラクション, 第 17 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2009), pp.121-122, 2009.

### その他の発表論文

[発表 1] 宮地充子, 前田俊行, 柴山悦哉, 鶴飼裕司, 大久保隆夫, 岡村久道: パネル討論「日本の情報基盤を安全にする - セキュリティ標準化の動向と今後」, 情報処理学会ソフトウェア工学・コンピュータセキュリティ・組込みシステム・情報セキュリティ心理学とトラスト研究グループ合同研究発表会, インターネットイニシアティブ本社, 2009 年 5 月.

[発表 2] Salikh Zakirov, Shigeru Chiba, Etsuya Shibayama: How to Select Superinstructions for Ruby, 情報処理学会第 76 回プログラミング研究会, 2009 年 10 月.

[発表 3] 田中哲朗: 「どうぶつしょうぎ」の完全解析, 情報処理学会ゲーム情報学研究会, Vol.2009-GI-22, No.3, pp. 1-8, 岩手, 2009 年 6 月.

- [発表 4] 岩藤 健弘, 関谷 貴之, 杉山 剛士, たかはし ゆういち: 学習管理システム CFIVE の運用, 平成 21 年度情報教育研究集会講演論文集, E2-4, 2009.
- [発表 5] 秋田 英範, 丸山 一貴, 関谷 貴之, 佐々木 馨, 増田 均: 学内向けメールサービスの運用と稼働状況について, 平成 21 年度情報教育研究集会講演論文集, pp.437-440, 2010.
- [発表 6] 平山慧, 丸山一貴, 寺田実: 手書きを用いた動画上の非同期コミュニケーション, 情報処理学会全国大会, 5ZF-8, 2010.
- [発表 7] 奥村俊也, 丸山一貴, 寺田実: 表データ操作を RDB で強化した Wiki システム, 情報処理学会全国大会, 5ZF-5, 2010.
- [発表 8] 佐藤和哉, 丸山一貴, 寺田実: CodeMusician: プログラム可聴化システム, 情報処理学会研究報告, Vol.2009-MUS-81, No.22, pp.4, 2009.
- [発表 9] 奥村俊也, 寺田実, 丸山一貴: 表データ操作を RDB で強化した Wiki システム, 情報処理学会研究報告, Vol.2009-HCI-134, No.10, 2009.
- [発表 10] 井桁正人, 寺田実, 丸山一貴: ScoutView: Web ページにおけるナビゲーション支援インタフェース, 情報処理学会研究報告, Vol.2009-HCI-133, No.8, 2009.

## 公開ソフトウェア

- [公開 1] 田中哲朗, 金子知適, 森脇大悟, 副田俊介, 林芳樹, 竹内聖悟: osl-for-csa (バージョン 0.6), コンピュータ将棋協会コンピュータ将棋選手権使用可能ライブラリ, 2010 年 1 月.
- [公開 2] 田中哲朗, 金子知適, 森脇大悟, 副田俊介, 林芳樹, 竹内聖悟: gpsshogi-for-csa (バージョン 0.6), コンピュータ将棋協会コンピュータ将棋選手権使用可能ライブラリ, 2010 年 1 月.

## 特記事項

- [特記 1] 柴山悦哉: ソフトウェアセキュリティ研究チームの活動概要, RCIS ワークショップ 2009, 秋葉原 UDX ビル UDX シアター, 2009 年 5 月.
- [特記 2] 喜連川優, 柴山悦哉, 伊藤実, 朴泰祐: 平成 20 年度論文賞の受賞論文紹介: 編集にあたって, 情報処理, Vol. 50, No. 7, p. 672, 2009.
- [特記 3] GPS 将棋: 第 19 回世界コンピュータ将棋選手権優勝, コンピュータ将棋協会, 2009 年 5 月.
- [特記 4] GPS 将棋: 第 5 回世界コンピュータ将棋最強決定戦優勝, 北陸先端科学技術大学院大学, 2010 年 1 月.

## 報道関連

- [報道 1] 読売新聞 2009 年 5 月 6 日朝刊: 「GPS 将棋」が初優勝.
- [報道 2] 週刊将棋 2009 年 5 月 13 日号: 革命的結果勢力図一変 - GPS 将棋が初優勝.
- [報道 3] 朝日新聞 2009 年 5 月 13 日夕刊: 世界コンピュータ将棋選手権 - 「GPS 将棋」が初優勝.
- [報道 4] NHK BS2、囲碁将棋ジャーナル 2009 年 5 月 16 日: 世界コンピュータ将棋選手権速報.
- [報道 5] NHK BS2、囲碁将棋ジャーナル 2009 年 5 月 30 日: 世界コンピュータ将棋選手権特集.
- [報道 6] 読売新聞 2009 年 6 月 2 日夕刊: 「GPS」が乱戦制し初優勝.
- [報道 7] 産経新聞 2009 年 7 月 26 日: 訪れるか将棋界「X デー」.

# プログラミング言語に関する研究

柴山 悦哉

## 1 概要

プログラミング言語は、プログラムを記述するために必須の道具であり、記述したプログラムの生産性・保守性、高い処理効率、汎用性などが求められる。しかし、これらの要求をすべて同時に高いレベルで満たすのは困難である。時代によって、また応用分野によっても異なる妥当なトレードオフポイントに対応し、さまざまなタイプのプログラミング言語に対する需要が生じている。

本年度は、以下の2種類のトレードオフについて、それぞれを改善するための研究を行った。

- プログラミング言語の表現力・柔軟性と言語処理系の実行性能のトレードオフ
- プログラミング言語の表現力と作成されるプログラムの理解可能性のトレードオフ

最初のものとして、コンパクトなプログラム記述と実行時の動的な機能の修正・追加を可能とする Ruby 言語を対象に、処理系の高速化に関する研究を行なった。二番目のものとして、Java 言語で記述されたオブジェクト指向プログラムの了解性を向上させるための支援機構に関する研究を行った。

## 2 動的言語処理系の高速化

### 2.1 背景

ソフトウェアに求められる機能は年々増加傾向にある。また、一度構築したソフトウェアに対し、一部の機能を修正したり、新しい機能を追加したりする作業を、迅速に行いたいという要求も年々強まっている。24時間365日稼働し続けるタイプのアプリケーションでは、稼働中のアプリケーションを止めずに機能を変更・追加したいという要請さえある。

このような要求に応えるための一つの考え方として、Ruby や Python のようなスクリプト言語を用いる場合がある。これらの言語は、C 言語などの従来の主流言語に比べ、以下のような特長を持つ。

- プログラムをコンパクトに記述するための言語機構が豊富
- プログラムの実行時に動的に機能を追加・修正することが可能

しかし、その一方で、スクリプト言語の実行性能はお世辞にも高いとは言えない。Ruby でプログラムを作成すると、C 言語を使った場合より、プログラミングに要する時間は短く、完成したプログラムのコードサイズも小さくなるが、実行速度は2桁くらい遅くなるのが相場である。

プログラム言語の表現力と柔軟性を確保するために、実行性能をある程度犠牲にするのは仕方がない。しかし、表現力や柔軟性を犠牲にしない範囲で、実効性能を高める努力は行うべきである。そこで、代表的な動的言語である Ruby を対象に、言語処理系の高速化に関する研究を行った。



表 1: 分岐予測ミスのコスト

	mandelbrot	nbody	partial_sum	spectral_norm
全実行サイクル ( $10^9$ cycles)	60	79	73	69
分岐予測ミスによるストール ( $10^9$ cycles)	1.6	1.3	1.0	0.4
分岐予測ミスのオーバーヘッド	2.6%	1.7%	1.4%	0.6%
全実行機械語命令 ( $10^9$ )	102	129	110	116
全実行仮想命令 ( $10^9$ )	0.9	1.0	0.8	1.8
間接分岐命令 ( $10^9$ )	1.7	2.1	1.8	2.3
予測ミス間接分岐命令 ( $10^9$ )	0.29	0.35	0.28	0.12
予測ミス率	17.0%	16.4%	15.5%	5.1%

## 2.2 内容

### 2.2.1 命令融合に関する従来知見

動的な言語の処理系は、伝統的にインタプリタをベースとしたものが多い。通常は、ソースプログラムを仮想命令（抽象命令と呼ぶこともある）列にコンパイルしてから、これらの命令をインタプリティブに実行する。Ruby 言語の処理系もこのような構成となっている。

インタプリタの高速化に関する研究は長年に渡って行われており、さまざまな手法が提案されている。その中の一つに命令融合という手法がある。これは、連続する複数の仮想命令を、一つの複合的な仮想命令に置換する手法である。

命令融合による高速化の有効性が高いのは、各仮想命令の本体を実行するコストに比べ、インタプリタが命令をディスパッチするコストが十分に大きい場合である。命令のディスパッチには、通常、間接分岐などの（分岐予測ミスが発生すると）遅延が大きくなる処理が必要であり、四則演算などの単純な処理より、どうしても多くの時間がかかりやすい。Ertl と Gregg による 2003 年頃の研究では、命令ディスパッチに要する時間が全実行時間の最大半分程度になると報告されている。

### 2.2.2 Ruby 処理系での分岐予測ミスの影響

まず、前節で述べた過去の知見がどれくらい正しいかを確認するために、Ruby 1.9 の処理系を対象に、命令ディスパッチのオーバーヘッドを測定した。命令ディスパッチに要する時間のうち大きな割合を占めるのは、分岐予測ミスにともないストールする時間と考えられる。そこで、CPU のサイクルカウンタを使ってストールしたサイクル数を計測した。Intel Core 2 Duo E8500 を用いた実験の結果を表 1 に記す。実験に用いた 4 種類のプログラム (mandelbrot, nbody, partial\_sum, spectral\_norm) は Ruby Benchmark Suite に含まれる数値計算中心のプログラムである。

結論から言うと、Ruby 1.9 の処理系を Core 2 Duo で実行した場合、分岐予測ミスによるオーバーヘッドはあまり小さくなく 0.6%–2.6% であった。分岐予測のミス率も 5.1%–17.0% と比較的 low、branch target buffer を使った単純な予測を行う以前の CPU に比べ、最近の CPU では予測精度が向上していることが伺える。

### 2.2.3 ナイーブな命令融合

前節の測定結果より、命令融合による命令ディスパッチ回数の削減効果はあまり期待できないことがほぼ明らかである。実験により確かめたところ、命令融合をナイーブに適用するだけでは、性能向上は高々 4% 程度にとどまった。この実験結果は、実行トレースから、連続して実行される頻度の高い 2 命令の対を抽出し、それらの命令を融合命令で置き換えて再度実行を行って得たものである。実験に用いた Ruby 1.9 の仮想機械では、各命令が C 言語で実装されている。命令ディスパッチの回数削

---

```

1 DEFINE_INSN opt_plus
2 ()
3 (VALUE a, VALUE b) /* input argumets */
4 (VALUE val)        /* return value */
5 {
6     ...
7     if (HEAP_CLASS_OF(a) == rb_cFloat &&
8         HEAP_CLASS_OF(b) == rb_cFloat &&
9         BASIC_OP_UNREDEFINED_P(BOP_PLUS)) {
10        val = DBL2NUM(RFLOAT_VALUE(a) + RFLOAT_VALUE(b));
11    } else {
12        PUSH(a);
13        PUSH(b);
14        CALL_SIMPLE_METHOD(1, idPLUS, a);
15    }
16 }

```

---

図 1: opt\_plus 命令の概要

減以外に、命令を融合することで、C 言語コンパイラによる最適化が進む可能性はある。しかし、実験結果を見る限り、その効果も限定的である。

#### 2.2.4 より効果的な命令融合

命令融合による命令ディスパッチ回数の削減効果はあまり大きくないことがわかった。しかし、命令融合の効果がこれに限定されるわけではない。表 1 で取り上げたような数値計算を行う Ruby プログラムの場合、メモリ管理オーバーヘッドの削減効果が大きい。Ruby 言語処理系の内部では、浮動小数点数をボックス化されオブジェクトとして扱う。そのため、四則演算等の結果として、新しい値が生成されるたびにヒープ上にオブジェクトが生成される。数値計算のプログラムでは、寿命の短いオブジェクトが大量に生成され、メモリ管理のオーバーヘッドが大きくなる。表 1 のベンチマークと同じ条件で測定すると、生成されるオブジェクトの 98.6%–99.99% が浮動小数点数オブジェクトであった。また、GC に要する時間が総実行時間の 2 割から 5 割程度を占めた。

図 1 は、加算を行う仮想命令 opt\_plus の概略を示したものである。Ruby は動的に型が定まる言語であるため、まず引数の型を調べ、その結果に応じて適切な処理 (e.g. 整数の加算、浮動小数点数の加算等) を行う。たとえば、7 行目の if 文により、二つの引数 (a, b) が浮動小数点数であるかどうかをチェックされ、そうである場合には 10 行目の処理が行われる。10 行目では、RFLOAT\_VALUE() によりボックス化された浮動小数点数 a, b の値を取り出し、加算を行ったのち、DBL2NUM() によりボックス化された浮動小数点数を新たに生成する。機械語レベルで浮動小数点数の加算を 1 回実行する時間より、オブジェクトの生成に要する時間の方が長くなる。

数値計算を行うプログラムでは、積和演算などのために、複数の算術演算命令が連続して実行されることが多い。そのような命令を融合し、さらに意味を考慮した最適化を行うとボックス化のオーバーヘッドを大幅に削減することができる。たとえば、積和演算の場合であれば、

```

tmp = DBL2NUM(RFLOAT_VALUE(b) * RFLOAT_VALUE(c));
...
val = DBL2NUM(RFLOAT_VALUE(a) + RFLOAT_VALUE(tmp));

```

を

```

...
val = DBL2NUM(RFLOAT_VALUE(a) + RFLOAT_VALUE(b) * RFLOAT_VALUE(c));

```

表 2: 算術命令の融合

	mandelbrot	nbody	partial_sum	spectral_norm
全実行サイクル ( $10^9$ cycles)	64	64	64	65
分岐予測ミスによるストール ( $10^9$ cycles)	1.4	0.8	0.9	0.3
分岐予測ミスのオーバーヘッド	2.2%	1.3%	1.3%	0.5%
全実行機械語命令 ( $10^9$ )	106	100	92	111
全実行仮想命令 ( $10^9$ )	0.9	0.9	0.7	1.6
間接分岐命令 ( $10^9$ )	1.8	1.7	1.6	2.2
予測ミス間接分岐命令 ( $10^9$ )	0.3	0.23	0.26	0.12
予測ミス率	17.2%	13.3%	17.0%	5.4%

と変換することで、オブジェクトの生成回数が半減し、GC のオーバーヘッドの大幅な削減が期待できる。

表 2 は、融合した算術命令の内部でこのような最適化を行い、表 1 と同じ実験をした結果である。mandelbrot 以外の三つのプログラムで、最大 20% 程度の性能向上がみられた。mandelbrot で性能向上がみられなかったのは、最内ループの内側に連続する算術命令が出現しなかったためである。プログラムを眺めると、

```
tr = zr zr - zizi + cr
```

のような有望な箇所は存在するが、これが、

```
getdynamic zr zr , 3
getdynamic zizi , 3
opt_minus
getdynamic cr , 3
opt_plus
setdynamic tr , 0
```

のように、二つの算術命令 (opt\_minus, opt\_plus) の間に変数アクセスする命令 (getdynamic) が割り込む形となったため、上で述べたような単純な命令融合では対応できなかった。

この問題は、ソースコードを

```
tr = cr + (zr zr - zizi)
```

と書き直すだけで回避でき、これで約 20% の性能向上が確認できた。他の方法としては、以下のようなものが考えられる。

- バイトコードコンパイラのコード生成部で、命令融合を意識した命令生成を行う。
- 連続する 3 命令以上の融合を行う。

## 2.3 具体的成果

Ruby 言語の仮想機械を現在主流の CPU で動かした場合、どのような処理にどれくらいの時間がかかっているかを測定し、インタプリタの最適化手法として従来知られていた命令融合によるディスクコストの削減があまり有効でないことを実験的に確かめた。また、数値計算を行うプログラムなどでは命令融合によるメモリ管理や型チェックのコスト削減に一定の有効性があることを示した。この結果を述べた論文が論文誌に掲載された [査読付 1]。



### 3 イベント駆動オブジェクト指向プログラムの理解支援

#### 3.1 背景

今日のソフトウェアの多くは、オブジェクト指向の考え方にに基づき設計・構築され、イベント駆動で動作している。このようなソフトウェア構成法は、部品間の結びつきを疎とし、部品の独立性を高める目的でよく使われる。こうすることで、部品の交換が容易となるだけでなく、部品間の構造的な関係もわかりやすくなる。

一方、既に存在するプログラムを修正するために、その（構造ではなく）挙動を理解する必要に迫られることがよくある。イベント駆動型のプログラムでは、イベントを介した制御の流れを追うのが必ずしも容易ではないため、挙動の理解に苦しむことがある。従来の構造化された命令型のプログラムであれば、部品としての柔軟性には欠ける反面、プログラムの字面と制御構造が直接的に対応していたため、制御の流れは追いやすかった。すなわち、制御の依存関係が静的に決まれば、柔軟性は低下するが、挙動はつかみやすい。今日主流のソフトウェアは、構造の了解性と柔軟性を得るために、挙動の了解性の一部を犠牲にしたと考えることもできる。

そこで、GUI プログラムを対象に、プログラム部品間の実行時の依存関係を解析し、プログラム理解を促進するためのモデルとそのモデルに基づくツールに関する研究を進めている。

#### 3.2 内容

今年度はイベント駆動プログラムの理解モデルに関する基礎的な研究を主に行った。このモデルは、以下の要素から構成される。

- プログラムの実行時トレースに対応する抽象度のトレースモデル
- トレースモデルを（プログラム理解のために）抽象化するための枠組み
- トレースモデルを抽象化するための操作群

トレースモデルは、初期イベントによって開始された一連の処理を表現するモデルである。ここでは詳細は省略するが、以下のようなトレース要素とトレース要素間の制御フローおよびデータフローを介した依存関係が表現される。

- メソッドおよび基本ブロックの実行
- 値やオブジェクトの生成
- 変数の値の更新
- 変数の参照、分岐など

プログラム理解を行うにあたって特に問題となるのは、ある一連の処理  $t_1$  により更新された変数の値を別の一連の処理  $t_2$  が参照するケースである。 $t_1$  が  $t_2$  に影響を及ぼす、あるいは  $t_2$  が  $t_1$  に依存すると言っても良いだろう。トレースモデルは、このような依存関係を原理的には表現できる。しかし、人間が理解しやすいような抽象化が施されていない。

そこで、トレースモデルの要素を、意味のあるまとまりに分割する抽象化の枠組みを定義する。トレースモデルではその要素間に依存関係を自然に定義することができるが、抽象化を施したのちの抽象モデルでは、意味的にまとまった要素のクラス間にその依存関係が拡張される。すなわち、 $e, e'$  をトレースモデルの要素、 $E, E'$  をそれぞれ  $e, e'$  を含む最小の抽象モデルの要素、 $\rightarrow, \Rightarrow$  をそれぞれトレースモデルと抽象モデルの依存関係とすると、以下の関係がなりたつ。

$$E \Rightarrow E' \text{ if } e \rightarrow e'$$

トレースモデルはプログラム実行トレースからほぼ自動的に生成可能である。一方、抽象モデルの方にはクラスタリングの仕方によって任意性があり、トレースモデルをどのように抽象化するか（トレースモデルのどの要素を同じクラスタに含めるか）を定めないと決定することができない。抽象化のためのクラスタリング方法を定義し、さらにそうやって定められた抽象モデルに対して、クエリを発行することで、プログラム理解は進められる。このようなクラスタリングとクエリのための暫定版プリミティブを今年度は定めた。

### 3.3 具体的成果

イベント駆動プログラムの意味理解のためのモデルを構築し、また実例に基づき、その有効性を議論した。そして、その結果について国際会議の査読付き論文で発表した [査読付 2, 査読付 3]。

## 4 成果要覧

### 著書／編集

[著書 1] 筧捷彦, 山口利恵, 近山隆, 石畑清, 松崎公紀, 柴山悦哉, 田中哲朗: 目指せ! プログラミング世界 — 大学対抗プログラミングコンテスト ICPC への挑戦 —, 近代科学社, 2009.

### 査読付論文

[査読付 1] Salikh Zakirov, Shigeru Chiba, Etsuya Shibayama: How to Select Superinstructions for Ruby, *IPSJ Transactions on Programming*, Vol. 3, No. 2, pp. 1-8, March 2010.

[査読付 2] Izuru Kume, Etsuya Shibayama: Feature Interactions in Object-Oriented Effect Systems from a Viewpoint of Program Comprehension, *Proceedings of 10th International Conference on Feature Interactions in Telecommunications and Software Systems*, pp. 123-130, June 2009.

[査読付 3] Izuru Kume, Etsuya Shibayama: A Conceptual Model for Comprehension of Object-oriented Interactive Systems, *Proceedings of International Conference on Software Engineering and Knowledge Engineering*, pp. 376-381, July 2009.

### その他の発表論文

[発表 1] 宮地充子, 前田俊行, 柴山悦哉, 鶴飼裕司, 大久保隆夫, 岡村久道: パネル討論「日本の情報基盤を安全にする - セキュリティ標準化の動向と今後」, 情報処理学会ソフトウェア工学・コンピュータセキュリティ・組込みシステム・情報セキュリティ心理学とトラスト研究グループ合同研究発表会, インターネットイニシアティブ本社, 2009年5月.

[発表 2] Salikh Zakirov, Shigeru Chiba, Etsuya Shibayama: How to Select Superinstructions for Ruby, 情報処理学会第76回プログラミング研究会, 2009年10月.

### 特記事項

[特記 1] 柴山悦哉: ソフトウェアセキュリティ研究チームの活動概要, RCIS ワークショップ 2009, 秋葉原 UDX ビル UDX シアター, 2009年5月.

[特記 2] 喜連川優, 柴山悦哉, 伊藤実, 朴泰祐: 平成 20 年度論文賞の受賞論文紹介: 編集にあたって, 情報処理, Vol. 50, No. 7, p. 672, 2009.

# ゲームプログラミングに関する研究

田中哲朗

## 1 概要

ゲームプログラミング研究の一環として、2003年から他研究室と共同でオープンソース将棋プログラムの開発を行っている。さまざまな改良を加えた結果、2009年5月に開催された第19回世界コンピュータ選手権、2010年1月に開催された第5回コンピュータ将棋世界最強決定戦で優勝することができた。また、11月におこなわれたトップアマとの公開対局に圧勝し、人間のトップに近づいていることを示すことができた。また、強いプログラムを作成するために行った評価関数の学習に関する研究をまとめた論文(共著)が高い評価を受けて情報処理学会の論文賞を受賞した。

また、「どうぶつしょうぎ」という新たに考案されたボードゲームを完全解析し、すべての状態の勝ち負け、勝つために必要な最大手数、勝ちを証明するための最小証明木の大きさなどの有用なデータベースを作成した。

## 2 オープンソース将棋プログラムの作成

### 2.1 背景

1997年に、チェスプログラムが初めて人間の世界チャンピオンを破るという事件が起きたが、将棋もあと数年で同様のレベルに達することが期待され、コンピュータ将棋に関する研究は一番面白い時期を迎えようとしている。

しかし、コンピュータ将棋の研究を始める際に超えなくてはならない障壁は小さくない。まず、将棋は駒の種類がチェスより多く、独自の持ち駒ルールや、打ち歩詰め、二歩等のルールなどのため、ルール通り指させるだけでも容易ではない。

また、ある程度の強さを持つプログラム上で評価しないと、意味のある結果が得られないが、終盤の詰め将棋など十分な強さを持つものをバグなしに作るのは難しい。

そのため、コンピュータ将棋研究の全体的なレベルアップを狙って、最初から公開する前提で将棋プログラムの開発を2003年からおこなっている。将棋プログラム名はGPS将棋で、開発メンバーは田中研究室の学生の他に、総合文化研究科の教員、学生、会社員などが加わっている。

### 2.2 内容

この将棋プログラム自体は2003年度から開発を開始し公開もしているが、2009年度は

- より多くの特徴に対応した高速な評価関数の学習アルゴリズム
- クラスタ並列化

などをテーマに改良をおこなった。

評価関数の学習に関しては、情報処理学会論文賞を受賞した[受賞1]を発展させて、多次元の特徴を持つ評価関数を高速に学習させるアルゴリズムを開発した。

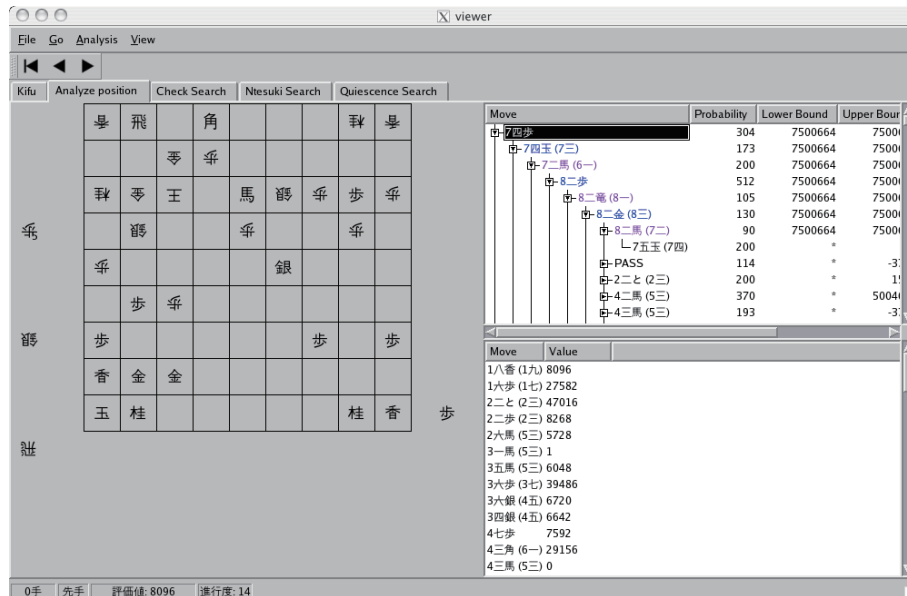


図 1: GPS 将棋の操作画面

## 2.3 具体的成果

ライブラリはBSDライセンスで公開し、将棋プログラム全体はGPLで公開する形にした。これにより、純粹に研究目的の人も、あるいは商品売って儲けようという人もどちらも自由に使えるようにしながら、「コンピュータ将棋選手権何位」といった売り文句は使えないという形で公開することができた。

作成したプログラムGPS将棋(図1)は、2009年5月に開催された第19回世界コンピュータ選手権、2010年1月に開催された第5回コンピュータ将棋世界最強決定戦で優勝することができた。また、11月におこなわれたトップアマとの公開対局に圧勝し、人間のトップに近づいていることを示すことができた。

## 3 Retrograde Analysis によるボードゲームの完全解析

### 3.1 背景

ゲームプログラミング研究において、特定のゲームに関する議論と、多くのゲームに通用する普遍的な議論を切り分けるのは難しい。また、ゲームの状態空間の大きさも、統計的な議論をするのには不適切なほど小さすぎるか、大容量化したディスク装置でも記録しきれないほど大きすぎるかのどちらかであり、共通の土俵で議論するのが困難であるという問題があった。

そんな中で、2008年12月に日本女子プロ将棋協会から「どうぶつしょうぎ」というボードゲームが発売され、ヒット作となった。「どうぶつしょうぎ」はゲーム名から分かるようにルールは将棋に類似しているが、児童への普及を主目的としているため以下のような簡潔なルールになっている。

**道具** 図2左のような3×4のボードを使う。各列は左からA,B,C、行は上から1,2,3,4と示され、各マスは[A1],[C4]のように示される。4行目、1行目はそれぞれのプレイヤー(先手、後手)の陣地と呼ぶ。駒は以下の4種類それぞれ2個の計8個使う。

**ライオン** 将棋の玉と同じ動きで8近傍に動ける

**ぞう** 斜めの4近傍に動ける。将棋の角と違って隣のマスにしか動けない。

**きりん** 上下左右の4近傍に動ける。将棋の飛車と違って隣のマスにしか動けない。

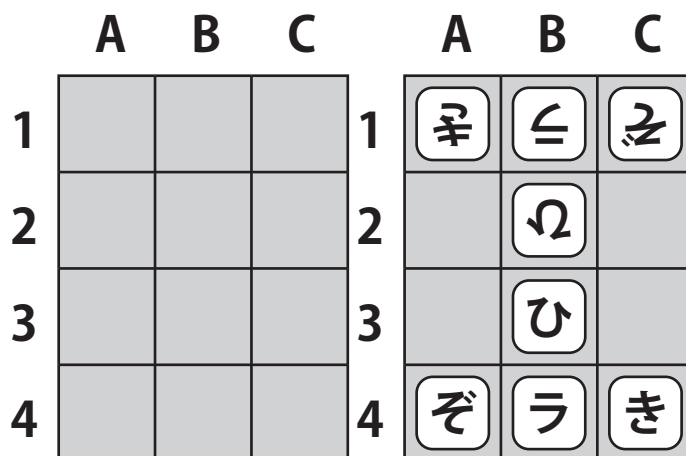


図2: 「どうぶつしょうぎ」のボードと初期局面

**ひよこ** 将棋の歩と同様にプレイヤーから見て前に1マス動ける。相手の陣地に入るとにわとりになる。ひよこの駒を裏返すとにわたりの絵柄になっている。にわとりは斜め後ろ以外の6方向に1マス動ける。

駒には女流棋士の藤田麻衣子一級のデザインによるイラストが書かれているが、本論文では代わりに駒の種類を表す平仮名、片仮名1文字で表現する(図3)。

**進行** プレイヤー2人で遊ぶゲームである。図2右の初期局面から、交互に1手ずつ駒を動かす。自分の手番では、以下のいずれかのプレイをおこなう。

- ボード上の自分の駒を一つ動かす。駒の動ける方向が盤面上で空きマスか敵の駒がある時に動かせる。敵の駒のあるマスに移動することは「捕まえる」と呼び、敵の駒がボード上から取り除かれ自分の持ち駒になる。
- 自分の持ち駒をボード上の空いているマスに置く(打つ)。敵のにわとりを捕まえた場合も持ち駒として打つ場合はひよことして打つ。

**ゲームの目標** 相手のライオンを捕まえるか、相手の陣地に自分のライオンが入り、直後に捕まえられることがなければ(トライと呼ばれる)勝ち。

**詳細** 同じ局面に3回目に到達すると引き分けになる。将棋の二歩、打ち歩詰め、1段目の歩打ち、王手千日手に該当する反則はない。

このゲームの解析を試みることにした。今後、ゲームの探索アルゴリズムの研究において、遺伝学におけるショウジョウバエのような役割をこのゲームが果たすことが可能だということが分かった。今回得たデータの公開が、今後のゲームプログラミング研究に大きく寄与することが期待される。

## 3.2 内容

### 3.2.1 盤面の表現と全局面の列挙

初期局面から開始して「手番のプレイヤーが相手のライオンを捕まえられる時は必ず捕まえて勝つ」、「ライオンがトライした時はゲームを終了する」として到達可能な局面(以下では単に初期局面から到達可能な局面と呼ぶ)のみを対象として解析をおこなう。

それぞれのマスは空白か各プレイヤーの5種類の駒が存在するかの11状態なので4bitで表現可能である。盤面のマスの数は $3 \times 4 = 12$ 個なので $4 \times 12 = 48$ bitで盤面の状態は表せる。持ち駒は各プレイヤー



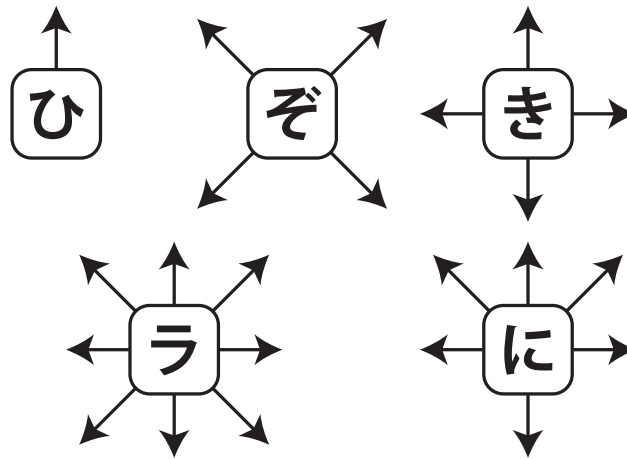


図 3: 駒の動き

3種類で駒の数は0, 1, 2の3種類なのでそれぞれ2bitの計 $6 \times 2 = 12$ bitで表せる。手番は固定しても一般性を失わないので、計60bitで盤面が表現できる。

左右を入れ替えた局面は同じと見なすことができるが、これは64bit整数として表現した時に小さい方の値で正規化することで自然に実現できる。

勝負のついた局面（以下では末端局面と呼ぶ）は、「手番のプレイヤーが敵のライオンを捕まえらるる」局面（以下では勝ち確定局面と呼ぶ）、「勝ち局面でなく敵のライオンが自陣にいる」局面（以下では負け確定局面と呼ぶ）とした。これは、トライ勝ちの時に即座に勝ちとならず、相手に手番が渡ってから相手の負けとなる分、手数が1手ずれるが、プログラムの作りやすさからこちらを採用した。

メモリ16GBのOpteron 2.6GHzマシンで初期局面から到達可能な局面を求めたところ約19分でプログラムが終了し、局面数が246,803,167と求まった。前節で求めた局面数と比較するとかなり減っていることが分かる。なお、このうちの半分以上は末端局面であり、末端局面を除いた局面数は99,485,568となった。

### 3.2.2 後退解析による局面の勝ち負けの確定

すべての局面を数え上げた後は、後退解析 (retrograde analysis) により、すべての局面の勝ち負けを求めることができる。後退解析は以下のようにおこなう。

1. 勝負のついた局面の集合から開始する。
2. 勝負のついた局面の1手前の局面を求める。
  - (手番のプレイヤーの) 負け局面から1手前の局面は (手番のプレイヤーの) 勝ち局面
  - 勝ち局面から1手前の局面の勝敗が未確定の時は、そこから可能な手がすべて勝ち局面に移行する時は、負け局面とする。
3. 操作を繰り返して、勝ち局面の集合も負け局面の集合がそれ以上増えなくなったら終了する。

メモリ16GBのOpteron 2.6GHzのマシンで約5.5時間で求めることができた。

## 3.3 具体的成果

### 3.3.1 局面の勝敗と手数

末端局面を除いた99,485,568局面のうち、手番の勝ちが56,474,473局面、引き分けは2,682,700局面、負けは40,328,395局面と分かった。

勝ち負けに要する手数も記録するようにしたところ、勝ち局面で勝つまでに要する手数の最大値は173であることが分かった。

### 3.3.2 初期局面の勝敗

初期局面である図2右は後手勝ちで78手要することが分かった。先手の4つの合法手それぞれへの後手の応手は以下ようになる。

**B2 ひよこ** 「同ぞう」で後手勝ち（初手から76手）。以下「B3ぞう」、「A2きりん」、「同ぞう」、「同ライオン」、「B3きりん」、「C2ぞう」、「B2きりん」、「同ライオン」、「A3ぞう」、「A2ライオン」、「C3きりん」、「B2ひよこ」、「B2ぞう」、「同ライオン」、「B3ひよこ」、「B1ライオン」、「A3ライオン」、「A2きりん」、「B4ライオン」のように進むのが最長の応答になる。ここまでの後手の応手は必然手（違う手を選ぶと先手勝ちか引き分けになる）だが、次は、「A3ぞう」（初手から76手で後手勝ち）、「A1きりん」（初手から84手で後手勝ち）、「C1ライオン」（初手から84手で後手勝ち）と複数の勝ち筋がある。

**C3きりん** 「A2きりん」以下後手勝ち（初手から78手）。「C4きりん」の後、「B3ひよこ」（初手から78手で後手勝ち）、「A1ライオン」（初手から82手で後手勝ち）、「A1きりん」（初期局面に戻る）、「C2ライオン」（初手から82手で後手勝ち）と複数の勝ち筋がある。

**C3ライオン** 「A2きりん」（初手から78手）、「B3ひよこ」（初手から78手）のどちらかで後手勝ち。

**A3ライオン** 「A2きりん」（初手から78手）、「B3ひよこ」（初手から78手）、「C2ライオン」（初手から82手）のどれでも後手勝ち。

### 3.3.3 合法手の数

末端局面を取り除いた99,485,568局面に関して合法手の平均を求めたところ、約9.435手となった。

合法手がなく、かつ末端局面でない局面は作ることにはできる。ただ、このような局面は初期局面から到達が不可能であることがわかった。

合法手の数の最大値は38で局面数は34あり、すべて手番の勝ち局面である。

### 3.3.4 ツークツワンク (ZugZwang)

ツークツワンクは主にチェスで使われている用語で、自分がパスをすることが許されれば勝ちだが、手を進めなくてはいけないために負けるという局面を指す。

初期局面から到達可能な全局面で以下の条件を満たす局面をすべて列挙した。

- 末端局面でない負け局面で
- 手番を入れ替えた局面が初期局面から到達可能で手番から見て負け局面になっている

その結果、21,839個のツークツワンク局面を見つけることができた。

その中には、手番を入れ替えた局面が元の局面と対称な局面も71個含まれている。容易に分かるように、初期局面もこの一つである。

なお、手番を入れ替えた局面が初期局面から到達可能でないツークツワンクもあり得るので、21,839個というのはツークツワンク局面の個数の下限を与えるに留まっている。

### 3.3.5 敵陣へのひよこ打ちの有効性

「どうぶつしょうぎ」においては将棋と違い、敵陣に（進めない）ひよこを打つ手が禁止されていない。これは、ルールを単純にするための選択だと考えられるが、将棋のように長い利きを持つ駒がない「どうぶつしょうぎ」では、多くの場合は敵陣へのひよこ打ちは良い手ではないが、有効な場合もあると考えられる。

	A	B	C
1			
2		ぞ	㇏
3	ラ	㇏	㇏
4	き		㇏ ひ

図 4: 敵陣へのひよこ打ちが有効な局面

そこで、進めないひよこ打ちのみが正解（勝ちの局面で勝ちを維持できる、引き分け局面で引き分けを維持できる）である局面を全局面の中から見つけたところ、68局面存在した。その一つを図4に示す。

ここでの勝つための唯一の手は「B1ひよこ」でそれ以外の手は負けとなる。「B1ひよこ」のあと、「同ライオン」は「B3ライオン」で、「B4きりん」は「同きりん」で、それ以外の手は「A2ライオン」で勝ちとなる。一方、「B1ひよこ」以外の手では「A1ぞう」が最も長い手延ばしだが、45手後に負けとなる。

敵陣へのひよこ打ちを認めないとゲーム全体への影響がどの程度あるかを確認するために、敵陣へのひよこ打ちを認めないルールの上で後退解析をおこなってみた。その結果、全体のうちの4,301局面で勝敗、引き分けが変わることが分かったが、その中には初期局面は含まれず、また、初期局面で勝ちに要する手数も78手で変化がないことが分かった。

### 3.3.6 詰みのある局面

将棋においては、攻める側が「王手」の連続で勝つ局面のことを「詰みがある」局面と呼び、「詰将棋」というパズルの対象として広く楽しまれている。

「どうぶつしょうぎ」でも、「相手がパスしたら次にライオンを捕まえられる」または「相手がパスしたら次にトライできる」手の連続で攻めて勝てる局面を「詰みのある局面」と定義することができる。これも後退解析ですべて求めることが可能である。

この結果、17,213,997個の「詰みのある局面」が存在し、勝ち局面の約30%が「詰みのある局面」であるということがわかった。また、手数の長い詰み局面はそれほどなく、最長でも23手詰めであることが分かった。

23手の詰みのある局面は4つあったが、いずれも詰むまでの手順は長いがパズル的な面白さはない。「面白い」詰み局面が存在するかどうかは更なる検討が必要になるだろう。

## 4 成果要覧

### 招待講演／招待論文

[招待1] 田中哲朗: 査読者はどこを見る? - 査読を通るゲーム研究論文を書くために, 第14回ゲームプログラミングワークショップ, 2009年11月



## 受賞関連

[受賞 1] 金子知適, 田中哲朗, 山口和紀, 川合慧:平成 20 年度論文賞, 情報処理学会, 2009 年 5 月

## 著書/編集

[著書 1] 笈捷彦, 山口利恵, 近山隆, 石畑清, 松崎公紀, 柴山悦哉, 田中哲朗: 目指せ!プログラミング世界——大学対抗プログラミングコンテスト ICPC への挑戦, 近代科学社, 2009 年 7 月

## その他の発表論文

[発表 1] 田中哲朗: 「どうぶつしょうぎ」の完全解析, 情報処理学会ゲーム情報学研究会, Vol.2009-GI-22, No.3, pp. 1-8, 岩手, 2009 年 6 月.

## 公開ソフトウェア

[公開 1] 田中哲朗, 金子知適, 森脇大悟, 副田俊介, 林芳樹, 竹内聖悟: osl-for-csa (バージョン 0.6), コンピュータ将棋協会コンピュータ将棋選手権使用可能ライブラリ, 2010 年 1 月.

[公開 2] 田中哲朗, 金子知適, 森脇大悟, 副田俊介, 林芳樹, 竹内聖悟: gpsshogi-for-csa (バージョン 0.6), コンピュータ将棋協会コンピュータ将棋選手権使用可能ライブラリ, 2010 年 1 月.

## 特記事項

[特記 1] GPS 将棋: 第 19 回世界コンピュータ将棋選手権優勝, コンピュータ将棋協会, 2009 年 5 月.

[特記 2] GPS 将棋: 第 5 回世界コンピュータ将棋最強決定戦優勝, 北陸先端科学技術大学院大学, 2010 年 1 月.

## 報道関連

[報道 1] 読売新聞 2009 年 5 月 6 日朝刊: 「GPS 将棋」が初優勝.

[報道 2] 週刊将棋 2009 年 5 月 13 日号: 革命的結果勢力図一変 - GPS 将棋が初優勝.

[報道 3] 朝日新聞 2009 年 5 月 13 日夕刊: 世界コンピュータ将棋選手権 - 「GPS 将棋」が初優勝.

[報道 4] NHK BS2、囲碁将棋ジャーナル 2009 年 5 月 16 日: 世界コンピュータ将棋選手権速報.

[報道 5] NHK BS2、囲碁将棋ジャーナル 2009 年 5 月 30 日: 世界コンピュータ将棋選手権特集.

[報道 6] 読売新聞 2009 年 6 月 2 日夕刊: 「GPS」が乱戦制し初優勝.

[報道 7] 産経新聞 2009 年 7 月 26 日: 訪れるか将棋界「X デー」.

# 教育支援システムの開発と運用

## —講義データベース—

関谷 貴之

### 1 概要

広い意味で教育を支援するシステムとして、教育用計算機システムと学習管理システム CFIVE に関する開発と運用を行っている。また、講義データベースと題して、講義のシラバスを用いてカリキュラムを分析する研究を行っている。その他、プログラミング教育の支援や成績データの分析なども行っている。

### 2 講義データベース

#### 2.1 背景

教材や論文のレポジトリ、オンラインシラバスという形で電子化されたコンテンツが大量に存在する現在、コンテンツの検索や類似度を算出する研究は広く行われている。これに対して、コンテンツが対象とする分野の範囲や、その分野の中でのコンテンツの位置づけを全体的に把握する研究は比較的少ないと考えられる。そこで、コンテンツに含まれる単語の分布を分析して可視化することで、カリキュラムの分析に活用するための研究を行っている。

#### 2.2 内容

2009年度は、第一に Latent Dirichlet Allocation(LDA) と Isomap を用いたこれまでの手法を改良し、異なるカリキュラムを同一の平面の上で比較可能とした。第二に、カリキュラムの分析結果を Web ブラウザで確認できるツールを開発した。

##### 2.2.1 カリキュラムの分析方法

本研究では、シラバスの文章が講義の特徴を表しているとの仮定の下で、多数のシラバスで構成されるカリキュラムの分析において、カリキュラムのマップを作成することで、カリキュラムの構造を可視化することを試みている。

異なるカリキュラムを可視化して分析するには、以下の条件を満たさなければならないと考える。

1. シラバスの可視化のための基準を設けること。
2. シラバス同士の関連度が可視化された空間内の距離として反映されること。
3. 異なるカリキュラムを同一のマップに示すこと。

この条件を満たすカリキュラムの分析手法を以下に説明する。また、その手法の概念図を図 1 に示す。

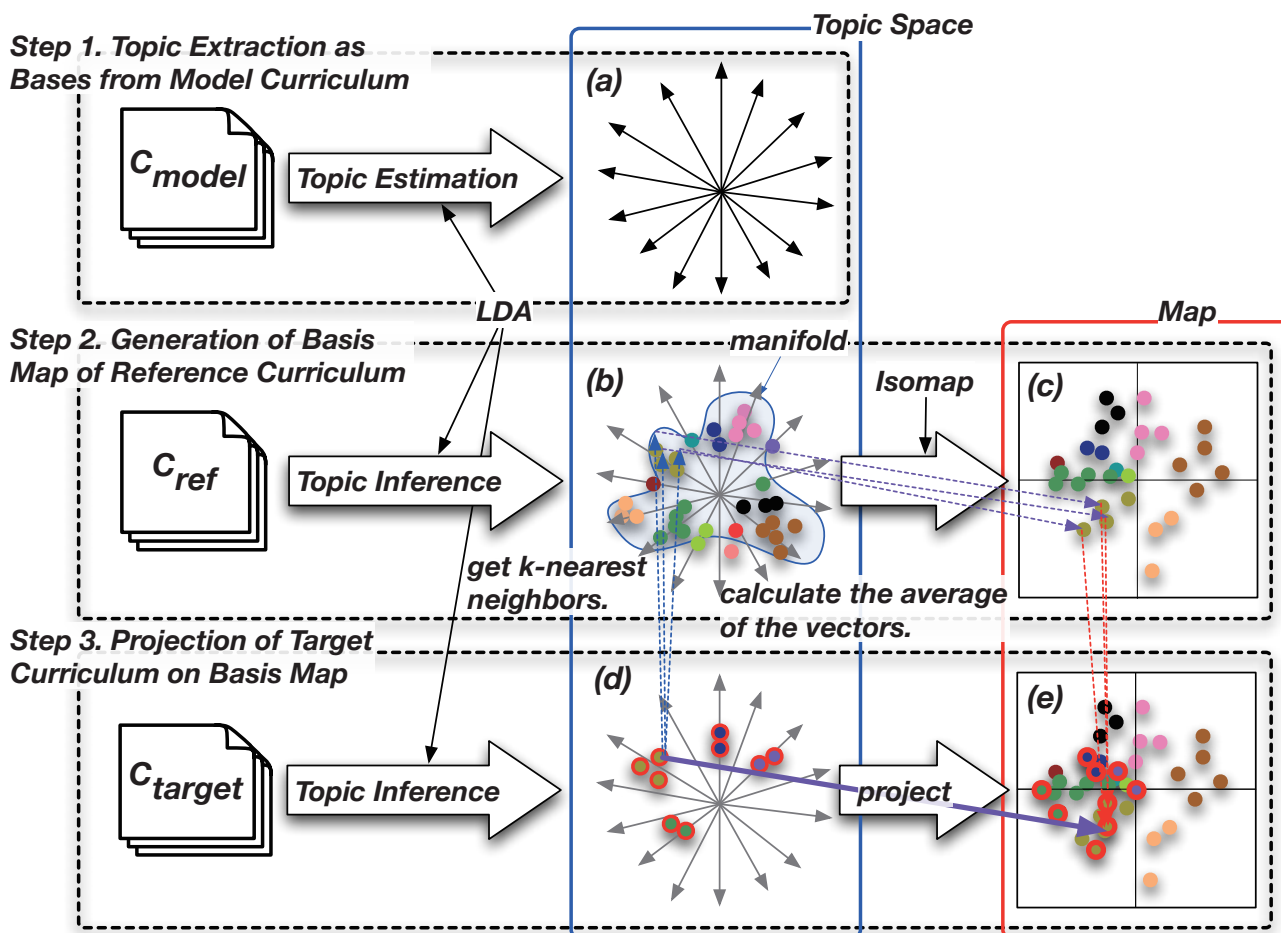


図 1: カリキュラム分析手法の概念図

- Step 1. 条件 1 を満たすために、分析するカリキュラムが扱う分野において基準となるカリキュラム“Model Curriculum”  $C_{model}$  を導入する。  $C_{model}$  を LDA で分析することで、分析の基準となるトピックを抽出する。ここで抽出したトピックは、今後分析対象となるカリキュラムの特徴を可視化するための高次元の空間“Topic Space”となる(図 1(a))。
- Step 2. 次に、分析するカリキュラムが扱う分野を有る程度広くカバーすると考えられる参照用のカリキュラム“Reference Curriculum”  $C_{ref}$  を導入する。まず  $C_{model}$  のトピックに基づいて、  $C_{ref}$  を分析することで、前述の Topic Space の中に  $C_{ref}$  を配置する(図 1(b))。Topic Space は高次元空間であり、カリキュラムの全体構造を目で見て判断することができない。そこで、高次元空間内の多様体構造を保持したまま、これを広げるようにして次元を縮退する手法 Isomap を用いる。その結果、条件 2 を満たすように  $C_{ref}$  を平面上に展開したマップが得られる(図 1(c))。
- Step 3. 次に分析対象となるカリキュラム“Target Curriculum”  $C_{target}$  を  $C_{ref}$  と比較する。最初に、  $C_{ref}$  と同様に  $C_{model}$  のトピックに基づいて、  $C_{target}$  を分析することで、前述の Topic Space の中に  $C_{target}$  を配置できる(図 1(d))。次に、  $C_{target}$  の個々のシラバスについて、Topic Space 内で近傍に有る  $C_{ref}$  のシラバスを抽出する。更に抽出したシラバスに対応する図 1(c) のマップ上のベクトルを合成することで、  $C_{target}$  のシラバスをマップ上に投影する。これによって、  $C_{target}$  を  $C_{ref}$  のマップ上に投影して比較が可能となり、条件 3 を満たすことができる(図 1(e))。

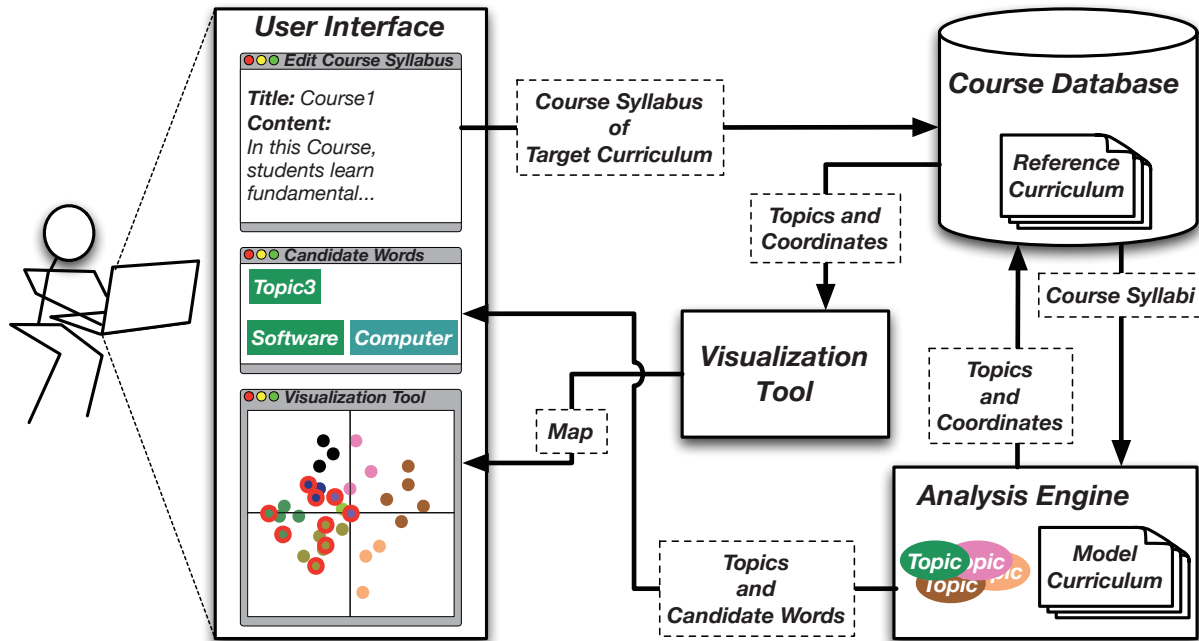


図 2: カリキュラム分析ツールの構成

## 2.2.2 カリキュラム分析ツール

2.2.1 節で述べた手法に基づいて、カリキュラムの分析を行う Web ベースのツール<sup>1</sup>を開発した。図 2 にツールのアーキテクチャを示す。また、図 3 にツールのスクリーンショットを示す。

利用者が“User Interface”から入力したタイトルや内容などの Target Curriculum の情報は、“Course Database”と呼ぶデータベースに保存される。Course Database には Reference Curriculum のシラバスや、可視化に必要なデータも保存される。

“Analysis Engine”は 2.2.1 節の手法を実装したカリキュラム分析用のエンジンである。Analysis Engine は LDA を用いたトピックの抽出や各シラバスの座標値の算出を行う他に、利用者が入力したシラバスと指定したトピックに基づいて、関連のある用語を提示することが出来る。

“Visualization Tool”は Analysis Engine で生成したマップを表示するツールである。Ajax の技術を用いて、表示されたマップを拡大・縮小・移動することができる。

## 2.2.3 カリキュラムの分析例

本研究で開発した手法とツールの利用例として、コンピュータ科学のカリキュラムを分析した。具体的には、Model Curriculum として ACM Education Board と IEEE Computer Society’s Education のタスクフォースがまとめた CS2008 を、Reference Curriculum として米国 MIT の OCW<sup>2</sup>で公開されているカリキュラムを、Target Curriculum として英国の遠隔教育による大学 The Open University(OU)<sup>3</sup>のカリキュラムを用いた。

CS2008 ではコンピュータ科学分野の内容を“Knowledge Area”と呼ぶ 14 個の領域に分けている。それぞれの Area はさらに 10 個程度の“Knowledge Unit”に分けられており、本研究ではこの Unit をシラバスと見なしている。表 1 は、CS2008 の Area や Unit と LDA で抽出したトピックとの関連を示している。

Area を示すアルファベット 2 文字の後ろに有る数字は Unit の数を示している。2 つ目の数字は、

<sup>1</sup><http://www.sekiya.ecc.u-tokyo.ac.jp/coursedb/>

<sup>2</sup><http://ocw.mit.edu/>

<sup>3</sup><http://www.open.ac.uk/>



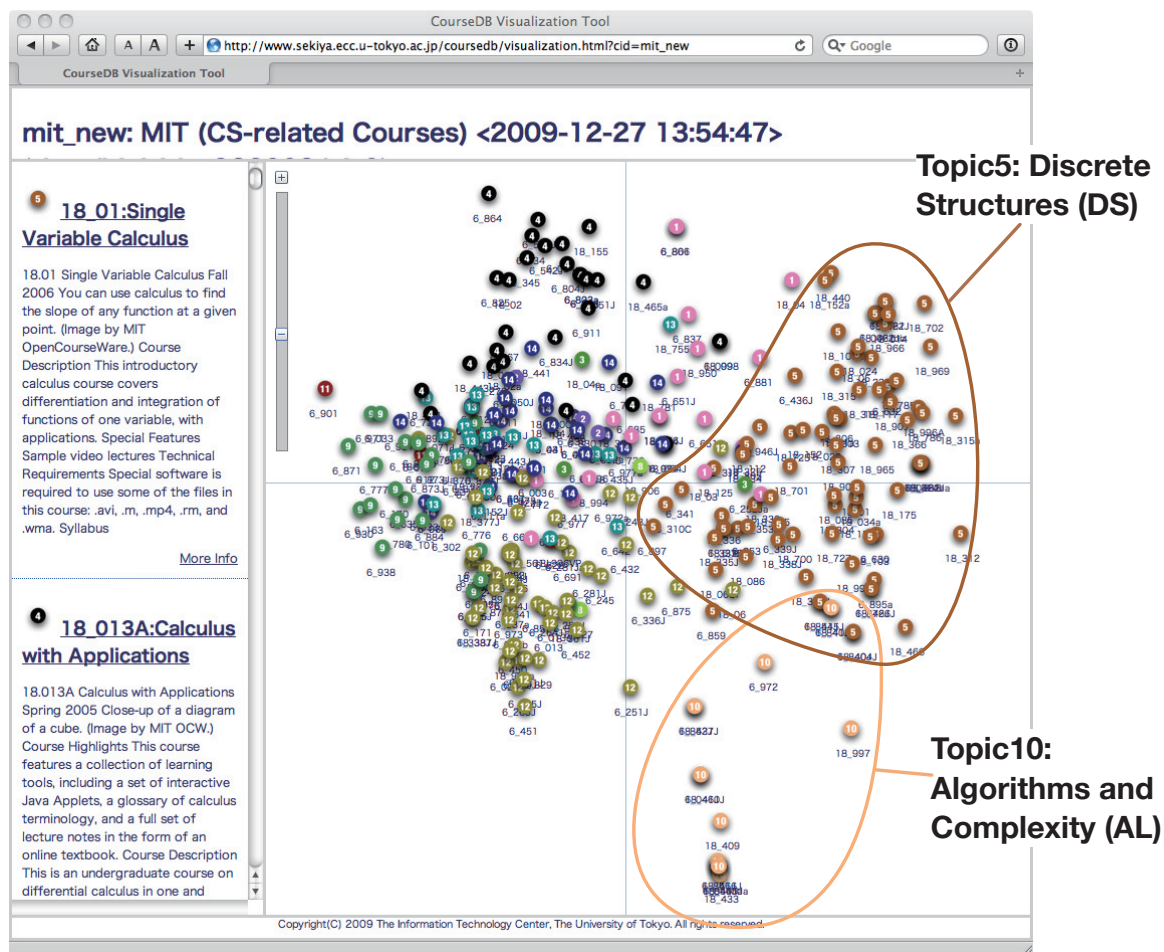


図 3: カリキュラム分析ツールのスクリーンショット (但し、Topic 5 と Topic 10 は別途書き足したものである。)

当該 Area の Unit の総数、1 つ目の数字はその中で当該 Topic に強く関連のある Unit の数である。ある Area について、ほとんどの Unit が特定の Topic に関連の有る場合、全体をボールド体で示している。例えば、GV(Graphics and Visual Computing) の全ての Unit が Topic 1 と強く関係があることから、ボールド体となっている。

2.2.1 節の手法を用いて得られた MIT のカリキュラムのマップを図 3 に示す。丸い記号とその下にある時間割番号が個々の講義に対応しており、関連の強いトピックに応じて色と数字を変えている。図 3 では MIT のカリキュラムの特徴として、比較的多くの講義が Topic 4, 5, 12 (DS, IS, NC and CN) と関連が強いこと、またその一方で Topic 6 (OU) に関わる講義が少ないことが判る。

図 4 は MIT と OU (白く縁取られた記号) のマップであり、OU の講義の多くが MIT のマップ上の左の部分にプロットされていることが判る。OU では、“Topic 12: Net Centric Computing (NC), Computational Science (CN)” や “Topic 14: Social and Professional Issues (SP)” と強く関連がある実学的な講義が多い一方で、“Topic 5: Discrete Structures (DS)” のような基礎理論的な講義が少ないことを意味している。これは、OU の学生の 70% が社会人で実学的な志向が強いことと一致している。

### 2.3 具体的成果

講義シラバスをもちいたカリキュラムの分析手法として、LDA と Isomap を用いた可視化手法を開発して、これを実装した Web ベースのツールを構築した。なお成果の一部は [査読付 1] で発表した。

表 1: CS2008 Knowledge Areas and Topics

Topic	Knowledge Areas and Numbers of Units
1	<b>GV: 13/13, HC: 2/10, IM: 1/15</b>
2	<b>IM: 9/15, HC: 1/10</b>
3	<b>PF: 8/8, HC: 1/10, IM: 1/15, PL: 1/11</b>
4	<b>IS: 11/11, IM: 2/15, HC: 1/10</b>
5	<b>DS: 6/6, HC: 2/10</b>
6	<b>OS: 12/14</b>
7	OS:2/14, HC:1/10
8	<b>PL: 10/11</b>
9	<b>SE: 14/14</b>
10	<b>AL: 11/11, IM:2/15</b>
11	HC:2/10
12	<b>NC: 9/9, CN: 3/3</b>
13	<b>AR: 10/10</b>
14	<b>SP: 11/11</b>

次に示すように、アルファベット大文字の2文字で、CS2008のKnowledge Areaを示す。GV: Graphics and Visual Computing, HC: Human-Computer Interaction, IM: Information Management, PF: Programming Fundamentals, IS: Intelligent Systems, DS: Discrete Structures, OS: Operating Systems, PL: Programming Languages, SE: Software Engineering, AL: Algorithms and Complexity, NC: Net Centric Computing, CN: Computational Science, AR: Computer Architecture, and SP: Social and Professional Issues.

### 3 成果要覧

#### 査読付論文

[査読付 1] T. Sekiya, Y. Matsuda, and K. Yamaguchi: Analysis of Curriculum Structure Based on LDA, ICEIT'09, Proceedings of International Conference on Education and Information Technology, pp.561–566, San Francisco, CA, USA, 2009.

[査読付 2] 山本 三雄, 関谷 貴之, 山口 和紀: 情報教育における理論と技術に関する研究の調査. 情報処理学会論文誌 (ジャーナル), Vol.50, No.10, pp. 2449 - 2461, 2009.

#### その他の発表論文

[発表 1] 岩藤 健弘, 関谷 貴之, 杉山 剛士, たかはし ゆういち: 学習管理システム CFIVE の運用, 平成 21 年度情報教育研究集会講演論文集, E2-4, 2009.

[発表 2] 秋田 英範, 丸山 一貴, 関谷 貴之, 佐々木 馨, 増田 均: 学内向けメールサービスの運用と稼働状況について, 平成 21 年度情報教育研究集会講演論文集, pp.437–440, 2010.

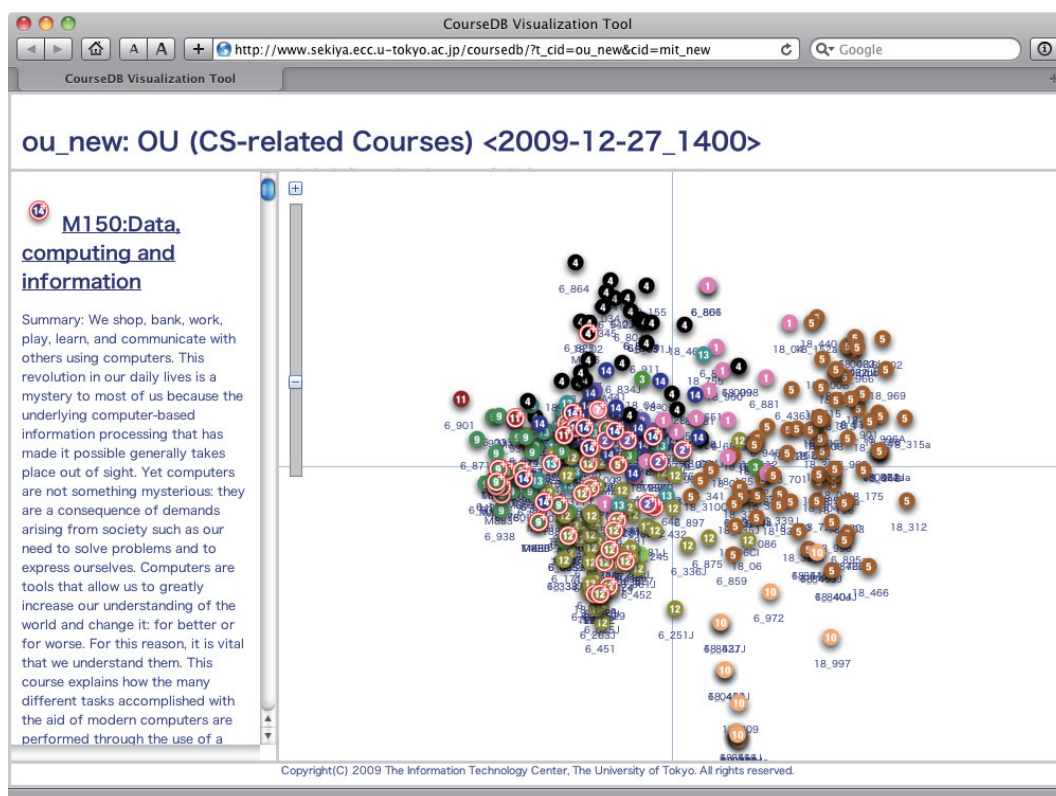


図 4: MIT と OU のカリキュラムのマップ

# Wikiとペンタブレットにおけるユーザインタフェースの研究

丸山 一貴

## 1 RDBとの連携によるWikiデータ入力の強化

### 1.1 背景

多数のユーザが非同期的にコンテンツを作成、更新していくプラットフォームとして、wikiが活用されている。代表的な利用例にWikipediaがある。コンテンツの中でも、特にリストアップされるようなデータについては「表」の形で保持されることが多く、wiki記法では“|”の記号でデータを区切って並べることで表現されている。

Consumer Generated Mediaでは、情報が最初から正確に揃っていることは稀であり、通常は誤りがあったり欠落したりしている。こうした情報を追加したり修正したりするのに、wiki記法を用いた編集操作は不向きである。また、表に登録されているデータ全てを表示するだけでなく、webページ(wikiのページ)ごとに必要なカラムだけを表示したり、特定のカラムでソートしたりといった操作も必要である。

著名なwiki実装であるPukiWikiには、これらの一部の機能を実現するプラグインは存在するものの、特に後者の問題を根本的に解決する手段はない。

### 1.2 内容

提案システムではwikiの表データをRDBに格納し、RDB操作に相当するwiki記法を導入することと、特定箇所のデータ操作を容易にするユーザインタフェースの実装により、前述の問題を解決した。

本システムのドキュメントモデルを図1に示す。wikiによる表データの入力や編集が行われる場合には、バックエンドでRDBと連携し、データの入力や更新を実施する。また、RDB単体では実施しにくいカラムの追加も、wikiによる入力の使い勝手を損なわないよう、列の追加と削除が容易にできる機能を実装している。

評価のため、10人のユーザによる実験を行った。wiki操作に慣れているユーザと不慣れたユーザが偏らないよう2つのグループに分け、表データを持ったページとしてWikipediaから「爆風」と「IntelCore2」の2つのページのドキュメント編集を行わせる。それぞれPukiWikiと提案システムで作業し、編集に要した時間を計測した(表1)。

2つの対象データについて異なる結果となっている。「爆風」では表データが比較的小さく、また表内のデータも日本語が登場するために編集対象のデータを探すのが容易であったため、被験者がより慣れているPukiWikiの方が時間が早かったと考えられる。一方、「IntelCore2」の方は表内に類似のデータが多数並ぶことや、数値のデータが多いことなどから提案手法が有利との結果が現れたと考えている。

なお、編集のしやすさについての主観的評価も集計しており、こちらでは提案システムが圧倒的に高評価となっている。



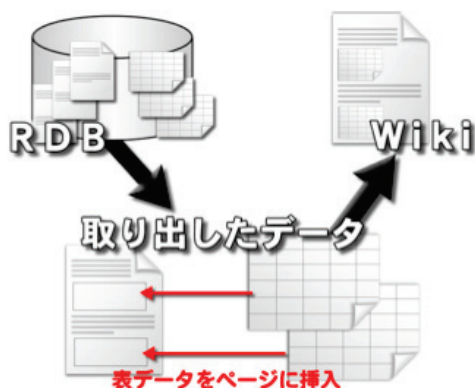


図 1: 提案システムのドキュメントモデル

表 1: 評価実験：平均編集時間

	PukiWiki	提案システム
「爆風」ドキュメント	4分59秒	5分29秒
「IntelCore2」ドキュメント	5分18秒	3分58秒

### 1.3 具体的成果

共著で論文を執筆し、共著者が発表を行った（成果要覧の [3], [5]）。

## 2 ペンタブレットの空中動作による3次元データ入力

### 2.1 背景

3次元データ入力を行うためには専用のハードウェア（PHANTOM など）が必要とされてきた。汎用品として安価に入手可能な入力デバイスとして静電誘導方式のペンタブレットがあるが、これは専用ペンの先端部が検知可能領域内にある場合に2次元座標が得られる装置である（図2）。

検知可能領域の深さ  $L$  はペンタブレットに固有の値であるため、ペン先端部の検知開始点や検知終了点（消失点）に関しては3次元座標が得られることになる。これを利用した3次元ベクトルデータ入力方式を提案するとともに、入力誤差等を測定する基礎実験を行った。

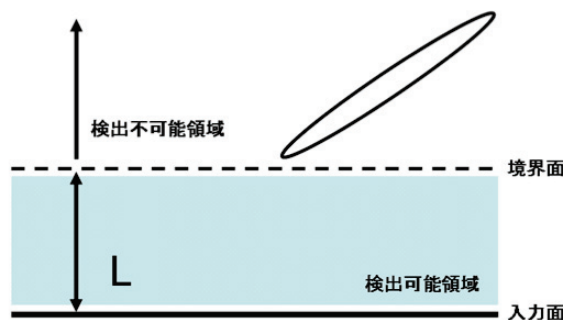


図 2: ペンタブレットの検知範囲

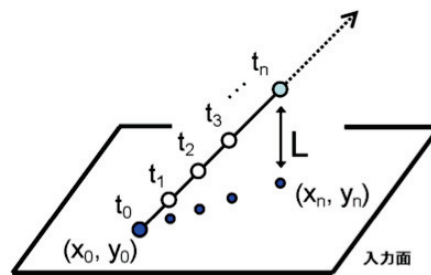


図 3: リリース→消失による 3 次元ベクトル

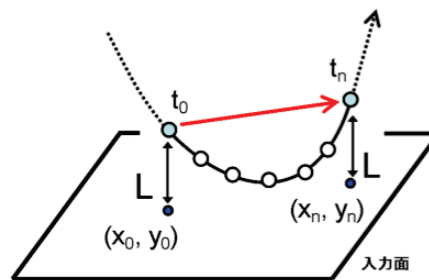


図 4: 空中ジェスチャ

## 2.2 内容

3次元ベクトル入力について、ペンがタブレット表面に接触した状態からリリースし、消失に至る場合を例にとって説明する。図3に示す通り、開始点である $t_0$ ではクリック状態からリリースに変化するため高さが0であると分かり、その座標は $(x_0, y_0, 0)$ と測定できたことになる。消失点 $t_n$ は、タブレットに固有の検知可能領域 $L$ の大きさを予め測定しておくことで $(x_n, y_n, L)$ と測定可能である。途中の点 $t_1, \dots, t_{n-1}$ では高さは測定不能だが、開始点から消失点までを直線状に移動したと見なす。なお、移動中も一定のサンプリングレートで2次元座標が取得されているため、ペン先端部の速度変化を直線上にマッピングしている。検知可能領域に入ってタブレット表面をクリックする、感知→プレスの場合も同様である。

タブレット表面に接触しない場合の応用として、空中ジェスチャの入力も可能である。この場合は図4に示すように、感知点 $t_0$ と消失点 $t_n$ が連続して観測されることになる。例えばタブレット上の空中で、ペン先端部が垂直な円運動を行った場合に相当する。

3次元ベクトル入力の所要時間と角度誤差(表2)を、空中ジェスチャ入力の成功率(図5)を測定した。表中のペンタブレットとは外付けのタブレット入力装置を用いた場合、液晶とはタブレット内蔵ディスプレイを使用した場合で、回転時とはディスプレイを90度回転させてポートレートの向きにした場合を意味する。

これらの入力方式を利用するインタフェースを設計する際には、これらのデータが有用である。例

表 2: 3次元ベクトル入力の方位角に関する評価結果

	所要時間 (ms)	角度誤差 (度)
ペンタブレット	223	13.2
液晶 (通常時)	281	12.9
液晶 (回転時)	252	8.7

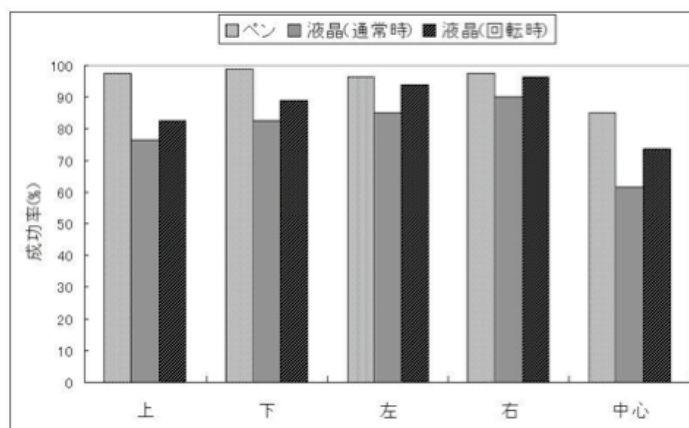


図 5: 空中ジェスチャ入力の成功率

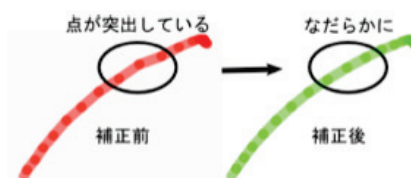


図 6: 手書きストロークの補正

例えばパイメニューを設計する場合には、前述の角度誤差を考慮して作成することが必要である。

### 2.3 具体的成果

共著で論文を執筆し、共著者が発表を行った（成果要覧の [1]）。

## 3 手書きを用いた動画上の非同期コミュニケーション

### 3.1 背景

ネットワークを介してユーザ同士が時間を共有しない非同期型のコミュニケーションが旧来より利用されている。古くは NetNews やメーリングリスト、比較的新しいものとして web 掲示板が代表的である。これらはいずれもテキストベースのやり取りであったが、言及の対象を動画に拡張して新しいコミュニケーションを創造したサービスとして、ニコニコ動画がある。

動画を対象として言及する場合には、再生時のタイムラインにおける特定の時刻を指定することと、その時間に発言を再生する必要があるが、ニコニコ動画ではそれを再生中の動画にオーバーレイによって表現している。これはテキストベースの表現にとどまっており、画面内の特定の場所を正確に指定したり、動画に合わせて動きを表現することは困難である。

そこで、手書き入力によって任意の時刻の任意の座標に、任意の曲線を描いてコミュニケーションを行うシステムを試作し、考察した。

### 3.2 内容

手書きで入力を行うことは一見簡単のように思えるが、実世界で手書きの字を綺麗に書けるとは限らないように、計算機上で入力する場合も思い通りに書くことは難しい。そこで、手書きストロークが意図せず折れ曲がったとみなされる場合には滑らかな曲線となるよう補正している（図 6）。

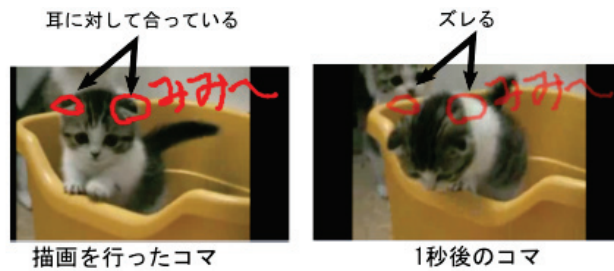


図 7: 時間経過によるアノテーションのズレとその緩和

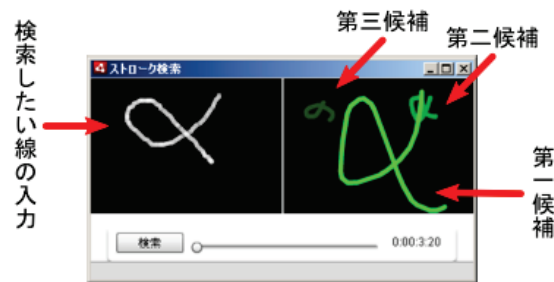


図 8: ストローク検索ウィンドウ

特定のフレームの、特定の座標に対して適切なアノテーションを行ったとしても、1秒程度後のフレームでは既にアノテーション対象が画面内で移動してしまい、アノテーションが不適切なものになってしまうことも多い。不適切なアノテーションの違和感を緩和するため、時間とともに半透明化していくよう工夫した（図7）。

テキストベースの場合、特定の入力を見つけ出すには入力のログを検索すればよいが、手書きストロークの場合にはそのままでは検索できない。しかし、長時間の動画をずっと見続けてストロークを探すことは現実的でない。そこで、ストロークの検索機能を実装した（図8）。

このような機能を持つシステムを実装して試用し、以下のような知見を得た。

- 自由度が高く、画面内の特定の箇所を示すアノテーションに有効
- 入力されるアノテーションは以下の3種類に分類される
  - 画面内の位置に対するアノテーション
  - 動画内容に対する言語表現のアノテーション
  - 頭出しを目的とした、検索キーとするための記号入力
- 非同期コミュニケーションによるアノテーション数の増加
- 具体的な説明を要するような言語表現には不適切
- 特定の動画再生区間にタイミングを合わせて入力するインタフェースが必要
- ストローク検索は有効であったが、複数ストロークからなる「文字」を検索するには改善が必要

### 3.3 具体的成果

共著で論文を執筆し、共著者が発表を行った（成果要覧の [2]）。

## 4 成果要覧

### 査読付

- [1] 梅林靖弘, 丸山一貴, 寺田実, “ペンタブレットの空中動作を利用したインタラクション”, 第17回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2009), pp.121–122, 2009.

### 発表

- [2] 平山慧, 丸山一貴, 寺田実, “手書きを用いた動画上の非同期コミュニケーション”, 情報処理学会全国大会, 5ZF-8, 2010.
- [3] 奥村俊也, 丸山一貴, 寺田実, “表データ操作を RDB で強化した Wiki システム”, 情報処理学会全国大会, 5ZF-5, 2010.
- [4] 佐藤和哉, 丸山一貴, 寺田実, “CodeMusician: プログラム可聴化システム”, 情報処理学会研究報告, Vol.2009-MUS-81, No.22, pp.4, 2009.
- [5] 奥村俊也, 寺田実, 丸山一貴, “表データ操作を RDB で強化した Wiki システム”, 情報処理学会研究報告, Vol.2009-HCI-134, No.10, 2009.
- [6] 井桁正人, 寺田実, 丸山一貴, “ScoutView: Web ページにおけるナビゲーション支援インタフェース”, 情報処理学会研究報告, Vol.2009-HCI-133, No.8, 2009.

# 学術情報研究部門

## 学術情報研究部門概要

中川 裕志

統計的機械学習とそのテキスト処理への応用

二宮 崇

自然言語処理のための基礎理論と応用  
—オンライン学習と英文読解支援—

吉田 稔

Web文書の分析・活用システムの研究  
—人名曖昧性解消システムの性能向上—

清田 陽司

これからの電子図書館を支える知識ナビゲーション技術の研究

清水 伸幸

知識発見と機械翻訳の自然言語処理要素技術  
—機械学習テンプレートライブラリー—



# 学術情報研究部門

## 概要

部門長 中川裕志

学術情報研究部門は、情報検索、図書館などに存在する学術情報資源への情報ナビゲーション、またこれらの応用研究の基礎となる機械学習、テキスト処理の研究を行っている。平成 21 年度は、教授 1 名、講師 1 名、助教 2 名、特任助教 1 名、また大学院生としては情報理工学系研究科博士課程 3 名、修士課程 5 名、学際情報学府博士課程 1 名修士課程 2 名が在籍し、具体的には以下にあげるような研究を行なった。この結果、査読論文 9 件、公開ソフト 3 件、招待講演等 10 件、受賞 6 件、その他発表 35 件の成果をあげた。

## 1 研究内容

### 1.1 情報ナビゲーション

学術情報へのナビゲーションは、膨大な学術情報を有する図書館の情報資源を効果的に検索する道筋を提示することを目的にした研究である。本研究部門では清田助教が中心になって数年にわたって力をいれてきた開発項目であり、研究部門で研究成果を元にベンチャー企業を立ち上げて実用化している。

学術情報研究部門で開発し、ベンチャー企業で事業化した図書館情報ナビゲーションシステム：リッテルナビゲータはすでに多数の大学図書館に納入されているが、2009 年度は「リサーチナビ」という名前で国立国会図書館へ導入された。その特徴は、Wikipedia などのいわゆる Web2.0 資源をレファレンスサービスに活用するところにあり、内外に高い評価を得ている。[招待 1,2,3,4,5,6,7,8] [査読付 2] [発表 1,6,9,19] [報道 1]

2009 年度における主な成果としては、リサーチナビの開発、項目間の構造のグラフィック表示機能の開発、Hadoop を利用したナビゲーションシステムの構想がある。[招待 9,10]

### 1.2 情報抽出

最近の情報検索に関する研究は、昨年度で終了した経産省の「情報大航海」プロジェクトや、科学研究費・特定領域研究「情報爆発」の成果にも見られるように、目的特化型の検索を狙うものが増えている。すでに Web のロングテールの重要性と、その検索の困難さは指摘されているところであるが、この問題を相手にした場合の方向性は、汎用のロングテール情報発掘ではなく、目的を固定した検索と抽出である。例えば、上記の特定領域研究「情報爆発」では、Web から抽出した大規模ディレクトリを用い、利用者の想定する問題に答える「鳥式」という鳥澤の研究、膨大なコンテンツで未踏分野を探す灘本のコンテンツホールの発見、Web で対立する意見を対比して俯瞰する藤井の研究などがある。我々は、より一般性のあるいくつかの提案とシステム開発を行った。具体的には、以下のようなテーマである。

- Web 人名検索結果の名寄せ。この研究では、第 2 回の WePS(Web People Search)というタスクで全世界からの参加 17 チーム中 2 位となった。[発表 12,23,35] [公開 1]
- テキスト中の数値表現検索システムに関する研究を進めた。[発表 27,30]



- 同義語・類義語および対訳抽出に関する研究。これに関しては、専門用語の日中対訳辞書の自動生成を行い、上位 3000 位まででは 90%の精度を得た。[査読付 6] [発表 22,31]
- 昨年に引き続き株価とテキストの関連分析に関する研究を続けている。[発表 25,26]
- 文書構造の解析を行った。[発表 28]
- Wikipedia からの意外性に沿った情報ランキングを、Wikipedia のカテゴリ構造を素性にして行った。[発表 2,8,13]
- 多言語ブログおよび多言語 Wikipedia の分析を行うツールの開発と実際のデータに関して評価、分析を行った。[査読付 7] [発表 3,4,5,7,10,14]
- スプログフィルタリング[査読付 5]

### 1.3 機械学習とテキスト処理

統計的機械学習における教師なし学習であるクラスタリングの基礎研究、少量のメモリで実行できるオンライン学習アルゴリズムの開発などに関して以下の研究を行った。

- ベイズ統計に基づくクラスタリングアルゴリズムにおいて、量子力学におけるトンネル効果のアイデアを適用した量子変分ベイズ法を提案した。世界的にも注目されているアイデアである。[査読付 4] [発表 18]
- 定評あるクラスタリングアルゴリズムである LDA(Latent Dirichlet Allocation)のオンライン学習化を行った。[発表 15]
- オンライン学習において性能の高さに定評のある Passive Aggressive アルゴリズムをサポートクラスという新規概念を導入して多クラス分類に適応するアルゴリズムの開発した。[発表 16,21]
- プライバシー保護データマイニングにおける結託耐性のあるアルゴリズムを実現するプロトコルを提案した。[発表 24]
- オンライン学習を応用した英文テキスト読解支援システムを開発し、評価した。[査読付 8,9] [発表 11,17,20,33,34] [公開 2,3]

## 2 図書館電子化研究部門 成果要覧

### 招待講演／招待論文

[招待 1] 清田 陽司: 図書館分類体系と Wikipedia を統合した情報探索支援システムの開発, 日本図書館研究会情報組織化研究グループ／情報知識学会関西部会 合同研究会, 大阪科学技術センター, 大阪市, 2009 年 7 月.

[招待 2] 清田 陽司: Wikipedia をいかに使いこなすか? -知識抽出、情報ナビゲーション、そしてトピック発見-, 第 114 回 Ku-librarians 勉強会, 京都大学, 京都市, 2009 年 8 月.

[招待 3] 清田 陽司: Wikipedia と図書館情報資源のマッシュアップ, 第 36 回生物医学図書館員研究会, 順天堂大学, 東京都, 2009 年 9 月.

[招待 4] 清田 陽司: リサーチ・ナビ検索システムの技術, 第 11 回図書館総合展/学術情報オープンサミット 2009 フォーラム企画, 国立国会図書館主催, パシフィコ横浜, 横浜市, 2009 年 11 月.

- [招待 5] 清田 陽司: 学生向けレファレンス支援ツールの可能性, 第 11 回図書館総合展/学術情報オープンサミット 2009 ミニ・フォーラム&プレゼンテーション企画, 紀伊國屋書店主催, パシフィコ横浜, 横浜市, 2009 年 11 月.
- [招待 6] 清田 陽司: 知識体系の新たな融合: 情報探索と件名標目表の活用をめぐる, 京都大学図書館機構 平成 21 年度第 2 回講演会「次世代 OPAC を考える ー目録情報の視点からー」, 京都大学, 京都市, 2009 年 11 月.
- [招待 7] 清田 陽司: リサーチ・ナビ検索システムの技術, 参考書誌研究, 国立国会図書館主題情報部 編, No. 71, pp. 33-53, 2009 年 11 月.
- [招待 8] 清田 陽司: Wikipedia を活用した新たな情報ナビゲーションシステムの提案, 薬学図書館, 日本薬学図書館協議会 編, Vol. 55, No. 1, pp. 51-59, 2010 年 1 月.
- [招待 9] 清田 陽司: Apache Hadoop による大規模データの研究活用事例, 産学連携共同研究シンポジウム「クラウド・コンピューティングと大学発ベンチャー」, 東京大学産学連携本部 / (株)ユニファイ・リサーチ共催, 東京大学情報学環福武ホール, 東京都, 2010 年 1 月.
- [招待 10] 清田 陽司: Hadoop による大規模分散データ処理, 産業技術大学院大学 InfoTalk 第 16 回, 産業技術大学院大学, 東京都, 2010 年 3 月.

## 受賞関連

- [受賞 1] 楊 斌, 中川裕志: 情報処理学会 50 周年記念大会 学生奨励賞: Anti-Collusion Privacy-Preserving Data mining, 5ZE-3, 2010 年 3 月 11 日.
- [受賞 2] 松島 慎, 清水伸幸, 吉田和弘, 二宮 崇, 中川裕志: 情報処理学会 50 周年記念大会 学生奨励賞: 多クラス識別問題におけるオンライン学習のための厳密な PA アルゴリズム, 3W-7, 2010 年 3 月 10 日.
- [受賞 3] 森井 正覚, 楊 斌, 佐藤 一誠, 中川 裕志: DEIM2010(第 2 回データ工学と情報マネジメントに関するフォーラム, 第 8 回日本データベース学会年次大会) 研究発表奨励賞: 複数グラフに対するプライバシー保護を考慮した統合的リンク解析. B5-1, 2010 年 3 月 1 日.
- [受賞 4] 吉田稔, 中川裕志: 人工知能学会 2009 年度全国大会優秀賞: 吉田稔, 中川裕志. Wikiwi: テキストマイニングによる Wikipedia 検索支援, 1C3-3, 2009 年 10 月.
- [受賞 5] 佐藤一誠, 吉田稔, 中川裕志: 人工知能学会 研究会優秀賞, 多重性を考慮したノンパラメトリックベイズグラフクラスタリングによる単語間関係抽出, SIG-DMSM-A801-07, 2009 年 4 月.
- [受賞 6] 江原 遥, 二宮 崇, 中川 裕志: NLP 若手の会 第 4 回シンポジウム 奨励賞, NLP 若手の会, 2009 年 9 月.

## 査読付論文

- [査読付 1] 吉田稔, 中川裕志, 寺田昭: コーパス検索支援のための動的同義語候補抽出, 人工知能学会論文誌, 25(1), pp.122—132. 2010 年.
- [査読付 2] Yoji Kiyota, Hiroshi Nakagawa, Satoshi Sakai, Tatsuya Mori and Hidetaka Masuda: Exploitation of the Wikipedia Category System for Enhancing the Value of LCSH, Joint Conference of Digital Library: JCDL 2009, pp.411—412, 2009.
- [査読付 3] Nobuyuki Shimizu and A. Haas: Learning to follow navigational route instructions, in Proceedings of the Twenty-first International Joint Conference on Artificial Intelligence (IJCAI 2009), pp. 1488—1493, 2009.

- [査読付 4] Issei Sato, Kenichi Kurihara, Shu Tanaka, Seiji Miyashita and Hiroshi Nakagawa: Quantum Annealing for Variational Bayes Inference, The 25th Conference on Uncertainty in Artificial Intelligence (UAI2009), <http://www.cs.mcgill.ca/~uai2009/proceedings.html>, Montreal, Canada, June 18—21, 2009.
- [査読付 5] Takayuki Yoshinaka, Tomohiro Fukuhara, Hidetaka Masuda, Hiroshi Nakagawa: User-Oriented Splog Filtering Based A User-Oriented Splog Filtering Based on a Machine Learning, BlogTalk 2009 The 6th International Conference on Social Software, Juje, Korea, Sept. 15—16, 2009.
- [査読付 6] Xiaorong Fan, Nobuyuki Shimizu and Hiroshi Nakagawa: Automatic Extraction of Bilingual Terms From A Chinese-Japanese Parallel Corpus, 3rd International Universal Communication Symposium (IUCS 2009), Session 1: Machine Translation 4, Dec 3-4, 2009.
- [査読付 7] Hiroyuki Nakasaki, Yusuke Abe, Takehito Utsuro, Yasuhide Kawada, Tomohiro Fukuhara, Noriko Kando, Masaharu Yoshioka, Hiroshi Nakagawa and Yoji Kiyota: Cross-Lingual Analysis of Concerns and Reports on Crimes in Blogs, Mining User-Generated Content for Security (MINUCS2009), Venice Italy, Dec.9, 2009.
- [査読付 8] Yo Ehara, Nobuyuki Shimizu, Takashi Ninomiya and Hiroshi Nakagawa: Personalized Reading Support for Second-language Web documents by Collective Intelligence, In the Proceedings of the 2010 International Conference on Intelligent User Interfaces (IUI2010), pp. 51—60, 2010.
- [査読付 9] Ayako Hoshino and Hiroshi Nakagawa: Predicting the Difficulty of Multiple-Choice Close Questions for Computer-Adaptive Testing, Research in Computing Science, Special issue: Natural Language Processing and its Applications, ISSN 1870—4069, pp. 279-292, the National Polytechnic Institute, Mexico, 2010.

## 公開ソフトウェア

- [公開 1] 小野真吾, 吉田稔, 中川裕志: “NAYOSE” Web の人名検索の結果をオンデマンドで同一人物ごとにまとめあげるクラスタリングを行うシステム, URL: <http://ianua7.r.dl.itc.u-tokyo.ac.jp:8080/nayose/servlet/Nayose>, (2010年更新).
- [公開 2] 江原遥: SocialDict, 英文の読解支援システム (<http://www.socialdict.com/>), 2009年11月.
- [公開 3] 江原遥: NEWikipedia, 英語版 Wikipedia の読解支援システム (<http://en.newwikipedia.org/>), 2010年3月.

## その他の発表論文

- [発表 1] 中野 幹生, 緒方 淳, 清田 陽司, 東中 竜一郎, 翠 輝久: 【パネル討論】音声インタフェースにおける Web テキスト処理技術の利用, 情報処理学会 第 191 回自然言語処理研究会 (NL-191-12), 東京工業大学, 東京都, 2009年5月.
- [発表 2] 野田 陽平, 清田 陽司, 中川 裕志: Wikipedia からの意外性のある情報の抽出, NLP 若手の会 第 4 回シンポジウム, 京都大学, 京都市, 2009年9月.
- [発表 3] 森 竜也, 増田 英孝, 中川 裕志, 清田 陽司: 多言語に展開する Wikipedia の特性調査, NLP 若手の会 第 4 回シンポジウム, 京都大学, 京都市, 2009年9月.
- [発表 4] 中崎 寛之, 阿部 佑亮, 宇津呂 武仁, 河田 容英, 福原 知宏, 神門 典子, 吉岡 真治, 中川 裕志, 清田 陽司: 特定トピックの日英ブログ収集・分析・類型化: 事例研究, 情報処理学会 第 194 回自然言語処理研究会 (NL-194-9), 愛媛大学, 松山市, 2009年11月.

- [発表 5] 佐藤 由紀, 横本 大輔, 中崎 寛之, 宇津呂 武仁, 吉岡 真治, 福原 知宏, 神門 典子, 中川 裕志, 清田 陽司: Wikipedia を介した関連ニュース・ブログの対応付け: Wikipedia エントリの分析, 情報処理学会 第 194 回自然言語処理研究会 (NL-194-10), 愛媛大学, 松山市, 2009 年 11 月.
- [発表 6] 坂井 哲, 増田 英孝, 清田 陽司, 中川 裕志: 国立国会図書館リサーチ・ナビにおけるテーマグラフの生成, 情報処理学会 第 96 回情報学基礎研究会 (FI-96-5), 秋葉原ダイビル, 東京都, 2009 年 11 月.
- [発表 7] 森 竜也, 増田 英孝, 清田 陽司, 中川 裕志: Wikipedia を活用した言語間差異比較システムの提案, 第 2 回データ工学と情報マネジメントに関するフォーラム/第 8 回日本データベース学会年次大会(DEIM2010), 淡路夢舞台国際会議場, 兵庫県, 2010 年 3 月.
- [発表 8] 野田 陽平, 清田 陽司, 中川 裕志: ネットワーク構造を利用した Wikipedia からの意外性のある情報の抽出, 情報処理学会 第 72 回全国大会 講演論文集 (3K-4), 東京大学, 東京都, 2010 年 3 月.
- [発表 9] 坂井 哲, 増田 英孝, 清田 陽司, 中川 裕志: Wikipedia と図書館情報資源による調べ方自動提示システム, 情報処理学会 第 72 回全国大会 講演論文集 (3K-5), 東京大学, 東京都, 2010 年 3 月.
- [発表 10] 森 竜也, 増田 英孝, 中川 裕志, 清田 陽司: Wikipedia における言語間の差異マイニング, 情報処理学会 第 72 回全国大会 講演論文集 (5ZN-6), 東京大学, 東京都, 2010 年 3 月.
- [発表 11] Ayako Hoshino, Hiroshi Nakagawa: Predicting the Difficulty of Multiple-Choice Cloze Questions for Computer-Adaptive Testing, 11th International Conference on Intelligent Text Processing and Computational Linguistics(CICLING2010), Iasi, Romania, Mar. 21—27, 2010.
- [発表 12] 池田雅紀, 小野真吾, 佐藤一誠, 吉田稔, 中川裕志: Web 上の人名検索結果の同姓同名問題における二段階クラスタリングを用いた再現率向上, 第 148 回データベースシステム・第 95 回情報学基礎合同研究発表会論文集. 2009-FI-95(6). 2009 年 7 月.
- [発表 13] 野田陽平, 清田陽司, 中川裕志: 意外性のある知識発見のための Wikipedia カテゴリ間の関係分析. 人工知能学会研究会資料 セマンティックウェブとオントロジー研究会 第 20 回 Wikipedia ワークショップ. 東京大学. 2009 年 12 月.
- [発表 14] 森竜也, 増田英孝, 清田陽司, 中川裕志: Wikipedia エントリ構造抽出ツール: Wik-IE, 人工知能学会研究会資料 セマンティックウェブとオントロジー研究会 第 20 回 Wikipedia ワークショップ. 東京大学, 2009 年.
- [発表 15] 佐藤一誠, 中川裕志: Latent Dirichlet Allocation における決定論的オンラインベイズ学習, 情報処理学会 第 193 回自然言語処理研究会, 2009 年 9 月.
- [発表 16] 松島慎, 清水伸幸, 吉田和弘, 二宮崇, 中川裕志: サポートクラスによる Passive-Aggressive アルゴリズムの多クラス化, 情報処理学会 第 192 回自然言語処理研究会, 2009-NL-192(12), pp. 1—6, 北見, 2009 年 7 月.
- [発表 17] 江原遥, 二宮崇, 清水伸幸, 中川裕志: 英語版 Wikipedia を対象としたユーザが知らない語を予測する読解支援システム, 言語処理学会 第 16 回年次大会 発表論文集, pp.864—867, 東京, 2010 年 3 月.
- [発表 18] 佐藤一誠, 中川裕志: Latent Dirichlet Allocation の量子アニーリング変分ベイズ学習, IBIS2009, 2009 年 11 月.



- [発表 19] 前田 朗, 中川裕志: 東京大学 OPAC Plus "言選 Web" - 関連学術用語による日本語文献情報への簡易ナビゲーションシステム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, デ-01, 2010 年 3 月.
- [発表 20] 江原 遥, 二宮 崇, 清水伸幸, 中川裕志: Web 文書中のユーザが知らない語を予測す読解支援システム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 5D-7, 2010 年 3 月.
- [発表 21] 松島 慎, 清水伸幸, 吉田和弘, 二宮 崇, 中川裕志: 多クラス識別問題におけるオンライン学習のための厳密な PA アルゴリズム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 3W-7, 2010 年 3 月.
- [発表 22] Xiaorong Fan, 清水伸幸, 中川裕志: Word Alignment Based Bilingual Terminology Extraction from a Chinese-Japanese Parallel Corpus, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 4W-6, 2010 年 3 月.
- [発表 23] 池田雅紀, 佐藤一誠, 吉田 稔, 中川裕志: 半教師有り学習に基づく Web 上の人物クラスタリングシステム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 6ZC-1, 2010 年 3 月.
- [発表 24] 楊 斌, 中川裕志: Anti-Collusion Privacy-Preserving Data mining, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 5ZE-3, 2010 年 3 月.
- [発表 25] 廣川敬真, 吉田稔, 山田剛一, 増田英孝, 中川裕志: 業種別による新聞記事と株価動向関係の解析, 言語処理学会第 16 回年次大会発表論文集, pp1070—1073, 東京, 2010 年 3 月.
- [発表 26] 吉田稔, 中川裕志, 松井藤五郎, 和泉潔, 石田智也, 中嶋啓浩: 取引高とニュース記事関連性の分析, 第 4 回ファイナンスにおける人工知能応用研究会(SIG-FIN), pp60—64, 東京, 2010 年 1 月.
- [発表 27] 吉田稔, 佐藤一誠, 中川裕志, 寺田昭: 接尾辞配列とディリクレ過程混合モデルを用いたテキスト中の数値表現マイニング, 情報処理学会第 96 回情報学基礎研究会, Vol.2009-FI-96.No.4, 東京, 2009 年 11 月.
- [発表 28] 吉田稔, 中川裕志: 確率モデルを利用した Web 文書からの見出し抽出, NLP 若手の会第 4 回シンポジウム, 2009 年 9 月.
- [発表 29] 吉田稔, 中川裕志: Wikiwi テキストマイニングによる Wikipedia 検索支援, 人工知能学会第 23 回全国大会, IC3-3, 高松, 2009 年 6 月.
- [発表 30] 吉田稔, 杉浦隆博, 山田剛一, 増田英孝, 中川裕志: テキストからの数値抽出による自動グラフ作成, 人工知能学会第 23 回全国大会, 3F2-NFC3-2, 高松, 2009 年 6 月.
- [発表 31] 吉田和弘, 吉田稔, 中川裕志: 文字列検索に基づく同義語・類義語抽出ツールとその性能評価, 第 191 回自然言語処理・第 76 回音声言語情報処理合同研究会, VolINL191/SLP76, No19, 東京, 2009 年 5 月.
- [発表 32] 松島慎, 清水伸幸, 佐藤一誠, 二宮崇, 中川裕志: SPA アルゴリズムの半教師あり学習への応用, 情報処理学会 第 76 回数理モデル化と問題解決研究会, 2009-MPS-76(21), pp. 1—8, 東京, 2009 年 12 月.
- [発表 33] 江原遥, 二宮崇, 中川裕志: Web 文書中の単語クリックログの解析から未知単語を予測する語義注釈システム, 情報処理学会 第 193 回自然言語処理研究会, 2009-NL-193(3), pp. 1—7, 京都, 2009 年 9 月.
- [発表 34] 江原遥, 二宮崇, 中川裕志: 語義注釈システムの単語クリックログからの言語能力情報の抽出, NLP 若手の会 第 4 回シンポジウム, 京都, 2009 年 9 月.

[発表 35] Masaki Ikeda, Shingo Ono, Issei Sato, Minoru Yoshida and Hiroshi Nakagawa: Person Name Disambiguation on the Web by TwoStage Clustering, 2nd Web People Search Evaluation Workshop (WePS 2009), 18th WWW Conference, Madrid, Spain, April 21, 2009.

### 特記事項

[特記 1] 清田 陽司: Google はなぜ的確に探せるのか(Web 連載), 第 5 回『悟空、秘技「分かち書き」を習う』/第 6 回『悟空、単語辞書を手に入れる』, ASCII.jp, アスキーメディアワークス, 2009 年 4 月～7 月.

[特記 2] 科学研究費補助金特定領域研究「情報爆発IT基盤」報告書(代表者) A01-05. 2009.

[特記 3] 吉田稔: 木接合文法, 言語処理学事典, 言語処理学会編, 共立出版, pp. 160—161, 2009 年 12 月.

[特記 4] 中川裕志: 言語処理概観: 古代から現代へ, 言語処理学事典, 言語処理学会編, 共立出版, pp. 21—24, 2009 年 12 月.

[特記 5] 二宮 崇: 単一化文法, 言語処理学事典, 言語処理学会編, 共立出版, pp. 158—159, 2009 年 12 月.

[特記 6] 二宮 崇: 深い構文解析, 言語処理学事典, 言語処理学会編, 共立出版, pp. 176—177, 2009 年 12 月.

[特記 7] 二宮 崇: 文法獲得, 言語処理学事典, 言語処理学会編, 共立出版, pp. 182—183, 2009 年 12 月.

### 報道関連

[報道 1] 日本経済新聞 2009 年(平成 21 年)5 月 13 日 13 面: 東大発 VB のリッテル、商品検索簡単に.

# 統計的機械学習とそのテキスト処理への応用

中川裕志

## 1 概要

2009 年度は、統計的機械学習、テキスト処理および Web 情報資源などに関するいくつかのテーマについて研究を行った。具体的には、以下に示すテーマの研究を行った。

- Web のサーチエンジンで人名検索を行った結果を異なる人物毎にクラスタリングする名寄せシステム。この研究においては競争型タスク WePS2 に参加し、全世界 17 参加チーム中 2 位の成績であった。[査読付 3] [発表 2] [発表 15]なお、この成果をまとめた論文が ACM SIGIR2010 に採録された。
- 重要な Web2.0 的情報資源である Wikipedia からの情報抽出に関する研究を行った。具体的には言語毎の比較研究[発表 4] [発表 16]、図書館情報検索用のターミノロジー抽出、[査読付 2] [発表 14]、意外性がある情報の発見[発表 3] [発表 13]に関する研究を行った。
- 中川研究室で開発した重要語抽出システム「言選 Web」を論文検索に応用する研究を行った。[発表 9]
- コーパスからのオンデマンド型同義語抽出システムを開発した[査読付 1] [受賞 2]。これは一昨年に行った情報大航海の共通技術プロジェクトの一環としての研究成果である。
- 多言語ブログからの意見抽出を筑波大学、情報学研究所、北大との共同研究で行った[査読付 7]。また、東京電機大学との共同研究として、スブログのフィルタリングの研究を行った[査読付 5]。
- 多重トピック文書の確率的生成モデルとして、量子トンネル効果のアナロジーでシミュレーテッドアニーリングを行う変分ベイズ法を提案した。このアルゴリズムは並列クラスタリングにも適用でき、機械学習のトップコンファレンスである UAI に採択された[査読付 4]。
- 複数のサイトからなるネットワーク環境で、自分のサイトのデータを公開せずにネットワーク全体のデータを利用してデータマイニングを行うプライバシー保護データマイニング (PPDM) の研究において、結託攻撃に耐える PPDM プロトコル[受賞 5] [発表 17]、およびクラスタリングを行う新規な PPDM プロトコルを提案した。[受賞 3]
- PassiveAggressive アルゴリズムと呼ばれるオンライン学習アルゴリズムにおいて、多クラス分類アルゴリズムの研究を行った。ここでは、処理中のデータが属するクラスの素性だけを更新する従来手法を原理的に改善し、処理中のデータが属する以外で関連するクラスを SVM のサポートベクトルに倣ってサポートクラスと名付け、これを数学的に求めるアルゴリズムを開発した。その結果、多クラス分類精度が大幅に改善した[受賞 4] [発表 6] [発表 11]。この成果はデータマイニングにおけるトップコンファレンスである SIAM SDM に採択された。また、オンライン LDA 手法を提案し、評価した[発表 5]。
- 日中対訳コーパスを用いた専門用語対訳の自動抽出の研究を行った[査読付 6] [発表 12]。これは、振興調整費「日中中日言語処理研究」の一環として行われた。

- Web 英文テキストを利用した英語穴埋め 4 択問題の自動生成システム開発[査読付 9] [発表 1] および利用者が知らない単語をオンライン学習によって学習し、その訳語を提示するシステム開発[査読付 8] [発表 7] [発表 10]などの、Web テキストを題材にして英語学習支援の研究を行った。

紙数の関係から、以下では、Wikipedia から意外性のある知識を発見する手法と実験的評価の研究について詳述する。

## 2 意外性のある知識発見のための Wikipedia カテゴリ間の関係分析

### 2.1 概要

本研究の目的は、Web 上の情報の信頼性と有益性の評価であるが、評価が定まりつつある Wikipedia の情報を対象にすることによって当面の信頼性を確保する。問題はむしろ有益性である。専門家は自分の専門分野については熟知しているから、既に自身の専門分野内の有益な情報は知っていると考えられるべきであろう。むしろ、自分の専門分野からかなり離れた分野における知識で自分の分野とつながりがある知識があれば、それらは有益と考えるであろう。専門家でない場合にも同じようなことが考えられ、概念的には遠く離れ、かつ関係が薄いとみなされていた 2 つの分野のかけ橋になるような知識が発見できれば、有益な知識を見出したと考えるだろう。言い換えれば、意外性のある知識ということになる。この研究では、上記のような考察により、Wikipedia において意外性のある記事を探すことを目的とする。

英語版 Wikipedia は 2008 年 8 月 11 日に 250 万記事、日本語版 Wikipedia は 2008 年 6 月 25 日に 50 万項目を超え、膨大な情報量を誇る百科事典として広く認知されている。日本語版 Wikipedia は 9 個の主要カテゴリの下に、サブカテゴリ、記事が関連付けられており、大規模なグラフ構造を成している。各項目はそれぞれ複数の親カテゴリを持っており、また、同義語はリダイレクトとして関係付けられている。Wikipedia の記事は、カテゴリシステムによってさまざまな観点からの分類がなされている。この特徴をうまく用いると、個別の記事からだけでは得られない意外な知識の発見につなげることができる。例えば、「オープンコーラ」という概念がある。「コーラ」は万人周知の飲料だが、その製造方法がオープンではない。一方、「オープンソース」という概念はソフトウェアの世界では良く知られたもので、コード内容が公開されている。この二つの概念、すなわち、飲料とソフトウェアという離れた概念のかけ橋として、製造方法が公開されているコーラという意味で「オープンコーラ」という概念があり、これは意外性が高いと考えられる。

以下では、このように直観的に規定した「意外性」という概念を、Wikipedia のカテゴリが持つネットワーク構造を利用して定義し、その定義に沿って意外な知識を Wikipedia から発掘する手法とシステムについて述べる。

### 2.2 利用する特徴情報

我々は、Wikipedia のカテゴリネットワークの構造情報のうち、上記の直観を実現するために有効であると問がられる以下の情報 (feature) に着目した。

- 各項目が属するカテゴリの数(親カテゴリの数)
- 各カテゴリが持つ子項目の数
- 各項目に関する、共通の親を 1 つ以上持つ項目の数(兄弟項目の数)
- 各カテゴリセットが持つ子の数
- 2 つのカテゴリ A,B の間の距離、すなわち A,B の共通親カテゴリを軽油した場合の A,B 間の最短パス長



### 2.3 機械学習による識別

上記の特徴情報を用いて意外性のある情報を発見する手法として第一に検討すべきものは、意外な記事の集合とそうでない記事の集合という教師データを作成し、SVM のような教師あり学習によって識別器を学習する方法である。教師データとしては 180 記事の意外性のある記事と、そうでない記事 180 記事を手で作成した。また、実際は意外性のない記事が圧倒的に多数なので、ランダムに記事を選択して意外でない記事とみなし、意外性なしの教師データの数を増やして実験してみた。SVM でこれらの教師データで学習し、4 分割の考査検定を行った結果を次の表に示す。

表1(意外性があると人手判断したデータ) / (意外性があると SVM が判定したデータ)

意外性のあるデータ(左)とないデータ(右)の数	180:180	180:1800	180:18000
線形カーネル	90.70%	4.17%	0%
2次多項式カーネル	90.70%	8.33%	0%
3次多項式カーネル	90.70%	12.50%	0%
ガウシアンカーネル	88.37%	4.17%	0%

この結果からみると、意外性のある記事が少数派(手作業の経験では 1000 分の 1 以下)の状況においては SVM による識別の方法は機能しないことが判明した。

### 2.4 回帰による順位付け

このように、意外性の有無で Wikipedia の記事を分類することが判明したので、意外性の程度によって記事を順位付ける方法を検討する。統計的機械学習の観点からすれば、識別ではなく回帰によるモデルを考えることになる。意外性に関する正例と負例は feature を各次元を持つ多次元空間中に分布する。そこで、これに意外性の有無を表す次元として有=1、無=0の値を持つ次元を加える。具体的には n 種類の feature で表現される各記事は次のようなベクトルで表現される。

意外性のあるデータ:  $(1, \text{feature}_1, \dots, \text{feature}_n)$

意外性のないデータ:  $(0, \text{feature}_1, \dots, \text{feature}_n)$

このデータを 1 次の直線で回帰すると以下のようなベクトル空間上でのイメージとなる。つまり、赤い円で囲まれた意外性のある記事と青い円で囲まれた意外性のない記事を上記にベクトルで表して線形回帰すると紫の回帰直線が得られる。図中に橙色の円で表現された新規の記事をこの紫色の回帰直線に射影し、その意外性を表す軸(図の縦方向の軸)の値を得れば、これがこの記事の意外性の度合いを表す数値を与える。

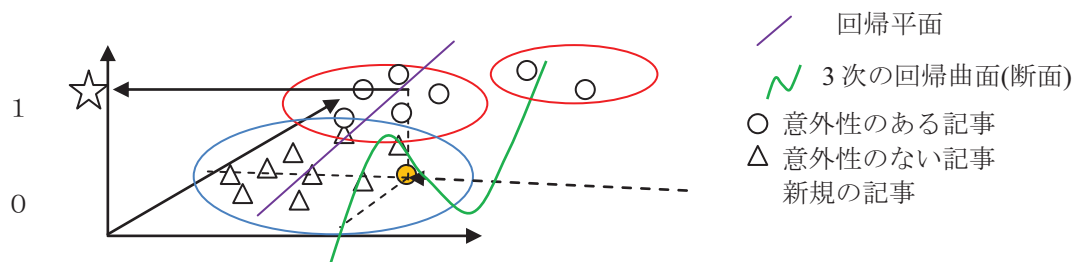


図1 意外性を表すベクトル空間における回帰

これは、1 次の線形回帰だが、3 次の曲線による回帰を用いれば緑色のような曲線で回帰するため、意外性の有無の記事がより複雑な状況に散在しても対処できる可能性がある。このような方法で具体的にデータを処理した結果の例として、マイケルジャクソンの例を以下に示す。

例:カテゴリ:マイケルジャクソンに対して、それと遠い別のカテゴリおよび両方のカテゴリに共通する記事を調べてみた。意外性のある記事数とない記事数の比を1:1, 1:10, 1:100の3種類の場合で線形回帰および3次関数で曲面回帰した場合の上位3位までを下の表に示す。

線形回帰した場合

	順位	カテゴリA	カテゴリB	AB共通の記事
意外:非意外 =1:1	1	マイケル・ジャクソン	エルヴィス・プレスリー	リサ・マリー・プレスリー
	2	マイケル・ジャクソン	ダンステクニック	ムーンウォーク
	3	マイケル・ジャクソンのアルバム	マイケル・ジャクソン	マイケル・ジャクソン
意外:非意外 =1:10	1	福祉	マイケル・ジャクソン	USA フォー・アフリカ
	2	億万長者	マイケル・ジャクソン	ジェイ・レノ
	3	ダンス	マイケル・ジャクソン	ムーンウォーク
意外:非意外 =1:100	1	福祉	マイケル・ジャクソン	USA フォー・アフリカ
	2	億万長者	マイケル・ジャクソン	ジェイ・レノ
	3	ダンス	マイケル・ジャクソン	ムーンウォーク

3次関数で曲面回帰した場合

	順位	カテゴリA	カテゴリB	AB共通の記事
意外:非意外 =1:1	1	福祉	マイケル・ジャクソン	USA フォー・アフリカ
	2	LGBTの人物	マイケル・ジャクソン	リサ・マリー・プレスリー
	3	ダンス	マイケル・ジャクソン	ムーンウォーク
意外:非意外 =1:10	1	福祉	マイケル・ジャクソン	USA フォー・アフリカ
	2	LGBTの人物	マイケル・ジャクソン	リサ・マリー・プレスリー
	3	ダンス	マイケル・ジャクソン	ムーンウォーク
意外:非意外 =1:100	1	福祉	マイケル・ジャクソン	USA フォー・アフリカ
	2	LGBTの人物	マイケル・ジャクソン	リサ・マリー・プレスリー
	3	ダンス	マイケル・ジャクソン	ムーンウォーク

これらの例で見られるように、回帰による方法は、意外性のある記事数とない記事数の比によらず安定した結果を得られる。特に3次曲線での回帰は安定度が高い。これ意外の順位においても同様の安定度が得られた。

### 3 成果要覧

#### 受賞関連

[受賞 1] 人工知能学会 研究会優秀賞: 佐藤一誠, 吉田稔, 中川裕志. 多重性を考慮したノンパラメトリックベイズグラフクラスタリングによる単語間関係抽出, SIG-DMSM-A801-07(7/24), 2009年4月.

[受賞 2] 人工知能学会 2009年度全国大会優秀賞: 吉田稔, 中川裕志. Wikiwi: テキストマイニングによるWikipedia検索支援, 1C3-3, (6/17), 2009年10月.

[受賞 3] DEIM2010(第2回データ工学と情報マネジメントに関するフォーラム, 第8回日本データベース学会年次大会) 研究発表奨励賞: 森井 正覚, 楊 斌, 佐藤 一誠, 中川 裕志. 複数グラフに対するプライバシー保護を考慮した統合的リンク解析. B5-1, 2010年3月1日.

[受賞 4] 情報処理学会 50周年記念大会 学生奨励賞: 松島 慎, 清水伸幸, 吉田和弘, 二宮 崇, 中川裕志. 多クラス識別問題におけるオンライン学習のための厳密な PA アルゴリズム, 3W-7, 2010年3月10日.

[受賞 5] 情報処理学会 50 周年記念大会 学生奨励賞:楊 斌. Anti-Collusion Privacy-Preserving Data mining, 5ZE-3, 2010 年 3 月 11 日.

### 査読付き論文リスト

[査読付 1] 吉田稔, 中川裕志, 寺田昭: コーパス検索支援のための動的同義語候補抽出, 人工知能学会論文誌, 25(1), pp.122-132, 2009.

[査読付 2] Yoji Kiyota, Hiroshi Nakagawa, Satoshi Sakai, Tatsuya Mori and Hidetaka Masuda: Exploitation of the Wikipedia Category System for Enhancing the Value of LCSH, Joint Conference of Digital Library:JCDL 2009, pp.411-412, 2009.

[査読付 3] Masaki Ikeda, Shingo Ono, Issei Sato, Minoru Yoshida and Hiroshi Nakagawa: Person Name Disambiguation on the Web by TwoStage Clustering, 2nd Web People Search Evaluation Workshop (WePS 2009), 18th WWW Conference, Madrid, Spain, April 21, 2009.

[査読付 4] Issei Sato, Kenichi Kurihara, Shu Tanaka, Seiji Miyashita and Hiroshi Nakagawa: Quantum Annealing for Variational Bayes Inference, The 25th Conference on Uncertainty in Artificial Intelligence (UAI2009) <http://www.cs.mcgill.ca/~uai2009/proceedings.html> , in Montreal, Canada, June 18-21, 2009.

[査読付 5] Takayuki Yoshinaka, Tomohiro Fukuhara, Hidetaka Masuda, Hiroshi Nakagawa: User-Oriented Splog Filtering Based A User-Oriented Splog Filtering Based on a Machine Learning, BlogTalk 2009 The 6th International Conference on Social Software, Juje, Korea, Sept. 15-16, 2009.

[査読付 6] Xiaorong Fan, Nobuyuki Shimizu and Hiroshi Nakagawa: Automatic Extraction of Bilingual Terms From A Chinese-Japanese Parallel Corpus, 3rd International Universal Communication Symposium (IUCS 2009), Session 1: Machine Translation 4, Dec 3-4, 2009.

[査読付 7] Hiroyuki Nakasaki, Yusuke Abe, Takehito Utsuro, Yasuhide Kawada, Tomohiro Fukuhara, Noriko Kando, Masaharu Yoshioka, Hiroshi Nakagawa, Yoji Kiyota: Cross-Lingual Analysis of Concerns and Reports on Crimes in Blogs, Mining User-Generated Content for Security (MINUCS2009), Venice Italy, Dec.9, 2009.

[査読付 8] Yo Ehara, Takashi Ninomiya, Nobuyuki Shimizu, Hiroshi Nakagawa: Personalized Reading Support for Second-language Web documents by Collective Intelligence, 2010 International Conference on Intelligent User Interfaces (IUI2010), Hong Kong, China, Feb.7-10, pp.51-60, 2010.

[査読付 9] Ayako Hoshino and Hiroshi Nakagawa: Predicting the Difficulty of Multiple-Choice Close Questions for Computer-Adaptive Testing, Research in Computing Science, Special issue: Natural Language Processing and its Applications, ISSN 1870-4069, pp. 279-292, the National Polytechnic Institute, Mexico, Mar. 2010.

### その他の発表論文リスト

[発表 1] Ayako Hoshino, Hiroshi Nakagawa: Predicting the Difficulty of Multiple-Choice Cloze Questions for Computer-Adaptive Testing, 11th International Conference on Intelligent Text Processing and Computational Linguistics(CICLING2010), Iasi, Romania, Mar. 21-27, 2010.

[発表 2] 池田雅紀, 小野真吾, 佐藤一誠, 吉田稔, 中川裕志: Web 上の人名検索結果の同姓同名問題における二段階クラスタリングを用いた再現率向上, 第 148 回データベースシステム・第 95 回情報学基礎合同研究発表会論文集, 2009-FI-95(6), 2009 年.

[発表 3] 野田陽平, 清田陽司, 中川裕志: 意外性のある知識発見のための Wikipedia カテゴリ間の関係分析, 人工知能学会研究会資料 セマンティックウェブとオントロジー研究会 第 20 回 Wikipedia ワークショップ, 東京大学, 2009 年.

- [発表 4] 森竜也, 増田英孝, 清田陽司, 中川裕志: Wikipedia エントリ構造抽出ツール: Wik-IE, 人工知能学会研究会資料 セマンティックウェブとオントロジー研究会 第 20 回 Wikipedia ワークショップ, 東京大学, 2009 年.
- [発表 5] 佐藤一誠, 中川裕志: Latent Dirichlet Allocation における決定論的オンラインベイズ学習, 情報処理学会 第 193 回自然言語処理研究会, 2009 年.
- [発表 6] 松島慎, 清水伸幸, 吉田和弘, 二宮崇, 中川裕志: サポートクラスによる Passive-Aggressive アルゴリズムの多クラス化, 情報処理学会 第 192 回自然言語処理研究会, 2009-NL-192(12), pp. 1-6, 2009 年.
- [発表 7] 江原遥, 二宮崇, 中川裕志: 機械学習による自動辞書引きを利用した英文の読解支援システム, 言語処理学会 第 15 回年次大会 発表論文集, pp. 885-888, 2009 年.
- [発表 8] 佐藤一誠, 中川裕志: Latent Dirichlet Allocation の量子アニーリング変分ベイズ学習, IBIS2009, 2009 年.
- [発表 9] 前田 朗, 中川裕志: 東京大学 OPAC Plus "言選 Web" - 関連学術用語による日本語文献情報への簡易ナビゲーションシステム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, デ-01, 2009 年.
- [発表 10] 江原 遥, 二宮 崇, 清水伸幸, 中川裕志: Web 文書中のユーザが知らない語を予測す読解支援システム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 5D-7, 2009 年.
- [発表 11] 松島 慎, 清水伸幸, 吉田和弘, 二宮 崇, 中川裕志: 多クラス識別問題におけるオンライン学習のための厳密な PA アルゴリズム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 3W-7, 2009 年.
- [発表 12] Xiaorong Fan, 清水伸幸, 中川裕志: Word Alignment Based Bilingual Terminology Extraction from a Chinese-Japanese Parallel Corpus, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 4W-6, 2009 年.
- [発表 13] 野田陽平, 清田陽司, 中川裕志: ネットワーク構造を利用した Wikipedia からの意外性のある情報の抽出, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 3K-4, 2009 年.
- [発表 14] 坂井 哲, 増田英孝, 清田陽司, 中川裕志: Wikipedia と図書館情報資源による調べ方自動提示システム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 3K-5, 2009 年.
- [発表 15] 池田雅紀, 佐藤一誠, 吉田 稔, 中川裕志: 半教師有り学習に基づく Web 上の人物クラスタリングシステム, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 6ZC-1, 2009 年.
- [発表 16] 森 竜也, 増田英孝, 中川裕志, 清田陽司: Wikipedia における言語間の差異マイニング, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 5ZN-6, 2009 年.
- [発表 17] 楊 斌, 中川裕志: Anti-Collusion Privacy-Preserving Data mining, 情報処理学会創立 50 周年記念(第 72 回)全国大会, 5ZE-3, 2009 年.

## 特記事項

- [特記 1] 科学研究費補助金特定領域研究「情報爆発IT基盤」報告書(代表者) A01-05, 2009 年.

# 自然言語処理のための基礎理論と応用

## —オンライン学習と英文読解支援—

二宮 崇

### 1 概要

本年度は、自然言語処理の基礎理論と応用に関する研究を行い、具体的には、オンライン学習の基礎理論および英文読解支援の研究を行った。オンライン学習は、データが逐次的に入力され、逐次的に最適化を行う学習手法の総称であり、非オンライン学習に比べ精度の面で劣るものの学習速度とメモリ効率の点で非常に優れていると考えられている。特に随時データを受け取ることを想定するストリーム処理における学習に適しており、世の中に大量に存するウェブテキストやセンシングデータに対する機械学習手法として期待されている。本研究では、非オンライン学習と同等以上の精度を達成するオンライン学習手法を実現した。また、学術情報の理解を支援するために、ウェブブラウザ上に表示される英文に対し訳語を自動的に付与する英文読解支援システムの研究も行った。具体的には、下記の研究内容について報告する。

**Passive-Aggressive 戦略の厳密解法によるオンライン学習** [受賞 2][発表 1][発表 2][発表 3] 本研究では Passive-Aggressive 戦略と呼ばれるオンライン学習の枠組における厳密解法による新しいオンライン学習手法を提案する。従来のアルゴリズムは Passive-Aggressive 戦略に対する緩和問題や近似解により導出されてきたが、本研究ではこの戦略における厳密解を閉じた式で導出することにより、新しいオンライン学習アルゴリズムを導出した。我々の提案するアルゴリズムは非オンライン学習よりも良い学習速度、メモリ効率を実現しながら、同時に非オンライン学習と同程度の精度を達成することに成功した。

**英文読解支援システム** [受賞 1][査読付 1][公開 1][公開 2][発表 4][発表 5][発表 6][発表 7] 英文読解を支援するために、自動的に辞書を引いて語義のアノテーションを与えるシステムを提案する。事前に与えられた利用者の語彙力に関する情報を訓練データとして識別器に与えることにより、利用者にとって未知の語彙を特定し、利用者の英語習得レベルに適応したアノテーションを行う。

その他の活動として、言語処理学会が編集する言語処理学事典の 3 項目に寄稿した[特記 1][特記 2][特記 3]。また、社会科学研究所図書職員の前田朗氏と協力して、図書館情報係が事務局となって開催された「図書系職員のためのアプリケーション開発講習会」に講師として参加した。詳細は、教育・サービス活動の章の図書館電子化部門、図書館関係システム運用・管理の講習会・研究会開催報告を参照されたい。

### 2 Passive-Aggressive 戦略の厳密解法によるオンライン学習

#### 2.1 背景

与えられた一定量の訓練データ全体に対し最適化を行う学習手法はバッチ学習と呼ばれ、SVM やロジスティック回帰(最大エントロピー法)のバッチ学習が自然言語処理や多くの問題に対し用いられ



表 1. 導出された SPA アルゴリズム

---

```

 $w_v^{(1)} = 0 \quad (\forall v \in Y)$ 
foreach  $i = 1, 2, \dots, N$  do
   $\ell_v = [1 - (w_v^{(i)} \cdot x_i - w_{y_i}^{(i)} \cdot x_i)]_+$  ( $v \neq y_i$ );
  Compute  $j$ -th class  $\sigma(j) \in Y \setminus \{y_i\}$  in descending order of  $\ell_v$  for  $\forall j \leq K - 1$ .
   $S := \emptyset$ ;
  while  $\sum_{j=1}^{|\mathcal{S}|} \frac{\ell_{\sigma(j)}}{|\mathcal{S}|+1} < \ell_{\sigma(|\mathcal{S}|)}$  do
     $S := S \cup \sigma(|\mathcal{S}|)$ ;
  endwhile
   $\tau_v := \ell_v - \sum_{u \in S} \frac{\ell_u}{|\mathcal{S}|+1}$  ( $v \in S$ )
   $\tau_v := 0$  ( $v \notin S$ )
   $w_v^{(i+1)} = w_v^{(i)} + \sum_{u \neq y_i} \tau_u x_i$  ( $v = y_i$ )
   $w_v^{(i+1)} = w_v^{(i)} - \tau_v x_i$  ( $v \neq y_i$ )
endforeach

```

---

てきた。バッチ学習に対し、訓練データを逐次的に受け取り、逐次的に学習を行う学習手法はオンライン学習と呼ばれる。オンライン学習は、時間がたつにつれて性質が変わっていくデータに対しても適用可能であるだけでなく、訓練データを受け取る毎に簡単なパラメータ更新を行うだけでよい場合が多いため計算時間やメモリの効率が良い点、また損失関数や誤分類するデータの総数に対し理論的な上限の保障ができる点など、多くの面から近年注目されている学習方法である。従来のオンライン学習は、バッチ学習に比べ精度が低いためあまり用いられてこなかったが、近年の研究によりバッチ学習に匹敵する精度が実現されることが示されつつある。大規模データに対する非常に複雑なモデルをバッチ学習で学習するには大規模メモリを搭載する高性能なコンピュータと非常に長い計算時間が必要とされるが、オンライン学習はそれらに比べ非常に少ないメモリと計算時間で学習することが可能である。そのため、バッチ学習と同程度以上の精度を達成するオンライン学習の実現が大きく期待されている。

## 2.2 内容

Crammer らによって提案された Passive-Aggressive アルゴリズム(PA アルゴリズム)は代表的なオンライン学習アルゴリズムであるが、多クラス識別問題においては、厳密な PA アルゴリズムの枠組みから外れた近似解法が用いられてきた。本研究は本来の PA アルゴリズムの厳密解をサポートクラスという概念を用いて導出し、これらに基づく識別関数の更新を効率的に行う為のアルゴリズムを提案する。

本稿で提案する Support Class Passive-Aggressive アルゴリズムでは、データを受け取る毎に、サポートクラスと呼ぶ適当なクラスの集合を定め、このサポートクラス全体に関してパラメータ更新を行う。サポートクラスを適切に定めてやることによって、更新後の分類器は、訓練データを正しく分類できるようになる。このアルゴリズムは、受け取った訓練データを一定の-margin で正しく分類することを強制するような損失関数を設計し、Passive-Aggressive アルゴリズムの枠組みで用いることによって自然に得られる。

### 2.2.1 問題設定

多クラス線形識別関数  $h_{w^{(i)}}$  とは、 $x \in \mathbb{R}^D$  に対し  $p_i \in Y = \{1, 2, \dots, K\}$  を返す関数であり、 $K$ 本の重みベクトル  $\{w_u^{(i)}\}_{u \in Y}$  ( $w_u^{(i)} \in \mathbb{R}^D$ ) を用いて、以下のように表されるとする。

$$p_i = h_{w^{(i)}}(x_i) = \operatorname{argmax}_{u \in Y} (w_u^{(i)} \cdot x_i)$$

ここで、 $w^{(i)} = \{w_u^{(i)}\}_{u \in Y}$  とした。一般的なオンライン学習の枠組みにおいてはこの関数の学習を次の工程を繰り返すことによって行う。

表 2: PA アルゴリズムと SPA アルゴリズム

	PA	SPA
Optimization	$\min_w \frac{1}{2} \sum_{u=1}^K \ w_u - w_u^{(i)}\ ^2$ $\text{s. t. } w_{y_i} \cdot x_i - w_{p_i} \cdot x_i \geq 1$	$\min_w \frac{1}{2} \sum_{u=1}^K \ w_u - w_u^{(i)}\ ^2$ $\text{s. t. } \forall u \neq y_i, w_{y_i} \cdot x_i - w_u \cdot x_i \geq 1$
Stepsize	$\tau = \frac{1 - (w_{y_i}^{(i)} \cdot x_i - w_u^{(i)} \cdot x_i)}{2\ x\ ^2}$	$\tau = 2\ x\ ^{-2} \left( \ell_v - \frac{1}{ S  + 1} \sum_{u \in S} \ell_u \right)$
Support class	$p_i$	$\sigma(k): \sum_{j=1}^{k-1} \ell_{\sigma(j)} < k \ell_{\sigma(k)}$

1.  $i$  番目のデータ  $x_i \in X \subset \mathbb{R}^D$  を受け取る。
2.  $i$  番目の識別関数  $h_{w^{(i)}}$  によって  $x_i$  に対する予測クラス  $p_i \in Y = \{1, 2, \dots, K\}$  を計算する。
3.  $i$  番目の正解クラス  $y \in Y = \{1, 2, \dots, K\}$  を受け取り、次の識別関数  $h_{w^{(i+1)}}$  を求める。

### 2.2.2 提案手法

PA アルゴリズムの枠組みにおけるオンライン学習の設定では、 $x_i$  に対応する正解ラベル  $y_i$  を受け取り、損失  $\ell(w; x_i, y_i) \geq 0$  を計算する。PA アルゴリズムの枠組みでは、損失関数に従い、次の識別関数に関する重みベクトルを以下の最適化問題を用いて特徴づける。

$$w^{(i+1)} = \min_w \frac{1}{2} \sum_{u=1}^K \|w_u - w_u^{(i)}\|^2 \quad \text{s. t. } \ell(w; x_i, y_i) = 0 \quad (1)$$

この最適化問題の解は、損失が最小値  $0$  をとるような重みベクトルの中で最もデータを受け取った時点での重みベクトルからの変更が少ない重みベクトルとなる。Crammer らは、 $K$  クラス識別問題における PA アルゴリズムで用いる損失関数を以下のように定義している。

$$\ell(w; x, y) = \max_{u \neq y} [1 - (w_y \cdot x - w_u \cdot x)]_+$$

$[\cdot]_+$  は  $\max(\cdot, 0)$  で定義される閾値関数である。この時、最適化問題(1)は  $K - 1$  本の線形制約式を持った凸最適化問題となる。Crammer らは計算の複雑性などを考慮してこれらの制約条件を緩和した問題の最適解を用いて識別関数を更新していた。本研究においては近似解を用いず厳密な最適解を解析的に導出し、その最適解を用いて識別関数を更新する。導出された解析解は以下のようになる。

$$w_v^{(i+1)} = w_v^{(i)} + \sum_{u \in S} \tau_u x_i \quad (v = y_i)$$

$$w_v^{(i+1)} = w_v^{(i)} - \tau_v x_i \quad (v \in S)$$

$$w_v^{(i+1)} = w_v^{(i)} \quad (v \notin S)$$

この  $S$  をサポートクラス集合と呼び、 $S$  の要素クラスすべてに対し重みベクトルの更新を行う。 $S$  の要素及び  $\tau_v$  の値は簡単なアルゴリズムで計算できる。これを本稿では SPA アルゴリズムと呼び概要を表 1 に示す。また、Crammer らによる手法との違いを表 2 に示す。Crammer らによる PA-I, PA-II アルゴリズムに関しても同様に厳密解による更新アルゴリズムを導出した。これを本稿では SPA-I, SPA-II アルゴリズムと呼ぶ。

### 2.2.3 実験

20Newsgroups1、Reuters2、USPS3 のデータセットを用いて実験を行った。20Newsgroups、Reuters は文書分類、USPS は文字認識のタスクになっている。PA、PA-I、PA-II、SPA、SPA-I、SPA-II 及びパーセプトロンに対しこれらを 40 回まで反復し、各時点でのテストセットに対する識別精度の比較実験を行った。そしてこれらのオンライン学習法のアルゴリズムに加えて、2 つのバッチ学習法、多クラ

表 3: 各データセットの誤識別率(%)

	PA	PA-I	PA-II	SPA	SPA-I	SPA-II	Perc	LR	M-SVM
USPS	6.82	6.83	6.82	6.02	5.73	4.95	6.45	4.98	4.83
Reuters	3.80	3.80	3.81	3.01	3.03	2.96	3.99	3.50	3.18
News20	22.98	21.49	21.02	20.01	16.84	15.83	22.63	17.79	15.94

表 4: 各データセットにおける所要計算時間(sec)

	PA	PA-I	PA-II	SPA	SPA-I	SPA-II	Perc	LR	M-SVM
USPS	5.5	5.5	5.5	5.9	5.9	5.9	1.5	53.6	68.5
Reuters	6.2	6.2	6.2	6.4	6.4	6.4	1.0	92.5	34.3
News20	15.7	15.9	15.8	16.8	17.1	17.2	7.9	405.6	75.7

スロジスティック回帰 (表中”LR”) 及び多クラス SVM (表中”SVM”) における精度も検証し比較した。結果を表 3 に示す。実験では 10 等分における交差検証を行い、超パラメータの調節はそのうちの一つを無作為に選んで行った。またこの時の所要時間を表 4 に示した。

提案手法はどのデータセットにおいてもそれぞれの既存の手法に比べ良い性能を示している。USPS ではバッチ学習の方が良い精度を得ているが、20Newsgroups および Reuters のデータセットでは SPA-I、SPA-II アルゴリズムがバッチ学習器よりよい精度を達成した。また、SPA アルゴリズムは各バッチ手法に比べ計算時間は格段に短く、PA アルゴリズムにおける計算時間とはほとんど差がないことがわかった。

### 2.3 具体的成果

本稿では、多クラス PA アルゴリズムの効率的な厳密解法を提案した。実験結果より、従来の主な学習手法であった SVM やロジスティック回帰のバッチ学習に比べ、非常に高速に学習し、精度もほぼ同等かそれ以上の性能を示した。結果として得られた更新を用いたアルゴリズムは優良であり、今後計算コスト的にも精度面においても優良な学習のためにオンライン学習が用いられうることを示唆できたと考えられる。

## 3 英文読解支援システム

### 3.1 背景

英語学習者の英文読解を支援する方法として、英文に対し語義のアノテーションを与えることが有効であることは主にコンピュータ支援語学学習 (CALL) の分野における研究によって示唆されてきた。しかし、CALL 分野における研究はアノテーションによる教育効果を測定することが主な目的となっており、実際に学習者のレベルに応じて自動的に語義のアノテーションを与えるシステムの研究は少ない。読解支援のツールとしては、マウスカーソルを英単語にあてることにより自動的にその単語の語義を表示するシステムが存在するが、この手法はコンピュータと向き合っている間は有効であるが、本や書類として印刷する場合には利用できない。

### 3.2 内容

本研究は、ユーザの英語レベルに合わせて英文に対し語義のアノテーションを自動的に与える手法を提案、および、ウェブ上で動作するシステムを開発、公開した[公開 1][公開 2]。システムは、登録ユーザのレベルに応じて、そのユーザにとって未知であると推測される英単語に日本語訳を自動的に付与する。次はウェブ上のニュースの一文に付与された日本語訳の例である。



Flags flew at half mast across the country and public entertainment was curtailed(切り詰める,短縮する,省略する,削減する) as a mark of respect, one week after the quake hit.

灰色の背景の領域の単語は辞書に登録されている単語を示しており、薄い灰色の領域の単語は、閲覧しているユーザにとって既知であると判断された単語であり、濃い灰色の領域の単語はそのユーザにとって未知であると判断された単語である。上記の例で“curtailed”が未知と判断された単語であり、その訳は“切り詰める、短縮する、省略する、削減する”と与えられている。ユーザは、システムが間違えて未知/既知を判断したときには、その単語をクリックすることにより、修正を与えることができる。これによってシステムはユーザ毎の単語の既知/未知を学習し、ユーザが使えば使うほどより良い推測がシステムによってなされる。良い推測がされることにより、必要最小な単語訳の付与が行われるようになる。つまり、未知の単語にはより確実に単語訳が付与され、既知の単語に対しては余計な付与が減ることが期待される。

システムが行うユーザの既知/未知単語予測には項目反応理論 (Item Reponse Theory, IRT) の一つであるラッシュモデルをロジスティック回帰として拡張した確率モデルを用いた。具体的な拡張は次のとおりである。ラッシュモデルにおいて定義される入力変数(素性)に加え、次の素性を確率モデルに追加した; (i) Google n-gram コーパスによる単語頻度(=ウェブコーパスの単語頻度)、(ii) SVL12000 に登録された単語難易度表。また、確率モデルの学習に数百人規模のユーザに対する語彙情報も用いた。ラッシュモデルと同様にユーザ ID の素性が入っているため、ユーザに適応した単語難易度の予測が行われることになり、単語頻度や単語難易度の情報により、より正確な推測がなされる。

### 3.2.1 問題設定

単語は全部で  $N$  個あるとし、ユーザ  $u$  は事前に  $n (< N)$  個の単語集合  $V$  に対し単語の既知/未知の情報を入力していると仮定する。本研究における目標は、利用者  $u$  に対し、単語集合  $V$  の情報から  $V$  以外の単語の既知/未知を予測することである。

本研究ではその予測のために次の外部リソースを用いる。

- SVL 単語難易度表
- Google n-gram コーパスから得られる単語の確率分布
- Smart.fm と呼ばれるコンピュータ支援語彙習得 (CAVOCA) システムから得られる数百人規模のユーザに対する語彙情報

SVL 単語難易度は単語頻度やネイティブスピーカーによる感覚的判断を参考に 12,000 語の単語に対し 12 段階の難易度を割り振ったものである。Google コーパスは、1 兆語分のウェブページにおける単語の出現頻度を 5 グラム (接続する 5 単語) まで記録したデータであり、1 グラム (単語) の出現頻度の対数を判別のための素性とした。Smart.fm からは 10,526 人分のデータを取得したが、Smart.fm ではユーザに提示されていない単語は未知として扱われているため、十分に学習したユーザの情報でないと未知と欠損 (=ユーザが既知/未知をつけていない状態) のラベルが混合されていることになる。そこで、学習した単語の 15%以上を既知としており、かつ、100 単語以上の単語を学習している利用者 675 人に対するデータのみ外部リソースとして利用した。

次に評価について説明する。Smart.fm から得られるデータに対する既知/未知の予測精度を評価することも可能ではあるが、Smart.fm から得られるデータは、多くの単語に対し既知/未知のラベルが欠損しており、データに偏りがある。そのため評価には Smart.fm とは別の既知/未知ラベルがつけられたデータを作成した。具体的には、学生 10 人に対し、SVL 中の 12,000 語について、単語を知っている度合いを 5 段階から選択してもらうことにより評価用データを作成し、本研究においては、最高の 5 をつけている単語のみを既知とし、残りの 1~4 までは未知とすることとした。この評価用デー

表 5: テストデータに対する単語の既知/未知の判別精度

	n=10	30	100	300	600
IRT	68.33%	73.69%	74.65%	74.65%	74.60%
LR (提案モデル)	<b>73.25%</b>	<b>77.89%</b>	<b>79.09%</b>	<b>80.03%</b>	<b>80.01%</b>

表 6: テストデータに対する単語の既知/未知の判別精度

	n=10	30	100	300	600
SVM (Linear)	74.78	<b>78.08</b>	78.88	79.20	79.27
SVM (RBF)	67.61	77.27	<b>79.16</b>	79.55	79.91
CW	73.77	72.40	75.06	75.60	75.82
SGD (公開システム)	<b>76.95</b>	77.94	78.52	78.46	78.39
LR	73.25	77.89	79.09	<b>80.03</b>	<b>80.01</b>

タから、10,000 語をテストデータとし、1,400 語を開発用テストデータとし、残りの 600 語を最初にユーザが与える既知/未知の情報とした。

### 3.2.2 手法

本研究では以下の 2 つの手法を比較した。

- IRT: IRT の一種であるラッシュモデルは次のように定義される。

$$p(y = 1|u, t) = (1 + \exp(-\theta_u + d_t))^{-1}$$

ただし、 $u$  はユーザ ID、 $t$  は単語 ID、 $\theta_u$  は各ユーザ  $u$  の語彙力を示すパラメータ、 $d_t$  は各単語  $t$  の難易度を表すパラメータである。実際に各ユーザに各単語の既知/未知に関するテストを行ってもらい、そのデータから各パラメータが学習される。得られたパラメータを用いることによって、学習用データにでてこない未知の単語についてユーザの既知/未知の推定が行える。また、ユーザの語彙力を表すパラメータが存在することにより、この確率モデルはユーザ適応している。つまり、単語の難易度を表すパラメータを学習する際にユーザの語彙力による偏りが補正されることになる。

- LR (提案手法): ラッシュモデルはロジスティック回帰の一種になっており、ロジスティック回帰における入力変数(素性)を増やすことにより、より正確な推定が行えるようになる。ロジスティック回帰は次の式で表される。

$$p(y = 1|u, t) = (1 + \exp(w \cdot \phi(u, t)))^{-1}$$

ただし、 $u$  はユーザー ID、 $t$  は単語 ID、 $\phi$  は  $u$  と  $t$  から得られる素性ベクトルで、 $w$  は  $\phi$  に対する重みベクトルである。素性として次のデータを用いた

- ユーザ ID (=u) (IRT と同じ素性)
- 単語 ID(=t) (IRT と同じ素性)
- Google n-gram コーパスにおける単語  $t$  の単語頻度の対数 (LR で新しく追加された素性)
- SVL 単語表における単語  $t$  の単語難易度 (LR で新しく追加された素性)

ユーザ毎の英語能力はユーザ ID に対する重み変数に学習され、単語の難易度は単語頻度の素性と SVL の素性に対する重み変数に学習されることが期待される。この手法はラッシュモデルと同様にユーザ適応の学習になっており、ユーザの英語能力と単語の一般的な難易度がそれぞれ別の変数に学習され、テストユーザに対する単語難易度はテストユーザの英語能

力と単語の一般的な難易度から推定される。識別器の訓練時にはユーザ毎の英語能力の差異がユーザ ID に対する重み変数で正規化されるため、テストユーザ以外の各単語の一般的な難易度が学習される。

テスト時には、少数の単語集合 (< 600 単語) に対するテストユーザの既知/未知情報と、テストユーザ以外のユーザから学習された各単語の一般的な難易度情報から、テストユーザに対する単語の既知/未知情報が推定される。実験では、Smart.fm のユーザ 675 人に対する 12000 単語の欠損データをテストユーザ以外のデータとして用いて識別器を学習した。

### 3.2.3 評価

表 5 は実験結果を示しており、IRT のモデルよりも提案手法の方が高い精度を実現している。ロジスティック回帰と我々の提案する素性を用いることでおおよそ 80%の精度を実現することがわかった。また、おおよそ 100 単語に関する既知/未知の情報を与えることで十分な学習が行えることもわかった。

我々のシステムはロジスティック回帰(LR)のオンライン学習である確率的勾配法(Stochastic Gradient Descent, SGD)で実際に学習されている。これはユーザが利用するごとに新しい学習データ(ユーザがある単語の既知/未知をシステムに知らせるクリック入力)がシステムに逐次的に入力され続けるためであって、このようなタスク環境では、全データから学習を行わなければならないバッチ学習を用いることは難しく、オンライン学習が非常に適している。表 6 は有名なバッチ学習手法とオンライン学習手法での性能比較を示している。バッチ学習としては、SVM の線形カーネル、SVM の RBF カーネル、ロジスティック回帰を評価し、オンライン学習では、Confidence Weighted Learning (CW)、我々のシステムで用いた SGD を比較した。SGD は LR に比べやや性能が劣るものの、実用上問題の無い差であることがわかった。

## 3.3 具体的成果

本研究では英文の読解を支援するために、英文に対し自動的に語義のアノテーションを与えるシステムを提案した。学生 10 人に、12,000 語について単語の既知/未知のラベルを与えたデータを作成し、このデータを用いてシステムの性能を評価した。外部リソースとして、Google n-gram コーパスによる単語頻度、SVL12000 単語の単語難易度表、Smart.fm と呼ばれるコンピュータ支援語彙習得システムから得られる 675 人分の既知/未知の欠損データを用いる手法を提案し、評価した。実験により、おおよそ 100 単語に対する既知/未知の情報を得ることでおおよそ 80%の精度で残りの単語の既知/未知を予測することが出来た。また、このようなシステムにおいてオンライン学習を用いることの有用性も確認できた。

## 4 成果要覧

### 受賞関連

[受賞 1] 江原 遥, 二宮 崇, 中川 裕志: NLP 若手の会 第 4 回シンポジウム 奨励賞, NLP 若手の会, 2009 年 9 月.

[受賞 2] 松島 慎: 情報処理学会創立 50 周年記念(第 72 回) 全国大会 学生奨励賞, 情報処理学会, 2010 年 3 月.

### 査読付論文

[査読付 1] Yo Ehara, Nobuyuki Shimizu, Takashi Ninomiya and Hiroshi Nakagawa: Personalized Reading Support for Second-language Web documents by Collective Intelligence, In the Proceed-

ings of the 2010 International Conference on Intelligent User Interfaces (IUI2010), pp. 51—60, 2010.

## 公開ソフトウェア

- [公開 1] 江原遥: SocialDict, 英文の読解支援システム (<http://www.socialdict.com/>), 2009年11月.  
 [公開 2] 江原遥: NEWikipedia, 英語版 Wikipedia の読解支援システム (<http://en.newwikipedia.org/>), 2010年3月.

## その他の発表論文

- [発表 1] 松島慎, 清水伸幸, 吉田和弘, 二宮崇, 中川裕志: サポートクラスによる Passive-Aggressive アルゴリズムの多クラス化, 情報処理学会 第 192 回自然言語処理研究会, 2009-NL-192(12), pp. 1—6, 北見, 2009年7月.  
 [発表 2] 松島慎, 清水伸幸, 佐藤一誠, 二宮崇, 中川裕志: SPA アルゴリズムの半教師あり学習への応用, 情報処理学会 第 76 回数理モデル化と問題解決研究会, 2009-MPS-76(21), pp. 1—8, 東京, 2009年12月.  
 [発表 3] 松島 慎, 清水伸幸, 吉田和弘, 二宮 崇, 中川裕志: 多クラス識別問題におけるオンライン学習のための厳密な PA アルゴリズム, 情報処理学会創立 50 周年記念(第 72 回)全国大会講演論文集, vol.2, pp. 467—468, 東京, 2010年3月.  
 [発表 4] 江原遥, 二宮崇, 中川裕志: Web 文書中の単語クリックログの解析から未知単語を予測する語義注釈システム, 情報処理学会 第 193 回自然言語処理研究会, 2009-NL-193(3), pp. 1—7, 京都, 2009年9月.  
 [発表 5] 江原遥, 二宮崇, 中川裕志: 語義注釈システムの単語クリックログからの言語能力情報の抽出, NLP 若手の会 第 4 回シンポジウム, 京都, 2009年9月.  
 [発表 6] 江原 遥, 二宮 崇, 清水伸幸, 中川裕志: Web 文書中のユーザが知らない語を予測する読解支援システム, 情報処理学会創立 50 周年記念(第 72 回)全国大会講演論文集, vol.2, pp. 77—78, 東京, 2010年3月.  
 [発表 7] 江原遥, 二宮崇, 清水伸幸, 中川裕志: 英語版 Wikipedia を対象としたユーザが知らない語を予測する読解支援システム, 言語処理学会 第 16 回年次大会 発表論文集, pp.864—867, 東京, 2010年3月.

## 特記事項

- [特記 1] 二宮 崇: 単一化文法, 言語処理学事典, 言語処理学会編, 共立出版, pp. 158—159, 2009年12月.  
 [特記 2] 二宮 崇: 深い構文解析, 言語処理学事典, 言語処理学会編, 共立出版, pp. 176—177, 2009年12月.  
 [特記 3] 二宮 崇: 文法獲得, 言語処理学事典, 言語処理学会編, 共立出版, pp. 182—183, 2009年12月.

# Web 文書の分析・活用システムの研究 — 人名曖昧性解消システムの性能向上 —

吉田 稔

## 1 概要

本年度は、主に Web 文書を対象とし、大量のテキストデータを活用するためのテキストマイニングタスクに関して、以下のような研究を行った。本稿では、このうち「Web 検索における同名曖昧性解消に関する研究」に関して詳細に解説を行う。

- テキスト中の数値表現検索システムに関する研究 [発表 4][発表 8]

テキスト中の数値表現を対象としたテキストマイニングシステムの研究。本年度は、数値の範囲検索を利用して、数値に対する検索支援を行うシステムの研究を行った。また、数値表現からの自動グラフ生成システムに関して、性能の分析を行った。

- 同義語・類義語抽出に関する研究 [受賞 1][受賞 2][査読付 1][発表 7][発表 9]

テキストから自動的に同義語・類義語を抽出するための研究。本年度は、特に、同義語・類義語をオンデマンドに抽出し、検索支援に用いるための研究について、システムの性能に関するより詳細な実験・検討を行った。

- Web 検索における同名曖昧性解消に関する研究 [発表 1][発表 6][発表 10]

Web における人名検索における人名の曖昧性解消のための研究。本年度は、新たなアルゴリズムの提案、論文検索システムへの応用を行った。また、国際ワークショップ（コンテスト）への参加を行い、17 チーム中 2 位の成績を取めた。

- 株価とテキストの関連分析に関する研究 [発表 2][発表 3]

企業に関するニュース記事と、その企業の株価等の指標の関連を分析する研究。本年度は、業種による単語の影響の違いの分析や、記事の話題によるクラスタリングを利用した取引高予測精度向上のための研究を行った。

- Web 文書構造の解析 [発表 5]

Web 文書のレイアウトを解析し、文書構造を抽出する研究。本年度は、レイアウトに対する新たな確率モデルの提案とその性能の検討を行った。

## 2 Web 検索における同名曖昧性解消に関する研究

### 2.1 背景

人名の検索は、Web 検索における主要なタスクの一つであるが、そこには常に曖昧性の問題が付いて回る。例えば、「吉田稔」という文字列だけでは、それがどの人を指しているのか特定することは難し



い。何故なら、「吉田稔」という名前を持つ人は全国に多く存在するからである。このため、人名での Web 検索では、検索結果が目的の人のページなのかどうかについて常に気を配る必要がある。**人名曖昧性解消**とは、このような問題に対し、同じ人物のページを自動的にまとめあげ、検索結果を見易くするというタスクである。

近年、人名曖昧性解消に関する国際ワークショップ WePS(Web People Search Workshop) が開催されており、各国から様々なシステムが参加してその精度を競い合っている。WePS を通じて、特に人名クエリと共起する固有名詞に着目した手法が効果を発揮することが明らかになっている。例えば、Bill Gates という検索クエリを考えたとき、検索結果の複数の Web 文書に Paul Allen という人名がそれと共起していれば、それらの文書は同一人物である可能性が高い。固有名詞の利用によって、高い精度で同一人物を判定できる半面、こうした固有名詞は、必ずしも文書に出現するとは限らず、網羅性という面で限界がある。

網羅性を上げるための方策として最初に上げられるのが、固有名詞以外の一般の単語（名詞）を用いることであるが、一般の単語は、人物との関連性が高い単語ばかりではないため、これらを手掛かりとすることによって、関連のない人物どうしを結びつけてしまうミスが発生しやすくなる。実際、WePS の結果でも、単純に一般の単語を用いることは、通常、人物曖昧性解消の性能の向上には貢献しないことがわかっている。

これに対し本研究では、一般の単語のうち、**有用な単語**に高い重みを与え、手掛かりとして用いる手法を提案する。そのために、固有名詞などの精度の高い特徴量を用いた、**第一段階クラスタ**（クラスタ＝文書のまとまり）を生成し、これら第一段階クラスタの結果を利用することで一般の単語に重みを付与する。この重みづけされた単語を用い**第二段階クラスタ**を生成する。このため、提案手法を、**二段階クラスタリング法**と呼ぶ。これにより、例えば、野球選手とコンピュータ科学者が同姓同名だった場合に、「ホームラン」や「ボール」、あるいは「メモリ」や「アルゴリズム」といった単語を重みの高い単語として人物の判定に利用することが可能になる。

## 2.2 内容

### 2.2.1 第一段階クラスタリング

まず、人物と関連の深い特徴量を利用して、精度の高い（間違いの少ない）クラスタを生成する。この際、先行研究で用いられていた「固有名詞」のほかに、「複合名詞」と「URL」も利用する。複合名詞は、ある専門分野の特徴的な概念を表していることが多い。先の Bill Gates の例で言えば、例えば chief software architect という複合語が複合名詞の例である。また、URL は、その人物に関連する Web ページを表現しているということで、人物の特定能力が高いと考えられる。

固有名詞としては、人名のほか、組織名と地名も用いる。ただし、これらのうち、一般的な（一般の文書に高い頻度で現れる）ものは使用しない。これは、例えば「中央区」のような、どこにでも現れる地名を使用することによるミスを防ぐためである。また、複合名詞の利用の際には、専門用語抽出アルゴリズム「言選 Web」を利用し、複合名詞のスコア付けを行っている。このスコアは、複合名詞を  $CW (= W_1 W_2 \cdots W_L)$ （ただし、 $W_i (i = 1, 2, \dots, L)$  は単名詞）、 $f(CW)$  を  $CW$  の頻度としたとき、

$$Score(CW) = f(CW) \cdot LR(CW), \quad (1)$$

ただし、 $LR(CW)$  は、

$$LR(CW) = \left( \prod_{i=1}^L (LN(W_i) + 1)(RN(W_i) + 1) \right)^{\frac{1}{2L}} \quad (2)$$

と定義される。ここで、 $LN(W_i)$ （あるいは  $RN(W_i)$ ）は、 $W_i$  の右（あるいは左）に接続する単語の頻度を表す。このスコアが閾値  $\theta_{CKW}$  を超えた複合名詞のみを使用する。

URL としては、HTML の<a>タグで囲まれた部分のほか、対象の Web 文書の URL そのものも使用する。また、組織名・地名と同様、あまりにも Web 上での頻度が高い URL は使用しない。

これらの特徴量を利用し、特徴量毎に、文書と文書の類似度を、オーバーラップ係数

$$Overlap(d_x, d_y) = \frac{|\mathbf{f}_x \cap \mathbf{f}_y|}{\max(\min(|\mathbf{f}_x|, |\mathbf{f}_y|), \theta_{overlap})}$$

で定義する。ただし、 $\mathbf{f}_x$  と  $\mathbf{f}_y$  は、文書  $d_x$  と  $d_y$  に共通する特徴（固有名詞など）の数である。また、 $\theta_{overlap}$  は、分母が小さくなりすぎることを防ぐためのパラメータである。

以上で定義される複数の類似度から、最終的な文書同士の類似度を定義するが、まず、固有名詞に関しては、「人名による類似度」「地名による類似度」「組織名による類似度」を、以下のようにして線形補間により組み合わせることにより固有名詞全体での類似度を得る。

$$sim_{NE}(d_x, d_y) = \alpha_P sim_P(d_x, d_y) + \alpha_L sim_L(d_x, d_y) + \alpha_O sim_O(d_x, d_y)$$

(ただし  $\alpha_P + \alpha_L + \alpha_O = 1$ 。また、これらの値は、学習データによる調整を経て、 $\alpha_P = 0.78$ ,  $\alpha_O = 0.16$ , and  $\alpha_L = 0.06$  と設定されている。)

さらに、こうして計算された「固有名詞による類似度」「複合名詞による類似度」「URL による類似度」に関して、以下のように最大値をとることで文書の類似度と定義する。

$$sim_{\max}(d_x, d_y) = \max(sim_{NE}(d_x, d_y), sim_{CKW}(d_x, d_y), sim_{URL}(d_x, d_y)) \quad (3)$$

これらの文書間類似度をもとに、文書のクラスタリング（纏め上げ）を行う。クラスタリングには、一般的な手法である「群間平均法による階層クラスタリング」を用いた。これは、「1 クラスタ = 1 文書」の状態から始めて、最も類似度の高いクラスタ同士を逐次併合してゆく手法である。クラスタ  $C_i, C_j$  間の類似度は、以下のようにクラスタ内文書の類似度平均を計算することで得られる。(ただし、 $sim_{\max}(d_x, d_y)$  は文書間類似度。)

$$sim(C_i, C_j) = \frac{1}{|C_i||C_j|} \sum_{d_x \in C_i} \sum_{d_y \in C_j} sim_{\max}(d_x, d_y) \quad (4)$$

### 2.2.2 第二段階クラスタリング

第二段階クラスタリングでは、第一段階で得られたクラスタを改良することを目指し、クラスタの改定を行う。そのための手法として、近年自然言語処理（情報抽出）の研究で提案されている Espresso アルゴリズムを採用する。Espresso は、最初にある名詞集合（あるいは名詞ペアの集合）が与えられると、それが出現しやすい言語的パターンを発見し、そのパターンをもとに類似する名詞（名詞ペア）を新しく獲得するという、**ブートストラップ法**と呼ばれる手法の一種である。例えば、(Bush, Regan) というペアをもとに、それらが出現する言語的パターンを発見し、それを利用して、類似のペアとして (Ford, Nixon)（注：どちらも succession という関係を持っている）が獲得できる。Espresso アルゴリズムは、行列により

$$\mathbf{i}^{(t+1)} = \frac{1}{|I||T|} \cdot \mathbf{M}^T \mathbf{M} \mathbf{i}^{(t)} \quad (5)$$

と表現できることが知られている。ここで、 $\mathbf{M}$  はインスタンスとパターンの関係の強さを表す行列、 $\mathbf{i}$  はインスタンスの列に対する重みを表すベクトルである。アルゴリズムは、与えられた名詞集合に



---

**Algorithm 1** ブートストラップ法による第二段階クラスタリング

---

**Procedure:**  $D, F, R_D^{(0)}$

**Step-1:** // Calculation of Feature-Document Matrix  $P$

$$P[f, d] = \begin{cases} \frac{1}{\max pmi} \log \frac{p(f,d)}{p(f)p(d)} & \text{if } \frac{p(f,d)}{p(f)p(d)} > 1 \\ 0 & \text{otherwise} \end{cases}, (f \in F, d \in D)$$

**where**  $\max pmi = \max(P[f', d'])$  ( $f' \in F, d' \in D$ )

**for**  $t \in 0, \dots, T - 1$  //  $T$ : Number of Iterations

**Step-2:** // Calculation of Feature-Cluster Matrix

$$R_F^{(t)} = \frac{1}{|D|} P R_D^{(t)}$$

**Step-3:** // Calculation of Document-Cluster Matrix

$$R_D^{(t+1)} = \frac{1}{|F|} P^T R_F^{(t)}$$

**endfor**

**Step-4:**

**for**  $C \in \mathcal{C}$  **do**  $C_d^{(T)} = \arg \max_C r_{d,C}^{(T)}$   
**where**  $\{C' | (C' \in \mathcal{C} \wedge |C'| > 1) \vee C_d^{(0)}\}$

**endfor**

Define  $\mathcal{C}^{(T)}$  based on  $C_d$ .

**return**  $\mathcal{C}^{(T)}$

---

従って初期化されたベクトル  $\mathbf{i}^{(0)}$  を用意し、これに  $M, M^T$ 、さらに正規化のための値  $\frac{1}{|T||T|}$  を乗算することにより  $\mathbf{i}$  を更新していく。

このアルゴリズムは、上記の名詞を**文書**、パターンを**単語**と見なすことにより、人名曖昧性解消の問題に適用することができる。第一段階クラスタリング結果の各クラスタ毎に、そのクラスタに入っている文書集合を表現した初期化ベクトルを用意し、各ベクトル毎に上記の更新を行っていく。複数ベクトルの更新を同時に行うため、ベクトル  $i$  のかわりに行列  $R$  を用いることに注意されたい。  $R$  の各列が各クラスタに対応するベクトルとなる。具体的には、  $R_D^{(t)} = \{r_{d,C}\}$  を定義する。 ( $\{r_{d,C}\}$  は、文書  $d$  とクラスタ  $C$  の関連の強さ。) これに加え、  $i$  番目の単語と  $j$  番目の文書の関係の強さを表す値を  $(i, j)$  要素に持つ行列  $P$ 、および単語-クラスタ行列  $R_F^{(t)} = \{r_{f,C}\}$  を定義する。  $P$  は、Espresso の定義式と同様に、自己相互情報量に従って定義されている。アルゴリズムは、  $R_D^{(t)}$  と  $R_F^{(t)}$  を、  $P$  および  $P^T$  を掛けることにより交互に更新していく。

具体的なアルゴリズムは Algorithm 1 に示す通りになる。第一段階クラスタリングによって得られたサイズ 2 以上のクラスタを初期クラスタ  $R_D^{(0)}$  として与え ( $d \in C$  なら  $r_{d,C}^{(0)} = 1$ 、それ以外なら 0)、どのクラスタにも属していない文書の所属をこのアルゴリズムで決定する。

表 1: WePS-2 データセットでの結果

Topic	BEP	BER	F <sub>B</sub>
<b>公式ベースライン手法</b>			
ALL IN ONE	0.43	1.00	0.53
ONE IN ONE	1.00	0.24	0.34
COMBINED	0.43	1.00	0.52
<b>第一段階クラスタリング</b>			
ORIGINAL	0.92	0.70	0.78
<b>二段階クラスタリング</b>			
TOPIC	0.94	0.70	0.79
CKW	0.87	0.77	0.81
<b>BOOTSTRAP</b>			
1-gram, $T = 1$	0.89	0.82	<b>0.85</b>
1-gram, $T = 2$	0.66	0.91	0.73
1-gram, $T = 3$	0.53	0.95	0.63
2-gram, $T = 1$	0.92	0.70	0.78
<b>WePS top 3</b>			
1st	0.87	0.79	0.82
2st	0.85	0.80	0.81
3st	0.93	0.73	0.81

最終的に、更新式は以下のように書ける。(  $|D|$  と  $|F|$  は定数であり、最終的な結果には影響しない。)

$$R_D^{(t+1)} = \frac{1}{|D||F|} \cdot P^T P R_D^{(t)} \tag{6}$$

### 2.3 具体的成果

最新のデータセットである WePS-2 データセットによる実験を行った。WePS の公式ベースライン手法と参加システムのベスト 3、さらに、他の二段階クラスタリング手法として、既存の手法である、トピックモデルを利用した手法 (TOPIC) と、二段階用語抽出手法 (CKW) を比較対象として結果を示したものが表 1 である。ここで、“BEP”、“BER”、 $F_B$  は WePS で用いられる評価指標であり、 $F_B$  がシステムの総合スコアとなる。“ORIGINAL” は第一段階クラスタリングのみの結果を示し、“BOOTSTRAP” が二段階クラスタリングを示す。また、 $T$  は二段階目の反復回数を示す。二段階クラスタリングによる手法が 0.85 と最も高い精度を示すことが確認された。ただし、2 回以上の反復は、精度を下げるということがわかった。また、本システムのプロトタイプシステムは、WePS-2 ワークショップに参加し、17 チーム中 2 位の成績を収めている。

## 3 成果要覧

### 受賞関連

[受賞 1] 吉田 稔, 中川 裕志: 第 23 回全国大会 全国大会優秀賞, 人工知能学会, 2009 年 9 月.

[受賞 2] 佐藤 一誠, 吉田 稔, 中川 裕志: データマイニングと統計数理研究会 研究会優秀賞, 人工知能学会, 2009 年 6 月.

## 査読付論文

[査読付 1] 吉田稔, 中川裕志, 寺田昭: コーパス検索支援のための動的同義語候補抽出, 人工知能学会論文誌, Vol.25, No.1, pp. 122–132, 2010 年 1 月

## その他の発表論文

[発表 1] 池田雅紀, 佐藤一誠, 吉田稔, 中川裕志: 半教師有り学習に基づく Web 上の人物クラスタリングシステム, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会講演論文集, 6ZC-1, 東京, 2010 年 3 月.

[発表 2] 廣川敬真, 吉田稔, 山田剛一, 増田英孝, 中川裕志: 業種別による新聞記事と株価動向の関係の解析, 言語処理学会 第 16 回年次大会 発表論文集, pp.1070–1073, 東京, 2010 年 3 月.

[発表 3] 吉田 稔, 中川 裕志, 松井 藤五郎, 和泉 潔, 石田 智也, 中嶋 啓浩: 取引高とニュース記事の関連性の分析, 第 4 回 ファイナンスにおける人工知能応用研究会 (SIG-FIN), pp.60–64, 東京, 2010 年 1 月.

[発表 4] 吉田稔, 佐藤一誠, 中川裕志, 寺田昭: 接尾辞配列とディリクレ過程混合モデルを用いたテキスト中の数値表現マイニング, 情報処理学会 第 96 回情報学基礎研究会, Vol.2009-FI-96, No.4, 東京, 2009 年 11 月.

[発表 5] 吉田 稔, 中川 裕志: 確率モデルを利用した Web 文書からの見出し抽出, NLP 若手の会 第 4 回シンポジウム, 2009 年 9 月.

[発表 6] 池田雅紀, 小野真吾, 佐藤一誠, 吉田稔, 中川裕志: Web 上の人名検索結果の同姓同名問題における二段階クラスタリングを用いた再現率向上, 情報処理学会 第 148 回データベースシステム・第 95 回情報学基礎合同研究発表会, Vol.2009-FI-95, No.6, 神戸, 2009 年 7 月.

[発表 7] 吉田 稔, 中川 裕志: Wikiwi: テキストマイニングによる Wikipedia 検索支援, 人工知能学会第 23 回全国大会, 1C3-3, 高松, 2009 年 6 月.

[発表 8] 吉田 稔, 杉浦 隆博, 山田 剛一, 増田 英孝, 中川 裕志: テキストからの数値抽出による自動グラフ作成, 人工知能学会第 23 回全国大会, 3F2-NFC3-2, 高松, 2009 年 6 月.

[発表 9] 吉田 和弘, 吉田 稔, 中川 裕志: 文字列検索に基づく同義語・類義語抽出ツールとその性能評価, 第 191 回自然言語処理・第 76 回音声言語情報情報処理 合同研究会, Vol.NL191/SLP76, No.19, 東京, 2009 年 5 月.

[発表 10] Masaki Ikeda, Shingo Ono, Issei Sato, Minoru Yoshida and Hiroshi Nakagawa: Person Name Disambiguation on the Web by TwoStage Clustering, Proceedings of the 2nd Web People Search Evaluation Workshop (WePS 2009) at the 18th International World Wide Web Conference (WWW2009), Madrid, Spain, April, 2009.

## 特記事項

[特記 1] 吉田稔: 木接合文法, 言語処理学事典, 言語処理学会編, 共立出版, pp.160–161, 2009 年 12 月.

# これからの電子図書館を支える知識ナビゲーション技術の研究

清田 陽司

## 1 概要

オンライン百科事典 Wikipedia は、Web 情報資源と図書館情報資源の性質の差異を考慮したとき、興味深い性質をもっている。Wikipedia には図書館情報資源ほど高い信頼性はないが、多くの参加者による素早い更新によってその欠点を補っている。また、図書館分類体系にはない最新の概念をカバーしている。本研究では Wikipedia 特有の性質に着目し、以下のトピックに取り組んだ。

- **図書館ナビゲーションシステム** [招待 1] [招待 2] [招待 3] [招待 4] [招待 5] [招待 6] [招待 7] [招待 8] [査読付 1] [発表 1] [発表 6] [発表 9] [報道 1]

以前から取り組んでいる図書館ナビゲーションシステム「Littel Navigator」の技術を応用し、国立国会図書館(NDL)がもつ調査に役立つ膨大な情報を効率的に探せるシステムを開発した。この成果は、NDL が 2009 年 5 月にサービスを開始した「リサーチ・ナビ」に利用されている。

- **Wikipediaを活用した言語間差異比較システムとその応用** [招待 9] [招待 10] [発表 2] [発表 3] [発表 4] [発表 5] [発表 7] [発表 8] [発表 10] [特記 1]

Wikipedia の各国語版の間に存在する差異を言語およびその言語を使う国や文化による興味・関心の違いとして捉え、比較を行うシステムを作成した。また、各国語版の Wikipedia を効率よく解析するため、分散処理ソフトウェア Hadoop を活用する取り組みを行った。

## 2 図書館ナビゲーションシステム

### 2.1 背景

Web 空間に流通している情報量の爆発的な増大は、個人が情報を探するときの前提を大きく変化させている。Web サーチエンジンによる検索は、何らかの調べ物をする際の大前提となっている。しかし、広大な情報の海の中を泳いで真に必要な情報と出会うことは簡単ではない。一方、過去の膨大な情報を収集・整理し、一般利用者に提供する役割を果たしてきた図書館には、広大な情報の海を泳ぐためのさまざまなノウハウが蓄積されている。ただし、現在の図書館サービスは紙媒体を前提として組み立てられており、電子化された情報を扱う仕組みは Web ほど洗練されているとはいえない。図書館がもつさまざまなノウハウを Web 空間に適切に提供していくための技術が切望されている。

### 2.2 内容

国立国会図書館(NDL)は、一般国民向けの情報探索支援サービスとして、蔵書検索システム NDL-OPAC、近代デジタルライブラリ、参考書誌紹介、調べ方案内、レファレンス協同データベース

など、多数のデータベースサービスを運営している。しかし、それぞれのデータベースがほとんど連携しておらず、利用者にとっての使いにくさが指摘されていた。

この問題を解決するため、NDL が保有する情報探索向けのさまざまなデータベースを統合して提供するポータルサイト「リサーチ・ナビ」の開発が行われた。我々の研究グループでは、以前より実運用されている図書館ナビゲーションシステム「Littel Navigator」の知見をもとに、リサーチ・ナビ検索システムの研究開発に取り組んだ。

システム開発にあたっては、昨年度の研究成果であるテーマグラフ生成機能を取り込むことによって、利用者がどの分野の情報資源にアクセスしたらよいかをナビゲートすることを目指した。また、Web で汎用的に利用されている有用なユーザインタフェース技術を積極的に採用した。具体的には、タブデザインによってユーザの用途に沿った検索結果を表示するようにしたり、Ajax 技術によって検索レスポンスを向上させるなどの工夫を盛り込んだ。

リサーチ・ナビ検索システムの構成を図 1 に示す。多様なデータベースの検索結果をスムーズに表示できるように、Web ブラウザ側に検索コントローラの機能をもたせ、得られた検索結果を順次インタフェースに反映していく構成となっている。

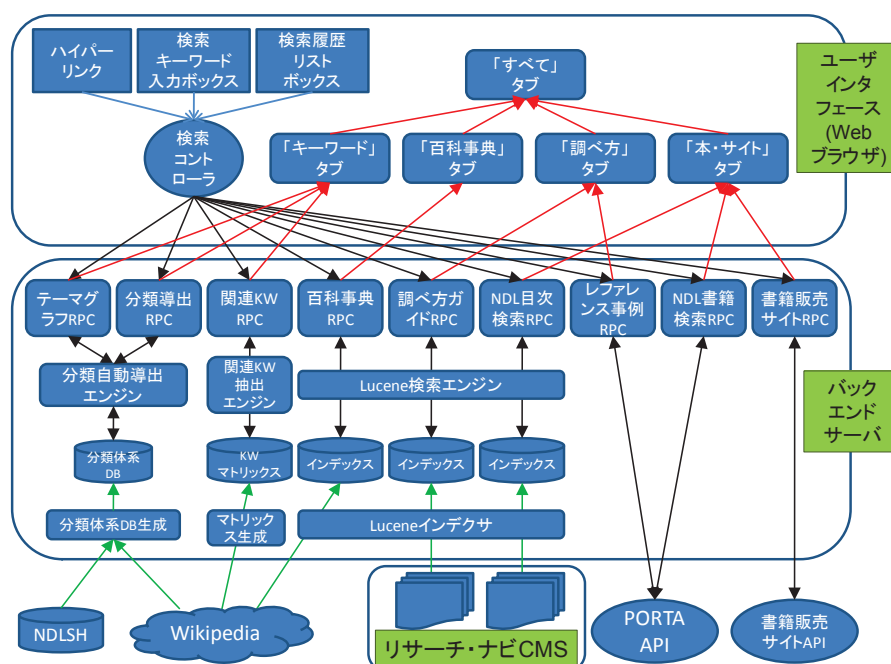


図 1 リサーチ・ナビ検索システムの構成

### 2.3 具体的成果

上記の取り組みの成果は、2009年5月に国立国会図書館のWebサイトにおいて「リサーチ・ナビ」<http://rnavi.ndl.go.jp/>として公開されている(図2)。ライブラリアンを中心としたユーザ層に文献調査を行うための基本的なツールとして受け入れられていて、週あたり10000回近い検索が行われている。

リサーチ・ナビ検索システムの概要については招待講演、招待論文などで発表した[招待4][招待6][招待7][発表6][発表9]。リサーチ・ナビ検索システムの基礎となった Littel Navigator については、図書館情報学分野の国際会議 JCDL2009 [査読付1]をはじめ、さまざまな機会において発表し

た[招待 1] [招待 2] [招待 3] [招待 5] [招待 8] [発表 1]。また、Littel Navigatorの技術は、ベンチャー企業において商品検索システムにも応用されている[報道 1]。



図 2 リサーチ・ナビのユーザインタフェース

### 3 Wikipedia を活用した言語間差異比較システムとその応用

#### 3.1 背景

Wikipedia が通常の百科事典と大きく異なるのは、一般のユーザ自身が記事を作成してコンテンツを育てていく点である。Wikipedia には多数の言語版が存在し、2010 年 1 月現在では 250 を超える言語版が運営されている。各言語版の編集はお互いに独立しており、記事の主題はユーザの自由に任されているので、ある事柄についての記事が特定の言語版にのみ存在していたり、ある話題に関する記事が他に比べて特に多く作成されている言語版があるなど、各言語版はそれぞれが異なる成長をしている。各言語版の差異をうまくとらえることができれば、言語およびその言語を使う国や文化による興味・関心の違いの比較が可能になることが期待される。

#### 3.2 内容

本研究では、言語間の違いをとらえるための手段として、Wikipedia の言語間リンクと、各々の言語版がもつカテゴリネットワークを利用することとした。具体的には、指定された言語版のカテゴリを起点としてページ構造を再帰的に探索し、カテゴリに属する記事とそのタイトル、およびカテゴリの下位カテゴリとそのタイトルを全て収集する探索アルゴリズムを考案した。さらに、このアルゴリズムをシステムとして実装し、任意のカテゴリ名を指定することで結果が得られる Web インタフェースを作成した。



### 3.3 具体的成果

作成した Web インタフェースを図 3 に示す。フランス語・ドイツ語・日本語・中国語の各言語版において、入力したカテゴリ名に関する記事がどの程度充実しているかを俯瞰することができる。表 1 に、カテゴリ名「欧州連合」に対して得られた結果を示す。ドイツ語版とフランス語版にページが多く、文化的・政治的な背景が Wikipedia に反映されていることが確認できる。

本研究の成果はいくつかの研究会において発表した[発表 3] [発表 7] [発表 10]。また、本研究の成果物として公開されている分散処理ソフトウェア Apache Hadoop ベースの Wikipedia 解析ツール Wik-IE を活用した研究取り組みの紹介 [招待 9] [招待 10] [発表 2] [発表 8]や、Wikipedia カテゴリ体系をブログ解析に応用した研究[発表 4][発表 5]も行っている。

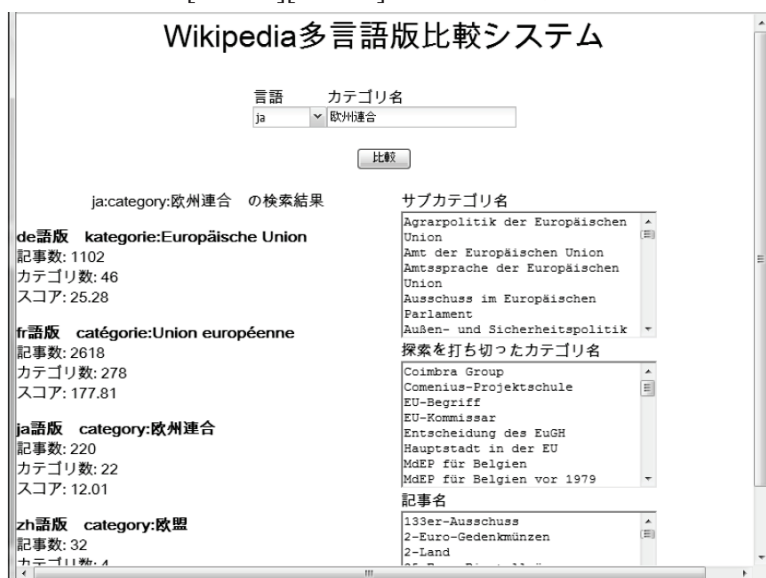


図 3 比較システムの Web インタフェース

表 1 カテゴリ名「欧州連合」の比較結果

言語	記事数	カテゴリ数
ドイツ語	1102	46
フランス語	2618	278
日本語	220	22
中国語	32	4

## 4 成果要覧

### 招待講演／招待論文

[招待 1] 清田 陽司: 図書館分類体系と Wikipedia を統合した情報探索支援システムの開発, 日本図書館研究会情報組織化研究グループ／情報知識学会関西部会 合同研究会, 大阪科学技術センター, 大阪市, 2009 年 7 月.



- [招待 2] 清田 陽司: Wikipediaをいかに使いこなすか? -知識抽出、情報ナビゲーション、そしてトピック発見-, 第 114 回 Ku-librarians 勉強会, 京都大学, 京都市, 2009 年 8 月.
- [招待 3] 清田 陽司: Wikipediaと図書館情報資源のマッシュアップ, 第 36 回生物医学図書館員研究会, 順天堂大学, 東京都, 2009 年 9 月.
- [招待 4] 清田 陽司: リサーチ・ナビ検索システムの技術, 第 11 回図書館総合展/学術情報オープンサミット 2009 フォーラム企画, 国立国会図書館主催, パシフィコ横浜, 横浜市, 2009 年 11 月.
- [招待 5] 清田 陽司: 学生向けレファレンス支援ツールの可能性, 第 11 回図書館総合展/学術情報オープンサミット 2009 ミニ・フォーラム&プレゼンテーション企画, 紀伊國屋書店主催, パシフィコ横浜, 横浜市, 2009 年 11 月.
- [招待 6] 清田 陽司: 知識体系の新たな融合: 情報探索と件名標目表の活用をめぐる, 京都大学図書館機構 平成 21 年度第 2 回講演会「次世代 OPAC を考える - 目録情報の視点から - », 京都大学, 京都市, 2009 年 11 月.
- [招待 7] 清田 陽司: リサーチ・ナビ検索システムの技術, 参考書誌研究, 国立国会図書館主題情報部 編, No. 71, pp. 33-53, 2009 年 11 月.
- [招待 8] 清田 陽司: Wikipediaを活用した新たな情報ナビゲーションシステムの提案, 薬学図書館, 日本薬学図書館協議会 編, Vol. 55, No. 1, pp. 51-59, 2010 年 1 月.
- [招待 9] 清田 陽司: Apache Hadoop による大規模データの研究活用事例, 産学連携共同研究シンポジウム「クラウド・コンピューティングと大学発ベンチャー», 東京大学産学連携本部/(株)ユニファイ・リサーチ共催, 東京大学情報学環福武ホール, 東京都, 2010 年 1 月.
- [招待 10] 清田 陽司: Hadoop による大規模分散データ処理, 産業技術大学院大学 InfoTalk 第 16 回, 産業技術大学院大学, 東京都, 2010 年 3 月.

## 査読付論文

- [査読付 1] Yoji Kiyota, Hiroshi Nakagawa, Satoshi Sakai, Tatsuya Mori, Hidetaka Masuda: Exploitation of the Wikipedia Category System for Enhancing the Value of LCSH, In Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries (JCDL 2009), pp. 411-412, Austin, Texas, USA, 2009 年 6 月.

## その他の発表論文

- [発表 1] 中野 幹生, 緒方 淳, 清田 陽司, 東中 竜一郎, 翠 輝久: 【パネル討論】音声インタフェースにおける Web テキスト処理技術の利用, 情報処理学会 第 191 回自然言語処理研究会 (NL-191-12), 東京工業大学, 東京都, 2009 年 5 月.
- [発表 2] 野田 陽平, 清田 陽司, 中川 裕志: Wikipediaからの意外性のある情報の抽出, NLP 若手の会 第 4 回シンポジウム, 京都大学, 京都市, 2009 年 9 月.

- [発表 3] 森 竜也, 増田 英孝, 中川 裕志, 清田 陽司: 多言語に展開する Wikipedia の特性調査, NLP 若手の会 第 4 回シンポジウム, 京都大学, 京都市, 2009 年 9 月.
- [発表 4] 中崎 寛之, 阿部 佑亮, 宇津呂 武仁, 河田 容英, 福原 知宏, 神門 典子, 吉岡 真治, 中川 裕志, 清田 陽司: 特定トピックの日英ブログ収集・分析・類型化: 事例研究, 情報処理学会 第 194 回自然言語処理研究会 (NL-194-9), 愛媛大学, 松山市, 2009 年 11 月.
- [発表 5] 佐藤 由紀, 横本 大輔, 中崎 寛之, 宇津呂 武仁, 吉岡 真治, 福原 知宏, 神門 典子, 中川 裕志, 清田 陽司: Wikipedia を介した関連ニュース・ブログの対応付け: Wikipedia エントリの分析, 情報処理学会 第 194 回自然言語処理研究会 (NL-194-10), 愛媛大学, 松山市, 2009 年 11 月.
- [発表 6] 坂井 哲, 増田 英孝, 清田 陽司, 中川 裕志: 国立国会図書館リサーチ・ナビにおけるテーマグラフの生成, 情報処理学会 第 96 回情報学基礎研究会 (FI-96-5), 秋葉原ダイビル, 東京都, 2009 年 11 月.
- [発表 7] 森 竜也, 増田 英孝, 清田 陽司, 中川 裕志: Wikipedia を活用した言語間差異比較システムの提案, 第 2 回データ工学と情報マネジメントに関するフォーラム/第 8 回日本データベース学会年次大会(DEIM2010), 淡路夢舞台国際会議場, 兵庫県, 2010 年 3 月.
- [発表 8] 野田 陽平, 清田 陽司, 中川 裕志: ネットワーク構造を利用した Wikipedia からの意外性のある情報の抽出, 情報処理学会 第 72 回全国大会 講演論文集 (3K-4), 東京大学, 東京都, 2010 年 3 月.
- [発表 9] 坂井 哲, 増田 英孝, 清田 陽司, 中川 裕志: Wikipedia と図書館情報資源による調べ方自動提示システム, 情報処理学会 第 72 回全国大会 講演論文集 (3K-5), 東京大学, 東京都, 2010 年 3 月.
- [発表 10] 森 竜也, 増田 英孝, 中川 裕志, 清田 陽司: Wikipedia における言語間の差異マイニング, 情報処理学会 第 72 回全国大会 講演論文集 (5ZN-6), 東京大学, 東京都, 2010 年 3 月.

## 特記事項

- [特記 1] 清田 陽司: Google はなぜ的確に探せるのか(Web連載), 第 5 回『悟空、秘技「分から書き」を習う』/第 6 回『悟空、単語辞書を手に入れる』, ASCII.jp, アスキーメディアワークス, 2009 年 4 月～7 月.

## 報道関連

- [報道 1] 日本経済新聞 2009 年(平成 21 年)5 月 13 日 13 面: 東大発 VB のリッテル、商品検索簡単に.

# 知識発見と機械翻訳の自然言語処理要素技術

## —機械学習テンプレートライブラリ—

清水伸幸

### 1 概要

本研究の目的は、英語、日本語、中国語など、自然言語で書かれたテキスト情報から、有益な知識を効率的に発見するための自然言語処理の基礎技術を発展させることにある。具体的には、自然言語処理の要素技術として、以下の5つのトピックについて研究を進めた。

- 英語読解サポートシステム
- 品詞タグ付け
- 言語ヘッジ検出
- 構文解析
- 談話理解
- 逐次学習

また、これらの成果を一般に公開するため機械学習テンプレートライブラリを開発した。

- 機械学習テンプレートライブラリ

### 2 英語読解サポートシステム

#### 2.1 背景

日本語のネイティブスピーカーにとって、第二外国語として英語を習得するには多大な労力が必要となる。特に、英語の文献を読む上で、辞書を引く作業は初学者には時間を取る大変な作業である。

#### 2.2 内容

この辞書引きの作業を機械的にサポートするため、クリックで辞書を引く UI をまず作成した。その上で、辞書を引く履歴から、逐次的な機械学習で、読者であるユーザの語彙力を自動的に推定し、ユーザが知らないと推定される単語の定義を UI 上に示す事で英語読解をサポートするシステムを構築した。機械学習のモデルとしては、テスト項目の難易度を定める **Item Response Theory (IRT)** を応用したロジスティック回帰モデルを利用し、単語とユーザの素性を入れた分類問題として定義した。結果として、数クリックでユーザに適応し、74%程度の精度でユーザの語彙を当てる事に成功した。クリック数を増やす事でこの精度は向上し、80%以上まで伸びる事が実験の結果、示された。

## 2.3 具体的成果

本研究の結果は[査読付 1]において掲載され、また、[発表 3]において発表した。

## 3 品詞タグ付け

### 3.1 背景

品詞タグを推定し、文章内において単語がはたしている動詞、名詞などの機能を確認する事は、自然言語処理において尤もよく使われる前処理であり、日本語においても、仮名漢字変換に使われる重要な基礎技術として、非常に重要な位置を占めている。そこで、近年高性能であるとされる幾つかの学習手法を系列分類学習に応用し、品詞タグ付けの精度の向上を目指す。

本研究テーマでは、近年発達した様々な最先端の線形分類器の学習手法による英語と日本語の品詞タガーを比較し、タガーの詳細なエラー解析をすることで、機械学習的、かつ、言語的な知見を得ることを目的とする。

### 3.2 内容

品詞タグ付けは、機械学習テンプレートライブラリ(Shimizu & Miyao, 2007)を用いて行った。機械学習テンプレートライブラリを用いることにより、様々な最先端の学習手法の比較を、同じ素性集合に対し容易に行うことが出来る。英語の品詞タグ付けは、前年の実験の結果での精度は 97.19%であった。今回、さらに素性を見直してチューニングを行ったところ、精度は 97.21%に達した。この精度は英語の品詞付け問題において *state of the art* であり、著者の知る限り同じテストセットで二番目に位置する。

また、日本語の日本語においても、標準データセットである *Kyoto Corpus 4.0* を用いた *Coarse Pos Tagging* のタスクにおいて、95.35%の精度を出した。この精度は、現在の品詞タガーの最高精度には及ばないものの、IPA などの辞書を利用しないタガーの中では最も性能が高い部類に含まれる。エラー解析の結果では、単語の切れ目に起因するエラーが多く見られる事が分かった。

### 3.3 具体的成果

本テーマでチューニングされた素性、機械的に予測した品詞タグは、次の言語ヘッジ検出のタスクにおいて利用された。また、学習済みのソフトウェアを機械学習テンプレートライブラリの一部に組み込まれた。

## 4 言語ヘッジ検出

### 4.1 背景

品詞ヘッジとは、ここでは話し手または書き手が述べている内容に対する不確実性を表す表現のことである。情報抽出など、自然言語処理で知識をテキストから抽出するにあたり、不確実性が高い情報とそうでない情報を区別するため、言語ヘッジの検出は重要なタスクである。

### 4.2 内容

本テーマでは、英語の言語ヘッジの検出が課題である *CoNLL2010* のタスクを系列分類問題として定式化し、機械学習テンプレートライブラリを使用して、キーワード、あるいはキーフレーズとなる曖昧表現を予測した。Shared Task ではこの問題に対して24チームが参加し、7位に相当する分類精度 85.08%を出した。1位の 86.35%との差は 1.2%ほどであるため、システムとしてはまずまずの性能を出したと言える。

### 4.3 具体的成果

本研究は CoNLL2010 Shared Task のホームページにおいて、システムの性能が公開されている。システムの概要は 2010 年度発表へ向けて投稿中である。

## 5 構文解析

### 5.1 背景

テキストマイニング、機械翻訳の双方において、文章の構文を解析し、係り受け関係を見つけることは自然言語処理の重要な基礎技術である。ソース言語からターゲット言語への翻訳や、テキストの意味解析、また、上記の言語ヘッジ検出タスクにおいても、ヘッジされた範囲を特定するには、文章の係り受け関係を正しく解析することが重要である。

### 5.2 内容

本研究テーマでは、英語の HPSG パーザを改良する上で重要な課題となる、構文木の構成方法について新たに Default Unification と呼ばれる決定論的な探索手法を新たに提案し、決定論的な方法意外に、ビームサーチやバックトラッキングとの比較を行った。

### 5.3 具体的成果

本研究は[査読付 3]において掲載された。

## 6 談話理解

### 6.1 背景

自然言語処理の中でも、文単位ではなく、談話の単位での処理は、自動要約を除いて大きな進展が多くない。特に、内容の理解にまで踏み込んだ研究は少ない。課題として、言語の理解そのものの評価が難しいことがあげられる。このため、オフィスで迷っている人を案内する、という設定で学部生から英語による道順案内の談話データを収集し、指示の通りに道順をたどれるか否かで評価することとした。コーパスの例としては次のものがあげられる。

"make first right then turn into second door on your left"

"Just head straight through the hallway ignoring the rooms to the left and right of you, but while going straight your going to eventually see a room facing you, which is north, enter it."

"Head straight. continue straight past the first three doors until you hit a corner. On that corner there are two doors, one straight ahead of you and one on the right. Turn right and enter the room to the right and stop within."

これらの道順案内を理解するため、意味論にわたる全部分を一つの機械学習システムで対応する手法を考案した。

### 6.2 内容

本研究では、道順を案内する談話コーパスに対して、スタート地点から、最終目的地到達に必要な動作のラベル付けを行い、機械学習の問題として定義した。この問題に対して、マルコフとセミマルコフモデルを用いた分析を行った。この際、モデルの粗な部分のラベルを分解し、単純なマルコフ仮定を置いた系列ラベリングではなく、セミマルコフ仮定を置いた可変長系列ラベリングを改良する



ことで、性能を向上させる事に成功した。評価としては、理解のレベルでテストするため、**Simulated Robot** が単純化されたオフィス空間を案内通りに目的地にたどり付けるか実験し、88%という高い目的達成率を得た。

### 6.3 具体的成果

本研究は[査読付 3]において掲載された。

## 7 逐次学習

### 7.1 背景

多くの機械学習ではデータ全体に対して1つの最適な関数を求めるため、大量のデータに対して学習を行う際に計算コストの面から困難が生じる場合があるが、逐次学習ではデータを1つずつ受け取りながら逐次的に識別関数を学習していく。近年のデータの爆発的な増加に伴い、教師ありデータを用いた機械学習においても、逐次的な学習を行う必要性が高まっている。

### 7.2 内容

Crammerらによって提案された **Passive-Aggressive** アルゴリズム(PA アルゴリズム)は代表的な逐次学習の一種だが、多クラス識別問題において、厳密な PA アルゴリズムの枠組みから外れた近似解法が用いられてきた。本研究は本来の PA アルゴリズムの厳密解をサポートクラスという概念を用いて導出し、これに基づく識別関数の更新を効率的に行うためのアルゴリズムを提案した。

### 7.3 具体的成果

本研究は[発表 1]と[発表 2]において発表し、また、[発表 2]においては学生奨励賞を受賞した。

## 8 機械学習テンプレートライブラリ

### 8.1 背景

近年、自然言語処理における機械学習の研究では、ほぼ同じデータ構造で、損失関数や最適化アルゴリズムが異なる手法が多く提案されている。本研究では、以下のような式で識別関数を定義する機械学習モデルを対象とする。

$$\eta(y|x) = f(w \cdot \Phi(x, y))$$

ここで、 $x, y$  は履歴とターゲット、 $\Phi$  は素性関数、 $w$  はパラメタ(重み)ベクトル、 $f$  は任意の関数である。例えば、文書分類では  $x$  は文書で  $y$  はラベル、品詞タグ付けでは  $x$  は文で  $y$  は品詞列、係り受け解析では  $x$  は文で  $y$  は係り受け木となる。

現在自然言語処理で利用されている多くの機械学習手法は、上記のように定式化することができる。そこで、木構造や連鎖構造など、言語の構造を表現するのに適したデータ構造と、そこに依存する計算(デコード、期待値計算など)と、学習アルゴリズム(損失関数、最適化アルゴリズム)を切り離し、どちらを改良しても残りの部分を利用可能なよう、設計した機械学習テンプレートライブラリを開発した。

### 8.2 内容

上記の成果のうち、品詞タグ付け、言語ヘッジ検出、逐次学習を機械学習テンプレートライブラリに実装した。さらに、学習器が利用する主なメソッドの一つ、デコード関数

- デコード (decode)

$\eta(y|x)$  を最大にする  $y$  を求める。逐次学習に必要で、テスト時の予測にも用いられる。系列ラベル問題では、ビタビアルゴリズムが代表的なデコード方法である。

を新たに係り受け解析用に定義した。この新たなデコード方法は係り受け問題を整数線形計画法で解くものであり、従来の MST 法では入れる事の出来なかった親と孫ノード間の素性を入れる事が可能になった。この整数線形計画問題を拡張する事で、依存構造を必要とする様々な自然言語処理のタスク (Semantic Role Labeling 等) が、エラーの入る余地が少ない同時推定 (joint inference) で行えるようになり、今後の研究の展望として非常に有望なコンポーネントがそろった。

また、系列ラベルの分類器には、新たに以下の関数を追加した。

- k ベスト (k-best)

$\eta(y|x)$  を最大にする  $y$  の中で、スコアが最も高くなる構造を  $k$  個求める。上記の成果の一つ、逐次学習で厳密な解を求める PA アルゴリズムに必要である。また、テスト時の予測でも、他の条件からデコードの解にエラーが発見される場合、 $k$  ベストがあれば、次点の  $y$  となる構造が利用できる。

この関数を追加する事で、新たな逐次学習手法が系列ラベル問題に適用できるようになり、また、 $y$  の構造に隠れ変数がある場合も、隠れ変数の周辺化を行うことが出来るようになる。これを用いて、Latent Variable Perceptron for Structured Classification を実装した。

また、テンプレートライブラリを以下の日本語の自然言語のタスクに対応するよう改良し、素性選択を行った。

- Kyoto Corpus 4.0 品詞付け (Coarse POS) --- 95.4% 辞書なし
- Kyoto Corpus 4.0 文節分け --- 99.6% 辞書なし
- Kyoto Corpus 4.0 依存構造解析 --- 90.5%
- IREX 固有表現抽出 --- 65.0% 辞書なし、単語区切りなし

上記のパーセンテージはそれぞれのタスクでの性能を表す。さらに、英語のコーパスであるが、CoNLL2007 コーパスの英語の係り受けについては Unlabeled Attachment で、88.7%を記録した。これらの数値は、現時点での最高性能ではないが、辞書を利用しない等の点で問題設定が若干違うため厳密な比較にはならない。ただし、それぞれの問題設定では概ね競争力のある結果と言える。学習には情報基盤センターのスーパーコンピューターを利用した。当該研究課題は、特任プロジェクト「機械学習テンプレートライブラリの拡張と実証実験」として進められ、情報基盤センターから援助を受けて実施した。

### 8.3 具体的成果

機械学習テンプレートライブラリは、ソースコードと、典型的なデータ構造・学習器についてバイナリパッケージを公開する予定である。本ライブラリは上記の成果の一つである言語ヘッジ検出に利用された。また、ライブラリの一部ではなく、目的が特化されているが、[査読付 3] において実装されたシステムは同じコードベースを基にしている。

## 9 成果要覧

### 査読付論文



- [査読付 1] Y. Ehara, N. Shimizu, T. Ninomiya and H. Nakagawa: Personalized Reading Support for Second-language Web documents by Collective Intelligence, in Proceedings of International conference on intelligent user interfaces (IUI 2010), Hong Kong, China, (Acceptance rate 22%) , 2010.
- [査読付 2] T. Ninomiya, T. Matsuzaki, N. Shimizu and H. Nakagawa: Deterministic shift-reduce parsing for unification-based grammars by using default unification, in Proceedings of European Association for Computational Linguistics (EACL 2009), pp. 603-611, (Acceptance rate 28%), 2009.
- [査読付 3] N. Shimizu and A. Haas: Learning to follow navigational route instructions, in Proceedings of the Twenty-first International Joint Conference on Artificial Intelligence (IJCAI 2009), pp. 1488-1493, (Acceptance rate 25.7%), 2009.

### その他の発表論文リスト

- [発表 1] 松島慎, 清水伸幸, 吉田和弘, 二宮崇, 中川裕志: サポートクラスによる Passive-Aggressive アルゴリズムの多クラス化, 情報処理学会 第 192 回自然言語処理研究会, 2009-NL-192(12), pp. 1—6, 2009 年 7 月.
- [発表 2] 松島慎, 清水伸幸, 吉田和弘, 二宮崇, 中川裕志: 多クラス識別問題におけるオンライン学習のための 厳密な PA アルゴリズム, 報処理学会創立 50 周年記念 (第 72 回) 全国大会, 2010 年 3 月.
- [発表 3] 江原遥, 二宮崇, 清水伸幸, 中川裕志: Web 文書中のユーザが知らない語を予測する読解支援システム, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会, 2010 年 3 月.



# ネットワーク研究部門

## ネットワーク研究部門概要

若原 恭

ネットワークの機能の高度化と設計・運用・保守・管理の合理化  
—アドホックネットワーク、車車間・路車間通信、P2Pネットワーク—

中山 雅哉

広域分散環境の高度基盤技術に関する研究

小川 剛史

実空間における状況認識およびデータ送受信に関する研究

関谷 勇司

連邦型クラウド構築のための技術研究ならびに  
広域分散サービス分析のための計測手法の研究

中村 文隆

P2Pネットワーク計測 ネットワークシミュレーションの信頼性 音声処理言語

加藤 朗

ネットワークの運用とセキュリティに関する研究



# ネットワーク研究部門

## 概要

部門長 若原 恭

情報通信ネットワークは、情報社会の発展とともに、経済・社会・文化等に係わる日々の活動に必要不可欠な社会基盤となっている。このようなネットワークへの要求は拡大する一方であり、これに応えるため、ネットワーク自体の構築と運用、ネットワーク利用基盤、ネットワーク応用システム、ネットワークセキュリティの維持確保等に関し、基礎技術から実用技術まで幅広く研究に取り組んでいる。2009年度の取り組みの概要は以下の通りである。

ネットワーク自体の構築と運用に関しては、ルータや基地局等のインフラ設備を必要としないアドホックネットワークやセンサーネットワークが今後の本格的なユビキタス社会において必須になるとの考えから、その基本課題の解決に取り組んだ。

アドホックネットワークについては、複数のノードから構成される情報転送ルートが、構成ノードの移動によって不安定になるという問題を解決するためのルーティング技術について検討した。特に、不安定となったルートを迅速に検知し修復することが可能な方法を考案し、その有効性を理論的に確認した。また、アドホックネットワークの応用の一例として、車車間・路車間通信によって、走行中の車両からインターネットにアクセスするための方法について検討を進めた。特に、車両からアクセスすべき基地局や利用すべき無線チャンネルを適切に選択する手法を考案し、シミュレーション実験により情報転送品質の向上と転送容量の拡大を達成できることを確認した。

センサーネットワークについては、センシングデータを基地局へ収集する際の消費電力を削減するため、放送型データ配信を用いたノード制御方式について検討した。具体的には、まず、ノードの故障に対応できるよう、これまでに考案した制御方式を機能拡張した。次に、シミュレーション実験を行って、この拡張手法により、ノード移動量を削減して消費電力を抑えるとともにデータ収集量を向上できるという特長を達成できることを実証した。更に、各ノードがセンシングするデータをサーバが予測し、プッシュ型放送を用いて予測値を配信して、予測値と異なるデータのみをノードから収集することによって、通信量を総合的に削減する方式についても検討を進め、その有効性を明らかにした。

ネットワークを利用するための基盤として、多量データの効率的な転送技術の研究に取り組んだ。一般に、各ノードが広帯域な環境でネットワークに接続されていても、遠く離れたノード同士でTCP方式を用いたデータ転送を行うと、帯域・遅延積が大きくなることにより、ネットワークの帯域を十分に利活用したデータ転送を行うことができなくなる。このLong fat pipe問題を解決するため、本来はエンドノード間で行われるデータ転送を、中継ノードを介して2つ以上のデータ転送に分割することで、仮想的にノード間の遅延を短縮させてデータ転送の高速化を図る方式について検討した。そして、PlanetLab環境を用いた実証実験を行い、この仮想データ転送技術が、世界中に点在したノード同士でTCPによるデータ転送を行う際に極めて有効であることを実証した。

昨今、ネットワーク基盤として、特定のサーバを持たず各ノードが対等な役割を持つ

P2P(Peer to Peer)アプリケーションの利用が増え続けている。特に、ファイル交換用P2Pネットワークは発展が著しく、トラフィック的にも大きな割合を占めるようになった。この結果、P2Pネットワークのトラフィックが、アンダーレイネットワークに与える影響が問題となっている。そこで、ISPにおいて適切にキャッシュを行うことによりP2Pネットワークから発生するトラフィックの削減を行うことを目的として、実稼働しているP2Pネットワーク上のコンテンツ人気度を測定するシステムを開発して実際に測定し、人気度特性にカットオフが存在する場合としない場合とがあることを確認した。カットオフ特性がある場合には、効率よいキャッシュ運用によってISP間トラフィックを削減できる見通しを得た。

ネットワークの利用に欠かせないもう一つの基盤はDNS(Domain Name System)であり、DNSに関し様々な検討を行った。まず、DNSでは一度問い合わせた情報をキャッシュするが、その保持時間TTLの動的調節手法を考案し、現場実験によって、負荷分散や攻撃に対応するための頻繁なDNSデータの更新を、キャッシュ情報の不整合なく実施可能であることを実証した。次に、インターネット上に存在する数万台のDNSサーバが連携することにより一つのデータベースを形成しているDNSの安定性と信頼性を向上させるためには、DNSの運用状態を把握する必要があることから、低コストかつ容易に運用・計測を行うことのできる分散計測環境を、計32台の計測機器を世界各地に設置して構築し、これを使って計測ならびに分析を行った。更に、DNSサーバのなりすまし対策としてDNS応答の署名連鎖に基づくDNSSECを実際に適用する際に、DNSサーバへの処理負荷や応答メッセージ長の増大の影響を評価するためのプロジェクトDURZに参画し、一つのRoot DNSサーバに実験機能を実装して評価用データの収集を開始した。

ネットワークの応用システムとしては、最近クラウドコンピューティングが急速に拡大している。この技術を用いてクラウド同士を相互接続し、広域に展開される連邦型クラウドの形成を目指した研究を開始した。具体的には、大学間においてリソースを融通しあうことによって、計画された停電やメンテナンス等に左右されることがない、運用コストを削減できる、各大学の運用ポリシーに柔軟に適応できる、仮想マシン環境のマイグレーションを柔軟に行うことのできる等の特長を持つ連邦型クラウドネットワークを、他の3大学と共同で開発し、基本的な実験を行って所期の性能を達成できることを確認した。

ネットワークの応用システムは極めて豊富になっているが、中でも、実際にはユーザが直接見ることのできない情報を現実空間に提示可能な拡張現実感(AR)技術は、ユーザのナビゲーションを行うシステム等で検討されており、今後の実用の期待が極めて大きい。しかし、ナビゲーションシステムでは、広大な現実空間に存在する膨大な量の注釈情報を、リアルタイムに変化するユーザの位置や視点に応じて検索し、適切な場所に効率よく提示することが実用上の鍵となる。そこで、特にヘッドマウンティッドディスプレイ(HMD)を利用する際、ユーザの様々なコンテキストに応じてクライアントへ提示する情報とその量を制御する方式について検討を行った。具体的には、「座る」・「歩く」・「走る」といったユーザの身体動作から、時々刻々と変化するユーザの「注釈情報への注意の向けやすさ」を識別し、これに基づいて提示情報の詳細度を切り替える手法を考案し、実証実験を行って、その有効性を確認した。

情報ネットワークの発展とともに、そのセキュリティ面の問題が顕在化してきている。上述の通り、ファイル交換用P2Pネットワークは発展が著しいが、交換されるファイルの中には意図的に改ざんされたものや著作権上違法なもの等も含まれ、適切なファイル交換が阻害されている。この問題を解決するため、ユーザの認証とファイル提供ユーザを集中的に管理する方式を考案し、その有効性を評価した。その結果、ネットワークに存在する悪意ユーザと悪意ファイルを次第に排除可能であることを実証でき、これによってファイル交換の安全化を達成できるとの見通しを得た。

## キャンパスネットワーク研究部門 成果要覧

### 受賞

[受賞 1] 石原 知洋, 関谷 勇司, 村井 純: 平成 21 年度情報処理学会論文賞, 情報処理学会, 2010 年 3 月.

[受賞 2] 中山 雅哉: 通信ソサイエティ活動功労賞, 電子情報通信学会, 2009 年 9 月.

### 著書／編集

[著書 1] 関谷 勇司 他: ロボット情報学ハンドブック, ナノオプトニクスエナジー, 2010 年 3 月.

[著書 2] 中村 文隆: グラフィカル言語 PureData による音声処理, CQ 出版社, 2009 年 9 月.

### 査読付論文

[査読付 1] Pakornsiri Akkhara, Yuji Sekiya and Yasushi Wakahara: Efficient Alarm Messaging by Multi-channel Cut-through Rebroadcasting based on Inter-Vehicle Communication, IAENG International Journal of Computer Science, 36:2, IJCS\_36\_2\_07, May 2009.

[査読付 2] Wei Tian, Yuji Sekiya, Yasushi Wakahara and Yoshiaki Tanaka: Centralized P2P Management System based on File and Peer Reputation, International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC2009), pp.1062-1065, Jul. 2009.

[査読付 3] Pakornsiri Akkhara, Yuji Sekiya and Yasushi Wakahara: Fast Dissemination of Alarm Message based on Multi-channel Cut-through Rebroadcasting for Safe Driving, in Ao Sio-Iong, Castillo Oscar, Huang Xu (Eds.): Intelligent Automation and Computer Engineering, Lecture Notes in Electrical Engineering, Springer for IMECS 2009 Paper Number:ICCSA\_70, Vol.52, 2009.

[査読付 4] 樽林勇氣, 中山雅哉: ボトルネック RTT の最小化に基づく TCP 分割手法の提案と PlanetLab 環境による実証評価, インターネットコンファレンス 2009, pp.25-31, 2009 年 10 月.

[査読付 5] Shinya Kitajima, Tomoki Yoshihisa, Takefumi Ogawa, Takahiro Hara, and Shojiro Nishio: A Broadcast-Based Data Gathering Method considering Energy Consumption for Sensor Networks, Proc. of International Workshop on Sensor Network Technologies for Information Explosion Era (SeNTIE2009), pp.536-541, May 2009.

[査読付 6] 廣瀬 康一, 小川 剛史, 清川 清, 竹村 治雄: 複数の仮想空間を操作するマルチビューポートインタフェースのフレームワークに関する評価, 日本バーチャルリアリティ学会論文誌, Vol.14, No.2, pp.203-211, 2009 年 6 月.

[査読付 7] Daisuke Takada, Takefumi Ogawa, Kiyoshi Kiyokawa, and Haruo Takemura: A Context-Aware AR Navigation System Using Wearable Sensors, Proc. of The 13th International Conference on Human-Computer Interaction (HCI2009), No.3, pp.793-801, Jul. 2009.

[査読付 8] 藤井 裕士, 小川 剛史: Integral Histogram を用いたカメラ画像からの高速物品検出手法の評価, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム(DICOMO2009) 論文集, pp.636-644, 2009 年 7 月.



- [査読付 9] 高田 大輔, 小川 剛史, 清川 清, 竹村 治雄: 身体動作に基づき提示情報を切り替えるコンテキストウェアなウェアラブル AR システム, ヒューマンインタフェース学会論文誌, Vol.12, No.1, pp.47-56, 2010 年 2 月.
- [査読付 10] Takefumi Ogawa, Tatsuya Shinjo, Shinya Kitajima, Takahiro Hara, and Shojiro Nishio: Node Control Methods to Reduce Power Consumption Using Push-Based Broadcast for Mobile Sensor Networks, Journal of Mobile Multimedia, Vol.6, No.2, pp.114-127, 2010.
- [査読付 11] Tomohiro Ishihara and Yuji Sekiya: DNS cache-record coordination Using Adaptive Cache TTL Adjustment, International Journal of Computer and Network Security, Vol.2, No.1, pp.66-69, Jan. 2010.

### その他の発表論文

- [発表 1] 田 偉, 関谷 勇司, 若原 恭, 田中 良明: 悪意ファイル排除のための P2P 管理システム, 電子情報通信学会・技術研究報告, IN2009-33, pp.49-54, 2009 年 7 月.
- [発表 2] 田 偉, 関谷 勇司, 若原 恭, 田中 良明: 悪意ピアと悪意ファイルの排除機能をもつ P2P 管理システムの評価, 電子情報通信学会・通信ソサイエティ大会, B-6-33, 2009 年 9 月.
- [発表 3] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間路車間通信を用いたストリーミング配信におけるパケット到達率最大化のための基地局選択手法, 電子情報通信学会・技術研究報告, NS2009-112, pp.49-54, 2009 年 11 月.
- [発表 4] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間・路車間通信におけるアクセス先基地局の選択及び無線通信チャンネル制御による帯域利用効率化手法, 電子情報通信学会・技術研究報告, NS2009-251, pp.501-506, 2010 年 3 月.
- [発表 5] Liang Zilu, Takefumi Ogawa and Yasushi Wakahara: An Automatic Route Recovery Method Based on Reserved Virtual Local Paths for Wireless Ad Hoc Network, 電子情報通信学会・総合大会, 情報・システムソサイエティ特別企画, ISS-P-134, 2010 年 3 月.
- [発表 6] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間・路車間通信によるインターネットアクセスにおいて無線帯域の利用効率向上を図る経路制御手法, 電子情報通信学会・総合大会, B-15-14, p.630, 2010 年 3 月.
- [発表 7] 江崎 浩, 中山 雅哉, 松浦 知史, 井上 博之, 落合 秀也, 山内 正人: JGN2plus SPARC における Live E! 研究活動, Live E! シンポジウム 2009 予稿集, pp.111-116, 2009 年 9 月.
- [発表 8] 樽林 勇氣, 中山 雅哉: ボトルネック RTT に基づく TCP 分割手法の提案と実証評価, 電子情報通信学会・総合大会, B-6-48, 2010 年 3 月.
- [発表 9] Yuji Sekiya: Behavior Analysis of DNS Anycast in 2008, The 2nd CAIDA/WIDE/CASFI Workshop, Apr. 2009.
- [発表 10] Yuji Sekiya: Gulliver Project - status update in 2009, ISMA 2010 AIMS-2 Workshop on Active Internet Measurements, Feb. 2010.
- [発表 11] 小林 正裕, 亀井 聡, 中村 文隆: 実データに基づく P2P キャッシュ方式の検討, 電子情報通信学会・総合大会, B-6-73, 2010 年 3 月.

### 特記事項

- [特記 1] 中村文隆: ネットワークシミュレーション概論, 電子情報通信学会 NS 研究会シミュレーショ

ンサマースクール, 2009年8月.

# ネットワークの機能の高度化と設計・運用・保守・管理の合理化

## － アドホックネットワーク、車車間・路車間通信、 P2P ネットワーク －

若原 恭

### 1 概要

情報通信ネットワークは、そのアプリケーションの急激な発展とともに、転送情報量の増大と規模的な拡大に加え、機能面の質的な変革も必要となる。このような認識のもと、情報通信に係わる各種リソースの最大限の利活用によって、複雑化・大規模化が進む一方の情報通信ネットワークの機能面における高度化のあり方を探求するとともに、そのようなネットワークの設計・運用・保守・管理をできる限り合理的なものにすることを目標にしている。具体的には、以下に示す諸技術に焦点を当て、理論と実験の両面から研究を進めた。

- アドホックネットワークの運用制御

主として移動可能な無線ノードから構成されるアドホックネットワークは、本格的なユビキタス社会において必須になると考えられている。そのようなアドホックネットワークでは、運用と制御に係わる多くの基本技術が課題として山積している。そのうち、ノードの移動によって切断され不安定となる情報転送ルートの効率的な維持方法について検討を進めた。

- 車車間・路車間通信を応用した車両からのインターネットアクセス

車社会と情報社会の相乗的な発展に伴って、車両からのインターネットへのアクセスの要求が高まっている。その実現法として、既の実証実験も行われており、アドホックネットワークの一例と見なすことが可能な車車間・路車間通信の応用を検討した。特に、車両の移動等によって情報転送の品質が劣化が避けられないという問題に対し、限られた無線リソースを有効に活用することによって解決する方法を中心に検討した。

- 安全なファイル交換用 P2P ネットワーク

昨今、音楽や映像等ユーザが持つファイルを自由に交換する目的に利用されるファイル交換用 P2P ネットワークの発展が著しく、トラヒック的にもインターネットで流れるパケットの大きな割合を占めるようになった。しかし、交換されるファイルの中には意図的に改ざんされたものや、本来は機密扱いのもの等も含まれ、適切なファイル交換を損なっている。そこで、このような問題を解決することを目標として、ユーザの認証とファイル提供ユーザの管理によって、ファイル交換用 P2P ネットワークの安全化を図り安心して利用できる方式の検討を進めた。

## 2 アドホックネットワークの運用制御

### 2.1 背景

アドホックネットワーク MANET(Mobile Ad-hoc Network)は、ルータやアクセスポイント等のインフラ設備を用いることなく、主として移動可能な無線端末ノードから一時的に構成されるネットワークである。例えば、イベント会場や旅行先での情報交換、走行中の車同士の間での情報交換等、近い将来と予想される本格的なユビキタス社会において重要な役割を果たすものと考えられている。このため近年研究が活発になってきているが、未解決の基本的な技術課題も多く、その一つとして、ノードの移動によって情報転送ルートの維持が困難になるという問題を解決するための検討を進めた。

### 2.2 内容

一般に MANET における任意の 2 個のノード間での情報授受では、直接電波が届かない場合、他のノードを中継するルートに従ってパケット単位で情報を転送する。しかし、ノードの移動によってルートが切断されることがあり、その対処は、送信側のノードからルートを設定し直すことが基本となる。しかし、ノードの移動によってはこのルートの再設定の頻度が高くなり、それに伴って情報転送の中断や情報の損失等品質の劣化が大きくなる。また、ルート設定に伴う制御情報の交換も多くなりトラフィック負荷も大きくなる。これらの問題を解決するため、概略次の原理に基づき、ルートの切断を検知し修復する方法を考案した。

- ・ルート周辺のノードは、常にルート内のノードが送受するパケットをモニタし続ける。
- ・このようなモニタにおいて、例えば送信ノードがデータパケットを送信した後一定時間以内に受信ノードがそれを中継しないことを認識した場合、受信ノードがデータパケットを受信できずルートが切断したと判断する。
- ・ルートが切断したと判断したモニタノードは、送信ノードにその旨伝えると同時に、受信ノードに対してデータパケットを送出することによってルートを修復する。

この原理に基づくルート切断の自動検知修復法によって、次の特長が得られる。

- ・ルート切断後直ちにデータパケットを再送でき、修復に要する時間が極めて小さい。
- ・実質的にルート全体を再設定するのではなく、切断したリンク付近でのローカルな処理でルートを修復させるので、ルート修復に伴う制御パケットのトラフィック負荷が極めて小さい。

しかしながら、この方法では、ルート周辺の全ノードが常にパケットをモニタする必要があるため、それに起因する消費電力や処理オーバーヘッドが大きくなるので、今後はモニタノードの削減等によってこれらの問題を解決できるよう改善を図っていく計画である。

### 2.3 具体的成果

前節で述べたルート切断の検知修復法は、今後の期待が大きいアドホックネットワーク MANET が現在抱えている基本問題の解決を目指すもので、机上検討によってその効果が明確であることを理論的に確認できた。従って、この方法は MANET の実用と発展に寄与するものと期待され、学会においてポスター発表等を行った。[発表 4, 5]。今後、改善を図るとともにシミュレーション実験等によって、その効果を定量的に明らかにしていく計画である。

### 3 車車間・路車間通信を応用した車両からのインターネットアクセス

#### 3.1 背景

移動手段として車の利用は拡大する一方であり、車両からのインターネットアクセスの需要が益々大きくなってきている。これを実現する方法としては、WiMAX を含む携帯電話ネットワークの活用も考えられるが、インフラ設備の投資回収の負担が大きいという問題がある。一方、昨今、運転の安全性や快適性の向上等を目的として、運転者や同乗者に様々な情報を提供するため車車間通信や路車間通信を活用する方向性が強くなってきており、一部では実証実験も行われている。このように別目的で実現される車車間通信や路車間通信を、車両からのインターネットアクセスに利用する方法は、インフラ設備の構築が比較的容易で投資回収も早く済むことが期待できる。そこで、道路に沿って基地局が設置され、他車両の中継による車車間通信に依って自車両から基地局までアクセスし、基地局からは光ファイバ経由でインターネットに接続できるという環境を想定して、車両からのインターネットアクセス技術について検討を行った。

#### 3.2 内容

車両からのインターネットアクセスを実現するための技術について、次の2段階に分けて検討した。

##### (1) 基地局の選択

一般に、車両からアクセス可能な基地局は複数あり、1本道では前後2台の基地局にアクセス可能となる。自車両から基地局まで直接電波が届かない場合は、他の車両の中継する必要があるが、それら車両の位置や動きによって情報転送の品質が大きく変化する。そこで、品質の観点から、アクセスすべき基地局を適切に選択する手法を検討した。具体的には、品質としてパケット到達率を取り上げ、自車両からアクセスする基地局までのルートにおけるパケット到達率が最大となるように基地局と中継車両を選択する方針とした。このため、各基地局は一定時間ごとに制御パケットをブロードキャストし、これを受信した車両は自車両の位置情報を追加して中継を繰り返す方式を考案した。各車両は、自車両と基地局の間に存在する全車両の位置情報を得て、その中からパケット到達率が最大となるよう中継車両を最適に選択するが、この処理を効率よく行うためダイナミックプログラミングを応用する方針とし、その具体的アルゴリズムを設計した。

図1は、最寄りの基地局を選ぶ基本手法に比較し、この考案方法によって、パケット到達率を大幅に向上できることの一例を示す。なお、横軸は、インターネットにアクセスする車両の位置を、基地局からの距離で表している。

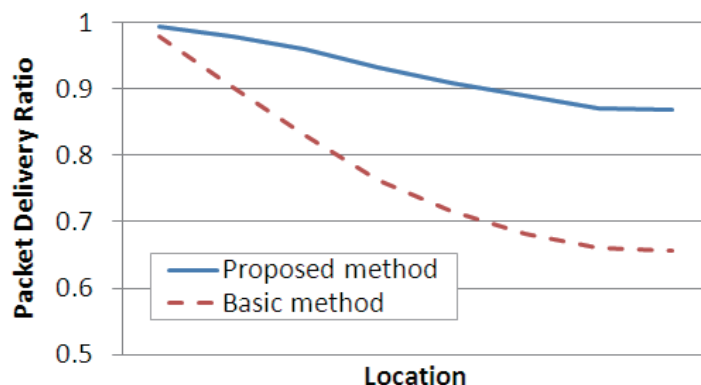


図1 考案法によるパケット到達率向上の効果の例

## (2) チャンネル割り当て

(1)では、利用できるチャンネルに制限がないことを前提としていたが、実際には無線リソースがかなり限定されるので、特にインターネットアクセスの需要が増えたときには適切なチャンネル選択によって転送できるトラフィック量を最大にすることが重要である。そこで、まず、基地局から一定周期で制御パケットを送出し、それを受けた車両は、送出ノードからの距離が長いほど小さな待ち時間後に中継して送出することとし、待ち時間中に他の車両が中継した場合は中継を中止する。その結果、中継車両が最少となるよう次の基地局まで制御パケットの中継を続けることとなり、この過程で中継した車両を、基地局が次に制御パケットを送出するまでの間のデータパケットの中継車両とし、情報転送ルートを構成する。次に、データパケットの送受信要求が生じた車両からアクセス可能な各基地局までのルートのうち、空きチャンネル容量が最小のリンクの空きチャンネル容量が最も大きいルートを選択する。更に、このルートにおいて基地局から順にチャンネルを割り当てることとし、このとき、いったん割り当てたチャンネルを優先して使用する。中継車両間の間隔が増減し、その結果中継車両の変更が必要となったときは、同時にチャンネルの再割り当てを行う。

この考案法によって、例えば、より近い基地局を選択してチャンネルを割り当てる基本手法に比較して、運べるトラフィック量を増大できる、ルートの設定に要する時間を小さくできる等のメリットが得られる。その効果の一例として、運べるトラフィック量についての比較評価結果を図2に示す。このように、特に車両密度が高く渋滞が発生するような場合やデータパケット量が多い場合に考案法の効果が特に大きいことが分かる。

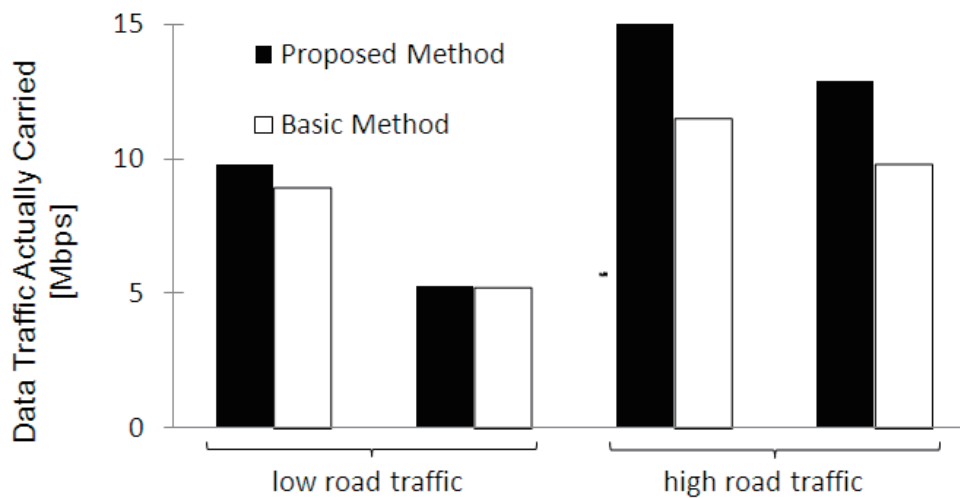


図2 運べるトラフィック量に関する比較評価結果の例

### 3.3 具体的成果

以上で示した通り、考案した車車間・路車間通信の応用によるインターネットアクセス法によって、転送情報量の増大や転送品質の向上等が実現できることを定量的に明らかにできた。このうち、基地局からの制御パケットのブロードキャストによる迅速かつ効率的なルート設定法については国際学会論文誌等で論文発表し、これ以外の技術に関する研究成果については国内で開催された学会の研究会や大会で論文発表を行った。[査読付 1, 3, 発表 3, 4, 6]



## 4 安全なファイル交換用 P2P ネットワーク

### 4.1 背景

昨今、ユーザが持つファイルを交換することが目的のファイル交換用 P2P ネットワークの発展が著しく、トラヒック的にもインターネットで流れるパケットの大きな割合を占めるようになってきた。しかし、交換されるファイルの中には、著作権上違法なもの、意図的に改ざんされたもの等が含まれ、この悪意ファイルの増大によって本来の適切なファイル交換が損なわれている。そこで、このような問題を解決することを目標として、交換する各ファイルの提供ユーザを特定可能とし、悪意ユーザと悪意ファイルを識別して排除することによって、ファイル交換用 P2P ネットワークの安全化を図る方式を検討した。

### 4.2 内容

ファイル交換における安全性を損なう悪意ファイルと悪意ユーザを排除するためには、P2P ネットワークへのユーザの参加時の認証に加え、悪意ファイルの特定と流通の停止、及び悪意ファイルを提供した悪意ユーザの特定と排除を行う機能が必要である。そこで、概略次の原理に基づく方式について検討を進めた。

- ・ユーザは P2P ネットワークの利用に際して認証を行う。また、各交換ファイルは提供ユーザが署名して初めてダウンロードできるようにする。
- ・ユーザがファイルをダウンロードした際にマニュアルでそのファイルが悪意か否かを判定する。悪意ファイルの場合、悪意ファイルや悪意ユーザを集中的に管理するセンターにその旨を通知する。
- ・センターでは、悪意ファイルを提供したユーザが、予め定めた悪意ユーザ判定条件を満たす場合は、そのユーザを悪意ユーザと判断し、悪意ユーザの P2P ネットワークへの参加を以降拒否する。
- ・ユーザがファイルをダウンロードする場合、その直前にこのセンターにアクセスし、それを提供したユーザが悪意ユーザか否か検査し、悪意ユーザが提供したファイルはダウンロードできなくする。

この原理に基づく P2P ネットワーク制御方式の有効性を実験評価するため、この方式に基づき悪意ファイル数や悪意ユーザ数等の時間的推移をシミュレーションした。図 3 にその結果の一例を示す。パラメータ  $\mu$  は悪意ユーザ判定のための閾値で、そのユーザが作成したファイルのうち悪意ファイルと判断されたものの割合が  $\mu$  を超えたとき悪意ユーザと判定される。図 3 から、考案した方式によって、P2P ネットワークにおけるダウンロード回数の増加に従って、そのネットワーク内の悪意ユーザが徐々に排除され、悪意ファイルが次第に低減していくことが確認でき有効性を実証できた。

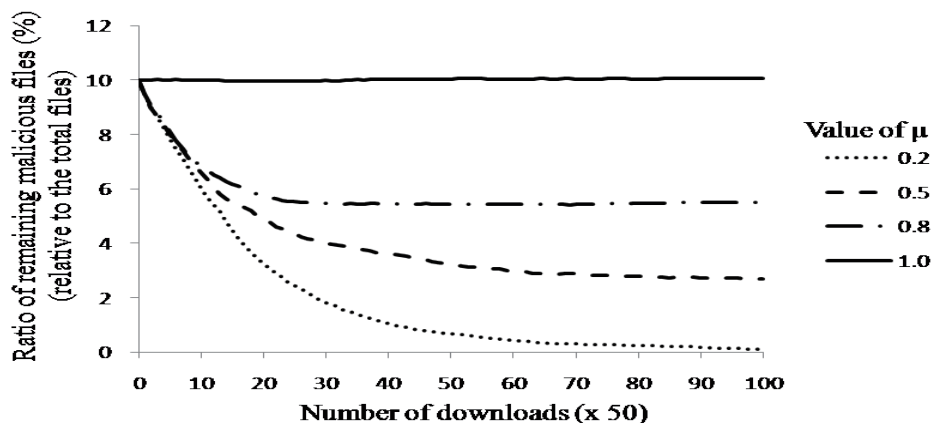


図 3 ダウンロードによって P2P ネットワークに残る悪意ファイルの推移の一例

### 4.3 具体的成果

安全なファイル交換を実現するために考案した P2P ネットワーク制御方式のアルゴリズムと定量的総合評価等に関する研究成果を国際学会や国内の研究会と大会で論文発表した。  
[査読付 2, 発表 1, 2]

## 5 成果要覧

### 査読付論文リスト

- [査読付 1] Pakornsiri Akkhara, Yuji Sekiya and Yasushi Wakahara: Efficient Alarm Messaging by Multi-channel Cut-through Rebroadcasting based on Inter-Vehicle Communication, IAENG International Journal of Computer Science, 36:2, IJCS\_36\_2\_07, May 2009.
- [査読付 2] Wei Tian, Yuji Sekiya, Yasushi Wakahara and Yoshiaki Tanaka: Centralized P2P Management System based on File and Peer Reputation, International Technical Conference on Circuits/Systems, Computers and Communications (ITC – CSCC2009), pp.1062-1065, Jul. 2009.
- [査読付 3] Pakornsiri Akkhara, Yuji Sekiya and Yasushi Wakahara: Fast Dissemination of Alarm Message based on Multi-channel Cut-through Rebroadcasting for Safe Driving, in Ao Sio-Iong, Castillo Oscar, Huang Xu (Eds.): Intelligent Automation and Computer Engineering, Lecture Notes in Electrical Engineering, Springer for IMECS 2009 Paper Number:ICCSA\_70, Vol.52, 2009.

### その他の発表論文リスト

- [発表 1] 田 偉, 関谷 勇司, 若原 恭, 田中 良明: 悪意ファイル排除のための P2P 管理システム, 電子情報通信学会・技術研究報告, IN2009-33, pp.49-54, 2009 年 7 月.
- [発表 2] 田 偉, 関谷 勇司, 若原 恭, 田中 良明: 悪意ピアと悪意ファイルの排除機能をもつ P2P 管理システムの評価, 電子情報通信学会・通信ソサイエティ大会, B-6-33, 2009 年 9 月.
- [発表 3] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間路車間通信を用いたストリーミング配信におけるパケット到達率最大化のための基地局選択手法, 電子情報通信学会・技術研究報告, NS2009-112, pp.49-54, 2009 年 11 月.
- [発表 4] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間・路車間通信におけるアクセス先基地局の選択及び無線通信チャネル制御による帯域利用効率化手法, 電子情報通信学会・技術研究報告, NS2009-251, pp.501-506, 2010 年 3 月.
- [発表 5] Liang Zilu, Takefumi Ogawa and Yasushi Wakahara: An Automatic Route Recovery Method Based on Reserved Virtual Local Paths for Wireless Ad Hoc Network, 電子情報通信学会・総合大会, 情報・システムソサイエティ特別企画, ISS-P-134, 2010 年 3 月.
- [発表 6] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間・路車間通信によるインターネットアクセスにおいて無線帯域の利用効率向上を図る経路制御手法, 電子情報通信学会・総合大会, B-15-14, p.630, 2010 年 3 月.

# 広域分散環境の高度基盤技術に関する研究

中山雅哉

## 1 概要

Internet に代表される広域分散環境は、今日の社会生活に不可欠な存在となっている。これまでは、コンピュータをノードとしてネットワークを構成することが中心であったが、最近では、気象センサーや振動センサーといった様々な小型デバイスもネットワークの構成要素として接続される様になってきた。これら小型デバイスから得られる情報は、社会生活への安心情報として広く活用されている。一方で、ネットワークに接続されるノードの規模が爆発的に拡大することで、各ノードから得られる情報をデータ統合ノードに集約する際の中継ネットワークの輻輳や遅延に関わる課題、データ統合ノードに集約された情報を広域に分散した多数の利用者に効率的に提供するための課題に関わる技術開発が不可欠なものとなっている。

本研究室では、これらのネットワーク基盤技術の研究に取り組んでいる。

## 2 自律分散ノード間での効率的な情報流通技術に関する研究

### 2.1 背景

今日の社会生活に不可欠な存在である広域分散環境 (Internet) には、最近では計算機だけでなく気象センサーや振動センサーなどの様々な小型センサーがノードとして接続される様になってきた。

これらの小型センサーは、自律分散ノードとして観測されたデータを集約ノードなどに常時送信する機能を持っており、集約ノードに集められた情報に基づいて、社会生活に対する安心情報としてネットワーク利用者に提供される様になってきている。

### 2.2 内容

ネットワークに接続される自律センサノードの規模が爆発的に拡大すると、各ノードから得られた情報をデータ統合ノードに集約する際の中継ネットワークの輻輳や遅延に関する問題や、データ統合ノードで得られた集約情報を広域に分散した多数の利用者に提供するための技術が必要不可欠なものとなる。

そこで、国内外におかれた各地の気象センサー情報を集約する際の課題を取り上げ、多くの利用者が利用できる実験研究ネットワーク JGN2plus の上で、その解決方法に関する研究を行っている [発表 1]。

また、各ノード装置が広帯域なネットワーク環境で接続される環境が構築されていても、距離の離れたノード同士で TCP 方式に基づいたデータ転送を行うと、帯域・遅延積の関係から、ネットワークの帯域を十分に使ったデータ転送を行うことができない問題が生じる。これは「long-fat-pipe 問題」と呼ばれている。

これに対して、本研究室では、本来はエンドノード装置間で行われるデータ転送を、中継ノードを介して2つ以上のデータ転送に分割することで、仮想的にノード間の遅延を短縮させてデータ転送の高速化を図る提案を行っている [査読付 1, 発表 2]。

## 2.3 具体的成果

この研究では、気象センサーを一例とした多様な自律分散ノードから得られる情報の集約に関わる課題と、その解決法に関する研究成果が得られているとともに、これらのセンサー情報が社会生活の安心情報として活用できる事例を、実験研究ネットワーク JGN2plus の上で実証することができた [発表 1]。

また、中継ノードを介した仮想データ転送技術は、世界中に点在したノード同士で TCP によるデータ転送を行う際により効果的であることを PlanetLab 環境を用いて実証実験により示すことができた [査読付 1, 発表 2]。

## 3 成果要覧

### 受賞関連

[受賞 1] 中山雅哉: 通信ソサイエティ活動功労賞, 電子情報通信学会, 2009 年 9 月 16 日.

### 査読付論文リスト

[査読付 1] 榎林勇気, 中山雅哉: ボトルネック RTT の最小化に基づく TCP 分割手法の提案と PlanetLab 環境による実証評価, インターネットコンファレンス 2009, pp.25-31, 2009 年 10 月.

### その他の発表論文リスト

[発表 1] 江崎浩, 中山雅哉, 松浦知史, 井上博之, 落合秀也, 山内正人: JGN2plus SPARC における Live E! 研究活動, Live E! シンポジウム 2009 予稿集, pp.111-116, 2009 年 9 月.

[発表 2] 榎林勇気, 中山雅哉: ボトルネック RTT に基づく TCP 分割手法の提案と実証評価, 電子情報通信学会総合大会 2010, B-6-48, 2010 年 3 月.

# 実空間における状況認識およびデータ送受信に関する研究

小川 剛史

## 1 概要

実世界から仮想世界を構築し、その実世界と仮想世界を融合させることで、人々のコミュニケーションを支援する研究を進めている。構築した仮想世界を可能な限り現実世界に近づけるためには、現実世界の状況を如何に把握しその情報を収集するのか、また、構築した仮想世界もしくは融合世界の規模が拡大し、空間に関する大量の情報を大多数のユーザに提供するためにはどのようにデータを配信するのかが問題となる。そこで、主に、実空間認識技術、センサネットワーク技術、ウェアラブル拡張現実感技術に関する研究を行った。以下では、下記の 2 テーマについて具体的に述べる。

- プッシュ型放送を用いた自律移動型センサネットワーク
- ウェアラブル拡張現実感システムのための注釈情報提示制御

## 2 プッシュ型放送を用いた自律移動型センサネットワーク

### 2.1 背景

センサネットワークでは、環境モニタリング、動植物の生体調査、建物内のセキュリティ管理などを目的として、センサノード(以下、ノードと呼ぶ)だけでネットワークを構築し、マルチホップ通信により基地局へデータを転送する。従来は、ノードの位置が固定された環境を想定したものが一般的であったが、広大な領域をセンシングするためには多数のセンサが必要となることや、人が入れない汚染地域などではセンサの設置自体が困難であることから、アクチュエータを備えた移動型ノードを用いたモバイルセンサネットワークが注目されている。移動型ノードを用いれば、人が直接センサを設置することが困難な場所であってもセンシングすることができ、さらに広大な領域であっても少数のノードでのセンシングが可能となる。しかし、ノード数に対して観測領域が非常に大きくなると、各ノードの無線通信範囲に他のノードが存在しない状況が発生する。この場合、基地局にセンシングデータを送信するために、各ノードは基地局と通信可能な位置まで移動する必要があり、多大な電力を消費することとなる。本研究では、図 1 に示すような疎な移動型センサネットワークにおいて、センシングデータを基地局へ収集する際の消費電力を削減するため、放送型データ配信を用いたノード制御方式について検討している。

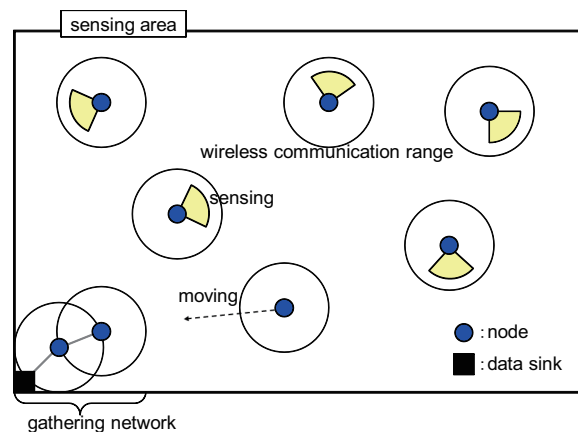


図 1 疎な移動型センサネットワーク



## 2.2 内容

### 2.2.1 ノードの移動制御方式

これまでに、プッシュ型放送から得られる基地局とマルチホップ通信を行っているノードの位置に関する情報を用いたり、移動中に通信可能となったノードと協調動作したりすることで、基地局とマルチホップ通信可能な一時的なネットワーク(収集ネットワーク)を動的に形成する SR-N(Shortest Route with Negotiation)方式、および移動コストの削減に特に注目して収集ネットワークのトポロジをあらかじめ静的に決定する MST(Moving-distance-based Static Topology)方式を提案してきた。今年度は、実環境への適用を考慮し、両手法をノード故障に対応できるように拡張した SR-N/NFD(SR-N with Node Failure Detection)方式と MST/NFD(MST with Node Failure Detection)方式を考案した。

#### (1) SR-N/NFD 方式

SR-N 方式では、データ収集時に全ノードで基地局までマルチホップ通信が可能な収集ネットワークを形成する。各ノードはメモリが一杯になった時点で移動を開始するが、その際、基地局からプッシュ型放送で得られる「収集ネットワークに接続しているノードの位置」を参考に可能な限り移動距離を短くするように目的地を決定する。移動中に他のノードと通信可能となれば、協調して移動(これらのノードをノード群と呼ぶ)することで互いの移動コストを低減する。構築した収集ネットワークは、すべてのノードが接続するまで維持することで、特に基地局から離れた場所をセンシングしているノードの移動コストを低減している。ノード故障時の問題点として、以下の3点が挙げられる。

- (i) センシング中、移動中にノードが故障すると、その故障を基地局は検知することができず、収集ネットワークを維持し続けなければならない。
- (ii) ノード群形成中のノードが故障すると、想定していた協調動作が不可能となる。
- (iii) 収集ネットワークに参加中のノードが故障すると、そのノードをマルチホップ通信の経路としていた他のノードがデータを基地局へ転送できなくなる。

SR-N/NFD 方式では、下記のようにノード故障に対応する。

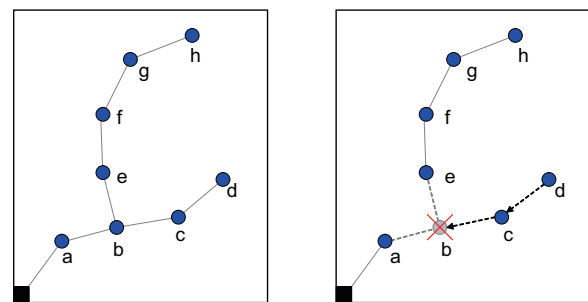
- (i) 基地局で各ノードが収集ネットワークに参加する時刻を予想し、その予想時刻を過ぎても収集ネットワークに参加しないノードは故障したと判断する。具体的には、収集ネットワークから離脱した位置からセンシング地点までの移動時間、センシング時間、センシング地点から収集ネットワークまでの移動時間を予測した。
- (ii) ノード群内のノードが故障した場合には、ノード群内の他のノードは、孤立ノードと同様に基地局からプッシュ型放送で得る情報をもとに動作する。また、ノードの故障を検知できるため、収集ネットワークに参加後、基地局へノードの故障を通知する。
- (iii) 収集ネットワークがノード故障により分断されると、故障ノードよりも葉側のノードは、ノード群となって動作し、基地局と通信可能なノードの通信可能位置まで移動する。

#### (2) MST/NFD 方式

MST 方式では、センシング終了後に各ノードがあらかじめ決定した接続位置に向かって移動し、その接続位置で収集ネットワークに参加する。

ノード故障時の問題点として、以下の2点が挙げられる。

- (i) センシング中、移動中にノードが故障すると、その故障を基地局は検知することができず、収集ネットワークを維持し続けなければならない。
- (ii) 収集ネットワークに参加中のノードが故障すると、そのノードをマルチホップ通信の経路としていた他のノードがデータを基地局へ転送できなくなる。



(a) ノード故障前

(b) ノード故障後

図2 収集ネットワークの再構成



MST/NFD 方式では、下記のようにノード故障に対応する。

- (i) 基地局で各ノードが収集ネットワークに参加する時刻を予想し、その予想時刻を過ぎても収集ネットワークに参加しないノードは故障したと判断する。MST方式では、SR-N方式と異なり、収集ネットワークのトポロジが固定であるため、より正確な参加時刻が予測できる。
- (ii) 収集ネットワークがノード故障により分断されると、故障ノードから葉ノードまでで最も短い枝に位置するノードが1つ分基地局側に移動することで収集ネットワークを修復する(図2)。

### 2.2.2 評価実験

提案手法の有効性をシミュレーション実験を行って確認した。シミュレーションで用いたパラメータを表1に示す。観測領域は2次元平面とし、基地局を領域の一角に配置した。各ノードのセンシング地点はランダムに割り当て、変更しないこととした。ノードの故障率を変化させてノード故障に対する各方式の性能を評価した。なお、ノード故障率とはシミュレーション中にノードが故障する確率を示しており、ノード故障率1とは、各ノードが必ず1回シミュレーション中に故障する状況である。ノードの故障を基地局が検知すると新たなノードを追加し、ノード数が減少しないようにした。

ノード故障と基地局におけるデータ収集量およびノードの移動距離との関係をそれぞれ図3、図4に示す。SR-N/NFD方式は、ノードが故障する場合でも、MST/NFD方式よりも効率よくセンシングデータをノードから収集できるのに対し、MST/NFD方式はノードの移動距離が短く、消費電力量を削減していることがわかる。また、SR-N/NFD方式はノードの故障率に大きな影響を受けるのに対し、MST/NFD方式はノード故障率が変化しても、性能には大きな影響がないことがわかった。

### 2.3 具体的成果

本研究で提案した方式を用いることで、疎なセンサネットワークにおいて移動コストを低減しつつ、スループットを向上させることが可能であることを確認した。これらの研究成果は国際ジャーナル[査読付6]にて発表した。本研究に関連して、固定のセンサネットワークでは、ノード数が膨大となった場合に、データ転送のための各ノードの消費電力が増加したり、無線帯域が不足したりする問題が発生する。各ノードがセンシングするデータをサーバが予測し、その予測値をプッシュ型放送を用いて配信し、予測値と異なるデータのみをノードから収集することで、通信量を削減する方式について検討を行った。この成果は、国際会議[査読付1]にて発表した。また、車車間・路車間通信を用いたインターネットアクセスに関する研究およびアドホックネットワークにおける通信経路制御に関する研究の成果を電子情報通信学会[発表1,2,3,4]にて発表した。

表1 パラメータ

パラメータ	値
領域の大きさ	2,000[m]×2,000[m]
ノード数	400
ノードの移動速度	1[m/s]
無線通信範囲	50[m]
通信速度	2[Mbps]
シミュレーション時間	200,000[s]

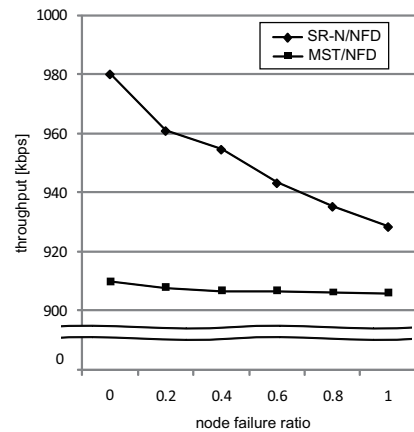


図3 基地局でのデータ収集量

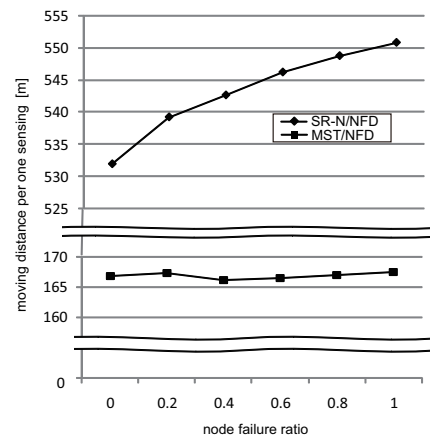


図4 ノード移動距離

### 3 ウェアラブル拡張現実感システムのための注釈情報提示制御

#### 3.1 背景

拡張現実感(AR)技術は、実際にはユーザが直接見ることのできない情報を現実空間に提示可能なため、ウェアラブルコンピューティングとの親和性が高く、これまでに、ウェアラブル AR 技術を用いてユーザのナビゲーションを行うシステムが構築されてきた。ナビゲーションシステムなどにおいて、広大な現実空間に配置する膨大な数の注釈情報を保持および管理する場合、リアルタイムに変化するユーザの位置や視点に応じて必要な注釈情報を検索し、正しい場所に提示する処理コストは大きく、一般的に非力なウェアラブル PC では処理が困難である。また、それらの注釈を複数のユーザが参照、修正するマルチユーザ環境を考えた場合、すべてのユーザが装着しているウェアラブル PC に保持したデータの一貫性を保証しなければならないため多大なコストが必要となる。これまで、ウェアラブル AR システムを構築するために必要な、サーバにおける注釈情報の管理方式とサーバからクライアントへ注釈情報を効率よく配信するための通信方式について検討してきた。

本研究では、クライアントへ配信した注釈情報について、ユーザの様々なコンテキストに応じて提示する情報量を制御する方式について新たに検討を行った。特にヘッドマウンテッドディスプレイ(HMD)を利用する際、周囲の実環境に対する注意が必要な場合には、その視認性を十分に確保することが重要となる。

#### 3.2 内容

##### 3.2.1 身体動作を利用した情報提示の切替え

従来システムでは、個々の注釈情報として画像と文字を混在して階層的に扱い、表示詳細度を切り替えていた。例えば、簡易モードではレストランの店舗名およびメニューの写真を注釈としてそのレストランの位置に表示し、詳細モードでは注目している注釈について、営業時間帯やメニューの価格帯、評判なども表示するといったことが可能である。しかし、HMD を用いた AR システムでは、常にすべての注釈および詳細情報を表示することは必ずしも好ましくない。

そこで、「座る」「歩く」「走る」といったユーザの身体動作から、時々刻々と変化するユーザの「注釈情報への注意の向けやすさ」を識別し、これに基づいて提示情報の詳細度を切り替える手法を提案する。例えば、「座っている(座位)」状態は、休憩や電車の待ち時間などであり、自動車や通行人との接触の可能性が低く、かつ数分以上持続するケースが多いと考えられる。したがって、文章などを含む詳細情報を提示したり、多数の注釈を同時に提示して実環境の視界を遮ることも許容されると思われる。

##### 3.2.2 加速度センサを用いた身体動作認識

各動作における両足の大腿部の動きに着目した。人間の大腿部は座っている時は地面に対してほぼ水平に、立っている時はほぼ鉛直に、自転車にまたかかっているかこいていない時は片足ずつ水平と鉛直に近い状態で静止している。歩いている時、走っている時、自転車に乗ってペダルをこいでいる時の大腿部の角度は周期的に変化し、その周期は走っている時と歩いている時では異なると考えられる。また、それらと比較して、ペダルをこいでいる時は足の動きはより滑らかであり、加速度の変化はより小さくなると考えられる。したがって、2 個の 3 軸加速度センサをユーザの両足大腿部前面に装着し、重力による静的な加速度とユーザの動作によって発生する動的加速度を計測することとした。大腿部の前後方向の加速度をサンプリングレート 100Hz で計測し、短時間フーリエ変換(STFT)によってウィンドウサイズ 200 サンプルで加速度の各周波数における振幅をスペクトルのパワーに変換する。ユ

表 2 ユーザコンテキストの認識精度

コンテキスト	再現率(%)	適合率(%)
座位	99.4	100
立位	97.8	99.6
歩行	90.8	99.9
走行	100	18.4
自転車	98.0	96.5

ユーザのコンテキストは、得られたパワースペクトルをサポートベクターマシン(SVM)を用いて認識させる。表 2 にユーザコンテキストの認識精度を示す。

### 3.2.3 プロトタイプシステムの実装と評価

ユーザの身体動作によって提示情報の切り替えを行うコンテキストウェアなウェアラブル AR システムを試作した。図 6 にシステム構成を示す。クライアントでは、GPS およびジャイロセンサを用いて、位置検出モジュールがユーザの位置・向きを計測する。コンテキスト認識モジュールは、加速度センサからユーザの大腿部の加速度を取得し、ユーザコンテキストを推定する。ネットワークモジュールがユーザの位置・向きの情報をサーバへ送信し、受信した情報をもとにサーバでは、データベースモジュールが適切な注釈情報を検索し、クライアントへ返信する。クライアントのキャッシュモジュールがサーバから受信した注釈情報を一時的に保存し、ユーザコンテキストに応じて必要な注釈情報をレンダリングモジュールが HMD などの表示デバイスに提示する。

提案手法の有効性を検証するため、実装したプロトタイプシステムの動作実験並びにユーザによる主観評価実験を行った。大阪大学内にサーバを設置し、ノート PC 及び各種センサを装備したユーザ(図 5)が梅田で注釈情報を閲覧した。注釈情報はグルメ情報のポータルサイトから梅田に存在するレストランの情報約 1800 件を取得し作成した。実験中、ユーザが座っている場合や立っている場合は注釈情報に加えて目の注釈情報についての詳細が、歩いている場合は注釈情報のみが表示され、走っている場合は一切の注釈情報が表示されないことを確認した。また、ユーザによる主観評価実験では、大阪大学豊中キャンパス内を 10~20 分程度散策させた。その際、基本的な移動経路は実験者が指示し、経路の脇にある建物などへは寄り道をして良いこととした。移動中は「座位」「立位」「歩行」「走行」の各動作を少なくとも一度は行うよう指示したが、どの場所でどのような行動をとるかは自由とした。AR システムに触れるのが初めての大学生 4 名と HMD 装着経験のある大学生 6 名に対して実験を行い、アンケートの結果から AR システムにおいてユーザのコンテキストによって提示情報を切り替えること、およびそのコンテキストとしてユーザの身体動作を用いることは多くの人にとって受け入れられる機能であることがわかった。

### 3.3 具体的成果

本研究で提案した方式を用いることで、ユーザの動きや状況に応じて、提示すべき注釈情報を制御することが有効であることを確認した。これらの研究成果はヒューマンインタフェース学会論文誌[査読付 5]および国際会議[査読付 3]にて発表した。また、本研究に関連して、仮想空間内でのユーザの移動や操作を支援するためのマルチビューポートインタフェースに関する研究成果を日本バーチャルリアリティ学会論文誌[査読付 2]にて、カメラ画像を用いた物体認識に関する研究成果を情報処理学会[査読付 4]にて発表した。

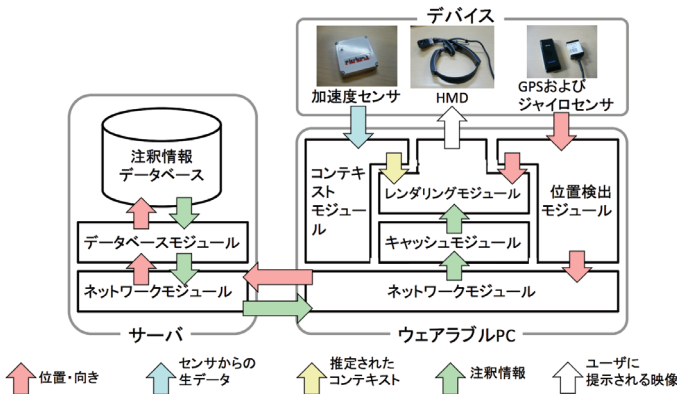


図 6 システム構成



図 5 動作実験の様子

## 4 成果要覧

### 査読付論文

- [査読付 1] Shinya Kitajima, Tomoki Yoshihisa, Takefumi Ogawa, Takahiro Hara, and Shojiro Nishio: A Broadcast-Based Data Gathering Method considering Energy Consumption for Sensor Networks, Proc. of International Workshop on Sensor Network Technologies for Information Explosion Era (SeNTIE2009), pp.536-541, May 2009.
- [査読付 2] 廣瀬 康一, 小川 剛史, 清川 清, 竹村 治雄: 複数の仮想空間を操作するマルチビューポートインタフェースのフレームワークに関する評価, 日本バーチャルリアリティ学会論文誌, Vol.14, No.2, pp.203-211, 2009年6月.
- [査読付 3] Daisuke Takada, Takefumi Ogawa, Kiyoshi Kiyokawa, and Haruo Takemura: A Context-Aware AR Navigation System Using Wearable Sensors, Proc. of The 13th International Conference on Human-Computer Interaction (HCI2009), No.3, pp.793-801, Jul. 2009.
- [査読付 4] 藤井 裕士, 小川 剛史: Integral Histogram を用いたカメラ画像からの高速物品検出手法の評価, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム(DICOMO2009) 論文集, pp.636-644, 2009年7月.
- [査読付 5] 高田 大輔, 小川 剛史, 清川 清, 竹村 治雄: 身体動作に基づき提示情報を切り替えるコンテキストウェアなウェアラブル AR システム, ヒューマンインタフェース学会論文誌, Vol.12, No.1, pp.47-56, 2010年2月.
- [査読付 6] Takefumi Ogawa, Tatsuya Shinjo, Shinya Kitajima, Takahiro Hara, and Shojiro Nishio: Node Control Methods to Reduce Power Consumption Using Push-Based Broadcast for Mobile Sensor Networks, Journal of Mobile Multimedia, Vol.6, No.2, pp.114-127, 2010.

### その他の発表論文

- [発表 1] 松岡 健太郎, 小川 剛史, 若原 恭: 車車路車間通信を用いたストリーミング配信におけるパケット到達率最大化のための基地局選択手法, 電子情報通信学会技術研究報告, Vol.109, No.273, NS2009-112, pp.49-54, 2009年11月.
- [発表 2] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間・路車間通信におけるアクセス先基地局の選択及び無線通信チャネル制御による帯域利用効率化手法, 電子情報通信学会技術研究報告, Vol.109, No.448, NS2009-251, pp.501-506, 2010年3月.
- [発表 3] 松岡 健太郎, 小川 剛史, 若原 恭: 車車間・路車間通信によるインターネットアクセスにおいて無線帯域の利用効率向上を図る経路制御手法, 電子情報通信学会総合大会講演論文集, B-15-14, 2010年3月.
- [発表 4] Zilu Liang, Takefumi Ogawa, and Yasushi Wakahara: An Automatic Route Recovery Method Based on Reserved Virtual Local Paths for Wireless Ad Hoc Network, 電子情報通信学会総合大会, ISS-P-134, 2010年3月.



# 連邦型クラウド構築のための技術研究ならびに 広域分散サービス分析のための計測手法の研究

関谷 勇司

## 1 概要

これまでに、インターネットを社会基盤のインフラとして利用するための研究を中心として行ってきた。インターネット上のすべてのアプリケーションにとっての基盤技術である DNS(Domain Name System) の安定運用ならびにその運用状態計測・分析に関する研究、ISP(インターネットサービスプロバイダー)間での柔軟なトラフィック交換を実現するための IX(インターネットエクスチェンジ)技術の研究、ならびに次世代インターネットプロトコルスタックである、IPv6, Mobile IP, IPsec といった技術の研究開発を行ってきた。本年度は、これらに加えて、クラウドコンピューティングに関する技術に重点をおき、研究を行った。

具体的には、本年度は以下の 3 項目に関する研究を重点的に行った。

1. 連邦型クラウド技術を用いた大学間クラウドの設計と構築  
普及しつつあるクラウドコンピューティングの技術を用いて、クラウド同士を相互接続して、広域に展開される連邦型クラウドの形成を目指した研究である。連邦型クラウドの要素技術研究を行うとともに、その実例としての大学間クラウドを、東京大学、慶應義塾大学、奈良先端科学技術大学院大学、北陸先端科学技術大学院大学との間で設計、構築した。
2. DNS の TTL 動的調整によるキャッシュの有効性評価に関する研究  
DNS にて長年存在する、キャッシュ不整合問題を解決して、より即時性のある負荷分散機能を提供するための、TTL 動的調整機構に関する研究を行った。これによって、規模性を損なうことなく、TTL を効率的に調整することのできる手法を提案、評価した。
3. DNS の分散計測手法とその結果を用いた分析に関する研究  
DNS の分散計測手法を提案し、その手法を用いた計測を行った。また、計測結果を用いて DNS の運用状態を分析し、2008 年度と 2009 年度の傾向と差異に関して考察を行った。その結果、BGP anycast による運用によって、DNS の到達性に変化が生じていることがわかった。また、日仏共同研究の枠組みに基づき、フランス側研究者と計測システムの相互接続に関する研究も行った。

## 2 連邦型クラウド技術を用いた大学間クラウドの設計と構築

### 2.1 背景

近年、クラウドコンピューティングと呼ばれる技術が普及し始めており、数々のサービスが登場し始めている。具体的には、Google App Engine や Amazon EC2、salesforce CRM といったものが商用サービスとして展開されており、これらは無償もしくは有償にて提供されている。特に Google App Engine や Amazon EC2 は、誰でも手軽に利用できるパブリッククラウドサービスの代表例となっている。一方、プライベートクラウドと呼ばれる、比較的小規模のクラウドコンピューティングも普及しつつある

しかし、これらのクラウドサービスは企業もしくは個人ユーザを意図して構築、提供されているサービスであり、大学における情報システム構築ならびに研究目的でのサーバ利用にて使用する場合には、使い易いサービスであるとは言い難い。大学においては、(1) 研究室毎に Web サーバやメールサーバ、DNS サーバを運用したい、等の要望や、(2)研究にて利用するための、制限の少ないサーバや実験ホストを、一時的に多数利用したい、といった要望が存在する。これらの要望を実現するために、データセンターへのサーバホスティングや、現存する商用クラウドサービスを用いることも可能である。しかし、その場合はコストの面や扱うサーバの自由度の点において、大学に適していないサービスとなっている。

そこで本研究では、大学における利用用途に適したクラウドのあり方とそれを実現するための要素技術に関する研究を行った。具体的には、大学ごとに存在するプライベートクラウドを、そのポリシーの差異や機器性能の差異、ネットワーク環境の差異を吸収し、各大学のクラウドを連結することによって、一つの広域クラウド網としてユーザに提供することを目指した。これは連邦型クラウドと呼ばれる形式であり、本年度は、この大学間連邦型クラウドを実現するための、要素技術研究を行った。

## 2.2 内容

本研究では、大学間においてリソースを融通しあうことで、計画された停電やメンテナンスといったサービス中断に左右されることなく、かつ運用コストを削減しつつ、各大学の運用ポリシーに柔軟に適應することのできる大学間クラウド網を構築することを目的とした。また、仮想環境上においてその信頼性を確保するために、各大学に分散した仮想化リソースを効率的に管理できるフレームワークも開発した。本年度は、このクラウドコントローラを設計、実装するための予備実験として、iSCSI を用いたストレージ性能評価と、大学間のネットワーク差異を吸収するための、ネットワーク移動透過性技術に関する考察、ならびに大学間クラウドコントローラの設計と実装を行った。

### 2.2.1 仮想ストレージ性能評価

まず、VM イメージの設置場所による I/O 性能の差分を調査した。以下に示す環境において、性能評価を行った。性能評価ベンチマークに用いたソフトウェア(iozone)の設定は以下の通りである。

- File Size = 64Kbytes ~ 2Gbytes
- Record Size = 4Kbytes~1024bytes
- MTU = 1500bytes

#### ①. VM ハイパーバイザーが iSCSI イニシエータとなり iSCSI ストレージに接続した場合(淡線)

接続したブロックデバイスをクラスタファイルシステム(OCFS2)でフォーマットし、マウントをしたパスに VM イメージを配置、VM 上でベンチマークを実行した。OCFS2 ではクラスタノードを自ノード以外に一つ追加し、クラスタノード上では特に他の I/O を実行しなかった。

#### ②. 個々の VM が iSCSI イニシエータとなり iSCSI ストレージに接続した場合(濃線)

接続したブロックデバイスを非クラスタファイルシステム(EXT4)でフォーマットし、マウントした先のパス上でベンチマークを実行した。

はじめに、Sequential Write の結果を図 1 に示す。空きメモリ容量と同等のファイルサイズまでは明らかに②の方が優位なパフォーマンスを見せた。ファイルサイズ 1GByte を越えてからはほぼ同じ結果となった。この差分は 1)の OCFS2 が書き込みの際にデータ保護のチェックがオーバーヘッドになったものと推測する。



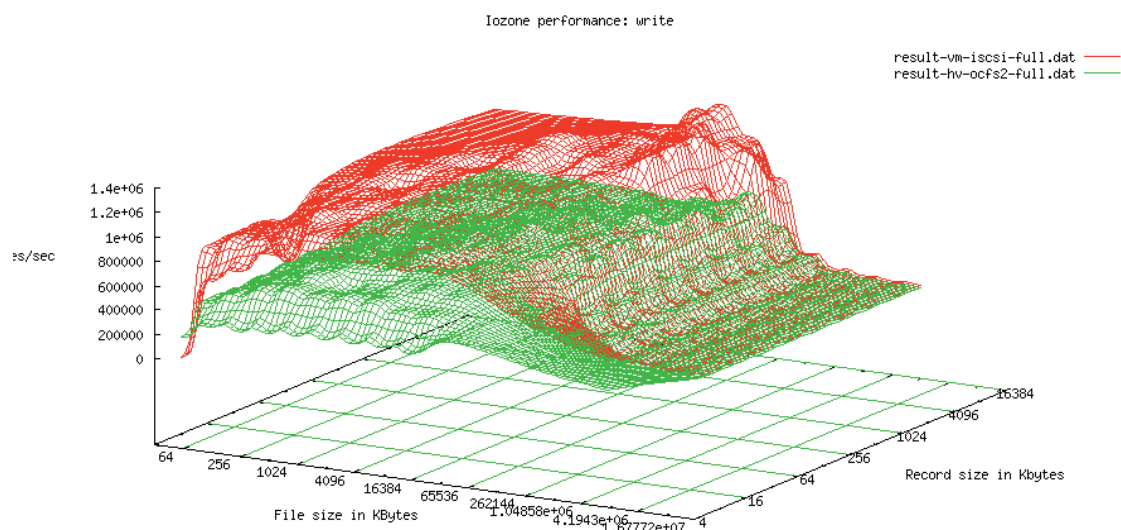


図 1 : iSCSI ストレージへの書き込み性能評価結果

次に、Sequential Read の結果を図 2 に示す。読み込みに関しては全般的に大きな差は見られなかった。読み込みに関してはデータ保護等の機能をファイルシステムが必要としないため、大きな差分が発生しなかったものと推測している。

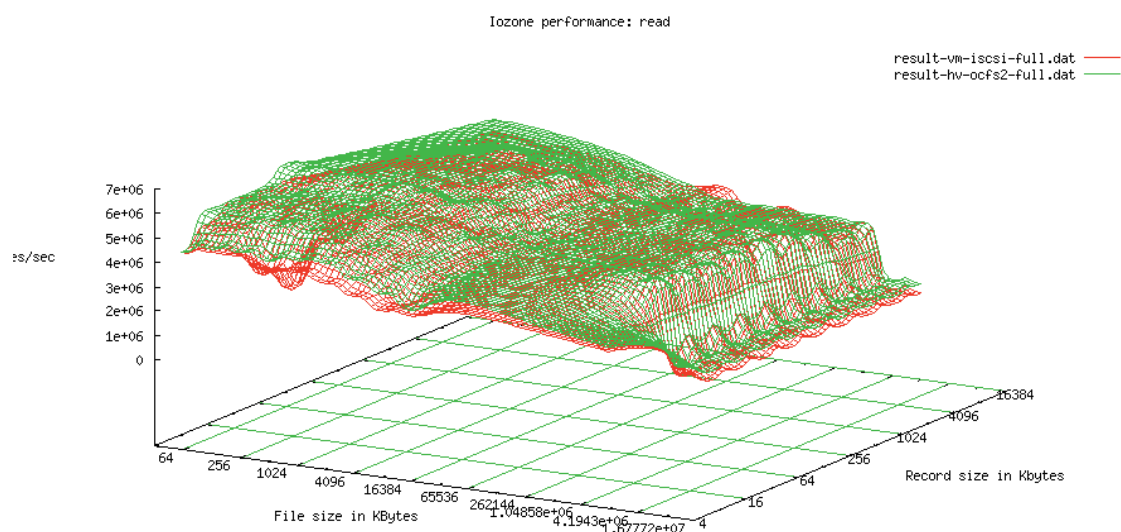


図 2 : iSCSI ストレージからの読み込み性能評価

上記の結果を比較すると、特に書き込み性能における差が顕著に現れ、VM イメージのコピー、ライブマイグレーションなど大容量データ転送が必要な場合に、データ配置先の最適化がクラウドサービスを楽しむユーザの待ち時間に直接的に影響を及ぼすことが分かった。

### 2.2.2 ネットワーク移動透過性に関する設計

現在、多くの仮想計算機技術が実用レベルで提供されており、その中には VM(仮想計算機)を、その親となっている計算機(VM ハイパーバイザー)から別の VM ハイパーバイザーに移動させる機能(マイグレーション)を提供しているものもある。

マイグレーションによって、VM がひとつのホスト計算機に集中することを回避できるが、同一サブネットワークに限られてしまい、他の位置に資源に余裕のあるホスト計算機が配置されていても活用ができない。また、計算資源を提供するクラウド環境と計算資源を活用するクライアントの間の通信遅延を抑えるためには、クラウド環境とクライアントがネットワーク的に近いのが理想である。クラウド環境の一部としてゲスト計算機を利用する場合、ゲスト計算機をより近いホスト計算機に移動させたいという要求が発生したとしても、前述した理由によりネットワークを越えるゲスト計算機の移動ができないのが現状である

これは、大学間クラウドを構築するにあたっての、大きな障害となる。大学間クラウドにおいては、各大学にて存在するクラウドを連結して連邦型クラウドを形成することを目的としており、大学間クラウド形成のために、各大学に共通したネットワークを準備することは、本来の目的にそぐわない。

したがって、VM ハイパーバイザー間における、ネットワーク環境の差異を吸収する技術が必要となる。そこで本研究では IETF による標準化技術である、NEMO(Network Mobility)を利用して、VM ハイパーバイザー間にネットワーク移動透過性を提供することを試みた。

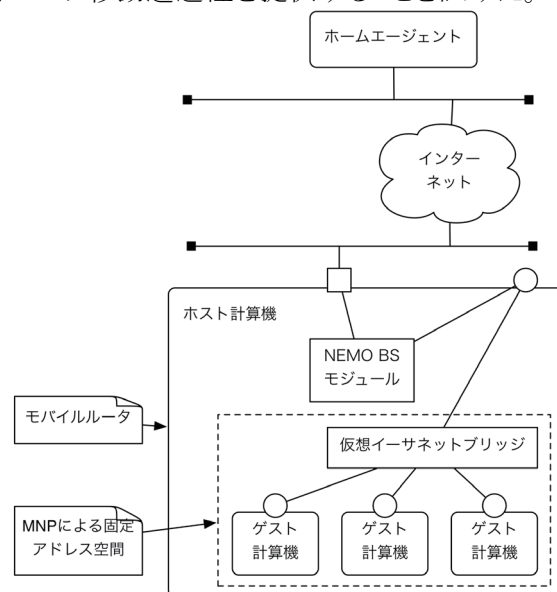


図 3: NEMO を利用した移動透過性の実現

IP モビリティ技術を使ってゲスト計算機のネットワーク環境を一定に保つ方法として、ゲスト計算機自体が MobileIP などのホスト移動通信機能を備える方式と、ホスト計算機が NEMO BS などを活用して透過的なネットワークをゲスト計算機に提供する方法がある。前者は、ゲスト計算機の改変 (IP モビリティ機能の導入) が必要になる代わりに、ゲスト計算機単体での移動が可能になり、計算機移動の細かい制御が期待できる。後者は、これまで利用してきたゲスト計算機がそのまま継続利用できる代わりに、ゲスト計算機の移動とホスト計算機の NEMO BS 移動ルータとしての移動が同期しなければならないという制限が発生する。今回は、既存のシステムで用いられているゲスト計算機をそのまま利用し続けるというシナリオを前提として、後者の方式を選択した。本研究で設計した、移動透過性を有する VM ハイパーバイザーの構成図を、図 3 に示す。本研究では、図 3 に示す通り、Linux kvm(Kernel Virtualization Mechanism)が提供する仮想計算機環境を用いた設計を提案する。この設計に基づき、現在移動透過性検証のための実装を行っている段階である。2010 年度は、この実装を完成させ、性能評価を含めたさらなる考察と分析を行う予定である。

## 2.3 具体的成果

本研究では、大学間においてリソースを融通しあうことで、計画された停電やメンテナンスといったサービス中断に左右されることなく、かつ運用コストを削減しつつ、各大学の運用ポリシーに柔軟に適應することのできる大学間クラウド網を構築した。また、その構築に必要な要素技術の研究を行った。その結果として、図 4 に示す大学間クラウドコントローラを構築し、大学間のユーザに対して研究者向けのクラウド環境を提供することができた。また、各大学の管理者に対して VM マイグレーションを柔軟に行うことのできる機能を提供することができた。

この研究成果を引継ぎ、2010 年度はより柔軟な連邦型クラウドを構築するための技術研究を行う予定である。

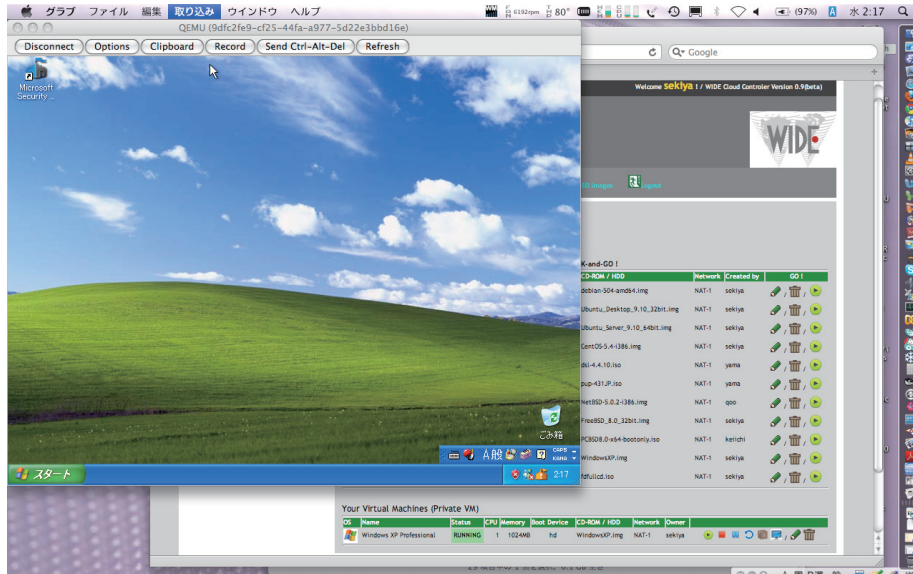


図 4: 大学間クラウドコントローラ実装

## 3 DNS の TTL 動的調整によるキャッシュの有効性評価に関する研究

### 3.1 背景

DNS には一度問い合わせた情報をキャッシュし、一定時間保持するという機能が存在する。この機能によって、DNS は規模性や耐障害性を確保している。現在の DNS 実装では、DNS のゾーン情報に明記された、固定された値の TTL(Time to Live)時間の間、キャッシュ情報を保持する仕組みとなっている。その一方で、元のゾーンに登録されているデータが更新された場合にも、一度キャッシュされている情報は、その TTL に明記されている時間を経過しない限り、古い情報を返し続けてしまうという問題も存在する。これは DNS の仕組みに由来する制限で、過去に何度も指摘されてきた問題である。近年、負荷分散に DNS を利用する事例が増加し、頻繁に DNS 登録情報が更新するために、TTL を非常に短い値に設定している例が見受けられる。確かに、TTL を短くすることによって DNS 登録情報の変更を即時に反映させることが可能となるが、逆に DNS の規模性を損なう結果となる。

この状況分析に関しては、2008 年度の研究にてデータ収集を行い、分析を行った。その結果をうけ、2009 年度では、この問題を解決するための TTL 動的調整手法を提案した。また、この手法を実際の DNS サーバに実装し、評価を行うことによって、ゾーンに記載されているデータと、キャッシュされたデータの不整合を低減できることを証明した。

### 3.2 内容

本研究にて提案した動的 TTL 調整手法は、DNS におけるキャッシュとゾーンデータ間でのデータ不整合問題を解決することができた。本手法のアルゴリズムを図 5 に示す。本手法では、ゾーンデータを保持する DNS サーバが、過去の更新頻度等を考慮、もしくは管理者が明示的にデータの更新間隔を指定することで、更新時刻に向けて動的にデータの TTL を減少させることで、データとキャッシュの間の不整合を解消することをねらう。

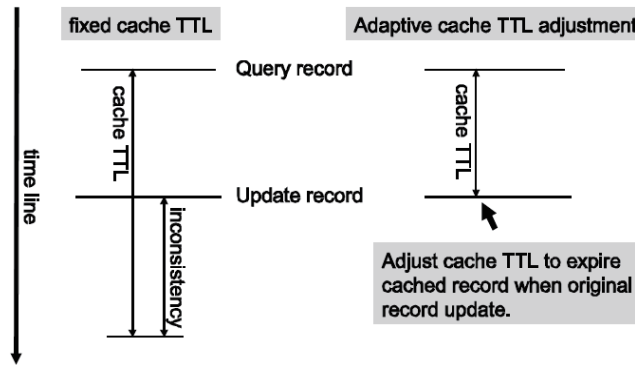


図 5：動的 TTL 調整手法

現在の DNS 実装では、データを保持する DNS サーバが、他の DNS サーバからデータの問い合わせを受けた場合、あらかじめ設定された一律の TTL をデータとともに返答する。しかし、本手法では、データを保持する DNS サーバ自身も、個々のデータ毎に TTL を内部で管理し、時間とともに TTL を減少させる。この手法を用いることで、問い合わせを受けた時間に応じた TTL を返答することができるため、DNS サーバ間でキャッシュされる時間の不整合が発生しない。

本研究では、提案手法を実際の DNS サーバに実装し、有用性を評価した。本手法では、前述の通り DNS サーバ間のデータ TTL の値を同期させることが可能となるが、その一方で、従来のキャッシュ機構に比べると、データ保持サーバへの問い合わせ数が増加する。このトレードオフを評価するために、ある Web サーバの A レコードに対する TTL を、本提案手法にて動的に調整した場合と、一律の TTL を返答した場合に関する比較を行った。具体的には、TTL の初期値を 3600 秒とし、1 時間ごとにこの Web サーバの A レコードを変更し、実際のクライアントが通信を試みる IP アドレスを、更新前のアドレスを持つサーバと更新後のアドレスを持つサーバ両方にてトラフィック監視を行うことで、計測を行った。

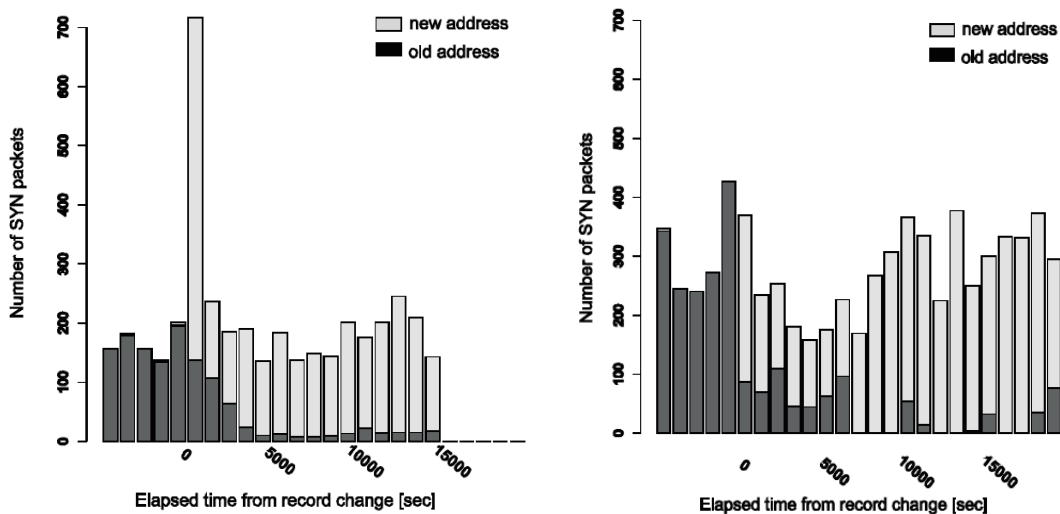


図 6：新旧アドレスへの通信数 v.s. 提案手法を適用した場合の通信数



従来手法にて TTL を返答した場合の、クライアントから Web サーバ新旧それぞれのアドレスへの通信数計測結果、ならびに本手法にて TTL を返答した場合の新旧アドレスへの通信数計測結果を図 6 に示す。X 軸の 0 の位置にてゾーンの A レコードデータを更新した。

両者の結果を比較すると、提案手法にて TTL を変更した場合には、A レコード更新後に旧アドレスへの通信が徐々にではなく急激に減少していることが見て取れる。一方、旧来の手法では、新しいアドレスへの変更後に新アドレスへの通信が増大していることが見て取れる。この更新後の新アドレスへの急激なアクセス増は、提案手法の場合には発生しておらず、なぜ旧来の手法の場合のみに発生しているのか、さらなる解析を行う必要がある。

また、DNS 問い合わせクエリ数の推移に関して、従来手法と本提案手法における比較を図 7 に示す。本提案手法において、一時的な問い合わせ数の増加を見てとることができる。このトレードオフに関しては、DNS の規模性にも影響する部分であるため、さらなる分析と考察が必要となる。

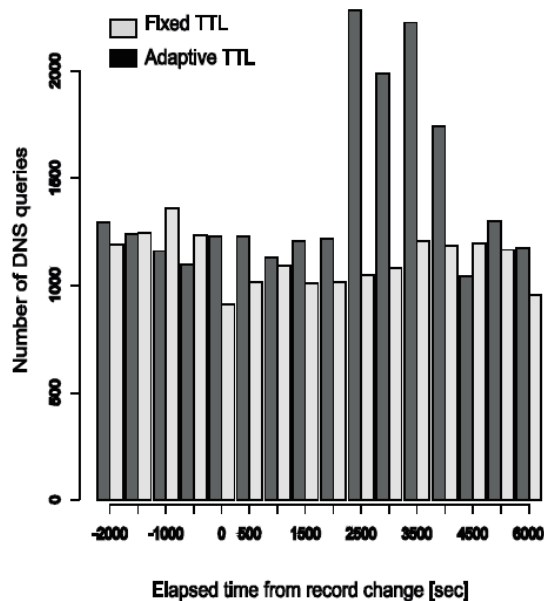


図 7: 両手法における DNS 問い合わせ数の推移

### 3.3 具体的成果

本研究では、TTL の動的調節手法を提案、検証することで、DNS サーバ間にて発生するキャッシュの不整合を解決できることを示した。これによって、DNS による負荷分散や攻撃に対応するための頻繁な DNS データの更新を、DNS のキャッシュ情報不整合に悩まされることなく行うことが可能となった。また、実際の Web サーバに対するアクセスに対し、本手法を導入した実験を行うことで、実環境での有用性を示すことができた。また、本手法を広域に適用するためにはさらなる分析が必要となるため、2010 年度も引き続き研究を継続していく。なお、本研究の成果は[査読付 4]にて発表した。

## 4 DNS の分散計測手法とその結果を用いた分析に関する研究

### 4.1 背景

DNS はインターネットのサービスを支える基盤技術である。DNS は世界最大規模の分散データベースであり、インターネット上に存在する数万台の DNS サーバが連携することにより、一つのデータベースを形成している。DNS の信頼性を向上させるためには、DNS の運用状態を把握する必要がある。しかし、単一の DNS サーバに対して、また単一の地点からだけ計測を行う手法では DNS 全体の運用状況を把握することはできない。DNS の系全体を計測するためには、多数地点から多数の DNS サーバに対して計測を行う必要がある。

そこで本研究では、(1) DNS がインフラを支える基盤技術としての信頼性を確保できているかを計測・分析する手法を確立する、(2) 提案手法を用いて DNS のインフラとしての信頼性を計測する、という2点を目標とした研究を行った。

本研究は 2006 年度から継続して行っており、(1)に関しては、世界規模の分散計測システムを構築することによって、計測手法を確立した。その計測手法を用いて、2008 年度から本格的な、主要 DNS サーバに対する分散計測を開始した。その初年度の結果を 2008 年度の年報にて報告した。その結果をうけ、本年度ではさらなる計測を行い、2008 年度の結果との比較分析を行った。

また、日仏共同研究の枠組みによって、フランス側研究者と分散計測に関する技術交流を行い、フランス側にて研究されている計測システムである、TopHat との相互乗り入れに関する議論も行った。この議論は 2010 年度も継続して行い、共同研究を続けていく予定である。

## 4.2 内容

2009 年度は、32 台の計測機器が世界各地に設置され、これらの計測機器を用いて、主要 DNS サーバの分散計測を行った。2010 年 3 月時点における、計測地点の一覧を表 1 に示す。強調文字の計測地点は、本年度に新たに設置された地点である。本研究での計測は、インターネット先進国のみならず、インターネット発展途上国に重点をおいた計測をめざしており、それを考慮した計測機器の配置が行われている。本研究の最新状況ならびに計測結果・成果は、<http://gulliver.nc.u-tokyo.ac.jp/> から取得することができる。

表 1: 計測機器設置拠点一覧

拠点番号	設置場所
1	Los Angeles, U.S.A.
2	Tokyo (WIDE), Japan
3	Sao Paulo, Brazil
4	Tokyo (WIDE), Japan
5	Osaka (WIDE), Japan
6	Osaka (WIDE), Japan
7	Paris, France
8	Brisbane, Australia
9	Bangkok, Thailand
10	Klong Luang, Thailand
11	Tokyo, Japan
12	Tokyo (The University of Tokyo), Japan
13	Tokyo, Japan
14	Seoul, Korea
15	Kuala Lumpur, Malaysia
16	San Diego, U.S.A.
17	Barcelona, Spain
18	Colombo, Sri Lanka
19	Napoli, Italy
20	Auckland, New Zealand
21	Bangkok, Thailand
22	Chicago, U.S.A.
23	Klong Luang, Thailand
24	<b>Taipei, Taiwan</b>
25	<b>Amsterdam, Netherlands</b>
26	Nairobi, Kenya
27	Nairobi, Kenya
28	<b>Singapore Singapore</b>
29	<b>Hong Kong, China</b>
30	<b>Pretoria, South Africa</b>
31	<b>Kathmandu, Nepal</b>
32	<b>Paris, France</b>



2009 年度における 12 組織の Root DNS サーバに対する到達性を、20 の計測拠点から計測した到達性に関する分布図を示す。縦軸は Root DNS サーバの組織名を示し、横軸は計測地点の国名を示す。2009 年度を通じた計測結果を用い、計測は約 5 分に一度の頻度で行われた。図 8 に、応答が得られるのにかかった時間(RTT)分布における中間値(ms)を示す。また、図 9 に Anycast Node ID の変化回数の分布を示す。図 8 が示す通り、東南アジア地域ならびにアフリカ地域の計測点からの到達性は、北米ならびに欧州といったインターネット先進国地域に比べ、応答時間が長いことがわかる。一方、2008 年度よりも 2009 年度の方が、これらの到達性が悪い地域であっても、応答時間が短くなっていることがわかる。さらに、図 9 が示す通り、Anycast Node ID の変化割合は、計測地域にはよらず、Root DNS サーバの種類に影響されていることがわかる。2008 年度よりも 2009 年度の方が、より Anycast Node ID の変化回数が増加しており、特に 2009 年度においては、到達性の悪い地域からの Anycast Node ID の変化回数が増加していることがわかる。これは、Root DNS サーバの BGP Anycast 運用がさらに普及し、インターネット発展途上地域においても、複数の Root DNS サーバが到達候補になっていることがわかる。

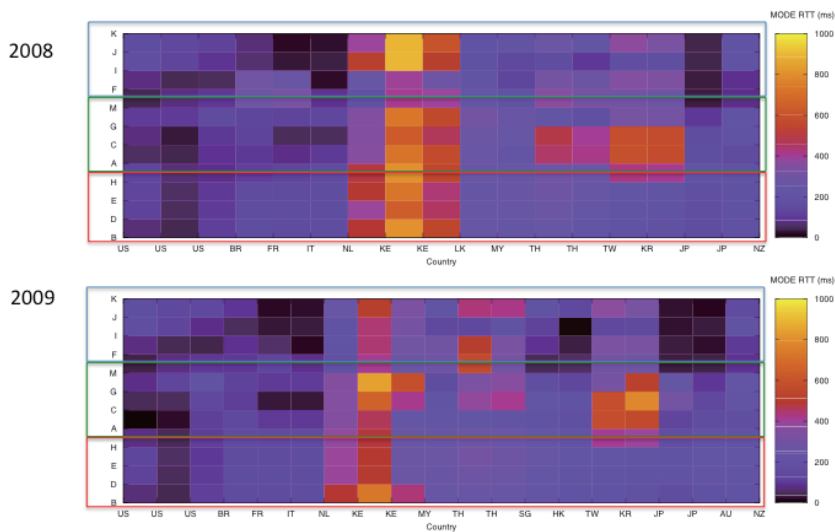


図 8：2008 年度ならびに 2009 年度における RTT 中間値の分布

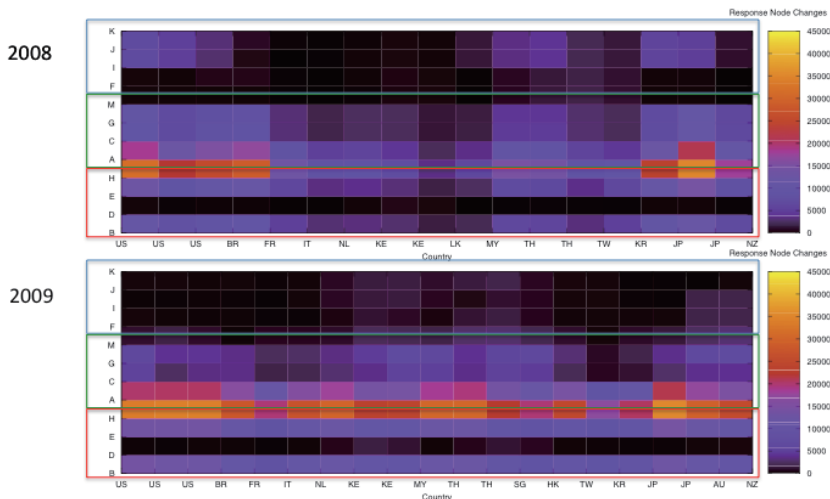


図 9：2008 年度ならびに 2009 年度における Anycast Node ID 変化数分布

### 4.3 具体的成果

本研究では、Root DNS サーバや ccTLD DNS サーバといった重要な DNS サーバに対する、世界各地、とくにインターネット発展途上地域からの到達性を示すことができた。これは類似研究においてもまだ行われていないデータ収集であり、この計測を 2 年間行うことで、それぞれの地域からの到達性の変化を分析することが可能となった。本研究における詳しい分析結果は、[発表 1]ならびに [発表 2]にて発表した。

## 5 成果要覧

### 受賞関連

[受賞 1] 石原知洋, 関谷勇司, 村井純 : 平成 21 年度情報処理学会論文賞, 情報処理学会, 2010 年 3 月

### 著書／編集

[著書 1] 関谷勇司 他: 「ロボット情報学ハンドブック」, ナノオプトニクスエナジー, 2010 年 3 月

### 査読付論文

[査読付 1] Pakornsiri Akkhara, Yuji Sekiya and Yasushi Wakahara: Efficient Alarm Messaging by Multi-channel Cut-through Rebroadcasting based on Inter-Vehicle Communication, IAENG International Journal of Computer Science, 36:2, IJCS\_36\_2\_07, May 2009.

[査読付 2] Wei Tian, Yuji Sekiya, Yasushi Wakahara and Yoshiaki Tanaka: Centralized P2P Management System based on File and Peer Reputation, International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC2009), pp.1062-1065, Jul. 2009.

[査読付 3] Pakornsiri Akkhara, Yuji Sekiya and Yasushi Wakahara: Fast Dissemination of Alarm Message based on Multi-channel Cut-through Rebroadcasting for Safe Driving, in Ao Sio-Iong, Castillo Oscar, Huang Xu (Eds.): Intelligent Automation and Computer Engineering, Lecture Notes in Electrical Engineering, Springer for IMECS 2009 Paper Number:ICCSA\_70, Vol.52, 2009.

[査読付 4] Tomohiro Ishihara and Yuji Sekiya: DNS cache-record coordination Using Adaptive Cache TTL Adjustment, International Journal of Computer and Network Security, Volume. 2, No. 1, pp. 66--69, Jan. 2010.

### その他の発表論文

[発表 1] Yuji Sekiya : Gulliver Project - status update in 2009, ISMA 2010 AIMS-2 Workshop on Active Internet Measurements, San Diego, U.S.A., Feb. 2010.

[発表 2] Yuji Sekiya : Behavior Analysis of DNS Anycast in 2008, The 2nd CAIDA/WIDE/CASFI Workshop, Seoul, Korea, Apr. 2009.

[発表 3] 田 偉, 関谷 勇司, 若原 恭, 田中 良明: 悪意ファイル排除のための P2P 管理システム, 電子情報通信学会・技術研究報告, IN2009-33, pp.49-54, 2009 年 7 月.

[発表 4] 田 偉, 関谷 勇司, 若原 恭, 田中 良明: 悪意ピアと悪意ファイルの排除機能をもつ P2P 管理システムの評価, 電子情報通信学会・通信ソサイエティ大会, B-6-33, 2009 年 9 月.

# P2P ネットワーク計測

## ネットワークシミュレーションの信頼性

### 音声処理言語

助教 中村 文隆

## 1 概要

## 2 P2P ネットワーク計測

### 2.1 背景

P2P 技術によるファイル共有、動画配信の普及にともない、P2P ネットワークのトラフィックが、アンダーレイネットワークに与える影響が問題となっている。

P2P ネットワークは、既存の IP ネットワーク上に構成されるオーバーレイネットワークであり、アンダーレイネットワークとしての IP ネットワークのトポロジを考慮しないトラフィックを生成する。その結果、IP ネットワーク上には”無駄な”トラフィックが発生し、ISP の通信帯域への圧迫が生じている。

本研究では、ISP において適切にキャッシュを行うことにより P2P ネットワークから発生するトラフィックの削減を行うことを目的として、実稼働している P2P ネットワーク上のコンテンツ人気度を計測するシステムを開発した。研究は NTT サービスインテグレーション基盤研究所と東京大学との共同研究として行い、P2P ネットワーク計測を東京大学が、計測したデータに基づくキャッシュ制御技術の検討を NTT サービスインテグレーション基盤研究所がそれぞれ担当した。

### 2.2 内容

P2P ネットワークは、任意のアンダーレイネットワーク上に構築されるオーバーレイネットワークであり、ファイル共有 P2P ネットワークとしては IP ネットワーク上に構築されている Winny や Share がよく知られている。これらの P2P ネットワークでは、ユーザーがダウンロードするファイルが一般に複数のノードに存在しているが、ファイルダウンロードにあたって P2P ネットワーク側では IP ネットワークへの影響を考慮することなくダウンロード先のノードが選択されている。

IP ネットワークへの影響の指標としては経路のホップ数やネットワーク全体のキャパシティなど複数の指標が考えられるが、ISP において考慮すべき重要な指標として、自 ISP と他 ISP との間に発生するトラフィック (ISP 間トラフィック) を本研究では対象とした。ある ISP から他 ISP へのトランジット回線維持にかかるコストが ISP において問題になっており、ISP 間トラフィックを削減することが ISP の提供する通信品質に寄与することがその理由である。具体的な ISP 間トラフィックの削減方法としては、P2P ネットワーク上のコンテンツを ISP 内でキャッシュし、ISP 間トラフィックを発生するダウンロード要求を ISP 内キャッシュからのダウンロードにおきかえる方式をとることとした。

上記のように P2P ネットワーク上のコンテンツをキャッシュする場合、キャッシュを効率的に運用するためにはダウンロードにより発生するトラフィックの全体に対する割合の高いものを優先してキャッシュ対象とする制御方式が必要となる。素朴に言えば、ファイルサイズ×ダウンロード頻度 (人気度) の大きなファイルを優先してキャッシュ対象とすることが考えられるが、Winny や Share のような P2P

ネットワークはネットワーク全体、あるいは主要な部分ネットワークにおける人気度についての情報を管理するノードが存在しないため、人気度の情報を知ることがキャッシュ方式の最適化において重要な要素となる。

## 2.3 具体的成果

前節で述べたように、P2P ネットワークにおけるコンテンツ人気度を計測するため、我が国における実稼働 P2P ネットワークとして最大のユーザー数をもつ Winny を対象として、Winny 上で流通するコンテンツの人気度を計測するシステムを開発した。

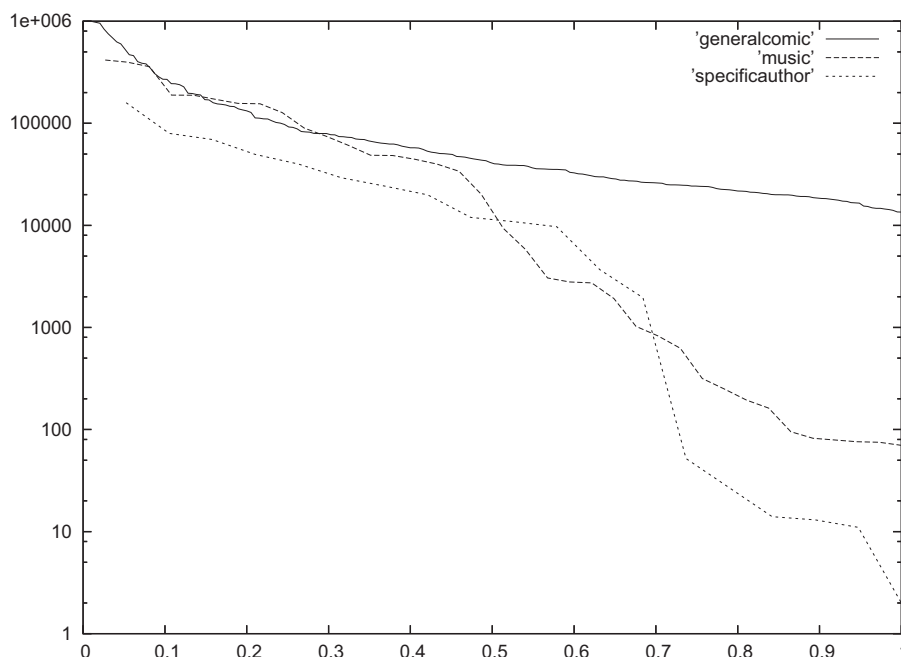


図 1: Winny ネットワークにおけるコンテンツ人気度

図 1 は、2010 年 2 月に行った測定の結果得られた人気度分布であり、横軸は該当する検索キーワードにマッチしたファイルの順位 (ファイル総数に対する割合として正規化)、縦軸はダウンロード回数を示す。実線はファイル総数が数千、点線・破線は数百のカテゴリで、マッチしたファイル数の違いにより、人気度分布にカットオフが存在する場合としない場合があることを示している。

人気度の分布にカットオフが存在する場合、キャッシュの制御としては分布がダンプする順位のファイルまでを優先してキャッシングし、効率的なキャッシュ装置の運用を図ることができる。カットオフが存在しない場合の制御アルゴリズムについては今後の課題として残っている。

本システムにより計測したデータに基づいて検討したキャッシュ制御方式については、電子情報通信学会 2010 年総合大会において発表を行った [発表 1]。

## 3 ネットワークシミュレーションの信頼性

### 3.1 背景

通信ネットワークの研究における性能評価においてネットワークシミュレータが利用される機会が増えてきているが、適切な利用を行わなければ結果の信頼性を確保することは難しい。

## 3.2 内容

通信ネットワークの研究において行われるシミュレーションの性質はさまざまであるが、この分野においては TCP/IP や ad hoc ネットワークなど、乱数を用いるモンテカルロシミュレーションとしての性格をもつものが少なくない。また、シミュレーション一般の話として、対象系の初期状態から過渡状態、定常状態の概念も重要であり、評価量の算出には対象系の定常状態での評価を行うべきである。

## 3.3 具体的成果

近年、通信ネットワーク分野において上記のような知見を欠いた論文、発表が散見されるようになってきているため、主に大学院生を対象として、ネットワークシミュレータが生成するシミュレーション結果を適切に扱い、信頼性のあるネットワークシミュレーションを行うためのチュートリアルを、電子情報通信学会 NS 研究会主催のサマースクールにおいて行った [特記 1]。

# 4 音声処理言語

## 4.1 背景

パーソナルコンピュータの性能向上に伴い、音声データの実時間処理機能を行うアプリケーションが数多く利用されている。しかし、従来、そのようなアプリケーションを開発する環境はある程度の専門的な知識とコーディングの技術を前提としており、エンドユーザが気軽に扱えるものではなかった。

## 4.2 内容

グラフィカルなプログラミング言語である PureData は、音声データおよびシリアル通信を用いたプログラミングに特化した言語であり、音声データの実時間処理を伴うアプリケーションの開発に適している。PureData を用いることで、従来の環境に比較して短時間でプログラミングを習得することができ、ユーザの自由な発想に基づいたさまざまなアプリケーションの開発を促すことができる。

## 4.3 具体的成果

音声アプリケーションの開発に興味のあるユーザを対象とし、初学者にも容易に PureData のプログラミングを習得できることを目的とした書籍の執筆を行った [著書 1]。

# 5 成果要覧

## 著書／編集

[著書 1] 中村文隆: グラフィカル言語 PureData による音声処理, CQ 出版社, 2009 年 9 月.

## その他の発表論文

[発表 1] 小林正裕, 亀井聡, 中村文隆: 実データに基づく P2P キャッシュ方式の検討, 電子情報通信学会 2010 年総合大会, 東北大学, 仙台, B-6-73, 2010 年 3 月.

## 特記事項

[特記 1] 中村文隆: ネットワークシミュレーション概論, 電子情報通信学会 NS 研究会シミュレーションサマースクール, 2009 年 8 月.

# ネットワークの運用とセキュリティに関する研究

加藤 朗

## 1 概要

既に社会基盤としての位置づけを確立しているインターネットにとって、その安定な運用は重要なテーマの一つである。さらに、単に安定に稼働しているだけでなくセキュリティ的な側面も重視されてきている。ここでは DNS の運用に関する問題について、従来から考察を行い、実際の DNS サーバの運用に利用しているほか、一般には気づかれにくい方法を用いた情報漏洩について、その可能性、検出法、および事前の検査法に関する検討を始めているので、それについても簡単に報告する。

## 2 DNSSEC の運用に向けて

### 2.1 背景

DNS はインターネット上の資源の名前付けおよびその解決を行うシステムである。資源の発見については、現在その殆んどがいわゆる検索エンジンに委ねられているが、得られた URL に実際にアクセスするためには引き続き DNS による IP アドレスへの変換が必要であり、DNS が停止した場合、IP パケット交換網としてのインターネットは機能していても、事実インターネットは使い物にならなくなるため、DNS は実際にはアーキテクチャ的には一アプリケーションであるにも関わらず基盤サービスとして位置づけられている。

DNS はキャッシュを多用することを前提にしているが、DNS メッセージを偽造することによって偽の情報をキャッシュに注入する攻撃に対する懸念が深刻化しており、その対策として署名連鎖によって DNS の応答の検証を行う DNSSEC の導入が急務になってきている。DNS の一部の部分木には DNSSEC の検証に必要な情報が登録されるようになってきているが、グローバルな導入はまだ実施されていない。

DNS に新たな機能を追加する際、最も注意しなければならないのは、そのことによって DNS が機能不全に陥ることがないようにすることである。特に、DNS の木構造の名前空間の Root ノードに対応する Root DNS サーバへのアクセスが出来なくなった場合には、DNS サーバのキャッシュにない TLD へのアクセスが出来なくなる他、最大 2 日程度でキャッシュ中の情報の寿命が尽きるため、インターネット全体が機能不全に陥ることになる。このような大規模障害を発生させないようにしながら、新機能の導入を図ることが求められている。

### 2.2 内容

現在、DNSSEC をサポートしている主要な DNS ソフトウェアは、各ゾーンの公開鍵が設定され、署名の検証が可能になっているかどうかに関わらず、DNSSEC で定義されている署名情報や鍵情報を要求する “DO” bit<sup>1</sup>を立てた問い合わせを行う。これは、仮にそのサーバが DNSSEC に対応していない

<sup>1</sup>DNSSEC OK の意味



としても、DNSSEC に対応した他のサーバやリゾルバから要求された場合には、DNSSEC 関連の情報提供も行わなければならないとの考え方に基づいている。このため、実際に DNSSEC による検証が可能になっているゾーンは .SE など少数であるにも関わらず<sup>2</sup>、Root DNS サーバに寄せられる問い合わせの 6 割程度のパケットには DO bit が立っている。従って、Root DNS サーバを含む DNS サーバは、それがおそらくは無駄になるということを予期しつつも、DNSSEC に対応した応答を返さなければならない。これによって、DNS サーバ側の CPU 負荷は増大し、また応答メッセージも大きくなるなどのオーバーヘッドの増大や、大きな応答メッセージによる副作用も懸念されている。

### 2.3 具体的成果

Root Zone の DNSSEC 対応は数年前から議論が行われてきているが、これが安全に Root DNS サーバに導入できるかどうか、ということが問題になっていた。そのため、検証はできないものの、DNSSEC を導入したのと同じ大きさの応答を返すように Root DNS サーバを順次切り替えていき、その都度トラフィック情報を収集・解析することによって、問題があるかどうかを注意深く観察することになった。このプロジェクトは DURZ (derivatively unvaridatable root zone) と呼ばれている。

まず、2010 年 1 月下旬に ICANN が運用する L サーバが DURZ 対応を行い、2 月中旬に VeriSign が運用する A サーバが、3 月初旬に筆者らが運用している M サーバが I サーバと同じ日に、残りのサーバの多くは 3 月下旬および 4 月上旬に移行しており、J サーバを残すのみである。Root DNS サーバへのアクセスに問題がある場合、順次アクセスを試み、最終的に J サーバに辿り着くことになるので、そのアクセス状況を解析することによって、問題の有無をある程度推定することができる。

この過程で、署名付き応答を要求した場合、サーバのソフトウェアバージョンによっては大きな応答を返すため、IP のフラグメント化が発生することがあった。フラグメント化は通常のインターネットでは受信ノードで再構成されるため、問題は発生しない筈であるが、一部の NAT はフラグメント化された IP パケットの処理に問題があり、当該 Root DNS サーバへの双方向のアクセスができなくなる問題が報告されている。最新のバージョンでは必要最小限の情報しか返さないため、現在はこの問題は表面化していない。しかし、実際に署名の検証ができるようになった場合、類似の問題が多発することが懸念され、引き続きトラフィックデータの解析による予測が必要となっている。

## 3 隠蔽通信路による情報漏洩

情報漏洩は、特に p2p による情報配布網上に、感染した計算機上にあるファイルを配布してしまうコンピュータウイルスの出現以来、大きな社会問題になっている。このような情報漏洩は、現実には、コストの問題や、移動式携帯端末と無線を利用したネットワークアクセスの普及によって、監視をきちんと実施することは必ずしも容易ではないが、ネットワークへの出入り口での監視によって、検出あるいは防止は原理的には可能である。

しかし、直接パケットを送出するのではなく、他のパケットやメッセージの一部に情報を埋め込むことによって実現される情報漏洩は、漏洩できる情報密度を高くすることはできないが、その存在を発見することは非常に難しくなる。通常発生しない宛先に対するパケットを送出することによって情報漏洩を凶った場合、仮に漏洩を凶る情報を暗号化したとしても、余計なパケットの発生は、情報漏洩の可能性について神経質に監視をしている場合には発見されやすい。しかしながら、通常は気づかない方法による情報漏洩が可能だったとすると、その検知は困難である。このような通信路を特に隠蔽通信路と呼ぶが、隠蔽通信路に関する一般的な対策はできないため、その手法を提案し、それぞれの手法に関する検知や対策を考えなければならない。

<sup>2</sup>2009 年夏時点

### 3.1 内容

情報漏洩を行うパケットを別に送出するのではなく、通常発生する通信に発見されにくい方法で漏洩すべき情報を少しづつ埋め込む方式は、パケット数の監視などの trivial な方法では発見することはできない。一度に埋め込むことができる情報量は限られているものの、通常頻繁に通信を実施するパケットを利用することができれば、それなりの量の情報漏洩は可能である。

“実用的”とは言いがたいものの、一例として IPv4 パケットの ID フィールドを考える。ID フィールドはフラグメント化されたパケットの再構成に用いられるが、フラグメント化されていないパケットに限定すれば、経路の途中で ID フィールドが変更されたとしても、通信には全く影響を与えないので、変更されたことを検出することはできない。従って、32bit の ID フィールドのうち、8bit 程度には漏洩情報を格納することが可能となる。一パケットについて 1byte 程度の情報しか埋め込むことはできないが、非常にアクセスが多い宛先、もしくはその宛先に至る経路上の機材が協調することによって、隠蔽通信路を構成することが可能になることが期待できる。

### 3.2 具体的成果

本研究に着手して日が浅いため、まだ具体的な成果を出すには至っていない。また、本手法による隠蔽情報路の作成は、各種定義されたプロトコルの予定外の使い方を駆使する必要がある。そのため、十分にデバッグされていない実装やプロトコルの解釈の微妙な食い違いが発覚する可能性も少なくないことが、副作用として考えられる。これらの問題は、インターネットの運用上問題となる場合もあり、研究成果として得られたとしても、セキュリティ上の理由により然るべき時期までは発表できない可能性もある。

隠蔽通信路を考える場合、既存のプロトコルで十分に定義されていないフィールドや機能などをうまく活用することで、その実現性を考えることができる。また同時に、仮定した隠蔽通信路が存在しているかどうかを検証することも必要な場面もある。検証方法はまだ確立されていないが、基本的にはネットワーク機器入力と出力を監視しておいて、様々なパターンでの入力を与え、対応した出力との差分を評価する方法が考えられている。また未定義の機能を使うことによって、隠蔽通信路としての利用はともかく、通常の方法では発見できない実装のバグに抵触する可能性もあり、引き続き、実機を用いた試験および検討が必要である。



# スーパーコンピューティング研究部門

## スーパーコンピューティング研究部門概要

### 中島 研吾

大規模並列シミュレーション手法に関する研究、教育

### 金田 康正

スーパーコンピューターに基づく大規模数値計算に関する研究

### 佐藤 周行

言語処理系とサービス体系におけるセキュリティ保証の研究  
Optimization Verifying Compiler,  
Document Carrying Authorization,  
and LoA of Services

### 田浦 健次郎

データ集約的な分散計算のための枠組み

### 黒田 久泰

シームレス高生産・高性能プログラミング環境  
—高効率・高可搬性ライブラリの研究開発—

### 松葉 浩也

クラスタからクラスタ外部への  
ネットワーク出力に関する研究

### 大島 聡史

GPUを用いた高速数値計算と  
GPUプログラミング環境に関する研究

### 堀 敦史

高効率かつ透過なファイルステージングシステム

### 片桐 孝洋

ソフトウェア自動チューニングおよび  
高性能数値計算ライブラリの研究とHPC教育

### 吉廣 保

先端的大規模計算利用サービス

### 渡辺 宙志

マルチスケールな現象の直接計算  
—大規模分子動力学法による気泡生成現象の研究—

### 鴨志田 良和

大規模分散計算環境を有効活用するソフトウェアの研究

### 藤田 肇

ブロードキャスト型単一IPアドレスクラスタにおける  
柔軟なリクエスト分散機構



# スーパーコンピューティング研究部門

## 概要

部門長 中島 研吾

計算科学が、理論、実験に続く「第三の科学」と呼ばれるようになって久しい。スーパーコンピューティングは計算科学を支える重要な基盤であり、ペタスケール、エクサスケールコンピューティングの時代を迎えて、スーパーコンピューティング研究部門の果たすべき役割は大きい。

スーパーコンピューティング研究部門は、2010年3月現在で12名のセンター専任教員(教授:2、准教授:1、助教:2、特任教授:1、同准教授:3、同講師:1、同助教:2)、3名の兼任教員(教授:1、准教授:2)を擁しており、大学院兼担教員による大学院生としては、新領域創成科学研究科基盤情報学専攻および工学系研究科電気系工学専攻から博士課程3名、修士課程3名の合計6名が在籍している。

専門分野は、計算機システムからコンパイラ、数値アルゴリズム、各種科学技術アプリケーション、セキュリティまで、また理論的研究から実用的研究まで多岐にわたっているほか、業務部門と協力してスーパーコンピューターシステム(SR11000/J2、HA8000 クラスタシステム)の運用にあたり、利用環境の向上、利用者拡大のための広報・普及活動を実施している。成果は各分野の学会において研究論文等として発表されており、高い評価を得ている。また各学会の役員、各種会合運営、セッションオーガナイザ等としても活発に活動している。

以下に研究、教育活動の概要を示す。

## 1 計算機システム

2008年度から、京都大学 学術情報メディアセンター、筑波大学 計算科学研究センターと共同で研究開発を進めている「シームレス高生産・高性能プログラミング環境」では、研究室レベルで使用されているPCクラスタからHA8000クラスタシステムに相当する大規模クラスタ、次世代スーパーコンピュータまでをシームレスに利用できるためのシステムソフトウェアの開発を行っており、学術的な成果のみならず、HA8000クラスタシステムの運用にあたっても有用な成果が得られている。

主な研究開発項目は以下の通りである:

- 高効率ファイルステージングシステムに関する研究
- 分散ワークフロー処理系に関する研究
- 分散ファイルシステムに関する研究
- データ集約的アプリケーションのアクセス解析に関する研究
- 分散環境における高性能通信に関する研究
- 資源の動的な伸縮が可能な大域アドレス空間に関する研究
- クラスタからクラスタ外部へのネットワーク出力に関する研究
- 大規模分散計算環境上のリアルタイムモニタリングシステムに関する研究
- ブロードキャスト型単一IPアドレスクラスタにおける柔軟なリクエスト分散機構(Distributed TCP Splicing, DTS)に関する研究

## 2 プログラム開発環境

以下の項目について、プログラムの性能改善のためのコンパイラ最適化の理論研究、プログラム開発環境に関する研究を実施した:



- 最適化検証つきコンパイラ(Optimization Verifying Compiler)に関する研究
- アルゴリズム変換を利用したコンパイラ最適化の拡張の研究
- GPU 向けプログラミング環境 OMP CUDA に関する研究
- 並列実効性能の高いプログラム開発環境に関する研究開発
- 

### 3 数値計算ライブラリ・アルゴリズム

SR110000/J2、HA8000 クラスタシステムを主たるターゲットとし、ペタスケール・エクサスケールシステムに向けて高性能数値計算法・ライブラリの研究開発を実施した。

PC からスーパーコンピュータに至る広範な計算機環境において、自動的に性能チューニングが達成できるソフトウェア(ソフトウェア自動チューニング)に関する、理論、方式、言語、およびライブラリに関する総合的研究を実施したほか、並列プログラミングモデルに関する研究も実施した。

この他、基礎的な数理アルゴリズムに関する研究も実施した。

主な研究開発項目は以下の通りである：

- 対称密行列用の超並列固有値ソルバの研究
- アルゴリズムと実装方式を自動チューニングする機能を有する疎行列反復解法ソルバの研究(1.で紹介した「シームレス高生産・高性能プログラミング環境」の一部)
- 並列実効性能の高い数値計算ライブラリに関する研究開発
- 並列多重格子法アルゴリズムに関する研究
- 並列プログラミングモデルに関する研究
- GPUを用いたアプリケーション高速化に関する研究
- “double-double”型データタイプ平方根の高速アルゴリズム
- 数学的な対象に対する証明の更正法
- 素数の数え上げの高速化

### 4 大規模並列シミュレーション

並列計算機を使用した大規模シミュレーションについて、アプリケーションの研究開発とそれを使用した科学的な研究の他、並列アプリケーション開発に必要な基盤技術の研究開発を実施した。

主な研究開発項目は以下の通りである：

- 沸騰現象を扱う大規模分子動力学法コードの研究開発
- 準安定液体における気泡核生成現象の研究
- 並列前処理手法、領域分割手法に関する研究
- 大規模並列形状処理に関する研究

### 5 サービス体系におけるセキュリティ保証

サービス体系におけるセキュリティ保障について、以下の研究を実施した：

- 認証におけるセキュリティ表現。
- LoA of Identity and Services

## 6 教育活動

全国共同利用機関を生かした人材育成の取り組みの一環として、本センターが所有するスーパーコンピュータを受講学生が活用できる新しい形態の全学的な教育プログラム(学際計算科学・工学人材育成プログラム)を、本センターが主導して実施している。工学系、理学系、情報理工学系、新領域創成科学の各研究科との緊密な協力のもと、2008年度から学部・修士課程を中心としたカリキュラム策定のためのワーキンググループを組織している。従来スーパーコンピューティング研究部門教員が担当している下記の3科目については、継続して実施した:

- スパコンプログラミング(工学部・工学系研究科共通科目)
- 並列計算プログラミング・先端計算機演習(大学院理学系研究科地球惑星科学専攻、理学系研究科大学院教育高度化プログラム)
- コンピュータ科学特別講義 I・II(有限要素法プログラミング)(大学院情報理工学系研究科コンピュータ科学専攻)

基礎的な並列プログラミング教育を目的とした、国内に例を見ないユニークな取り組みとして「お試しアカウント付きスパコン利用講習会」がある。2009年度は同講習会を4回実施した。

## 7 共同研究プロジェクトの推進

共同研究プロジェクトとしては、業務部門と協力して、「2009年度 T2K オープンスパコン(東大)共同研究プロジェクト」を実施した。本共同研究プロジェクトは、HA8000 クラスタシステムの利用環境の改善を目的として2009年4月～2010年3月の間実施され、「64ノード(1,024コア)」程度を使用する大規模計算を大量に行う研究を対象とする。この共同研究プロジェクトで採択された研究グループは、様々なシミュレーションのアルゴリズムの開発、プログラムの高速化に関する研究を本センターのスタッフと共同で実施する。研究成果は「T2K オープンスパコン(東大)」上でのその成果をライブラリ、HPC ミドルウェア等のアプリケーション開発環境整備にフィードバックすることにより、利用環境の向上に資することを最終的な目標とする。2009年度は8件の応募があり、7件が採用された。

## 8 広報・研究会活動

業務部門と協力して、広報誌「スーパーコンピューティング」を6回発行した他、特集号として、「2008年度公募型プロジェクト成果報告書(2010年2月)」を1回発行した。

部門の研究活動として、2009年度は先進スーパーコンピューティング環境研究会(Advanced Supercomputing Environment, ASE)(略称:ASE研究会)を2回開催した。

## スーパーコンピューティング研究部門 成果要覧

### 招待講演／招待論文

- [招待 1] 中島研吾:並列前処理手法と領域分割, マルチコア時代の戦略, 分野横断型研究会「アルゴリズムによる計算科学の融合と発展」, 筑波大学計算科学研究センター, 2009.
- [招待 2] 中島研吾, 大学におけるスーパーコンピューティング『共同利用・共同研究拠点』時代を迎えて, 日本 IBM 科学技術計算ソリューションセミナー, 2009.
- [招待 3] Nakajima, K., Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, The 12th International Specialist Meeting on Next Generation Models on Climate Change and Sustainability for High Performance Computing Facilities, 2010.
- [招待 4] Nakajima, K., Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, 2nd International Kyoto Forum on Krylov Subspace Methods, 2010.
- [招待 5] 金田康正: 計算機科学が拓く未来, 電子情報通信学会, 集積回路研究専門委員会 (LSI の未来を考えるワークショップ) 2010 年 3 月 6 日-8 日, 香川県香川郡直島町琴弾地, ベネッセハウス, 2010 年 3 月.
- [招待 6] 大島聡史: GPGPU の魅力と可能性と本当のところ, ICT スクール特別講演, 2009 年 12 月.
- [招待 7] Takahiro Katagiri: Auto-tuned Sparse Iterative Solver Toward Petascale Era, Third French-Japanese Workshop -- Petascale Applications, Algorithms, and Programming (PAAP) --, Shiran-Kaikan Hall Annex, Kyoto University, Apr. 2009.
- [招待 8] 片桐孝洋: ソフトウェア自動チューニング入門～ペタフロップス, 10 万並列を達成するための実装方式から数値アルゴリズムの自動最適化技術～, 第 38 回数値解析シンポジウム-NAS2009-, チュートリアル, 熱川ハイツ, 2009 年 6 月.
- [招待 9] 片桐孝洋: ソフトウェア自動チューニング技術の最新動向～マルチコア, ヘテロジニアス, 10 万並列な環境に向けた新しい最適化技術～, 第 9 回 ANS 研究会, 京都大学学術情報メディアセンター北館 3 階講習室, 2009 年 6 月.
- [招待 10] Takahiro Katagiri: Xablib: A Sparse Iterative Solver with a Generalized Auto-tuning Interface, and Overview of Auto-tuning Studies in Japan, 米国エネルギー省 (DOE) 主催, Center for Scalable Application Development Software (CScADS) Summer Workshops, Workshop on Libraries and Auto-tuning for Petascale Applications, Granlibakken Resort and Conference Center, Tahoe City, CA, USA (Organizers: Jack Dongarra, Keith Cooper, Rich Vuduc, Kathy Yelick), Aug. 2009.
- [招待 11] 片桐孝洋: ソフトウェア自動チューニングの最新動向 ～数値計算ライブラリへの適用を例にして～, 京都大学グローバル COE, 知識循環社会のための情報学教育研究拠点, 京都大学吉田キャンパス工学部総合校舎 406 号室, 2009 年 10 月.
- [招待 12] 片桐孝洋: マルチコア・超並列時代の並列固有値ソルバ実装法 ～ブロック化, マルチキャストを中心に, 日本応用数学会 3 部会連携「応用数理セミナー」, 国立情報学研究所, 資料集, pp. 16-35, 2009 年 12 月.

- [招待 13] 片桐孝洋:ソフトウェア自動チューニング:パソコンからスパコンまでの先進最適化技術 ～ 数値計算ライブラリを中心に～, 情報処理学会東北支部, 第 350 回研究講演会, 秋田県立大学本荘キャンパス 講義棟 K321 教室, 2009 年 12 月.
- [招待 14] Takahiro Katagiri, Takao Sakurai, Hisayasu Kuroda, Ken Naono, and Kengo Nakajima: Development of numerical computation policy facility and its effect on Xabclib, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), Yokohama, JAPAN), OW-1 Innovative Algorithms for Multicore and Hierarchical Computational Environments (Organizers: H. Okuda, K. Nakajima and T. Aoki), Mar. 2010.
- [招待 15] 嶋志田良和: クラスタ監視ソフト VGXP の開発・運用を通しての評価と課題, 第 9 回 PC クラスタシンポジウム, 2009 年 12 月.

## 著書／編集

- [著書 1] 藤田肇:ディペンダブル・シングルシステムイメージ OS, ソフトウェアデザイン 2010 年 2 月号, pp.89-92, 技術評論社.

## 査読付論文

- [査読付 1] 中島研吾: 並列反復法と自動チューニング—マルチコア時代の並列プログラミングモデル—, 特集: 科学技術計算におけるソフトウェア自動チューニング, 情報処理 50-6, 情報処理学会, pp.517-522, 2009.
- [査読付 2] Kengo Nakajima: Parallel Multistage Preconditioners by Extended Hierarchical Interface Decomposition for Ill-Conditioned Problems, Proceedings of International Conference on Parallel Computing (ParCo2009), 2009 (in press).
- [査読付 3] Kengo Nakajima: Flat MPI vs. Hybrid: Evaluation of Parallel Programming Models for Preconditioned Iterative Solvers on “T2K Open Supercomputer”, IEEE Proceedings of the 38th International Conference on Parallel Processing (ICPP-09) (Second International Workshop on Parallel Programming Models and Systems Software for High-End Computing (P2S2)), pp.73-80, 2009.
- [査読付 4] 中島研吾: 拡張階層型領域間境界分割に基づく悪条件問題向け並列前処理手法, 「ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2010」論文集, pp.91-98, 2010.
- [査読付 5] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰: OpenATLib: 数値計算ライブラリ向け自動チューニングインターフェース, 2010 年ハイパフォーマンスコンピューティングと計算科学シンポジウム (HPCS2010), HPCS2010 論文集, pp. 99-106, 2010 年 1 月.
- [査読付 6] Sato, Hiroyuki: Service Framework Based on Grades of IdPs and SPs, Proceedings of Security and Management 2009, pp.379–385, 2009.
- [査読付 7] Sato, Hiroyuki: Analyzing Semantics of Documents by using a Program Analysis Method, Proceedings of 33rd IEEE Computer Software and Applications Conference (COMPSAC), pp.373–382, 2009.
- [査読付 8] Nishimura, Takeshi, Sato, Hiroyuki: Analysis of a Security Incident of Opensource Middleware- Case Analysis of 2008 Debian Incident of OpenSSL-, in MidArch 2009, Proceedings of 9th International Symposium on Applications and the Internet, pp.247–250, 2009.

- [査読付 9] Sato, Hiroyuki: N § ": Reflecting Local Risk Assessment in LoA, On the Move 2009 (LNCS5871), pp.833–847, 2009.
- [査読付 10] Sato, Hiroyuki: Idiom Recognition and Program Scheme Recognition based Program Transformation for Performance Tuning – Beyond Compiler Optimizations–, Proceedings of Parallel and Distributed Computing, Applications and Technologies 2009, pp.272–279, 2009.
- [査読付 11] Nan Dun, Kenjiro Taura, and Akinori Yonezawa: Gmount: An ad hoc and locality-aware distributed file system by using ssh and fuse, In Proceedings of the IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID 2009), May 2009.
- [査読付 12] Nan Dun, Kenjiro Taura, and Akinori Yonezawa: Paratrac: A fine-grained profiler for data-intensive workflows, In Proceedings of ACM Conference on High Performance Distributed Computing (HPDC), 2010. to appear.
- [査読付 13] Ken Hironaka, Hideo Saito, and Kenjiro Taura: High performance wide-area overlay using deadlock-free routing, In Proceedings of International ACM Symposium on High Performance Distributed Computing (HPDC2009), pp. 81-90, 2009.
- [査読付 14] Takeshi Shibata, Sung~Jun Choi, and Kenjiro Taura: File-access patterns of data-intensive workflow applications and their implications to distributed filesystems: In Proceedings of The Third International Workshop on Data Intensive Distributed Computing, 2010. to appear.
- [査読付 15] 吉富翔太, 田浦健次郎:メッセージ衝突を防止する適応的な収集操作アルゴリズム, 電子情報通信学会論文誌(情報爆発特集号), J93-D(6), 2009年6月.
- [査読付 16] 原健太郎, 田浦健次郎, 近山隆:DMI:計算資源の動的な参加/脱退をサポートする大規模分散共有メモリインタフェース, 情報処理学会論文誌, 3(1):1--40, 2010年3月.
- [査読付 17] 柴田剛志, 田浦健次郎:トポロジ情報を用いた効率的かつ漸近安定な大容量ブロードキャスト, 先進的計算基盤システムシンポジウム(SACSYS2009)予稿集, pp. 71-78, 2009.
- [査読付 18] 柴田剛志, 田浦健次郎:トポロジ情報を用いた効率的かつ漸近安定な大容量ブロードキャスト, 情報処理学会論文誌コンピューティングシステム(ACS), 2(3), pp.47-57, 2009.
- [査読付 19] 吉富翔太, 弘中健, 田浦健次郎:メッセージ衝突を防止する適応的な収集操作アルゴリズム, 先進的計算基盤システムシンポジウム(SACSYS2009)予稿集, 2009.
- [査読付 20] Hiroya Matsuba, Yutaka Ishikawa: Aggregate Router: An Efficient Inter-Cluster MPI Communication Facility, IPSJ Transactions on Advanced Computing Systems Vol.2 No.3 (ACS 27), Sep. 2009.
- [査読付 21] 松葉浩也, 堀敦史, 石川裕:高性能クラスタのための高速汎用ステージングソフトウェア, 先進的計算基盤システムシンポジウム(SACSYS), pp. 239-246, May 2009.
- [査読付 22] Kazuki Ohta, Hiroya Matsuba, Yutaka Ishikawa: Improving Parallel Write by Node-Level Request Scheduling, IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2009), 2009.
- [査読付 23] Atsushi Hori, Yoshikazu Kamoshida, Hiroya Matsuba, Kazuki Ohta, Takashi Yasui, Shinji Sumimoto, Yutaka Ishikawa: On-Demand File Staging System for Clusters, IEEE International Conference on Cluster Computing (Cluster) 2009, 2009.
- [査読付 24] Shinji Sumimoto, Kohta Nakashima, Akira Naruse, Kouichi Kumon, Takashi Yasui, Yoshikazu Kamoshida, Hiroya Matsuba, Atsushi Hori, Yutaka Ishikawa: The Design of Seamless MPI Computing Environment for Commodity-based Clusters, Euro PVM/MPI09, 2009.



- [査読付 25] 堀敦史, 鴨志田良和, 松葉浩也, 安井隆, 住元真司, 石川裕: リングトポロジーによる MPI-IO 書込の高速化技法, SACSIS 2010, 2010 年 5 月 (発表予定).
- [査読付 26] 片桐孝洋: ペタフロップス環境における小規模行列用対称密行列固有値ソルバに向けて - 逆変換の改良 -, 2010 年ハイパフォーマンスコンピューティングと計算科学シンポジウム (HPCS2010), HPCS2010 論文集, pp. 27-34, 2010 年 1 月.
- [査読付 27] Takahiro Katagiri and Shoji Itoh: A Massively Parallel Dense Symmetric Eigensolver with Communication Splitting Multicasting Algorithm, 9th International Meeting on High Performance Computing for Computational Science (VECPAR'2010), Proceedings of VECPar'2010, Berkeley, California, USA, June 2010. (採録決定)
- [査読付 28] H. Watanabe: Non-equilibrium Relaxation Analysis on Two-dimensional Melting, Progress of Theoretical Physics, Supplement, No.178, pp. 41-48, 2009.
- [査読付 29] Balazs Gerofi, Hajime Fujita, Yutaka Ishikawa: Live Migration of Processes Maintaining Multiple Network Connections, コンピュータシステム・シンポジウム(ComSys2009), pp. 1-10, Nov. 2009.
- [査読付 30] Balazs Gerofi, Hajime Fujita, Yutaka Ishikawa: Live Migration of Processes Maintaining Multiple Network Connections, IPSJ Transactions on Advanced Computing Systems Vol.3 No.1 (ACS 29), pp. 1-12, Mar. 2010 ([査読付 29] と同時投稿).

## 公開ソフトウェア

- [公開 1] Nan Dun: paramark: high fidelity parallel file system benchmark.  
<http://code.google.com/p/paramark/>.
- [公開 2] Nan Dun: paratrac: a tracking suite for parallel/distributed applications.  
<http://code.google.com/p/paratrac/>.
- [公開 3] Kenjiro Taura: GXP: Grid and cluster shell.  
<http://www.logos.t.u-tokyo.ac.jp/gxp/>.
- [公開 4] Atsushi HORI: SCore Cluster System Version 7, <http://www.pccluster.org/ja/score7.html>, Feb 2010.

## その他の発表論文

- [発表 1] Kengo Nakajima: Parallel Multistage Preconditioners by Hierarchical Interface Decomposition on "T2K Open Super Computer (Todai Combined Cluster)" with Hybrid Parallel Programming Models, The 15th International Conference on Finite Elements in Flow Problems (FEF09), MS15: Parallel and Hierarchical Algorithms for Accelerating System Matrix Solvers, 2009.
- [発表 2] Kengo Nakajima: Evaluation of Hybrid Parallel Programming Models for Finite-Element Applications on T2K Open Supercomputer (Todai Combined Cluster), Third French-Japanese Workshop Petascale Applications, Algorithms and Programming (PAAP), 2009.
- [発表 3] 中島研吾: T2K オープンスパコン(東大)におけるハイブリッド並列プログラミングモデルの最適化, 日本計算工学会第 14 回計算工学講演会, 2009.



- [発表 4] 中島研吾:講義紹介:並列計算プログラミング, 先端計算機演習(地球惑星科学専攻), スーパーコンピューティングニュース(東京大学情報基盤センター) 11-3, 2009.
- [発表 5] 中島研吾:コンピュータ科学特別講義 I「科学技術計算プログラミング(有限要素法)」, スーパーコンピューティングニュース(東京大学情報基盤センター) 11-3, 2009.
- [発表 6] 中島研吾, 片桐孝洋:マルチコアプロセッサにおけるリオーダーリング付き非構造格子向け前処理付反復法の性能, 情報処理学会研究報告(HPC-120-6), 2009.
- [発表 7] 中島研吾:T2K オープンスパコン(東大)チューニング連載講座番外編 Hybrid 並列プログラミングモデルの評価(I), スーパーコンピューティングニュース(東京大学情報基盤センター) 11-4, 2009.
- [発表 8] 中島研吾:拡張階層型領域間境界分割に基づく並列前処理手法, 日本応用数理学会「行列・固有値問題の解法とその応用」研究部会(MEPA), 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP 仙台 2009), 2009.
- [発表 9] 中島研吾:マルチコア時代の前処理付き反復法と性能チューニング, 日本応用数理学会 2009 年年会, 2009.
- [発表 10] 中島研吾:T2K オープンスパコン(東大)チューニング連載講座番外編 Hybrid 並列プログラミングモデルの評価(II), スーパーコンピューティングニュース(東京大学情報基盤センター) 11-5, 2009.
- [発表 11] Kengo Nakajima:Flat MPI vs. Hybrid: Evaluation of Parallel Programming Models for Preconditioned Iterative Solvers on Multicore Clusters, The Fourth International Workshop on Automatic Performance Tuning (iWAPT2009), 2009
- [発表 12] 中島研吾:講習会:科学技術計算のためのマルチコアプログラミング入門, スーパーコンピューティングニュース(東京大学情報基盤センター) 11-6, 2009.
- [発表 13] 中島研吾:OpenMP/MPI ハイブリッド並列プログラミングモデルの多重格子法への適用, 情報処理学会研究報告(HPC-124-7), 2010.
- [発表 14] Kengo Nakajima:Framework for Development of Parallel Codes in “Integrated Predictive Simulation System for Earthquake and Tsunami Disaster”, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), MS 30: Coupling and Regriding Tools for Supporting Parallel Multi-physics Modeling - Part I of II, 2010.
- [発表 15] Kengo Nakajima:Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), MS 55: Joint JSIAM-SIAM Minisymposium: Parallel Programming Models and Algorithms for Multicore Clusters and GPGUs - Part II of III, 2010.
- [発表 16] Kengo Nakajima:Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), 2010.
- [発表 17] 片桐孝洋, 櫻井隆雄, 黒田久泰, 直野健, 中島研吾:汎用的自動チューニングインターフェースをもつ疎行列反復解法ライブラリ, 2009 年先進的計算基盤システムシンポジウム(Symposium on Advanced Computing Systems and Infrastructures (SACSIS)), 広島国際会議場, SACSIS2009 論文集, pp. 116-117, 2009 年 5 月.
- [発表 18] 片桐孝洋, 齊藤竜彦, 古村孝志, 中島研吾:3 次元津波伝搬シミュレーションにおけるコード最適化一手法, 情報処理学会研究報告 2009-HPC-120, 2009 年 6 月.

- [発表 19] 片桐孝洋, 櫻井隆雄, 黒田久泰, 直野健, 中島研吾: OpenATLib:汎用的な自動チューニングインターフェースの設計と実装, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 情報処理学会研究報告 2009-HPC-121, 2009 年 8 月.
- [発表 20] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰: OpenATLib を利用した疎行列ライブラリの開発と評価, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 情報処理学会研究報告 2009-HPC-121, 2009 年 8 月.
- [発表 21] 片桐孝洋, 黒田久泰, 中島研吾: Xabclib : 汎用な自動チューニングインターフェース OpenATLib を利用した疎行列反復解法ライブラリ, 第 5 回 ASE 研究会, 2009 年 8 月.
- [発表 22] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰: クリロフ部分空間法に対する自動チューニングについて, 加速法フォーラム, 2009 年 12 月.
- [発表 23] Takahiro Katagiri, Hisayasu Kuroda and Kengo Nakajima: Xabclib: An Iterative Solver with a General Auto-tuning Interface "OpenATLib" --- Development of Numerical Computation Policy Function, 6th ASE Seminar, Information Technology Center, The University of Tokyo, Feb. 2010.
- [発表 24] Takahiro Katagiri, Takao Sakurai, Hisayasu Kuroda, Ken Naono and Kengo Nakajima: Xabclib: An Iterative Solver with a General Auto-tuning Interface "OpenATLib", International Workshop on Peta-Scale Computing Programming Environment, Languages and Tools (WPSE 2010), Kyoto, Feb. 2010.
- [発表 25] Takao Sakurai, Ken Naono, Hisayasu Kuroda, Takahiro Katagiri, and Kengo Nakajima: OpenATLib: A General Auto-tuning Interface for Numerical Solvers, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: The State-of-the-art of Auto-tuning Technologies: Adaptation to Advanced Computer Environment and Numerical Libraries - Part I of II, MS6, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.
- [発表 26] Takahiro Katagiri, Hisayasu Kuroda and Kengo Nakajima: A Branchless Segmented Scan Method for Sparse Matrix-vector Multiplications, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: Parallel Programming Models and Algorithms for Multicore Clusters and GPGPUs - Part III of III, MS62, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.
- [発表 27] 清野善裕, 吉田仁, 金田康正: Riemann Zeta 関数の零点計算の高速化, 日本応用数理学会 2009 年度年会, p.333, 2009 年 9 月.
- [発表 28] 清野善裕, 吉田仁, 金田康正: Riemann zeta 関数の零点とその計算の基礎について, 第 8 回「代数学と計算」研究集会(AC2009), 2009 年 12 月.
- [発表 29] 筒井直機: 多倍精度演算の高速化とその周辺関数に関する研究について, 東京大学グローバル COE セキュアライフ・エレクトロニクス平成 21 年度博士課程学生報告書, pp.120-121, 2009.
- [発表 30] 雪下洋輔, 佐藤周行: シンボリックアサインメントグラフを用いたコンパイラ最適化器, 第 26 回ソフトウェア科学大会, 3C-2, 2009.
- [発表 31] 西村健, 佐藤周行: 2008 年の Debian OpenSSL インシデントにみるオープンソースソフトウェアのセキュリティ分析, コンピュータセキュリティシンポジウム 2009, 2009.

- [発表 32] 吉富翔太, 田浦健次郎:メッセージ衝突を防止する適応的な集合通信, 情報処理学会研究報告 HPC-121 (SWoPP 2009), 2009.
- [発表 33] 原健太郎, 田浦健次郎, 近山隆:DMI: 計算資源の動的な参加/脱退をサポートする大規模分散共有メモリインタフェース, 情報処理学会研究報告 HPC-121 (SWoPP 2009), 2009.
- [発表 34] 長沼翔, 田浦健次郎:大規模分散環境におけるバンド幅測定アルゴリズム, 情報処理学会研究報告 HPC-121 (SWoPP 2009), 2009.
- [発表 35] 住元真司, 中島耕太, 成瀬彰, 久門耕一, 安井隆, 鴨志田良和, 松葉浩也, 堀敦史, 石川裕: シームレスな MPI 環境を実現する MPI-Adapter の設計と性能評価, 情報処理学会研究報告 (SWoPP2009), 2009.
- [発表 36] 堀敦史, 鴨志田良和, 松葉浩也, 太田一樹, 安井隆, 住元真司, 石川裕: ファイルステージングシステム Catwalk の MPI-IO 実装, 情報処理学会研究報告(SWoPP 2009), 2009.
- [発表 37] 安井隆, 清水正明, 住元真司, 太田一樹, 鴨志田良和, 松葉浩也, 堀敦史, 石川裕: ファイルキャッシュシステムの有効性向上に向けた科学技術計算アプリケーションの I/O 特性評価, 情報処理学会研究報告(SWoPP 2009), 2009.
- [発表 38] 大島聡史, 平澤将一, 本多弘樹: OpenMP and auto-tuning framework for GPU, IWA-COM-II(2nd International Workshops on Advances in Computational Mechanic), p.41, 2010.
- [発表 39] 平澤将一, 大島聡史, 本多弘樹: GPU コンピューティング向け中間言語の研究, 2010.
- [発表 40] Atsushi HORI: Parallel File IO Using Ring Communication Topology, WPSE2010, Feb.2010.
- [発表 41] 片桐孝洋, 黒田久泰:マルチコア環境における自動チューニング機能付き疎行列反復解法ライブラリ, 第 14 回日本計算工学会講演会, オーガナイズドセッション, 「科学技術計算における自動チューニングの挑戦:マルチコア環境への適用」, 東京大学生産研究所計算工学会講演論文集, pp. 167-170, 2009 年 5 月.
- [発表 42] 片桐孝洋:ペタスケール計算を目指した MRRR 法を用いた固有値ソルバの開発, 第 14 回日本計算工学会講演会, オーガナイズドセッション, 「解けたらうれしい固有値問題」, 東京大学生産研究所, 計算工学会講演論文集, pp. 185-188, 2009 年 5 月.
- [発表 43] 片桐孝洋, 黒田久泰:マルチコア環境における密および疎行列ソルバの自動チューニング機構の評価, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ (SWoPP2009), フォレスト仙台, 日本応用数理学会「行列・固有値問題の解法とその応用」研究部会, 2009 年 8 月.
- [発表 44] 黒田久泰, 片桐孝洋, 須田礼仁:省電力のための基本演算ライブラリの実装と評価, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 日本応用数理学会「行列・固有値問題の解法とその応用」研究部会, 2009 年 8 月.
- [発表 45] 片桐孝洋, 黒田久泰:マルチコア・超並列計算機時代の自動チューニング機能付き疎行列反復解法ソルバ, 日本応用数理学会 2009 年度年会, 大阪大学豊中キャンパス, 年会予稿集, オーガナイズドセッション:ペタスケール環境を目指す数値計算ライブラリと自動チューニング技術, A2-1, 2009 年 9 月.
- [発表 46] Takahiro Katagiri: Model for Software Automatic Tuning, Poster Presentation, Forth International Workshop on Automatic Performance Tuning (iWAPT2009), Proceedings of iWAPT2009, pp. 74, Oct. 2009.

- [発表 47] 片桐孝洋:ソフトウェア工学としての展開,「自動チューニング技術の現状と応用に関するシンポジウム」, 東京大学弥生講堂アネックスセイホクギャラリー, 2009年10月.
- [発表 48] 渡辺宙志, 鈴木将, 伊藤伸泰: 気泡生成過程の微視的仕事と非平衡性, 日本物理学会第65回年次大会, 2010年3月.
- [発表 49] 能川知昭, 伊藤伸泰, 渡辺宙志: 2次元弾性円盤系における並進、回転対称性の破れ, 日本物理学会第65回年次大会, 2010年3月.
- [発表 50] 渡辺宙志、伊藤伸泰: 気泡生成のシミュレーション解析, 地震研共同利用研究集会「火山現象の数値計算研究」, 2009年11月.
- [発表 51] 渡辺宙志: 気液混相流における熱伝導と核生成, 日本機械学会第22回計算力学講演会, 2009年10月.
- [発表 52] Hiroshi Watanabe: Huge-scale molecular dynamics simulation on gas-liquid multi-phase flow, 8th Workshop on Stochastic Analysis on Large Scale Interacting Systems, 2009年10月.
- [発表 53] 渡辺宙志, 轟木義: 定温ダイナミクスとその幾何学的理解, 日本物理学会秋季大会, 2009年9月.
- [発表 54] 渡辺宙志, 鈴木将, 伊藤伸泰: 気液混相流における熱伝導と核生成, 日本物理学会秋季大会, 2009年9月.
- [発表 55] 能川知昭, 伊藤伸泰, 渡辺宙志: 相境界面運動の分子動力学シミュレーションによる解析, 日本物理学会秋季大会, 2009年9月.
- [発表 56] 山田賢太, 島田尚, 渡辺宙志, 伊藤伸泰: レナードジョーンズ液滴の衝突破砕シミュレーションによるクラスタ解析, 日本物理学会秋季大会, 2009年9月.
- [発表 57] 渡辺宙志: Pressure in liquid-gas interface, 第11回 ACP ワークショップ, 2009年9月.
- [発表 58] 渡辺宙志: 動的ロードバランスを備えた並列分子動力学法, 第9回 ACP ワークショップ, 2009年4月.
- [発表 59] 鴨志田良和, 田浦健次朗: 並列アプリケーションの性能を損なわないポーリング型のモニタリング, 情報処理学会研究報告 2009-HPC-121, 仙台, 2009年8月.
- [発表 60] 藤田肇, 石川裕: レイヤー7 負荷分散のための TCP 接続移送機構, 情報処理学会研究報告(OS-111), 2009年4月.
- [発表 61] Balazs Gerofi, Hajime Fujita, Yutaka Ishikawa: TCPmig: Migration of processes with TCP sockets in Single IP Address Cluster, 情報処理学会研究報告(OS-111), 2009年4月.
- [発表 62] 中澤仁, 松野裕, 菅谷みどり, 塙敏博, 前田俊行, 藤田肇, 石綿陽一, 杵渕雄樹, 高村博紀, 三浦信一, 山田浩史, 松田元彦, 倉光君郎, 石川裕: オペレーティングシステムおよび実システムにおけるディペンダビリティの評価と見積り, 第7回ディペンダブルシステムワークショップ (DSW2009), pp. 27-41, 2009年7月.
- [発表 63] 加藤真平, 藤田肇, 中澤仁, 松田元彦, 前田俊行, 杵渕雄樹, 塙敏博, 三浦信一, 石綿陽一, 松野裕, 高村博紀, 山田浩史, 吉田哲也, 倉光君郎, 菅谷みどり, 石川裕: ディペンダブルシステム向けベンチマークフレームワークの提案, 第7回ディペンダブルシステムワークショップ (DSW2009), pp. 171-178, 2009年7月.
- [発表 64] 加藤純, 藤田肇, 石川裕: 単一 IP アドレスクラスタにおける耐故障機構の設計と実装, 情報処理学会研究報告(OS-112 SWoPP 2009), 2009年8月.



## 特記事項

- [特記 1] 中島研吾: パネリスト, 「PC クラスタから次世代スパコンへ –アカデミックと企業連携のあり方–」, 第 9 回 PC クラスタシンポジウム, 2009.
- [特記 2] 中島研吾: パネリスト, 「シームレスな計算機環境を目指して」, PC クラスタワークショップ in 京都 2010, 2010.
- [特記 3] 中島研吾: General Council Member, IACM (International Association for Computational Mechanics)
- [特記 4] 中島研吾: 日本計算工学会 評議員
- [特記 5] 中島研吾: 日本応用数理学会 評議員
- [特記 6] 中島研吾: 日本応用数理学会 学会誌編集委員(幹事)
- [特記 7] 中島研吾: 日本応用数理学会 「行列・固有値問題の解法とその応用」研究部会運営委員
- [特記 8] 中島研吾: 日本機械学会計算力学技術者認定事業 計算力学技術者 1 級・2 級(固体力学分野の有限要素法解析技術者)認定試験 試験問題作成委員(能力開発促進機構 計算力学技術者認定事業委員会 固体力学分野 試験問題作成 WG 委員)(2005.4~)
- [特記 9] 中島研吾: 情報処理学会 論文誌(コンピューティングシステム)編集委員
- [特記 10] 中島研吾: Associate Editor, SIAM Journal for Scientific Computing
- [特記 11] 中島研吾: Member of Organizing Committee, 15th International Conference in Finite Elements in Flow Problems (FEF 2009), 2009.
- [特記 12] 中島研吾: Member of Program Committee, 2009 IEEE International Workshop on HPC and Grid Applications (IWHGA2009), 2009.
- [特記 13] 中島研吾: 日本計算工学会第 14 回計算工学講演会 実行委員, 2009.
- [特記 14] 中島研吾: Member of Program Committee, Parallel Computing 2009 (ParCo 2009), 2009.
- [特記 15] 中島研吾: Member of Technical Papers Committee, 2009 IEEE International conference on high performance computing, networking, storage, and analysis (SC09), 2009.
- [特記 16] 中島研吾: ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2010, プログラム委員, 2010.
- [特記 17] 中島研吾: Member of International Program Committee, IASTED International Conference on Parallel and Distributed Computing and Networks (PDCN 2010), 2010.
- [特記 18] 中島研吾: Member of Organizing Committee, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), 2010.
- [特記 19] 中島研吾: Co-organizer of Mini-Symposium, Parallel and Hierarchical Algorithms for Accelerating System Matrix Solvers, 15th International Conference in Finite Elements in Flow Problems (FEF 2009), 2009.
- [特記 20] 中島研吾: セッションオーガナイザ, 日本計算工学会第 14 回計算工学講演会「大規模シミュレーションと並列前処理手法」, 2009.
- [特記 21] 中島研吾: セッションオーガナイザ, 第 23 回数値流体力学シンポジウム「並列計算, グリッド計算」, 2008.

- [特記 22] 中島研吾: Co-organizer of Mini-Symposium, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), Coupling and Regriding Tools for Supporting Parallel Multi-physics Modeling Part I-II, 2010.
- [特記 23] 中島研吾: Co-organizer of Mini-Symposium, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), Joint JSIAM-SIAM Minisymposium: Parallel Programming Models and Algorithms for Multicore Clusters and GPGPUs Part I-III, 2010.
- [特記 24] 中島研吾: Co-organizer of Workshop, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), Innovative algorithms for multicore and hierarchical computational environments, 2010.
- [特記 25] 中島研吾: 学際計算科学・工学人材育成プログラム, <http://nkl.cc.u-tokyo.ac.jp/CSEedu/>
- [特記 26] 中島研吾: 並列計算プログラミング, 先端計算機演習, 東京大学大学院理学系研究科地球惑星科学専攻, <http://nkl.cc.u-tokyo.ac.jp/09e/>
- [特記 27] 中島研吾: コンピュータ科学特別講義 I・II, 東京大学大学院情報理工学系研究科コンピュータ科学専攻, <http://nkl.cc.u-tokyo.ac.jp/09s/>, <http://nkl.cc.u-tokyo.ac.jp/09w/>
- [特記 28] 黒田久泰, 直野健, 岩下武史: 情報処理学会誌「情報処理」, 大特集: 科学技術計算におけるソフトウェア自動チューニング, 6. 自動チューニング機能付き数値計算ライブラリ, Vol. 50, No 6, pp.505-511, 2009年6月.
- [特記 29] 松葉浩也: A STUDY OF HIGH-QUALITY NETWORK TRANSMISSION ON COMPUTER CLUSTERS, 東京大学学位論文博士(情報理工学), 2010年3月.
- [特記 30] GPU チャレンジ 2010, SACSIS2010 - 先進的計算基盤システムシンポジウム, 2010年3月.
- [特記 31] 堀敦史: SCore7 最新状況, 第九回 PC クラスタシンポジウム講演, 2008年12月.
- [特記 32] 堀敦史: Catwalk Tutorial, PC クラスタワークショップ in 京都講演, 2010年2月.
- [特記 33] 堀敦史: SCore7 最新状況, PC クラスタワークショップ in 京都講演, 2010年2月.
- [特記 34] 片桐孝洋: スーパーコンピュータ若手利用者推薦(試行)制度 2009年度(前期)4件(新規3件, 継続1件), 2009年度(後期)4件(新規1件, 継続3件).
- [特記 35] 片桐孝洋: 東京大学情報基盤センタースーパーコンピューティング部門主催, Advanced Supercomputing Environment (ASE)研究会: 2009年8月21日第5回研究会, 2010年2月12日第6回研究会の2回実施.
- [特記 36] 片桐孝洋: お試しアカウント付き並列プログラミング講習会(試行): 2009年7月1日~2日, 2009年9月7日~8日, 2010年3月17日の3回実施.
- [特記 37] 片桐孝洋: ミニシンポジウム・オーガナイザー, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: The State-of-the-art of Auto-tuning Technologies: Adaptation to Advanced Computer Environment and Numerical Libraries, MS6 and MS42, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.
- [特記 38] 片桐孝洋: 情報処理学会誌「情報処理」, 大特集: 科学技術計算におけるソフトウェア自動チューニング, 片桐孝洋: 編集にあたって, pp. 476, 弓場敏嗣, 片桐孝洋: ソフトウェア自動チューニングの枠組み, pp. 478-482, 片桐孝洋: ソフトウェア自動チューニング記述のための計算機言語, pp. 494-498, Vol. 50, No 6, 2009年6月.



- [特記 39] 大島聡史: GPU チャレンジ 2010, SACSIS2010 - 先進的計算基盤システムシンポジウム, 2010年3月コンテスト終了.
- [特記 40] 大島聡史: これからの並列計算のための GPGPU 連載講座(I) GPU と GPGPU の歴史と特徴, スーパーコンピューティングニュース, Vol.12, No.1, pp.33-46, 2010.
- [特記 41] 大島聡史: これからの並列計算のための GPGPU 連載講座(II) GPGPU プログラミング環境 CUDA 入門編, スーパーコンピューティングニュース, Vol.12, No.2, pp.14-28, 2010.
- [特記 42] 鴨志田良和: クラスタシステム上での並列プログラミングコンテスト実行委員, 2009年3月～5月.
- [特記 43] 鴨志田良和: 第2回クラスタシステム上での並列プログラミングコンテスト実行委員, 2009年7月～12月.
- [特記 44] 鴨志田良和: お試しアカウント付き並列プログラミング講習会(試行): 2009年7月, 2009年9月, 2010年3月.
- [特記 45] 鴨志田良和: 学術グリッド基盤の構築・運用技術に関する研究共同研究者, 平成21年度学際大規模情報基盤共同利用・共同研究拠点公募型共同研究(試行), 2009年.
- [特記 46] 鴨志田良和: グリッド配備・運用タスクフォース.
- [特記 47] 鴨志田良和: NAREGI アーキテクトタスクフォース.

## 報道関連

- [報道 1] 片桐孝洋: 東京大学新聞, 2009年9月22日(火)発刊, 第2480号, 「10分でゲノム解読する時代に—ゲノム解読をスパコンで, 人材育成が急務, 情報悪用を防げ」, 2009年9月.
- [報道 2] 片桐孝洋, 黒田久泰, 中島研吾: 科学技術政策研究所発行, 「科学技術動向」11月号特集, 「数値シミュレーションにおけるソフトウェア研究開発の動向—並列分散型のハードウェアとソフトウェア自動チューニング—」, 2009年11月.

# 大規模並列シミュレーション手法に関する研究、教育

中島 研吾

## 1 概要

有限要素法を中心とした、大規模並列シミュレーション手法に関する研究、教育を実施した。主な実施項目を以下に示す：

- 並列前処理手法、領域分割手法に関する研究

科学技術シミュレーションにおいて、有限要素法等の離散化を使用して得られる大規模疎行列を係数行列とする連立一次方程式を、並列計算機上で反復法によって解く場合の前処理手法、領域分割手法について研究を実施した。実問題において得られる悪条件な行列では適切な前処理手法、領域分割手法の選択が重要である。有限要素法における悪条件問題を中心に、問題、要素の特性に基づいた前処理・領域分割手法(選択的 Fill-in、選択的オーバーラップ)を提案し、様々な実問題における検証を実施したほか、Hierarchical Interface Decomposition (HID)の実問題向け拡張に関する研究を実施した [査読付 1~4][発表 1~11]。

- 並列プログラミングモデルに関する研究

近年プロセッサのマルチコア化が進み、並列プログラミングモデルとして、ノード(ソケット)内に OepnMP などのスレッド、ノード(ソケット)間にメッセージパッシング(MPI)を適用する「Hybrid」並列プログラミングモデルが再び脚光を浴びている。多重格子法を中心として、T2K オープンスパコン(東大)を使用して様々な Hybrid 並列プログラミングモデルの評価、最適化を実施した[招待 1, 3, 4]、[発表 13, 15, 16]。

- 大規模並列形状処理に関する研究

有限要素法による大規模並列問題を対象とした、メッシュ生成、可視化、適応格子、動的負荷分散に関する研究を実施した。特に、本年度は六面体要素を使用した並列有限要素法による、不均質複雑断面における単位すべり応答関数計算コードを対象として、適合型局所細分化のためのフレームワークのプロトタイプを開発し、負荷分散におけるデータマイグレーション部分の性能に注目した評価を T2K オープンスパコン(東大)を使用して実施した[発表 14]。

- HPC 教育

「学際計算科学・工学人材育成」委員会(委員長:松本洋一郎副学長)の活動の一環として全学的な HPC 教育プログラムの策定を実施したほか、試験的な講義、演習を実施した [特記 25~27]。内外の講演会において、本教育プログラムに関する紹介を実施した[招待 2]。

## 2 ハイブリッド並列プログラミングモデルの多重格子法への適用

### 2.1 背景

近年、マルチコアプロセッサの普及、大規模システムにおけるコア数の増加を背景として、ハイブリッド (Hybrid) 並列プログラミングモデルが脚光を浴びるようになり、Flat MPI (または Pure MPI) との優劣に関する議論が盛んとなっている。Hybrid 並列プログラミングモデルはメッセージパッシングによる「coarse-grain parallelism」と、ディレクティブによる「fine-grain parallelism」の融合であり、一般的には MPI と OpenMP を組み合わせたスタイルである。

著者は、有限要素法に基づく三次元弾性静力学問題向けシミュレーションで使用されている前処理つき反復法に OpenMP/MPI ハイブリッド並列プログラミングモデルを適用し、T2K オープンスパコン(東大) (以下「T2K(東大)」)の 512 コアを使用した評価を実施した。OpenMP/MPI ハイブリッド並列プログラミングモデルは適切な NUMA control の組み合わせにより、OpenMP/MPI ハイブリッド並列プログラミングモデルが Flat MPI と同等かそれを上回る性能を発揮することがわかった。更に、First Touch Data Placement, 連続メモリアクセスのためのデータ再配置を適用することにより、特にコア当たり問題規模が小さい場合の性能が改善されることが明らかとなった。

本研究では、OpenMP/MPI ハイブリッド並列プログラミングモデルを、並列多重格子前処理付き反復法を使用した、三次元有限体積法に基づく不均質場における地下水流れ問題シミュレーションに適用した。多重格子法 (multigrid) は、大規模問題におけるスケーラブルな手法として注目されているが、T2K(東大)のようなマルチコア・マルチソケットクラスタにおいて、OpenMP/MPI ハイブリッド並列プログラミングを適用し、評価した例は無い。本研究では T2K(東大)の他、米国ローレンスバークレイ国立研究所 National Energy Research Scientific Computing Center (NERSC) の有する「Cray-XT4」の 1,024 コアまでを使用して Flat MPI と OpenMP/MPI ハイブリッド並列プログラミングモデルの評価を実施した。

First Touch Data Placement, 連続メモリアクセスのためのデータ再配置, 適切な NUMA control の組み合わせにより、OpenMP/MPI ハイブリッド並列プログラミングモデルが Flat MPI と同等かそれを上回る性能を発揮することがわかった。特に問題規模が大きくなり、コア数が増加すると、OpenMP/MPI ハイブリッド並列プログラミングモデルの優位性は顕著である。

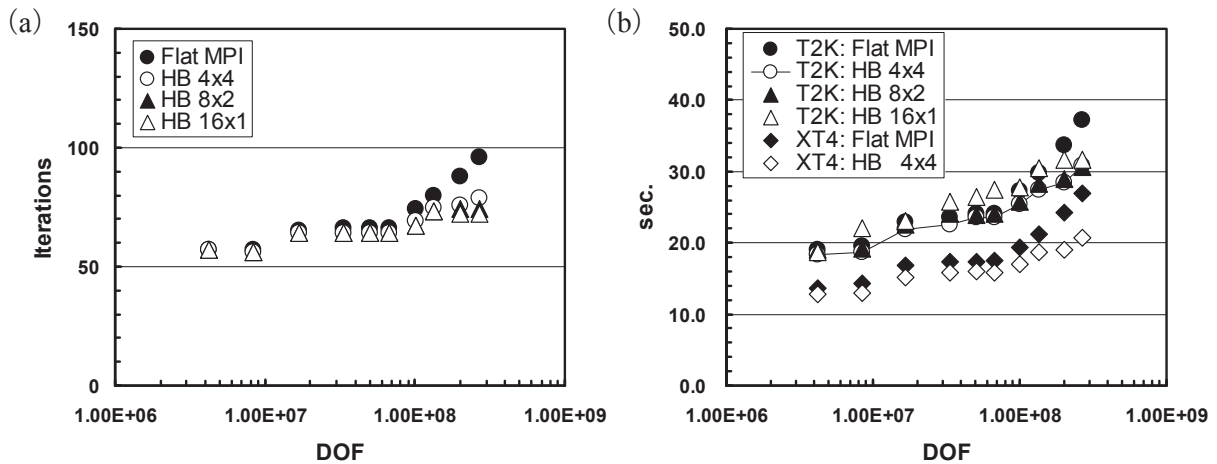


図1 MGCG法の計算性能((a)反復回数, (b)MGCGソルバー計算時間)(最適化されたソルバーを適用, CM-RCM(2), 16~1,024 コア)(Weak Scaling, 262,144 セル/コア, 最大問題規模: 268,435,456セル)(不均質多孔質媒体中の三次元地下水流れ)

### 3 東京大学における全学的 HPC 教育の試み「学際計算科学・工学人材育成プログラム」

#### 3.1 背景

計算科学・工学、ハードウェアの急速な進歩、発達を背景に、「第3の科学」としての大規模並列シミュレーションへの期待は、産学において一層高まっている。最新ハードウェアを駆使して、大規模並列シミュレーションプログラムを開発するための体系的な教育プログラムの事例は、世界的に見ても少ない。本学では2008年初頭から、「T2K オープンスパコン、次世代スーパーコンピュータ等を駆使した大規模シミュレーションを実施し、新しい科学を開拓する人材」を育成するための全学的な教育プログラムとして「学際計算科学・工学 人材育成プログラム」構想が検討されている。平尾公彦副学長をヘッドとして、情報基盤センターが中心となって、関連部局(理学系・工学系・情報理工学系・新領域創成科学各研究科、生産技術研究所、気候システム研究センター)の緊密な協力のもとに作業が進められている。2008年度冬学期から既に試験的な講義が開講されている。

#### 3.2 並列計算プログラミング・先端計算機演習(理学系研究科)

「学際計算科学・工学人材育成プログラム」の一環として、大学院理学系研究科地球惑星科学専攻で「並列計算プログラミング・先端計算機演習」を実施した[特記25]。2009年度夏季集中講義として、2009年8月に6回の講義と演習を実施した。本講義・演習は、21世紀COEプログラム「多圏地球システムの進化と変動の予測可能性(観測地球科学と計算地球科学の融合拠点の形成)」(2003年度~2007年度)において2004年度より開講されたもので、2008年度からは「理学系研究科大学院教育高度化プログラム」に認定され、夏季集中講義として開講された。

本講義・演習は、科学技術計算プログラミングに必須の項目である「SMASH (Science-Modeling-Algorithm-Software-Hardware)」を、できるだけ幅広くカバーし、広い視野を持った人材を育成することを最終的な目標とし、特に、MPI (Message Passing Interface) を使って、差分法、有限要素法、境界要素法等によるアプリケーションを並列化する能力を身につける、ことを重視している。

MPI には 400 以上の関数があるが、実際、科学技術シミュレーションで必要になるのは 10 程度である。本講義・演習では、まず、科学技術シミュレーションの手法を 2 種類(局所的手法、大域的手法)に分類し、それぞれを「並列化」するために必要な最小限の MPI 関数(1 対 1 通信、グループ通信)について習得する。MPI の基本的な考え方である SPMD (Single-Program Multiple-Data) を実際のアプリケーションの並列化を通して学ぶことに主眼を置いている。

講義では、アプリケーションの中身をよく理解し、並列化するための局所データ構造をどのように考えるか、というような話題を中心に進める。アプリケーションの中身、アルゴリズムの特性、をよく理解していれば、「並列化」というのはそれほど難しいことではない。従って講義の中では「並列計算」に直接結びつくものは少なく、むしろ基礎的なモデリングやアルゴリズムに関する基本的な項目について、学習することを重視している。ごく基本的なプログラミングと数値解析に関する知識があれば、これまで、自分でアプリケーションプログラムを開発した経験が無くても充分受講できる内容である。

プログラミング能力を高めるためには実習が必須である。本講義・演習では T2K オープンスパコン(東大)を使用したプログラミング実習を実施した。2009 年度は合計 50 名の受講者があった。2010 年度は同様の内容で 8 月 16 日～27 日の間に集中講義として実施する。

### 3.3 コンピュータ科学特別講義 I・II (科学技術計算プログラミング(有限要素法))(情報理工学系研究科)

「学際計算科学・工学人材育成プログラム」の一環として、大学院情報理工学系研究科地球惑星科学専攻で「コンピュータ科学特別講義 I・II : 科学技術計算プログラミング(有限要素法)」を 2009 年夏・冬学期に実施した[特記 26]。

偏微分方程式の数値解法として、様々な科学技術分野のシミュレーションに使用されている有限要素法について、背景となる基礎的な理論から、実用的なプログラムの作成法まで、連立一次方程式解法などの周辺技術も含めて講義を実施し、情報基盤センター教育用計算機システム (ECCS2008) を使用したプログラミングの実習を行なった。題材としては一次元及び三次元弾性静力学を扱った(夏学期: I)。更に、冬学期(II)では、並列計算プログラミングのためのデータ構造の考え方、並列アルゴリズム、並列プログラムの作成法に関する講義、T2K オープンスパコン(東大)実習を実施した。

有限要素法は計算機と深い関係にあり、計算機の発展とともに進歩してきた分野であるが、本学の各学部、研究科において実施されている有限要素法関連の講義は、理論、アルゴリズムに関する教育が中心で、プログラミングまでカバーしているものはほとんどない。また、有限要素法は最終的には疎行列を係数とする大規模な連立一次方程式を解くことに帰着されるため、疎行列を係数とする行列解法と密接な関係を持っている。有限要素法を学ぶためには、背景となる物理、変分法など



の基礎的な理論の他に、疎行列解法、特にプログラミングのためには疎行列の係数格納法に習熟することが不可欠であるが、疎行列解法まで含んだ教育を実施している講義はない。

担当者は、計算力学が専門であるが、数値線形代数、特に実用問題向けの前処理付並列反復法の研究に長年従事しており、疎行列解法と関連したプログラミング技術の教育を実施してきた。本講義は有限要素法そのものだけでなく、連立一次方程式解法についても学習できる非常にユニークな試みと言える。

大規模並列シミュレーションにおいては、科学・工学と計算機科学の専門家の密接な協力が必要である。本講義は、単に並列アプリケーション開発技術を習得するだけでなく、特にコンピュータ科学専攻の学生がアプリケーション側のニーズを把握し、両分野の融合領域を開拓する問題意識を育てることを目的としている。2009年度は夏学期(12名)と冬学期(6名)の受講者があった。2010年度は数理情報学専攻「科学技術計算 I・II」として開講の予定である。

## 4 成果要覧

### 招待講演

[招待 1] 中島研吾「並列前処理手法と領域分割, マルチコア時代の戦略」, 分野横断型研究会「アルゴリズムによる計算科学の融合と発展」, 筑波大学計算科学研究センター, 2009.

[招待 2] 中島研吾「大学におけるスーパーコンピューティング『共同利用・共同研究拠点』時代を迎えて」, 日本 IBM 科学技術計算ソリューションセミナー, 2009.

[招待 3] Nakajima, K., Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, The 12th International Specialist Meeting on Next Generation Models on Climate Change and Sustainability for High Performance Computing Facilities, 2010.

[招待 4] Nakajima, K., Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, 2nd International Kyoto Forum on Krylov Subspace Methods, 2010.

### 査読付論文リスト

[査読付 1] 中島研吾: 並列反復法と自動チューニング—マルチコア時代の並列プログラミングモデル—, 特集: 科学技術計算におけるソフトウェア自動チューニング, 情報処理 50-6, 情報処理学会, pp.517-522, 2009.

[査読付 2] Kengo Nakajima: Parallel Multistage Preconditioners by Extended Hierarchical Interface Decomposition for Ill-Conditioned Problems, Proceedings of International Conference on Parallel Computing (ParCo2009), 2009 (in press).



[査読付 3] Kengo Nakajima: Flat MPI vs. Hybrid: Evaluation of Parallel Programming Models for Preconditioned Iterative Solvers on “T2K Open Supercomputer”, IEEE Proceedings of the 38th International Conference on Parallel Processing (ICPP-09) (Second International Workshop on Parallel Programming Models and Systems Software for High-End Computing (P2S2)), pp.73-80, 2009.

[査読付 4] 中島研吾: 拡張階層型領域間境界分割に基づく悪条件問題向け並列前処理手法, 「ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2010」論文集, pp.91-98, 2010.

## その他発表論文リスト

[発表 1] Kengo Nakajima: Parallel Multistage Preconditioners by Hierarchical Interface Decomposition on “T2K Open Super Computer (Today Combined Cluster)” with Hybrid Parallel Programming Models, The 15th International Conference on Finite Elements in Flow Problems (FEF09), MS15: Parallel and Hierarchical Algorithms for Accelerating System Matrix Solvers, 2009.

[発表 2] Kengo Nakajima: Evaluation of Hybrid Parallel Programming Models for Finite-Element Applications on T2K Open Supercomputer (Today Combined Cluster), Third French-Japanese Workshop Petascale Applications, Algorithms and Programming (PAAP), 2009.

[発表 3] 中島研吾: T2K オープンスパコン(東大)におけるハイブリッド並列プログラミングモデルの最適化, 日本計算工学会第 14 回計算工学講演会, 2009.

[発表 4] 中島研吾: 講義紹介: 並列計算プログラミング, 先端計算機演習(地球惑星科学専攻), スーパーコンピューティングニュース(東京大学情報基盤センター) 11-3, 2009.

[発表 5] 中島研吾: コンピュータ科学特別講義 I 「科学技術計算プログラミング(有限要素法)」, スーパーコンピューティングニュース(東京大学情報基盤センター) 11-3, 2009.

[発表 6] 中島研吾, 片桐孝洋: マルチコアプロセッサにおけるリオーダーリング付き非構造格子向け前処理付反復法の性能, 情報処理学会研究報告(HPC-120-6), 2009.

[発表 7] 中島研吾: T2K オープンスパコン(東大)チューニング連載講座番外編 Hybrid 並列プログラミングモデルの評価(I), スーパーコンピューティングニュース(東京大学情報基盤センター) 11-4, 2009.

[発表 8] 中島研吾: 拡張階層型領域間境界分割に基づく並列前処理手法, 日本応用数理学会「行列・固有値問題の解法とその応用」研究部会(MEPA), 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP 仙台 2009), 2009.

[発表 9] 中島研吾: マルチコア時代の前処理付き反復法と性能チューニング, 日本応用数理学会 2009 年年会, 2009.

- [発表 10] 中島研吾:T2K オープンスパコン(東大)チューニング連載講座番外編 Hybrid 並列プログラミングモデルの評価(II), スーパーコンピューティングニュース(東京大学情報基盤センター) 11-5, 2009.
- [発表 11] Kengo Nakajima:Flat MPI vs. Hybrid: Evaluation of Parallel Programming Models for Preconditioned Iterative Solvers on Multicore Clusters, The Fourth International Workshop on Automatic Performance Tuning (iWAPT2009), 2009
- [発表 12] 中島研吾:講習会:科学技術計算のためのマルチコアプログラミング入門, スーパーコンピューティングニュース(東京大学情報基盤センター) 11-6, 2009.
- [発表 13] 中島研吾:OpenMP/MPI ハイブリッド並列プログラミングモデルの多重格子法への適用, 情報処理学会研究報告(HPC-124-7), 2010.
- [発表 14] Kengo Nakajima:Framework for Development of Parallel Codes in “Integrated Predictive Simulation System for Earthquake and Tsunami Disaster”, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), MS 30: Coupling and Regridding Tools for Supporting Parallel Multi-physics Modeling - Part I of II, 2010.
- [発表 15] Kengo Nakajima:Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), MS 55: Joint JSIAM-SIAM Minisymposium: Parallel Programming Models and Algorithms for Multicore Clusters and GPGPUs - Part II of III, 2010.
- [発表 16] Kengo Nakajima:Parallel Multigrid Solvers using OpenMP/MPI Hybrid Programming Models on Multi-Core/Multi-Socket Clusters, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), 2010.

## 特記事項

- [特記 1] パネリスト, 「PC クラスタから次世代スパコンへ -アカデミックと企業連携のあり方-」, 第 9 回 PC クラスタシンポジウム, 2009.
- [特記 2] パネリスト, 「シームレスな計算機環境を目指して」, PC クラスタワークショップ in 京都 2010, 2010.
- [特記 3] General Council Member, IACM (International Association for Computational Mechanics)
- [特記 4] 日本計算工学会 評議員
- [特記 5] 日本応用数理学会 評議員
- [特記 6] 日本応用数理学会 学会誌編集委員(幹事)
- [特記 7] 日本応用数理学会 「行列・固有値問題の解法とその応用」研究部会運営委員

- [特記 8] 日本機械学会計算力学技術者認定事業 計算力学技術者 1 級・2 級(固体力学分野の有限要素法解析技術者)認定試験 試験問題作成委員(能力開発促進機構 計算力学技術者認定事業委員会 固体力学分野 試験問題作成 WG 委員)(2005.4~)
- [特記 9] 情報処理学会 論文誌(コンピューティングシステム)編集委員
- [特記 10] Associate Editor, SIAM Journal for Scientific Computing
- [特記 11] Member of Organizing Committee, 15th International Conference in Finite Elements in Flow Problems (FEF 2009), 2009.
- [特記 12] Member of Program Committee, 2009 IEEE International Workshop on HPC and Grid Applications (IWHGA2009), 2009.
- [特記 13] 日本計算工学会第 14 回計算工学講演会 実行委員, 2009.
- [特記 14] Member of Program Committee, Parallel Computing 2009 (ParCo 2009), 2009.
- [特記 15] Member of Technical Papers Committee, 2009 IEEE International conference on high performance computing, networking, storage, and analysis (SC09), 2009.
- [特記 16] ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2010, プログラム委員, 2010.
- [特記 17] Member of International Program Committee, IASTED International Conference on Parallel and Distributed Computing and Networks (PDCN 2010), 2010.
- [特記 18] Member of Organizing Committee, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), 2010.
- [特記 19] Co-organizer of Mini-Symposium, Parallel and Hierarchical Algorithms for Accelerating System Matrix Solvers, 15th International Conference in Finite Elements in Flow Problems (FEF 2009), 2009.
- [特記 20] セッションオーガナイザ, 日本計算工学会第 14 回計算工学講演会「大規模シミュレーションと並列前処理手法」, 2009.
- [特記 21] セッションオーガナイザ, 第 23 回数値流体力学シンポジウム「並列計算, グリッド計算」, 2008.
- [特記 22] Co-organizer of Mini-Symposium, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), Coupling and Re gridding Tools for Supporting Parallel Multi-physics Modeling Part I-II, 2010.
- [特記 23] Co-organizer of Mini-Symposium, SIAM 14th Conference on Parallel Processing for Scientific Computing (PP10), Joint JSIAM-SIAM Minisymposium: Parallel Programming Models and Algorithms for Multicore Clusters and GPGPUs Part I-III, 2010.

[特記 24] Co-organizer of Workshop, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), Innovative algorithms for multicore and hierarchical computational environments, 2010.

[特記 25] 学際計算科学・工学人材育成プログラム, <http://nkl.cc.u-tokyo.ac.jp/CSEedu/>

[特記 26] 並列計算プログラミング, 先端計算機演習, 東京大学大学院理学系研究科地球惑星科学専攻, <http://nkl.cc.u-tokyo.ac.jp/09e/>

[特記 27] コンピュータ科学特別講義 I・II, 東京大学大学院情報理工学系研究科コンピュータ科学専攻, <http://nkl.cc.u-tokyo.ac.jp/09s/>, <http://nkl.cc.u-tokyo.ac.jp/09w/>

# スーパーコンピューターに基づく 大規模数値計算に関する研究

金田 康正

## 1 概要

本研究報告は 2009 年度中に行った、5 種類の研究テーマ（それぞれの研究テーマは主に研究室の大学院博士学生の研究テーマ。）に関する研究概要報告である。本報告では、頁数や研究内容の理解難易度を考えて、以下にある 4 種類の項目について研究全体の背景を述べた後、それぞれの研究テーマの説明を以下に行う。

- “double-double”型データタイプ平方根の高速アルゴリズム
- 素数の数え上げの高速化
- pi.super-computing.org 上での  $\pi$  並びに関連する情報公開
- 並列実行性能の高い数値計算ライブラリー及びプログラム開発環境に関する研究開発

## 2 背景

毎年の報告で繰り返し記述している事であるが、本センタースーパーコンピューティング研究部門教員の研究における主たる役割として、「利用者サービスに提供している現有のスーパーコンピューターの安定運用並びに、スーパーコンピューターの可用性向上の為の各種問題点の摘出」、また「利用者が将来要求する高性能計算需要をスーパーコンピューターの将来動向を見据えた上で安定的に導入・提供する事柄に関する情報収集や研究活動」等がある。

また教育における役割として、「工学系研究科電気系工学専攻並びに新領域創成科学研究科基盤情報学専攻の情報基盤センターに割り当てられている学生定員に対する兼任教員としての、同専攻修士・博士学生の教育指導」も、前者に劣らず主要な役割となっている。ちなみに 2008 年度は博士 1 年 1 名、修士 2 年 0.5 名、修士 1 年 0.5 名、計 3 名の教育指導を行った。（なお 0.5 名は 10 月入学の修士学生を意味する。）これまでに理学系研究科情報科学専攻の兼担として修士・博士課程の学生指導を引き受ける事はあったが、毎年一定数の学生教育を引き受けるという状況にはなっていなかった。学生定員が文部科学省定員として本センタースーパーコンピューティング研究部門とネットワーク研究部門に措置されてからは、ほぼ毎年一名の学生が進学してくる様になり、中・長期的な研究テーマに取り組む事が可能となった。なお情報基盤センターに対する学生定員が情報基盤センターになって初めて実現したという事、また 2008 年度からは兼担先が、新領域創成科学研究科から工学系研究科へ移っている事を記録としてここに記しておく。

私の研究室の従来からの指導方針として、研究テーマは各人の興味に基づき選択させている。また研究効率を考えると、競争的研究資金を獲得せずとも基本的に一人で実施できる（研究費のかからない）研究テーマを選ぶとともに、たとえ時間がかかったとしても誰かが実際に使う（=役に立つ）成果物を得る事を心がけている。2005 年度以降の年報はそれまで発行していたセンター全国共同利用部門が発

行している年報の記述内容と変更になったが、それまでの年報の記載内容もそれぞれの教員の研究・教育活動報告であった事を考え、報告内容の継続性が重要であるという事を考慮しつつ、これまでとは記載する研究テーマの順序を変更し、まずは研究指導の一環で研究室学生が得た成果を記載し、その次に学生が関係しない研究テーマの研究成果を記載するという流れで、2009年度の研究・教育活動報告を以下に記す。

### 3 研究内容

#### 3.1 “double-double”型データタイプ平方根の高速アルゴリズム

本研究は本研究室の博士1年の筒井直機君が行った研究である。彼がこれまでに得た成果の一部は、修士論文[文献1]で発表済みである。この研究はその後の研究であり、概要を以下に示す事にする。

##### 3.1.1 はじめに

並列計算機の性能向上や数値計算法の進展は大規模科学技術計算の大きな鍵となっている。浮動小数点演算において計算規模が増すに従って多くの計算を必要とし、計算誤差が増大する。その解決策の1つとして倍精度数より有効桁数の多い倍精度を2つ用いて表される“double-double”精度のデータ型の4倍精度数がある。しかし4倍精度演算はソフトウェアによって実現されているため、倍精度演算に比べ多くの計算時間を必要とする。当研究室ではこれまで4倍精度四則演算に関する高速化について研究[文献2, 文献3]を行ってきたが、4倍精度数学関数の高速化を目指し、2009年度は4倍精度平方根関数の高速化を試みた。

##### 3.1.2 “double-double”精度4倍精度数

本研究で使用している“double-double”精度4倍精度数のデータ型は128bit長であるが、IEEE754-2008[文献4]で定義されている単精度データ型や倍精度データ型の規格のように固有のデータ型が用意されているわけではなく、2つの倍精度データ型を使って表現している。具体的には仮数部の桁数を増やすことで精度を上げていることから、指数部が倍精度数と同じために絶対値の大きい数は表現できない。しかしこの方法を採用することで4倍精度用の特別命令が不要で、比較的高速かつ高精度に計算ができるようになっている。

4倍精度数 $x$ がある時、上位桁、下位桁で構成されている4倍精度数を $(x_H, x_L)$ のように表す。浮動小数点の仮数部（浮動少数が0あるいはNot a Number以外であれば、必ず絶対値が1以上、2未満とする）と指数部を $\text{Sig}()$ 、 $\text{Exp}()$ という関数で表現できる時、 $x = \text{Sig}(x) \times \text{Exp}(x)$ と表現することができる。また、 $x = (x_H, x_L) = \text{Sig}(x_H) \times \text{Exp}(x_H) + \text{Sig}(x_L) \times \text{Exp}(x_L)$ とする。なお倍精度数は仮数部53ビット分の数字を表現可能であるが、“double-double”精度4倍精度数は上位部分と下位部分の符号を別にするので、倍精度の倍の106ビット分ではなく、107ビット分の数字を表現できるということに注意しておく。そのため、通常“double-double”型4倍精度数は $\text{Exp}(x_H) - \text{Exp}(x_L)$ は54以上となる。ただしgccが実現しているdouble-double演算アルゴリズムは54未満になることがある。この“double-double”精度の浮動小数点数の上位数と下位数の間にある見えざるビットを本稿ではインビジブルビットと定義する。

##### 3.1.3 誤差を考慮した浮動小数点演算について

通常、演算には誤差が発生するが、今回使用する誤差を考慮した浮動小数点演算アルゴリズムは図2の通りである。例えば浮動小数点同士の加算 $a + b$ は、 $a + b = \text{fl}(a + b) + \text{err}(a + b)$ と表せる。ここで $\text{fl}(a + b)$ は $a + b$ の浮動小数点演算、 $\text{err}(a + b)$ は $a + b$ の演算で生じる誤差を示す。

$\text{Quick-TwoSum}(a, b)$ は、Dekker[文献5]の加算アルゴリズムで $s = \text{fl}(a + b)$ 、 $e = \text{err}(a + b)$ を示す。ただし引数 $a$ と $b$ に制約がある。その制約とは図1のように浮動小数点数 $a$ と $b$ が与えられた時、“ $a$



の仮数部の1を示す最後のビットが示す桁が“ $b$ の仮数部の最後のビットが示す桁”以上になることである。また  $\text{TwoProd-FMA}(a,b)$  は、 $p = \text{fl}(a \times b)$ 、 $e = \text{err}(a \times b)$  を示す。

### 3.1.4 通常平方根アルゴリズム

平方根アルゴリズムとして Newton-Raphson 法などの反復法を使う。Newton-Raphson 法は与えられた関数  $f(x)$  が単調連続で変曲点がなく、且つ導関数  $f'(x)$  が求められる時に利用できる数値解法で、二次の収束を示す。今曲線上の点  $(x_n, f(x_n))$  での接線と、 $x$  軸との交点を  $(x_{n+1}, 0)$  とすると、

$$f'(x_n) = -\frac{f(x_n)}{x_{n+1} - x_n} \tag{1}$$

なので、下位の適当な推定値から初めて反復公式

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n = 0, 1, 2 \dots) \tag{2}$$

を繰り返すと漸近的に解に達する。数字  $a$  の平方根を求める場合、

$$f(x) = x^2 - a \tag{3}$$

のように定義されるため、

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} \quad (n = 0, 1, 2 \dots) \tag{4}$$

$$= \frac{1}{2} \left( x_n + \frac{a}{x_n} \right) \tag{5}$$

となる。また4倍精度四則演算の命令数最小アルゴリズム [文献2, 文献3] の各命令の個数は、表1の通りで、4倍精度平方根アルゴリズムは Newton-Raphson 法では初期値を上位桁の倍精度平方根とした時、最低でも2回更新をする必要とするが、4倍精度の加算2個、乗算2個、除算2個となるため、倍精度浮動小数点の加減乗算と FMA を54個、時間が掛かる除算を4個の命令を含むことになる。

表 1: Floating Point Operations of Four Quadruple Operation Short Algorithms.

	±	×	÷	FMA
Quadruple Addition/Subtraction	11	0	0	0
Quadruple Multiplication	1	1	0	3
Quadruple Division	7	1	2	2

### 3.1.5 高速4倍精度平方根アルゴリズム

4倍精度平方根は  $sq = \text{fl}(\sqrt{a})$  を計算する。図3に示す3種類のアロリズムを作成した。ここで  $sq$ 、 $a$  はそれぞれ倍精度2つで構成された4倍精度数であり、 $(a^H, a^L)$  は  $a = a^H + a^L$  を表している。

これらアルゴリズムは Newton-Raphson 法を基に改良を重ねたものである。このセクションでは各アルゴリズムについて述べる。

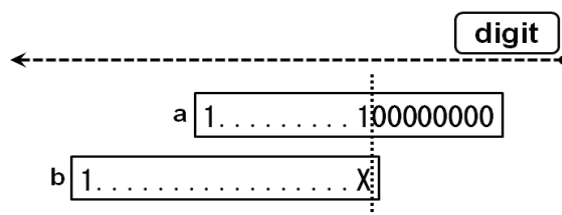


図 1: Quick Two Sum Algorithm Range.

```

誤差を考慮した演算アルゴリズム
: Quick-TwoSum(a, b) {
01 : s ← a ⊕ b
02 : e ← (a ⊖ s) ⊕ b
03 : return (s, e)
: }
: TwoSum(a, b) {
01 : s ← a ⊕ b
02 : v ← s ⊖ a
03 : e ← (a ⊖ (s ⊖ v)) ⊕ (b ⊖ v)
04 : return (s, e)
: }
: TwoProd-FMA(a, b) {
01 : p ← a ⊗ b
02 : e ← fl(s × a - p)
03 : return (p, e)
: }

```

図 2: Algorithms with error detection.

● 高速平方根アルゴリズム A

図3からわかるように、提案する全ての高速平方根アルゴリズムの初期値  $x_0$  を  $\text{fl}(\sqrt{a^H})$  としているため、式の上では  $x_0 = \text{fl}(\sqrt{a^H})$  となり、Newton-Raphson 法の解の更新式 (4) を変形させると、

$$x_1 = x_0 + \frac{a - x_0^2}{2x_0} \tag{6}$$

$$= x_0 + \frac{a^H + a^L - x_0^2}{2x_0} \tag{7}$$

$$= x_0 + \frac{a^L - (x_0^2 - a^H)}{2x_0} \tag{8}$$

となる。そのため、行 04 までに

$$sq^H \leftarrow \text{fl}(\sqrt{a^H}) \tag{9}$$

$$sq^L \leftarrow \frac{a^L \ominus \text{fl}(sq^{H^2} - a^H)}{2 \otimes sq^H} \tag{10}$$

から  $sq^H$  と  $sq^L$  が求まり、行 05 にて Quick-TwoSum を使うことで正規化を行っている。

● 高速平方根アルゴリズム B

アルゴリズム B はアルゴリズム A を使って  $sq^H$  を求めてから、もう一度アルゴリズム A を適用することによって、アルゴリズム A より精度を高めている。

● 高速平方根アルゴリズム C

アルゴリズム C はアルゴリズム A を使って一度解  $sq$  を求め、アルゴリズム A の  $sq^L$  の足りない精度を修正するアルゴリズムである。ただし更新式 (4) の  $2x_n$  の乗算と除算は繰り返し使うため、 $\frac{0.5}{x_n}$  を乗算の形で再利用することで高速化を図っている。また  $sq^L$  の誤差を修正する前と後では  $sq^H$  の値に変更はない点を考慮することで、全体の演算量を削減している。改変されたアルゴリ

高速平方根アルゴリズム A

```

: Fast-Square-Root-A(a) {
01 :  $sq_0^H \leftarrow \text{fl}(\sqrt{a^H})$ 
02 :  $t \leftarrow \text{fl}(sq_0^H \times sq_0^H - a^H)$ 
03 :  $d \leftarrow a^L \ominus t$ 
04 :  $d \leftarrow d \oslash (2.0 \otimes sq_0^H)$ 
05 :  $(sq^H, sq^L) \leftarrow \text{Quick-TwoSum}(sq_0^H, d)$ 
06 : return  $(sq^H, sq^L)$ 
: }

```

高速平方根アルゴリズム B

```

: Fast-Square-Root-B(a) {
01 :  $sq_0^H \leftarrow \text{fl}(\sqrt{a^H})$ 
02 :  $t \leftarrow \text{fl}(sq_0^H \times sq_0^H - a^H)$ 
03 :  $d \leftarrow a^L \ominus t$ 
04 :  $d \leftarrow d \oslash (2.0 \otimes sq_0^H)$ 
05 :  $sq^H \leftarrow sq_0^H \oplus d$ 
06 :  $(t_0, t_1) \leftarrow \text{TwoProd-FMA}(sq^H, sq^H)$ 
07 :  $d \leftarrow ((a^H \ominus t_0) \ominus t_1) \oplus a^L$ 
08 :  $sq^L \leftarrow d \oslash (2.0 \otimes sq^H)$ 
09 : return  $(sq^H, sq^L)$ 
: }

```

高速平方根アルゴリズム C

```

: Fast-Square-Root-C(a) {
01 :  $sq_0^H \leftarrow \text{fl}(\sqrt{a^H})$ 
02 :  $c \leftarrow 0.5 \oslash sq_0^H$ 
03 :  $t \leftarrow \text{fl}(sq_0^H \times sq_0^H - a^H)$ 
04 :  $d \leftarrow a^L \ominus t$ 
05 :  $d \leftarrow c \otimes d$ 
06 :  $(sq_1^H, sq_1^L) \leftarrow \text{Quick-TwoSum}(sq_0^H, d)$ 
07 :  $sq^H \leftarrow sq_1^H$ 
08 :  $(t_0, t_1) \leftarrow \text{TwoProd-FMA}(sq_1^H, sq_1^H)$ 
09 :  $(t_2, t_3) \leftarrow \text{TwoProd-FMA}(sq_1^H, sq_1^L)$ 
10 :  $t_2 \leftarrow 2.0 \otimes t_2$ 
11 :  $t_3 \leftarrow 2.0 \otimes t_3$ 
12 :  $t_0 \leftarrow t_0 \ominus a^H$ 
13 :  $t_3 \leftarrow \text{fl}(sq_1^L \times sq_1^L + t_3)$ 
14 :  $t_0 \leftarrow t_0 \ominus a^H$ 
15 :  $(t_0, t_1) \leftarrow \text{Quick-TwoSum}(t_0, t_1)$ 
16 :  $(t_0, t_2) \leftarrow \text{Quick-TwoSum}(t_0, t_2)$ 
17 :  $t_0 \leftarrow t_0 \oplus (t_1 \oplus (t_2 \oplus t_3))$ 
18 :  $d \leftarrow c \otimes t_0$ 
19 :  $sq^L \leftarrow sq_1^L \ominus d$ 
20 : return  $(sq^H, sq^L)$ 
: }

```

図 3: Fast Square Root Algorithm

ズム A で 4 倍精度数  $sq_1$  を計算している。下位桁に誤差を含む 4 倍精度数  $x_n$  を 4 倍精度数  $a$  の平方根の近似値とすると、更新式 (4) の右辺の  $\text{fl}(x_n^2 - a)$  は浮動小数点で表現した場合、図 4 のようになるので演算しやすい形に変形すると次のようになる。

$$\text{fl}(x_n^2 - a) \tag{11}$$

$$= \text{fl}(x_n^{H^2} + 2x_n^H x_n^L + x_n^{L^2} - a^H - a^L) \tag{12}$$

$$\begin{aligned} &= \text{fl}(\text{fl}(x_n^{H^2}) + \text{err}(x_n^{H^2}) + 2 \cdot \text{fl}(x_n^H x_n^L) \\ &\quad + 2 \cdot \text{err}(x_n^H x_n^L) + \text{fl}(x_n^{L^2}) + \text{err}(x_n^{L^2}) \\ &\quad - a^H - a^L) \end{aligned} \tag{13}$$

$$\begin{aligned} &= \text{fl}(\text{fl}(x_n^{H^2}) + \text{err}(x_n^{H^2}) + 2 \cdot \text{fl}(x_n^H x_n^L) + 2 \cdot \text{err}(x_n^H x_n^L) \\ &\quad + \text{fl}(x_n^{L^2}) - a^H - a^L) \end{aligned} \tag{14}$$

このように考えることによって、アルゴリズム C の行 08~17 で  $\text{fl}(x_n^2 - a)$  を演算している。

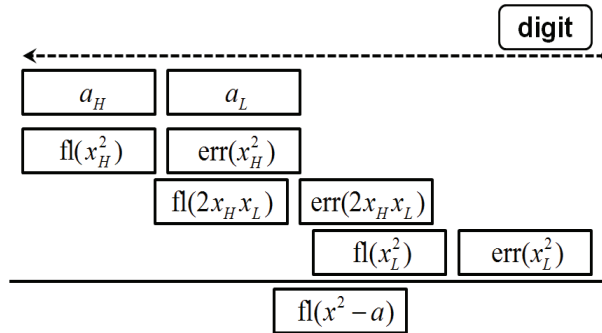


図 4: Difference between  $a$  and  $x^2$ .

また行 06 で出た誤差を含む解  $sq_1$  において最大 4 ビットの誤差がある。詳しくは表 2 に示す。

表 2: Error bit Number of Quadruple Square Root Algorithms 2.

Error bit Number	normalized gcc	$sq_1$ in Alogrithm C
0	9999895	6291698
1	100	3427868
2	4	272614
3	0	7688
4	0	132
5	1	0

### 3.1.6 アルゴリズムにおける理論上の誤差について

ここでは提案したアルゴリズムが理論上図 3 のどこでどれほどの誤差が発生したかを示す。入力  $a$  は 4 倍精度数  $a = (a^H, a^L)$  とし、 $Exp(a^H) - Exp(a^L) = E_d$  とする。3.1.2 節で示したとおり、正規化された 4 倍精度数であれば  $E_d$  は 54 以上である。また 4 倍精度で表現可能な最も精度のよい平方の数を  $sq_t$  とする。

- アルゴリズム A

行 01 で計算された  $sq_0^H \leftarrow \text{fl}(\sqrt{a^H})$  は  $sq_t^H$  と最大 1 ビットの誤差が発生する。 $sq_0^H$  に誤差が無

い場合、行 04 で計算される  $d$  と行 05 で計算される解の下位  $sq^L$  と一致する。また仮数部が無限ある浮動小数点数を基準に考えた時、インビジブルビットの影響が  $sq_0^H$  の仮数部の 53 ビット目に存在するため、誤差が 1 ビット存在しているように見える。そのため行 04 における浮動小数点除算の誤差 1 ビットも発生するため、 $sq^L$  は、 $sq_t^L$  と比べ最大 2 ビットの誤差が発生する。逆に  $sq_0^H$  に誤差がある場合、 $sq_0^H$  の仮数部の 1~52 ビット目の値が正しいことを指名しているため、行 04 で計算した  $d$  の仮数部の先頭ビットは  $sq_t^H$  の 53 ビット目か 54 ビット目の桁に該当する。(53 ビット目の桁に対する繰り上がり処理が発生しないため、55 ビット目の桁に一致することはない。) ただし 53 ビット目の桁と一致する場合は仮数部は  $d$  の仮数部の先頭ビット以外はすべて 0 でなければならず、確率上ほとんど考えなくてよい。54 ビット目の桁と一致する場合は行 05 で解の上位数の 53 ビット目に繰り上がり処理を行い、解の下位数の先頭ビットを 55 ビット目以下になるように正規化を行う。そのため、アルゴリズム A は上から 52+51 の 103 ビット範囲の精度しか保証できないことがわかる。

- アルゴリズム B

アルゴリズム B はアルゴリズム A の 1 度目の誤差修正の時点で解の上位部分と下位部分に誤差修正演算がまたがる部分を改良したアルゴリズムである。まず行 05 で解の上位桁を厳密に求める。ただしこの時仮数部 103 ビット分の精度が保証された浮動小数点数を倍精度浮動小数点に落とし込んでいるため、インビジブルビットの部分まで浮動小数点数の精度が保証されている。つまりこの時解の上位の浮動小数点数の仮数部上位 53 ビット目までが精度保証されているというわけではなく、解の下位の浮動小数点数の仮数部の 1 ビット目より前の部分で精度保証されていることとなる。そのため行 08 の浮動小数点除算演算の除数の 1 ビット誤差と浮動小数点除算の誤差 1 ビットより、最悪  $sq$  は下位 2 ビット分に誤差が発生することがわかる。

- アルゴリズム C

アルゴリズム C はアルゴリズム A の行 04 やアルゴリズム B の行 04 と行 08 の除算部分にあたる部分を乗算に変換し再利用する形で浮動小数点除算の回数を削減し、アルゴリズム A をベースに精度を向上させたアルゴリズムである。そのため行 01 で 52 ビット、行 05 で 51 ビット、行 18 で 51 ビット保証されているため、アルゴリズム C で算出された  $sq$  は 52 + 51 + 51 の合計 154 ビット範囲の精度が保証されるアルゴリズムとなる。具体的にはまず最初に行 01 で 52 ビット保証できる。2 番目で行 02 の除算で下位 1 ビット分の誤差、行 05 の乗算で下位 1 ビット分の誤差が発生し、下位 2 ビット分の誤差発生するため、行 05 の演算結果  $d$  は 53-2 の 51 ビット保証されることになる。3 番目では 2 番目と同様、行 02 の除算で下位 1 ビット分の誤差、行 18 の乗算で下位 1 ビット分の誤差が発生し、下位 2 ビット分の誤差発生するため、行 05 の演算結果  $d$  は 53-2 の 51 ビット保証されることになる。そのため理論的に精度が保証される範囲は仮数部 154 ビット範囲であることが分かる。

また計算機実験ではアルゴリズム C は誤差が発生しなかったが、これらの誤差解析により“double-double”精度で表現できる最大精度の解が演算できるアルゴリズムというわけではなく、場合によっては誤差が発生しうることが理解できる。具体的には“double-double”精度の特性により最大 2 ビットの誤差が発生する可能性がある。解の上位桁は行 06 で誤差なく演算することができる。その時の解の下位桁  $sq_t^L$  が 0.0 になる可能性がある。従って行 19 で解の下位桁  $sq^L$  を演算した時、アルゴリズム B の時と同じように解の下位桁に 2 ビットの誤差が発生することがある。

アルゴリズム C より精度を保証したい場合はアルゴリズム B のように解の上位数を Newton-Raphson 法を 1 回更新し、厳密に計算した後に解の下位数を Newton-Raphson 法を 2 回更新し、厳密に計算することで“double-double”精度で表現できる最大精度の解が演算できる。

### 3.1.7 まとめ

今回“double-double”型データタイプの平方根の高速アルゴリズムを3つ設計し、実験を行うとともに理論的にその解の精度を解析した。アルゴリズム A と B の共に演算数は少ないが、gcc より精度を確保できなかった。ところがアルゴリズム C は“double-double”精度で表現できる最大精度の解を算出することはできないが、その精度は計算機実験上 gcc よりはるかに精度が良く、 $10^7$  回の実験で全て正しい解を算出することができたことに加え、理論的な解析でさらに精度を高めたい場合の対策方法が判明した。また gcc で採用されているアルゴリズムでは、今回の計算機実験では最大 5 ビットの誤差が発生し、正規化処理がされていないことがわかった。今後の展開としてはアルゴリズム C の問題点を改善したアルゴリズムの作成と計算機実験。さらに並列性を考慮したアルゴリズムの改良が考えられる。

## 3.2 素数の数え上げの高速化

本研究は本研究室の博士 3 年の吉田仁君が博士課程の研究テーマとして行っている研究である。以下にその概要を示す。

### 3.2.1 はじめに

1 とそれ自身しか因数を持たない素数は整数環を構成する素元であり、暗号学や数論において重要な役割を示す数である。素数それ自体が現在でも Goldbach 予想や双子素数問題を始めたとした様々な未解決問題を持つ研究対象である。またこれらの未解決問題への 1 つの解答方法として、具体的な素数を用いた数値実験プログラムによる検証も行われており、規模の大きな実験を効率的に行うため、高速に素数リストの生成を行う必要がある。

素数列を生成する高速なアルゴリズムとしては Eratosthenes の篩が広く知られているものの、演算量の削減や、キャッシュメモリの利用、といったアルゴリズム以上の高速化について言及されることは殆ど無い。そこで本報告では Eratosthenes の篩を用いた素数のリストアップ処理の高速化について述べる。具体的には、キャッシュメモリーを意識した分割領域で篩を行う、ループアンローリングを行うといった一般的な高速化処理の他、自明な合成数を処理から外す、篩う素数による処理の変更といった Eratosthenes の篩ならではの高速化を行った。

最後に、Eratosthenes ではなく 2 次形式を用いた素数の高速な列挙方法として知られる primegen[文献 7]に含まれるプログラム primespeed と、本論文で述べた高速化を行ったプログラムとを用い、 $3 \times 10^{11}$  の幅に含まれる素数を求めるのにかかる時間を計測した結果、低い領域で 1.8 倍、高い領域では 58 倍以上も高速になり、本手法の高速性が示された。

### 3.2.2 Eratosthenes の篩

- 概要

Eratosthenes[文献 8] は自然数の集合  $\mathbb{N}$  から合成数を除き、素数のみを抽出するアルゴリズムである。

1. 1 を除く。
2. 残っている最小の数  $p$  は素数である。
3.  $p$  の倍数を除く。
4. 2. に戻る。

という手順で行われる。

- 実装

2、3、5 の倍数を取り除くことで 30 あたり 8 個の数を判断するフラグで十分となる。具体的には

$$\{a_i\}_{0 \leq i \leq 7} = \{1, 7, 11, 13, 17, 19, 23, 29\} \quad (15)$$



とした時、 $30k + a_i$  で表される数だけで十分である。そのため、この 8 フラグを Byte 内の各 bit に対応させることでメモリーを効率的に使うことができる。また、 $p, q \in \{a_i\} \pmod{30}$  と  $pq \in \{a_i\} \pmod{30}$  は同値となるため、 $\{a_i\} \pmod{30}$  は乗算で閉じている。

篩を行うにあたり、素数リストを上で述べたような形で保持しているとすると、与えられる情報は  $p$  や  $q$  自体ではなくそれらの bit インデックスである。ここで篩い落とす合成数を計算する素直な方法として、

1.  $p, q$  を求める。
2.  $pq$  を計算する。
3.  $pq$  に対応する bit を求める。

という手順を踏むと非常に計算量が大きい。そのため、 $p, q$  の bit 位置から直接  $pq$  の bit 位置を高速に求めるために表を用いて計算を行う。 $n$  以下の数を素数  $p = 30i + a_j$  で篩うとき、表 3 と表 4 の値を用いることで図 6 のようなアルゴリズムを構成できる。ループの初期値は  $p^2$  になるようにしている。

表 3:  $A_j$  and  $B_{i,j}$ , where  $\Delta k = A_j k + B_{i,j}$

	$i \setminus j$	0	1	2	3	4	5	6	7
$A_j$		6	4	2	4	2	4	6	2
$B_{i,j}$	0	0	0	0	0	0	0	0	1
	1	1	1	1	0	1	1	1	1
	2	2	2	0	2	0	2	2	1
	3	3	1	1	2	1	1	3	1
	4	4	3	1	2	1	3	3	1
	5	5	2	2	2	2	2	4	1
	6	6	3	1	4	1	3	5	1
	7	7	4	2	4	2	4	6	1

表 4: Indices of  $a_i a_j \pmod{30}$  as  $C_{i,j}$

$i \setminus j$	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	5	4	0	7	3	2	6
2	2	4	0	6	1	7	3	5
3	3	0	6	5	2	1	7	4
4	4	7	1	2	5	6	0	3
5	5	3	7	1	6	0	4	2
6	6	2	3	7	0	4	5	1
7	7	6	5	4	3	2	1	0

フラグ bit 列は  $\pi(\sqrt{n})$  個の素数で篩いをかけるため、同じ領域を何度も舐めることになる。これを高速化するためにメモリアクセス速度を高める必要がある。篩う全領域を適当な Window size で分割し、1つ1つの分割領域がキャッシュメモリに収まるようにすることでこれを実現することができる。正確にはこの Window size を Cache size より若干小さくすることで、篩にかける素数を読み込む領域などにより篩う領域が Cache out しないようにすることができる。

具体的なアルゴリズムは図5に示したようになる。この中では素数  $p$  での篩を行う際に計算量の削減を目的として Window 内の篩い開始位置を条件分けし、 $x < p^2$  の篩を避けているが、この分岐にかかる計算コストが篩う範囲の縮小により削減される計算量より小さくなる場合はこの分岐を無くし、**then** 節にある代入のみ行うと良い。

```

W := Window size
for B = 0, W, 2W, .., n/30 :
  flag[0..W - 1] = {11111111} in binary
  foreach primes ≤ √n as p := 30i + aj :
    k ← (30i + 2aj)i + ⌊aj2/30⌋
    if k < B then :
      (k, l) ← ((i - B) mod p, 1)
    else
      (k, l) ← (k - B, j)
    end of if
    while k < 0 :
      k ← k + Al × i + Bj,l
      l ← (l + 1) mod 8
    end of while
    while k < W :
      Set 0 in Cj,l-th bit of flag[k]
      k ← k + Al × i + Bj,l
      l ← (l + 1) mod 8
    end of while
  end of foreach
  CheckBits(flag)
end of for

```

図5: Algorithm of the sieve of Eratosthenes with area assigned

計算機資源を有効に使うために並列化を行う必要がある。Eratosthenes の篩では篩う領域を分割することで並列化することができる。図5のアルゴリズムを使うことで限られた領域のみを篩にかけることができるため、篩う領域を離散的に各プロセッサに担当させることができる。

また各プロセスに振り分けられるタスク量はできるだけ均等になるようにすると効率が良い。そのため計算する領域全体をブロック分割するのではなく、ブロック・サイクリック分割する。

### 3.2.3 高速化

- 初期ビット列の利用  
フラグを作るにあたって2、3、5の倍数は除いているが、その上で7、11、13などの小さな素数

```

k ← (30i + 2aj)i + ⌊aj2/30⌋
l ← j
while k ≤ n/30 :
    Set 0 in Cj,l-th bit of flag[k]
    k ← k + Al × i + Bj,l
    l ← (l + 1) mod 8
end of while
    
```

図 6: Algorithm of sieving with  $p$

で篩う作業もまた大量の計算量を要する。具体的には素数  $p$  での篩に要する計算量は  $O(n/p)$  となる。

また  $p_4 = 7$  から  $p_K$  までの素数で篩った結果のフラグ状態は  $\prod_{k=4}^K p_k$  毎に繰り返しになることから 1 周期分だけ具体的な篩を行い、それより後はこれをコピーすることでこの大きな計算量を削減することができる。

メモリーコピーに要する時間と  $p_K$  での篩に要する時間との違いや、1 周期分のデータを保持する領域 ( $\prod_{k=4}^K p_k$  Byte) から上限  $p_K$  を決める必要がある。例えば  $p_K = 23$  とする場合には初期ビット列の保存領域の容量は

$$\prod_{k=4}^7 p_k = 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 = 7,436,429 \text{ [Byte]}$$

であり、同様に  $p_K = 29$  とする場合は

$$\prod_{k=4}^8 p_k = 7,436,429 \times 29 = 215,656,441 \text{ [Byte]}$$

であることから、このいずれかが現実的に使いやすい数と考えられる。

- 篩ビットの検出

図 6 と図 5 にある **for each prime** では、簡単に考えると  $\sqrt{n}$  以下の素数を表すビットを全て確認する事になる。素数定理 [文献 8] から、数  $x$  付近に素数が現れる確率は  $(\ln x)^{-1}$  である。本実験で対象とする  $n = 3 \times 10^{16}$  までの素数リストを作り上げるためには篩の中で  $\sqrt{n} \simeq 1.73 \times 10^8$  までの、約  $2\sqrt{n}/\ln n$  個の素数を用いる必要がある。このとき、 $\ln n \simeq 37.9$  から、19 個に 1 個程度しか素数が現れないということになる。そのため、全ビットのフラグ状態を確認するアルゴリズムでは条件分岐の約 80% が無駄に終わることになる。

そこで、立っているビットのみを検出するアルゴリズム [文献 9] (図 7) を利用することでその計算コストを下げる事ができる。なお図 7 内での演算子 **and**、**or**、**xor** はビット演算、 $\sim(B)$  はビット反転演算子、**popCount**( $b$ ) はビット数を数える関数である。

- ループアンローリング

領域分割した Eratosthenes の篩では、各素数  $p$  について 1 つの Window 内で  $W/p$  回ループを回すことになる。  $p$  が小さな素数である場合、3.2.3 節で述べたように処理に時間がかかることになるが、あまり多くの素数を使って初期ビット列を作成するとビット列の格納に必要な容量が指数関数的に大きくなってしまふ。そこで、次点の策として篩それ自体を高速に行う。篩処理のループ

```

Get bits of B from LSB:
while B ≠ 0:
    n := popCount((~B) and (B - 1))
    process using n
    B ← B and (B - 1)
end of while

Get bits of B from MSB:
while B ≠ 0:
    b ← B
    for k in {1, 2, 4, 8, 16, 32} :
        b ← b or (b >> k)
    end of for
    n := 63 - popCount(b)
    process using n
    B ← B xor (1 << n)
end of while

```

図 7: Algorithm of get only '1' bits in 64 bit unsigned integer

では 30pByte 毎に同じ位置の bit を篩い落とすことになるので、8 段階のループアンローリングを行うことで計算のパイプライン化し、同時にループ全体が規則的な処理となる。

一方で、アンローリングをしていない篩処理では、ループ 1 回で平均約  $\frac{8}{30}$ pByte 分進むことになる。つまり  $p \simeq WindowSize$  ではループ内を高々 8 回、それ以上の素数ではより少ない回数しか実行する必要がない。そのため、ある程度より大きな素数での篩を行う場合はループ展開を行わず、元の形のまま処理する方が高速になることが期待される。

- 篩位置の保存

3.2.3 節において、アンローリングをしない場合、各素数での篩い位置を求めるために除算と、処理を行わないループを回す必要がある。また、処理ループにおいて担当範囲を超えループを出た状態は本来、次の担当範囲での篩い開始位置を指すはずである。そのことから、この位置を保持しておくことで次の領域の処理において初期位置の計算に必要なコストが削減することができる。

この保持領域については、 $\pi(\sqrt{10^{17}}) = 17082663$  個の値を保持する必要があり、1 個のデータを保存するために 8Byte 使うことで 130MB の領域が必要になる。

この篩い位置の保存については、篩う Byte 位置を示す値  $k$  と、bit 位置を示す値  $l$  の 2 つの値を保持する必要がある。1Byte 当たり 30 の数を示すこと、1Byte 内に 8bit しかないことから、各値の範囲は

$$\{k \leq \lceil 10^{17}/30 \rceil = 33333333333333340 \leq l < 8 \quad (16)$$

であり、その格納に必要な bit 数はそれぞれ 52 bit と 3 bit である。この 2 つの値を組み合わせ

$$store(k, l) = 8k + l \quad (17)$$

という形で保持することでプログラムで用いやすい型を保持しつつ格納に必要な容量を 8Byte に、つまり全体で 130MB に納めることができる。

### 3.2.4 実験

実験環境となるスーパーコンピューター SR16000 のスペックは表 5 に示した通りである。この 1 ノード内で実験を行ったため、通信環境については記載していない。

表 5: SR16000 specs

CPU	16 CPUs / node CPU Spec POWER6 3.5GHz (Dual Core) L2 Cache: 4MB / core
Memory	64 GiB / node

- 各手法の効果

まず各高速化処理を行うことでどの程度速くなるかを示す。比較対象として 3.2.2 節に記した篩領域の分割と並列化とを行った Eratosthenes の篩を実装したプログラムを用いている。

このプログラムと、それに対し各高速化方法を実装したプログラム全 16 種を用意し、そのそれぞれについて

$$1 < x < 3 \times 10^{11}, 3 \times 10^{16} < x < 3 \times 10^{16} + 3 \times 10^{11}$$

の範囲において  $3 \times 10^9$  ずつの幅 100 箇所にある素数を洗い出すために要した時間を計測し、その平均値を見る。6 回ずつ実験を行ったので全体としては 600 回計測の平均値となる。表 6 にその結果を示す。表内の A~D のアルファベットは 3.2.3 節内各節の高速化手法を行ったことを示している。

表 6: Time to sieve  $3 \times 10^{12}$  numbers [sec]

Change	Low	High
	12.49 ± 3.64	30.26 ± 4.58
A	11.99 ± 1.66	28.07 ± 0.21
B	12.35 ± 3.57	30.11 ± 4.37
A B	12.51 ± 1.97	27.97 ± 0.35
C	12.55 ± 3.69	30.29 ± 4.58
A C	13.42 ± 1.99	28.13 ± 0.27
B C	12.30 ± 3.63	17.59 ± 3.94
A B C	12.64 ± 2.03	16.07 ± 0.39
D	12.33 ± 1.94	37.98 ± 0.82
A D	13.04 ± 2.24	37.20 ± 0.55
B D	11.87 ± 2.27	27.56 ± 1.84
A B D	12.61 ± 2.16	26.51 ± 0.30
C D	11.86 ± 2.04	38.00 ± 0.84
A C D	12.72 ± 2.18	37.14 ± 0.42
B C D	12.17 ± 2.21	15.67 ± 1.86
A B C D	12.81 ± 2.28	14.76 ± 0.22

結果として、低い領域の篩では各高速化手法での違いは誤差程度であるが、高い領域の篩では各手法による高速化が明らかである。また、手法 D は単体で適用させると計算速度が遅くなるが他の手法と組み合わせることで高速になることが分かる。最終的に全高速化手法を適用させたものは、領域分割したものに比較して計算時間が 48.8% に短縮されている。

表 7: Time to sieve  $3 \times 10^{11}$  numbers [sec]

Change	Low	High
my program	1281	1476
primegen	2328	> 86400

- primegen との比較

本手法により実装された Eratosthenes の篩と、高速な素数列挙アルゴリズムとして知られる primegen[文献 7] とを比較した結果を表 7 に示した。primegen は並列実装されていないこと、低い領域の篩は分割担当領域により計算時間に差がでることから、全領域にあたる  $3 \times 10^{11}$  の幅を篩う時間を計測、比較した。なお、primegen による  $[3 \times 10^{16}, 3 \times 10^{16} + 3 \times 10^{11}]$  の篩では実行時間制限の 24 時間を超えたため正確な計測はできなかった。

測定結果を比較すると、時間測定ができた低い領域の篩でも 1.8 倍、測定できていない高い領域は 58 倍以上高速であることが分かる。

### 3.2.5 まとめ

素数のリストアップを行うにあたり、Eratosthenes の篩を実装した。実装手段にさまざまな工夫を行った結果、 $[3 \times 10^{16}, 3 \times 10^{16} + 3 \times 10^{11}]$  の篩を行う速度が 2.1 倍高速になった。

また、高速な素数列挙プログラムとして知られる primegen と比較したところ、低い領域でも 1.8 倍、高い領域では 58 倍以上高速に素数列挙できることが示された。

今後の研究としては primegen についても並列化や高速化を行い、再度比較する。primegen で用いられているアルゴリズムは特殊なため本手法で述べたものと同様の高速化手法は適用しにくいいため、こちらに適用できる高速化手法を調査する必要がある。

### 3.3 pi.super-computing.org 上での $\pi$ 並びに関連する情報公開

80 年代に入ってからこれまで 4 半世紀以上に渡り実施してきている円周率計算に関する学問的意義他については、一般向けの啓蒙・報告書 [文献 10] やその報告書中で引用している文献、また一般聴取者向けの講演で述べており繰り返すことになるので、本報告では省略する。なお 2009 年 12 月 31 日に、フランス人により、家庭でも購入可能なパソコンを 131 日間かけて、2 兆 6999 億 9999 万桁の記録が出たことが、2010 年になってから一部マスコミ誌にニュースとして掲載された。

さて毎年の年報で報告しているように、1995 年 6 月末から <http://www.super-computing.org/> 上で円周率の値、並びに関連する (anonymous ftp による) 情報の公開を継続しているが、情報公開開始以降これまでにおおよそ 15 年が経過した事になる。公開当初、一般公開中のものは 2 億桁までの  $\pi$  と  $1/\pi$ 、並びにそれらに関連する統計値であり、特別な使用同意契約を FAX でやりとりして締結する事で、アクセス可能な期間は最大 1 ヶ月であったが、40 億桁の  $\pi$  並びに  $1/\pi$  がアクセス出来る様になっていた。しかしそれまで公開していた [www.cc.u-tokyo.ac.jp](http://www.cc.u-tokyo.ac.jp) (sun ws) マシンの不調が度重なった為、2000 年 9 月中旬にサーバーマシンを変更したついでに、40 億桁の  $\pi$  並びに  $1/\pi$  が自由にアクセス出来る様にしていると同時に、要求元情報、総データ転送量等の統計情報を保存している。

この統計情報によると、2009 年度中に転送された総データ量は 2008 年度実績値と比べて約 1 割増加し、2006 年度、2007 年度と同じレベルまでに回復している。また合わせてアクセス回数は減少したが、アクセスサイト数とアクセス元のアドレスのそれぞれも増加している。これまでに各種マスメディアでの報道のとおり、2002 年度後半に新しい円周率計算の世界記録を樹立する事が出来たので、その時に得られた数値を利用し自由にアクセスできるデータ量及びデータ種の増加させる価値は十分にある。2007 年度まではこの種の計算桁数を Web で公開する為のハードウェア整備の為の競争的研究資金の獲得が困難という事情もあったが、2008 年度後半にその目処がたちハードウェアの準備



が行えた事により、公開桁数の大幅な増加を 2009 年度には実施できる見込みとなった。しかし、過去に行った計算結果やプログラム他の大量の情報の移動を求められたことにより、ディスクスペースが枯渇してしまい、2009 年度の公開が行えなくなってしまった。その為、2010 年度（以降）に、大容量ディスクスペースの確保を行って、公開桁数の大幅な増加を実施する事になる。2009 年度中での計算データの公開を実現させたいと考えている。

さて 2009 年度中の利用統計は 71,146 サイト（1995 年度は 748 サイト、1996 年度は 2,052 サイト、1997 年度は 5,604 サイト、1998 年度は 9,615 サイト、1999 年度は 16,396 サイト、2000 年度は 15,619 サイト、2001 年度は 15,354 サイト、2002 年度は 25,529 サイト、2003 年度は 35,652 サイト、2004 年度は 60,728 サイト、2005 年度は 64,312 サイト、2006 年度は 73,469 サイト、2007 年度は 81,664 サイト、2008 年度は 70,740 サイト）（DNS に名前が載っているサイトとしては 60,547 サイト（1995 年度は 641 サイト、1996 年度は 1,627 サイト、1997 年度は 4,654 サイト、1998 年度は 8,302 サイト、1999 年度は 14,092 サイト、2000 年度は 13,500 サイト、2001 年度は 13,518 サイト、2002 年度は 21,795 サイト、2003 年度は 29,503 サイト、2004 年度は 46,434 サイト、2005 年度は 54,459 サイト、2006 年度は 62,896 サイト、2007 年度は 71,500 サイト、2008 年度は 60,135 サイト））から総計 125,299 回（1995 年度は 5,012 回、1996 年度は 9,296 回、1997 年度は 16,127 回、1998 年度は 24,495 回、1999 年度は 40,457 回、2000 年度は 41,191 回、2001 年度は 37,188 回、2002 年度は 58,243 回、2003 年度は 67,851 回、2004 年度は 108,379 回、2005 年度は 116,351 回、2006 年度は 129,497 回、2007 年度は 147,086 回、2008 年度は 134,180 回）ファイルがアクセスされ、転送された総データ量は 122,689.0 MB（1996 年度は 6,918MB、1997 年度は 13,880MB、1998 年度は 14,455MB、1999 年度は 22,353MB、2000 年度は 55,392.6MB、2001 年度は 59,866.2MB、2002 年度は 111,082.7MB、2003 年度は 86,500.9MB、2004 年度は 120,464.2 MB、2005 年度は 167,199.9 MB、2006 年度は 126,747.6 MB、2007 年度は 124,604.4 MB、2008 年度は 109,411.6 MB）と一昨年度と同レベルに回復した。

これらのアクセス元のアドレス（u-tokyo.ac.jp の jp に対応する部分）種別は総計で 124 種類（ac ad ae al am an ar arpa at au aw ba be bg bh biz bm bo br bs by ca cc ch cl cn co com cr cu cx cy cz de dk do ec edu ee eg es eu fi fm fr ge gh gov gr gt gy hk hr hu id ie il in ir is it jo jp ke kh kr kz la lt lu lv ma mc md mil mk mu mx my mz na nc net ni nl no np nu nz org pa pe pf ph pk pl pt py ro rs ru sa se sg si sk su sv th to tr tt tv tw ua uk us uy ve vn ws yu za zw。なお 1996 年度は 50 種類、1997 年度は 59 種類、1998 年度は 65 種類、1999 年度は 78 種類、2000 年度は 68 種類、2001 年度は 76 種類、2002 年度は 82 種類、2003 年度は 89 種類、2004 年度は 104 種類、2005 年度は 115 種類、2006 年度は 126 種類、2007 年度は 125 種類、2008 年度は 123 種類）と、増加した。なおこれらの統計情報値が昨年度を上回った理由は不明である。

なお昨年末に不完全ながら記録が更新となったために、2010 年度中の更新を試みる価値があると考えている。その為の計算リソースの確保をまず国内設置マシンで試みたが、先方の政治的理由もあって失敗した。2010 年度は、海外に目を向けて、計算リソースの確保を試みると同時に、実際に計算を実施することを考えている。また関連して、Windows 以外の Apple コンピューター用の GUI インターフェースを持つプログラムをほぼ完成し、2010 年度早々に一般公開を予定している。

### 3.4 並列実行性能の高い数値計算ライブラリー及びプログラム開発環境に関する研究開発

分散メモリー型並列型計算機の性能を高度に引き出すプログラムを書く事は、過去における失敗経験量に大きく依存するとともに、職人芸的要素に基づく活動はなかなか評価を受けにくい。その結果プログラミングを生業としている、あるいは計算科学を専門としない多くの研究者にとって、高性能プログラム作成を行う事に対する動機付が弱く、プログラム作成効率を高める為の努力や、プログラムそのものの実行効率を極限まで高めようとする事に興味を持つとともに、長期にわたり没頭する事に喜びを感じる学生が生まれる事は多くは無い。ところが研究の最前線の進展速度を高める為には、少なくとも多くの研究者にとって使いやすく、高い実行性能が出せ、ノード数に関するスケラビリティ

が高く、広範囲の機能を有する数値計算ライブラリーの提供が必須と考える。その様な目的を持った数値計算ライブラリーが、一部ベンダーあるいは独立ベンダーから提供・販売、また個別に研究開発がなされてはいるが、その導入コストが高かったり、利用者が要求する機能を提供していなかったり、実現している機能間の性能に問題があったり、という問題が存在する。計算機を利用した研究・開発における研究環境のインフラ整備として、高い実行効率を容易に出す事ができるとともに、プログラムの保守管理が容易となる様なプログラム開発環境の開発や提供が今後必要となると考える。

この様な考え方に基づく研究として、分散メモリー型計算機および共有メモリー型計算機、PC クラスタからスカラー並列計算機あるいは（ハード及びソフト）ベクトル並列計算機までの広範囲のアーキテクチャーのどの様なマシンに対しても高い並列実行性能を発揮でき、広範囲の機能をカバーし、使いやすく実行効率の高い数値計算ライブラリーの研究開発、また簡単なプログラミングで実行効率が高いプログラム開発環境の研究開発、等がある。

この研究テーマに関連し、「次世代スーパーコンピューティング応用言語研究会（Next Generation Computing Application Language Consortium（略称 NSCAL）」を 2005 年 11 月に立ち上げて、研究室外のメンバーと活動を開始し、2009 年度は合計で 3 回の研究会を開催した。また、この研究の延長として、実際の研究に応用できるようにするための、さらなる機能強化点が存在しないかの検討を、外部の共同研究者と共同で開始した。実際の具体例については昨年度報告したので、今年度は報告しないこととする。

## 4 具体的成果

2009 年度は博士 3 年 1 名、博士 2 年 1 名、博士 1 年 0.5 名、修士 2 年 0.5 名の学生が研究室に存在した。なお、博士 1 年の学生は民間企業へ就職した。

2009 年度末に、並列数値計算についての教科書の最終校正を終了し、2010 年 4 月に 272 ページの本として発刊となる事が確定した。

また本報告では詳しくは述べなかったが、圧電トランス関係の活動と、新しい言語処理系について電子通信学会研究会主催のワークショップ [発表 4] での招待講演がある。また、鶴岡高等専門学校における特別講義の実施がある。

研究室学生の対外発表として Riemann Zeta 関数の零点計算に関するもの [発表 1]、[発表 2] と、工学系研究科の GCOE での発表 [発表 3] 他がある。

一般社会への貢献として、2009 年 11 月に政府が実施した、事業仕分けの仕分け人として貢献した。また大阪大学サイバーメディアセンターの外部評価における外部評価委員として貢献した。

獲得した外部資金として 2009 年度から開始の、科学研究費、基盤研究 (C)、課題番号 21500031、研究課題「科学技術計算に有用な高速 4 倍長演算に関する研究」、研究期間 3 年間。総額 3,400,000 円 (+ 間接経費 1,020,000 円) がある。

これらを具体的成果とする。

### その他の発表論文リスト

[発表 1] 清野善裕, 吉田仁, 金田康正: Riemann Zeta 関数の零点計算の高速化, 日本応用数理学会 2009 年度年会, p.333, 09/28/2009.

[発表 2] 清野善裕, 吉田仁, 金田康正: Riemann zeta 関数の零点とその計算の基礎について, 第 8 回「代数学と計算」研究集会 (AC2009), 12/03/2009.

[発表 3] 筒井直機: 多倍精度演算の高速化とその周辺関数に関する研究について, 東京大学グローバル COE セキュアライフ・エレクトロニクス平成 21 年度 博士課程学生報告書 pp.120-121.

- [発表 4] 金田康正（招待講演）：計算機科学が拓く未来, 電子情報通信学会, 集積回路研究専門委員会 (LSIの未来を考えるワークショップ) 2010年3月6日-8日, 香川県香川郡直島町琴弾地 ベネッセハウス.

## 引用文献関連

- [文献 1] 筒井直機：Study on the Fast Quadruple Precision Operations with Programmability and Portability プログラマビリティとポータビリティを考慮した高速な4倍精度演算に関する研究, 東京大学新領域創成科学研究科基盤情報学専攻修士課程修士論文, 2009年8月.
- [文献 2] 永井貴博、吉田仁、黒田久泰、金田康正：SR11000モデルJ2における4倍精度関和演算の高速化, 情報処理学会論文誌：コンピューティングシステム, Vol.48 No.SIG13(ACS19)pp.214-222, 2007.
- [文献 3] Takahiro Nagai, Hitoshi Yoshida, Hisayasu Kuroda, and Yasumasa Kanada：Fast Quadruple Precision Arithmetic Library on Parallel Computer SR11000/J2：ICCS 2008, Part I, LNCS 5101, pp. 446-455, 2008.
- [文献 4] IEEE 754-2008, IEEE Standard for Floating-Point Arithmetic
- [文献 5] T. J. Dekker. A floating-point technique for extending the available precision. Numerische Mathematik, 10:224-242, 1971.
- [文献 6] HITACHI SR16000/L1、<http://www.hitachi.co.jp/>.
- [文献 7] A. Atkin and D. Bernstein, Prime sieves using binary quadratic forms, 1999. [online]. Available: [citeseer.ist.psu.edu/atkin99prime.html](http://citeseer.ist.psu.edu/atkin99prime.html)
- [文献 8] H. Riesel, Prime Numbers and Computer Methods for factorization, Birkhäuser, 1994.
- [文献 9] H. S. Warren, Hacker's Delight, Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc. 2002.
- [文献 10] 金田康正：計算機による円周率計算, 数学文化, pp. 72-83, 日本評論社, Vol. 1, No. 1, 2003.

# 言語処理系とサービス体系におけるセキュリティ保証の研究

## Optimization Verifying Compiler, Document Carrying Authorization, and LoA of Services

佐藤周行

### 1 概要

われわれは、「検証付最適化コンパイラ」のもと、最適化に際してセマンティクスを保証するための検証の枠組と、最適化の結果、それが性能を確実に上げていることを検証する枠組を研究してきた。これは言語処理系を中心としてセマンティクスを保証するための枠組の研究だったが、さらにサービス体系全般に敷衍し、複数のテーマについて「保証」をテーマに研究を行っている。2008年度は特に以下の研究を行った。

1. プログラムの性能改善のためのコンパイラ最適化の理論の研究。中間表現を再構築し、データフローをベースにした中間表現を提案した。ここで、最適化はある種のグラフ変換として表現され、検証の枠組が提供される。
2. アルゴリズム変換を利用したコンパイラ最適化の拡張の研究。コンパイラ最適化は、プログラム解析をベースにしているが、それを拡張し、イディオム変換とその高階化を通して、より強力な枠組が構築できることを示した。
3. 認証におけるセキュリティ表現。セキュアな認証を用いて、認可を表現する枠組みの研究を行った。プラットフォームは言語処理系に限らず、一般に言語で表現されているもの一般である。特にワークフローを証明するための理論的な枠組みを提唱した。
4. ID の LoA の研究。ID の LoA は、強固かつ合理的な認証基盤を構築する上で必須の研究対象である。東大で運用中のサーバ証明書審査体制において認証強度の criteria を定めるための評価軸を定めることから、LoA を通して IdP と SP のインタラクション全体を統制する枠組みの提案を行った。

### 2 プログラムの性能改善のためのコンパイラ最適化の理論の研究

#### 2.1 背景

コンパイラの研究の中心である最適化において、従来、研究の展開はアドホックなノウハウの蓄積によってなされてきた。しかし、副作用を許容する最適化、性能的にトレードオフを持つ最



適化群の登場など、最適化が複雑・高度になった現在では、理論的基盤を持たなければ最適化適用の正しさの証明さえ危うくなってきている。事実 Gnu-C では、いまだに大量のバグレポートがなされている。

検証付きコンパイラ (Verifying Compiler) は、もともと Hoare の提唱したプログラムの正当性を併せて証明するコンパイラシステムとして 21 世紀の課題として提唱した概念である。本研究で目指すのは、コンパイラ最適化における上のような危機的状況に対応すべく最適化をその正しさの証明を込めて適用するコンパイラの構築であり、**最適化検証つきコンパイラ (Optimization Verifying Compiler)** と呼ぶべきものである。

従来の国内外における構成要素としての研究成果にはアセンブラの型付けによる動作正当性の証明 (Necula)、最適化の適用前後での意味保存の証明 (Necula)、最適化適用条件の時相論理等での形式化 (D Jones) を含む。さらに広義にはモデル検査の応用として動作正当性を証明することも含む。

現在、最適化検証付きコンパイラが検証することのひとつの柱は最適化適用の前後でプログラムの意味が保存されることである (Translation Validation)。証明を最適化の適用と同時に行うことを、アセンブラ言語に対する型理論として展開する。最適化適用はデータフロー解析を本質的に用いることから、データフロー理論を包含した型理論の構築につなげることができる。この最大の意義はデータフロー型の最適化の正当性を形式的に証明する体系を提示することである。また研究の基礎として、動作同一性証明に適したプログラム実行の適切な抽象化ができることになる。

また、最適化検証付きコンパイラはプログラム実行の抽象化を通して、最適化適用の最大の目的である性能向上の理論的解析のベースになる。現実的に妥当な数理的な性能モデルを仮定すると、その中で適用最適化が「真に」最適なコードを出すかどうかの検証ができる。(従来の最適化は、もとのコードと比較して、性能が良くなるかもしれないということを定性的にしか示せなかった。自動チューニング (Atlas 他) は、没理論でノウハウを適用することを目指しているものである。また、アーキテクチャ研究ではエミュレータが実質的なモデルの役割を果たしてきた。しかし、アンローリングなどのパラメトリックな最適化を解析するには数理的なモデルが必須である。) 性能モデルと、各種整数計画法の適用により、複数の最適化の組み合わせのどれが真に最適かを決定することができる。これはまさに性能面からみた検証であり、検証というプロセスが従来のアドホックな最適化研究に与えるインパクトの大きさを証明するものである。理論的な検証付きコンパイラが具体的な性能モデルの概念と結合して具体的な性能保証をすることは、理論と実践の両面での貢献が大きいものである。

## 2.2 内容

目的としてあげたゴールに対して 2009 年度、本研究では具体的に以下のようにテーマを設定した。

### 1. 最適化の正しさを検証するためのよりよい理論的枠組の研究

プログラムの最適化の検証は、Translation Validation をベースにした枠組で論じられるのが一般的であるが、変換に際して、プログラムポイントの対応を取る必要があり、適用性が広くない。それに対し、中間表現そのものを再構築し、本質的な拡張を行うことを目指した。具体的には、データフロー言語をベースにしてグラフ型の中間表現を設計し、最適化をグラフ変換として表現した。

## 2.3 具体的成果

[発表 1]において、データフロー言語をベースにしてグラフ型の中間表現を設計し、最適化をグラフ変換として表現することに成功した。いくつかの最適化をグラフ変換として具体的に定義しなおし、表現力の検証を行った。

# 3 アルゴリズム変換を利用したコンパイラ最適化の拡張の研究

## 3.1 背景

コンピュータアーキテクチャの複雑化に伴い、プログラムのチューニングの難度が増している。性能要素が複雑に絡み合うので、単純なモデルでは性能予測と実測のずれが大きくなっていることが最大の要因である。そのため、過去に省みられなくなったアルゴリズムをもう一度適用したり、さらに新しいテクニックで変換することが結果的に有効である場面が多くなってきた。このように、アルゴリズムレベルからの検討が要請される中、コンパイラ最適化の限界も見えてきた。プログラム解析をベースとしたコンパイラ最適化では、アルゴリズムの変換に対応することが難しい。より柔軟な変換ルールのもとで、最適のアルゴリズム、ルーチンを探し出すことが求められ始めている。

## 3.2 内容

ここでは、従来から行われてきたイディオム認識を、中間表現を適切に設定することによって、ループレベルのアルゴリズムに適用した。理論を項書き換え理論によって定式化し、イディオム認識の表現力の強さを検討した。さらに、これを高階に適用して、パターンの認識までできる枠組を構築した。

## 3.3 具体的成果

[査読付 5]において、イディオム認識を高階に適用し、パターンやテンプレートの認識によるアルゴリズム認識の拡張に成功した。

# 4 認証におけるセキュリティ表現

## 4.1 背景

コンピュータが情報システムの中核をなすようになると、データへのアクセス権限の証明と、認証（アクセス元の本人性証明）が重要になっている。それは一つとしてアクセス権設定のためのロールモデルの研究を生み出し、もうひとつはデジタルデータを対象にした厳密な証明をもとめる PKI を生み出した。ロールモデルとして RBAC (NIST Standard CS1.1) が実用化され、また組織のロールの研究から ORBAC(<http://www.orbac.org/>) が研究されるなどこの分野の成長は著しい。一方、PKI は、現在電子署名や暗号化メール (RFC2633 等)、さらに認証に使われるようになった。しかし、これらの学術的な成果が一般的な情報システムに組み込まれるには、まだ時間がかかるというのが共通の理解である。



他方において、サービス志向アーキテクチャにおけるサービス構築技術においてシステムどうしが通信を行なうワークフローの定式化が WS3C(<http://www.w3c.org/>) によって進行している。とくに WS-\*の枠内において WS-BPEL によりビジネスプロセスを記述する流れが決定的なものになっている。ワークフローの定式化は Grid その他、複数のプロセスのインタラクションにより計算が進行していくものにおいても重要である。ここでは WS-BPEL にとどまらない言語が提案、研究、実装されている。しかし、ここでのワークフローはロールモデル中の複雑な意思決定機構を反映できるようなものではなかった。

組織における意思決定をワークフローに反映させるためには、従来の貧弱なワークフロー表現では対応できないことはもはや明らかである。

## 4.2 内容

現在のサービス構築技術、またサービス志向アーキテクチャにおいて、ワークフローの表現が本質的になっている。現実のワークフローによるプロセスの進行においては、進行をアクティベートさせるものが重要である。とくに人間による「承認」が本質的に重要である。これが署名や印鑑で担保されることにより現実の世界は動いている。情報システム、およびその結合体は、ワークフローと進行のアクティベータの実現により、はじめてある程度の複雑度を持ったサービスを構築できる。ところが、現状は、要素的なサービス（DB アクセス、計算等）の記述は進んでいても、それらを結合させるワークフローの記述力は十分でない。

しかし、要素技術は存在する。それがロールモデルであり、PKI の特に電子署名技術による認証（証明）機能の提供である。不足しているのは、一定の動作がある言明を証明していることの形式化とその表現である。本研究では、ワークフロー、特にプロセスのアクティベーションに関する記述を定式化し、加えて PKI の電子署名技術を使って一連のプロセスに（実世界上での）セマンティクスを与えるためのフレームワークを確立する手法を研究した。

情報化以前においては、これら「承認」に基づくアクティベーションはドキュメントの流通によって行なわれてきた。ここではドキュメントが組織の意思を体現し、署名や押印がそれを承認するものとして扱われてきた。本研究は、特にこれに着目する。すなわち、ドキュメントが持つ「意思をセマンティクスとして表現し、PKI による電子署名で承認を表現する。これによって電子化されたドキュメントをワークフローの中に組み込むことが可能になる。

## 4.3 具体的成果

[査読付 2] において、分散環境中にワークフローが検証可能なドキュメントをやりとりする枠組みを提唱した。これは検証可能なプログラムコードが分散環境を飛び交う PCC(Proof Carrying Code) 等と並列に論じられるべきものである。

PCC では検証器がコードの証明を受け取る。われわれの枠組みでは検証器はある要求を満たすドキュメントをモデルとして受け取る。証明の生成が困難であるのと同じように、「ある要求を満たす」（上司の許可を得ていることを証明している等）ことを充足するドキュメントの作成は業務におけるワークフローの完成と同等の困難さを持つ。したがってモデルとしてのドキュメントを受け渡すことが重要である。

## 5 LoA of Identity and Services

### 5.1 背景

認証基盤で、ID のライフサイクル管理は基本的なものである。その構築において NIST の基準にも見られるように、「ID の保証基準」を定め、それによってコンテンツへのアクセスを制御する技術が一般的になってきた。この背景には、合理的なコストでセキュリティを担保することへの強い要請がある。

PKI プロジェクトで実施しているクライアント証明書発行のための UTCA と、サーバ証明書発行審査のための TRA は、強い認証強度をもつように CP/CPS を作成し、実際に運用しているにもかかわらずその認知度は一般的ではない。これはひとつには監査その他標準的な認証強度の認定プロセスを経ていないことが原因であるが、もうひとつには認証強度の criteria が定められていないことも原因としてあげられる。

### 5.2 内容

この研究では、認証強度の criteria を定めるための評価軸を定めることを行っている。特にサーバ証明書において FQDN のすべてのドメインについてまで保証を行うための審査体制の構築の研究を行った。また、これをもとに、EV 証明書と WTCA 証明書の中間の強度を持つ証明書を合理的なコストで発行するための基準についても検討を行った。

加えて、認証強度を保証することでどのようなサービスが展開可能かの研究を実際にシステムを構築することで行った。

認証強度の criteria は、組織内の統制のみならず、組織間のデータ通信にも有用である。本年度は特に証明書のパス検証を拡張する枠組を研究した。

### 5.3 具体的成果

[査読付 1] において、認証の LoA の拡張を行った。ここでは、IdP と SP 両方に LoA を与え、ふたつの LoA の対応を組織のセキュリティポリシーで与える枠組みを提唱している。これにより、従来行われている IdP のみに LoA を与える枠組を本質的に拡張することが可能になった。さらに、[査読付 4] において、異なる方式をひとつの LoA にまとめていることを緩和し、 $\pm\epsilon$  を付加する方法について研究した。

なお、研究の途中で Open source software のアップグレードのコストに注目せざるを得ない状況に至った。2008 年に起こった、OpenSSL に対する Debian の「悪い」改変はその典型的な例である。[査読付 3, 発表 2] において、この改変の影響を調査した。ここでは、自明な並列を用い、対象となる鍵生成を高速に行ったことがベースとなっている。

さらに、組織間で合理的なコストで信用を構築するために、証明書のパス構築の概念拡張を試みた。成果は、以下の形で発表する予定である。

- kubo, A., Sato, H.: Design of Graded Trusts by Using Dynamic Path Validation, to appear in IFIP Trust Management 2010.

また、クラウドのセキュリティモデル、特にトラストモデルについても検討を開始している。初期の成果は、以下の形で発表する予定である。

- Sato, H., Kanai, A., Tanimoto, S.: A Cloud Trust Model in Security Aware Cloud, to appear in SAINT 2010

## 6 成果要覧

### 査読付論文リスト

- [査読付 1] Sato, Hiroyuki: Service Framework Based on Grades of IdPs and SPs, Proceedings of Security and Management 2009, 379–385, 2009.
- [査読付 2] Sato, Hiroyuki: Analyzing Semantics of Documents by using a Program Analysis Method, Proceedings of 33rd IEEE Computer Software and Applications Coonference (COMPSAC), 373–382, 2009.
- [査読付 3] Nishimura, Takeshi, Sato, Hiroyuki: Analysis of a Security Incident of Opensource Middleware – Case Analysis of 2008 Debian Incident of OpenSSL–, in MidArch 2009, Proceedings of 9th International Symposium on Applications and the Internet, 247–250, 2009.
- [査読付 4] Sato, Hiroyuki:  $N \pm \epsilon$ : Reflecting Local Risk Assessment in LoA, On the Move 2009 (LNCS 5871), 833–847, 2009.
- [査読付 5] Sato, Hiroyuki: Idiom Recognition and Program Scheme Recognition based Program Transformation for Performance Tuning – Beyond Compiler Optimizations–, Proceedings of Parallel and Distributed Computing, Applications and Technologies 2009, 272–279, 2009.

### その他の発表論文リスト

- [発表 1] 雪下洋輔, 佐藤周行: シンボリックアサインメントグラフを用いたコンパイラ最適化器, 第26回ソフトウェア科学会大会, 3C-2, 2009.
- [発表 2] 西村健, 佐藤周行: 2008年のDebian OpenSSL インシデントにみるオープンソースソフトウェアのセキュリティ分析, コンピュータセキュリティシンポジウム 2009, 2009.

# データ集約的な分散計算のための枠組み

田浦 健次郎

情報理工学系研究科

スーパーコンピュータ部門(兼務)

## 1 概要

ゲノム処理、天文学、生物医学、素粒子物理、環境モニタリングなど多くの分野で、大量に蓄積されたデータとその処理が必須の方法論となっている。また、実験データが複数の拠点に分散して収集・蓄積されることから、拠点をまたがったデータの共有や協調処理が不可欠となる。クラウド環境のような、オンデマンドに利用できる計算資源を利用する場合も、ローカルな環境とクラウドを同様に協調させることが不可欠であり、いずれの場合も分散したデータを容易に、かつ高性能に処理できる枠組みの構築が非常に重要な課題である。

我々はそのような枠組みを構築するため、以下のような研究を進めている。

**分散ワークフロー処理系:** SSH、TORQUE、Sun Grid Engine など様々なインタフェースでアクセスされる計算資源を、管理者による設定なく、個々のユーザレベルで統合して利用でき、かつ各拠点へのインストールも不要な並列シェル GXP を開発している。その上で make によってワークフローを並列実行できる処理系 GXP make を開発し、オープンソースのソフトウェアとして公開している [公開 3]。

**分散ファイルシステム:** 拠点間でのデータ共有も、管理者による設定なく、個々のユーザレベルで可能とする分散ファイルシステムを設計、実装している。ユーザレベルファイルシステム FUSE と、その上で SSH のみを用いてファイル共有が可能な SSHFS をベースとし、スケーラビリティを高めた分散ファイルシステム GMount を設計・実装した [査読付 1]。また、並列ファイルシステムのためのベンチマーク paramark を実装し、オープンソースのソフトウェアとして公開している [公開 1]。

**データ集約的アプリケーションのアクセス解析:** 分散ファイルシステムの設計や、ワークフロースケジューラに対する知見を得るため、5つの実アプリケーションを含むデータ集約的なプログラムレポジトリを構築し<sup>1</sup>、それらのファイル I/O 挙動を集積した [査読付 4]。そのために、再コンパイルや再リンクを行うことなく、ファイル I/O を解析する枠組み paratrac を設計・実装し [査読付 2]、オープンソースのソフトウェアとして公開している [公開 2]。

**分散環境における高性能通信:** 拠点をまたがった環境ではネットワーク性能が非均質で遅延も大きい。また、環境は複雑な上、利用可能な性能も動的に変化する。クラウド環境のように一般のユーザには構成すら未知な環境を使う必要もある。そのようなで、データの集約 (gather) や配信 (broadcast) などを、適応的に行うための、ネットワーク構成の推論、性能測定、非均質環境下での通信アルゴリズムに関する研究を行っている [査読付 7, 査読付 8, 査読付 9, 査読付 5, 発表 1, 査読付 3, 発表 3]。

<sup>1</sup><http://www.intrigger.jp/wiki/index.php/Applications>



**資源の動的な伸縮が可能な大域アドレス空間:** 科学技術計算において、UPC、Global Arrays、Co-Array Fortran、XcalableMP などの、大域的なアドレス空間をサポートするプログラミングモデルが、MPI よりも生産性の高いモデルとして浸透しつつある。我々は、動的なスレッド生成・消滅のサポート、計算資源の増減、スレッドの移動など、より動的な環境に適応した大域アドレス空間に基づくプログラミングモデル DMI を提案し、処理系を設計・実装している [発表 2, 査読付 6]。

本報告の残りでは、上記から分散ファイルシステムに関して述べる。

## 2 分散ファイルシステム GMount

### 2.1 動機・概要

GMount は、独立に管理された拠点間で、分散共有ファイルシステムを「即席で」、ad-hoc に構築し、データ集約的計算のためのストレージとして利用可能にするシステムである。

一般に計算機間でファイルを共有するために、NFS などのネットワークファイルシステムが用いられる。また、高性能計算のためには、ディスクやネットワークバンド幅を増強するために、ファイルサーバを複数束ねる事ができる、並列ファイルシステムが用いられる (Lutre など)。しかし並列ファイルシステムの性能・安定性を達成することは難しく、一般に高コスト要因となる。それは、容量プランニングにあたり、ファイルサーバをアクセスする「潜在的な」クライアント数に見合ったサーバ数および、ネットワークバンド幅を用意しなくてはならないからである。そうしてもなお、拠点内におけるファイル共有しか達成できない。

一方、データ集約的な計算では、処理すべきデータは実験装置などに近い拠点に蓄積され、それを処理するのに必要な計算資源は資源管理システムを経由するなどして、その時々利用可能な計算資源が利用されるのが通常である。それらは多くの場合管理ドメインも異なり、統一的なソフトウェアでデータを透明にアクセスできるような基盤が整備されている場合は少ない。

GMount はそのような環境でも、必要に応じて任意のホスト間でファイルを即座に共有する事を可能にする。まず、ビルディングブロックとして、ユーザレベルでファイルシステムの構築を可能とする枠組み FUSE<sup>2</sup>を用いる。また、SSH によるログインが可能な、任意の遠隔ホストをマウントする事ができる、SSHFS ファイルシステム<sup>3</sup>をベースとする。

GMount は、互いに SSH ログイン可能なホスト間で、互いの共有ディレクトリを重ね合わせた単一のファイルシステムを即座に構築する。短く述べれば、以下(\*)を達成する。

- サーバ  $h_1, \dots, h_n$  がそれぞれ  $d_1, \dots, d_n$  を共有ディレクトリとして提供している。
- この時、 $d_1, \dots, d_n$  をすべて重ね合わせたディレクトリを、グローバルな共有ディレクトリとして、全クライアントに提供する。

特に、SSH によるアクセスさえ受け付ければ任意のホストをサーバとする事ができるため、計算に必要なホスト間で柔軟にファイルシステムを構築することが可能である。また、サーバとなるために必要な条件は SSH によるアクセスを受け付ける事だけなので、計算に参加するホストをその場でファイルサーバとして用いることができる。

この性質から、GMount はデータ集約的なワークフローへ、入力データを容易にインポートする手段、実行中の一時ファイルを保存する手段として有効に利用できる。また、計算に参加するホスト自身をファイルサーバとして用いることで、ワークフロースケジューラと連動したローカリティの高い実行をする余地が生まれる。

<sup>2</sup><http://fuse.sourceforge.net/>

<sup>3</sup><http://fuse.sourceforge.net/sshfs.html>

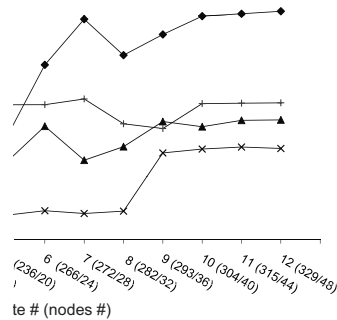


図 1: 拠点数 (12 まで)、ノード数 (329 まで) に応じたマウント処理時間

## 2.2 成果

(\*)を達成する単純な方法としては、実際にすべてのホスト間で接続を作り、マウントを行う方法があるが、明らかにスケーラビリティの限界がある。GMount ではホストを階層的にグループ化し、必要な接続 (マウント) の個数を少なくしている。

また、一般に並列ファイルシステムでは、単一のメタデータサーバにすべてのファイルの位置情報を持たせているが、GMount では、クライアントの近傍から探索するという手法をとっている。精密な評価は今後の課題だが、この方式には、

- データのインポート (マウント) 時にファイルの位置情報を集約する必要がない
- 近傍のファイル参照に対して、広域のトラフィックが発生しない

などの利点があり、広域ファイルシステムで有望な方式として検討している。

これらにより、GMount は、

- 多数のホストに対しても短いファイルシステム構築 (実際にはマウント・接続処理を行う) 時間
- 広域環境であっても、局所性の高い (近傍を参照する確率の高い) ワークロードに対する、良好なメタデータアクセス性能

を達成している。

図 1 は、拠点数 12 まで、ファイルサーバ・クラインとノード数を 329 まで増やしたときの、ファイルシステム構築時間 (全ホストが同じディレクトリを見られるよう、マウント処理を行うのにかかった時間) を示しており、12 拠点/329 ノードの場合でも 10 秒ほどで終了している。

図 2 は、アクセスの局所性 (アクセスするファイルが同一の拠点内にある確率) に応じて、GMount におけるメタデータ性能がどのように変わるかを示している。

比較として図 3 は Gfarm における結果を示している。単一のメタデータサーバを用いる Gfarm においては、メタデータサーバ、クライアント、ファイルサーバがすべて同一の拠点にある場合 (左) は非常に高速だが、そうでない場合 (右) は、たとえアクセス自体は局所的であっても、性能は同じである事がわかる。

## 2.3 今後

GMount のアプローチでは、実際の性能はホスト間にどのような接続関係を構築するかによって大きく左右される。また、場合により複数ホップを経由してメタデータへのアクセスが行われる。したがって接続関係やそのグラフの半径などを、実際のネットワーク性能に適応して定める必要がある。また、複数ホップ数を経由したメタデータアクセスを行う際のオーバーヘッドが大きいなどの問題点もある。

これらについてさらなる研究を行い、実際のデータ集約的なワークロードでの評価・有用性の検証を行うことが今後の課題である。



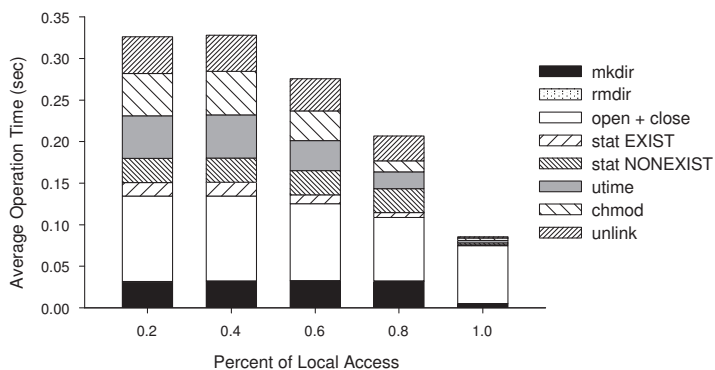


図 2: アクセスの局所性に応じたメタデータ性能 (GMount の場合)。局所性に応じて性能が向上する。

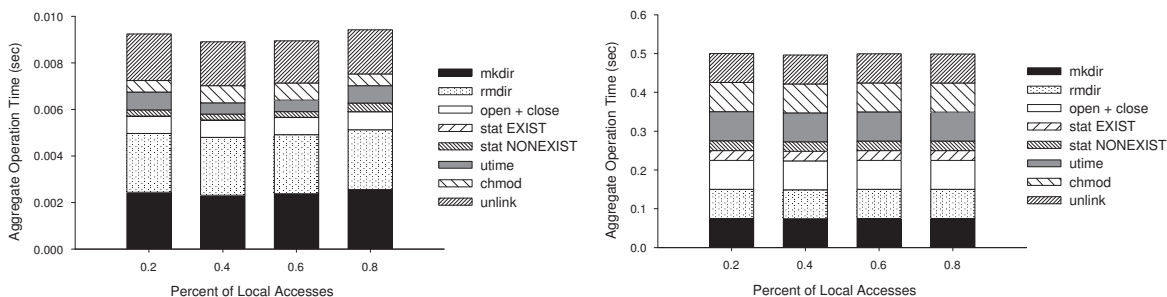


図 3: アクセスの局所性に応じたメタデータ性能 (Gfarm の場合)。単一メタデータサーバに基づく Gfarm では、クライアント、対象ファイル、メタデータサーバがすべて近くにある場合の性能はよい (左)。一方、クライアントと対象ファイルが近くにあっても、メタデータサーバが遠い場合性能はメタデータサーバとの通信で律速される (右)

### 3 分散環境におけるデッドロックフリーオーバーレイ

多拠点にまたがる広域環境で並列計算を行うには様々な問題がある。

**接続性:** NAT や firewall など通信が制限されている場合が多い

**スケーラビリティ:** 単純に、通信が発生したホスト対間で接続を作ってしまうと、プロセスのファイルディスクリプタ数、ルータのセッション数などの制限によって通信がエラーになったり、それらをクリアしても接続ごとに無駄にメモリを消費するなど、スケーラビリティの障害になる。

**衝突による性能悪化:** 大規模なネットワークでは異なるホスト対間の通信が、同じ中間リンクを共有し、そのバンド幅が限られているため、それらが衝突してパケットロスを引き起こし、全体性能が悪化する。

これらの問題を統一的に解決する枠組みとして、並列計算におけるオーバーレイネットワークを研究している。つまり、ホスト間の接続を少数に制限し、直接接続されていないホスト間は、複数の接続上をルーティングすることで通信を行う。

オーバーレイネットワークは安易に構築すると、中間ノードにおいてメモリオーバーフローもしくはデッドロックを引き起こす。まず、中間ノードが、単純に受け取れるデータをすべて受け取ると、メモリオーバーフローとなる。一方、単純にある程度以上のデータはバッファリングしないというア

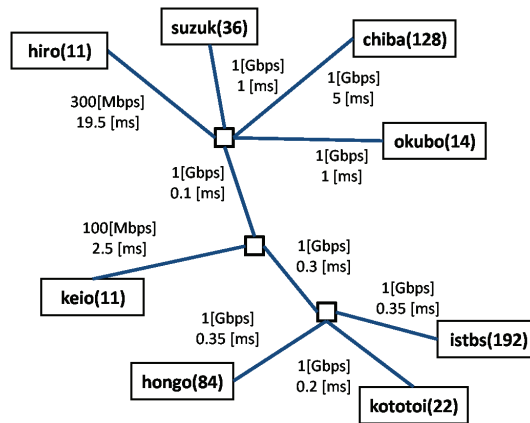


図 4: 実験およびシミュレーション環境

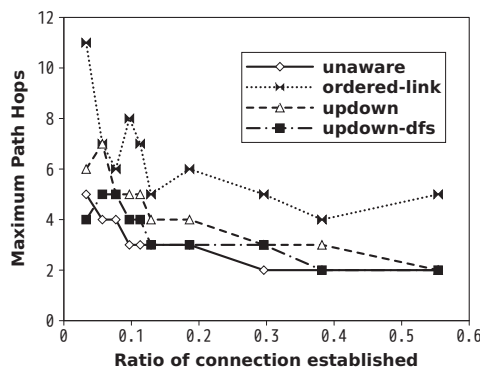


図 5: 各手法による最大ホップ数。updown-dfs は、デッドロックを考慮しない方式からの性能劣化が少ない。

ルゴリズムは、デッドロックを引き起こす。これは、並列計算機用のネットワークの文脈で研究されてきた、デッドロックフリーなルーティング (updown ルーティング、ordered-link ルーティングなど) を用いることで、原理的には解決できる問題であり、我々もその手法を用いている。しかしながら、単純にデッドロックフリールーティングを採用すると、それによる回り道が生じ、通信性能が劣化してしまう。

我々は、

- ノード間の物理的な距離に基づく接続の選択アルゴリズム
- 可能なデッドロックフリールーティングのうち、回り道による性能劣化の少ない方式 (updown-dfs)

を提案し、8 拠点 515 ノードからなる実環境 InTrigger<sup>4</sup> (図 4) 及び、シミュレーションによる評価を行った。

図 5 は、接続密度 (可能な接続のうちどれだけの割合を実際に接続するか) を様々に変え、ノード間の最大ホップ数を示している。“unaware” ラベルのついたグラフが、単純に最短パスを選択するもので、性能の上限を与えている (もちろんこれは、デッドロックに陥る可能性がある)。提案手法である updown-dfs が、ほとんどの接続密度で、unaware とほぼ同じホップ数を示している。図 6 は、デッドロックフリールーティングによって選択されたパスのバンド幅 (パス上のリンクバンド幅の最小値) が、デッドロックを考慮しないもの (性能の上限値) と比べてどれだけ劣っているかを示している。

<sup>4</sup><http://www.intrigger.jp>

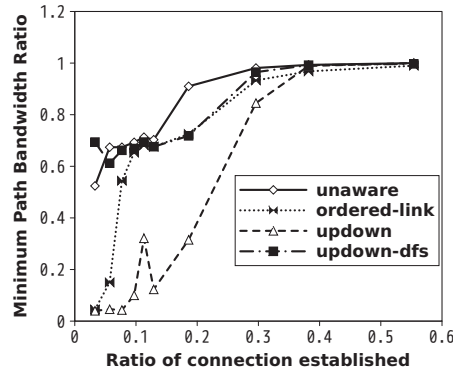


図 6: 各手法による最小バンド幅比 (実際のバンド幅/最短パスのバンド幅)。updown-dfs が良い性能を示している。

ホスト間の接続を制限する事は、MagPie や MPICH-G2 などの広域環境用の MPI で拠点間の接続を制限するなど、部分的には取り入れられてきた考え方であり、極めて自然である。しかしそれらは、手動で与えられた設定 (拠点内、拠点間) と固定された規則にしたがって接続を管理しているに過ぎず、本研究のように、与えられた任意の環境に対し、性能が優れた接続方法、ルーティング方法を適応的に構築するものではなかった。

現時点では、本研究はアルゴリズムの提案と評価にとどまっており、実際の並列アプリケーションを記述するだけの通信 API として実装されていない。今後の課題としては、本アルゴリズムを接続管理に用いつつ、ソケット API などの汎用的な API を実装し、広域環境での並列処理の基盤として用いるソフトウェアを実装することがあげられる。そしてその上に広域分散ファイルシステムなどの、大容量のトラフィックを不規則に発生させる基盤ソフトウェアを構築する。

## 4 その他: HA8000 システムへの Lustre ファイルシステムの導入

センターにおける HA8000 のサービス向上業務の一貫として、ファイルシステムの性能向上を行うべく、各種システムのベンチマーキング、製品調査、動作検証、新ファイルシステムの仕様策定などを、本センターの鴨志田らとともにに行った。入札の結果 Lustre ファイルシステムが導入され、2010 年 4 月から稼働している。

## 5 成果要覧

### 査読付論文

[査読付 1] Nan Dun, Kenjiro Taura, and Akinori Yonezawa. Gmount: An ad hoc and locality-aware distributed file system by using ssh and fuse. In *Proceedings of the IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID 2009)*, May 2009.

[査読付 2] Nan Dun, Kenjiro Taura, and Akinori Yonezawa. Paratrac: A fine-grained profiler for data-intensive workflows. In *Proceedings of ACM Conference on High Performance Distributed Computing (HPDC)*, 2010. to appear.

[査読付 3] Ken Hironaka, Hideo Saito, and Kenjiro Taura. High performance wide-area overlay using deadlock-free routing. In *Proceedings of International ACM Symposium on High Performance Distributed Computing (HPDC2009)*, pages 81–90, 2009.

- [査読付 4] Takeshi Shibata, Sung Jun Choi, and Kenjiro Taura. File-access patterns of data-intensive workflow applications and their implications to distributed filesystems. In *Proceedings of The Third International Workshop on Data Intensive Distributed Computing*, 2010. to appear.
- [査読付 5] 吉富翔太, 田浦健次朗. メッセージ衝突を防止する適応的な収集操作アルゴリズム. *電子情報通信学会論文誌 (情報爆発特集号)*, J93-D(6), June 2009.
- [査読付 6] 原健太朗, 田浦健次朗, 近山隆. Dmi: 計算資源の動的な参加/脱退をサポートする大規模分散共有メモリインタフェース. *情報処理学会論文誌*, 3(1):1-40, 3月 2010.
- [査読付 7] 柴田剛志, 田浦健次朗. トポロジ情報を用いた効率的かつ漸近安定な大容量ブロードキャスト. In *先進的計算基盤システムシンポジウム (SACIS2009) 予稿集*, pages 71-78, 2009.
- [査読付 8] 柴田剛志, 田浦健次朗. トポロジ情報を用いた効率的かつ漸近安定な大容量ブロードキャスト. *情報処理学会論文誌コンピューティングシステム (ACS)*, 2(3):47-57, 2009.
- [査読付 9] 吉富翔太, 弘中 健, 田浦 健次朗. メッセージ衝突を防止する適応的な収集操作アルゴリズム. In *先進的計算基盤システムシンポジウム (SACIS2009) 予稿集*, 2009.

## 公開ソフトウェア

- [公開 1] Nan Dun. paramark: high fidelity parallel file system benchmark. <http://code.google.com/p/paramark/>.
- [公開 2] Nan Dun. paratrac: a tracking suite for parallel/distributed applications. <http://code.google.com/p/paratrac/>.
- [公開 3] Kenjiro Taura. GXP: Grid and cluster shell. <http://www.logos.t.u-tokyo.ac.jp/gxp/>.

## その他の発表論文

- [発表 1] 吉富翔太, 田浦健次朗. メッセージ衝突を防止する適応的な集合通信. In *情報処理学会研究報告 HPC-121 (SWoPP 2009)*, 2009.
- [発表 2] 原健太朗, 田浦健次朗, 近山隆. DMI: 計算資源の動的な参加/脱退をサポートする大規模分散共有メモリインタフェース. In *情報処理学会研究報告 HPC-121 (SWoPP 2009)*, 2009.
- [発表 3] 長沼翔, 田浦健次朗. 大規模分散環境におけるバンド幅測定アルゴリズム. In *情報処理学会研究報告 HPC-121 (SWoPP 2009)*, 2009.

# シームレス高生産・高性能プログラミング環境

## — 高効率・高可搬性ライブラリの研究開発 —

黒田 久泰

### 1 概要

文部科学省の次世代 IT 基盤構築のための研究開発「e-サイエンス実現のためのシステム統合・連携ソフトウェアの研究開発」における選定課題「シームレス高生産・高性能プログラミング環境(研究代表者:石川裕教授)」に取り組んでいる。本課題の目標は、PC クラスタから情報基盤センターに設置されているスーパーコンピューターまでをシームレスに利用できるプログラミング環境を実現することである。サブテーマとして「高性能並列プログラミング言語処理系の研究開発」、「高生産並列スクリプト言語の研究開発」、「高効率・高可搬性ライブラリの研究開発」があり、私は「高効率・高可搬性ライブラリの研究開発」に関連して自動チューニング(AT)機構を含む数値計算ライブラリの開発に取り組んでいる。

### 2 高効率・高可搬性ライブラリの研究開発

#### 2.1 背景

数値計算ソフトウェアの性能チューニングコストの増大が問題となっている。これは、PC クラスタからスーパーコンピューターまで様々なアーキテクチャー毎にチューニングしなくてはならないためである。また、実行時に与えられるデータセットの内容により、最適な数値アルゴリズムや実装方式の選択を行うことでそのためのオーバーヘッドを加味してもなお大幅な速度向上が得られる可能性もある。そのため、自動チューニング(AT)機構を含む数値計算ライブラリの役割は大きい。また、これまでのAT機能付きライブラリでは、汎用的なユーザーインターフェースが存在しないので、AT機能の再利用性が低く、開発コストが増大する傾向にあった。そこで、ライブラリ開発者やユーザー自身がAT機能付きライブラリを容易に開発したり、機能を追加したりできるようなAPIを提供することの意味も大きい。

#### 2.2 内容

汎用的自動チューニングAPIとしてOpenATLibの設計及び開発を行った。OpenATLibは数値計算ライブラリにおけるAT機能を規格化してAPI化し、その参照実装を提供する。現在、疎行列ソルバに特化したAT機能、及び実行時のAT機能を中心に機能開発を行っている。2008年度に開発したOpenATLib α版に対し、2009年度は機能を高度化したOpenATLib β版を開発した。このOpenATLib β版では、新たにユーザーからのAT方針の指定機能である数値計算ポリシーが設定できるようになった。提供機能として、演算速度、メモリ量、演算精度の3種が指定可能であり、それぞれAT戦略が異なる。詳細については片桐の報告を参照のこと。



## 2.3 具体的成果

2009 年度に開発を行った OpenATLib  $\beta$  版は PC クラスタコンソーシアムの Score Cluster System Version 7 download page (<http://www.pccluster.org/ja/score7.html>) からダウンロードできる。ソースプログラムのダウンロードページからは OpenATLib  $\beta$  版のソースプログラムも単体でダウンロードすることができるようになっているため管理者権限のない一般ユーザーのアカウントでもインストールして利用できるようになっている。本ライブラリの性能についての評価等は片桐の報告を参考のこと。

## 3 成果要覧

### 招待講演／招待論文

- [招待 1] Takahiro Katagiri, Takao Sakurai, Hisayasu Kuroda, Ken Naono, and Kengo Nakajima: Development of numerical computation policy facility and its effect on Xablib, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), Yokohama, JAPAN), OW-1 Innovative Algorithms for Multicore and Hierarchical Computational Environments (Organizers: H. Okuda, K. Nakajima and T. Aoki), Mar. 2010.

### 査読付論文リスト

- [査読付 1] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰: OpenATLib: 数値計算ライブラリ向け自動チューニングインターフェース, 2010 年ハイパフォーマンスコピューティングと計算科学シンポジウム (HPCS2010), HPCS2010 論文集, pp. 99-106, 2010 年 1 月.

### その他の発表論文リスト

- [発表 1] 片桐孝洋, 黒田久泰: マルチコア環境における自動チューニング機能付き疎行列反復解法ライブラリ, 第 14 回日本計算工学会講演会, オーガナイズドセッション, 「科学技術計算における自動チューニングの挑戦: マルチコア環境への適用」, 東京大学生産研究所計算工学会講演論文集, pp. 167-170, 2009 年 5 月.
- [発表 2] 片桐孝洋, 櫻井隆雄, 黒田久泰, 直野健, 中島研吾: 汎用的自動チューニングインターフェースをもつ疎行列反復解法ライブラリ, 2009 年先進的計算基盤システムシンポジウム (Symposium on Advanced Computing Systems and Infrastructures (SACISIS), 広島国際会議場, SACISIS2009 論文集, pp. 116-117, 2009 年 5 月.
- [発表 3] 片桐孝洋, 櫻井隆雄, 黒田久泰, 直野健, 中島研吾: OpenATLib: 汎用的な自動チューニングインターフェースの設計と実装, 2009 年並列／分散／協調処理に関する『仙台』サマー・ワークショップ (SWoPP2009), フォレスト仙台, 情報処理学会研究報告 2009-HPC-121, 2009 年 8 月.
- [発表 4] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰: OpenATLib を利用した疎行列ライブラリの開発と評価, 2009 年並列／分散／協調処理に関する『仙台』サマー・ワークショップ (SWoPP2009), フォレスト仙台, 情報処理学会研究報告 2009-HPC-121, 2009 年 8 月.
- [発表 5] 片桐孝洋, 黒田久泰: マルチコア環境における密および疎行列ソルバの自動チューニング機構の評価, 2009 年並列／分散／協調処理に関する『仙台』サマー・ワークショップ (SWoPP2009), フォレスト仙台, 日本応用数理学会「行列・固有値問題の解法とその応用」研究部会, 2009 年 8 月.
- [発表 6] 片桐孝洋, 黒田久泰, 中島研吾: Xablib: 汎用な自動チューニングインターフェース OpenATLib を利用した疎行列反復解法ライブラリ, 第 5 回 ASE 研究会, 2009 年 8 月.



- [発表 7] 黒田久泰, 片桐孝洋, 須田礼仁: 省電力のための基本演算ライブラリの実装と評価, 2009年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 日本応用数学会「行列・固有値問題の解法とその応用」研究部会, 2009年8月.
- [発表 8] 片桐孝洋, 黒田久泰: マルチコア・超並列計算機時代の自動チューニング機能付き疎行列反復解法ソルバ, 日本応用数学会 2009年度年会, 大阪大学豊中キャンパス, 年会予稿集, オーガナイズドセッション: ペタスケール環境を目指す数値計算ライブラリと自動チューニング技術, A2-1, 2009年9月.
- [発表 9] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰: クリロフ部分空間法に対する自動チューニングについて, 加速法フォーラム, 2009年.
- [発表 10] Takahiro Katagiri, Hisayasu Kuroda and Kengo Nakajima: Xablib: An Iterative Solver with a General Auto-tuning Interface "OpenATLib" --- Development of Numerical Computation Policy Function, 6th ASE Seminar, Information Technology Center, The University of Tokyo, Feb. 2010.
- [発表 11] Takahiro Katagiri, Takao Sakurai, Hisayasu Kuroda, Ken Naono and Kengo Nakajima: Xablib: An Iterative Solver with a General Auto-tuning Interface "OpenATLib", International Workshop on Peta-Scale Computing Programming Environment, Languages and Tools (WPSE 2010), Kyoto, Feb. 2010.
- [発表 12] Takao Sakurai, Ken Naono, Hisayasu Kuroda, Takahiro Katagiri, and Kengo Nakajima: OpenATLib: A General Auto-tuning Interface for Numerical Solvers, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: The State-of-the-art of Auto-tuning Technologies: Adaptation to Advanced Computer Environment and Numerical Libraries - Part I of II, MS6, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.
- [発表 13] Takahiro Katagiri, Hisayasu Kuroda and Kengo Nakajima: A Branchless Segmented Scan Method for Sparse Matrix-vector Multiplications, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: Parallel Programming Models and Algorithms for Multicore Clusters and GPGPUs - Part III of III, MS62, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.

## 特記事項

- [特記 1] 黒田久泰, 直野健, 岩下武史: 情報処理学会誌「情報処理」, 大特集: 科学技術計算におけるソフトウェア自動チューニング, 6. 自動チューニング機能付き数値計算ライブラリ, Vol. 50, No 6, pp.505-511, 2009年6月.

# クラスタからクラスタ外部へのネットワーク出力に関する研究

松葉 浩也

## 1 概要

計算機分野における「クラスタ」とは多数の計算機からなる高性能コンピュータであり、今日では並列計算機や大規模サーバなどを構成する一般的な形態のひとつとなっている。このようなクラスタからインターネットに対してデータを送出する場合、多数の計算機から出力されるデータがクラスタ内のネットワーク装置の中で競合しパケットロスが多発することがあり、このようなパケットロスは通信性能低下の原因の一つとなる。本年度はこのようなクラスタから外部への通信を安定的に行うためのソフトウェア技術を研究、開発し、従来手法より高速な通信方式を確立することに成功した。

## 2 クラスタからクラスタ外部へのネットワーク出力に関する研究

### 2.1 背景

計算機センターに設置される大型計算機は歴史的には専用設計されたスーパーコンピュータであることが多かったが、近年では汎用部品によるクラスタ型のスーパーコンピュータも導入されるようになってきている。本センターで稼働中の HA8000 クラスタシステムもそのようなクラスタ型スーパーコンピュータである。

クラスタ型のコンピュータは構成要素となるそれぞれの計算機に一般的な設計のサーバコンピュータが使用され、Linux などの単体でもサーバとして運用に耐える機能を備えたオペレーティングシステムが使用される。そのため、クラスタからインターネットなどの外部計算機と通信する際にはそれぞれのノードが備える通信機能を利用するのが最も単純な方法であり、実際に多くのクラスタではそのような運用がなされている。

ところがクラスタの各ノードが協調することなく個々にネットワーク送信を行った場合、複数のコンピュータが同時に送信する可能性があり、合計の送信速度がインターネット接続の回線速度を越え

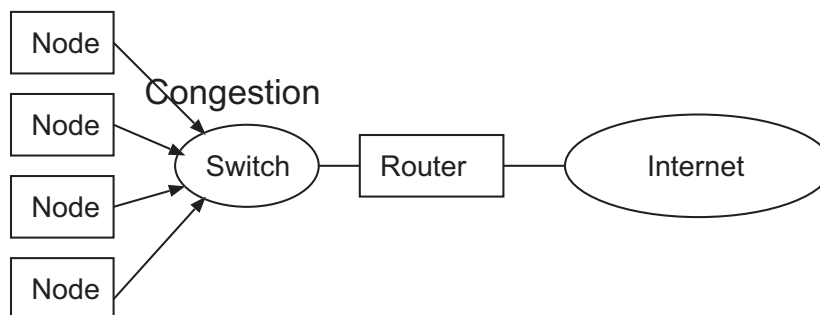


図 1: 同時出力によるネットワーク混雑

てしまうことがある。このような場合は図1のように、各ノードへのネットワーク接続を収容するスイッチにおいて通信パケットが破棄されることとなり、再送処理等のために通信性能が低下する。

並列計算のための通信を対象にこのパケットロスの問題を解決する手法は過去に研究しており 2007年度の年報に報告済みである（文献 [査読付 1] も同内容）。この方法は通信を行うすべてのプロセスに通信ライブラリの使用を要求するため、並列計算における MPI のように、デファクトスタンダードの通信ライブラリが存在する分野では適用が容易であるが、UNIX のシステムコールを直接利用していることが多い一般的なサーバプログラムには不向きである。そこで本年度はアプリケーションに完全な透過性を提供しつつ、安定した高速な送信を可能とするクラスタ向け送信システムを研究した。なお、本報告は [特記 1] にある報告者による学位論文の内容の一部と重複する。

## 2.2 内容

図2にあるように提案手法ではクラスタ内に I/O サーバと呼ばれる特別な計算機を配置し、送信データはこの I/O サーバを経由するようにする。そして、クラスタの各ノードと I/O サーバの間は Myrinet やロスレスの 10Gbps Ethernet など、データセンター向け的高速ネットワークで通信する。

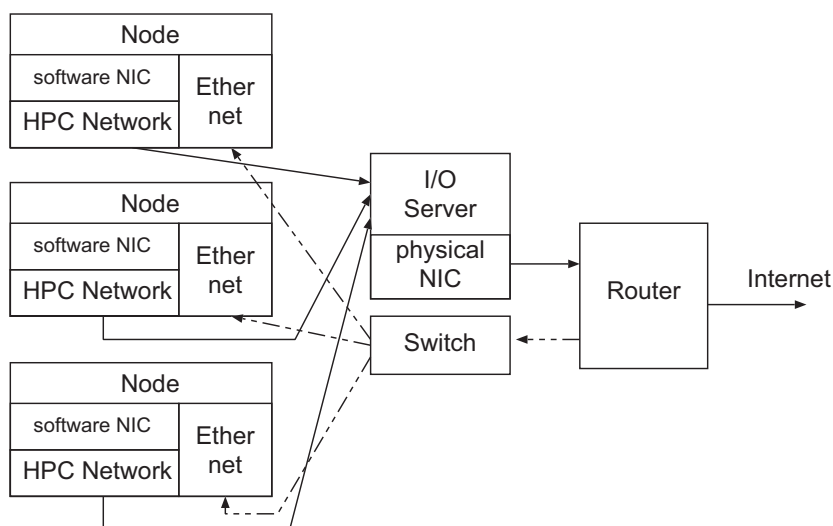


図 2: 提案手法概要

Myrinet やロスレスの 10G Ethernet はそれ自体がハードウェアフロー制御機能を持ちパケットロスに強い設計となっているため、これらのネットワークを経由するだけでも性能向上が期待できるが、提案手法は単にネットワークトポロジを規定しているにとどまらず、以下の新規性がある。

### 新規フロー制御プロトコルによる利用者別帯域制御

クラスタノードと I/O サーバの間は新たに提案したフロー制御プロトコル (SFCP) を用いて通信する（プロトコルの詳細は文献 [特記 1] 参照）。このプロトコルを用いることで Myrinet やロスレスの 10G Ethernet が提供するハードウェアレベルのフローコントロールに頼ることなく、ソフトウェアレベルで送信過多状態を防ぐ。ハードウェアのフロー制御のみでは一部ユーザーが I/O サーバに対して大量にデータ送信を行った場合、無関係な利用者のデータまで止められてしまうが、提案したソフトウェアレベルのフロー制御は利用者別の独立したフロー制御が可能である。

### I/O サーバにおける帯域制御

I/O サーバは単にデータの中継を行うだけでなく、インターネットに出力するデータの送出スピードを調整する機能を持つ。これにより、インターネットで利用可能な帯域を越える速度で送出することを防ぎ、パケットロスを防止する。さらにこの機能は前述のフローコントロールと連携し

ており、I/O サーバは外部に送出する速度以上でクラスターノードからの流入がないよう調整する。これにより帯域調整によるパケットロスも防止する。

図3に提案手法の具体的設計を示す。Applications nodes とあるのがクラスターの一般ノードである。一般ノードの application, TCP, および IP の各レイヤーは通常のオペレーティングシステムとまったく同じであり、ネットワークアプリケーションはオペレーティングシステムが提供するネットワーク通信機能呼び出すことで通信を行う。提案手法は IP より下の層である。TCP, IP などのプロトコル処理が行われた後、送信パケットはネットワークデバイスドライバに対して送信命令を出す。提案手法はここに新しいドライバを準備しており（図中の Distributed NIC Driver）、このドライバはパケットを物理 Ethernet デバイスに送出する代わりに前述の SFCP プロトコルを用いて I/O サーバに転送する。そして、I/O サーバが帯域制御を行った上で I/O サーバの Ethernet インタフェースを用いてインターネットに送出する。

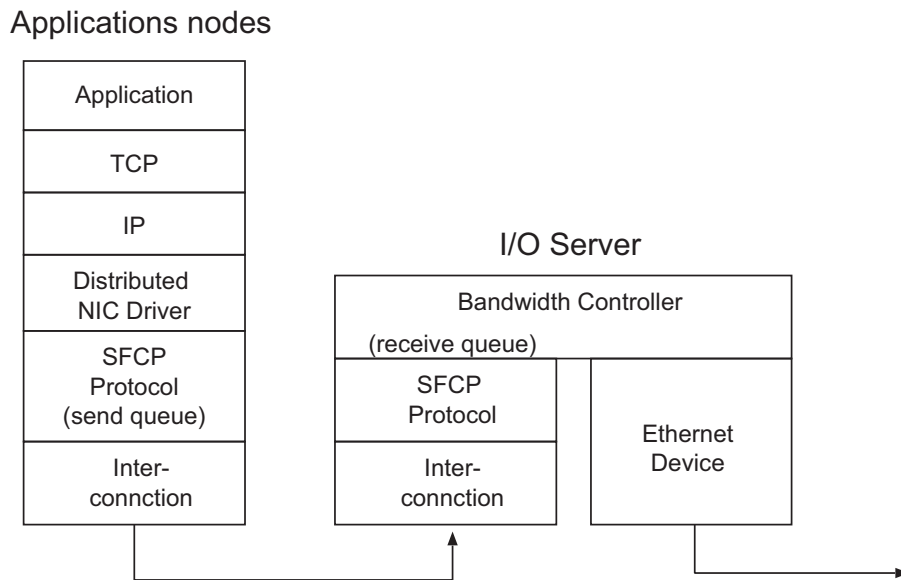


図3: 提案手法の設計

クラスター一般ノードにはネットワークデバイスドライバとして提案手法のドライバを組み込むのみであり、アプリケーションやオペレーティングシステムへの変更は不要である。これにより「背景」で述べたアプリケーション透過性が実現されている。実装は Linux オペレーティングシステムのカーネルモジュールとして行った。一般ノードはもちろん、I/O サーバの機能もカーネルモジュールとして実現されており、オペレーティングシステムに対する特殊な変更は不要である。

### 2.3 具体的成果

提案手法を図4の環境で評価した。クラスターとして3台のサーバコンピュータを用い、I/O サーバとの間を Myrinet-10G ネットワークで接続した。クラスター外部のネットワークとして、人工的な 10ms の通信遅延を挿入するルーターと一台のクライアントマシンを利用した。実験はインターネットで 500Mbps が利用可能であるという想定で行った。提案手法においては I/O サーバの帯域調整機能で 500Mbps に出力帯域を制限している。比較のため既存の方法の代表として、ネットワークスイッチの QoS 機能で 500Mbps の帯域制限を設定する方法を採用し、同じ評価プログラムを実行した。

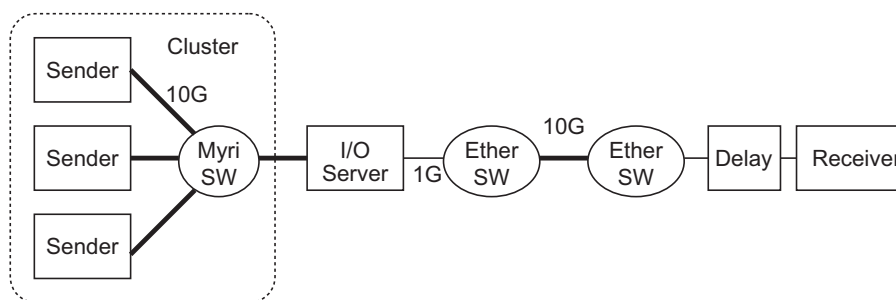


図 4: 実験環境

## バンド幅測定

ファイル転送のような連続的な通信パターンの代表として、連続ストリームによるバンド幅測定プログラムを用い、通信性能を評価した。このベンチマークでは3台のサーバから1台のクライアントに対して同時に通信を行い、外部ネットワークを流れるトラフィックの状態を5msごとに取得して通信の状況を監視した。前述のようにこのベンチマークは3本のストリームが500Mbpsの帯域を分けあうように設定されたネットワーク上で実行されている。

結果を図5に示す。この図に示されているように提案手法では設定された500Mbpsの通信速度が安定的に実現されている。一方、図6に示されている既存手法(QoSスイッチによる帯域制御)では上限の設定が500Mbpsであるにもかかわらず、実際に得られているのは最大でも200Mbps程度である。これはQoS機能がパケットを破棄し、TCPプロトコルがそれをネットワークの混雑とみなして送出レートを抑えるために発生する現象である。

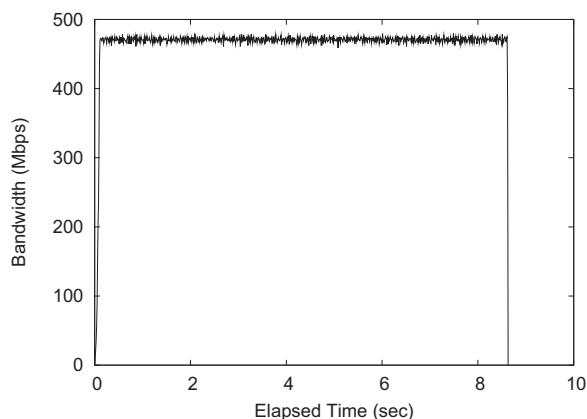


図 5: バンド幅テストの結果 (提案手法)

## HTTP サーバベンチマーク

単純送信のバンド幅が高いことは通信の基本性能として重要であるが、実際のネットワークは Web サーバからの送信データも多く含まれている。Web サーバからの送信データは、HTTP ヘッダのみの応答や、非常に小さなレイアウト用画像ファイルの転送が含まれており、バンド幅測定のような連続送信とは異なった通信パターンを持つ。このような通信パターンにおける提案手法の性能を確認するために、<http://www.cnn.com> のトップページを構成するファイル群のファイルサイズを調べ、それとまったく同じサイズのダミーデータを配置した3台の Web サーバからページ一式を繰り返しダウンロードするベンチマークを作成し、前節のバンド幅テストと同様の環境で実行した。

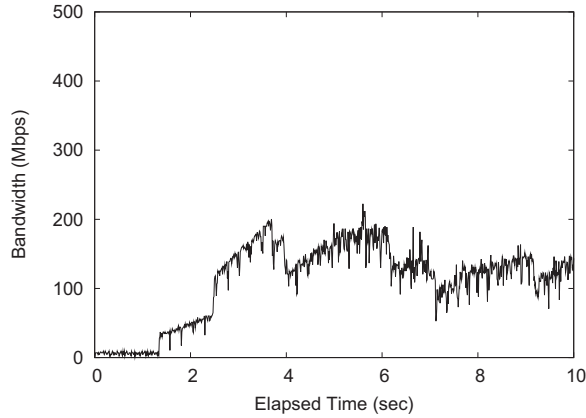


図 6: バンド幅テストの結果 (QoS スイッチによる帯域制限)

結果を表 1 に示す。Total Time は 3000 回トップページを取得するのに要した時間であり、提案手法が既存手法を 8.8%上回っている。Average Time は 1 ページを構成するファイル一式を取得するのに要した時間の平均であり、29.7%の高速化を実現していることが示されている。図 7 および 8 は通信の状況である。提案手法は設定した 500Mbps 近い性能が安定して得られており、QoS スイッチによる帯域制限よりも有効に与えられたバンド幅が活用できていることがわかる。

表 1: HTTP Server Performance Results

	Total Time	Average Time
提案手法	74.3 sec.	2.32 sec.
QOS スイッチ	80.9 sec.	3.01 sec.

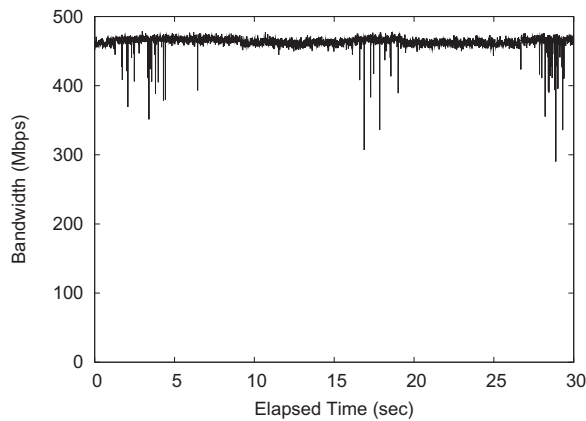


図 7: HTTP トラフィックのサンプリング (提案手法)

### 3 成果要覧

#### 査読付論文

[査読付 1] Hiroya Matsuba, Yutaka Ishikawa: Aggregate Router: An Efficient Inter-Cluster MPI Communication Facility, IPSJ Transactions on Advanced Computing Systems Vol.2 No.3 (ACS 27),



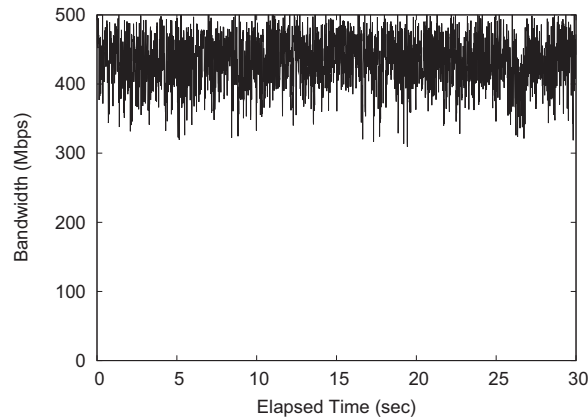


図 8: HTTP トラフィックのサンプリング (QoS スイッチによる帯域制限)

September 2009

- [査読付 2] 松葉 浩也, 堀 敦史, 石川 裕: 高性能クラスタのための高速汎用ステージングソフトウェア, 先進的計算基盤システムシンポジウム (SACSIS), pp. 239-246, May 2009
- [査読付 3] Kazuki Ohta, Hiroya Matsuba, Yutaka Ishikawa: Improving Parallel Write by Node-Level Request Scheduling, IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2009), 2009
- [査読付 4] Atsushi Hori, Yoshikazu Kamoshida, Hiroya Matsuba, Kazuki Ohta, Takashi Yasui, Shinji Sumimoto, Yutaka Ishikawa: On-Demand File Staging System for Clusters, IEEE International Conference on Cluster Computing (Cluster) 2009, 2009
- [査読付 5] Shinji Sumimoto, Kohta Nakashima, Akira Naruse, Kouichi Kumon, Takashi Yasui, Yoshikazu Kamoshida, Hiroya Matsuba, Atsushi Hori, Yutaka Ishikawa: The Design of Seamless MPI Computing Environment for Commodity-based Clusters, Euro PVM/MPI09, 2009

### その他の発表論文

- [発表 1] 住元真司, 中島耕太, 成瀬彰, 久門耕一, 安井隆, 鴨志田良和, 松葉浩也, 堀敦史, 石川裕: シームレスな MPI 環境を実現する MPI-Adapter の設計と性能評価, 情報処理学会研究報告 (SWoPP 2009), 2009
- [発表 2] 堀敦史, 鴨志田良和, 松葉浩也, 太田一樹, 安井隆, 住元真司, 石川裕: ファイルステージングシステム Catwalk の MPI-IO 実装, 情報処理学会研究報告 (SWoPP 2009), 2009
- [発表 3] 安井 隆, 清水 正明, 住元 真司, 太田 一樹, 鴨志田 良和, 松葉 浩也, 堀 敦史, 石川 裕: ファイルキャッシュシステムの有効性向上に向けた科学技術計算アプリケーションの I/O 特性評価, 情報処理学会研究報告 (SWoPP 2009), 2009

### 特記事項

- [特記 1] 松葉 浩也: A STUDY OF HIGH-QUALITY NETWORK TRANSMISSION ON COMPUTER CLUSTERS, 東京大学学位論文 博士 (情報理工学) 2010 年 3 月 12 日

# GPUを用いた高速数値計算とGPUプログラミング環境に関する研究

大島 聡史

## 1 概要

本報告ではGPU(Graphics Processing Unit)を用いた高性能計算に関する研究について述べる。

GPUは画像処理用のハードウェアとして発展してきた並列計算向けのハードウェアである。ビデオカードの形態で家電量販店やPCショップにて広く販売されているGPUは、HPCの分野においては並列演算性能の高い新たな計算ハードウェアとして近年注目を集めている。GPUはCPUと比較して1パッケージ(CPU1プロセッサおよびGPUカード1枚)あたり10倍以上高い理論演算性能を持つ。またGPUはCPUと比較して消費電力や占有体積が大きい、消費電力あたりや占有体積あたりの演算性能が高く、さらにCPUと同程度の価格で購入できる。これらの利点から、今日では様々なアプリケーションに対するGPUの利用が始まっている。GPUを用いた汎用演算はGPGPU(General-Purpose computation using GPUs)やGPUコンピューティングと呼ばれており、現在のHPC分野において注目を集めているトピックの一つとなっている。

GPGPUは一過性の流行を超えて広く利用されつつある技術である。今日ではGPUを搭載したPCの普及に伴いGPGPU計算環境が普及しているのみならず、GPUを特定の演算を高速化するためのアクセラレータとして搭載するスーパーコンピュータも登場している。そのため、現在本情報基盤センターが所有するスーパーコンピュータにはGPUは搭載されていないものの、スーパーコンピュータ向けのアプリケーションをGPUに移植することやGPGPU研究の推進および情報発信を行うことには意義がある。

本研究における活動は以下の3つに大別される。

- GPUを用いたアプリケーション高速化に関する研究
- GPU向けのプログラミング環境に関する研究
- GPGPUの啓蒙と教育に関する活動

## 2 GPUを用いたアプリケーション高速化に関する研究

### 2.1 背景

現在GPUは様々な対象問題においてCPUを超える性能を発揮している。しかしGPUはCPUと比べて並列処理に特化した特殊なアーキテクチャを採用しているため、

- 高い並列性を持つプログラムでないと高い性能が得られない
- 分岐処理を多く含むような複雑なプログラムでは高い性能が得られない
- 単精度浮動小数点演算性能に特化している(倍精度浮動小数点演算性能が低い、ただし改善されつつある)

といった問題(難しさ)を抱えている。そのため、全てのプログラムが GPU によって高速化できるわけではない。また、GPU はディスクやネットワークに直接接続できないハードウェアであり、アプリケーション全体の性能向上には CPU との適切な役割分担や協調処理が重要である。以上から、より多くのアプリケーションを GPU 上に実装すること、および実装の際に得られた最適化の知識や技術を蓄積および公開することは非常に重要である。

## 2.2 内容

アプリケーションの高速化と GPU 向け最適化技術の開発を目指して、GPU を用いたアプリケーション高速化を行った。本年度は E-サイエンスプロジェクトのライブラリグループに参加し、当グループで力を入れている対象問題の 1 つである疎行列ベクトル積の GPU 化に取り組んだ。

疎行列ベクトル積は、ゼロ要素を多く含む疎行列をベクトルと掛け合わせる問題である。疎行列ベクトル積の計算は密行列の行列ベクトル積と同様に行うことも可能であるが、大規模な疎行列を直接メモリに格納すると計算に不要な要素が多くなり、メモリ効率の低下や性能の低下を招いてしまう。そのためゼロ要素を削除して格納したデータ構造を用いて計算を行うのが一般的である。しかしデータ格納形式と対象問題の非ゼロ要素の配置によっては、計算時に不連続なメモリアクセスが多くなり、キャッシュの効果が得られずに高い性能が得られないことも多い。

一方 GPU は、不連続なメモリアクセスを非常に苦手としているハードウェアである。しかしながら疎行列ベクトル積の計算においては、データ格納形式と対象問題の非ゼロ要素の配置によっては不連続アクセスが多いが並列性は高いという CPU よりは GPU に適した計算が行われることもあり、GPU が CPU より高い性能を得られる可能性がある。

以上を踏まえて、GPU を用いた疎行列ベクトル積の高速化に取り組んだ。今年度は主に以下の 2 つの実装に注力した。

1. 現在広く普及している疎行列データ格納形式である CRS(Compressed Row Storage) 形式を用いた行列ベクトル積(CRS 法)の GPU 向け実装に取り組んだ。GPU における CRS 法の実装自体は既に行われている研究であるが、対象問題の形状と GPU 向け性能最適化手法(性能最適化パラメタ)の関係を確認および調査する目的で実装を行い、性能を測定した。
2. 特定形状の行列において特に高い性能が期待できる新しいデータ格納形式を用いた行列ベクトル積の GPU 化に取り組んだ。CRS 法は行単位での並列計算を容易に行うことができる。しかし、特定の行のみに非ゼロ要素が多いような問題形状においてはスレッド・プロセスごとの実行時間に大きなばらつきが生じてしまい、全体の実行時間が短くならないという問題がある。一方、SegmentedScan 法(以下 SS 法)および E-サイエンスプロジェクト内で開発された SS 法の改良法(BSS 法)は、計算を各スレッド・プロセスへ均等に割り当てることで並列実行時の性能を改善することができる。本研究では BSS 法の GPU 化を行った。

## 2.3 具体的成果

CRS 法の実装においては、GPU に対する問題割り当て(スケジューリング)方法の最適化や GPU が持つキャッシュが有効なメモリの活用などを行い、性能を調査した。これまでに、いくつかの対象問題においては GPU(GeForceGTX285)1 枚でマルチコア CPU(XeonW3520 4 コア)1 プロセッサを上回る性能を得ることに成功している。ただし、問題ごとに GPU に与えるべき最適な性能パラメタが違っており、また問題によっては CPU の方が高い性能を得られている。そのため、常に最大の性能を得るためには対象問題にあわせて GPU に与えるパラメタや計算対象の選択(CPU を使うか GPU を使うか)を行う必要がある。今後は最適化パラメタと性能の関連性をより詳細に分析し、常に最適なパラメタを選択する自動チューニングなどに取り組む予定である。

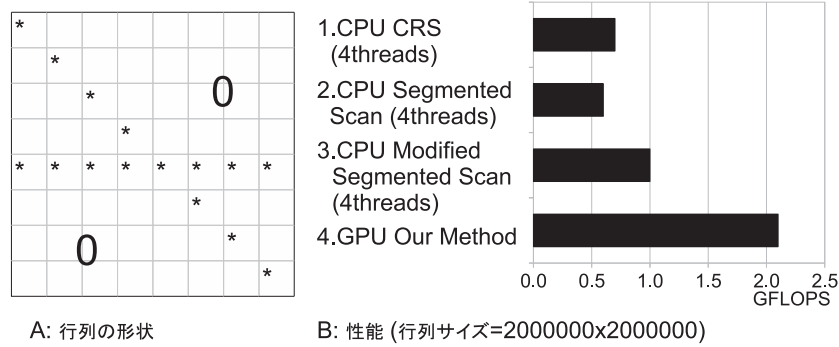


図 1: BSS 法の性能評価において対象とした行列の形状と性能

BSS 法の実装についても、BSS 法が得意とする問題形状において GPU1 枚でマルチコア CPU1 プロセッサを上回る性能を得ることに成功した (図 1)。本実装においては、並列計算に向いている計算のみを GPU に割り当てる CPU と GPU の問題分割と、GPU 上の論理的な並列度の最適化が性能に大きな影響を持つことを確認している。今後は GPU に対する問題割り当て (スケジューリング) の改善や、より高い性能を得られる最適化パラメタの調査などに取り組む予定である。

### 3 GPU 向けのプログラミング環境に関する研究

#### 3.1 背景

GPU を利用するためには、GPU を利用するように作られたライブラリを利用するか、自ら GPU プログラミングを行う必要がある。現在 GPU プログラミングには NVIDIA 社製 GPU 向けのプログラミング環境 CUDA が広く用いられているが、CUDA を利用するには GPU のアーキテクチャや実行モデルなどに関する理解が必要である。そのため、GPU は並列処理により高い性能を得られるハードウェアである一方、これまでに CPU 向けの並列処理を行ってきたプログラマにとっては習得や利用に手間がかかるハードウェアとなっている。

また CUDA を習得したプログラマにとっても、CUDA には様々な性能最適化手法や最適化パラメタがあるため、CUDA プログラミングには手間がかかるのが現状である。たとえば前章の疎行列ベクトル積においては、性能に影響を与えうるパラメタが多数確認できていたため、パラメタを調整して連続実行するスクリプトを作成して総当たりに性能調査を行った。CUDA には公式に提供されているプロファイラが存在するものの、これを用いて性能向上を達成するためにはアプリケーションプログラマ自身がプロファイル結果を分析しつつプログラムの修正と再実行を繰り返す必要がある。

#### 3.2 内容

GPU プログラミング環境の習得や利用 (最適化) の手間を削減することができるシステムの開発に取り組んでいる。なお本開発は JST CREST プロジェクト「ULP-HPC: 次世代テクノロジーのモデル化・最適化による超低消費電力ハイパフォーマンスコンピューティング」の一部として進めている。

SMP・PC クラスタ・スーパーコンピュータなど既存の並列計算機においては pthread・OpenMP・MPI などの並列化プログラミング環境が利用されてきた。そのため、既存の並列化プログラミング環境を用いた GPGPU プログラミングが可能となれば、多くのアプリケーションが容易に GPU 上で実行可能となることや、多くのプログラマが容易に GPU を利用可能となることが期待できる。

以上の考えに基づく具体的な実装として、GPU (CUDA) 向けの OpenMP 処理系 OMPCUDA を作成した。OMPCUDA は既存の CPU 向け OpenMP 処理系 OMNI OpenMP Compiler (以下 OMNI) をベースにしている。OMNI はプログラマの記述した OpenMP プログラムを独自の中間表現に変換し、中間



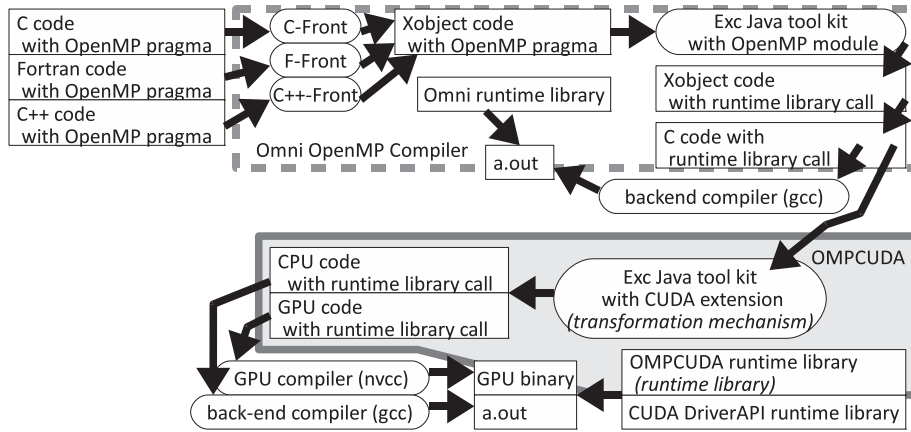


図 2: OMPCUDA の構造

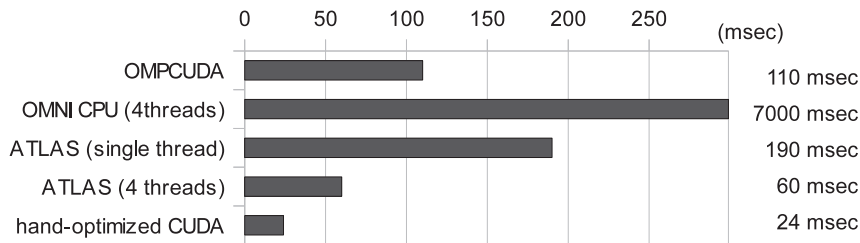


図 3: OMPCUDA の性能 (単精度行列積、行列サイズ 1024\*1024)

表現レベルでプログラム解析や書き換えを行いC言語プログラムを出力し、バックエンドコンパイラ (gcc など) によってコンパイルを行い、独自の実行時ライブラリ (スレッド関連の各種処理を担当) とリンクして実行可能ファイルを生成する (図 2)。OMPCUDA においては、OMNI の中間表現レベルでの処理を変更して CUDA 向けのソースコードを出力するプログラム変換機構と、CUDA の API を呼び出す独自の実行時ライブラリを実装し、OpenMP プログラムの並列実行部分を GPU 上で実行できるようにした。

さらに、OMPCUDA において OpenMP の指示子では容易に指定できない最適化パラメタの指定を手動で行えるようにするため、また通常の CUDA プログラムにおいて性能最適化パラメタの選択を容易にするため、新たな指示子を用いた最適化機構の検討を行っている。本件については具体的な成果としてはまとまっていないため、次節では記述を割愛する。

### 3.3 具体的成果

現在の OMPCUDA は典型的な OpenMP プログラムである for ループの並列化が可能である。性能確認のために同一ソースコードの単純な行列積プログラムを CPU と GPU それぞれで実行したところ、OMPCUDA は CPU よりも高い性能を得ることができた (図 3)。

一方で現在の OMPCUDA は OpenMP プログラムを単純に GPU へと割り当てているため、さらなる最適化の余地は大きい。例としては、CPU-GPU 間におけるデータ転送の削減に取り組む必要がある。現在の OMPCUDA は並列実行部分の前で全ての共有変数を送受信してしまうため、SPECOMP2001 の swim など、大きな配列を用いる並列実行部分を繰り返し実行するようなプログラムでは著しく性能が低下してしまう。今後も性能向上を目指して開発を続けていく予定である。

なお、OMPCUDA は 2010 年 6 月に筑波にて開催される国際ワークショップ IWOMP2010 への論文採録が決定している。(掲載が本報告の対象期間外のため成果要覧には掲載していない。) また ULP-HPC

プロジェクトにおいてはソースコードレベルではなくバイナリレベルでの共通化を目指す研究[発表2]も行っている。

## 4 GPGPUの啓蒙と教育に関する活動

### 4.1 背景

GPU(GPGPU)の利用報告数や研究報告数は年々増加している。しかしより多くのアプリケーションに活用されて成果を得るためには、またさらに利用者が増加し研究が一層活発になるためには、GPGPUの啓蒙と教育に関する活動が必要である。

### 4.2 内容

GPUプログラミングコンテストの実行委員として活動した。高校生を対象としたプログラミングスクールに招待され、特別講演およびプログラミング指導を行った。本センターの発行物「スーパーコンピューティングニュース」にてGPGPUの連載を行った。

### 4.3 具体的成果

社団法人情報処理学会が主催するシンポジウムSACISIS2010におけるGPUプログラミングコンテスト「GPUチャレンジ2010」の実行委員[特記1]として活動した。多数のチームに参加していただき、啓蒙と教育の効果は大きかった。(コンテストは2010年3月までに行われた。SACISIS2010自体は2010年5月開催である。)

高校生を対象とした文部科学省の委託事業「ICTスクール」にて特別講演[招待1]を行い、GPGPUの魅力や可能性、最新の研究トピック等について紹介した。GPGPUにより興味を持ってもらうことができ、啓蒙と教育の効果は大きかった。

本センターの発行物「スーパーコンピューティングニュース」にてGPGPUの連載を行った[特記2, 特記3]。2010年度も継続して連載を続ける予定である。

## 5 成果要覧

### 招待講演／招待論文

[招待1] 大島聡史: GPGPUの魅力と可能性と本当のところ, ICTスクール特別講演, 2009年12月27日.

### その他の発表論文

[発表1] 大島聡史, 平澤将一, 本多弘樹: OpenMP and auto-tuning framework for GPU, IWACOM-II(2nd International Workshops on Advances in Computational Mechanic), p.41, 2010.

[発表2] 平澤将一, 大島聡史, 本多弘樹: GPUコンピューティング向け中間言語の研究, pp.(後日発行予定), 2010.

### 特記事項

[特記1] GPUチャレンジ2010, SACISIS2010 - 先進的計算基盤システムシンポジウム, 2010年3月コンテスト終了.

[特記2] 大島聡史: これからの並列計算のためのGPGPU連載講座(I) GPUとGPGPUの歴史と特徴, スーパーコンピューティングニュース, Vol.12, No.1, pp.33-46, 2010.

[特記3] 大島聡史: これからの並列計算のためのGPGPU連載講座(II) GPGPUプログラミング環境CUDA入門編, スーパーコンピューティングニュース, Vol.12, No.2, pp.14-28, 2010.



# 高効率かつ透過なファイルステージングシステム

堀 敦史

## 1 概要

クラスタはコストパフォーマンスの高い科学技術計算用コンピュータとして広く使われるようになった。そのため、研究室レベルで使われる数十ノード程度の比較的小型なクラスタから T2K 東大にあるような千台規模のクラスタと幅広い性能のダイナミックレンジがある。そこで動くプログラムの大半は MPI 通信ライブラリを用いて記述されていることから、一見、プログラムの可搬性が高く、研究室レベルのクラスタで開発したソフトウェアを大規模クラスタの上で動かすことは容易であるように思われる。しかしながら実際には、システムソフトウェアのバージョンや、MPI ライブラリの細かい差異、並列ファイルシステムの違いなどが存在し、違いを全く意識せずに使う事はできないのが現状である。本研究では、クラスタのファイル IO に着目し、分散あるいは並列ファイルシステムの違いを隠蔽すると同時に、効率的な並列ファイルのアクセス手法について提案するものである。このため、実現にあたっては可搬性を重視し、使用に際し、システムの運用を中断したり、特権的な処理の必要がないことも重要となる。

昨年度は、既存の分散あるいは並列ファイルシステム上で動作するファイルステージングシステム、Catwalk、を研究開発した。これまでのファイルステージングシステムでは、ユーザが並列プログラムの実行に際し、必要なファイルを陽に記述する必要がある。このため分散あるいは並列ファイルシステムに比べ使い勝手の悪いものとなっていた。Catwalk のファイルステージングはオンデマンドで起動される。このため、ユーザが陽にステージングの記述をする必要がなく、分散あるいは並列ファイルシステムと同等の使い勝手を実現することができた。また、リング状の分散プロセス構造とすることで、既存の NFS 分散ファイルシステムよりも高い性能を発揮することが確認された。

本年度の研究では、並列ファイル書込に着目し、昨年度に開発した Catwalk に改良を施し、既存の並列ファイルシステムに比肩できる性能を引き出すことを目的とした。改良されたシステムは、MPI-IO の実装のひとつである ROMIO の ADIO デバイスとして実装された。実装された ADIO でバイスは Catwalk-ROMIO と呼ばれ、既存の並列ファイルシステム PVFS2 と Lustre と性能比較された。その結果、Catwalk-ROMIO では 1 Gbps の Ethernet で接続された 1 台のファイルサーバを用いているにも関わらず、10 Gbps の Myri10G で結ばれた複数のファイルサーバで構成される PVFS2 や Lustre を上回る性能を発揮する場合があることが確認された。

## 2 背景

並列ファイルシステムは、基本的に複数のファイルサーバが計算ノードと高速なネットワークで接続されている環境で、ファイルサーバの台数効果による高速化を目指すものである。一方で、並列アプリケーションからの並列ファイルアクセスは、多くの場合、並列ジョブの個々のプロセスが対象とするファイルの異なった部分を同時にアクセスするため、不連続なアクセスとなる。また並列度が高くなるにつれ、一回のアクセスの粒度が小さくなる傾向もある。ところがディスク単体の性能は、大き

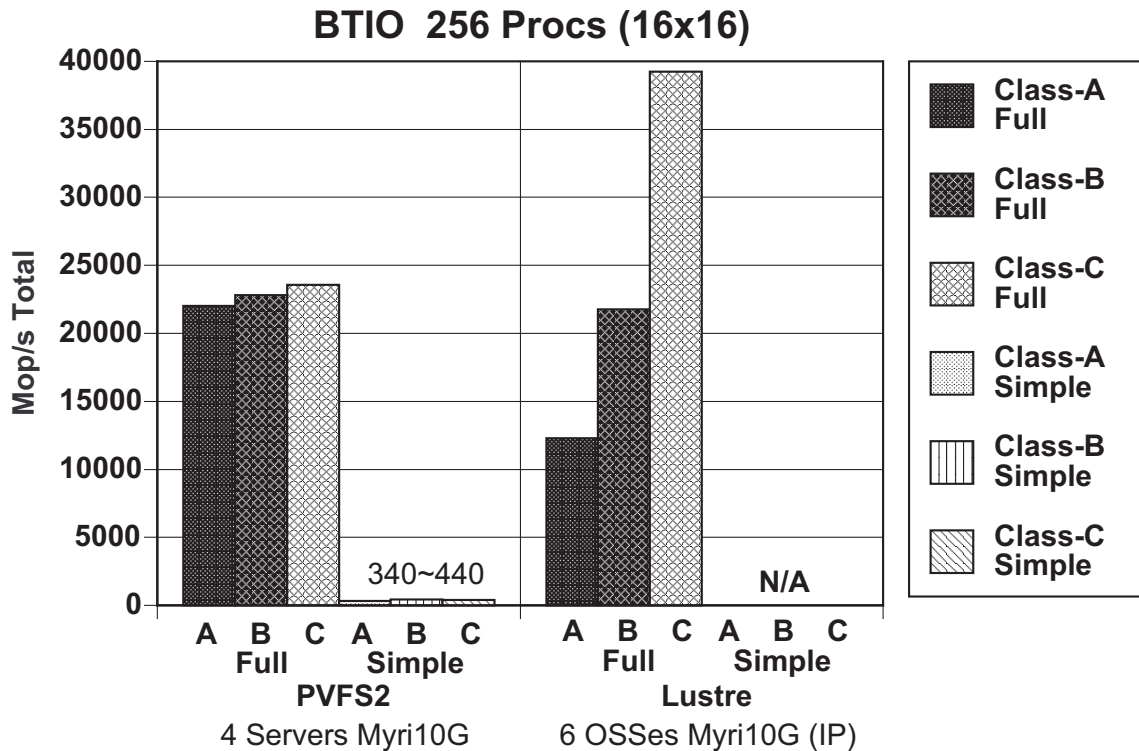


図 1: PVFS2 と Lustre の BTIO 性能

な粒度での逐次アクセスが最も性能が高い。このため、並列アプリケーションが必要とする並列 IO と、実際に記録されるディスクの間には大きなギャップが存在し、高い並列度における並列ファイルアクセスを高速化するには、なんらかの工夫が必要となる。

この問題は比較的古くから良く知られており、two-phase I/O や data sieving などが代表的な研究として良く知られている。これらの技法は MPI-IO の ROMIO に既に組込まれている。図 1 は、ROMIO 経由で PVFS2 上のファイルに書込んだ場合の性能と、ROMIO を経由しないで Lustre 上のファイルを並列書込んだ場合の性能を、並列ファイル IO ベンチマークのひとつである BTIO (NAS 並列ベンチマーク) を用いて計測した結果である。

BTIO は問題の大きさにより A、B、C と 3つのクラスがあり、クラス A で 419.43 MiB、クラス B で 1697.93 MiB、クラス C で 6802.44 MiB のファイルを書き出す。またファイルのアクセス方法についても simple と full の 2種類があり、simple では非 collective な 並列 IO、full では collective な 並列 IO が用いられている。図 1 の計測では T2K (東大) の計算ノードを 16 台用いた。PVFS2 では計算とは別な計算ノード 4 台をファイルサーバとして用い、Lustre では専用の 6 台のファイルサーバ (OSS) を用いている。また PVFS2 および Lustre において、計算ノードとファイルサーバは Myri10G (10Mbps) を用いている。

このグラフから、PVFS2 および Lustre とも full アクセス方式での性能が高いが、simple での性能が全く出ていない。特に Lustre では ROMIO の内部エラーとなってしまう計測不可能な状況であった。BTIO-full では、前述した tw-phase I/O や data sieving の技法を用いて最適化がされているため、高い性能を発揮することが出来ているが、これらの技法は collective な IO 操作においてのみ適用可能で、非 collective な IO 操作で記述された simple の場合では、全く性能が引き出せていないことが分かる。換言すれば、PVFS2 や Lustre では、大きな粒度で整列された並列書込に対して性能が出るようにチューニングされていると言うことができる。

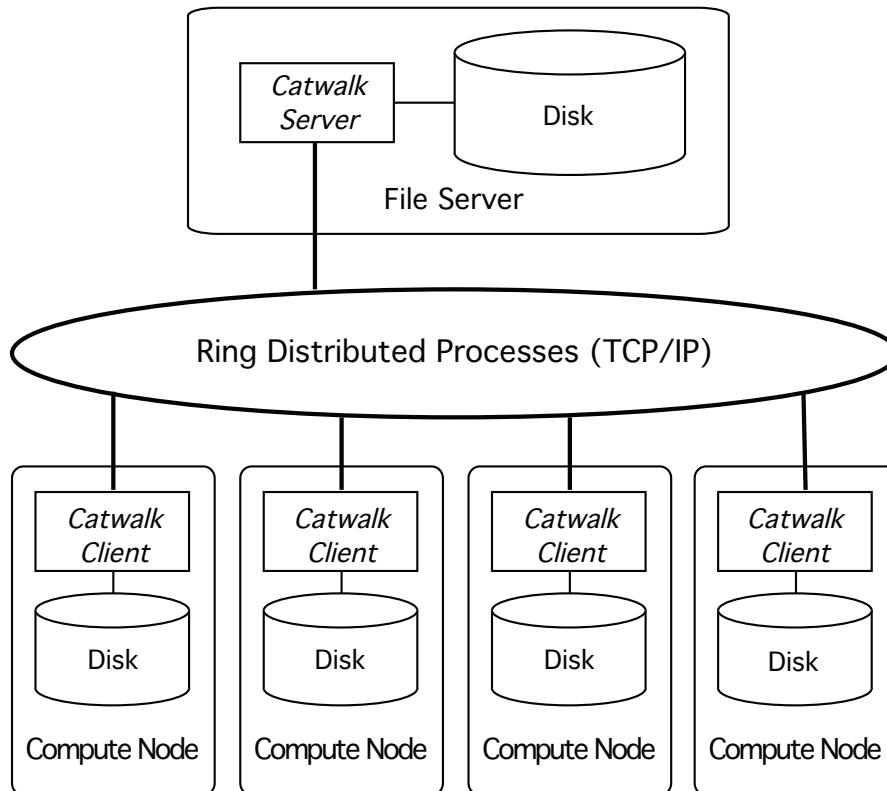


図 2: Catwalk-ROMIO の分散リングプロセス構造

### 3 内容

以下、Catwalk-ROMIO の実装の詳細と、その性能を PVFS2 および Lustre と比較した評価結果について記す。

#### 3.1 実装

Catwalk-ROMIO は、図 2 に示すようなリング状の分散プロセス構造を持つ。ステージングの元となるファイルサーバおよびステージングの先となる計算ノード上にそれぞれひとつの Catwalk プロセスが生成される。隣接するプロセスとは TCP/IP により通信が可能になっている。これらのプロセスは MPI-IO の `MPI_File_open()` 関数呼出しにより生成され、`MPI_File_close()` 関数呼出しにより消滅する。以下の説明で、ファイルサーバ上の Catwalk-ROMIO プロセスを「サーバプロセス」、計算ノード上の Catwalk-ROMIO プロセスを「クライアントプロセス」と呼ぶ。

Catwalk-ROMIO の読込処理を図 3 に示す。この処理は、`MPI_File_open()` 時にリングプロセス構造を作るという点を除いて、Catwalk と同じである。

1. ユーザプロセスが `MPI_File_open()` 関数により読込 `open` を要求した場合、図 2 にあるような Catwalk-ROMIO の分散プロセス構造が生成される。
2. ファイルの読込 `open` の要求がリングに沿って同期され、最終的にサーバノードに到達する。
3. サーバノードでは、要求のあったファイルを開き、その内容をリングに沿って転送する。
4. クライアントプロセスは受け取ったデータを次のクライアントに渡すと同時にローカルディスクのファイルに書込む。

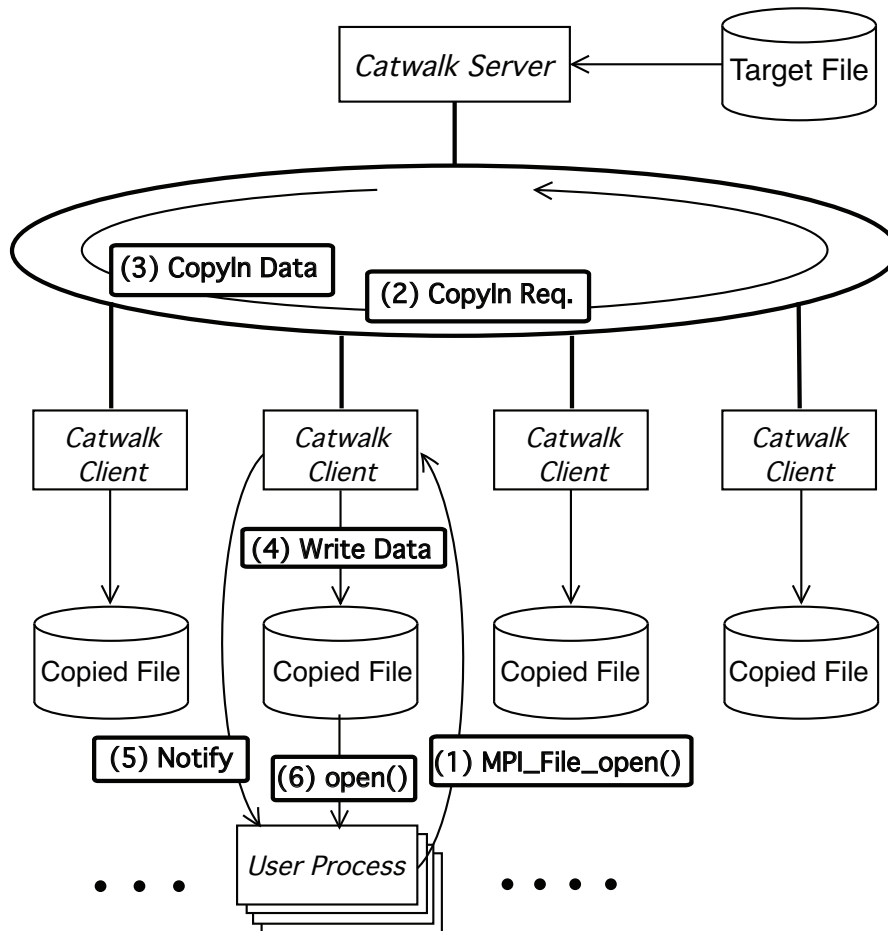


図 3: 並列ファイルの読込手順

5. 全てのデータを受け取ったクライアントプロセスは、ローカルディスクのファイルを close し、終了をユーザプロセスに通知する。
6. ファイル転送の完了を受け取ったユーザプロセスは、新たに生成されたローカルディスクのファイルを open し、MPI\_File\_open() 関数が終了する。以後、このファイルに対する読込は全てローカルディスクにコピーしたファイルへのアクセスとなる。

Catwalk-ROMIO をリング状の分散プロセス構造としたのは、1) ファイルのブロードキャストをパイプライン処理しバンド幅を稼ぐ、2) 逐次的な処理のためプログラミングが比較的容易、という理由からである。このようなパイプライン処理では計算ノード数に比例したパイプライン段数となり、ファイルサーバからのデータ出力バンド幅が一定にも関わらず、理論上はファイルが十分に大きければ計算ノード数に比例したバンド幅が得られるという利点がある。また、バッファキャッシュの flush による影響を避けるため、計算ノード上で書込むファイルを O\_DIRECT フラグで open している点も同じである。

Data sieving の手法では、ファイルの読込要求が細粒度で不連続な場合、それらを包含するようなより大きな粒度かつ連続な範囲で、実際にディスクからデータを読み込み、実際に必要とされる領域のみをユーザプロセスに渡すという方式である。この観点で Catwalk-ROMIO での読込方式を見ると、複数の読込要求を包含する単位が、ファイル全体になっていると見なすことができよう。

Catwalk-ROMIO の各ノードにおける読込処理を図 4 に示す。図 2 に示したリング分散プロセス構造は、読込処理と同様、MPI\_File\_open() 関数呼出時に生成され、MPI\_File\_close() 関数呼出

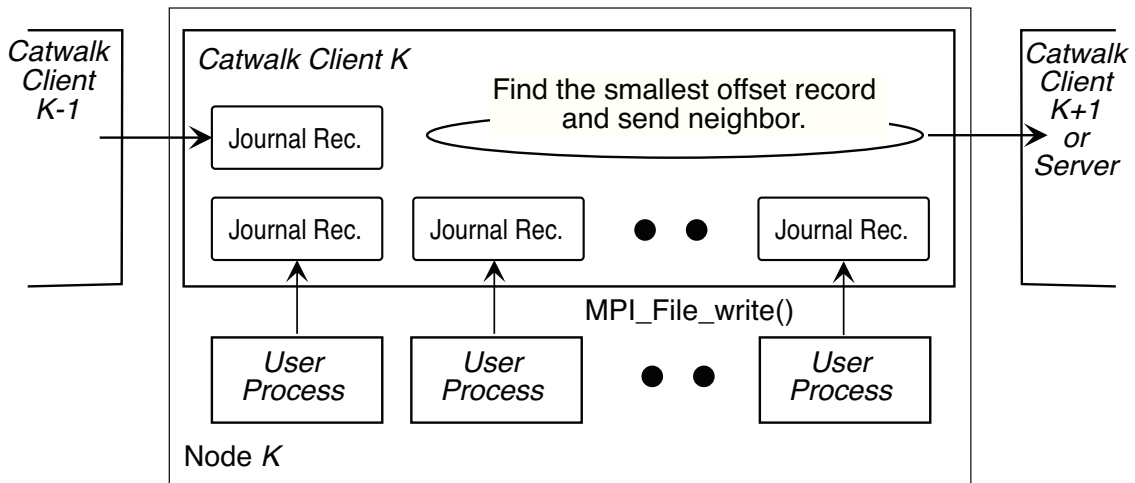


図 4: 並列ファイルの書込手順

しで消滅する。

ここでひとつの計算ノードでの処理に注目する。複数のユーザプロセスは `MPI_File_write()` 関数により並列ファイルへの書込要求が発行される。この要求はクライアントプロセスに伝えられ、ジャーナルデータとして取り扱われる。ノード内の複数のプロセスからのジャーナルデータと、リングトポロジ的にひとつ前のプロセスからの要求は、内部の有限バッファにいったん蓄えられる。この時、書込要求のファイル内オフセットの値に従って、ジャーナルデータはソートされる。また、連続する複数の書込要求は、可能ならばひとつのジャーナルデータとしてマージされる。このようにユーザプロセスからの書込要求をソートあるいはマージすることで、最終的にファイルサーバ上でのファイルへの書込ができるだけ大きな粒度かつ連続になるようにしている。

残念ながら、このバッファは有限長であるので、溢れが生じる可能性がある。この時、バッファ内にある最もオフセットが小さいジャーナルデータをリングトポロジ的に次のノードに送る。`MPI_File_close()` が呼ばれた時点で、バッファ内の全てのジャーナルデータはリングの先頭から順に次のノードへと転送され、またこのときに可能であればソートとマージを行い、さらに次のノードへと転送し、最終的に書込の全てのデータはファイルサーバに送られる。有限バッファを用いて書込要求をソートすることから、最終的に得られるソートの結果は完全とはならない場合がある。しかしながら、目的は不連続な書込の回数を減らすことなので、不完全なソートでも問題とはならない。

前述した two-phase I/O では、並列ファイルの IO を、最初のフェーズで各プロセスからの要求を集め、可能な範囲でソートあるいはマージし、次のフェーズで複数のファイルサーバに書込まれる。この時、複数の IO 要求は collective な操作の範囲でソート、マージが行われる。一方、Catwalk-ROMIO の場合では、内部バッファに要求が順次蓄えられるため、必ずしも collective な要求である必要がない。また、要求のソートおよびマージはリングトポロジに沿ったパイプライン的に処理され、処理が明確なフェーズに分かれていないという点でも two-phase I/O と異なっている。

### 3.2 評価

図 5 は先の図 1 に、同じ環境で計測した Catwalk-ROMIO のグラフ（右端）を追加したものである。Full の場合では、Catwalk-ROMIO は、PVFS2 や Lustre に比べ、ファイルサーバが 1 台で、かつネットワークも 1/10 のバンド幅しかないという貧弱なハードウェア環境でありながら、PVFS2 や Lustre のおよそ数分の 1 程度の性能を引き出している。一方、simple の場合では、計測不能であった Lustre は言うに及ばず、PVFS2 の 10 倍以上の性能を引き出していることが分かる。

このことから、Catwalk-ROMIO は、細粒度の不連続な書込要求パターンを、リングトポロジに沿った



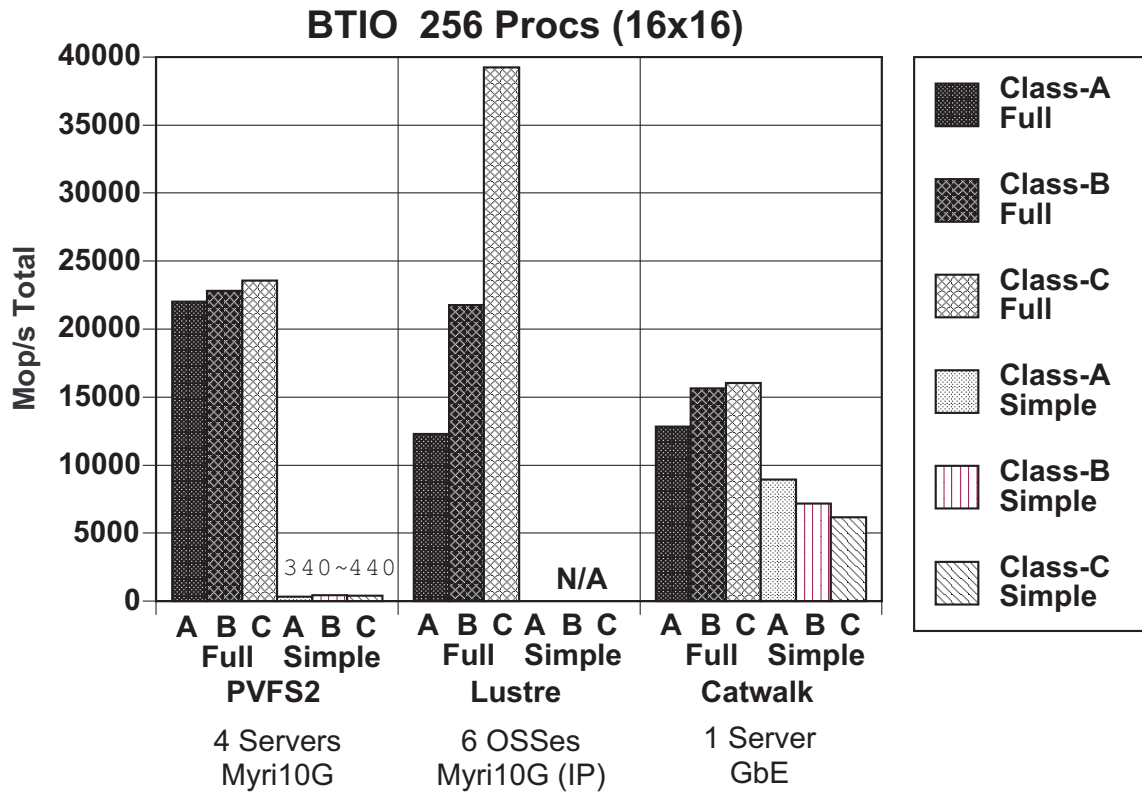


図 5: Catwalk-ROMIO の BTIO 性能

パイプラインにより旨く改善し、より大きな粒度で、より連続な書込要求パターンに変換することに成功していると考えられる。Catwalk-ROMIO は、もともとの設計条件として、Catwalk-ROMIO が用いるファイルシステムに依存しない。これまでの計測において Catwalk-ROMIO の下位ファイルシステムとして Linux の（逐次）ファイルシステムを用いてきた。ここで、simple で計測不能だった Lustre を Catwalk-ROMIO の下位ファイルシステムとすることで、改善できるかどうかを調べたのが図 6 のグラフである。

このグラフの真ん中が、Lustre を下位ファイルシステムとした場合の BTIO の性能である。比較のため、左のグラフは Lustre 単体、右のグラフは Catwalk-ROMIO 単体での性能である。注目すべきは full の場合でも、問題サイズが小さいクラス A や B で Lustre の性能が大きく向上しているのが分かる。また、計測不能であった simple の場合でも、Catwalk-ROMIO 単体での性能には及ばないものの、図 5 の PVFS2 よりも高い性能となった。この結果から、Catwalk-ROMIO によりファイルアクセスパターンを改善することができた、ということができる。

以上の評価結果から、Catwalk-ROMIO は、より貧弱なハードウェア環境においても、専用のハードウェアを用いた並列ファイルシステムに比肩できる性能を示すことができたと考えられる。ハードウェアの依存性が少ないということは、Catwalk-ROMIO が有効に活用できる間口が、PVFS2 や Lustre よりも広いと考えることができよう。

Catwalk-ROMIO の読込はファイルを全て各計算ノードに転送する。一方、並列ファイルシステムでは、必要な部分だけを各計算ノードに転送する。このため読込に関しては Catwalk-ROMIO の方が不利である。また BTIO はファイルの並列書込しか計測しない。このため、これまでに示した結果に、ファイルの読込時間は全くカウントされていない。しかしながら、BTIO は計算や IO の結果が正しいかを検証するために、書き出したファイルを計測後に読込む。このため、BTIO が出力する計測時間ではなく、プログラムの実行時間全体を計測することで、ファイルの読み書き性能を比較することが



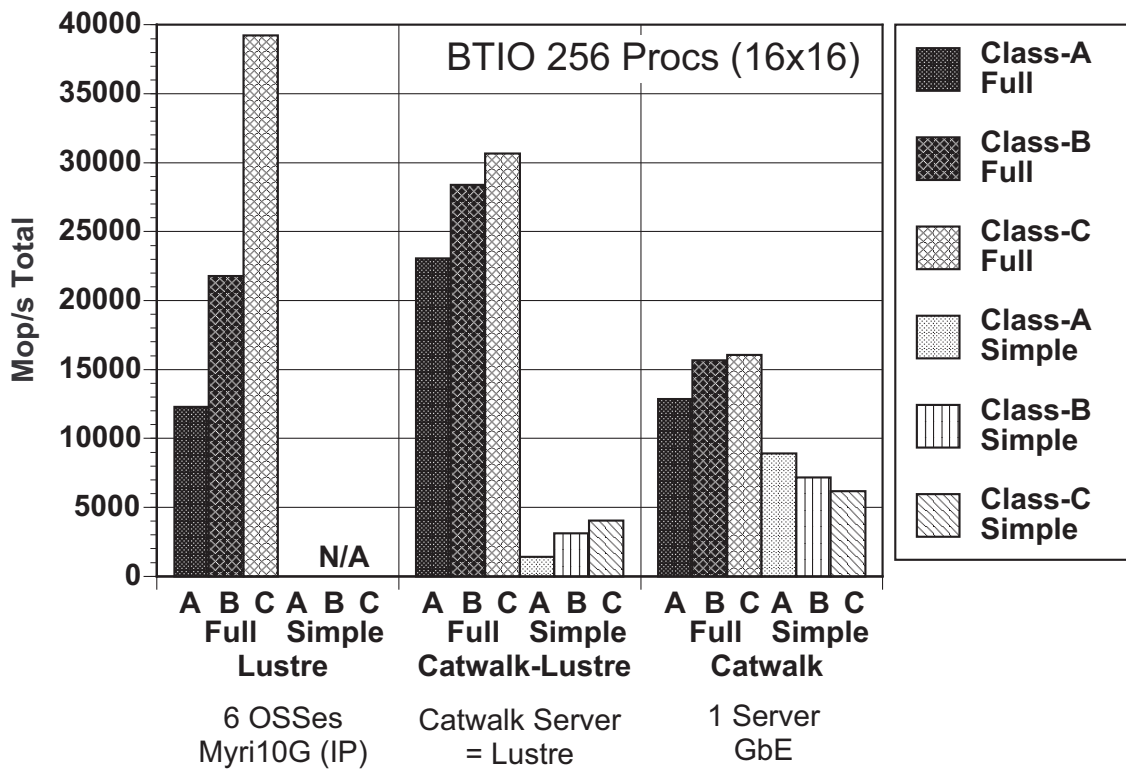


図 6: Lustre 上での Catwalk-ROMIO の BTIO 性能

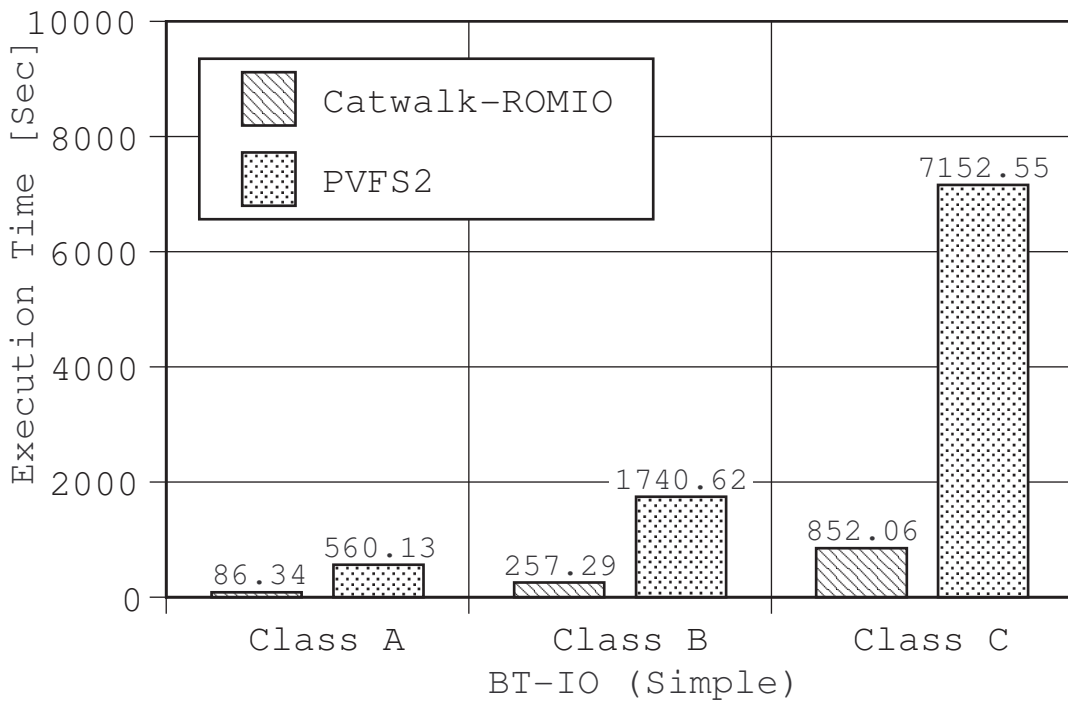


図 7: BTIO 総実行時間の比較

できる。

図 7 は PVFS2 と Catwalk-ROMIO で BTIO (Siple) の実行時間を比較したものである。このグラフでは縦軸が実行時間であるため、小さい程性能が良いことになる。このグラフから分かるように、Catwalk-ROMIO は読込時間を含んでも、まだ PVFS2 より高い性能を発揮していることが確認された。

### 3.3 具体的成果

本研究で開発されたソフトウェア Catwalk および Catwalk-ROMIO は PC クラスタコンソーシアム<sup>1</sup> (情報基盤センターもコンソーシアムの会員である) で開発されている SCore クラスタシステムソフトウェアに組み込まれ、オープンソースとして一般に配布されている [公開 1]。また、Catwalk および Catwalk-ROMIO については SCore の新機能として、PC クラスタコンソーシアムが主催するシンポジウム [特記 1] やワークショップ [特記 3] で、既に発表済みである。また Catwalk および Catwalk-ROMIO の使い方に関するチュートリアル [特記 2] もおこなった。

## 4 成果要覧

### 査読付論文

[査読付 1] Atsushi HORI, Yoshikazu KAMOSHIDA, Hiroya MATSUBA, Kazuki Ohta, Yutaka ISHIKAWA: On-Demand File Staging System for Linux Clusters, IEEE Cluster 2009, 2009.

[査読付 2] 堀敦史 鴨志田良和 松葉浩也 安井隆 住元真司 石川裕: リングトポロジーによる MPI-IO 書込の高速化技法, SACSIS 2010 (2010 年 5 月発表予定) .

### 公開ソフトウェア

[公開 1] Atsushi HORI: SCore Cluster Sytem Version 7, <http://www.pccluster.org/ja/score7.html>, Feb 2010.

### その他の発表論文

[発表 1] 堀敦史 鴨志田良和 松葉浩也 安井隆 住元真司 石川裕: ファイルステーキングシステム Catwalk の MPI-IO 実装, SWoPP2009, 2009.

[発表 2] Atsushi HORI: Parallel File IO Using Ring Communication Topology, WPSE2010, Feb.2010.

### 特記事項

[特記 1] 堀敦史: SCore7 最新状況, 第九回 PC クラスタシンポジウム講演, 2008 年 12 月 10 日.

[特記 2] 堀敦史: Catwalk Tutorial, PC クラスタワークショップ in 京都講演, 2010 年 2 月 18 日.

[特記 3] 堀敦史: SCore7 最新状況, PC クラスタワークショップ in 京都講演, 2010 年 2 月 19 日.

<sup>1</sup><http://www.pccluster.org>

# ソフトウェア自動チューニングおよび 高性能数値計算ライブラリの研究と HPC 教育

片桐 孝洋

## 1 概要

本報告は、以下の研究・教育・業務報告に大別される。

- ソフトウェア自動チューニング研究
  - 研究室の PC から国家が提供するスーパーコンピュータに至る広範な計算機環境で、自動的に性能チューニングができるソフトウェア(ソフトウェア自動チューニング)の理論、方式、言語、およびライブラリに関する総合的研究。
- ペタコンに向けた超並列高性能数値計算ライブラリ研究開発
  - 対称密行列用の超並列固有値ソルバの研究。
  - アルゴリズムと実装方式を自動チューニングする機能を有する、疎行列反復解法ソルバの研究開発(e-サイエンスプロジェクトによる研究開発)。
- 高性能計算(HPC)における学部・大学院教育およびスーパーコンピュータユーザ教育業務
  - センターのスーパーコンピュータを用いて行う並列プログラミングの授業。
    - ◇ 東京大学工学部・工学系研究科共通科目:スパコンプログラミング(1)(I)
    - ◇ 駒場の教養学部学生を対象に行う全学ゼミ:スパコンプログラミング研究ゼミ
  - 並列プログラミング教育および新規ユーザ獲得支援業務。
    - ◇ 若手利用者推薦(試行)
    - ◇ Advanced Supercomputing Environment(ASE)研究会
    - ◇ お試シアカウント付き並列プログラミング講習会(試行)
  - 共同研究プロジェクト
- スーパーコンピュータによる社会貢献業務
  - 文部科学省 研究開発施設共用等促進費補助金(先端研究施設共用促進事業)「先端的大規模計算利用サービス」
  - 東京大学情報基盤センター企業利用支援制度

## 2 ソフトウェア自動チューニング研究

### 2.1 背景

近年の数値計算ソフトウェアは、性能に影響を及ぼす計算機上やアルゴリズム上のパラメタが多数存在する。これは平易な利用の妨げになるばかりか、高性能を達成するため時間を浪費し、職人芸的な知識を必要とする。この背景から、本研究では以下を目的とする。

1. ソフトウェア工学的観点から、ソフトウェア自動チューニング(以降、単に自動チューニング、ATとよぶ)という新しい汎用的ソフトウェア・パラダイムの研究をおこなう。AT研究により、パラ

メタチューニング作業が自動化され、最適プログラムが自動選択される。このことで、ソフトウェアの再利用性を格段に高め、コスト削減を実現することができる。

2. AT機能を付加した実用的な数値計算ソフトウェアの研究開発を行う。日本発となる、自由入手可能なソフトウェアの普及を目指す。
3. AT機構をソフトウェアに実装するための計算機言語、ミドルウェア、システムソフトウェア、およびAT支援ツールの研究開発を行う。

## 2.2 内容

本研究により、数値計算や並列処理を専門としない利用者に対し、容易に利用できる高性能数値ソフトウェアが初めて開発できるようになる。世界に先駆け、数値計算ライブラリに限定しないATパラダイムやAT技術の提案を行ってきた。

## 2.3 具体的成果

本研究は2002年から継続して成果を出してきている。2009年度の主な研究成果は以下である。

1. **疎行列ソルバの実行時AT方式研究[発表1、8、11]**: 疎行列反復解法は、入力疎行列の形状により最適な実装方式と数値アルゴリズムが異なる。数値計算ライブラリのレベルでは、実行時でないユーザーが入力する疎行列形状が不明であり最適化ができない。そこで本研究は、実行時のAT機能を数値計算ライブラリに実装することで、従来では達成不可能な高速化を実現した。本研究は、愛媛大学黒田久泰准教授との共同研究である。
2. **汎用的ATインターフェースおよびATライブラリの開発(e-サイエンスプロジェクト)**  
[招待4][査読付2][発表6、7、16、17、18]: いままでのAT機能付きライブラリでは、汎用的なユーザーインターフェースが存在しないので、AT機能の再利用性が低い。そこで、汎用的ATインターフェースOpenATLibを設計し、ATに必要な機能の一部を実現した。本研究は、文部科学省、e-サイエンス実現のためのシステム統合・連携ソフトウェアの研究開発、「シームレス高生産・高性能プログラミング環境」、サブテーマ「高生産高可搬性ライブラリに関する研究」(研究代表者: 石川裕教授)の一環で行われた。本研究は、日立製作所中央研究所との共同研究である。
3. **超省電力数値計算ライブラリの研究[発表10]**: 低電力を実現するシステム上において、数値計算ライブラリレベルで行う省電力化方式の基礎研究である。なお本研究は、JST-CREST、「ULP-HPC: 次世代テクノロジーのモデル化・最適化による超低消費電力ハイパフォーマンスコンピューティング」(研究代表者: 東京工業大学 松岡聡 教授)の支援により、愛媛大学黒田久泰准教授、東京大学須田礼仁教授との共同研究である。
4. **ペタフロップス環境に向けた超並列固有値ソルバの研究開発[招待6][査読付1、3]**  
[発表2]: ペタフロップス環境は10万コア以上を有する超並列環境である。このような超並列環境に向けた固有値ソルバのアルゴリズム、実装方式、通信アルゴリズム、AT機能を総合的に研究する。

情報処理学会学会誌「情報処理」2009年6月号において、「科学技術におけるソフトウェア自動チューニング」特集号[特記5]が出版された。また、文部科学省のシンクタンクの特集記事として取り上げられる[報道2]など、日本国内においてAT分野の研究の注目度は高い。また、米国DoEにおいてATをテーマとするサマーセミナーに、日本人研究者で唯一招待される[招待4]など、海外においても我々のグループによるAT研究は注目されている。

### 3 高性能数値計算ライブラリ研究開発

#### 3.1 背景

深い階層のメモリ構造と非均質なメモリアクセスのマルチコア CPU が広く浸透し、PC でもマルチコア CPU の利用は避けられない。複雑化した計算機環境では、コンパイラ最適化のみに依存することはできない。人手によるチューニングがいまだになされているが、性能挙動が複雑になり人手のチューニングさえ困難となってきた。一方、研究室レベルの PC クラスタから国家レベルで提供されるスーパーコンピュータまで、コードを改変することなく透過的に移植でき性能維持ができることは、低いコストのソフトウェア構築に不可欠な技術項目となる。本研究では数値計算ライブラリの AT を取り扱うが、以下の2点を目的とする。

第1に、従来の AT ライブラリは AT 機能を搭載しているが、別の開発者が類似の AT 機能を新規開発する場合、AT 機能のプログラム再利用性がない。そこで本研究では、AT 機能の Application Programming Interface (API)を規格化することで AT ライブラリにおけるコード再利用問題の解決を目的とする AT 用 API ライブラリ OpenATLib を開発した。

第2に、従来の AT ライブラリでは実行速度のみ最適化し、利用メモリ量や演算精度に対しての AT 適用が不可能であった。そこで我々は、エンドユーザにより指定される AT 方針である「数値計算ポリシー」の設定機能を新規開発した。この機能説明と効果について一部紹介する。

#### 3.2 内容

##### (1) OpenATLib:汎用的自動チューニングAPI

OpenATLib は数値計算ライブラリにおける AT 機能を規格化して API 化し、その参照実装を提供することを目的に開発された。現在、疎行列ソルバに特化した AT 機能、および実行時の AT 機能を中心に機能開発を行っている。我々は既に開発した OpenATLib  $\alpha$  版に対し、機能を高度化した OpenATLib  $\beta$  版を開発した。この  $\beta$  版の主要関数と機能説明を表1に示す。

表1 OpenATLib  $\beta$  版機能

関数名	機能の説明
OpenATI_DAFRT	Krylov部分空間のリスタート周期を増加するか判定
OpenATI_DSRMV	CRS形式において、倍精度演算 対称行列用の疎行列ベクトル積で、最もよい実装方式の選択
OpenATI_DURMV	CRS形式において、倍精度演算 非対称行列用の疎行列ベクトル積で、最もよい実装方式の選択
OpenATI_DSRMV_Setup	OpenATI_DSRMVのための初期データ設定処理
OpenATI_DURMV_Setup	OpenATI_DURMVのための初期データ設定処理
OpenATI_DAFGS	Gram-Schmidt直交化処理において、4種類の実装から最もよい実装方式の選択
OpenATI_BLDATA	デフォルトデータ設定処理 (Fortran言語のBlock data format)
OpenATI_LINEAR SOLVE	数値計算ポリシーを適用する連立一次方程式ソルバ用メタ・インタフェース
OpenATI_EIGENSOLVE	数値計算ポリシーを適用する固有値ソルバ用メタ・インタフェース

OpenATLib  $\alpha$  版の最大の特徴は、いくつかの疎行列反復解法中に現れ収束性と実行速度に大きく影響するアルゴリズム・パラメタである<リスタート周期>を、残差履歴を実行時にモニタすることでリスタート周期の増減を実行時に決定する API である OpenATI\_DAFRT の提供にあった。



OpenATL\_DAFRT の判断基準は、MM比という基準を用いる。β版についてもこの機能について変更は無いが、以下の新規機能を追加している。

- 疎行列-ベクトル積 API: OpenATL\_D(S|U)RMV

自動的に選択される実装方式として、以下の機能を新規開発した。

1. 非零要素数を考慮し並列実行時の計算負荷をバランス化させる方式(対称・非対称)
2. 零要素について並列実行時の加算を省く方式(対称)
3. 新方式の Branchless Segmented Scan 方式(非対称)

- Gram-Schmidt 直交化 API: OpenATL\_DAFGS

1. 古典 Gram-Schmidt (CGS)
2. DGKS
3. 修正 Gram-Schmidt (MGS)
4. ブロック化古典 Gram-Schmidt (BCGS)

- 数値計算ポリシーを適用するためのメタ・インターフェース:

OpenATL\_LINEAR SOLVE と OpenATL\_EIGEN SOLVE

(2) 演算ポリシーと演算精度ポリシー指定

OpenATLib β版では、エンドユーザからの AT 方針の指定機能である数値計算ポリシー機能が提供されている。この数値計算ポリシー機能は、メタ・インターフェース API の 2 種に対して作用する。提供機能として、演算速度、メモリ量、演算精度の 3 種が指定可能であり、AT 戦略が異なる。

ここでは、エンドユーザにより演算精度最適化が指定された場合の処理(演算精度ポリシーの AT 戦略)について解説する。演算精度ポリシーの AT 戦略は、以下にまとめられる。

- I. ソルバ内で利用される収束判定値 EPS は、ユーザが与えた要求精度(RAP)について最初の呼び出しについて同一(EPS=RAP)とする。
- II. ソルバ内で収束後、真の残差(RRV)を計算する。ここで真の残差とは、解ベクトルを  $x$ 、行列を  $A$  とすると連立一次方程式の場合は  $|Ax-b|/|b|$ 、固有値問題の場合は  $\lambda$  を固有値とすると  $|Ax-\lambda x|/|\lambda|$  の最大値(複数固有値・固有ベクトルを求める場合)である。
- III. もし、 $RAP < RRV$  である場合、 $EPS = EPS / 10$  として再度ソルバを呼び出す(これを、精度反復改良ループとよぶ)

なおソルバ中に用いられる Gram-Schmidt 直交化方式は、直交精度が高い DGKS 法を精度反復改良ループで強制的に採用する。以上の I~III の AT 戦略は、実行時間が増加する可能性がある。ただしエンドユーザからの要求誤差を(収束する行列であれば)必ず満たす戦略であり、演算精度要求にかなうものである。一般に、 $RAP < RRV$  になることを「偽収束」と呼ぶが、本 AT 戦略により(ソルバが正しく収束する問題である限り)偽収束は生じない。かつ、いままで手動におこなっていた EPS の調整作業が自動化される点に大きなメリットがある。

### 3.3 具体的成果

(1) Xabclib と OpenATLib

演算精度ポリシーが実装された OpenATLib β版を利用し、疎行列に対する対称標準固有値問題の解法であるリスタート付き Lanczos 法、および連立一次方程式の解法である GMRES(m)法を実装し、数値計算ライブラリ化した。本ライブラリを、Xabclib と呼ぶ。OpenATLib β版を用いた Xabclib は、Xabclib β版となる。

(2) α版の評価詳細

OpenATLib α版を利用して構築された Xabclib α版と、Xabclib β版の性能を比較する。Xabclib α版では、数値計算ポリシー機能と、MGS 以外の 3 種の Gram-Schmidt 直交化方式が実装されていない。そこで、α版の AT 機能を最大限に利用したうえで、演算精度ポリシーの AT 戦略を模倣した精度反復改良ループを構築し、最大実行時間制限を 1000 秒として実行した場合の実行時間と演算精度について評価する。なお、ユーザにより与えられた精度は  $1.0E-8$  とする。



(3) 性能評価環境

テスト行列として、フロリダ大学疎行列コレクションから、固有値ソルバに対して 21 種 (対称行列) を選んだ。

評価対象の計算機は東京大学情報基盤センターが所有する T2K オープンスパコン (東大版) の 1 ノード (16 コア) である。計算ノードは、4 ソケットの Quad-Core AMD Opteron Processor 8356 (2.3GHz) で構成され、ノードあたりのメモリ量は 32GB である。OpenATLib は OpenMP を利用して並列化されている。

コンパイラは Intel Fortran Compiler Professional Version 11.0 であり、コンパイラオプションは -O3 -m64 -openmp -mcmmodel=medium である。

(4) 演算精度ポリシー機能の効果

図 1 に、固有値ソルバ (OpenATLib の OpenATL\_EIGENSOLVE を利用した場合) における評価結果を載せる。なお、計算する固有値は、絶対値の大きいほうから 10 個であり、固有ベクトルはそれに付随するものを計算した。精度に関しては残差のほかに、固有ベクトルの直交性をフロベニウスノルムで評価した。

従来方式では、以下の行列に対し  $\alpha$  版はエンドユーザの要求精度を満たしていなかった。

- T3dl : 残差 : 2.24E-01、直交性 : 1.27E-02
- Ga41As41H72 : 残差 : 2.52E-08、直交性 : 4.81E-09

図 1 から、演算精度指定をしたのにもかかわらず、 $\alpha$  版よりも  $\beta$  版の実行速度が高速化された。この理由は、新規実装方式による疎行列-ベクトル積の実効効率の向上と Gram-Schmidt 再直交化方式の高精度化による反復回数の減少が理由と考えられる。なお、図 1 の  $\alpha$  版では、 $\beta$  版の精度反復改良ループを模倣しているため、エンドユーザの精度要求を満たしている。また、実行時間について、21 種の行列の平均速度向上は 2.24 倍であり、最大速度向上は 8.92 倍 (Ga41As41H72) であった。

以上より、本数値計算ポリシー機能の追加によりユーザからの要求精度をより満たすようになった。加えて、速度向上の効果も確認できた。したがって、評価したテスト行列の範囲内で本機能は有効である。

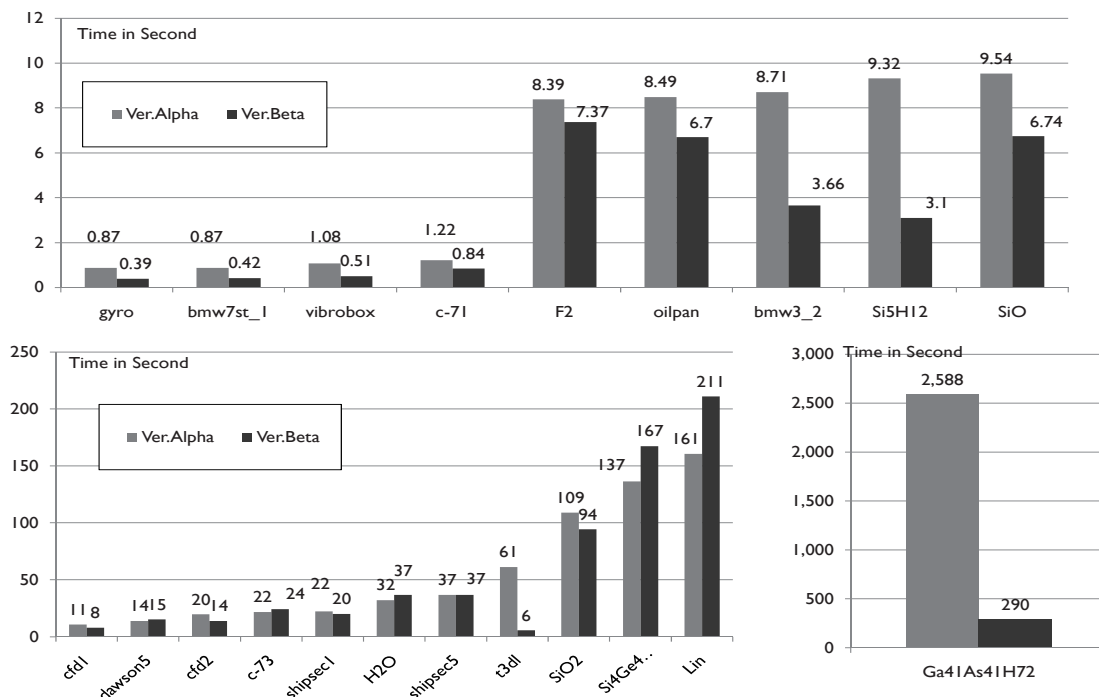


図 1 数値計算ポリシー (演算精度) の効果

## 4 HPC 教育関連業務

### 4.1 背景

東京大学情報基盤センター特任准教授として 2007 年 4 月 1 日に赴任後、教育業務として並列プログラミング基礎の講義と、本センターのスーパーコンピュータを用いた演習を、東京大学工学部・工学系研究科の共通科目「スパコンプログラミング(1)および(I)」として、通年(夏学期・冬学期)各 1 コマ行っている。また、駒場キャンパスの教養学部の学生に工学部と同様の授業を教えることで、天才プログラマを早期から育成することを目指す、全学ゼミ「スパコンプログラミング研究ゼミ」を開講している。本講義は、「学際計算科学・工学人材育成プログラム」の一環で開講されている。

内容は、通信ライブラリ MPI を利用し情報基盤センターに設置されている T2K オープンスパコン(東大版)4 ノード(64 コア)を受講生に無料で利用させる。スーパーコンピュータの利用方法の講習を含む。情報基盤センターのスーパーコンピュータの将来的な新規ユーザ獲得も狙った講義である。

### 4.2 内容(HPC 教育業務)

- **スーパーコンピュータ若手利用者推薦(試行)**(2007 年 10 月から実施)
  - **概要:**概ね 35 歳以下の若手研究者(学生を含む)を対象とした利用者向けのスーパーコンピュータ若手利用者推薦(試行)(以降、若手推薦(試行))による課題公募である。スーパーコンピューティング研究部門の教員により審査の上、採択された課題の計算機利用負担金(半年分)をセンターが負担する。
  - **詳細:**年 2 回公募し、年間で 4 件程度の優れた研究提案を採択する。継続申請と再審査の上で、最大で 1 年間の無料利用ができる。採択者には、報告書の提出、研究成果の発表の際に若手推薦(試行)を利用したことの明記、およびセンターが発行する「スーパーコンピューティングニュース」誌の原稿執筆を採択の条件とする。
- **お試しアカウント付き並列プログラミング講習会(試行)**(2008 年 3 月から実施)
  - **趣旨:**センターに設置されたスーパーコンピュータの利用者における利用促進、および利用を検討している新規ユーザを対象とした、並列プログラミング講習会(MPI および OpenMP を用いるもの。)を実施した。
  - **対象者:**センターのスーパーコンピュータのアカウントを所有しているユーザのみならず、センターのスーパーコンピュータの利用資格を有し、今後の利用を検討している方(大学教員、および学生など)、および企業所属の方も対象にする。
- **先進スーパーコンピューティング環境(ASE)研究会**(2008 年 3 月から実施)
  - **概要:**センターに設置されたスーパーコンピュータに関する最新の研究成果をユーザが知ることは、スーパーコンピュータの利用成果を創出するために必要である。また、センターのユーザが個別に所有する情報に関して、ユーザ間での情報交流の活性化も必要である。この趣旨から、スーパーコンピューティング部門が主催する、スーパーコンピュータ環境に関する研究会を行う。
- **共同研究プロジェクト(2009 年 1 月に実施開始)**
  - **概要:**「T2K オープンスパコン(東大版)」の利用環境を向上することを目的とし、様々なシミュレーションのアルゴリズムの開発、プログラムの高速化に関する研究を本センターのスタッフと共同で実施する。本共同研究プロジェクトの対象となる研究は「64 ノード(1024 コア)」程度を使用する大規模計算を大量に行う研究とする。利用者は、本センターとプログラムの並列化、高速化に関する研究を共同で実施する。

### 4.3 具体的成果

- 東京大学における HPC 教育(2009 年度)
  - 工学部共通科目(夏学期・冬学期)において、登録人数合計 40 名、単位取得者数 9 名を達成した。
  - 教養学部の全学ゼミにおいては、登録 9 名、合格 4 名であった。
- スーパーコンピュータ若手利用者推薦(試行)(2009 年度、前期・後期)[特記 1]
  - 本学、電気通信大学、埼玉大学、静岡大学、九州大学から合計 8 件の課題を採択した。
- お試しアカウント付き並列プログラミング講習会(試行)(2009 年度は 3 回実施[特記 3])
  - 他大学、研究機関、企業を含む、合計 55 名の参加登録があった。
- 先進スーパーコンピューティング環境研究会(2009 年度は 2 回実施[特記 2])
  - 招待講演として、ドイツの Hamburg-Harburg 工科大学から Jens-Peter Zemke 博士をお呼びし、線形反復解法において近年注目されている IDR(s)法に関する講演を行った。
  - 参加者合計は 38 名に達し、高性能計算分野において最新情報の交換、および研究者間の交流に貢献した。
- 共同研究プロジェクト
  - 主に 2 件の共同研究プロジェクトを担当した。
  - そのうち 1 件について、3 次元津波伝搬シミュレーションのコードで高速計算できる手法の開発に成功した。本手法は、情報処理学会の研究会で発表を行った[発表 4]。

## 5 スーパーコンピュータによる社会貢献業務

### 5.1 背景

本業務は、文部科学省 研究開発施設共用等促進費補助金(先端研究施設共用促進事業)「先端的大規模計算利用サービス」(以降、大規模利用サービス)の支援に関する業務、および東京大学情報基盤センターが民間企業にスーパーコンピュータを有償利用に供するための業務である。

### 5.2 内容

大規模利用サービスは、各大学が開発した応用ソフトウェア(シミュレーションソフトウェア等)と各大学が持つ国内有数のスーパーコンピュータを、それらの利用支援サービスと一体化して民間企業に提供することを目的にしている。

東京大学情報基盤センター業務は、スーパーコンピュータを民間企業に供する場合、企業ユーザの利用資格を審査したうえで、限定資源に対し有償利用を行う。

### 5.3 具体的成果

- (大規模利用サービス)2009 年度、課題申込書・募集・成果報告に関する業務補助
- (大規模利用サービス)2009 年度、第 1 期、第 2 期の課題選定業務
- 東京大学情報基盤センターにおける企業利用者に対する支援業務

## 6 成果要覧

### 招待講演／招待論文

[招待 1] Takahiro Katagiri: Auto-tuned Sparse Iterative Solver Toward Petascale Era, Third French-Japanese Workshop -- Petascale Applications, Algorithms, and Programming (PAAP) --, Shiran-Kaikan Hall Annex, Kyoto University, Apr. 2009.

- [招待 2] 片桐孝洋:ソフトウェア自動チューニング入門～ペタフロップス, 10 万並列を達成するための実装方式から数値アルゴリズムの自動最適化技術～, 第 38 回数値解析シンポジウム-NAS2009-, チュートリアル, 熱川ハイツ, 2009 年 6 月.
- [招待 3] 片桐孝洋:ソフトウェア自動チューニング技術の最新動向～マルチコア, ヘテロジニアス, 10 万並列な環境に向けた新しい最適化技術～, 第 9 回 ANS 研究会, 京都大学学術情報メディアセンター北館 3 階講習室, 2009 年 6 月.
- [招待 4] Takahiro Katagiri: Xabclib: A Sparse Iterative Solver with a Generalized Auto-tuning Interface, and Overview of Auto-tuning Studies in Japan, 米国エネルギー省(DOE)主催, Center for Scalable Application Development Software (CScADS) Summer Workshops, Workshop on Libraries and Auto-tuning for Petascale Applications, Granlibakken Resort and Conference Center, Tahoe City, CA, USA (Organizers: Jack Dongarra, Keith Cooper, Rich Vuduc, Kathy Yelick), Aug. 2009.
- [招待 5] 片桐孝洋:ソフトウェア自動チューニングの最新動向 ～数値計算ライブラリへの適用を例にして～, 京都大学グローバル COE, 知識循環社会のための情報学教育研究拠点, 京都大学吉田キャンパス工学部総合校舎 406 号室, 2009 年 10 月.
- [招待 6] 片桐孝洋:マルチコア・超並列時代の並列固有値ソルバ実装法 ～ブロック化, マルチキャストを中心に, 日本応用数学会 3 部会連携「応用数理セミナー」, 国立情報学研究所, 資料集, pp. 16-35, 2009 年 12 月.
- [招待 7] 片桐孝洋:ソフトウェア自動チューニング:パソコンからスパコンまでの先進最適化技術 ～数値計算ライブラリを中心に～, 情報処理学会東北支部, 第 350 回研究講演会, 秋田県立大学本荘キャンパス 講義棟 K321 教室, 2009 年 12 月.
- [招待 8] Takahiro Katagiri, Takao Sakurai, Hisayasu Kuroda, Ken Naono, and Kengo Nakajima: Development of numerical computation policy facility and its effect on Xabclib, 2nd International Workshops on Advances in Computational Mechanics (IWACOM-II), Yokohama, JAPAN), OW-1 Innovative Algorithms for Multicore and Hierarchical Computational Environments (Organizers: H. Okuda, K. Nakajima and T. Aoki), Mar. 2010.

## 査読付論文リスト

- [査読付 1] 片桐孝洋: ペタフロップス環境における小規模行列用対称密行列固有値ソルバに向けて –逆変換の改良–, 2010 年ハイパフォーマンスコンピューティングと計算科学シンポジウム(HPCS2010), HPCS2010 論文集, pp. 27-34, 2010 年 1 月.
- [査読付 2] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰:OpenATLib: 数値計算ライブラリ向け自動チューニングインターフェース, 2010 年ハイパフォーマンスコンピューティングと計算科学シンポジウム(HPCS2010), HPCS2010 論文集, pp. 99-106, 2010 年 1 月.
- [査読付 3] Takahiro Katagiri and Shoji Itoh: A Massively Parallel Dense Symmetric Eigensolver with Communication Splitting Multicasting Algorithm, 9th International Meeting on High Performance Computing for Computational Science (VECPAR'2010), Proceedings of VECPAR'2010, Berkeley, California, USA, June 2010. (採録決定)

## その他の発表論文リスト

- [発表 1] 片桐孝洋, 黒田久泰:マルチコア環境における自動チューニング機能付き疎行列反復解法ライブラリ, 第 14 回日本計算工学会講演会, オーガナイズドセッション, 「科学技術計算におけ



- る自動チューニングの挑戦:マルチコア環境への適用」, 東京大学生産研究所計算工学会講演論文集, pp. 167-170, 2009年5月.
- [発表 2] 片桐孝洋:ペタスケール計算を目指した MRRR 法を用いた固有値ソルバの開発, 第 14 回日本計算工学会講演会, オーガナイズドセッション, 「解けたらうれしい固有値問題」, 東京大学生産研究所, 計算工学会講演論文集, pp. 185-188, 2009年5月.
- [発表 3] 片桐孝洋, 櫻井隆雄, 黒田久泰, 直野健, 中島研吾:汎用的自動チューニングインターフェースをもつ疎行列反復解法ライブラリ, 2009 年先進的計算基盤システムシンポジウム (Symposium on Advanced Computing Systems and Infrastructures (SAC SIS)), 広島国際会議場, SAC SIS2009 論文集, pp. 116-117, 2009年5月.
- [発表 4] 片桐孝洋, 齊藤竜彦, 古村孝志, 中島研吾:3 次元津波伝搬シミュレーションにおけるコード最適化一手法, 情報処理学会研究報告 2009-HPC-120, 2009年6月.
- [発表 5] 中島研吾, 片桐孝洋:マルチコアプロセッサにおけるリオーダーリング付き非構造格子向け前処理付反復法の性能, 情報処理学会研究報告 2009-HPC-120, 2009年6月.
- [発表 6] 片桐孝洋, 櫻井隆雄, 黒田久泰, 直野健, 中島研吾:OpenATLib:汎用的な自動チューニングインターフェースの設計と実装, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 情報処理学会研究報告 2009-HPC-121, 2009年8月.
- [発表 7] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰:OpenATLib を利用した疎行列ライブラリの開発と評価, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 情報処理学会研究報告 2009-HPC-121, 2009年8月.
- [発表 8] 片桐孝洋, 黒田久泰:マルチコア環境における密および疎行列ソルバの自動チューニング機構の評価, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 日本応用数学会「行列・固有値問題の解法とその応用」研究部会, 2009年8月.
- [発表 9] 片桐孝洋, 黒田久泰, 中島研吾:Xabclib : 汎用な自動チューニングインターフェース OpenATLib を利用した疎行列反復解法ライブラリ, 第 5 回 ASE 研究会, 2009年8月.
- [発表 10] 黒田久泰, 片桐孝洋, 須田礼仁:省電力のための基本演算ライブラリの実装と評価, 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP2009), フォレスト仙台, 日本応用数学会「行列・固有値問題の解法とその応用」研究部会, 2009年8月.
- [発表 11] 片桐孝洋, 黒田久泰:マルチコア・超並列計算機時代の自動チューニング機能付き疎行列反復解法ソルバ, 日本応用数学会 2009 年度年会, 大阪大学豊中キャンパス, 年会予稿集, オーガナイズドセッション:ペタスケール環境を目指す数値計算ライブラリと自動チューニング技術, A2-1, 2009年9月.
- [発表 12] Takahiro Katagiri: Model for Software Automatic Tuning, Poster Presentation, Forth International Workshop on Automatic Performance Tuning (iWAPT2009), Proceedings of iWAPT2009, pp. 74, Oct. 2009.
- [発表 13] 片桐孝洋:ソフトウェア工学としての展開, 「自動チューニング技術の現状と応用に関するシンポジウム」, 東京大学弥生講堂アネックスセイホクギャラリー, 2009年10月.
- [発表 14] 櫻井隆雄, 直野健, 片桐孝洋, 中島研吾, 黒田久泰:クリロフ部分空間法に対する自動チューニングについて, 加速法フォーラム, 2009年12月.
- [発表 15] Takahiro Katagiri, Hisayasu Kuroda and Kengo Nakajima: Xabclib: An Iterative Solver with a General Auto-tuning Interface "OpenATLib" --- Development of Numerical Computation

Policy Function, 6th ASE Seminar, Information Technology Center, The University of Tokyo, Feb. 2010.

[発表 16] Takahiro Katagiri, Takao Sakurai, Hisayasu Kuroda, Ken Naono and Kengo Nakajima: Xabclib: An Iterative Solver with a General Auto-tuning Interface "OpenATLib", International Workshop on Peta-Scale Computing Programming Environment, Languages and Tools (WPSE 2010), Kyoto, Feb. 2010.

[発表 17] Takao Sakurai, Ken Naono, Hisayasu Kuroda, Takahiro Katagiri, and Kengo Nakajima: OpenATLib: A General Auto-tuning Interface for Numerical Solvers, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: The State-of-the-art of Auto-tuning Technologies: Adaptation to Advanced Computer Environment and Numerical Libraries - Part I of II, MS6, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.

[発表 18] Takahiro Katagiri, Hisayasu Kuroda and Kengo Nakajima: A Branchless Segmented Scan Method for Sparse Matrix-vector Multiplications, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: Parallel Programming Models and Algorithms for Multicore Clusters and GPGPUs - Part III of III, MS62, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.

## 特記事項

[特記 1] スーパーコンピュータ若手利用者推薦(試行)制度 2009 年度(前期)4 件(新規 3 件, 継続 1 件), 2009 年度(後期)4 件(新規 1 件, 継続 3 件).

[特記 2] 東京大学情報基盤センタースーパーコンピューティング部門主催, Advanced Supercomputing Environment (ASE)研究会: 2009 年 8 月 21 日第 5 回研究会, 2010 年 2 月 12 日第 6 回研究会の 2 回実施.

[特記 3] お試しアカウント付き並列プログラミング講習会(試行):2009 年 7 月 1 日~2 日, 2009 年 9 月 7 日~8 日, 2010 年 3 月 17 日の 3 回実施.

[特記 4] ミニシンポジウム・オーガナイザー, 14th SIAM Conference on Parallel Processing for Scientific Computing (PP2010), Joint JSIAM -- SIAM Minisymposium: The State-of-the-art of Auto-tuning Technologies: Adaptation to Advanced Computer Environment and Numerical Libraries, MS6 and MS42, Grand Hyatt Seattle, Seattle, Washington, USA, Feb. 2010.

[特記 5] 情報処理学会誌「情報処理」, 大特集: 科学技術計算におけるソフトウェア自動チューニング, 片桐孝洋: 編集にあたって, pp. 476, 弓場敏嗣, 片桐孝洋: ソフトウェア自動チューニングの枠組み, pp. 478-482, 片桐孝洋: ソフトウェア自動チューニング記述のための計算機言語, pp. 494-498, Vol. 50, No 6, 2009 年 6 月.

## 報道関連

[報道 1] 東京大学新聞, 2009 年 9 月 22 日(火)発刊, 第 2480 号, 「10 分でゲノム解読する時代に—ゲノム解読をスパコンで, 人材育成が急務, 情報悪用を防げ」, 2009 年 9 月.

[報道 2] 科学技術政策研究所発行, 「科学技術動向」11 月号特集, 「数値シミュレーションにおけるソフトウェア研究開発の動向—並列分散型のハードウェアとソフトウェア自動チューニング—」, 2009 年 11 月.



# 先端的大規模計算利用サービス

吉廣 保

## 1 概要

先端的大規模計算利用サービスの施設共用技術指導研究員として、大学で開発された応用ソフトウェアを、スーパーコンピュータへ移植する作業をなど、企業がスーパーコンピュータを利用するために必要な支援を行った。また、研究対象が「施設利用の高度化に関する研究」に限定されているため、本年度もこれまで開発してきたグラフィカル・ユーザ・インタフェースで、情報基盤センターのスーパーコンピュータを利用できるような整備を行った。

## 2 文部科学省 研究開発施設共用等促進費補助金(先端研究施設共用促進事業) 先端的大規模計算利用サービス

### 2.1 背景

独立行政法人・大学等の研究機関等の保有する先端的な研究開発に係る施設及び設備(以下「先端研究施設」)の中には、広範な分野や多様な研究等に活用されることにより、様々な研究等に大きな成果をもたらす可能性のあるものが多数あるが、利用者支援等の共用のための体制が整備されていないことから十分に活用されていない。「研究開発施設共用等促進費補助金(先端研究施設共用促進事業)」は、それらの先端研究施設の共用を促進し、基礎研究からイノベーション創出に至るまでの科学技術活動全般の高度化を図るとともに国の研究開発投資の効率化を図るために2009年度から文部科学省が開始した事業である。

本事業を通じて、産学官の研究者による戦略的かつ効率的な研究開発や、研究機関・研究分野を越えた横断的な研究開発活動を推進することにより、継続的に産学官の知の融合によるイノベーションを加速していくことを目指している。

全国共同利用施設であるスーパーコンピュータを有する7大学(北海道大学、東北大学、東京大学、名古屋大学、京都大学、大阪大学、九州大学)は、2007年7月より2009年3月まで文部科学省先端研究施設共用イノベーション創出事業の支援を受け「先端的大規模計算シミュレーションプログラム利用サービス」を通じ、社会貢献の一環として、大学で開発された応用ソフトウェアとスーパーコンピュータの民間企業への提供を実施してきた。

2009年度からは、「先端研究施設共用促進事業」の支援を受け「先端的大規模計算利用サービス」として、より幅広い研究課題に対してスーパーコンピュータの提供を実施している。

本事業の特徴は、各大学が開発した応用ソフトウェア(シミュレーションソフトウェア等)と各大学が持つ国内有数のスーパーコンピュータを、それらの利用支援サービスと一体化して民間企業に提供することにある。このようにサービス、ソフトウェア、ハードウェアの3つを同時に支援する試みは、国内外を通じて初めてのものである。この事業の目的は、大学の持つ知と施設によって我が国の経済発展に貢献することである。

## 2.2 内容

施設共用技術指導研究員として、「革新的シミュレーションソフトウェアの研究開発 (RSS21)」で開発されたソフトウェアを 7 大学の計算機へインストール・整備・サポートを行う。また、RSS21 ソフトウェアのチューニングを行うための非常勤研究員と連携してチューニング作業を行う。ASP (Application Service Provider) 事業を考えている企業に関しての全般的なサポートを行う。

「施設利用の高度化に関する研究」に関してのみ研究活動が行えるため、グラフィカル・ユーザ・インタフェース (GUI) によるスーパーコンピュータの操作性向上に関する研究等に取り組む。

## 2.3 具体的成果

昨年度から引き続き、RSS21 ソフトウェアを HA8000 クラスタシステム上で利用する企業に対しサポートを実施した。平成 21 年度は、特に PHASE と ABINIT-MP のサポートを行った。

「施設利用の高度化に関する研究」に関しては、これまで作成してきた GUI を基に、汎用的に使用できるインタフェースの実装についてと、データの暗号化について設計、試作を行った。本事業や、センターで実施している HA8000 クラスタシステムの企業利用支援では、企業側からの要望として、データの安全性が求められることが多い。その解決策の 1 つとして、ユーザが設定する暗号化の方式を使用してデータエンコードを行い、センターのディスク上には常に暗号化されたデータのみが存在し、実際にデータを使う際に復号化を行いつつ計算を行う仕組みについて、いくつかの手法を模索している。

先端研究施設共用促進事業としては、2009 年 2 月 2 日から 3 月 6 日までで 2009 年度第 1 期公募を行い、6 件 (継続課題を含む、うち東京大学は 1 件) の採択課題が 4 月 1 日より計算機利用を開始した。また、7 月 9 日より 9 月 4 日まで第 2 期公募を行い、12 件 (継続課題を含む、うち東京大学は 1 件) の採択課題が 10 月 1 日より計算機利用を開始した。2009 年度に予定していた採択件数を下回った大学が多かったため、10 月 1 日より 12 月 4 日までで第 3 期公募を行った。第 3 期では、1 件 (東京大学は 0 件) の採択課題が平成 21 年 1 月 4 日より計算機利用を開始した。

2008 年度に採択された課題で、新規利用拡大利用が終了した企業が、HA8000 クラスタシステムの企業利用支援を使用して ASP 事業を開始したいと希望したため、ASP 事業で利用するソフトウェアのライセンスサーバ問題などについてサポートを行い、良好な評価を得た。また、契約に関する打ち合わせなどのサポートも行った。

# マルチスケールな現象の直接計算 — 大規模分子動力学法による気泡生成現象の研究 —

渡辺 宙志

## 1 概要

本研究の目的は、相転移や流れを伴うマルチスケールな現象を粒子を用いた大規模計算により再現し、その非平衡特性を明らかにすることである。具体的には、大規模粒子系を計算するための並列化技法の開発、およびその計算コードを用いた気泡生成過程の研究を行っている。今年度は、以下に挙げる二つのトピックについて取り組んだ。

- 大規模分子動力学法コードの開発

沸騰現象は発電所などの熱機関において中心的な役割を果たすが、強い非平衡現象であること、相転移を含む移動境界問題であることから扱いがきわめて難しい。特に軽水炉などで問題となる気液二相流は、熱伝導や流動などの輸送現象に加え沸騰という相転移も複雑にからみあっており、スケールの異なる現象の同居が研究の妨げとなってきた。このようなマルチスケールな問題については、現象をそのスケールにより分類し、計算するスケールよりも小さいスケールについては粗視化された構成式を用いることで研究が行われてきた。しかし、流体を構成する全ての粒子の運動を分子動力学法により計算すれば、粒子法においては界面は自発的に形成され、かつ気液界面における摩擦や熱輸送といった実験では調べづらいパラメータを入力として与える必要も無い。そこで、気液相転移を持つ標準的な粒子間ポテンシャルである Lennard-Jones 粒子について、MPI を用いた並列コードを開発、その並列化効率を測定した。

- 準安定液体における気泡核生成現象の研究

沸騰とは、温度や圧力などが変化し、単相として不安定となった準安定な液体から気泡が核生成する過程である。こういった準安定相からの核生成過程は一般的に古典核生成論 (Classical Nucleation Theory, CNT) という理論の枠組みで記述されるが、気泡核生成は圧力を介した気泡間相互作用や気泡成長時の強い非平衡性などにより理論的な解析が困難である。特に古典核生成論が予言する核生成率と実測値の大きなずれが問題となってきた。そこで、分子動力学法による全粒子計算により準安定状態の液相を作り、そこからの気泡核生成過程を観察することで、理論の検証を行い、これまで観測が困難であった非平衡性や気泡間相互作用を明らかにする。

## 2 大規模分子動力学法コードの開発

### 2.1 背景

気液混相流は流れと相転移がカップルした典型的なマルチスケール、マルチフィジックス問題であり、工学応用上重要な現象でありながら、移動境界問題であること、相転移により密度が大きく変化すること、さらに相転移と流れがカップルしていることが数値計算を困難なものとしている。これらの系の計算は主にナビエ・ストークス方程式などの構成方程式を連立させて行われてきたが、相境界の扱い、特に相境界

の厚さをゼロとしていることと、相境界での輸送係数が実験で決めづらいことなどが問題視されてきた。それに対し、系を構成する粒子の挙動をすべて追う全粒子計算が注目されている。全粒子計算では、粒子間ポテンシャルのみを与えることで、相転移や熱輸送、相境界の移動などがすべて自発的に行われる。近年の計算機の能力の向上および数値計算アルゴリズムの改良により、沸騰流を再現するのに必要と見込まれている  $10^{12}$  粒子に到達、さらに増大を続けている。これだけの計算を行うためには並列化が必須となるが、どれだけの大きさの計算機において、どれだけの粒子数がどの程度の並列化効率で扱えるかが問題となる。そこで、大規模並列化分子動力学法コードを開発し、その並列化効率、および計算コストの詳細について調べる。

## 2.2 内容

二体相互作用をする粒子系において、単純領域分割による MPI 並列化を行い、その並列化効率を調べた。粒子系のポテンシャルには、カットオフのある Lennard-Jones ポテンシャルを用いた。その関数形は以下の通りである。

$$V(r) = 4\epsilon \left( \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 + c_2 \left( \frac{r}{r_c} \right)^2 + c_0 \right) \quad (1)$$

ただし、 $r$  が粒子間距離、 $r_c$  はカットオフ長さである。係数

$c_2, c_0$  は、力とポテンシャルがカットオフのところでゼロになる条件  $V(r_c) = V'(r_c) = 0$  から決まる。

カットオフを入れた場合、力の計算をする前に、相互作用している粒子のペアのリストを構築する必要がある粒子のペアを直接探すと計算量が  $O(N^2)$  となるため、メッシュ探索を用いることで  $O(N)$  に落とす。メッシュ探索とは、全系をある長さの格子に区切り、その格子に粒子を登録することで近所の粒子を探す方法であり、大きく分けて排他的格子 (Exclusive Mesh) と、非排他的格子 (Non-exclusive Mesh) の二種類がある。排他的格子は、一つの格子にはたかだか一つの粒子しか入らないようにする方法であり、純粋に格子の数だけ配列を用意して、もし粒子がその格子に存在していたら粒子番号を登録する。非排他的格子は、複数の粒子が入る程度の大きさに格子を作る方法である。一般に剛体粒子のような相互作用距離が短いケースでは排他的格子が、Lennard-Jones ポテンシャルのように相互作用距離が比較的長いケースでは非排他的格子の方が効率が良いことが多い。ここでは非排他的格子を採用した。

粒子数が多くなると、相互作用ペアリストの構築にかかる時間も長くなる。分子動力学法においてもっとも計算コストが高いのは力の計算だが、ペアリストの構築にはメモリを広範囲にわたってアクセスする必要があるため、近年の計算速度に対するメモリ転送性能 (いわゆる Byte per FLOP) の低下により、相対的な計算コストが高くなりつつある。そこで、相互作用をしている粒子を探す範囲を長くして、ペアリストが有効な時間内はリストを再構築しないことで高速化を図る。ペアリストの有効期限のチェックには、系に存在する粒子の最高速度により判定する方法のほか、リストを作ったときの粒子の位置を覚えておいて、粒子の変位が有効範囲を超えないという条件で有効期限を決める方法もある。密度がある程度高い場合、粒子は頻繁に衝突し、速度に比して変位はさほど増えないため、位置によって有効期限を決めたほうが効率的である。ここでは、両方の判定法を組み合わせることでペアリストの寿命を伸ばしつつ、ペアリストの有効範囲判定の計算コストを下げている。

並列化は MPI を用いた単純領域分割によって行った。各プロセスは自分に割り当てられた領域を担当し、境界にいる粒子の情報を交換しながら計算を行う。通信は、1) 境界からはみ出した粒子の交換、2) 境界領域の粒子の情報、3) ペアリストの有効期間確認、の三種類存在する。相互作用が短距離で

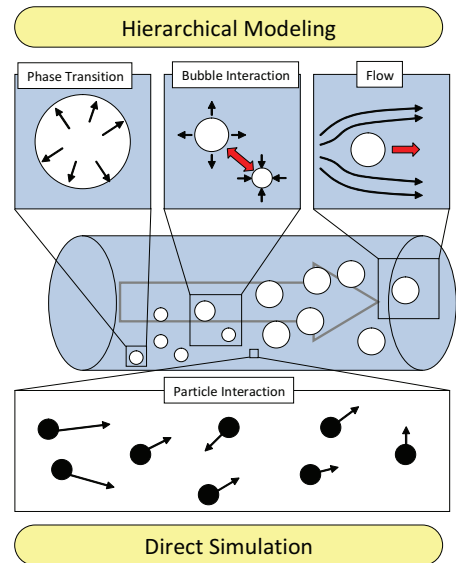


図 1: 分子動力学法によるスケール層のある物理現象の全粒子直接計算。



あるため、基本的には通信は隣接する領域を担当するスレッド間でのみ行われるが、ペアリストの有効期限確認には大域通信 (具体的には MPI\_Allreduce) が必要となる。

以上のアルゴリズムを実装したコードを開発し、実際に並列化効率の測定を行った。カットオフ距離を  $r_c = 2.5$  とし、プロセスあたりの粒子数を約 50 万個として固定してプロセス数を増やす weak scaling を行った。密度は 0.5 に固定し、最初に FCC 格子に配置しておき、初期速度 1 を与えて時間発展させる。そして 150 ステップだけ緩和させた後の 1000 ステップを計算するのに必要な時間から計算速度を決める。計算速度は、MUPS という単位を用いる。MUPS は百万粒子更新毎秒 (Millions Update per Second) の略で、1 秒間に百万粒子を一度数値積分できると 1 となる計算速度の単位である。以上の内容を東京大学情報基盤センターの HA8000 において測定した。1 ノードあたり 16 プロセスとして、1 ノードから 64 ノードまで測定した結果を図 2 に示す。最大 1024 プロセス、5 億粒子を超える計算を行い、1 ノード占有に比較した並列化効率は 64 ノードで 86.1% であった。通信による並列化効率の低下は、3次元の全方向にノード間通信をとまなう 8 ノードから頭打ちとなり、以後ほとんど低下していないことがわかる。これは、通信がレイテンシではなく、ほぼバンド幅にボトルネックがあることを示唆する。また、ここまでの規模の計算では大域通信の影響はほとんど見られないが、今後スレッド数が増えてくると、全体の同期の時間の影響は無視できなくなると考えられる。この同期コストをどのように減らすかも今後の課題である。

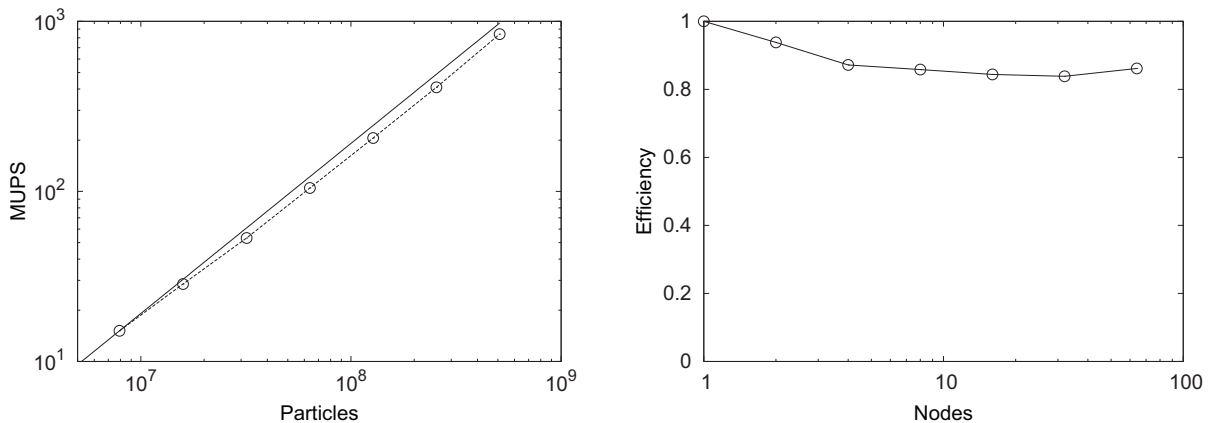


図 2: HA8000 において、1 ノードあたり 16 プロセス、プロセスあたり粒子数を約 50 万個に固定してノード数を増やす weak scaling を行った。(左) 粒子数と速度。64 ノードの計算で、5 億 1008 万 2164 粒子、841MUPS の速度を達成した。(右) 並列化効率。1 ノード占有計算を基準とし、1,2,4,8,16,32,64 ノードにおける計算時間の比を表示している。64 ノードで 86.1% の並列化効率であった。

### 2.3 具体的成果

本テーマについて二件の口頭発表を行った [発表 5, 発表 11]。また、九州大学情報基盤研究開発センター先端的計算科学研究ベンチマーク課題として採択され、同センター SR16000 において、42 ノード、2688 プロセス、最大 13 億粒子の計算も行い、並列化効率を測定した。

## 3 準安定液体における気泡核生成現象の研究

### 3.1 背景

沸騰現象は発電所などの熱機関において中心的な役割を果たすが、強い非平衡現象であること、相転移を含む移動境界問題であることから扱いがきわめて難しい。特に軽水炉などで問題となる気液二相流は、熱伝導や流動などの輸送現象に加え沸騰という相転移も複雑にからみあっており、スケールの

異なる現象の同居が研究の妨げとなってきた。沸騰流はマイクロな相転移とマクロな流動がカップルしたマルチスケールな現象であるが、そのマイクロな素過程である気泡発生過程についても分かっていないことが多い。過熱状態にされた、もしくは飽和蒸気圧以下に減圧された液体は、液相単相では不安定となり、準安定状態となる。このような準安定液体は、やがて気泡が発生することで相共存状態へと遷移していくが、このような準安定液体における気泡核生成率を予言するのが古典核生成論である。しかし、実際に実験や数値計算によって観測された気泡生成率は古典核生成論による予言と数桁以上異なることが知られており、その原因はまだよくわかっていない。そこで、分子動力学法による全粒子計算において気泡生成現象を再現し、そのずれの原因を探る。

### 3.2 内容

ある温度や圧力において、液相単相の平衡状態が実現されているとする。これを、定圧条件で温度を上げる、定積条件で密度を下げるなどの操作を行い、気液共存線を横切ると液相は準安定状態となり、気泡核生成が起きる。その後、定圧条件では最終的に気相単相が、定積条件では気液共存状態が実現される。一般に温度を上げることにより起きる発泡現象を沸騰 (boiling)、減圧により生じる発泡現象をキャビテーション (cavitation) と呼び、区別される。シミュレーションによる実現が簡単なのは、定積条件において急減圧することである。急減圧を行った直後、液体は準安定状態となり、化学ポテンシャルは同じ温度、同じ密度における気体よりも高くなる。しかし、液相から気相にいたるにはエネルギーバリアがあり、すぐに気泡は発生できない。このエネルギーバリアを見積もるには、気泡を生成するための仕事を求める必要がある。気泡生成の仕事  $W$  は次のように記述される。

$$W = \gamma A - \Delta PV + V\Delta\mu, \quad (2)$$

ただし  $\gamma$  は界面張力、 $A$  が気泡面積、 $V$  が気泡体積、 $\Delta P$  は気泡内外の圧力差、 $\Delta\mu$  は気相と液相の化学ポテンシャル差である。右辺第一項が界面が存在することによる自由エネルギーの損、第二項が気泡がまわりの液体にする仕事、第三項がその温度、密度において気泡の方が安定であることによる化学ポテンシャルの得を表す。一般にこの生成過程は非平衡過程であるため、生成仕事  $W$  は状態の関数として求めることはできないが、気泡形状は成長過程を通してほぼ球であり、全ての状態変化が準静的に行われると仮定することで圧力差がつねに Young-Laplace の関係式

$$\Delta P = \frac{2\gamma}{R} \quad (3)$$

を満たすと要請できる。ただし  $R$  は気泡半径である。これは圧力緩和が非常に早く、その時間スケールが気泡成長に比して無視できるほど短いという力学平衡の仮定である。しかし、実際の気泡成長でこの力学平衡が成り立っているという保証はない。そこで、分子動力学法により気泡生成過程を再現し、圧力緩和と気泡成長の時間スケールを直接測定する。

まず、系を液相が単相で安定である状態 ( $T = 0.9$ ,  $\rho = 0.7$ ) で緩和させ、そこから急減圧することで準安定状態とする。減圧はある瞬間に粒子半径を一様に小さくすることで行った。これは系を一様に急膨張させたことと等価である。減圧の度合いが強ければ、系はスピノーダル線を超えるために気相が絶対安定となり、減圧の瞬間から多数の気泡が発生する (図 3)。また、減圧強さが弱ければ、エネルギーバリアから決まる待ち時間の後に (多くの場合において単一の) 気泡が生成する (図 4)。以後、この領域を核生成領域と呼び、この領域において音速と気泡成長時間の観測を行った。

圧力緩和の時間スケールは、ほぼ系の音速と同程度と考えられることから、系の音速を測定する。系の音速  $c$  は、圧力を  $P$ 、密度を  $\rho$  として  $c = \sqrt{dP/d\rho}$  で与えられる。そこで、系の圧力を密度の関数として測定した結果を図 5 左に示す。ここから、準安定領域における音速は  $c \sim 3.0$  と求められた。

核生成領域において、成長する気泡の時間発展を観測した。時刻  $t = t_0$  に成長をはじめた気泡体積の時間発展を  $V(t)$  とすると、近似的に

$$V(t) = V(0) \exp((t - t_0)/\tau), \quad (4)$$



で表される。ただし $\tau$ は気泡成長の時間スケールである。この気泡体積の時間発展の様子を図5の右に示す。同じ条件で減圧を行っても、異なる気泡生成時刻 $t_0$ に気泡成長が始まるが、その時間スケール $\tau$ はほぼ一定であることがわかる。今回の観測の条件では、 $\tau = 15$ と見積もられた。音速が $c = 3$ 程度であったため、音速が系を横切る時間スケールはおよそ32程度となり、気泡成長の時間スケールと同等かそれより遅いことがわかる。したがって力学平衡の仮定には疑問が残る。このような状況を正しく取り扱うためには、古典核生成論を気泡内外の圧力が釣り合っていない状態に拡張する必要がある。

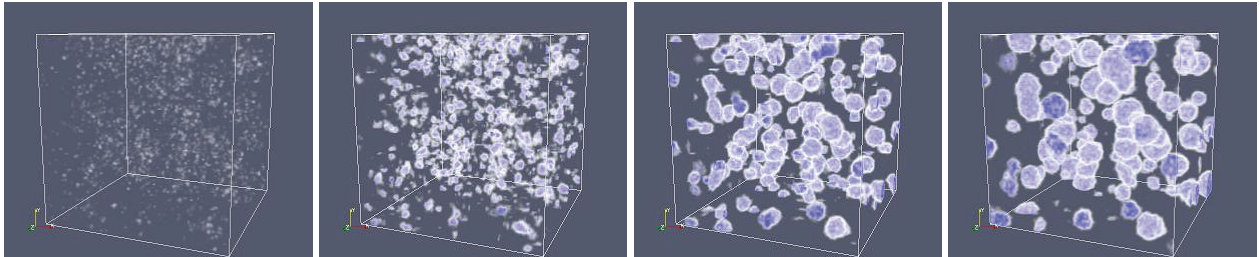


図3: 気泡生成過程の可視化例 (スピノーダル領域)。966万3597個のLennard-Jones粒子を安定な液相に緩和させた後に急減圧することで気泡が生成する。局所密度を反転させることで気泡の可視化を行った。この例ではスピノーダル点を越えたため、多重気泡生成が起きている。

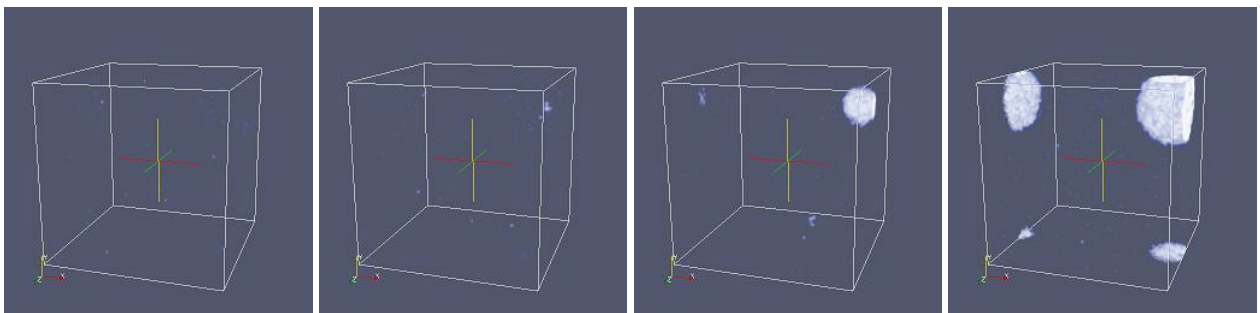


図4: 気泡生成過程の可視化例 (核生成領域)。減圧強さが弱い場合にはエネルギーバリアのため、気泡生成までに待ち時間があり、単一の気泡が生成される。

### 3.3 具体的成果

本テーマについて5件の口頭発表を行った [発表1, 発表2, 発表3, 発表4, 発表7]。

## 4 成果要覧

### その他の発表論文

- [発表1] 渡辺宙志, 鈴木将, 伊藤伸泰: 気泡生成過程の微視的工作と非平衡性, 日本物理学会第65回年次大会, 2010年3月.
- [発表2] 能川知昭, 伊藤伸泰, 渡辺宙志: 2次元弾性円盤系における並進、回転対称性の破れ, 日本物理学会第65回年次大会, 2010年3月.
- [発表3] 渡辺宙志, 伊藤伸泰: 気泡生成のシミュレーション解析, 地震研共同利用研究集会「火山現象の数値計算研究」, 2009年11月.

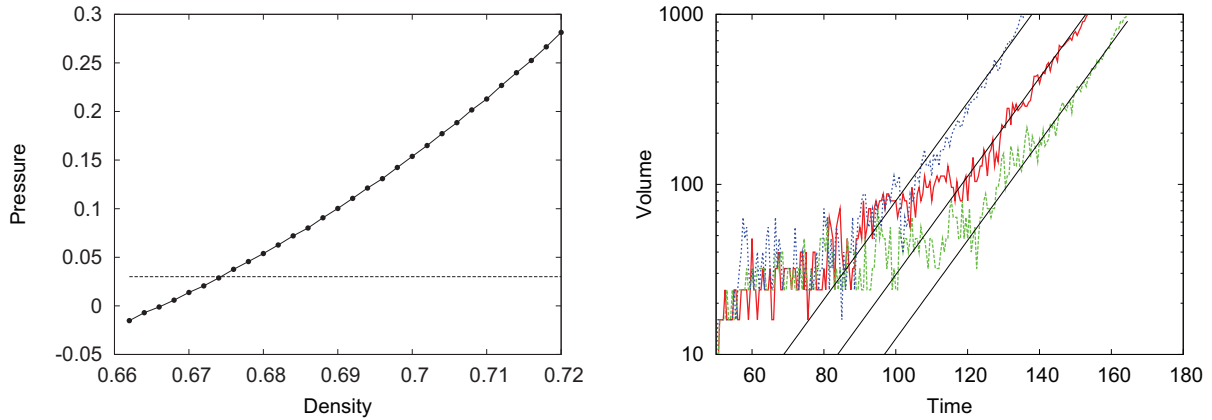


図 5: (左) 圧力の密度依存性。温度は  $T = 0.9$ 。その温度における飽和蒸気圧を破線で示してある。 $\rho = 0.675$  よりも低密度側では圧力は飽和蒸気圧よりも低く、系が準安定状態となかったことがわかる。ここから音速は  $c \sim 3$  と見積もられた。(右) 気泡成長過程。異なる初期条件の 3 サンプルを表示している。それぞれ異なる待ち時間により生成が始まるが、成長速度の時間スケールは共有していることがわかる。

[発表 4] 渡辺宙志: 気液混相流における熱伝導と核生成, 日本機械学会 第 22 回計算力学講演会, 2009 年 10 月.

[発表 5] Hiroshi Watanabe: Huge-scale molecular dynamics simulation on gas-liquid multi-phase flow, 8th Workshop on Stochastic Analysis on Large Scale Interacting Systems, 2009 年 10 月.

[発表 6] 渡辺宙志, 轟木義: 定温ダイナミクスとその幾何学的理解, 日本物理学会秋季大会, 2009 年 9 月.

[発表 7] 渡辺宙志, 鈴木将, 伊藤伸泰: 気液混相流における熱伝導と核生成, 日本物理学会秋季大会, 2009 年 9 月.

[発表 8] 能川知昭, 伊藤伸泰, 渡辺宙志: 相境界面運動の分子動力学シミュレーションによる解析, 日本物理学会秋季大会, 2009 年 9 月.

[発表 9] 山田賢太, 島田尚, 渡辺宙志, 伊藤伸泰: レナードジョーンズ液滴の衝突破碎シミュレーションによるクラスタ解析, 日本物理学会秋季大会, 2009 年 9 月.

[発表 10] 渡辺宙志: Pressure in liquid-gas interface, 第 11 回 ACP ワークショップ, 2009 年 9 月.

[発表 11] 渡辺宙志: 動的ロードバランスを備えた並列分子動力学法, 第 9 回 ACP ワークショップ, 2009 年 4 月.

### 査読付論文

[査読付 1] H. Watanabe: Non-equilibrium Relaxation Analysis on Two-dimensional Melting, Progress of Theoretical Physics, Supplement, No. 178, pp. 41-48, 2009.

# 大規模分散計算環境を有効活用するソフトウェアの研究

鴨志田 良和

## 1 概要

クラスターベースの大規模分散計算環境をより扱いやすく、より効率よく利用可能にするためのソフトウェアの研究を行った。具体的には、以下の項目について実施した。

- 大規模分散計算環境上のリアルタイムモニタリングシステム

モニタリングシステムが情報収集を行う負荷が、大規模な並列処理に与える影響について調査し [発表 1]、HA8000 クラスターシステム上で試験運用した [招待 1]。また、公開しているモニタリングソフトウェア VGXP に改良を施し、GXP への統合を行った [公開 1]。

- スパコン運用支援

本センターのスーパーコンピューティング部門でサービスを提供する SR11000、HA8000 クラスターシステムの各種運用支援を行った。さらに、スパコン利用促進のため、スパコン講習会 [特記 3] での講義の実施、プログラミングコンテスト [特記 1]、[特記 2] の開催に貢献した。また、全国共同利用施設であるスパコンセンターを設置する 7 大学と筑波大学、東京工業大学、国立情報学研究所によるグリッド配備・運用タスクフォース [特記 5] に参加し、HA8000 クラスターシステムに導入した NAREGI ミドルウェアの試験運用を行った。

## 2 大規模分散計算環境上のリアルタイムモニタリングシステム

### 2.1 背景

計算機上で実行されているアプリケーションの性能(実行時間)は、その計算機上の OS やデーモンプロセスなどの動作に影響されることがある。実際どのような影響があるのかを端的に示したのが、図 1 である。通信をしないローカルな計算と全プロセスの同期を含む通信を繰り返し行うようなアプリケーションでは、どこか一つの CPU に、計算に直接関係しないプロセス(Noise)がスケジュールされると、処理全体が遅延する。OS 等が原因で定期的に行われるべきアプリケーションの同期処理の時間がずれることから、このような現象は OS jitter とも呼ばれる。ノード数が増えるほど、OS jitter によって並列処理が遅れる可能性は高くなり、図 1 の灰色の領域で表わされる、無駄な待ち時間も増加することになる。

各計算機で適当な時間間隔でシステム情報をポーリングして取得するようなモニタリングシステムを実装する場合も、このことに十分注意する必要がある。モニタリングシステムが取得する情報は、システム全体の統計情報など、取得にほとんど負荷がかからないものが多いが、より詳細な情報を取得することができれば、アプリケーションの性能解析などに役立てることができる。しかし、モニタリングシステムがより多くの仕事をすることが必要になる場合には、モニタリング用のデーモンプロセスが並列アプリケーションに与える影響が、より甚大なものになると考えられる。より多くの情報

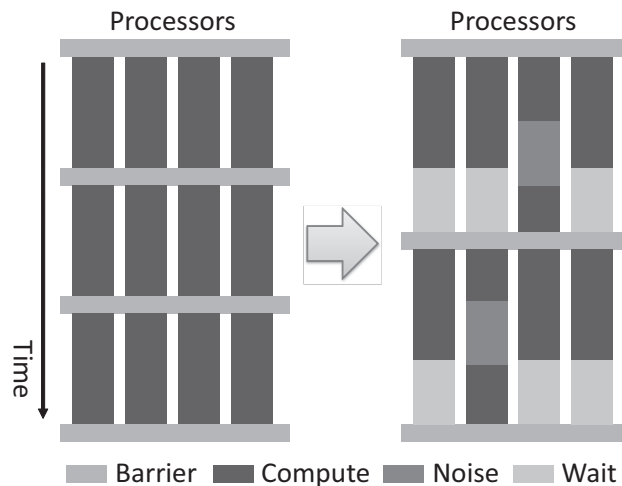


図 1: 同期が並列アプリに与える影響

を収集しつつ、システムへの負荷がより低いモニタリングを行うためには、モニタリングプロセスがシステムやアプリケーションにどのような影響を与えるかを把握することが重要である。以上の背景のもと、以降の章ではモニタリングが並列アプリケーションに与える影響を減らす可能性があるいくつかの方法について、実環境で評価を行い、その結果を報告する。

## 2.2 内容

### 2.2.1 予備的実験

本研究ではまず、モニタリングプロセスが並列アプリケーションに、実際にどの程度の影響を与えるのかを調べるための予備的実験を行った。モニタリング用のプロセス (pinfo) は、我々が開発・公開しているモニタリングソフトウェアである VGXP で用いているシステム情報取得用のプロセスと同様のものを使用した。pinfo は、定期的に実行され、ロードアベレージやメモリ使用量などのシステムの統計情報と、ノード上で動作している各プロセスの使用 CPU 時間などの統計情報を取得する。このプロセスをノードあたり 1 つ起動する。本稿で行われたすべての実験では、0.5 秒に 1 回の頻度でこの情報取得処理を実行した。

### 2.2.2 アプリケーションの概要

実験に使用したアプリケーションは、不均質な物性値分布を有する立方体形状における三次元静的弾性問題を、並列有限要素法 (Finite-Element Method、FEM) によって解くもので、連立一次方程式の解法には SGS (Symmetric Gauss-Seidel) を前処理手法とし、共役勾配法 (Conjugate Gradient、CG) 法を使用している。

アプリケーション内では以下に述べるような通信が行われる。階層型領域間境界分割 (HID) によって各自の担当領域を分担したプロセスは、繰り返しごとに、隣接領域を担当するプロセスと境界部分のデータを交換する。このために 1 対 1 通信による送受信が 8 回程度行われる。その後、全プロセスで MPI\_Allreduce を 3 回実行する。この処理を収束するまで (実験で与えた入力データでは 1,250 回程度) 繰り返している。全体の同期を伴う Allreduce 処理は、ベクトル同士の内積計算のために必要な処理であり、この頻繁な同期処理があるため、デーモンプロセスの動作が実行時間に大きな影響を与えることが予想される。

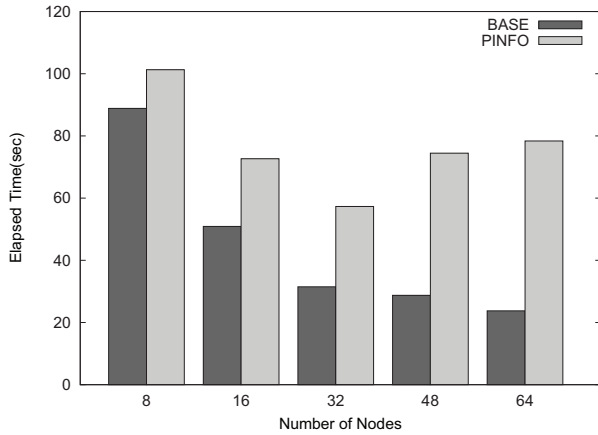


図 2: モニタリングプロセスの有無によるアプリケーション実行時間の違い  
BASE/PINFO はモニタリングプロセスがないある場合

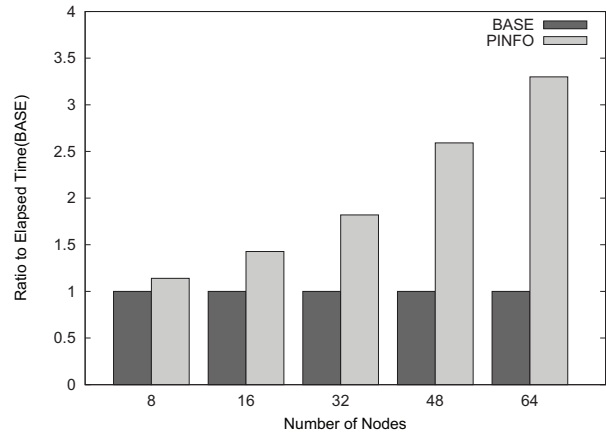


図 3: モニタリングプロセスの有無によるアプリケーション実行時間の違い  
(BASE に対する PINFO の割合)

### 2.2.3 実行環境の概要

実験は HA8000 クラスタシステム上で行った。各ノードは AMD quad-core Opteron(2.3GHz)4 ソケット、合計 16 コアから構成されており、ノードあたりの記憶容量は 32GB(一部 128GB)である。HA8000 クラスタシステムは内部でクラスタ群に分かれている。各クラスタ群内のノード間は Myrinet-10G(1 リンクあたり 1.25GB/sec × 双方向)で接続されている。ノード A 群は各ノード 4 本 (5.00GB/sec × 双方向)、ノード B 群は 2 本 (2.50GB/sec × 双方向)である。本章の実験では、ノード A 群のうちの 8 ノード (合計 128 コア) から 64 ノード (合計 1024 コア)を使用した。アプリケーションのコンパイルには日立製専用コンパイラ (FORTRAN90)を使用した。各ノード 16 コアを全てを使用した。

### 2.2.4 予備的実験の結果

図 2 の BASE が、ノード数を 8 ノード (128 プロセス) から 64 ノード (1024 プロセス) まで変化させて、アプリケーションを実行した時の実行時間である。一方、PINFO は、このアプリケーションと同時に、pinfo を 0.5 秒に 1 回の頻度で実行した結果である。図 3 には、PINFO の BASE に対する割合を示した。このように、システムのプロセスが定期的に動作しているだけで、アプリケーションの実行時間が大きく遅延することが分かる。しかも、8 ノードの場合は数%の遅延にとどまるのに対して、64 ノードの場合は元の 3 倍以上の時間がかかることから分かるように、モニタリングプロセスが与える影響は、ノード数が増加するに従って、より大きなものになる。

### 2.2.5 実行時間遅延の原因

実行時間の遅延を、計算を行うアプリケーションとモニタリングプロセスを次のようにモデリングすることで見積もることとする。まず、アプリケーションは、一定時間 ( $C$ ) の計算をした後に同期を行う処理を  $L$  回繰り返すものとする。同期にかかる時間も  $C$  に含まれるものとする。また、モニタリングプロセスは、各ノードでランダムなタイミングでスタートし、一定間隔 ( $F$ ) でスケジュールされ、一定時間 ( $M$ ) 実行されるものであるとする。あるループ中にモニタリングプロセスがスケジュールされる確率を  $p$  とし、モニタリングプロセスによってループが遅延する時間は  $M$  に等しいと仮定すると、ループ 1 回の実行時間 ( $e$ ) は、

$$e = C + pM$$



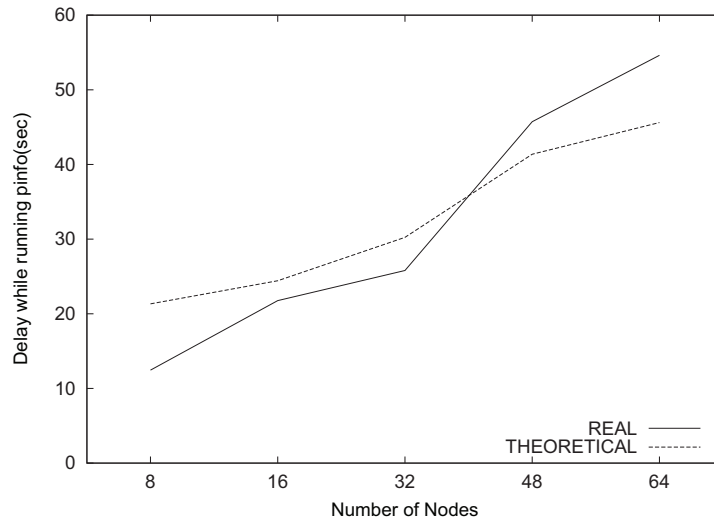


図 4: モニタリングプロセスによるアプリケーションの遅延時間

とあらわすことができる。(2つ以上のもの他リングプロセスが同じループにスケジュールされる場合、遅延時間は  $M$  より長くなる場合もあるが、単純にするため考慮しないこととする。) 一方、 $p$  は、

$$p = 1 - \left(1 - \frac{e}{F}\right)^N$$

と書けるので、この2式から  $p$  を求めることができる。

pinfo 単体での実測値に基づき、 $M = 15(\text{msec})$ 、 $F = 500(\text{msec})$  とし、残りを pinfo なしでアプリケーションを実行したとき (BASE) の結果から決定し、理論的な遅延時間 ( $LMp$ ) を計算したものと、実際に観測された遅延時間を図 4 に示す。

理論値と実測値がおおむね一致しているため、上に述べた現象が、遅延の主な原因であると考えられる。モニタリングプロセスがスケジュールされる頻度を減少させずに、アプリケーションの遅延を改善する方法は、2つ考えられる。まずはランダムなタイミングでスタートせず、各ノードのモニタリングプロセスが、同期的に同じ時間にスケジュールされるようにすれば、アプリケーションの1度あたりの繰り返しでプロセスが遅延する確率が下がり、影響を少なくすることが可能である (図 5)。また、遅延の要因の一つである、モニタリングプロセスが1回スケジュールされるたびに実行される時間を減らすことでも、遅延時間を削減することが可能である。以下では、この2つの最適化手法について述べる。

### 2.2.6 タイミングをそろえる

アプリケーションの同期処理が遅延する現象は、あるひとつの CPU の処理の遅れが、その完了を待機している他の多数のプロセスに影響を与えることで、より深刻になる。逆に、ある CPU の処理が遅れたときに、他の CPU でも遅延が発生していれば、相対的に影響は小さくなると言える。このために、モニタリングプロセスをノード間で同時にスケジュールさせることを検討する。ここで考えるモニタリングプロセスは、一定時間ごとに情報収集を実行するものであるから、最初の情報収集のタイミングがすべてのノードで揃っていれば、その後もすべてのノードで同時にプロセスが実行されると期待できる。具体的には以下のように動作させる。

- 動作間隔を  $I = 0.5(\text{sec})$  とする。
- まず、すべてのノードでほぼ同時に現在時刻を取得して、ノード  $n$  でのこの結果を  $S_n$  とする。
- 情報取得プロセスは、1回の情報取得を行った後、 $fmod(t_n - S_n, I) == 0$  となるまで sleep する。



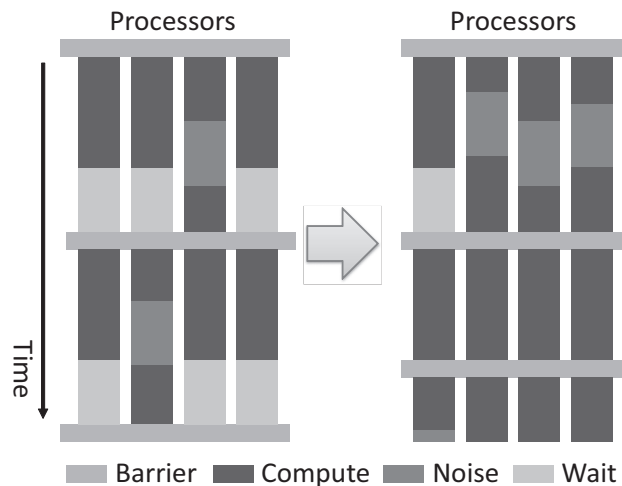


図 5: モニタリングプロセスの同期によるアプリケーション遅延の削減

ほぼ同時に現在時刻を取得するためには、MPI プログラムで各ノード 1 プロセスずつを含むコミュニケータを作成して MPI\_Barrier を実行した後に、各ノードのローカル時刻を取得することにした。各ノードで MPI\_Barrier から復帰した直後と同時にローカル時刻を取得しても、厳密には、物理的に同じ時刻を得られるわけではない。通信遅延の分布が既知の場合に利用することができる Probabilistic Clock Synchronization のような時刻同期プロトコルも提案されているが、アプリケーションの繰り返しにかかる時間や、モニタリングプロセスが情報取得にかかる時間と比べると、この時刻のずれは無視できる程度であると考え、今回は上記の厳密でない方法を用いることとした。

また、クロックの進む速度の違いなどの理由により、単に初期時刻をずらすだけでは並列処理のグローバル時刻を表すには不十分であるという報告もあるが、本稿で実行するアプリケーションの実行時間は短いため、このことは性能に大きな影響を与えないと考えている。

### 2.2.7 モニタリングプロセスの軽量化

モニタリング用の情報取得を行うプログラム pinfo が 1 回の情報収集で実行される時間は、HA8000 クラスタシステムで 200 程度のプロセスが動いている標準的な状態では、15 msec から 20 msec である。この大部分は、計算機上で動いている、プロセスごとの統計情報を取得する部分である。この処理では、/proc ファイルシステムから各プロセスの情報を取得する（これが、ps コマンドなどでも行われている最も普通の方法である）が、プロセスの情報は、プロセスごとに別々のファイルで提供されるため、プロセスの数だけ open、read、close システムコールが実行される。実際の情報取得にかかる時間は少ないのだが、このシステムコールのオーバーヘッドのために、プロセス数にほぼ比例する実行時間がかかり、1 度の情報取得あたり、10 msec 台の時間を必要としている。

この、プロセスごとの情報取得の負荷を減らすための方法として、Linux カーネルの taskstats インタフェースを利用する方法について述べる。

### 2.2.8 taskstats インタフェース利用

上記の方法は情報の正確性のある程度犠牲にして情報収集処理を高速化する方法であったが、Linux のプロセスアカウンティング機能の一部を使用することで、情報の正確性を変えることなく高速化することが可能である。Linux カーネルソース内のドキュメントのひとつである、Documentation/accounting/taskstats.txt によると、Linux Kernel の taskstats インタフェースは、netlink ソケット (Linux カーネルとユーザ空間の通信を、ソケット通信を用いて行うインタフェース) を利用して、プロセス ID を指定してリクエストを送り、そのプロセスに関する情報をレスポンス

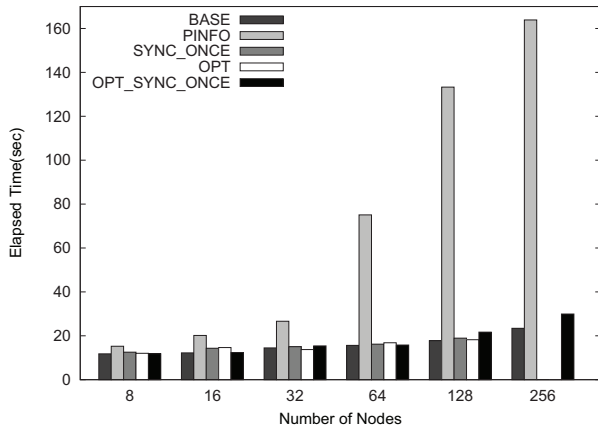


図 6: テストプログラムの実行時間

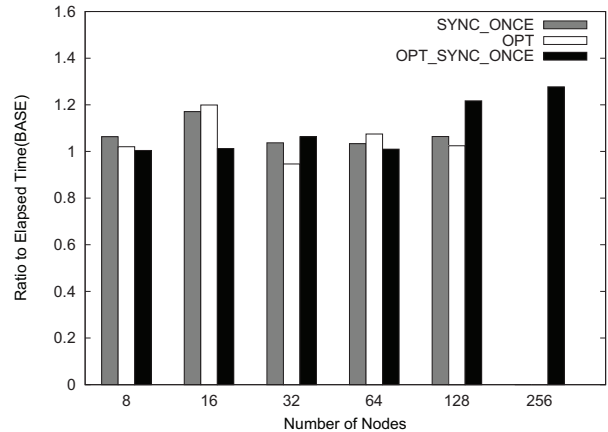


図 7: テストプログラムの実行時間 (BASE に対する割合)

として受け取るという機能を提供している。このインターフェースを利用することにより、プロセスごとにファイルをオープンする必要が最小限に抑えられるので、モニタリングプロセスの CPU 使用時間を大幅に少なくすることができる。(ただし、プロセスの一覧を取得する部分は依然として /proc ファイルシステムに頼る必要がある。)

この taskstats インタフェースを利用してプロセス毎の CPU 時間を取得する負荷を計測したところ、1 回の情報取得にかかる時間は 1 msec から 3 msec 程度となり、これまでの 10 分の 1 程度に抑えられることが確認できた。

### 2.2.9 実験

上記の最適化を適用したモニタリングシステムを作成し、並列計算への影響を評価した。評価の際に使用したプログラムは、予備実験で使用したアプリケーションをより単純化したものを用いた。具体的には、各プロセスが約 25ms 程度ローカルな計算をした後、全プロセスで Allreduce を行う処理を 4096 回繰り返すものとした。

このプログラムを、HA8000 クラスタシステムの A 群の、8 ノードから 256 ノードまでを使用して実行した時の時間を図 6 に示す。

グラフで、BASE はモニタリングプロセスがない場合、PINFO は最適化前の pinfo を使用した場合、SYNC\_ONCE は 2.2.6 で述べたタイミングをそろえる最適化を行った場合、OPT は pinfo に 2.2.8 で述べた「taskstats インタフェース利用」の最適化を行った場合である。OPT\_SYNC\_ONCE 等は、OPT と SYNC\_ONCE の組み合わせである。256 ノードの場合は、実験のために割り当てた時間の制約により、一部のケースしか試すことができなかった。

最適化前の pinfo を使用した場合、128 ノード、あるいは 256 ノード使用時は、実行時間が元の 7 倍を超えており、影響が甚大であることが分かる。図 7 には PINFO を除く各ケースについて、BASE に対する割合を示したものである。最適化を施すと、遅延への影響は多くても 20% 程度に抑えられていることが分かった。しかしながら、最適化前の pinfo と比べると大きく削減できたものの、20% の実行時間増加というのは決して無視できる程度のものではない。またこの実験だけでは、観測結果に揺れが大きく、ノード数の増加に従って遅延への影響が大きくなる現象は PINFO 以外では明確に表れたとは言えなかった。

次に、予備実験で用いたアプリケーションを、最適化した pinfo と同時に実行した場合の結果を図 8 に示す。PINFO+NICE は最適化前の pinfo を nice をかけて実行したときの結果である。単に nice をかけて優先度を下げるだけでは遅延削減の効果は限定的であることが分かる。ノード数が増えるほど、

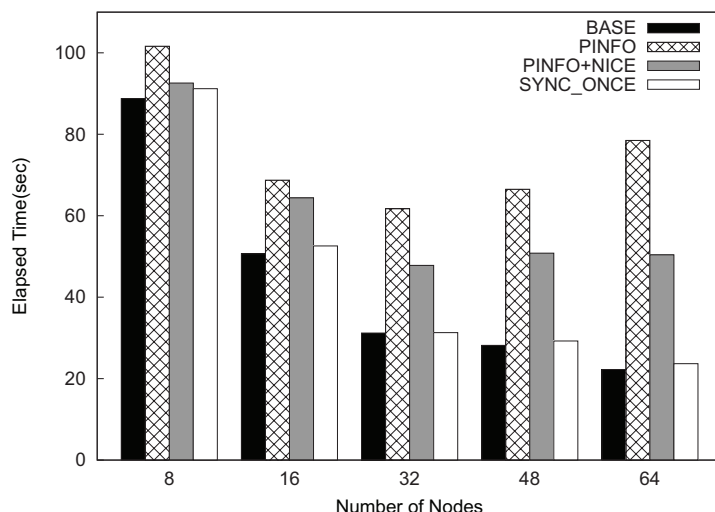


図 8: アプリケーションの実行時間

nice の効果は少なくなる。また、SYNC\_ONCE 最適化だけでも 64 ノード時の遅延は約 6% に削減できていることが観察された。

今後は、収集したデータの解析をさらに深く行い、遅延が発生する際の詳細な仕組みを明らかにし、また、他の OS プロセスや、Ganglia 等の他のモニタリングシステムの動作の特性についても調査を行う予定である。また、実験で観測された、実行時間のゆれについてもより詳細な解析を行う予定である。

### 2.3 具体的成果

上記の、システムプロセスが並列処理に与える影響についての調査について、[発表 1]において発表を行い、HA8000 クラスタシステムで VGXP の試験運用を行った際に得られた知見について、[招待 1]において講演を行った。

また、評価に用いたモニタリングプロセスは VGXP の一部であり、VGXP は、公開ソフトウェア GXP [公開 1] の機能の一部に組み込んで公開している。VGXP は、文部科学省科学研究費補助金特定領域研究「情報爆発に対応する新 IT 基盤研究プラットフォームの構築」の支援班が提供する分散計算機環境 InTrigger において、公式のモニタリングツールとして利用されている。

## 3 スパコン運用支援

### 3.1 背景

2008 年 6 月より運用開始した HA8000 クラスタシステムは、PC クラスタベースのシステムであり、無料の試行期間の利用者に対して行ったアンケートによれば、半数が本学のスーパーコンピュータの利用経験がなく、さらにその半数はスーパーコンピュータ自体の利用経験がない利用者である。このような新規利用者を含むスパコン利用者の利用実態を把握し、幅広い利用者の新規開拓とシステム運用の改善を継続的に行っていくことが重要である。

### 3.2 内容

日常的にスパコンシステムの運用支援を行っている中で、特に以下に挙げる活動などを実施している。

- 質問対応  
SR11000, HA8000 の利用方法等に関する相談・質問に他の職員・教員と共同で対応した。

- 利用促進  
並列プログラミング講習会やプログラミングコンテストを通してスパコンの効率的な利用方法を案内した。
- ソフトウェアの導入支援  
新規導入した Lustre ファイルシステムの導入・試験の支援を行った。また、講習会では Lustre の利用方法の案内も行った。
- グリッド配備・運用  
他大学と高速のネットワークで接続する2つのグリッド環境 (T2K グリッド、CSI グリッド) の配備と運用を行った。

### 3.3 具体的成果

スパコン利用促進のため、スパコン講習会 [特記 3] での講義や、プログラミングコンテスト [特記 1]、[特記 2] の開催を行った。CSI グリッドでは、全国共同利用施設であるスパコンセンターを設置する 7 大学と筑波大学、東京工業大学、国立情報学研究所によるグリッド配備・運用タスクフォース [特記 5] と、NAREGI の今後のアーキテクチャを検討するための NAREGI アーキテクトタスクフォース [特記 6] に参加し、議論を行った。グリッドの配備については、学際大規模情報基盤 共同利用・共同研究拠点の試行プロジェクト [特記 4] として位置づけ、計算資源の提供を行っている。

## 4 成果要覧

### 招待講演／招待論文

[招待 1] 鴨志田良和: クラスタ監視ソフト VGXP の開発・運用を通しての評価と課題, 第 9 回 PC クラスタシンポジウム, 2009 年 12 月.

### 公開ソフトウェア

[公開 1] Kenjiro Taura, Yoshikazu Kamoshida and Nan Dun: GXP: Grid and cluster shell, <http://www.logos.t.u-tokyo.ac.jp/gxp/>.

### その他の発表論文

[発表 1] 鴨志田良和, 田浦健次朗: 並列アプリケーションの性能を損なわないポーリング型のモニタリング, 情報処理学会研究報告 2009-HPC-121, 仙台, 2009 年 8 月.

### 特記事項

[特記 1] クラスタシステム上での並列プログラミングコンテスト 実行委員, 2009 年 3 月～5 月.

[特記 2] 第 2 回クラスタシステム上での並列プログラミングコンテスト 実行委員, 2009 年 7 月～12 月.

[特記 3] お試しアカウント付き並列プログラミング講習会 (試行): 2009 年 7 月, 2009 年 9 月, 2010 年 3 月.

[特記 4] 学術グリッド基盤の構築・運用技術に関する研究 共同研究者, 平成 21 年度 学際大規模情報基盤共同利用・共同研究拠点 公募型共同研究 (試行), 2009 年.

[特記 5] グリッド配備・運用タスクフォース.

[特記 6] NAREGI アーキテクトタスクフォース.

# ブロードキャスト型単一IPアドレスクラスタにおける 柔軟なリクエスト分散機構

藤田 肇

## 1 概要

ブロードキャスト型の単一IPアドレスクラスタサーバにおいて、クライアントからのリクエスト内容に応じて接続先サーバを決定可能にするリクエスト分散機構 DTS(Distributed TCP Splicing) の設計と実装を行った。DTS はサーバ OS カーネル内のパケットフィルタとして実現されるため、クライアント側システムやサーバアプリケーション、サーバ側 TCP/IP スタックの変更を必要とせずこのようなリクエスト分散を可能にする。また、リクエストの処理が担当サーバノードに割り当てられた後は各サーバノードが他のノードによる中継を必要とせずクライアントと通信を行うため、他のノードの故障に影響されず通信を継続することが可能である。

## 2 DTS: Distributed TCP Splicing

### 2.1 背景

インターネットを用いたサービスは既に我々の生活にとって不可欠なインフラとなっており、サービスを支えるサーバの性能向上と信頼性向上は重要な課題である。

インターネット上のサーバを高性能かつ高信頼にするための技術として、著者は単一IPアドレスクラスタの研究を行っている。これは複数の計算機からなるクラスタに1つのIPアドレスを割り当ててサーバとして機能させるもので、クライアントには1つのIPアドレスしか見えないため、クライアント側では実際のサーバ構成を気にすることがない。よってクライアントに意識させずにサーバクラスタ内のノード間で負荷を分散させたり、故障したノードを置き換えたりすることが可能になる。

単一IPアドレスクラスタを実現する手法として、著者はブロードキャスト型とよばれる手法を用いている。これは、クラスタサーバを構成するすべてのノードに同一のIPアドレスを設定し、さらにクラスタ宛に送られてくるパケットが全ノードにブロードキャストされるようにスイッチやルータを設定しておくことによって単一IPアドレスクラスタを実現する手法である。ブロードキャスト型単一IPアドレスクラスタでは、クラスタ外からのパケットを中継するための特別なノードが必要ないため、あるサーバノードの故障が他のノードの通信に影響を及ぼさないという特徴がある。著者はこのようなブロードキャスト型単一IPアドレスクラスタにおいてクライアントからのTCP接続を任意のノードに割り当てる手法を研究し、その一部を2008年度年報にて報告した。

2008年度年報にて報告した手法は、クライアントからサーバに対するリクエストが送られる前、TCP接続が確立する時点で接続先サーバノードを決定し割り振っていた。これに対して、クライアントからのリクエストに応じて担当サーバノードを決定するとよい場合がある。TCPを用いた代表的なアプリケーションであるWebを例にすると、特定のノードにしか存在しないファイルを要求された場合にはそのノードに処理を担当させないと、ファイルが見つからないというエラーが返ることになる。そこで2009年度は、クライアントからのリクエスト内容に応じて接続先サーバノードを決定するための手法の研究を行った。



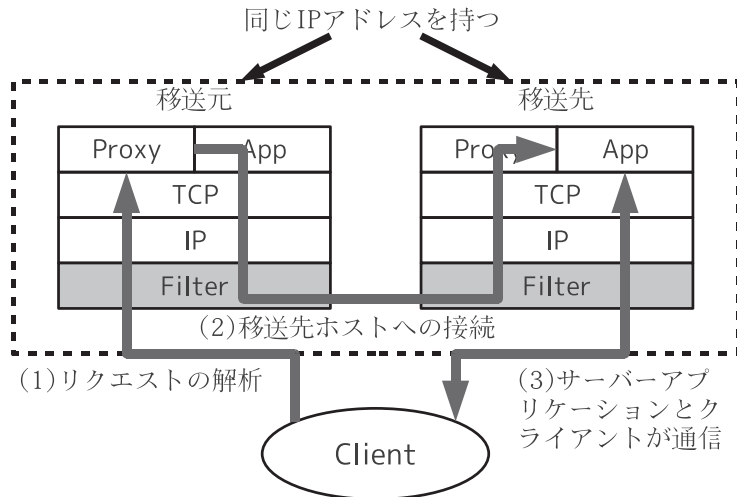


図 1: DTS の概要

なお、本研究は JST CREST のプロジェクト「実用化を目指した組込みシステム用ディペンダブル・オペレーティングシステム」の一部として行われたものである。

## 2.2 内容

ブロードキャスト型単一 IP アドレスクラスタにおいてクライアントからのリクエスト内容に応じた TCP 接続の割り当てを実現する機構 DTS(Distributed TCP Splicing) を提案した。これはクライアントから TCP を用いて送られてきたリクエストの内容を解析し、それに適した適切なサーバノードに TCP 接続の処理を委譲することで、複数のサーバノード間での処理の分散を図るものである。DTS は以下の特徴をもつ。

- リクエスト内容に応じてクライアントからの TCP 接続を適切なサーバノードに振り分けることが可能
- TCP 接続の振り分けが完了した後は、他のノードによるパケットの中継や変換を必要とすることなく通信を継続することが可能
- DTS を利用するためにクライアント側のアプリケーションや TCP スタックの改変が必要ないのはもちろん、サーバ側アプリケーションやサーバ側 OS の TCP スタックの改変も不要

DTS の構造を図 1 に示す。図には 2 台のサーバノードがクライアントと通信する様子を示した。各サーバノード上には、DTS の機能を実現するためのパケットフィルタ (図中“Filter”) がインストールされる。このパケットフィルタは OS カーネル内で動作し、ネットワークデバイスと IP スタックの間に位置する。また各サーバノードでは、実際にクライアントにサービスを提供するサーバアプリケーション (図中“App”) と、クライアントからのリクエスト内容を解析するためのプロキシ (図中“Proxy”) が動いている。前述したように、DTS はブロードキャスト型単一 IP アドレスクラスタを前提としている。そのため、各サーバノードは同一の IP アドレスを持ち、クライアントからサーバに向かうパケットは全サーバノードにブロードキャスト配送されるように設定されている。

クライアントからのリクエスト内容に応じて TCP 接続をあるノードに割り当てるためには、サーバノードの 1 つとクライアントとの間で一旦 TCP 接続を確立させ、クライアントから送られてくるリクエスト内容を解析する必要がある。そこで、DTS ではまずサーバノードの任意の 1 台でクライアントとの TCP 接続を確立させ、クライアントからのリクエストを受信する。この段階では、クライアント



```
int request_handover(int socket,
                    const struct sockaddr *dest, socklen_t addrlen,
                    const char *request_buf, size_t buflen)
```

図 2: TCP 接続移送を要求するシステムコール

表 1: DTS TCP オプションのフィールド

フィールド	長さ	説明
オプション種別	8 ビット	現在のところ値 32(10 進数)を使用
Len	8 ビット	このオプションの長さ(オクテット単位) オプション種別及び Len フィールド(合計 2 オクテット)も含む
DTSraddr	32 ビット	クライアントの IP アドレス
DTSrport	16 ビット	クライアントのポート番号
DTSinitseq	32 ビット	接続移送元のプロキシがクライアントに送った SYN+ACK セグメントのシーケンス番号
DTSremotewin	32 ビット	TCP 接続移送開始時点でのクライアントのウィンドウサイズ
DTShandoverlen	32 ビット	移送先にこれから送られるデータの長さ

との TCP 接続はサーバーアプリケーションではなく、プロキシとの間で確立するよう、パケットフィルタにおいてポート番号の変換が行われる。

プロキシはユーザーランドのプログラムとして動作し、通常のサーバプログラムのようにソケット API を用いて接続の待ち受けやデータの受信を行う。リクエストを解析し、実際に処理を担当べきノードが決定できたら、OS カーネルに対して TCP 接続を担当ノードに移送するよう要求するシステムコールを発行する(図 2)。このシステムコールには、既にプロキシが読み取ったリクエスト内容を含むバッファへのポインタを渡す。

TCP 接続の移送が要求されると、カーネルは移送先のノードに対して新しい TCP 接続を張り、request\_handover システムコールに渡されたリクエスト内容と、既にプロキシの TCP ソケットの受信バッファ内に入っているデータを送信する。このとき、移送元から移送先に TCP 接続を作成する際、最初の SYN セグメントに表 1 に示す TCP オプションを付加し、これによって移送先の DTS パケットフィルタにこれから移送される TCP 接続の情報を通知する。

移送先のノードでは、移送元からのパケットのヘッダをパケットフィルタによって書き換え、発信元をクライアントの IP アドレスおよびポート番号にする。これによって、移送先 OS の TCP/IP スタックからは、その接続要求が直接クライアントから来ているように見える。移送先ノードでは、このようにして送られてきたデータをプロキシではなく、サーバーアプリケーションに直接渡す。移送元からすべてのデータを受信すると、移送先ではそれ以後クライアントとの間でのパケットのやりとりを開始する。すなわち、移送先ノードの TCP/IP スタックから見ると、移送元ノードからのパケットも、クライアントからのパケットも同一の TCP 接続に属するものに見える。つまり、移送元-移送先の TCP 接続と、移送先-クライアントの 2 つの TCP 接続を透過的に接合(splice)していることになる。

2 つの TCP 接続を同一視して透過的に接合するためには、宛先/発信元の IP アドレスとポート番号の調整だけでなく、TCP ヘッダ中のシーケンス番号と確認応答番号が連続するように値を調整しなければならない。調整の具体例を図 3 に示す。まず、クライアントがサーバに対して TCP 接続要求を出す。これは、SYN ビットの立った TCP セグメントを送ることによってなされる(図中セグメント 1)<sup>1</sup>。

<sup>1</sup>サーバ側はブロードキャスト型クラスタであるので、本来、セグメント 1 は移送元ノードだけでなく移送先ノードにも届いているが、図では煩雑になるため省略している。セグメント 1 は移送先ノードに届くが、パケットフィルタによって受け入れられるべきではないと判断され、破棄される。

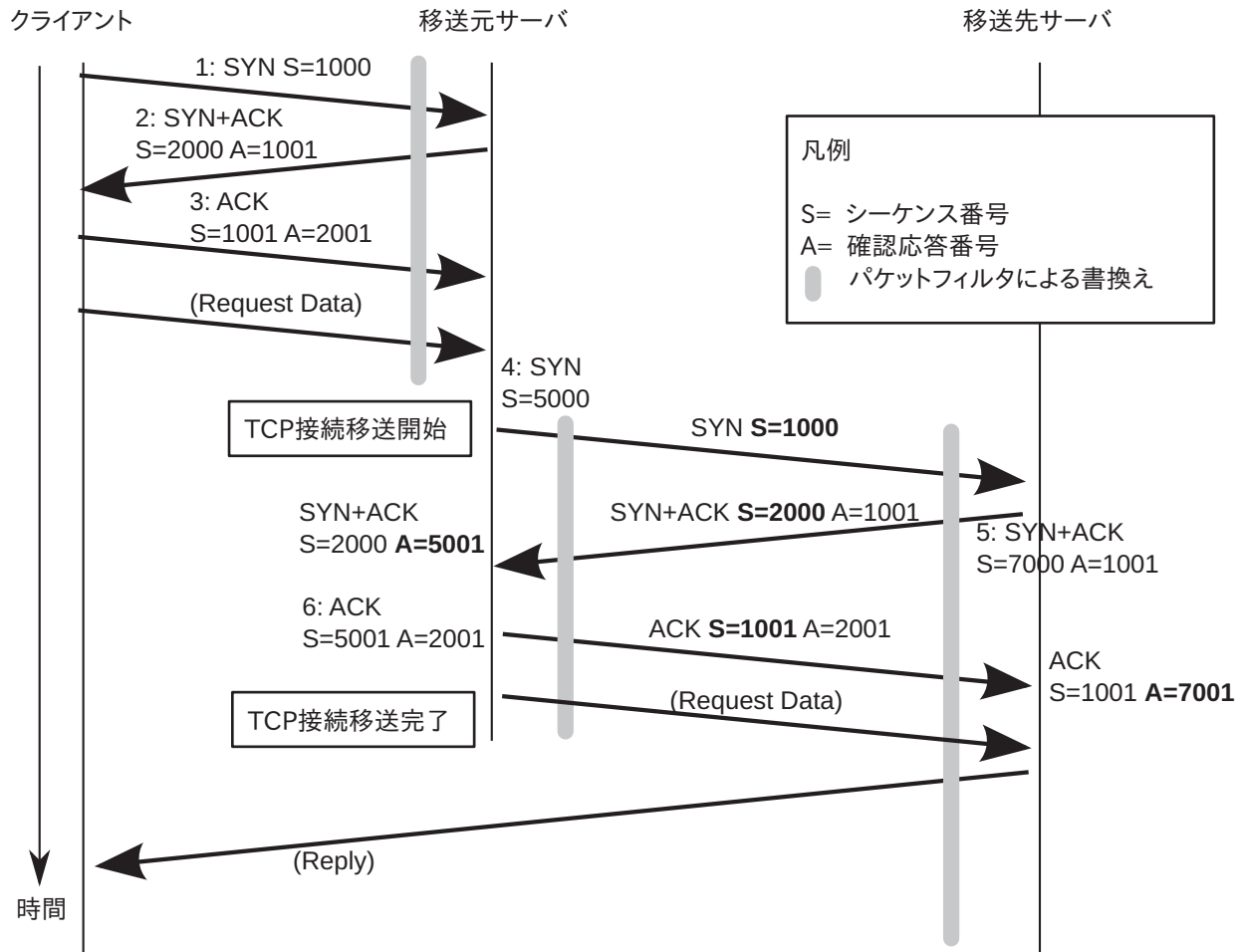


図 3: パケットの流れとシーケンス番号の書き換え (太字は書換えられた箇所を示す)

セグメント 1 に続き、TCP の 3-way handshake の手順にしたがってセグメント 2, 3 が交換され、TCP 接続が確立する。ここで、移送元サーバで実際にはデータを受け取るのはプロキシである。TCP 接続が確立したのち、クライアントはリクエスト内容を含むデータセグメントを送る。

プロキシが移送先を判断し、`request_handover` システムコールを発行すると、移送先への TCP 接続が試みられ、4 番の TCP セグメントが送出される。このとき、このセグメント 4 のシーケンス番号は移送元サーバの TCP スタックがランダムに決定するが、このセグメント 4 が移送元サーバから送出される際に、セグメント 1 のシーケンス番号 (図の例では 1000) と等しくなるように書換えが行われる。このセグメントは移送先サーバに到着すると、さらに発信元 IP アドレスおよびポート番号がクライアントのものに書換えられる。よって、移送先サーバの TCP/IP スタックにとっては、このセグメントは 1 番のセグメントとまったく同じものに見える。TCP 接続要求を受け取った移送先ノードは次に 5 番の packets を返送するが、これも同様にシーケンス番号はランダムに決定されているので、送出時に 2 番のセグメントのシーケンス番号 2000 と一致するように書換える。また、セグメント 5 はクライアント宛の packets として移送先から送出されるが、これが移送元サーバに戻るよう、アドレス及びポート番号の変換を行う。さらに 5 番のセグメントが移送元サーバに受信される際には、4 番の送出時にシーケンス番号を書換えたのに対応して確認応答番号を 1001 から 5001 に書換える。

同様にしてリクエストのデータすべてを移送先ノードが受信し終わると、TCP 接続の移送処理は完了し、以後移送先ノードから送出される packets は宛先を変えずに直接クライアントへ向けて送信される。この時点で、移送先サーバノードが送出する TCP セグメントは packets フィルタでの調整を経て、クライアントが受信を期待するシーケンス番号になっている。逆も同様に、クライアントから送られてくるセグメントは移送先サーバが期待するシーケンス番号となっており、クライアントと実際に通信しているノードが切り替わっているにも関わらず、クライアントから見て途切れずに通信が継続できる。また、移送の完了後は移送元サーバは通信に関与しないため、仮に移送元サーバに障害が発生してもクライアントと移送先サーバの間の通信は正常に継続できる。

## 2.3 具体的成果

提案手法 DTS を Linux カーネル内で動作する packets フィルタモジュールとして実装した。これは完全に動的ロード可能なモジュールとして実装されているため、カーネル本体への静的パッチは必要ない。DTS の設計と実装に関して、情報処理学会 OS 研究会で発表した [発表 1]。さらに DTS を応用したデモシステムとして、分散ファイルサーバを構築した。これは複数の WebDAV サーバを DTS を用いて単一サーバとして見せるものである。また、JST CREST プロジェクトの取り組みを紹介する特集の一部として、単一 IP アドレスクラスタに関する研究について、技術評論社 Software Design 誌に記事を執筆した [著書 1]。

## 3 成果要覧

### 査読付論文

[査読付 1] Balazs Gerofi, Hajime Fujita, Yutaka Ishikawa: Live Migration of Processes Maintaining Multiple Network Connections, コンピュータシステム・シンポジウム (ComSys2009), pp. 1-10, November 2009.

[査読付 2] Balazs Gerofi, Hajime Fujita, Yutaka Ishikawa: Live Migration of Processes Maintaining Multiple Network Connections, IPSJ Transactions on Advanced Computing Systems Vol.3 No.1 (ACS 29), pp. 1-12, March 2010 ([査読付 1] と同時投稿).

## その他の発表論文

- [発表 1] 藤田肇, 石川 裕: レイヤー 7 負荷分散のための TCP 接続移送機構, 情報処理学会研究報告 (OS-111), 那覇, Apr 2009.
- [発表 2] Balazs Gerofi, Hajime Fujita, Yutaka Ishikawa: TCPmig: Migration of processes with TCP sockets in Single IP Address Cluster, 情報処理学会研究報告 (OS-111), 那覇, Apr 2009.
- [発表 3] 中澤仁, 松野裕, 菅谷みどり, 埜敏博, 前田俊行, 藤田肇, 石綿陽一, 杵渕雄樹, 高村博紀, 三浦信一, 山田浩史, 松田元彦, 倉光君郎, 石川裕: オペレーティングシステムおよび実システムにおけるディペンダビリティの評価と見積り, 第 7 回ディペンダブルシステムワークショップ (DSW2009), pp. 27-41, 函館, July 2009.
- [発表 4] 加藤真平, 藤田肇, 中澤仁, 松田元彦, 前田俊行, 杵渕雄樹, 埜敏博, 三浦信一, 石綿陽一, 松野裕, 高村博紀, 山田浩史, 吉田哲也, 倉光君郎, 菅谷みどり, 石川裕: ディペンダブルシステム向けベンチマークフレームワークの提案, 第 7 回ディペンダブルシステムワークショップ (DSW2009), pp. 171-178, 函館, July 2009.
- [発表 5] 加藤 純, 藤田 肇, 石川 裕: 単一 IP アドレスクラスタにおける耐故障機構の設計と実装, 情報処理学会研究報告 (OS-112 SWoPP 2009), 仙台, Aug 2009.

## 著書／編集

- [著書 1] 藤田肇: ディペンダブル・シングルシステムイメージ OS, ソフトウェアデザイン 2010 年 2 月号, pp.89-92, 技術評論社.



PART 4

各部門

# 教育・サービス活動報告

情報メディア教育部門

図書館電子化部門

ネットワーク部門

(全国共同利用)

スーパーコンピューティング部門





## 情報メディア教育部門



情報基盤センター1階大演習室(1)と教育用計算機システム標準端末



# 情報メディア教育部門

## 概要

部門長 柴山 悦哉

情報教育支援係長 増田 均

情報メディア教育部門の主要なミッションは、教育研究目的で利用可能なさまざまな機器、ソフトウェア、サービスなどを、東京大学に所属する学生および教職員に対して提供することである。そして、これに付随する管理運用、ユーザサポート、システム設計、研究開発などの作業も行っている。現在提供中のシステムやサービスには以下のようなものがある。これらは、研究室、学科、部局等で個別に運用管理を行うことが困難であったり、大学全体で共有して管理運用を集約化する方が費用面等で有利であったりするものである。

**教育用計算機システム(ECCS):**本郷、駒場、柏の3キャンパスに分散配置した 1,176 台の iMac 端末と 186 台の Windows 端末を約 24,000 名の登録ユーザに提供している。これらの端末には、Office アプリケーションや数式処理システムなどの商用アプリケーションと多数のフリーソフトウェアが搭載されており、授業での利用をはじめとした教育目的を中心にさまざまな用途に利用されている。一般ユーザが直接触れることはないが、これらを一元管理するために、ブートサーバ、ファイルサーバなどを用いた集中管理体制を取っている。端末の他にも、プリントサービス、メールサービス、遠隔ログインサービスなどを各ユーザに対して提供している。また、新規利用者向け講習会などのソフトなサービスにも力を入れている。

**メールホスティング:**学内のドメインを単位としたメールホスティングサービスを提供しており、470 近くのドメインに所属する 16,000 名以上のユーザが利用している。ECCS ユーザ向けのメールサービスと合わせると、延ユーザ数は 40,000 名以上に達する。

**Web ホスティング:**学内の 650 近くのドメインに対して、Web のホスティングサービスを提供している。

**認証サービス:**ECCS の登録ユーザ情報を活用した認証サービスも一部で行っており、無線 LAN、SSL-VPN などがこれに該当する。さらに一部の部局に対しては、部局所有の端末に ECCS の認証情報を提供するサービスも行っている。

**ストリーミングサービス:**映像配信用のストリーミングサーバを利用したサービスの提供を行っており、UTOCW、TODAITV、卒業式中継をはじめさまざまな用途に利用されている。また、映像作成のための機材の提供も行っている。

**CFIVE :** 講義支援のためのオープンシステムとして CFIVE の開発と運用を行っており、学生証を用いた出欠管理、課題の出題とレポートの提出、教材のオンライン配信などを支援する機能を提供している。CFIVE は、140 以上の講義で利用されている。

**遠隔講義支援:**テレビ会議システムを備えた遠隔講義室、多地点接続装置などを用意しており、遠隔講義や遠隔会議に利用されている。

# 教育用計算機システム運用報告

情報教育支援係

情報リテラシー教育支援係

## 1 運用報告

### 1.1 Mac(iMac, Xserve)

教育用計算機システム全体では標準端末として iMac 端末(1,176 台)を設置している。iMac 端末はブートサーバ(Apple Xserve)から OS(Mac OS X)のイメージを読み込んで起動する。1 台のブートサーバは iMac 端末 30 数台を制御し、33 台(駒場 20 台、本郷 13 台)のサーバを設置している。また、iMac 端末にインストールされている Parallels Desktop というソフトウェア上で Windows Vista を起動し、Windows 環境を利用することができる。

システムイメージの更新は OS のバージョンアップ、セキュリティアップデート、アプリケーションのインストールやアップデートに伴って行っている。

2009 年度の主な更新は以下の通りである。なお、不具合の修正や広報の表示の際にもイメージの更新を行うことがあった。

#### iMac 端末(Mac 環境)

##### 2009 年 4 月 27 日 iMac 端末のシステム更新

- iTunes 8.1.1
- Microsoft Office 2004 for Mac 11.5.4
- Microsoft Office 2008 for Mac 12.1.7
- MATLAB 2009a インストール
- MATLAB 2008a 削除
- Firefox 3.0.9
- Thunderbird 2.0.0.21
- VLC 0.9.9a

##### 2009 年 5 月 18 日 iMac 端末のシステム更新

- Mac OS X 10.5.7
- Firefox 3.0.10
- Adobe Reader 9.1.1

##### 2009 年 6 月 25 日 iMac 端末のシステム更新

- Microsoft Office 2008 for Mac 12.1.9
- Microsoft Office 2004 for Mac 11.5.5
- Adobe Reader 9.1.2
- Firefox 3.0.11
- Thunderbird 2.0.0.22
- Java for Mac OS X 10.5 アップデート 4
- iTunes 8.2
- QuickTime Player 7.6.2

##### 2009 年 8 月 5 日 iMac 端末のシステム更新

- Mathematica 7.0.1
- Microsoft Office 2008 for Mac Service Pack 2 (12.2.0)
- Adobe Reader 9.1.3

- FireFox 3.0.12
  - FireFox 3.5.2 インストール
  - Flash Player 10.0.32.18
  - iTunes 8.2.1
- 2009年9月24日 iMac 端末のシステム更新
- Mac OS X 10.5.8
  - Security Update 2009-005
  - Safari 4.0.3
  - Java for Mac OS X 10.5 アップデート 5
  - iTunes 9.0.0
  - QuickTime 7.6.4
  - Remote Desktop クライアントアップデート 3.3.1
  - Parallels Desktop 4.0 for Mac
- 2009年11月2日 iMac 端末のシステム更新
- iTunes 9.0.1
  - MS Office 2008 for Mac 12.2.1
  - Adobe Reader 9.2
  - Firefox 3.0.15
  - Firefox 3.5.4
  - Thunderbird 2.0.0.23
  - The Unarchiver 2.1
  - Shockwave Player 11.5.1.601
  - Flip4Mac 2.3.0.14
- 2009年11月27日 iMac 端末のシステム更新
- Safari 4.0.4
  - Security Update 2009-006
  - iTunes 9.0.2
  - Microsoft Office 2004 for Mac 11.5.6
  - Microsoft Office 2008 for Mac 12.2.3
  - FireFox 3.5.5
  - Shockwave Player 11.5.2.602
- 2009年12月24日 iMac 端末のシステム更新
- Java for Mac OS X 10.5 アップデート 6
  - Graphviz 2.26.0 インストール
  - FireFox 3.0.16
  - FireFox 3.5.6
  - Flash Player 10.0.42.34
  - VLC media player 1.0.3
- 2010年2月5日 iMac 端末のシステム更新
- セキュリティアップデート 2010-001
  - iTunes 9.0.3
  - Adobe Reader 9.3
  - Firefox 3.6 インストール
  - Firefox 3.5.7
  - Firefox 3.0.x 削除
  - Shockwave Player 11.5.6.606



- 2010年3月15日-16日、3月25日 iMac用ブートサーバのOSバージョンアップ
  - MacOSX 10.5(Leopard)から10.6(Snow Leopard)へのバージョンアップ
- 2010年3月29日 iMac 端末のOSバージョンアップ
  - MacOSX 10.5(Leopard)から10.6(Snow Leopard)へのバージョンアップ

iMac 端末(Windows 環境)

- 2009年4月21日 iMac 端末のWindows 環境のシステム更新
  - Windows Update / Office Update
  - ウイルスバスター2008 パターンファイル等 update
  - Firefox 3.0.8
- 2009年5月27日 iMac 端末のWindows 環境のシステム更新
  - Windows Update / Office Update
  - ウイルスバスター2008 パターンファイル等 update
  - Firefox 3.0.10
  - Adobe Reader 9.1.1
- 2009年6月29日 iMac 端末のWindows 環境のシステム更新
  - Windows Update / Office Update
  - ウイルスバスター2008 パターンファイル等 update
  - Firefox 3.0.11
  - Thunderbird 2.0.0.22
  - Adobe Reader 9.1.2
  - Java Version 6 Update 14
- 2009年8月5日 iMac 端末のWindows 環境のシステム更新
  - Windows Update / Office Update
  - ウイルスバスター2008 パターンファイル等 update
  - Firefox 3.0.12
  - Firefox 3.5.1 インストール
  - Flash Player 10.0.32.18
  - Shockwave Player 11.5.1.601
- 2009年9月24日 iMac 端末のWindows 環境のシステム更新
  - Parallels Desktop 4.0 for Mac アップデート対応
  - Windows Update
  - ウイルスバスター2008 パターンファイル等 update
- 2009年10月27日 iMac 端末のWindows 環境のシステム更新
  - Windows Update
  - ウイルスバスター2008 パターンファイル等 update
  - Adobe Reader 9.2
  - Firefox 3.0.14
  - Firefox 3.5.3
  - Java Version 6 Update 16
  - Thunderbird 2.0.0.23
- 2009年12月4日 iMac 端末のWindows 環境のシステム更新
  - Windows Update
  - ウイルスバスター2010 インストール
  - ウイルスバスター2008 削除
  - EViews 6 インストール
  - EViews 5 削除

- Firefox 3.0.15
- Firefox 3.5.5
- Java Version 6 Update 17
- Shockwave Player 11.5.2.602

2010年1月18日 iMac 端末の Windows 環境のシステム更新

- Windows Update
- ウイルスバスター2010 パターンファイル等 update
- Adobe Reader 9.3
- Firefox 3.0.17
- Firefox 3.5.7
- Flash Player 10.0.42.34

2009年3月30日 iMac 端末の Windows 環境のシステム更新

- Windows Update / Office Update
- ウイルスバスター2008 パターンファイル等 update
- Firefox 3.0.7
- Thunderbird 2.0.0.21
- Adobe Reader 9.1
- Java Version 6 Update 13
- Flash Player 10.0.22.87
- Shockwave Player 11.5.0.595

Unix サーバ

2010年3月16日 Unix サーバの OS バージョンアップ

- MacOSX 10.5(Leopard)から 10.6(Snow Leopard)へのバージョンアップ

## 1.2 CAD 端末(Windows)

CAD 端末は、CAD のアプリケーションを必要とする講義などで用いることを想定した Windows Vista ネットブート型シンクライアントシステムの端末であり、本郷・駒場に合計 186 台設置している。

Windows において、ディスクイメージファイルの更新時には OS のセキュリティアップデートとウイルス対策ソフトのパターンファイル更新を必須に行い、その他にソフトウェアの更新などを行った。

2009 年度の主な更新は以下の通りである。

CAD 端末

2009年4月21日 CAD 端末のシステム更新

- Windows Update / Office Update
- ウイルスバスター2008 パターンファイル等 update
- MATLAB 2009a インストール
- MATLAB 2008a 削除
- R 2.8.1 インストール
- Firefox 3.0.8
- Java Version 6 Update 13
- Shockwave Player 11.5.0.595

2009年5月29日 CAD 端末のシステム更新

- Windows Update / Office Update
- ウイルスバスター2008 パターンファイル等 update

- Firefox 3.0.10
- Adobe Reader 8.1.5
- VLC media player 0.9.9
- 2009年7月2日 CAD 端末のシステム更新
- Windows Update / Office Update
- ウイルスバスター2008 パターンファイル等 update
- Mathematica 7.0.1
- Adobe Reader 8.1.6
- Firefox 3.0.11
- Thunderbird 2.0.0.22
- Java Version 6 Update 14
- QuickTime 7.6.2
- Shockwave Player 11.5.0.600
- 2009年8月4日 CAD 端末のシステム更新
- Windows Update / Office Update
- ウイルスバスター2008 パターンファイル等 update
- FireFox 3.0.12
- FireFox 3.5.1 インストール
- Flash Player 10.0.32.18
- Shockwave Player 11.5.1.601
- 2009年9月18日 CAD 端末のシステム更新
- Windows Update(Windows Vista Service Pack 1 を含む)
- ウイルスバスター2008 パターンファイル等 update
- 2009年12月15日 CAD 端末のシステム更新
- Windows Update
- ウイルスバスター2010 インストール
- ウイルスバスター2008 削除
- EViews 6 インストール
- Adobe Reader 8.1.7
- FireFox 3.0.15
- FireFox 3.5.5
- Java Version 6 Update 17
- Thunderbird 2.0.0.23
- QuickTime 7.6.5
- VLC media player 1.0.3
- Flash Player 10.0.42.34
- Shockwave Player 11.5.2.602
- 2010年1月27日 CAD 端末のシステム更新
- Windows Update
- ウイルスバスター2010 パターンファイル等 update
- Adobe Reader 8.2
- FireFox 3.0.17
- FireFox 3.5.7
- Java Version 6 Update 18
- Shockwave Player 11.5.6.606

2010年3月25日 CAD 端末のシステム更新

- Windows Update (Internet Explorer 8 を含む)
- ウイルスバスター2010 パターンファイル等 update
- Adobe Reader 8.2.1
- FireFox 3.6.2 インストール
- FireFox 3.5.8
- FireFox 3.0.17 削除
- Flash Player 10.0.45.2
- JMP 8.0.2
- RealPlayer SP
- Thunderbird 3.0.3

## 2 サービス関係

### 2.1 分散配置端末の配置変更

9月24日、農学生命科学研究科により、端末の利用時間延長と端末利用率の向上を目的として、農学部7号館B棟2階238号室に設置していたiMac端末16台が農学生命科学図書館3階PC端末室2に移設された。この移設に係る費用は農学生命科学研究科が負担した。

### 2.2 プリンタの配置変更

3月24日駒場地区、3月25日日本郷地区において、使用頻度の高いプリンタと使用頻度の低いプリンタを次の通り配置替えした。

<駒場地区>

場所	台数	配置換	場所	台数
駒場図書館1階メディアパーク	2台	↔	情報教育棟1階事務室	1台
			情報教育棟1階業務室	1台
情報教育棟1階相談室	2台	↔	情報教育棟2階小演習室1	1台
			情報教育棟3階小演習室2	1台
情報教育棟2階大演習室1	1台	↔	情報教育棟2階大演習室1	1台
情報教育棟2階大演習室3	1台	↔	情報教育棟3階大演習室2	1台

<本郷地区>

場所	台数	配置換	場所	台数
総合図書館2階メディアプラザ2	2台	↔	薬学図書館4階情報処理室	1台
			福武ホール地下1階	1台
総合図書館3階メディアプラザ3	2台	↔	情報基盤センター4階業務室	1台
			情報基盤センター1階自習室前	1台
経済学部1階101	1台	↔	経済学部4階407	1台
文学部法文2号館1階文学部学生ホール	1台	↔	文学部法文1号館1階111	1台

### 2.3 ファイル容量、メール容量制限値の緩和

3月10日より、iMac 端末の OS バージョンアップおよびファイルサーバの使用状況から、ファイル容量とメール容量の制限値が緩和された。詳細は次の通り。

- ・ファイル容量制限値
  - 学生 1.5GB → 2.5GB
  - 教職員 4GB → 6GB
- ・メール容量制限値
  - 全ユーザ 200MB → 300MB

## 3 部局別実利用者数

教育用計算機システムを一度でも利用したことのある学部学生、大学院生、研究生、教職員の月毎累積の部局別実利用者数である。比較のため、2009年度(各月右側)と2008年度(同左側)を掲せてある。

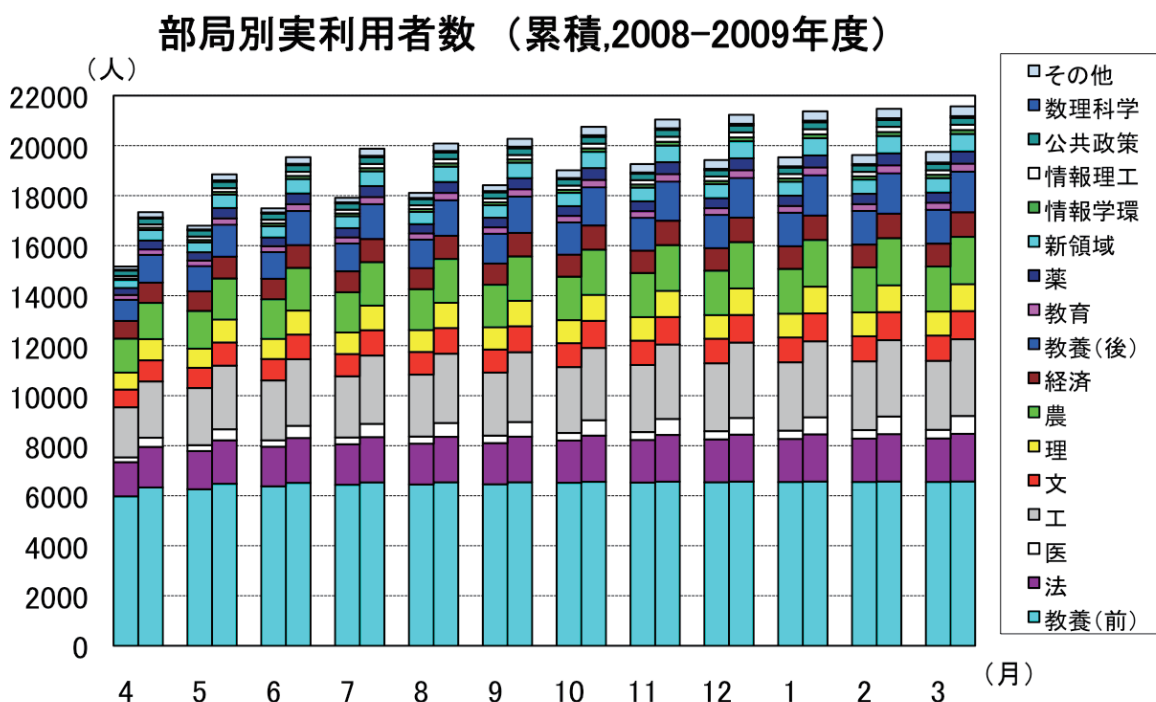


図 1. 部局別実利用者数

### 3.1 端末利用状況

情報メディア教育部門では、教育用計算機システムを管理しており、その端末等を本郷キャンパスと駒場キャンパスに配置している。本郷地区では、情報基盤センター(浅野地区)の演習室に端末等を集中配置すると共に、総合図書館、留学生センター、福武ホール、各学部/研究科等の 15 箇所にも端末を分散配置している。駒場地区では、情報教育棟に端末等を集中配置すると共に駒場図書館にも端末を分散配置している。また、柏地区では柏図書館に端末を分散配置している。

これらの配置場所のうち、利用の多い本郷の総合図書館と駒場の情報教育棟と駒場図書館のiMac 端末利用状況(2009年12月16日)は図2から図6の通りである。

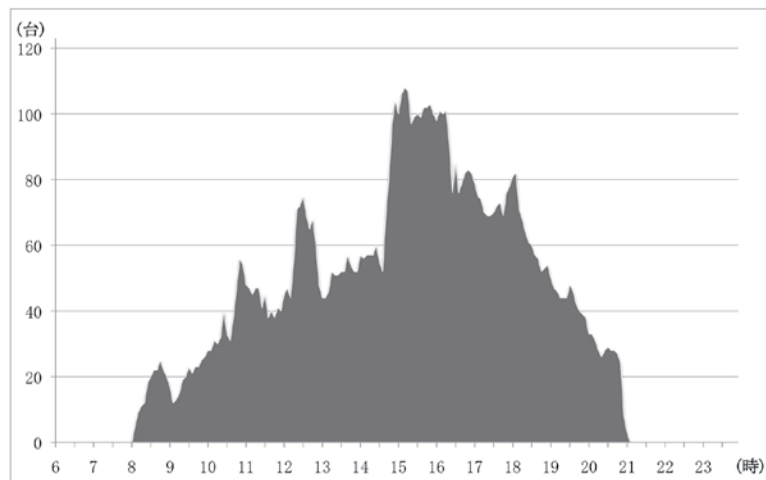


図2. 駒場情報教育棟 1階自習室 (123台)

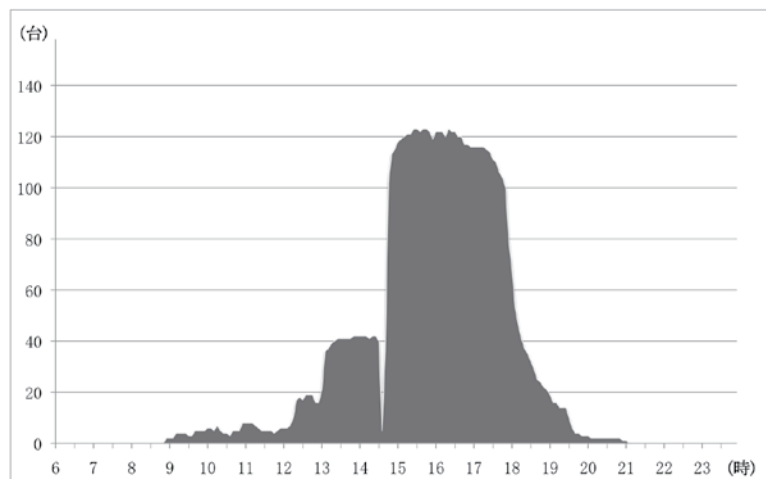


図3. 駒場情報教育棟 2階大演習室1 (158台)

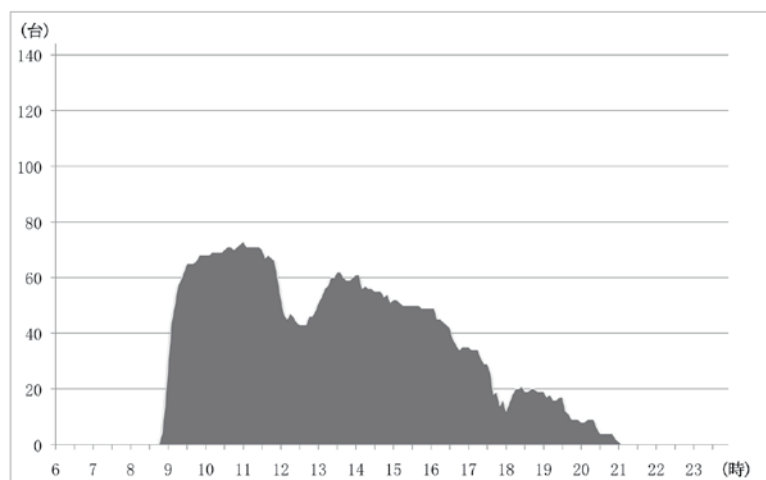


図4. 駒場情報教育棟 3階大演習室2 (144台)



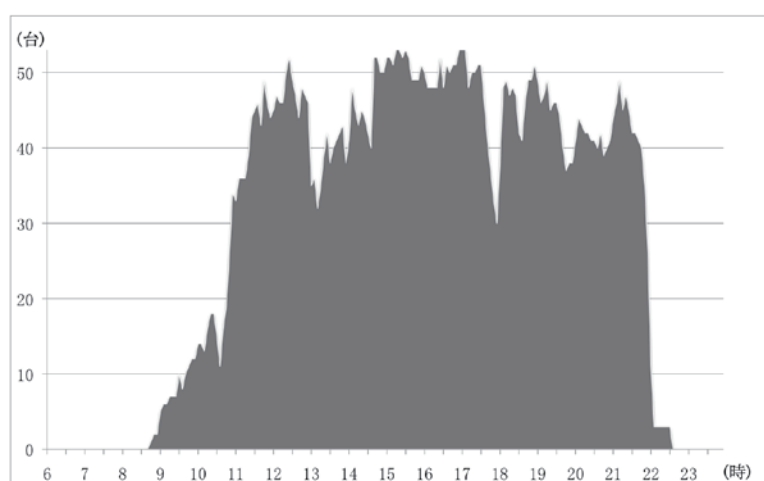


図 5. 駒場図書館 1F-メディアパーク (53 台)

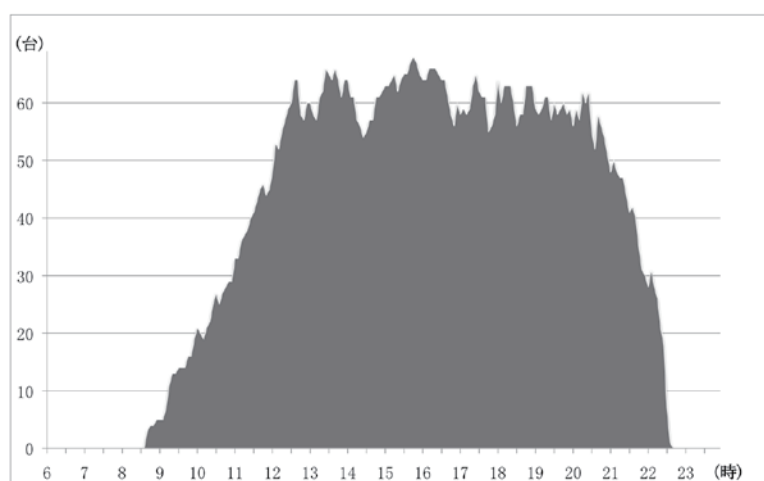


図 6. 総合図書館 2F-メディアプラザ 2, 3F-メディアプラザ 3 (69 台)

#### 4 柏キャンパス出張講習会

学部学生、大学院生および研究生が教育用計算機システムの利用を新規に申し込む際には、新規利用者向け講習会を受講する必要がある。しかし、柏キャンパスに所属する学部学生、大学院生については、本郷への移動は時間がかかるため、受講者の負担が大きく、また、遠隔キャンパスとしては所属人数が多いという事情を考慮して出張講習会を開催している。

2009 年度の実施状況は以下の通りである。

開催日	講習会受講者数
4 月 15 日(水)	101 名
4 月 24 日(金)	66 名
5 月 12 日(火)	34 名
10 月 14 日(水)	24 名

## 5 柏キャンパスでのアカウント発行業務委託の業務追加について

柏キャンパス在籍の学生の便宜をはかるため、2008年10月1日より柏図書館にて ECCS のアカウントの代行配付を行うこととなり、柏図書館と情報基盤センターの間で覚書を取り交わしている。

当初の業務は

- ・教育用計算機システム利用者登録済通知書(新規)の配付
- ・IC カードリーダ読み取り方式による登録済通知書(継続)の配付

であったが、2009年4月より

・IC カードリーダ読み取り方式によるパスワード再発行通知書の配付を、2009年6月より

・代理受け取りによる登録済通知書(継続)の配付の業務を追加した。

# 学内組織向けメールサーバ(MAILHOSTING)運用報告

情報教育支援係

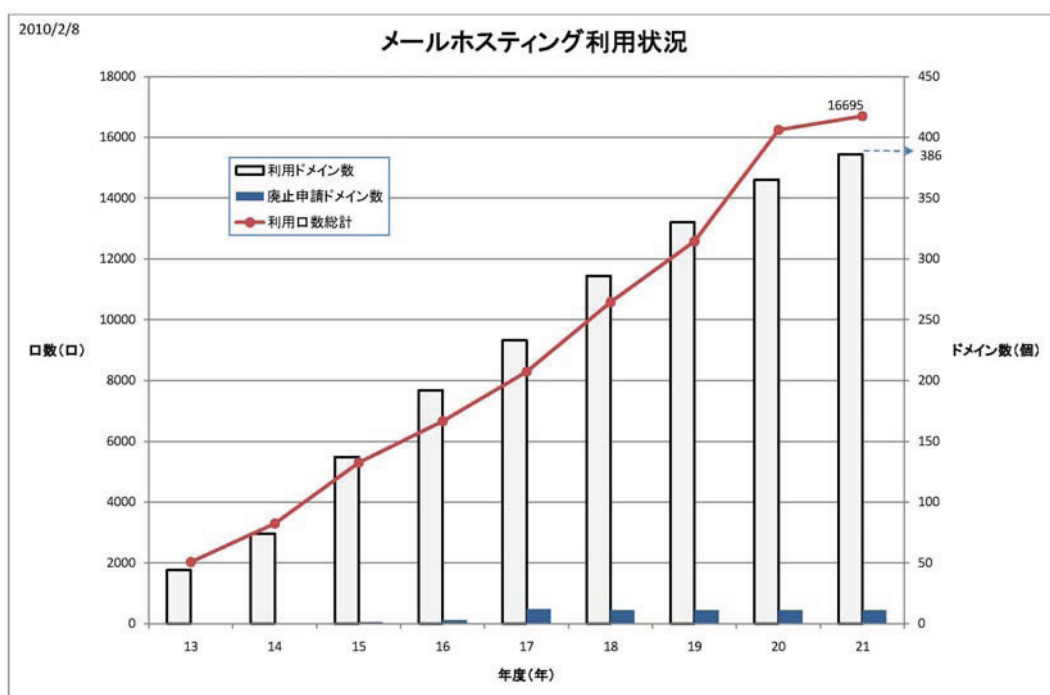
情報リテラシー教育支援係

## 1 経過

2009年6月30日 第33回情報メディア教育専門委員会において、利用負担金免除制度の適用継続を希望していた1組織に対し、制度の適用継続が認められた。

## 2 参加組織

平成21年度末現在、本サービスを利用している組織は466組織であり、それらの申込み口数の合計は17,380口である。平成20年度末に比べ、92組織、1,450口増加している。



### 3 広報

本年度に行った主な広報は、以下のとおりである。

メールホスティング M

2009年

- 4月07日 4/7(火) Active!mail 用ファイルサーバメンテナンスのお知らせ
- 4月08日 学内組織向けメールサーバの利用負担金について
- 4月11日 4/11 M サーバのファームウェアアップデートによる停止について
- 5月12日 自動返信が機能しない障害
- 5月13日 mh001 サーバ障害
- 6月16日 Mサーバ(mh002,mh004)障害
- 7月14日 20090624 mh001 障害情報
- 7月21日 QuotaOver ポリシーの設定変更実施について
- 7月22日 @gmail.com 宛メールの遅延について
- 7月28日 学内組織向けメールサーバー (MAILHOSTING)を申込み/利用する上で注意して

いただきたいこと

- 9月09日 9/9(水) Active!mail 用ファイルサーバメンテナンスのお知らせ
- 9月15日 9/15(火) Active!mail 用ファイルサーバメンテナンスのお知らせ
- 10月10日 メンテナンス作業に伴うメールサービス停止のお知らせ(mh001)
- 10月26日 本日(10/26)発生したメール送信障害について
- 11月01日 メンテナンス作業に伴うメールサービス停止のお知らせ(mh002)
- 11月15日 メンテナンス作業に伴うメールサービス停止のお知らせ(mh003)
- 11月18日 メールボックスが満杯(FULL)という警告が出る場合の対処方法
- 11月28日 メンテナンス作業に伴うメールサービス停止のお知らせ(mh004)
- 11月30日 11/30(月) メールングリスト(Mailing List/ML)サーバの障害について

2010年

- 1月26日 ECCS 各種メールサービスの電子証明書更新について
- 3月09日 MailHosting M と MailHosting D の違いは何ですか?
- 3月09日 Active! mail で件名と本文が一致しない現象について
- 3月09日 使用しているドメイン名で、メールングリスト(mailman)を使う方法
- 3月09日 UUENCODE 形式の添付ファイルが Active!mail で開けない/文字化けする
- 3月09日 Thunderbird 2 の設定例 (IMAP)
- 3月09日 メールで送信できるサイズについて制限を設定するには
- 3月09日 メール送受信容量制限について
- 3月09日 1フォルダあたりのメール数上限について
- 3月09日 メールの送信について
- 3月09日 複数の宛先にメールを転送したいのですが
- 3月09日 メール転送設定でメッセージの保存を行う時の注意事項
- 3月09日 ユーザのアカウントをどのように作れば良いのですか?
- 3月09日 パスワードの変更方法
- 3月09日 Administration Suite の迷惑メールフィルタ設定方法
- 3月09日 管理用各種ツールについて
- 3月10日 利用をやめた場合の DNS 設定削除について
- 3月10日 申請方法について
- 3月11日 mailman の廃止申請について
- 3月11日 mailman で Subject に番号をつける方法
- 3月17日 停電に伴うサーバ停止について

## メールホスティング D

### 2009 年

- 4 月 08 日 学内組織向けメールサーバの利用負担金について
- 4 月 22 日 4/22(水)12:00～12:30 メールシステムメンテナンス作業
- 6 月 16 日 M サーバ(mh002,mh004)障害
- 7 月 14 日 ECCS メールおよびメールホスティング D サーバのアクセス障害
- 7 月 28 日 学内組織向けメールサーバー (MAILHOSTING)を申込み/利用する上で注意して

### いただきたいこと

- 11 月 09 日 MailSuite のモバイルメール機能障害
- 11 月 18 日 メールボックスが満杯(FULL)という警告が出る場合の対処方法
- 11 月 30 日 11/30(月) メーリングリスト(Mailing List/ML)サーバの障害について

### 2010 年

- 1 月 26 日 ECCS 各種メールサービスの電子証明書更新について
- 3 月 09 日 MailHosting M と MailHosting D の違いは何ですか?
- 3 月 09 日 メール送受信容量制限について
- 3 月 10 日 利用をやめた場合の DNS 設定削除について
- 3 月 10 日 Thunderbird 2 の設定例 (IMAP)
- 3 月 11 日 届かないメールが、Greylisting によって切断されたものかどうか確認できないか
- 3 月 11 日 /etc/aliases 相当の機能はありますか?
- 3 月 11 日 mailman の廃止申請について
- 3 月 11 日 mailman で Subject に番号をつける方法
- 3 月 16 日 使用しているドメイン名で、メーリングリスト(mailman)を使う方法
- 3 月 16 日 申請方法について
- 3 月 17 日 停電に伴うサーバ停止について

# WEB PARK サービス運用報告

## 電子教材係

### 1 システム更新

WEB PARK システムは導入から4年が経過したため、システムの更新を行い、2010年2月22日より新システムによるサービスを開始した。

#### 1.1 システム調達

以下の日程により、システム調達を行った。調達名は「Web ホスティングシステム・ストリーミングシステム」である。なお、今回の調達は、ストリーミングサーバ2台を含めて行った。3社から応札があり、開札の結果、キャノン IT ソリューションズに決定した。

公告(文部科学省のページに掲載)	2009年9月17日(木)
提案書提出期限	2009年10月9日(金)
開札	2009年10月23日(金)
納品	2010年2月19日(金)

#### 1.2 システム更新作業

以下の日程により、システム更新作業を行った。

システム搬入作業	2009年12月24日(木)
利用者にテスト環境提供	2010年1月15日(金)
新システムによるサービス開始	2010年2月22日(月)

#### 1.3 システム構成

システム構成は、以下のとおりである。なお、調達にはこれら以外にストリーミングサーバ2台(学外向け、学内向け)を含む。

##### 1.3.1 ハードウェア

Webサーバ	2台 (アクティブ、アクティブ)
ファイルストレージ	1台
負荷分散装置	2台 (アクティブ、スタンバイ)
ネットワークスイッチ	2台
コンソール	1式
無停電電源装置	1式

##### 1.3.2 ソフトウェア

新システムに導入された主なソフトウェアのバージョンは以下のとおりである。(かっこ内は旧システムのバージョンである。)

- OS Red Hat Enterprise Linux 5.4 64bit kernel 2.6.18  
(Red Hat Enterprise Linux ES (v. 3 for x86) kernel 2.4.21)
- Apache (httpd) 2.2.3 (2.0.46)



- PHP 5.1.6 (4.3.2)
- Perl5.8.8 (5.8.0)
- Ruby 1.8.5 (1.6.8)
- Python 2.4.3 (2.2.3)

### 1.4 システム変更点

システムの主な変更点は以下のとおりである。

- Web サーバ二重化による高可用性の実現(これにより一部の機器が故障してもサービスが継続可能となった。また個々の機器の性能が向上しただけでなく、負荷分散装置を用いた2台のサーバ構成にしたことで、以前よりも遥かに多数のアクセスを処理できるようになった)
- 他組織のコンテンツ・ログの参照抑止(これまでは他組織の情報も参照可能であった)
- SSL 環境の提供(2010年4月以降にサービス開始予定。ただし park-ssl.itc.u-tokyo.ac.jp でのアクセス時のみ)
- ssh によるログインの抑止(従来は学内からは許可していたが抑止することとし、コンテンツのアップロード、ダウンロードのみ可能とすることとした)

## 2 利用組織数

2009年度に利用を開始した組織数は88、利用を中止した組織数は10であった。2010年3月末現在の総利用組織数は649である。利用組織数の変化を図1に示す。

なお、2009年6月30日に行われた第33回情報メディア教育専門委員会において、利用負担金免除制度の適用継続を希望していた5組織に対し、制度の適用継続が認められた。

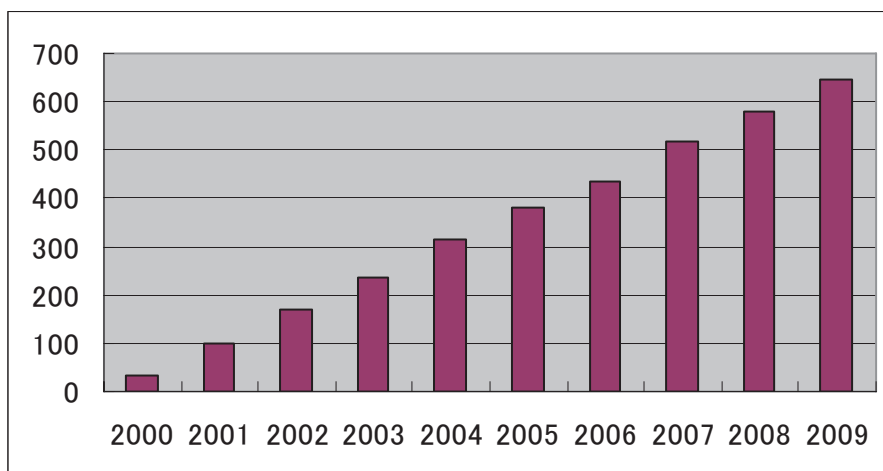


図1 利用組織数

部局別利用組織数は以下のとおりである。(2010年3月末現在)

法学系研究科／法学部	33
医学系研究科／医学部	36
工学系研究科／工学部	182
人文社会系研究科／文学部	1
理学系研究科／理学部	9
農学生命科学研究科／農学部	101
経済学研究科／経済学部	3

総合文化研究科／教養学部	78
教育学研究科／教育学部	2
薬学系研究科／薬学部	1
新領域創成科学研究科	81
情報理工学系研究科	13
情報学環・学際情報学府	1
公共政策大学院	1
東洋文化研究所	1
社会科学研究所	1
生産技術研究所	2
海洋研究所	20
先端科学技術研究センター	16
低温センター	1
アイトープ総合センター	2
環境安全研究センター	5
留学生センター	1
生物生産工学研究センター	2
アジア生物資源環境研究センター	9
医学教育国際協力研究センター	1
情報基盤センター	6
機構・委員会等	31
その他	9

### 3 利用者管理システム

WEB PARK サービスの利用手続きについて、これまでは、新規申請、継続申請、各種情報の更新等は、申請書による方法をとっていた。このため、センター側でのデータ入力の手間や、メールアドレス等の誤入力について支障をきたしていた。そこで、各種情報を利用者自身に入力、更新してもらうための利用者管理システムを作成し、新規申請については 2010 年 3 月 31 日から、継続等については同 4 月上旬から利用してもらうこととした。

関連 URL <http://park.itc.u-tokyo.ac.jp/>

# 遠隔講義支援サービス運用報告

## 電子教材係

### 1 運用報告

2009 年度においては、遠隔講義支援に関わる設備等は次の利用回数があった。なお、会議等の利用状況については、学外との接続利用のみの回数である。

	会議等(回数)	講義等(講座数)
1) 本郷 情報基盤センター413 遠隔講義室	36	10
2) 本郷 情報基盤センター405 遠隔スタジオ	1	-
3) 駒場 情報教育棟 E49 遠隔講義室	15	5
4) 貸出用テレビ会議システム	0	-
5) MCU:テレビ会議システム多地点接続制御装置	35	-
6) インターネットライブ中継	2	4

以下にその詳細を記す。

### 2 IP 遠隔会議システム

#### 2.1 会議等利用状況 (学外接続利用のみ抜粋)

##### 2.1.1 本郷 情報基盤センター413 遠隔講義室

2009年 04月 13日(月) 15:30-17:30 情報理工学系研究科 土本康生  
 2009年 04月 16日(木) 13:00-15:00 情報基盤センター 加藤 朗  
 2009年 04月 20日(月) 15:00-18:00 情報理工学系研究科 土本康生  
 2009年 04月 21日(火) 09:00-18:00 情報理工学系研究科 土本康生  
 2009年 04月 30日(木) 16:00-18:00 情報基盤センター 関谷勇司  
 2009年 05月 07日(木) 12:00-14:00 情報基盤センター 中山 雅哉  
 2009年 05月 13日(水) 09:00-18:00 工学系研究科 今野 雅  
 2009年 06月 17日(水) 12:30-18:00 情報基盤センター 中島研吾  
 2009年 06月 25日(木) 15:30-18:00 情報基盤センター 柴山悦哉  
 2009年 07月 01日(水) 09:00-18:00 情報基盤センター 片桐 孝洋  
 2009年 07月 02日(木) 15:00-18:00 薬学系研究科 佐藤 剛  
 2009年 07月 08日(水) 10:00-17:00 新領域創成科学研究科 小林 一三  
 2009年 07月 09日(木) 09:00-17:00 情報基盤センター 西澤 明生  
 2009年 07月 17日(金) 13:00-14:45 新領域創成科学研究科 小野 亮  
 2009年 07月 21日(火) 13:00-17:00 情報基盤センター 片桐 孝洋  
 2009年 07月 23日(木) 09:00-18:00 情報基盤センター 片桐 孝洋  
 2009年 07月 24日(金) 09:00-18:00 情報基盤センター 片桐 孝洋  
 2009年 08月 03日(月) 15:00-17:00 情報基盤センター 西澤 明生  
 2009年 08月 21日(金) 10:00-18:00 情報基盤センター 中川 裕志  
 2009年 08月 28日(金) 13:00-15:00 情報基盤センター 西澤 明生  
 2009年 09月 03日(木) 09:30-17:30 情報基盤センター 片桐 孝洋

2009年09月04日(金) 09:30-17:30 情報基盤センター 片桐 孝洋  
 2009年09月16日(水) 14:00-17:00 情報基盤センター 中山 雅哉  
 2009年09月28日(水) 15:30-18:00 情報基盤センター 柴山 悦哉  
 2009年10月22日(木) 18:00-20:00 情報基盤センター 丸山一貴  
 2009年10月27日(火) 10:00-12:00 情報基盤センター 中山雅哉  
 2010年01月20日(水) 15:00-18:00 情報基盤センター 関谷勇司  
 2010年01月26日(火) 15:30-18:00 情報基盤センター 柴山悦哉  
 2010年02月12日(金) 13:00-18:00 情報基盤センター 片桐孝洋  
 2010年02月16日(火) 10:00-17:00 情報基盤センター 中山雅哉  
 2010年02月17日(水) 14:00-18:00 生産技術研究所 中野美由紀  
 2010年02月18日(木) 10:00-17:00 情報基盤センター 中山雅哉  
 2010年02月22日(月) 10:00-17:00 情報基盤センター 中山雅哉  
 2010年02月23日(火) 15:00-17:00 情報基盤センター 西澤明生  
 2010年03月03日(水) 11:00-15:00 情報基盤センター 大橋公一郎  
 2010年03月17日(水) 09:00-17:00 情報基盤センター 片桐孝洋

### 2.1.2 本郷情報基盤センター405 遠隔スタジオ

2010年02月01日(月) 17:00-20:00 情報基盤センター 米山浩

### 2.1.3 駒場情報教育棟 E49 遠隔講義室

2009年04月03日(金) 13:00-16:00 総合文化研究科 玉井哲雄  
 2009年04月03日(金) 16:00-18:00 総合文化研究科 エリックシルバーマン  
 2009年06月11日(金) 13:00-14:30 大学院情報学監 佐藤 洋一  
 2009年07月08日(金) 10:20-12:30 総合文化研究科 金子邦彦  
 2009年08月05日(水) 14:00-15:30 情報基盤センター 岩藤健弘  
 2009年08月07日(金) 10:00-16:30 情報基盤センター 電子教材係  
 2009年10月27日(火) 13:00-14:00 新領域創成科学研究科 三谷 啓志  
 2009年10月31日(土) 09:00-15:30 総合文化研究科 山本 泰  
 2009年12月04日(金) 10:00-12:00 総合文化研究科 玉井哲雄  
 2009年12月10日(木) 14:30-18:00 総合文化研究科 木村 理子  
 2009年12月22日(火) 12:30-14:00 総合文化研究科 滝浪 幸次郎  
 2010年01月08日(金) 16:00-20:00 東アジア・リベラルアーツ・イニシアティブ (EALAI) 石井 弓  
 2010年01月25日(月) 15:45-16:30 総合文化研究科 古城 佳子  
 2010年01月08日(金) 16:00-20:00 東アジア・リベラルアーツ・イニシアティブ (EALAI) 崔 蘭英  
 2010年03月01日(月) 12:00-12:50 総合文化研究科 古城 佳子

## 2.2 講義等利用状況

### 2.2.1 本郷 情報基盤センター413 遠隔講義室

2009年04月20日 - 07月13日 毎週月曜日 13:00-14:45  
 新領域創成科学研究科 先端エネルギー工学輪講  
 柏キャンパスと接続

2009年04月22日 - 23日  
 新領域創成科学研究科メディカルゲノム 医科学のための情報生命学 II 集中講義  
 白金台キャンパス、柏キャンパスと接続

2009年08月05日

情報基盤センター 第69回 コンピュータ・ネットワーク利用セミナー  
駒場遠隔講義室と接続

2009年09月07日 - 08日

情報基盤センター お試シアカウント付き並列プログラミング講習会  
駒場遠隔講義室と接続

2009年09月14日 - 15日

情報基盤センター 科学技術計算のためのマルチコアプログラミング入門  
駒場遠隔講義室と接続

2009年10月05日 - 2010年02月01日 毎週月曜日 13:00-14:45

新領域創成科学研究科 先端エネルギー工学輪講  
柏キャンパスと接続

2009年10月02日 - 2010年01月29日 毎週金曜日 10:15 - 11:55

工学系研究科 電気系工学専攻 融合情報学輪講 I、II  
柏キャンパスと接続

2009年12月03日 09:00-18:00

医科学研究所遺伝子動態分野 メディカルゲノム集中講義「RNA 医科学」  
白金台キャンパス、柏キャンパスと接続

2009年12月22日 12:30-17:30

新領域メディカルゲノム 生命情報システム学  
白金台キャンパス、柏キャンパスと接続

2010年02月10日

関東 JGN2Plus(超高速ネットワーク)懇話会セミナー 眼科遠隔診療実験  
山梨大学と接続

## 2.2.2 駒場情報教育棟 E49 遠隔講義室

2009年04月08日 - 07月16日 毎週水曜日 16:00-18:00

学術俯瞰講義「いのち」のシステムを解き明かす ―急展開する生命科学  
教養学部 18号館ホール

2009年04月20日 - 07月13日 毎週月曜日 16:00-18:00

学術俯瞰講義「死すべきものとしての人間――生と死の思想」  
教養学部 18号館ホール

2009年04月09日 - 07月02日 毎週木曜日 15:00-18:00

「聞く力をつける」

2009年06月11日 - 07月23日 毎週木曜日 13:00-14:30

情報学環 佐藤洋一

2009年10月09日 - 12月11日 毎週金曜日 12:00-16:30

東アジア・リベラルアーツ・イニシアティブ(EALAI) ソウル大学との共同授業(E-lecture)

### 3 MCU サービス

#### 3.1 会議等利用状況

2009年04月09日(木) 13:00-18:00 情報基盤センター 関谷 勇司 3地点同時接続  
 2009年04月16日(木) 17:30-21:00 情報基盤センター 佐藤周行 3地点同時接続  
 2009年05月28日(木) 16:00-18:00 理学系研究科 鈴木 隆敏 3地点同時接続  
 2009年06月26日(金) 16:00-19:00 理学系研究科 鈴木 隆敏 3地点同時接続  
 2009年07月02日(木) 09:00-15:00 情報基盤センター 米山 浩 3地点同時接続  
 2009年07月30日(木) 15:30-18:00 理学系研究科 鈴木 隆敏 7地点同時接続  
 2009年08月12日(水) 09:30-18:00 情報基盤センター 関谷 勇司 3地点同時接続  
 2009年08月27日(木) 16:00-18:30 理学系研究科 鈴木 隆敏 4地点同時接続  
 2009年09月04日(金) 11:00-17:00 情報基盤センター 米山 浩 3地点同時接続  
 2009年09月24日(木) 16:00-18:30 理学系研究科 鈴木 隆敏 4地点同時接続  
 2009年10月05日(月) 10:00-12:00 工学系研究科 原 辰徳 3地点同時接続  
 2009年10月21日(水) 18:00-26:00 本部国際系 清水 響子 5地点同時接続  
 2009年10月27日(火) 16:45-19:30 理学系研究科 鈴木 隆敏 4地点同時接続  
 2009年11月05日(木) 11:00-12:00 工学系研究科 原 辰徳 3地点同時接続  
 2009年11月06日(金) 13:00-18:00 理学系研究科 藤代 知子 6地点同時接続  
 2009年11月24日(火) 17:15-20:00 理学系研究科 鈴木 隆敏 4地点同時接続  
 2009年11月19日(木) 09:00-10:00 新領域創成科学研究科 小貫 元治 8地点同時接続  
 2009年12月02日(木) 18:00-20:00 工学系研究科 新井 民夫 3地点同時接続  
 2009年12月04日(金) 09:00-10:00 新領域創成科学研究科 小貫 元治 2地点同時接続  
 2009年12月10日(木) 14:00-17:00 情報基盤センター 関谷 勇司 7地点同時接続  
 2009年12月18日(金) 17:00-翌日 19:00 新領域創成科学研究科 小貫 元治 8地点同時接続  
 2009年12月22日(火) 18:45-21:00 理学系研究科 鈴木 隆敏 5地点同時接続  
 2010年01月08日(金) 17:00-翌日 19:00 新領域創成科学研究科 小貫 元治 8地点同時接続  
 2010年01月20日(水) 15:00-18:00 情報基盤センター 関谷 勇司 6地点同時接続  
 2010年01月22日(金) 17:00-翌日 19:00 新領域創成科学研究科 小貫 元治 8地点同時接続  
 2010年01月26日(火) 16:45-20:00 理学系研究科 佐藤 将春 5地点同時接続  
 2010年02月01日(月) 13:30-17:30 宇宙線研究所 黒田 和明 6地点同時接続  
 2010年02月16日(火) 13:00-18:00 情報基盤センター 西澤明生 3拠点同時接続  
 2010年02月22日(月) 12:00-18:00 情報基盤センター 関谷 勇司 5拠点同時接続  
 2010年02月23日(火) 16:45-20:00 理学系研究科 佐藤 将春 5拠点同時接続  
 2010年03月01日(月) 13:15-17:30 宇宙線研究所 黒田 和明 8拠点同時接続  
 2010年03月15日(月) 18:00-20:00 工学系研究科 新井 民夫 3拠点同時接続  
 2010年03月23日(火) 18:00-20:00 工学系研究科 新井 民夫 3拠点同時接続  
 2010年03月30日(火) 16:45-20:00 理学系研究科 佐藤 将春 5拠点同時接続  
 2010年03月31日(水) 18:00-20:00 工学系研究科 新井 民夫 3拠点同時接続

#### 3.2 講義等利用状況

2009年04月20日 - 07月17日 毎週月曜日 16:30-18:20  
 新領域創成科学研究科 複雑理工学専攻 3地点同時接続

2009年06月30日(火) 15:00-17:30  
 新領域創成科学研究科 高橋 成雄 3地点同時接続



2009年10月05日 - 2010年01月25日 毎週月曜  
新領域創成科学研究科 江尻 晶 3地点同時接続

#### 4 インターネットライブ中継

2009年04月08日 - 07月16日 毎週水曜日 16:20-17:50  
学術俯瞰講義「「いのち」のシステムを解き明かすー急展開する生命科学ー」  
(配信範囲: 学内のみ)

2009年04月20日 - 07月13日 毎週月曜日 16:20-17:50  
学術俯瞰講義「死すべきものとしての人間——生と死の思想」  
(配信範囲: 学内のみ)

2009年07月02日(木) 15:00-18:00  
東京大学グローバル COE「学融合に基づく医療システムイノベーション」  
留学生のためのオリエンテーション/セミナー  
(配信範囲: 制限なし)

2009年10月05日 - 2010年01月25日 毎週月曜日 16:20-17:50  
学術俯瞰講義「歴史とは何か」  
(配信範囲: 学内のみ)

2009年10月08日 - 2010年01月21日 毎週木曜日 16:20-17:50  
学術俯瞰講義「数学を創るー数学者達の挑戦ー」  
(配信範囲: 学内のみ)

2010年03月24日 9:00-12:15、03月25日 9:00-12:15  
平成21年度 東京大学 学位記授与式・卒業式  
(配信範囲: 制限なし)

関連 URL <http://elearn.itc.u-tokyo.ac.jp/>

# CFIVE 運用報告

## 電子教材係

### 1 経過

CFIVE はプログラムソースを公開している学習管理システムであり、2004 年 4 月より東京大学でのサービスを開始した。サービス開始以来 CFIVE を利用する講義数は順調に増加しており、2005 年度は 35 講義、2006 年度は 63 講義、2007 年度は 104 講義、2008 年度は 126 講義、2009 年度は 143 講義で利用された。以下に改良された機能、利用された講義名、広報等を示す。

### 2 新たに追加・改善された機能

2009 年度に CFIVE に新たに追加・改善された機能は、以下の通りである。

#### 追加機能

- ・テスト実施時の制限時間設定
- ・ユーザ ID 変更時のデータ移行機能
- ・課題ダウンロード履歴参照機能
- ・受講期間終了後のコース削除機能
- ・システム障害調査プログラム
- ・テスト実施時の解答バックアップ機能

#### 機能改善

- ・テスト機能の処理フロー改善
- ・コース再登録時のユーザ登録不具合の改善
- ・課題登録画面の修正、課題提出ボタン作成
- ・掲示版の RSS リーダの改善
- ・WebDAV の改善
- ・テストプレビュー時のエラー対応
- ・提出期限前の課題に評価出来る様に仕様変更
- ・出席一覧表示をユーザ ID 順にソートする様に仕様変更

### 3 夏学期に CFIVE を利用した講義

夏学期(2009 年 4 月－2009 年 9 月)に CFIVE を利用した講義は以下の通りである。

1	教養学部(前期課程)	ALESS ライティングセンター(英語二列 PA 履修者)	板津 木綿子他
2	教養学部(前期課程)	はいぱーワークブック	はいぱーワークブックプロジェクト
3	教養学部(前期課程)	プログラム構成論(月 5)	山口和紀
4	教養学部(前期課程)	モンゴル語初級(火 5,木 5)	木村理子
5	教養学部(前期課程)	宇宙科学実習 II(火 4,5)	鈴木建
6	教養学部(前期課程)	宇宙科学実習 II(月 4,5)	鈴木建
7	教養学部(前期課程)	英語2列(C)(金 3)	アルヴィなほ子
8	教養学部(前期課程)	英語2列(C)(木 2)	アルヴィなほ子
9	教養学部(前期課程)	英語2列(C)(木 4)	アルヴィなほ子
10	教養学部(前期課程)	英語二列(R)(火 3)	奥 聡一郎

11	教養学部(前期課程)	英語二列(R)(火 4)	奥 聡一郎
12	教養学部(前期課程)	英語二列C(C)(火 3)	丹治愛
13	教養学部(前期課程)	英語二列C(C)(火 3)	広瀬友紀
14	教養学部(前期課程)	英語二列C(C)(火 4)	丹治愛
15	教養学部(前期課程)	英語二列C(C)(月 3)	広瀬友紀
16	教養学部(前期課程)	英語二列C(C)(月 4)	広瀬友紀
17	教養学部(前期課程)	英語二列C(C)(木 2)	広瀬友紀
18	教養学部(前期課程)	英語二列 C(C)(木 3)	河合祥一郎
19	教養学部(前期課程)	英語二列C(C)(木 4)	広瀬友紀
20	教養学部(前期課程)	英語二列 C(C)(木 4)	河合祥一郎
21	教養学部(前期課程)	英語二列 R(R)(火 4)	河合祥一郎
22	教養学部(前期課程)	英語二列 R(R)(火 5)	河合祥一郎
23	教養学部(前期課程)	英語二列R(R)(金 1)	荒木純子
24	教養学部(前期課程)	英語二列R(R)(金 2)	荒木純子
25	教養学部(前期課程)	英語二列 R(木 2)	小林薫
26	教養学部(前期課程)	英語二列 R(木 3)	小林薫
27	教養学部(前期課程)	学術俯瞰講義「いのち」のシステムを解き明かす(水 5)	福田裕穂、石浦章一
28	教養学部(前期課程)	学術俯瞰講義「死すべきものとしての人間」(月 5)	島蘭進、村松真理子
29	教養学部(前期課程)	環境エネルギーゲーミング(木 4)	飯田 誠
30	教養学部(前期課程)	環境エネルギー科学基礎概論(木 5)	飯田 誠
31	教養学部(前期課程)	基礎現代化学(火 1)	永田敬
32	教養学部(前期課程)	基礎現代化学(水 1)	永田敬
33	教養学部(前期課程)	近現代史 I(木 5)	伊熊幹雄
34	教養学部(前期課程)	現代国際社会論(金 5)	旭英昭
35	教養学部(前期課程)	国際関係論 I(火 5)	大庭三枝
36	教養学部(前期課程)	社会 I(金 4)	和田毅
37	教養学部(前期課程)	初修外国語スペイン語(公開講座.全員登録)	スペイン語部会全教員
38	教養学部(前期課程)	小型風力発電を通じて環境エネルギーを考えるー2(木 3)	飯田 誠
39	教養学部(前期課程)	情報(火 2)	植田一博
40	教養学部(前期課程)	情報(火 2)	中村政隆
41	教養学部(前期課程)	情報(火 3)	増原英彦
42	教養学部(前期課程)	情報(火 3)	中谷多哉子
43	教養学部(前期課程)	情報(火 4)	玉井哲雄
44	教養学部(前期課程)	情報(火 4)	佐藤洋一
45	教養学部(前期課程)	情報(共通)	丹羽清ほか
46	教養学部(前期課程)	情報(金 1)	山口和紀
47	教養学部(前期課程)	情報(金 2)	蔡 東生
48	教養学部(前期課程)	情報(金 2)	広田光一
49	教養学部(前期課程)	情報(金 3)	蔡 東生
50	教養学部(前期課程)	情報(金 3)	藤垣裕子
51	教養学部(前期課程)	情報(金 4)	藤垣裕子
52	教養学部(前期課程)	情報(金 4)	田浦健次朗
53	教養学部(前期課程)	情報(月 1)	山口 泰
54	教養学部(前期課程)	情報(月 3)	河内谷幸子
55	教養学部(前期課程)	情報(月 3)	柴山悦哉
56	教養学部(前期課程)	情報(月 4)	田村肇
57	教養学部(前期課程)	情報(月 4)	玉井哲雄

58	教養学部(前期課程)	情報(月 5)	竹内郁雄
59	教養学部(前期課程)	情報(水 2)	原田至郎
60	教養学部(前期課程)	情報(水 2)	開 一夫
61	教養学部(前期課程)	情報(水 3)	田中哲朗
62	教養学部(前期課程)	情報(水 3)	日暮英治
63	教養学部(前期課程)	情報(水 4)	原田至郎
64	教養学部(前期課程)	情報(水 4)	美馬秀樹
65	教養学部(前期課程)	情報(木 1)	植田一博
66	教養学部(前期課程)	情報(木 2)	植田一博
67	教養学部(前期課程)	情報(木 2)	佐藤洋一
68	教養学部(前期課程)	情報(木 3)	中村政隆
69	教養学部(前期課程)	情報(木 3)	辰己 丈夫
70	教養学部(前期課程)	情報科学概論 I(火 5)	増原英彦
71	教養学部(前期課程)	図形科学 II(火 4)	奈尾信英
72	教養学部(前期課程)	図形科学 II(火 3)	金井崇
73	教養学部(前期課程)	図形科学 II(金 3)	堤江美子, 横山ゆりか
74	教養学部(前期課程)	図形科学 II(金 4)	横山ゆりか, 堤江美子
75	教養学部(前期課程)	図形科学 II(月 3)	田中一郎, 道川隆士
76	教養学部(前期課程)	図形科学 II(月 4)	道川隆士, 田中一郎
77	教養学部(前期課程)	図形科学 II(水 3)	金井崇柏原賢二
78	教養学部(前期課程)	図形科学 II(水 4)	柏原賢二金井崇
79	教養学部(前期課程)	図形科学 II(木 2)	横山ゆりか
80	教養学部(前期課程)	図形科学 II(木 3)	深野暁雄・横山ゆりか
81	教養学部(前期課程)	図形科学演習 II(月 2)	山口泰
82	教養学部(前期課程)	生命科学(木 1)	佐藤直樹
83	教養学部(前期課程)	生命科学 I(月 3)	佐藤直樹
84	教養学部(前期課程)	西洋古典学 I	小林薫
85	教養学部(前期課程)	全学自由研究ゼミナール 情報システム利用入門(金 5)	玉井哲雄金子知適
86	教養学部(前期課程)	地球環境論(木 5)	丸山康司 飯田誠
87	教養学部(前期課程)	法と社会(月 2)	岩田太
88	教養学部(後期課程)	国際協力	旭英昭
89	教養学部(後期課程)	数理情報学 I 演習(木 3)	齊藤宣一
90	教養学部(後期課程)	特殊講義 IV	吉原真里
91	総合文化研究科	アメリカ宗教文化論	ロビンズ ロジャー
92	総合文化研究科	ラテンアメリカ社会論 I	和田毅
93	総合文化研究科	平和プロセスと国際協力2	旭英昭
94	総合文化研究科	北米・中南米地域文化演習 I	ロビンズ ロジャー
95	工学部	計算機工学概論	矢入健久
96	文学部	哲学特殊講義	ロビンズ ロジャー
97	その他	計算数学 I(金 3)	一井信吾
98	その他	計算数理演習(金 3)	齊藤宣一
99	その他	国際文化演習	ロビンズ ロジャー

## 4 冬学期に CFIVE を利用した講義

冬学期(2009年10月ー2010年3月)に CFIVE を利用した講義は以下の通りである。

1	教養学部(前期課程)	ALESS ライティングセンター(英語二列 PA 履修者)	板津 木綿子他
2	教養学部(前期課程)	はいばーワークブック	はいばーワークブックプロジェクト
3	教養学部(前期課程)	初修外国語スペイン語(公開講座,全員登録)	スペイン語部会全教員
4	教養学部(前期課程)	近現代史 II(月 2)	森 まり子
5	教養学部(前期課程)	情報科学概論 II(月 2)	玉井哲雄
6	教養学部(前期課程)	法 II(月 3)	西村 弓
7	教養学部(前期課程)	情報科学(月 4)	玉井哲雄
8	教養学部(前期課程)	経済政策(月 5)	河村哲二
9	教養学部(前期課程)	英語二列 P(火 2)	河合祥一郎
10	教養学部(前期課程)	情報科学(火 2)	山口 泰
11	教養学部(前期課程)	英語二列(P)(火 3)	奥聡一郎
12	教養学部(前期課程)	英語二列(P)(火 4)	奥聡一郎
13	教養学部(前期課程)	情報科学(火 4)	山口和紀
14	教養学部(前期課程)	情報科学(水 4)	田中哲朗
15	教養学部(前期課程)	英語二列 R(木 2)	小林薫
16	教養学部(前期課程)	英語二列 P(P)(木 3)	アルヴィなほ子
17	教養学部(前期課程)	英語二列 R(木 3)	小林薫
18	教養学部(前期課程)	地域文化論 II(木 4)	アルヴィなほ子
19	教養学部(前期課程)	近現代史 II(木 5)	伊熊幹雄
20	教養学部(前期課程)	モンゴル語初級(木 5)	木村理子
21	教養学部(前期課程)	学術俯瞰講義「数学を創る」(木 5)	岡本和夫
22	教養学部(前期課程)	英語二列 P(金 1)	河合祥一郎
23	教養学部(前期課程)	英語二列 P(P)(金 1)	荒木純子
24	教養学部(前期課程)	英語二列 P(P)(金 2)	荒木純子
25	教養学部(前期課程)	情報科学(金 3)	上田哲郎
26	教養学部(前期課程)	情報科学(金 4)	増原英彦
27	教養学部(前期課程)	学術俯瞰講義「歴史とは何か」(月 5)	吉田元夫, 森山工
28	教養学部(前期課程)	宇宙素粒子物理学のフロンティア(水 4)	早戸良成
29	教養学部(前期課程)	コンピュータ及び演習(水 1)	阿久津好明
30	教養学部(前期課程)	神話学(木 1)	小林薫
31	教養学部(後期課程)	米州地域文化論 I/米州関係論	ドボルザーク、グレッグ
32	教養学部(後期課程)	中東近現代史 II(金 3)	森 まり子
33	教養学部(後期課程)	情報システム科学 V(水 2)	山口 泰
34	教養学部(後期課程)	情報システム科学実習 II(水 4)	山口和紀,金子知適
35	総合文化研究科	データ分析(水 1)	和田毅
36	システム創成学科	システム制御工学(火 1)	大和 裕幸
37	農学生命科学研究科	国際環境管理学特論(月 4)	高橋太郎
38	農学生命科学研究科	農学国際特論 II(金 3・4)	高橋太郎
39	工学部	ソフトウェア第一/機械ソフトウェア演習	森武俊, 広田光一, 杉田直彦
40	工学部	プログラミング基礎 D(金 3,4,5)	白山晋
41	工学部	プログラミング基礎演習(金 5)	川原圭博, 峯松信明, 伊庭齊志
42	農学部	情報工学	中村典裕
43	情報基盤センター	コンピュータネットワーク研修	情報基盤センター
44	その他	北米・中南米地域文化演習 I/比較現代政治 II(水 2)	和田毅

## 5 組織毎の利用状況

組織毎にまとめた 2009 年度の利用状況は、以下の通りである。

開講組織	講義数
教養学部(前期課程)	117
教養学部(後期課程)	7
総合文化研究科	5
工学部	4
文学部	1
農学部	4
情報基盤センター	1
その他	4
合計	143

## 6 広報

2009 年度に行った主な広報は、以下の通りである。

お知らせ

1	WebDAV 機能の不具合について	2009/04/14
2	「履修中のコース」の表示(学生向け)	2009/04/21
3	CFIVE ログインについて	2009/04/21
4	ECCS アカウント変更があった学生の対応	2009/05/15
5	CFIVE メンテナンスのお知らせ	2009/05/20
6	CFIVE メンテナンスのお知らせ	2009/08/28
7	CFIVE に関する情報について	2009/09/09
8	CFIVE Support ではレポート提出の受付は行っておりません。(学生向け)	2009/09/09
9	課題・教材が参照できない(学生向け)	2009/09/09
10	課題作成の際の注意について(教職員向け)	2009/09/09
12	推奨ブラウザについて	2009/09/16
13	サポートへの問い合わせ(学生向け)	2009/09/16
14	CFIVE 利用規則, 運用ポリシーについて	2009/09/16
15	CFIVE 利用について	2009/10/19
16	平成 16 年度、17 年度データの削除	2010/03/19
17	CFIVE 新規利用について(教職員向け)	2010/03/30
18	CFIVE 新規利用について(新入生向け)	2010/03/31





## 図書館電子化部門



データベースリーフレット



ネットでアカデミック 学術情報へのアクセスガイド



東京大学 OPAC



# 図書館電子化部門

## 概要

専門員 本多 玄

図書館電子化部門は本学の学生・教職員の学習・教育・研究活動にとって必要不可欠な学術情報をデジタルコンテンツとしてネットワークを介して提供している。

また、本学において日々生産・蓄積される研究成果を電子化し社会に向けて発信するための事業、さらにそれら電子的学術情報を利用者が十分に利活用できるよう講習会等による情報リテラシー支援事業を附属図書館とも連携しながら行なっている。

図書館電子化部門にはこれらの目的に沿って「図書館システム係」「デジタル・ライブラリ係」「学術情報リテラシー係」の3係が置かれている。

図書館電子化部門が提供しているサービス・事業の主なものとして以下がある。

- 東京大学 OPAC
- GACoS
- 東京大学機関リポジトリ
- 学位論文データベース
- 文献管理ツール「RefWorks」
- 東京大学で利用できる電子ジャーナル検索
- 貴重書の電子化・公開
- 「ネットでアカデミック」等の発行
- 各種講習会、情報探索ガイダンス等の実施

2009年度は上記に関連して以下の事業を実施した。

- 次期図書館システムの検討
- 東京大学機関リポジトリのコンテンツ増強
- 学内貴重資料の電子化支援
- 学位論文データベースへのコンテンツ追加
- 情報探索・データベース講習会の実施
- 「ネットでアカデミック」や「図書館利用ガイド」等印刷出版物の内容改訂
- 「GACoS」の内容メンテナンスやメールマガジン「Litetopi」の発行
- 図書系職員のアプリケーション開発能力を高めるための支援

# デジタルコンテンツサービス

## デジタル・ライブラリ係

### 1 運用報告

#### 1.1 東京大学学術機関リポジトリの構築

東京大学学術機関リポジトリ～UT Repository～とは、東京大学で生産されたさまざまな研究成果を電子的な形態で集中的に蓄積・保存し、学内外に公開することを目的としたインターネット上の発信拠点である。平成 16 年度から附属図書館と連携し学術機関リポジトリの構築を行い、平成 18 年 4 月 1 日から「東京大学学術機関リポジトリ(UT Repository)」としてサービスを開始した。

(<http://repository.dl.itc.u-tokyo.ac.jp/>)

平成 21 年度は、昨年度に引き続き、国立情報学研究所の最先端学術情報基盤(CSI)構築事業による外部資金を調達・活用するなどして、附属図書館と連携協力して次の業務を行った。



##### 1.1.1 コンテンツ作成と著作権許諾処理

紀要、学位論文、学術雑誌論文を中心としてコンテンツの収集と作成を行い、3 月末には約 20,000 件のデータを公開しており、前年度以前と比較しても順調に増加している。

機関リポジトリでは本文そのものを PDF 等で公開するため、単に電子化作業を行うだけではなく、著作権の許諾処理を必須として作業を行っている。著者本人や共著者の許諾はもちろんであるが、学術雑誌掲載論文では出版者の許諾も必要であるため、機関リポジトリへの掲載許諾情報を調査し、必要に応じて学協会等へ個々に許諾の確認を取る作業を行った。また、紀要等のバックナンバーの電子化にあたり、著者への許諾を電子メールと郵送により実施した。

#### 1.2 電子ジャーナルのゲートウェイサービス

ゲートウェイとして、SSL-VPN Gateway サービス(試行版)を実施、3 月末現在で約 50 件のデータベースや電子ジャーナルのパッケージに、外部から当該ゲートウェイを介してアクセス可能である。

電子ジャーナルの提供を円滑に行うための電子ジャーナルリンク集データベース(約 8, 400 件)のメンテナンス及び維持・管理と、本学が契約するデータベースを経由して利用できる電子ジャーナルとフリーの電子ジャーナルについての管理ソフトを導入した E-Journal Portal



(約 45,000 件)の学内へのサービスを行っている。

なお、上記 SSL-VPN Gateway サービスを提供しているサーバは6月に更新を行った。

### 1.3 東京大学学位論文要旨データベースの構築

学位論文データベースは従来冊子体で刊行されていた「東京大学博士学位論文の内容の要旨と審査結果の要旨」を電子化して公開しているものである。

平成 21 年度は平成 19 年度論文要旨約 1,300 件の電子化を行った。昨年度同様に、提出電子ファイルを全面的に利用し、データベース作成の効率化と経費の削減を図った。

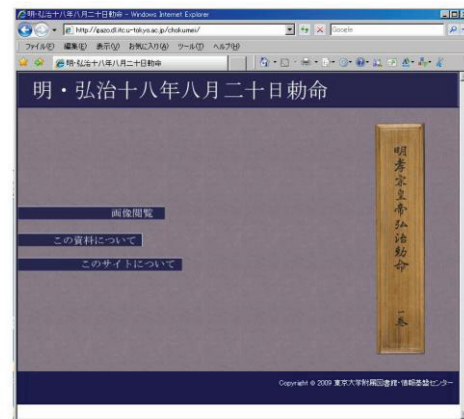
(<http://gakui.dl.itc.u-tokyo.ac.jp/>)

### 1.4 貴重資料のデータベース化支援

#### 1.4.1 『明・弘治十八年八月二十日勅命』

この史料は中国明時代における辞令書の一種で、弘治 18 年(1505)8 月 20 日に南京広西道監察御史王欽の父母に官職・称号を授ける旨が記されているものである。明代中期の文書史料の実例は多くないため、明代の文書様式・官僚制度等を理解する上で貴重なものとなっている。この現物は、総合図書館鷗外文庫(森鷗外の旧蔵書)の一書として所蔵されている。なお、この電子版は平成 20～23 年度科学研究費基盤研究(A)「東国地域及び東アジア諸国における前近代文書等の形態・料紙に関する基礎的研究」(課題番号:20242016)による研究成果の一部として作成された。また、資料解説は経済学研究科助手の小島浩之先生にご執筆いただいている。

(<http://gazo.dl.itc.u-tokyo.ac.jp/chokumei/>)



#### 1.4.2 観世アーカイブ

これは、本学所蔵の資料ではないが、観世宗家・観世文庫が所蔵する、室町時代以降約 6000 点の電子アーカイブです。科学研究費基盤研究(A)「観世文庫所蔵能楽関係資料のデジタル画像化と解題目録作成に向けた総合的研究」(課題番号:18202006)の調査・研究成果として、構築・公開したものである。

世阿弥自筆本等もあり、能楽の資料として価値が高いだけでなく、当時の日本語(仮名)に関しても学術的価値の高いものを見ることができる。

(<http://gazo.dl.itc.u-tokyo.ac.jp:8080/kanzegazo/index.html>)





# 学術情報リテラシー支援

## 学術情報リテラシー係

### 1 運用報告

#### 1.1 情報探索ポータルサイト GACoS(Gateway to Academic Contents System)の運用

電子的な学術情報にアクセスするためのポータルサイトとして、GACoS(日本語・英語・中国語・韓国語版)を継続的に構築、運用した。

(<http://www.dl.itc.u-tokyo.ac.jp/gacos/>)

#### 1.2 リテラシー支援資料の作成・改訂

##### 1.2.1 「ネットでアカデミック」

東京大学における文献収集のガイドブック「ネットでアカデミック 2010」(日本語・英語・中国語・韓国語版)を作成した。Web 版文献管理ツール RefWorks 等の新規サービスの解説を加えるなど、内容の改訂を行なった。また、PDF 版を GACoS 上で公開した。

(<http://www.dl.itc.u-tokyo.ac.jp/gacos/net.html>)



##### 1.2.2 リーフレット「文献探しのヒント」

東京大学で利用できるデータベース等のサービスについて、各々の概要や機能等を紹介したリーフレット 5 種(日本語・英語版)を作成した。平成 21 年度は、「目的に合ったデータベースを探すには？」といったタイトルを、前年度より大きく表示するなどの改良をした。PDF 版を GACoS 上で公開した。

(<http://www.dl.itc.u-tokyo.ac.jp/gacos/literacy.html>)

リーフレット「文献探しのヒント」5 種類は次のとおり。

- ・ 目的に合ったデータベースを探すには？ (GACoS)
- ・ 複数のデータベースを横断検索するには？  
(東京大学学術論文横断検索・UT Article Search)
- ・ 図書館の本や雑誌を検索するには？  
(東京大学 OPAC)
- ・ 自宅からデータベースや電子ジャーナルを使うには？ (SSL-VPN Gateway サービス)
- ・ 雑誌論文を入手するには？  
(電子ジャーナル、東京大学学術論文リンク・UT Article Link ほか)



### 1.2.3 「図書館利用ガイド」

東京大学の図書館を初めて利用する学生等を対象とした、「図書館利用ガイド 2010」(附属図書館学術情報リテラシー教育部会編集、日本語・英語版)の作成に協力し、「第 2 部 インターネットリソースへのアクセス」を担当した。

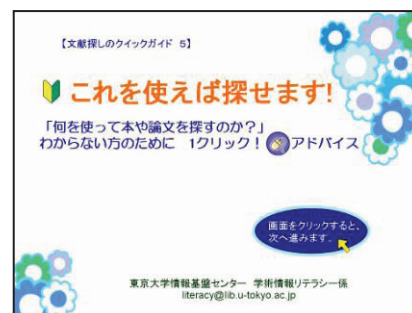
### 1.2.4 クイックガイド

前年度試行公開したアニメーション教材のクイックガイドを、新たに3種作成した。また、全6種類の英語版も作成し、GACoS 上で本格的に公開した。

クイックガイド6種類は以下のとおり。

- ・目的に合ったデータベースを探すには？
- ・自宅からデータベースや電子ジャーナルを使うには？
- ・「UT Article Link」はどう使うの？
- ・電子ジャーナルで論文を入手するには？
- ・これを使えば探せます！
- ・参考文献リストの見かた

(<http://www.dl.itc.u-tokyo.ac.jp/gacos/literacy.html>)



## 1.3 Litetopi メールマガジンの発行

東京大学構成員を対象に、新規サービスや講習会の情報を発信する「Litetopi (リテトピ) メールマガジン」を、年 31 回発行した。購読登録者は約 1,500 人に達した。バックナンバーは GACoS 上で公開した。

## 1.4 Web 版文献管理ツール RefWorks の提供

インターネット上で使える文献管理ツール RefWorks を導入し、新たにサービスを開始した。

## 2 講習会・説明会開催報告

### 2.1 情報探索ガイダンスの実施概要

データベース等の電子的学術資料を学習・教育・研究に効率的に利用してもらうことを目的とし、情報探索ガイダンス各コースを開催した。平成 21 年度は計 157 回開催し、1,881 人の参加があった。このうち、いくつかのコースでは、電子教材係の協力の下、テレビ会議システムを用いて、柏図書館に中継を行った。(以下の表で会場欄に※印のあるもの)。

### 2.2 平成 21 年度情報探索ガイダンス開催実績

#### 2.2.1 テーマ別ガイダンス

1 つのテーマについて検索実習を交えながら解説する、テーマ別ガイダンスを 7 コース実施した。平成 21 年度は、Web 版文献管理ツール RefWorks、国内新聞記事、海外新聞記事の 3 コースを新たに実施した。特に RefWorks 講習会は好評であった。

コース名	月日	回数	人数	内容
日本の論文を探すには?	5/18,5/26,7/10, 10/9,11/4,12/4, 1/21,2/5	8	48	CiNii の使い方を中心に、日本の雑誌論文を検索できるデータベースを紹介
電子ジャーナルを利用するには?(11 月)	5/19,5/29,7/15, 10/23,11/19,	9	52	電子ジャーナルの探し方や、代表的な出版社の電子ジャーナルサイトを例にと

より「電子ジャーナルで論文を入手」に改称)	12/8,1/20,2/18,3/3			って紹介
自宅から検索するには?	9/14,10/7,11/10,12/9,1/29,2/10,3/4	7	57	学外からも利用できるデータベース、電子ジャーナルの利用手続きやアクセス方法
文献検索早わかり	5/8,5/21,7/24,9/16,9/29,10/21,11/25,12/10,1/26,2/12	10	68	図書や電子ジャーナル、雑誌論文、新聞記事など、各種の文献検索方法をコンパクトに紹介
文献リストをサクッと作成～RefWorksを使うには	5/22,7/3,7/28,9/15,9/30,10/13,10/29,11/13,11/24,12/16,1/15,2/24,3/12	13	120	RefWorksを使った文献検索結果の保存・整理方法、参考文献リストの作成方法
国内の新聞記事を探すには?	2/4	1	9	朝日新聞、毎日新聞、読売新聞の記事の検索方法
海外の新聞記事を探すには?	3/5	1	5	ProQuest Newspapers、LexisNexis Academicなどの海外新聞記事の検索方法
合計		49回	359名	

### 2.2.2 データベースユーザトレーニング

昨年度開催の Web of Science + EndNote Web コースのほか、新たに PubMed + RefWorks のコース等を開催した。いずれも概ね好評であった。

コース名	月日	回数	人数	内容
Web で文献の「検索」から「整理」まで～Web of Science + EndNote Web～	10/28,11/27,2/2	3	29	Web of Science の検索方法と、EndNote Web を使った文献検索結果の保存・整理方法、参考文献リストの作成方法
Web で文献の「検索」から「整理」まで～PubMed + RefWorks～	2/9	1	15	PubMed の検索方法と、RefWorks を使った文献検索結果の保存・整理方法、参考文献リストの作成方法
Web of Science を使うには?	3/9	1	5	Web of Science の検索方法、引用情報の見方など
EndNote Web を使うには?	3/10	1	13	EndNote Web を使った文献検索結果の保存・整理方法、参考文献リストの作成方法
合計		6回	62名	

### 2.2.3 分野別データベースユーザトレーニング

データベースの提供元から講師を招き、各分野の主要なデータベースの講習会を実習形式で実施した。前年度同様、一部のコースについては駒場図書館、柏図書館でも実施した。

#### 総合分野

コース名	月日	時間	人数	会場
JapanKnowledge+	6/2(火)	15:00-16:00	8	総合図書館

Littel Navigator	6/2(火)	16:15-17:15	10	総合図書館
ヨミダス歴史館	6/3(水)	15:00-16:00	10	総合図書館
聞蔵IIビジュアル	6/8(月)	15:00-16:00	8	総合図書館
毎日 News パック	6/10(水)	13:45-14:45	6	総合図書館
JapanKnowledge+	6/11(木)	13:30-14:30	5	駒場図書館
ヨミダス歴史館	6/11(木)	14:45-15:45	7	駒場図書館
LexisNexis Academic	6/12(金)	15:00-16:30	14	総合図書館
EndNote Web	6/26(金)	16:30-17:30	19(4)	総合図書館※
EndNote Web	6/30(火)	16:30-17:30	19	柏図書館
毎日 News パック	7/1(水)	13:30-14:30	6	駒場図書館
聞蔵IIビジュアル	7/1(水)	14:45-15:45	5	駒場図書館
RefWorks	3/31(水)	14:00-15:00	21	医学図書館

## 人文社会科学系分野

コース名	月日	時間	人数	会場
Eighteenth Century Collections Online+GVRL	6/1(月)	13:30-14:30	4	駒場図書館
Early English Books Online	6/1(月)	14:45-15:45	7	駒場図書館
PIO(Periodicals Index Online)/PAO(Periodicals Archive Online)	6/3(水)	16:15-17:15	11	総合図書館
eol	6/4(木)	15:00-16:00	15	総合図書館
InfoTracCustom	6/5(金)	15:00-16:00	12	総合図書館
LRC+GVRL	6/5(金)	16:15-17:15	12	総合図書館
ASSIA/LLBA	6/10(水)	15:00-16:30	10	総合図書館
Lexis.com	6/12(金)	13:15-14:45	13	総合図書館
Business Source Elite /SocINDEX	6/18(木)	15:00-16:30	13	総合図書館
Web of Science(人文科学系)	6/19(金)	13:00-14:30	6	駒場図書館
JSTOR	6/22(月)	15:00-16:00	13	総合図書館
PsycINFO/PsycARTICLES	6/23(火)	15:00-16:00	5	総合図書館
CNKI	6/24(水)	15:00-16:30	15	総合図書館
Web of Science(人文科学系)	6/26(金)	10:30-12:00	9	総合図書館
Westlaw International/ Westlaw Japan	9/29(火)	15:00-16:30	9	総合図書館

## 自然科学系分野

コース名	月日	時間	人数	会場
Engineering Village	6/9(火)	15:00-16:30	17(4)	総合図書館※
MathSciNet	6/15(月)	16:00-17:00	16	数理科学研究科
MathSciNet	6/16(火)	15:00-16:00	16	数理科学研究科
SciFinder Scholar	6/17(水)	15:00-16:30	17	総合図書館
Web of Science(自然科学系)	6/19(金)	14:45-16:15	8	駒場図書館
SciFinder Scholar	6/25(木)	15:00-16:30	9	柏図書館
Web of Science(自然科学系)	6/26(金)	13:30-15:00	22(9)	総合図書館※

BIOSIS Previews	6/26(金)	15:15-16:15	9(3)	総合図書館※
Web of Science(自然科学系)	6/30(火)	13:30-15:00	9	柏図書館
BIOSIS Previews	6/30(火)	15:15-16:15	11	柏図書館
AGRICOLA Plus Text	7/30(木)	15:00-16:30	8	農学生命科学図書館
合計	39回		434名	

※柏図書館に中継したコース。人数欄に( )で中継参加者を内数で示す。

### 2.2.4 総合図書館オリエンテーション

総合図書館オリエンテーションに協力し、春季は「OPAC 入門」、秋季は「データベース講習会」として、各 30 分のコースを担当した。

コース名	月日	回数	人数	内容
OPAC 入門	4/8~4/17	8	79	OPAC、電子ジャーナルの探し方
OPAC 入門(英語)	4/13,4/15	2	15	OPAC、電子ジャーナルの探し方
留学生ガイダンス (日本語)	4/9	2	18	OPAC、電子ジャーナルの探し方
留学生ガイダンス (英語)	4/14,5/25	2	5	OPAC、電子ジャーナルの探し方
秋季留学生ガイダンス (日本語)	10/8	2	22	OPAC、電子ジャーナル、Web of Science の利用方法
秋季留学生ガイダンス (英語)	10/15,10/19, 10/29	6	60	OPAC、電子ジャーナル、Web of Science の利用方法
合計		22回	199名	

### 2.2.5 留学生向け情報探索ガイダンス

「留学生向け情報探索ガイダンス」を韓国語および中国語で開催した。韓国語コースについては、入学者の多い春季に試行的に開催した。両コースともに外国人講師が講習を担当した。

コース名	月日	回数	人数	
留学生向け情報探索ガイダンス	韓国語コース	5/14	1	7
	中国語コース	11/6	1	8
合計		2回	15名	

### 2.2.6 事務職員向け情報探索ガイダンス

事務職員が業務をする上で役立つと思われるデータベースの紹介を行ない、多くの参加を得た。

月日	時間	人数	内容
8/5	16:00-17:00	15	辞書・事典ツール、新聞・雑誌記事、官報、統計、企業情報の 検索方法や図書館サービス
8/11	11:00-12:00	18	
合計	2回	33名	

### 2.2.7 秘書さんのための論文の探し方講習会

事務職員講習会参加者の要望を受け、研究補助従事者を対象とした講習会を新たに実施し、大変好評であった。

月日	時間	人数	内容
11/12	13:30-15:00	19	辞書・事典ツール、雑誌論文、新聞記事の探し方など



## 2.2.8 出張講習会

利用者(教員、学生等)からの依頼に合わせて授業や研究室単位などで行なう、出張講習会を実施した。

研究科・研究室名など	月日	時間	人数	講習内容	場所
学際情報学府	4/10	11:30-12:30	77	JapanKnowledge+、OPAC、電子ジャーナル、CiNii、Web of Science 等	福武ホール ラーニングシ アター
教養学部「基礎演習」	4/17	9:45-10:30	24	JapanKnowledge+、OPAC、Webcat Plus、CiNii 等	駒場情報教育 棟 4 階 中演習室 1
	4/17	10:40-11:40	21		
文学部現代文芸論研究室「現代文芸論入門」	4/23	10:25-11:50	22	電子ジャーナル、JSTOR、LRC、文学部図書室ツアー等	総合図書館
(工学系・情報理工学系)文献検索早わかりコース	5/12	15:00-16:00	13	OPAC、Webcat、電子ジャーナル、CiNii、Web of Science、Engineering Village 等	工学部 3 号館 3F プロジェク トルーム 3C
	5/18	15:00-16:00	8		
	5/20	15:00-16:00	9		
	10/16	15:00-16:00	23		
	10/20	15:00-16:00	16		
	10/27	15:00-16:00	10		
理学部生物学科 動物学コース・植物学コース	5/27	15:00-16:00	6	JapanKnowledge+、Web of Science、PubMed、生物学図書室案内等	総合図書館
文学部 Academic Writing(I): Introductory	5/28	15:00-16:00	18	JapanKnowledge+、OPAC、Webcat、CiNii、Web of Science、PIO 等	総合図書館
	11/5	15:00-16:00	14		
工学部社会基盤「国際プロジェクト」	7/2	15:00-16:00	13	電子ジャーナル、Web of Science、Engineering Village 等	総合図書館
人文社会系研究科文化資源学専攻	8/7	15:30-17:00	6	JapanKnowledge+、CiNii、PIO、Web of Science、新聞記事、RefWorks 等	総合図書館
	8/14	13:00-14:30	4		
新領域創成科学研究科 国際協力学専攻	8/14	15:00-16:30	8	JapanKnowledge+、OPAC、CiNii、Web of Science、RefWorks 等	総合図書館



農学生命科学研究科 附属千葉演習林	9/17	10:30- 12:30	12	電子ジャーナル、 CiNii、農学生命科学 図書室利用案内等	農学生命科学 研究科附属千 葉演習林 清 澄学生宿舎
工学部精密工学科	10/1	14:15- 15:30	43	JapanKnowledge+、 OPAC、電子ジャー ナル、CiNii、Web of Science、Engineering Village 等	工学部 14 号 館 3 階計算機 演習室 (326 号室)
社会科学研究所	10/9	15:00- 15:30	9	新聞・雑誌記事、官 報、統計、企業情報の 検索方法	社会科学研究所図書室 端末コーナー
工学系研究科化学シス テム工学専攻 飯塚・水 流研究室)	10/22	10:30- 12:00	10	OPAC、電子ジャーナ ル、Web of Science、 RefWorks 等	工学部 8 号館 710 号室
文学部社会心理学調 査実習 1	10/28	10:00- 11:00	16	OPAC、Webcat、電子 ジャーナル、CiNii、 PsycINFO、Web of Science 等、総合図書 館、文学部図書室ツア ー	総合図書館
	10/28	11:00- 12:00	15		
教育学部教育資料調 査演習	11/2	9:00- 10:30	26	NDL-OPAC、OPAC、 CiNii、ERIC 等	駒場キャンパ ス情報教育棟 4 階 中演習室 2
工学部・工学系研究科 建築学専攻 藤井研究室	11/11	10:30- 12:00	11	JapanKnowledge+、 OPAC、電子ジャーナ ル、CiNii、Web of Science、RefWorks 等	総合図書館
in 駒場 II 秘書さんのた めの論文の探し方講習 会	11/20	15:30- 16:30	19	JapanKnowledge+、 OPAC、電子ジャーナ ル、CiNii、Web of Science、新聞記事等	先端科学技術 研究センター4 号館 3 階 316 号室
医者なら誰でも知って いる!?～医学系文献検 索入門	11/30	14:00- 15:00	12	電子ジャーナル、医中 誌 Web、PubMed、 Ovid MEDLINE 等	医学図書館 (1階マルチメ ディアコーナ ー)
	11/30	15:30- 16:30	17		
	11/30	17:30- 18:30	15		
教養学部後期課程 「英語」クラス	12/7	14:40- 16:10	11	LRC、JSTOR 等	駒場図書館 会議室
教養学部 「アメリカ近代史」	12/18	10:30- 12:10	25	American: History & Life、Eary American Newspapers、CiNii 等	駒場図書館 会議室
	合計	31 回	533 名		

### 2.2.9 セミナー・説明会

研究活動支援に資するため、セミナー「学術雑誌の評価指標を考える」を開催した。また、新規サービスの RefWorks 導入説明会、トライアルサービス Business Source Premier の説明会を開催した。いずれも、提供元から講師を招いて実施し、多くの参加を得た。

講習会名	月日	時間	人数	会場
セミナー 「学術雑誌の評価指標を考える」	5/13	15:00-16:30	67(8)	総合図書館※
RefWorks 説明会	5/14	15:30-17:00	63(20)	総合図書館※
	5/15	15:30-17:00	73	医学図書館
	5/20	15:30-17:00	13	駒場図書館
Business Source Premier トライアル説明会	12/10	16:00-17:00	11	総合図書館
合計	5回		227名	

※柏図書館に中継したコース。人数欄に( )で中継参加者を内数で示す。

# 図書館関係システム運用・管理

## 図書館システム係

### 1 運用報告

図書館システム係では、「附属図書館学術情報システム」を中心に附属図書館 Web サーバも含めた図書館関係のサーバ群の運用管理を行なっている。

#### 1.1 附属図書館学術情報システム

「附属図書館学術情報システム」は、いわゆる全学の図書館業務システムでもあり、また利用者に OPAC や MyLibrary 等のサービスを提供するシステムでもある。

2009 年度は、2011 年 3 月にシステムリプレイスが控えていることから、特にシステム改修等は行なわず、人事情報システムとの連携を想定した試験環境の構築を行なった。

また、昨年度から継続して「次期システム検討ワーキンググループ」を開催し、次期図書館システムのコンセプトの検討を行ない、その後引き続き「次期システム仕様検討ワーキンググループ」を発足させ、図書館職員と連携しながら仕様策定委員会に諮る叩き台の検討を行なった。

#### 1.2 その他の図書館関係サーバ

セキュリティパッチの摘要、インシデント対応、ハードウェア障害への対応等、日常的にサーバの保守・管理作業を行なっている。

### 2 講習会・研究会開催報告

図書館システム係では「図書系職員のためのアプリケーション開発講習会」を主催している。本講習会は 2006 年度から継続しているが、その成果であるアプリケーションはすでに約 30 件ほどになっており、特にそのうちの「図書館間返送管理システム」は 2009 年度の業務改善総長賞を受賞した。

**「図書系職員のためのアプリケーション開発講習会」成果**

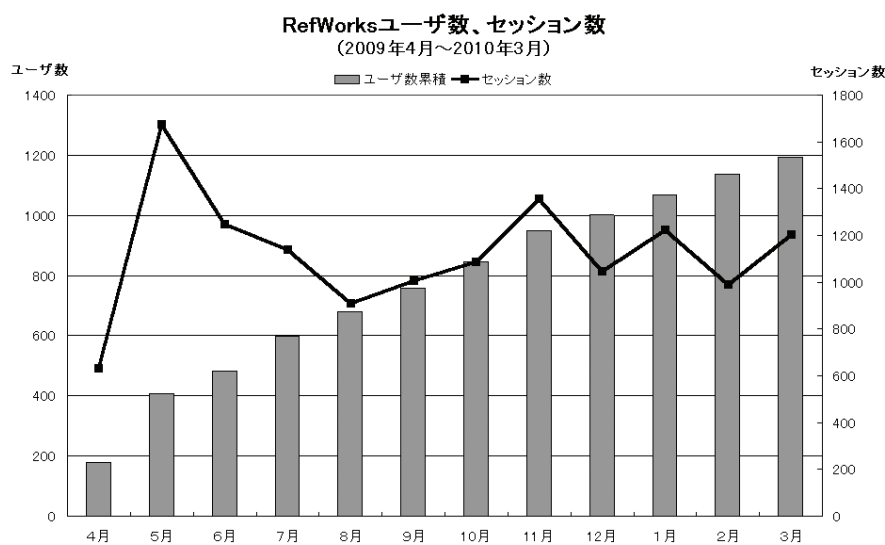
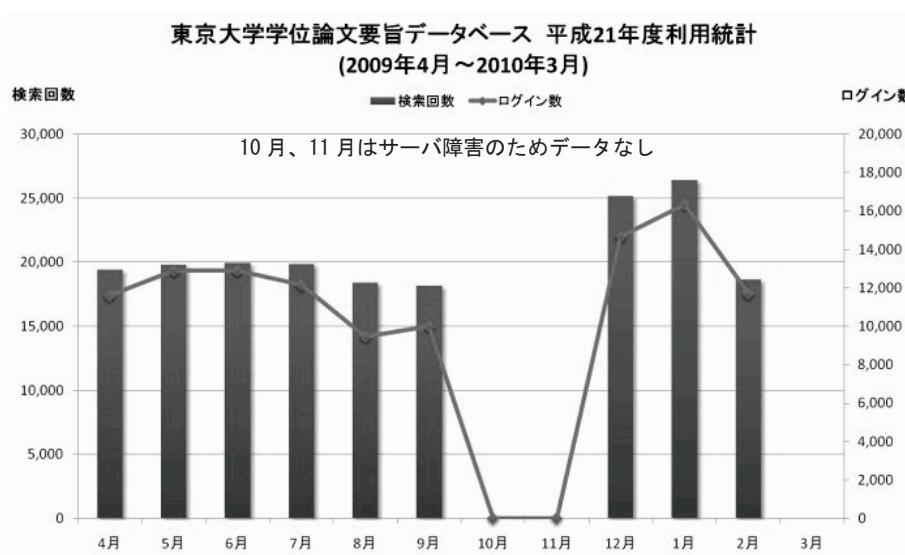
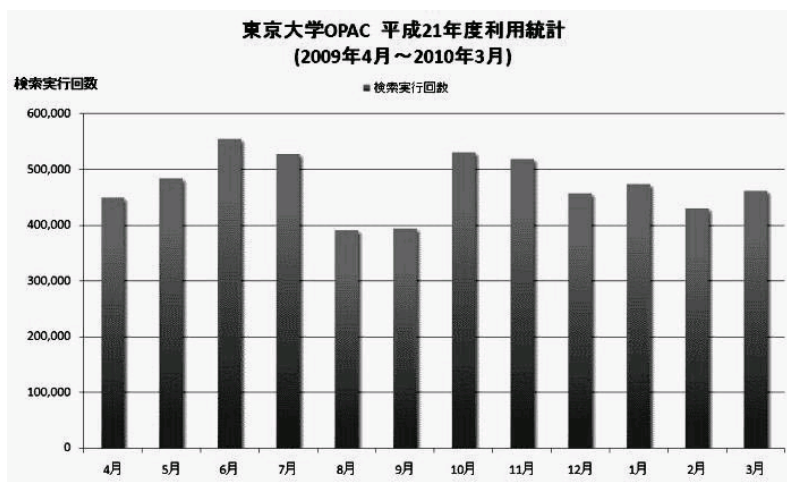
「図書系職員のためのアプリケーション開発講習会」は、東京大学情報基盤センターが主催する東京大学の図書系職員向けの講習会です

このページでは、講習会の受講生が開発したアプリケーションを公開しています。

1. 利用者向けサービス

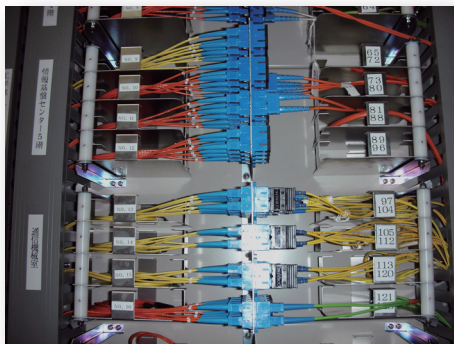
サービス名及びリンク	解説	開発者
東京大学版 LibX	LibXは、Webブラウザ「Firefox」に各大学図書館向けの付加機能をつけるプラグインソフトです。「東京大学版LibX」は文字となり、その東京大学版になります。Webブラウザと連携した東大OPACの検索機能、Amazonからの東大OPACリンク、UT Article Link 対応などさまざまな機能があります。(参考)「東大版LibXが自動生成するリンク」 <small>Update 2009.07.22</small>	赤津 実
iGoogleガジェット集	GoogleのポータルGoogle上で使用できる、自作のiGoogleガジェット集です。OPACリンク集の「東大図書館キート」ガジェットと、東京大学生命科学構造化センターのCSLS Search(P4M4)に専門用語の日本語表示機能などを搭載した検索システムXのガジェットがあります	近藤 真 留子
東京大学OPACウィジェット	ウィジェット(ガジェット)とは、アプリケーションソフトの一部としてか、デスクトップ上に常駐させる小画面のアプリケーションのことです。ウィジェットには種類がいくつもありますが、Macintosh OS XのDashboard用の東京大学OPACウィジェットを作成しました。	田口 忠 祐
iPhone用東京大学OPAC	iPhone及びiPod touch用に画面デザインをした、東京大学OPACです。 <i>New!</i>	田口 忠 祐
UT Article Search	東京大学のデータベースの横断検索=UT Article Search(学内限定)では、横断検索を行うデータベースを数多くの中から選択する必要があります。このMy UT Article Searchは、使用するPCごとに検索対象データベースの選択結果を保持するものです。東京大学内限定限定のツールですが、プログラムのソースを変更することで他大学のCentral Searchにも対応できます。Firefox版(使用版 2008.07.30値)とAdobe AIR(実験版)の2種の実装を行いました。	前田 朗
東大OPACから東大生協位置情報へのリレー検索	東京大学OPACの書籍詳細表示画面から、東京大学生協位置情報検索の確認をシームレスに行なえるツールです。これは、東大図書館の所蔵資料で自分用に欲しい本を、キャンパス内の生協で即購入したいという利用者のニーズに対応するものです。IE7やFirefoxブラウザでのご利用が可能です。なお、和書限定になりますので、その点はご了承ください。 <small>Update 2010.03.18 従来のブックマークレットに加えて、アドオン版も公開</small>	小川 雅 弘

### 3 サービス統計

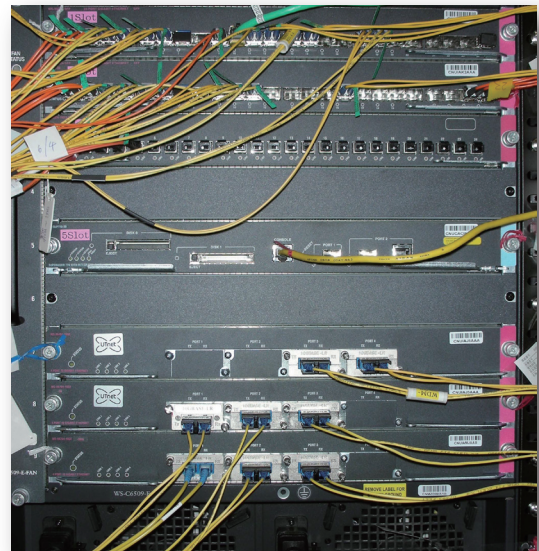




## ネットワーク部門



光ケーブルが集線されているsprayボックス



UTnet3の代表的な機器(レイヤ3スイッチ)





# ネットワーク部門

## 概要

部門長 若原 恭

副事務長 早野裕士

### ■東京大学情報ネットワークシステム(UTnet3)の運用管理

本学の情報ネットワークシステム UTnet3(University of Tokyo network system 3)は、各建物内の支線ネットワーク(支線)及び建物間接続や学外との接続のための基幹ネットワーク(基幹)とから構成されている。情報基盤センターの本ネットワーク部門は、以下に示す通り、基幹の運用管理に加え、これに関連する業務も含めて取り組んだ。

#### 1. 基幹ネットワークの運用管理

基幹の中核を占めるネットワーク機器は、本郷地区、駒場Ⅰ地区、駒場Ⅱ地区、柏地区、白金地区、中野地区の各ハブサイトに設置したギガビットスイッチ(L3)であり、基幹の基本トポロジーは情報基盤センターを中心にしたスター型になっている。基幹のギガビットスイッチ相互間の信頼性向上のための回線の二重化等の増強を実施した。

#### 2. 基幹ネットワークと支線ネットワークとの接続

各建物にレイヤー2 スイッチを設置し、基幹のギガビットスイッチ(L3)と支線を接続しているが、ネットワーク利用の増大への対応と高機能化のため、2009 年度も 2008 年度に引き続き、レイヤー2 スイッチの順次更新を行った。また、支線ネットワークから 10Gbps による接続要求に伴い、基幹のギガビットスイッチ(L3)へ接続できる環境を整備した。

#### 3. VLAN 対応

UTnet3 では仮想 LAN(VLAN: Virtual LAN)に対応している。VLAN によって、部局や研究室が複数の建物やキャンパスに分散して配置されている場合でも部局や研究室のまとまりごとに同一のサブネットに収容することを可能としている。本学では、建物の新設や組織変更等が少なくなく、それに伴った VLAN に関する要求は非常に多く、それら要求に応じてネットワーク機器の設定変更や増設等を実施した。

#### 4. キャンパス間接続

キャンパス間接続については、本郷地区と柏地区間を 1Gbps から 10Gbps に高速化した。

#### 5. 遠隔地施設接続

理学部小石川植物園、日光植物園、総合研究博物館小石川分館、柏の葉ロッジを公衆イーサ網や SINET3 回線で UTnet 接続した。また、本部国際系事務室(角川本郷ビル)は、構内光ケーブルを敷設して、UTnet 接続した。

#### 6. 無線 LAN サービス

山上会館、武田先端知ビル内の武田ホール、弥生講堂・一条ホール、柏地区の柏図書館内メディアホール、医学部教育研究棟14階 鉄門記念講堂で無線 LAN サービスを提供した。

## 7. その他関連業務

UTnet3 基幹ネットワークの運用管理に関連する業務として、全学法定点検のための停電措置による基幹のギガビットスイッチ(L3)の停止を回避するために HUB サイト内に設置している無停電電源装置(UPS:Uninterruptible Power Supply)のバッテリー増強、ドメイン名の割り当て、DNS(Domain Name System)のサービス、UTnet 光ファイバケーブル専用利用の各業務に取り組んだ。

### ■セキュリティ対応

最近のネットワークにおいてはセキュリティへの対応が必須となっており、本学においてもセキュリティインシデントは日常茶飯事と言っても過言ではないほどに頻繁に発生している。特にコンピュータへの不正侵入の試み、電子メールや Web ページを通してのウィルス感染、P2P(Peer to Peer)、SPAM メール等が問題となっている。そこで、本学のセキュリティの維持確保に向け、以下の通り、セキュリティ対応業務を実施した。

#### 1. 不正侵入の監視

学外と学内との間の通信に関し、基幹部分において不正侵入の試みを検知し事前に予防する仕組みとして、UTnet3 では不正侵入検知装置(IDS: Intrusion Detection System)と異常トラフィック監視システムを導入して監視し、その結果に応じて必要な対処を実施した。

#### 2. ウィルス感染への対応

メールや Web ページ等を通してのウィルス感染に関する対策として、情報基盤センターでは、適切なウィルス対策ソフトウェア製品を学内利用者の希望に応じて配布し、予防に役立てた。また、新しいウィルス対策ソフトウェアの導入による検知対象範囲の拡大等、ウィルス対策の強化を積極的に進めた。

#### 3. SPAM 対策

全学的に激増する SPAM メールに対して、全学的な対策として、本部事務組織と連携協力して、メールサーバ単位での SPAM 対策の運用を実施した。

### ■ 東京大学情報システム緊急対応チーム(UT-CERT)との連携

UTnet3 ではセキュリティに関する対応として、基幹部分のギガビットスイッチ(L3)で、全学的に問題となる事象や各部局の個別の事象についてのフィルタを設定している。東京大学情報システム緊急対応チーム(UT-CERT)から委託業務を受け、このフィルタを活用することによって、全学的なセキュリティ対策を行った。

### ■ 学内ソフトウェアライセンス

学内に利用者の多いソフトウェアについて、全学サイトライセンスを取得し、そのライセンス管理と配布サービスを行った。

### ■ 関係委員会

情報ネットワークに関する事項について、全学的視点から企画、立案及び審議を行う情報基盤センターネットワーク専門委員会が下記のとおり行われた。

2009年 8月 3日 第33回情報基盤センターネットワーク専門委員会

2009年 10月 26日 第34回情報基盤センターネットワーク専門委員会

2010年 2月 8日 第35回情報基盤センターネットワーク専門委員会

主な検討事項および承認事項

- ・基幹ネットワークの構成変更について

- 10G 接続の負担金について
- 光ファイバケーブル専用利用の業務利用について
- 情報セキュリティポリシー関連
- 対外接続について
- 遠隔施設接続について

# 東京大学情報ネットワークシステム(UTnet3)の運用管理

## ネットワーク係

### 1 運用報告

本学の情報ネットワークシステムは UTnet(University of Tokyo network system)と称し、これまでの更新経緯に応じて現在は UTnet3 と呼ばれている。UTnet3 は支線ネットワーク(支線)と基幹ネットワーク(基幹)とから構成されている。支線は、各建物内に設置されたネットワークで、当該部局によって運用管理されている。基幹は、支線との相互接続及び学外との接続のために設置されたネットワークで、情報基盤センターの本ネットワーク部門が運用管理している。本部門では、以下に示す通り、基幹ネットワークの運用管理に加え、これに関連する業務も含めて取り組んだ。

#### 1.1 基幹ネットワークの運用管理

本郷地区の基幹には、5ヶ所の HUB サイト(情報基盤センター、電話庁舎、附属図書館、工学部 8 号館、農学部 3 号館)がある。各 HUB サイトにはコア用レイヤ 3 スイッチ(L3SW)を設置し、基幹は情報基盤センターを中心にしたスター型の構成としている。駒場 I 地区、駒場 II 地区、柏地区、白金地区、中野地区の各郊外地区についても、L3SW を設置した HUB サイトから専用回線等を介して本郷地区と接続している。このような基幹の運用管理について、主に以下の取り組みを行った。

- －本郷地区の L3SW の機能拡張(CPU 増強、10Gbps(以下、Gbps を G と略す)I/F 増強、1G I/F 増強、OS 更新)
- －柏地区のアップリンク(SINET 接続)への 10G I/F を追加
- －工学部 8 号館、農学部 3 号館 HUB サイト設置の無停電電源装置のバッテリー増設
- －短周期ポーリングが可能なトラフィック可視化・監視ツールの導入

以上の通り監視の強化・安定化運用の対策等を実施し、基幹に対する学内からの要求に応えることによって、本学の研究・教育の一層の円滑な推進を図った。基幹のネットワーク構成を図 1 に示す。

また、本郷地区と、駒場 I 地区、駒場 II 地区、柏地区、白金地区、中野地区の各キャンパス間のトラフィック量の推移を図 2 に示す。

#### 1.2 基幹ネットワークと支線ネットワークとの接続

基幹と各支線の接続については、各建物に設置したエッジ用レイヤ 2 スイッチ(L2SW)で行っており、この L2SW は HUB サイトの L3SW から伸ばした光ファイバで L3SW と接続している。UTnet3 導入(2001 年度)当初には L2SW を約 200 台設置した。以来、撤去や新設で台数の増減はあったものの、今年度は新設が多く約 40 台の L2SW を更新し、現在では約 220 台のエッジ用 L2SW を設置している。これらの L2SW は老朽化したものから優先的に更新しており、支線との接続速度がギガビットに対応した機種を配置している。また、更新時の I/F 使用量に応じて、I/F 数の少ない省電力機種も配置している。

#### 1.3 VLAN 対応

部局や研究室が、複数の建物やキャンパスに分散配置されたり、他部局の建物内に入居するケースがある。UTnet では、部局や研究室と同一のサブネットを VLAN で延長することができる。今年度実施した基幹ネットワークの構成変更及びサブネットの割当は表 1 に示す。

## 1.4 キャンパス間接続及び対外ネットワーク接続

以下の通り、学内キャンパス間接続及び対外ネットワークとの接続に関して、変更作業を実施し、高機能化・高速化を達成した。

－2009年5月に、本郷地区～柏地区間の回線(SINET 経由)を10Gに増速

SINET 柏ノードに10G接続のI/Fを追加し、1GのVLANパスを10Gに変更した。

なお、中野地区に所在していた海洋研究所が柏地区に移転(2010年2月～3月)したために、2月以降のトラフィックが減っている。

## 1.5 遠隔地接続

以下の通り、遠隔研究施設等の接続構成変更を行った。

－2009年10月に、理学部小石川植物園をNGN(10Mbps)で新規接続

－2010年2月に、総合研究博物館小石川分館をNGN(10Mbps)で新規接続

－2010年2月に、角川本郷ビルをUTnetの光ファイバ直結で新規接続

－2010年2月に、理学部附属植物園日光分園をSINET 経由 B フレッツ VPN で新規接続

－2010年3月に、柏インターナショナルロッジをNGN(10Mbps)で新規接続

## 1.6 その他関連業務

### (1)ドメイン名の割当等

ネットワーク専門委員会の承認を経て、新規割当て4件、廃止1件を実施した。詳細を表2に示す。

### (2)UTnet 光ファイバケーブル専用利用

建物間を横断するUTnet 光ファイバケーブル専用利用の変更、及び割当を行った。結果を表3に示す。

### (3)無線 LAN サービス

山上会館、武田ホール(武田先端知ビル)、一条ホール(弥生講堂)、鉄門記念講堂、柏図書館内のメディアホールにおいて無線 LAN サービスを提供した。それらの部局別の申請利用状況を表4に示す。

### (4)公衆無線 LAN 接続サービス

UTnet が整備した無線 LAN サービス拠点において、通信事業者(キャリア)の提供する公衆無線 LAN 接続サービスも利用できるように、アクセスポイントの相乗り試行運用を行っている。2006年10月からソフトバンクテレコム社が提供するBBモバイルポイントが相乗りに参加していたが、2010年3月からはライブドア社が提供するlivedoor wirelessも相乗りに参加した。利用者は、各社指定の接続方法にて公衆無線 LAN を利用できる。

### (5)トラフィック可視化・監視ツールの導入

従来UTnetの全て(約250台)のL2SW/L3SWをフリーソフトのMRTGで監視していたが、グラフ表示が重く、メンテナンスにも手間がかかっていた。この問題を解決するため、過去(日時指定)に遡ったグラフ表示、1分という短周期でのポーリングによるSNMP採取が可能で、メンテナンスが手軽、グラフ表示も軽く、見たいI/Fを即時に表示できる市販のツールを導入した。これにより、障害切分に必要な時間が短縮し、各I/Fのトラフィック流量確認が容易になった。

## 2 講習会・研究会開催報告

[第69回コンピュータ・ネットワーク利用セミナー“最近のセキュリティ事情”]

学内の活動にはなくてはならないPCやインターネットを使う上で、今、もう一度あるべきセキュリティを考え直すために、最近のセキュリティ事情についてセミナーを行った。参加者は、情報基盤センターの関係者を含め45名。

・期間 2009年8月5日(水)



- ・場所 情報基盤センター4F 遠隔講義室、駒場情報教育棟 4F 遠隔講義室、  
柏図書館メディアホール

プログラムは次の通り

1. 最近のセキュリティ事情

一般社団法人 JPCERT コーディネーションセンター 早貸 淳子

[第7回 UTnet Meeting]

直近のネットワーク動向や管理の問題点に関して情報交換を行うための集まりとして開催している UTnet Meeting を行った。参加者は、情報基盤センターの関係者を含め 52 名。

- ・期間 2009 年 10 月 30 日(金)
- ・場所 農学部弥生講堂・一条ホール

プログラムは次の通り

- |                             |                |
|-----------------------------|----------------|
| 1. UTnet update             | 友西 大           |
| 2. セキュリティ報告(ウイルス、迷惑メール)     | 椿山惣一郎          |
| 3. UT-CERT 報告               | 中山雅哉           |
| 4. NAT ルータ利用時の留意点           | 工学系研究科 三木修次    |
| 5. 本部事務システムのシンクライアントの導入について | 本部情報化推進 G 櫻井清隆 |
| 6. フリーディスカッション              |                |

[第70回コンピュータ・ネットワーク利用セミナー“Chem Bio Office 活用法セミナー”]

化学構造式を描画する方法(ChemBioDraw)から、ChemBioOffice に含まれるその他のソフトウェアの活用法について、新バージョン(v12/2010)の情報を交えて紹介を行った。参加者は、情報基盤センターの関係者を含め 3 名。

- ・期間 2009 年 11 月 10 日(水)
- ・場所 情報基盤センター4F 遠隔講義室

プログラムは次の通り

1. Chem Bio Office 活用法 CambridgeSoft 久枝秀次

[平成 21 年度東京大学技術職員研修(コンピュータ関係)]

本学職員のうち、研究室やセンター等においてコンピュータやコンピュータ・ネットワークの初心者レベルの利用者、及びそれらの運用管理業務やセキュリティ対策業務に携わる初心者レベルの管理者を対象に、セキュリティ対応を含むコンピュータやコンピュータ・ネットワークの利用に関する研修を、以下の通り実施した。参加者は 17 名であった。

- ・期間 2009 年 11 月 24 日(火)～11 月 26 日(木)(3 日間)
- ・場所 情報基盤センター4F 講義室

◇外部講演

[第31回全国共同利用情報基盤センター研究開発連合発表講演会]

旧 7 帝大の情報基盤センターと国立情報学研究所に勤務する教職員が、日頃の成果を互いに発表しあい技術の研鑽に資するために開催している発表講演会。16 セッションのうち 1 つで講演を担当した。

- ・期間 2009 年 11 月 6 日(金)
  - ・場所 国立情報学研究所 12F 会議室
1. 迷惑メール対策サービスの試行運用

小菌隆弘

[学術認証フェデレーション及び SINET サービス説明会]

全国の大学等教育研究機関向けに、学術認証フェデレーション及び SINET3 への理解を深めることを目的として、国立情報学研究所(NII)が主催した全国説明会を本センターの共催で行った。東京会場

で開催された説明会にて、3セッションある「SINET3 利用事例紹介」のうち1つで講演を担当した。参加者は、約130名。

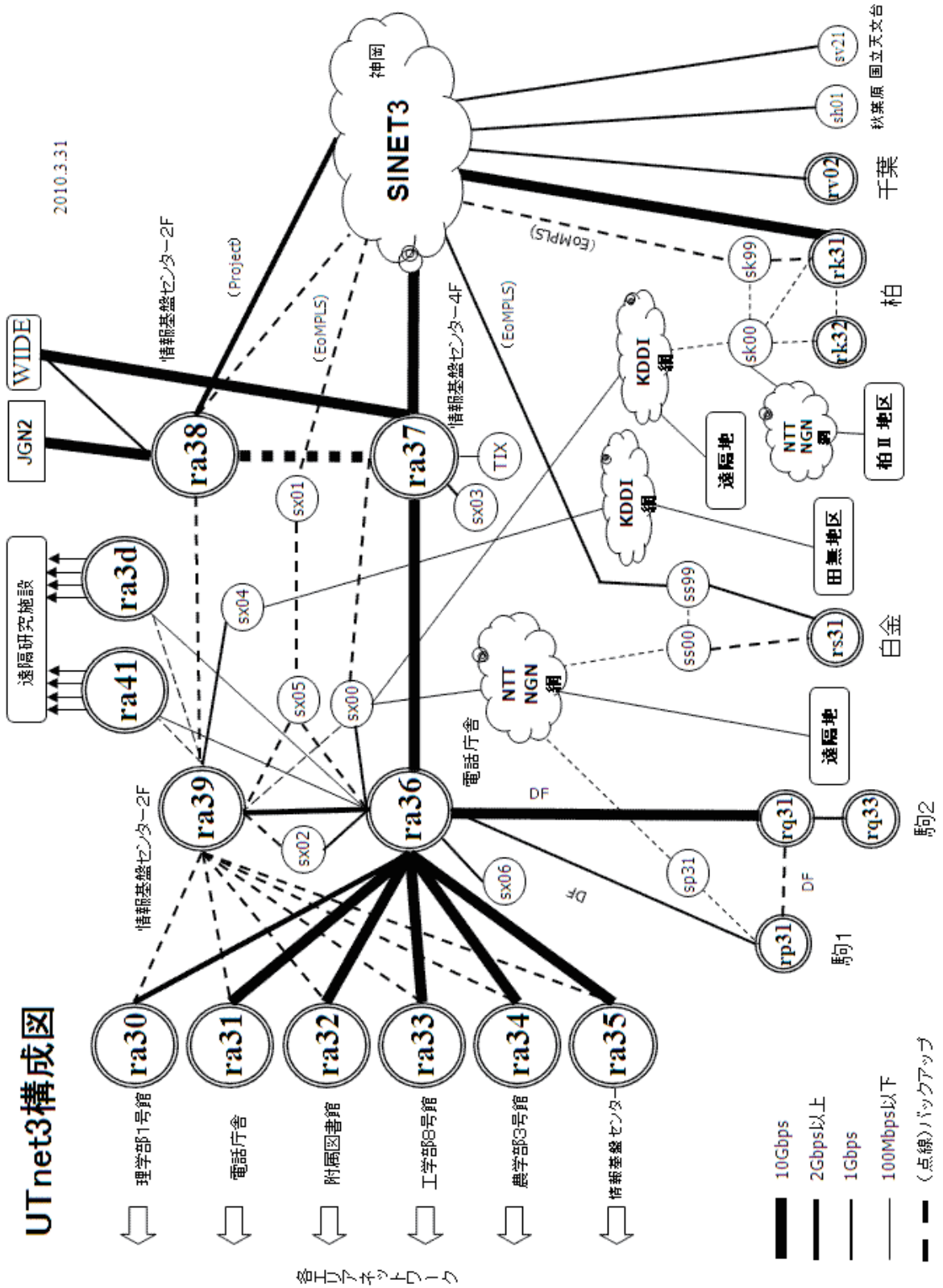
( <http://www.sinet.ad.jp/inform/news-1/20100126/> )

・期間 2009年12月8日(火)

・場所 東京大学小柴ホール

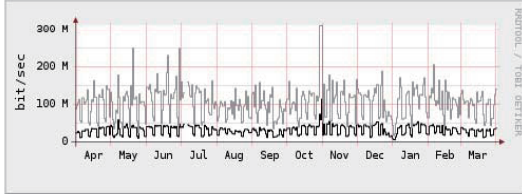
1. 東京大学情報ネットワークシステム UTnet3 の運用について

下田哲郎

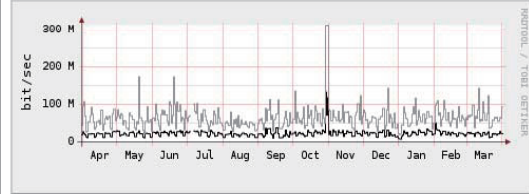


## キャンパス間回線トラフィック

駒場1 → 本郷

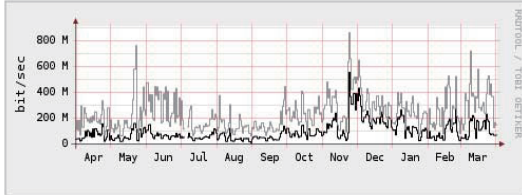


駒場1 ← 本郷

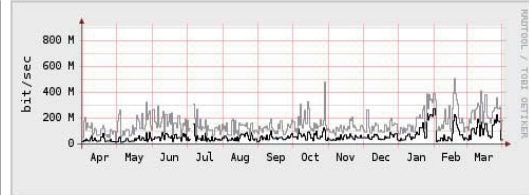


1Gbps MAX

駒場2 → 本郷

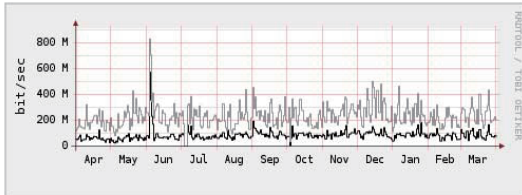


駒場2 ← 本郷

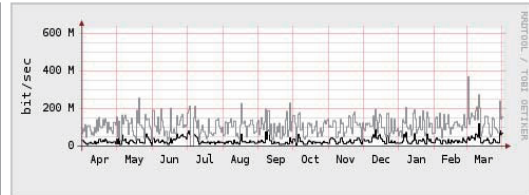


10Gbps MAX

柏 → 本郷

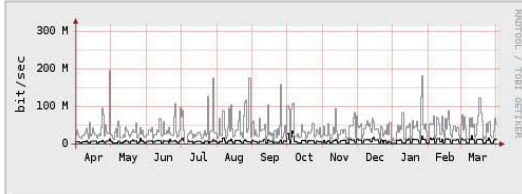


柏 ← 本郷

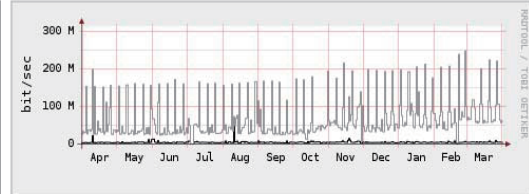


10Gbps MAX: SINET 経由 (5月20日までは1Gbps MAX: SINET MPLS 経由)

白金 → 本郷

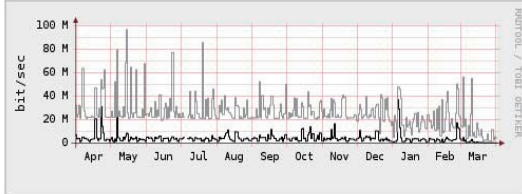


白金 ← 本郷

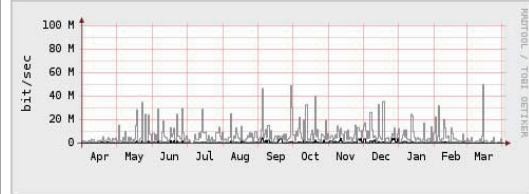


1Gbps MAX: SINET MPLS 経由

中野 → 本郷



中野 ← 本郷



100Mbps MAX

— 1日平均トラフィック(本郷-各キャンパス間)  
 - - 1日最大トラフィック(本郷-各キャンパス間)  
 凡例:

図2 UTnet のキャンパス間トラフィック

表 1 基幹ネットワークの構成変更及びサブネットワークの割当

項番	申請部局名	設置場所	変更内容
1	大学院理学系研究科	理学部1号館	建物間 VLAN の申請
2	大学院医学系研究科	医学部2号館	建物間 VLAN の申請
3	大学院理学系研究科	理学部1号館	サブネットアドレスの申請
4	大学院農学生命科学研究科	農学部3号館	建物間 VLAN の申請
5	本部事務組織	医学部1号館、学士会分館	建物間 VLAN の申請
6	情報基盤センター	情報基盤センター	建物間 VLAN の申請
7	社会科学研究所	学士会分館	建物間 VLAN の申請
8	生産技術研究所	生産技術研究所	サブネットアドレスの申請
9	大学院理学系研究科	理学部1号館	サブネットアドレスの申請
10	本部事務組織	教養学部8号館	建物間 VLAN の申請
11	情報基盤センター	情報基盤センター・物性研本館	建物間 VLAN の申請
12	柏地区事務部	新領域基盤棟	建物間 VLAN の申請
13	情報基盤センター	工学部7号館、工学部14号館	建物間 VLAN の申請
14	本部事務組織	第2本部棟	建物間 VLAN の申請
15	大学院総合文化研究科	教養学部14号館	基幹ネットワークの構成変更
16	本部事務組織	学術交流研究棟	建物間 VLAN の申請
17	本部事務組織	経済学研究科棟・農学部3号館・ 教育学部・薬学部総合研究棟	建物間 VLAN の申請
18	情報基盤センター	社会科学研究所	建物間 VLAN の申請
19	本部事務組織	本部アネックス	建物間 VLAN の申請
20	大学院情報理工学系研究科	情報基盤センター	建物間 VLAN の申請
21	大学院医学系研究科	医学部1号館	基幹ネットワークの構成変更
22	本部事務組織	本部棟	サブネットアドレスの申請
23	社会科学研究所	理学部研究 A 棟	建物間 VLAN の申請
24	本部事務組織	本部棟	基幹ネットワークの構成変更及び サブネットアドレスの申請
25	大学院教育学研究科	情報基盤センター	基幹ネットワークの構成変更及び サブネットアドレスの申請
26	情報基盤センター	工学部1号館	建物間 VLAN の申請
27	情報基盤センター	理学部化学館	建物間 VLAN の申請
28	大学院新領域創成科学研究科		サブネットアドレスの申請
29	本部事務組織	本部棟	基幹ネットワークの構成変更
30	本部事務組織	情報基盤センター	建物間 VLAN の申請
31	本部事務組織	本部棟	建物間 VLAN の申請
32	本部事務組織	附属病院・分子細胞生物学研究所・ 教養学部アドミニストレーション棟・ 低温センター・生産技術研究所・ 海洋研究所・教育学部・情報学環・ 東洋文化研究所・社会科学研究所・ 史料編纂所・総合図書館・薬学部本館・ 工学部列品館	建物間 VLAN の申請
33	本部事務組織	附属高等学校	建物間 VLAN の申請
34	情報基盤センター	情報基盤センター	基幹ネットワークの構成変更及び サブネットアドレスの申請
35	大学院法学政治学研究科	法文1号館・法学部(3,4)号館・ 法学政治学総合教育棟・史料編纂所・ 弥生地区総合研究棟・第2本部棟・ 医学部1号館・情報基盤センター・ 駒場地区情報教育棟	基幹ネットワークの構成変更



36	大学院法学政治学研究科		サブネットアドレスの申請
37	情報基盤センター	情報基盤センター	建物間 VLAN の申請
38	本部事務組織	総合研究博物館・法文1号館	建物間 VLAN の申請
39	数物連携宇宙研究機構	数物連携宇宙研究機構研究棟	サブネットアドレスの申請
40	数物連携宇宙研究機構	数物連携宇宙研究機構研究棟	基幹ネットワークの構成変更及びサブネットアドレスの申請
41	生産技術研究所	生産技術研究所	基幹ネットワークの構成変更
42	情報基盤センター	新領域基盤棟	基幹ネットワークの構成変更及びサブネットアドレスの申請
43	本部事務組織	受変電設備棟	建物間 VLAN の申請
44	柏地区事務部	数物連携宇宙研究機構研究棟	建物間 VLAN の申請
45	大学院理学系研究科	通信機器室	建物間 VLAN の申請
46	本部事務組織	医学部生命科学実験棟	建物間 VLAN の申請
47	本部事務組織	教養学部101号館・柏地区総合研究棟	建物間 VLAN の申請
48	大学院人文社会系研究科	法文(1,2)号館	建物間 VLAN の申請
49	大学院総合文化研究科	教養学部9号館	建物間 VLAN の申請
50	本部事務組織	新領域基盤棟	建物間 VLAN の申請
51	情報基盤センター	情報基盤センター・駒場情報教育棟	基幹ネットワークの構成変更
52	情報基盤センター	工学部(1,4,14)号館	建物間 VLAN の申請
53	情報基盤センター	情報基盤センター	サブネットアドレスの申請
54	生産技術研究所	生産技術研究所	基幹ネットワークの構成変更及びサブネットアドレスの申請
55	本部事務組織	向ヶ岡ファカルティハウス	建物間 VLAN の申請
56	本部事務組織	総合研究博物館小石川分館	建物間 VLAN の申請
57	本部事務組織	教養学部8号館	建物間 VLAN の申請
58	本部事務組織	角川本郷ビル	建物間 VLAN の申請
59	海洋研究所	大気海洋研究所研究棟・海洋研究所	基幹ネットワークの構成変更
60	大学院理学系研究科		サブネットアドレスの申請
61	本部事務組織	総合研究博物館小石川分館	建物間 VLAN の申請
62	先端科学技術研究センター	先端研(1,4)号館	建物間 VLAN の申請
63	情報基盤センター	情報基盤センター	サブネットアドレスの申請
64	地震研究所	地震研究所1号館	基幹ネットワークの構成変更
65	大学院総合文化研究科	教養学部10号館	基幹ネットワークの構成変更
66	本部事務組織	教養学部101号館・駒場保健センター・柏保健センター	建物間 VLAN の申請
67	大学院情報理工学研究科	第2本部棟	建物間 VLAN の申請
68	大学院情報理工学研究科	第2本部棟	サブネットアドレスの申請
69	大学院新領域創成科学研究科	新領域基盤棟・柏地区総合研究棟	建物間 VLAN の申請
70	史料編纂所	医学部2号館	建物間 VLAN の申請
71	大学院法学政治学研究科	医学部2号館	建物間 VLAN の申請
72	本部事務組織	物性研本館・大気海洋研究所研究棟・本部棟・法学部4号館	建物間 VLAN の申請
73	本部事務組織	柏2地区インターナショナルロジ	建物間 VLAN の申請
74	大学院理学系研究科	通信機器室・情報基盤センター	建物間 VLAN の申請
75	大学院法学政治学研究科	法文(1,2)号館・法学部3号館	建物間 VLAN の申請
76	情報基盤センター	情報基盤センター	建物間 VLAN の申請



表 2 ドメイン名の割当

項番	ドメイン名	申請部局名	備考
1	lsn.u-tokyo.ac.jp	本部事務組織	生命科学研究ネットワーク
2	aviation.u-tokyo.ac.jp	本部事務組織	総括プロジェクト機構 航空イノベーション総括寄付講座
3	aori.u-tokyo.ac.jp	海洋研究所	大気海洋研究所
4	htpc.u-tokyo.ac.jp	新領域創成科学研究科	高温プラズマ研究センター(廃止)
5	asnet.u-tokyo.ac.jp	東洋文化研究所	日本・アジアに関する教育研究ネットワーク

表 3 UTnet 光ファイバケーブル専用利用の変更及び割当

項番	申請部局	利用区間	種類	
1	医科学研究所	ゲノムセンター-白金地区総合研究棟	GI2 芯	新規
2	大学院情報理工学研究科	工学部1号館-情報基盤センター	SM2 芯	新規
3	情報基盤センター	物性研本館-新領域環境棟	SM2 芯	新規
4	情報基盤センター	物性研本館-新領域環境棟	SM2 芯	廃止
5	生産技術研究所	先端研15号館-生研 C 棟	SM2 芯	新規
6	生産技術研究所	先端研15号館-生研 C 棟	SM2 芯	廃止
7	本部事務組織	総合研究博物館- 法文(1,2)号館・経済学研究科棟・教育学部・ 工学部(1,2,3,5,6,8,14)号館・理学部1号館・ 農学部3号館・医学部2号館・ 薬学部総合研究棟	GI 各 1 芯	新規
8	情報基盤センター	総合図書館-農学部図書館	SM2 芯	新規
9	大学院情報理工学研究科	情報基盤センターHUB サイト-工学部6号館	SM2 芯	新規
10	本部事務組織	工学部8号館- 工学部(1,2,3,5,6,14)号館	GI7 芯	新規
11	大学院経済学研究科	総合図書館-赤門総合研究棟	SM4 芯	新規

表 4 無線 LAN サービス利用の申請部局・組織別申請件数

項番	申請部局・組織	山上 会館	武田 ホール	弥生 講堂	柏 図書館	鉄門記念 講堂
1	サステナビリティ学連携研究機構	3				1
2	地球観測データ統合連携研究機構		2	1		1
3	数物連携宇宙研究機構			1	5	
4	政策ビジョン研究センター					1
5	大学発教育支援コンソーシアム推進機構	1				
6	国際連携本部	1				
7	男女共同参画オフィス	1				
8	国際系	1				
9	大学総合教育研究センター			1		
10	空間情報科学研究センター	1				
11	インテリジェント・モデリング・ラボラトリー		1			
12	法学政治学研究科	3				
13	医学系研究科	1				3
14	工学系研究科	11	9	2		2
15	人文社会系研究科	3				
16	理学系研究科		1	1		
17	農学生命科学研究科			5		
18	経済学研究科	1				
19	総合文化研究科					1
20	薬学系研究科					2
21	新領域創成科学研究科	1			6	
22	情報理工学系研究科	3	1			1
23	公共政策大学院	1				
24	東洋文化研究所	1				
25	社会科学研究所	1				
26	生産技術研究所	2				
27	分子細胞生物学研究所	1				
28	東大 GCOE「機械システム・イノベーション国際拠点」		2			
29	ライフサイエンス統合データベースセンター		1			
30	宇宙航空研究開発機構		1			
31	株式会社ユニバーサルエネルギー研究所		1			
32	科学技術振興機構		1	1		
33	国立西洋美術館			1		
34	北海道大学大学院情報科学研究科 GCOE			1		
35	理化学研究所					1

# セキュリティ対応

## 分散システムセキュリティ支援係

### 1 運用報告

最近のネットワークにおいては、高速な接続性の実現とともに、セキュリティへの対応も重要となっている。本学においてもセキュリティインシデントは日常茶飯事と言っても過言ではないほどに頻繁に発生している。特にコンピュータへの不正侵入の試み、電子メールや Web ページを通じてのウイルス感染は問題で、感染を受けた側の被害はもとより、学内や学外への不正アクセスの踏台となることも珍しくない。しかし、一方で、セキュリティを厳重にすることは利用者の利便性の低下にもつながる。一般論としていえば、セキュリティ対策の要点は安全性と利便性のトレードオフといえるが、本学のように多種多様な部局が存在している環境では、このトレードオフのバランス点を一つに収束させることは容易ではない。このような配慮のもと、本学のセキュリティの維持確保に向け、以下の通り、セキュリティ対応業務を実施した。

#### 1.1 不正侵入の監視

本学内のコンピュータに対して不正なアクセスを試みている可能性を検出し、必要ならば当該コンピュータが収容されているネットワークの管理者と連絡をとって不正侵入を防ぐべく対応を行った。このような不正アクセスの検知には、以下の2システムを活用した。

- 不正侵入検知システム (IDS :Intrusion Detection System)

これは、通常のネットワークアクセスでは発生しないようなアクセスパターン (例えば多数のホストを順にスキャンしていく動作など)を対象として検知し、警告を発するシステムである。マカフィー社の I-4000 と M-8000 を併用し、学内と学外との通信をタップして監視を行っている。2007 年度から引き続き P2P 通信の監視を行い、P2P 通信情報を各部局のセキュリティ担当者宛に電子メールでの連絡を行っている。また、7 月より I-4000 において附属病院及び地震研究所のローカルネットワークの通信情報を部局の端末で監視するためのサービスを試験的に開始した。

- 異常トラフィック監視システム

これは、サービス妨害攻撃またはサービス不能攻撃などと呼ばれる、インターネット経由での不正アクセスの 1 つで、大量のデータや不正パケットを送りつける行為などの不正な攻撃を検知できるシステムである。実際には、沖電気工業社製 Secure Traffic Probe (ソフトウェア)を導入し、学内と学外との通信をタップして、ポート単位での異常なトラフィックの変化について監視した。2009 年度においては、10Gbps 用のトラフィックキャプチャシステム (SIRIUS LINE-RATE CAPTURE STORAGE)の購入及び設置を行った。

#### 1.2 ウイルス対策

Windows パソコンやサーバマシン等における、ウイルスの脅威と感染被害を未然に防ぐ有効な手段として、コンピュータウイルス対策ソフトウェアがある。情報基盤センターでは、昨年度に引き続きトレンドマイクロ社の製品についてサイトライセンス契約を拡大継続することによって、学内への安価な配布サービスを推し進めた。Mac OS (Intel 版 Mac) 対応として Sophos Anti-Virus for Mac、及び多言語対応 (基本的には留学生向け)として、Windows OS の簡体中国語と繁体中国語用の Sophos Anti-Virus for Windows の配布サービスも引き続き行っている。また、2009 年 10 月より、特に USB メモリ等のウイルス対策用として ESET Smart Security 4.0 (Canon IT ソリューションズ)の配布を開始した。

### 1.3 迷惑メール(SPAMメール)対策

4月より全学的な迷惑メール対策の正式運用を開始した。7月末に迷惑メール対策システムのバージョンアップ及びバックアップ機器を含めた構成変更を行った(図1の円グラフ参照)。2009年12月から東京大学職員メールシステムの本格運用が始まり、これについても本システムの設定変更を行い迷惑メール対策を開始した。2010年3月末時点での参加数は56ドメイン、約11750メールアドレスである。

## 2 講習会・研究会開催報告

(東京大学情報ネットワークシステム(UTnet3)の運用管理の2.を参照)

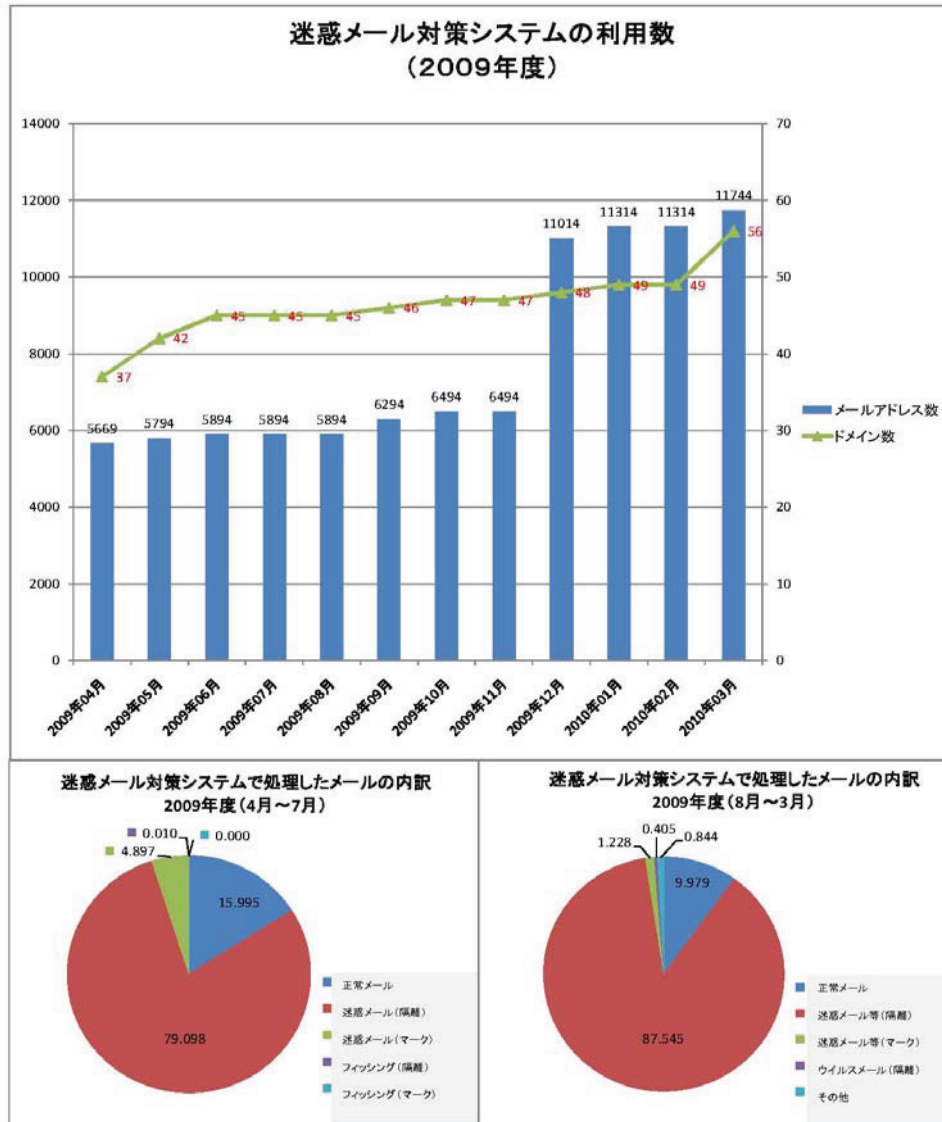
## 3 サービス統計

2009年度のコンピュータウイルス対策ソフトウェアの申請状況は表1の通りである。

表1 ウイルス対策ソフトウェアの申請状況

ソフトウェア名	申請件数	部局数	申請ライセンス数
ウイルスバスターコーポレートエディション	11	7	737
ウイルスバスター2009(日本語版)	634	43	23,023
ウイルスバスター2009(英語版)	100	25	1,012
Sophos Anti-Virus for Windows	24	14	219
Sophos Anti-Virus for Mac OSX	73	25	1,252
Eset Smart Security 4.0	2	2	1,424
Server Protect	51	21	184
InterScan VirusWall	9	8	33
InterScan VirusWall エンタープライズエディション	2	2	9

2009年度の迷惑メール対策システムの運用状況は図1の通りである。



※ 隔離とは、迷惑メールと判定され、サーバに保存されたメール  
 マークとは、迷惑メールと判定されたことを表す特定の文字列をサブジェクトに追加し、利用者に送信されたメール

図1 迷惑メール対策システムの運用状況

# 東京大学情報システム緊急対応チーム(UT-CERT)との連携

## 分散システムセキュリティ支援係

### 1 運用報告

情報基盤センターでは、UTnet3 の基幹ネットワークでセキュリティ対策を実施していることから、ネットワークを介したセキュリティインシデントについて、東京大学情報システム緊急対応チーム UT-CERT(University of Tokyo-Computer Emergency Response Team)から以下の委託業務を受け、UT-CERT と連携協力しながら、全学的なセキュリティ対策を実施した。

- セキュリティ関連情報の収集と学内への注意喚起等  
メーカーやセキュリティ情報サイト等から、セキュリティの脆弱性やウイルスに係わる情報を収集し、情報基盤センターおよび UT-CERT の web ページへの掲載や部局 CERT へのメールによって、セキュリティ対応に関し注意喚起を行った。
- 学内と学外との通信におけるネットワークを介したインシデントの監視  
IDS や異常トラフィックシステムで学内と学外との通信を監視し、セキュリティインシデントの検知を行った。
- セキュリティインシデントの部局 CERT への連絡  
学内と学外との通信の監視や UT-CERT への学内外からの連絡等によって判明したセキュリティインシデントについて、該当する部局 CERT 担当者へメールで対処を依頼した。また、インシデントレポートシステムでの報告を求めた。
- P2P 通信情報の部局 CERT への連絡  
P2P の通信の監視や通信情報を、該当する部局 CERT 担当者へメールで連絡を行った。また、部局 CERT からの P2P 通信に対する詳細情報提供等の依頼に関する対応を行った。
- インシデントレポートシステムの運用管理  
部局 CERT から、発生したセキュリティインシデントの内容や対処等を定型的な書式で報告できるように、インシデントレポートシステムを開発し、運用管理した。
- インシデントレポートの集計  
部局 CERT から報告されたインシデントレポートについて、毎月集計し UT-CERT のホームページ (<http://park.itc.u-tokyo.ac.jp/ut-cert/>) に掲載した。
- セキュリティインシデントの当該 IP アドレスのネットワーク遮断  
部局 CERT に連絡したセキュリティインシデントのうち、発見時点から事象が連続している場合は、緊急措置として、当該 IP アドレスのホストによる通信を UTnet 機器でネットワーク遮断した。
- セキュリティインシデントの部局 CERT の対策への連携協力  
部局 CERT に連絡したセキュリティインシデントのうち、具体的な対処方法等について、部局 CERT から UT-CERT に協力の依頼があった場合、UT-CERT 及び部局 CERT と連携協力して対応した。

### 2 講習会・研究会開催報告

(東京大学情報ネットワークシステム(UTnet3)の運用管理の 2. を参照)



### 3 サービス統計

2009年度のセキュリティ対策の中で扱ったインシデント件数は図1のとおりである。

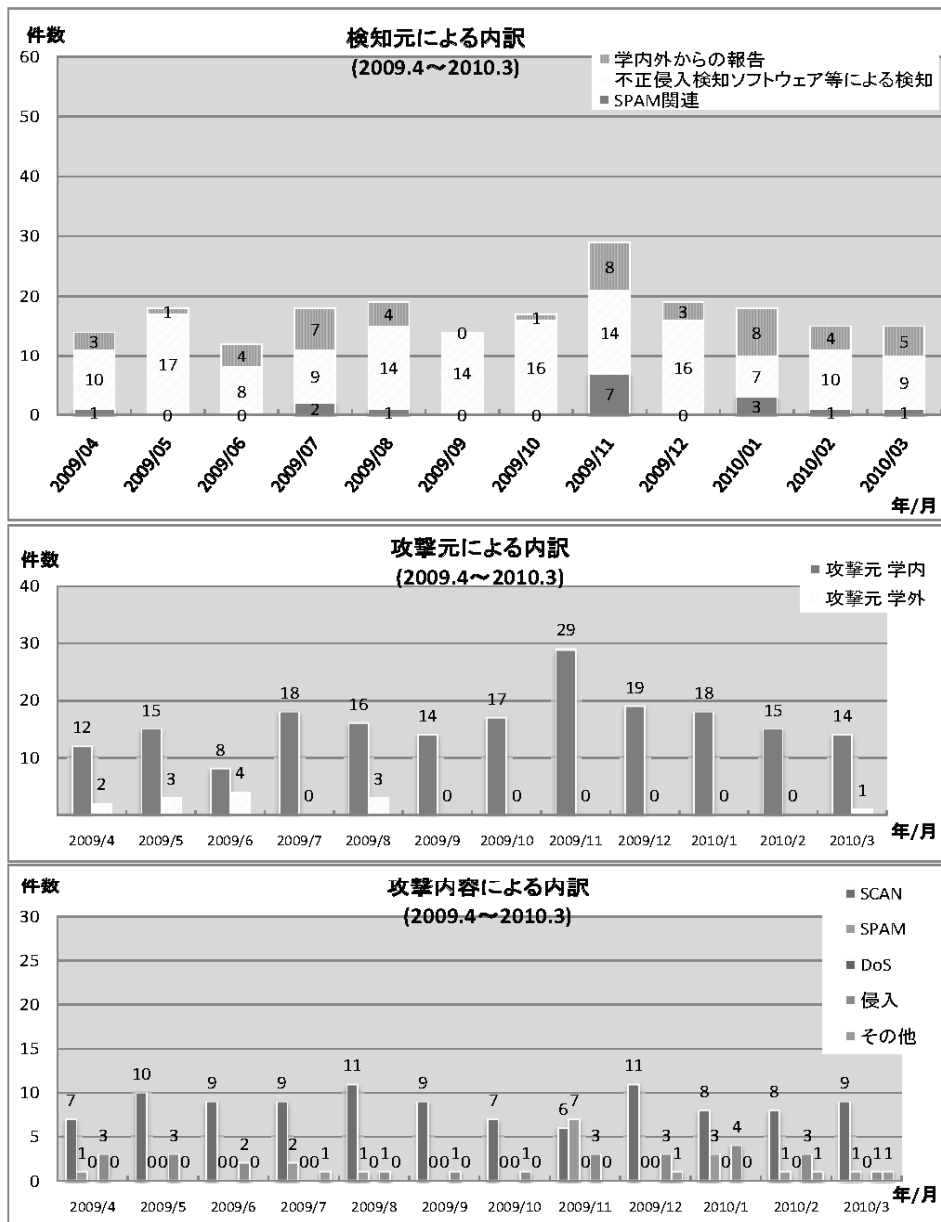


図 1 セキュリティインシデント件数

# 学内ソフトウェアライセンス

## 分散システムセキュリティ支援係

### 1 運用報告

学内での研究教育の円滑な推進を達成することを目的として、利用が多いソフトウェアのうち、可能な範囲でサイトライセンス契約を提供業者との間で結び、安価でかつ容易に利用できるよう図っており、具体的には、以下の各種ソフトウェアの配布サービスとライセンスの管理を行った。表1にソフトウェアライセンスの利用状況を示す。

(1)ウィルス対策ソフトウェア

Windows パソコン、ファイルサーバ、メールサーバ等のコンピュータウィルス対策ソフトウェア。

(2)Pro/ENGINEER

3次元のCADのソフトウェア。

(3)JMP

統計解析ソフトウェア。

(4)SAS9

統計解析ソフトウェア。

(5)MATHEMATICA

数値計算や数式処理用のソフトウェア。

(6)ChemOffice

化学・バイオ分野で必要とされる様々なツールを1つにまとめた統合化学ソフトウェア。

(7)LabVIEW

計測・制御ハードウェアとの通信や、データの解析、結果の共有、システムの分散化したデータを処理するためのグラフィカルプログラミングソフトウェア。

(8)Buzzsaw

設計レビューと建設プロセスを自動化・標準化しながら、プロジェクトに関連するドキュメントや設計情報を一元管理することができるソフトウェア。

(9)Adobe CLP

Adobe社のAcrobat、Photoshop等の製品が安価に購入できるライセンス。2010年1月にAdobe CLPの契約更新を行った。

表1 ソフトウェアライセンスの利用状況

ソフトウェア名	申請件数	部局数
Pro/ENGINEER	19	10
JMP	21	13
SAS9	57	14
MATHEMATICA	129	17
ChemOffice	21	11
LabVIEW	56	11



## スーパーコンピューティング部門



ベクトル並列型スーパーコンピュータ HITACHI SR11000モデルJ2



スーパーコンピュータ HITACHI HA8000クラスシステム  
(T2Kオープンスパコン (東大版))



# スーパーコンピューティング部門

## 概要

部門長 中島 研吾

主査 平野 光敏

### 1 スーパーコンピュータシステムの運用

スーパーコンピューティング部門では、学術研究および教育に供することを目的として、全国の大学・研究機関等に在籍する大学教員、大学院学生、および卒業研究や授業を目的とした学生に対して、スーパーコンピュータシステムを用いた高度かつ大規模な計算サービスを提供している。

当部門では Hitachi SR11000/J2 (128 ノード, 2,048 コア, ピーク性能 18.84TFLOPS), Hitachi HA8000 クラスタシステム (T2K オープンスパコン(東大)) (952 ノード, 15,232 コア, ピーク性能 140.1TFLOPS) の 2 システムを有している。2009 年度末の稼働率はともに 90%に達しており、学内外の幅広い利用者に研究・教育のために利用されている。

2005 年 3 月より稼働を開始した SR11000/J2 は 2010 年度末にリプレース時期を迎える予定であり、次期機種の見込みが始められている。

2008 年 6 月より稼働した HA8000 クラスタシステムは、2009 年度はファイルシステムの性能改善のため、新ファイルシステムの導入、ネットワーク増強を実施した。増強したネットワークの負荷試験をかねて Linpack の再測定を実施した。前回(2008 年 5 月)は、768 ノードを利用して 82.98TFLOPS であったが、今回は全 952 ノードのうち 944 ノードを使用して 101.74TFLOPS (ピーク性能の 73.2%) を達成し、2009 年 11 月の TOP500 リストで世界 45 位(国内 3 位)となった。

利用者拡大、サービスの向上を目的として以下のような様々な試みを実施している:

- SR11000/J2
  - 先端的大規模計算利用サービス(\*)
  - 広報活動(\*)
  - 利用の手引きの整備
  - 成果登録システムの整備
  - 利用に関するアンケート
- HA8000 クラスタシステム
  - 研究プロジェクトの推進(\*)
  - 先端的大規模計算利用サービス(\*)
  - 企業ユーザー利用支援(\*)
  - 講習会(\*)
  - 広報活動(\*)
  - ファイルシステムの増強(\*)
  - 利用の手引きの整備
  - 成果登録システムの整備

(\*)を付した各項目については別項にて紹介する。



## 2 研究プロジェクトの推進

研究プロジェクトとしては、「2009 年度 T2K オープンスパコン(東大)共同研究プロジェクト」、「スーパーコンピュータ若手利用者推薦(試行)」を実施した。

「2009 年度 T2K オープンスパコン(東大)共同研究プロジェクト」は HA8000 クラスタシステムの利用環境の改善を目的として、2008 年度に引き続き 2009 年 4 月～2010 年 3 月の間実施された。本共同研究プロジェクトでは「64 ノード(1,024 コア)」程度を使用する大規模計算を大量に行う研究を対象とする。この共同研究プロジェクトで採択された研究グループは、様々なシミュレーションのアルゴリズムの開発、プログラムの高速化に関する研究を本センターのスタッフと共同で実施する。研究成果は「T2K オープンスパコン(東大)」上でのその成果をライブラリ、HPC ミドルウェア等のアプリケーション開発環境整備にフィードバックすることにより、利用環境の向上に資することを最終的な目標とする。2009 年度は 8 件の応募があり、7 件が採用された。

「スーパーコンピュータ若手利用者推薦(試行)」は、概ね 35 歳以下の若手研究者(学生を含む)を対象としており、採択された課題の計算機利用負担金(半年分)をセンターが負担する。年 2 回公募し、年間で 4 件程度の優れた研究提案を採択する。継続申請と再審査の上で、最大で 1 年間の無料利用ができる。2009 年度は、前期・後期各 4 件、合わせて 8 件の課題を採択した。

## 3 先端的大規模計算利用サービス

文部科学省が 2009 年度から開始した補助金事業「研究開発施設共用等促進費補助金(先端研究施設共用促進事業)」から支援を受け、全国共同利用施設である 7 大学(北海道大学、東北大学、東京大学、名古屋大学、京都大学、大阪大学、九州大学)が連携して、スーパーコンピュータシステム及び応用ソフトウェア(シミュレーションソフトウェア等)を社会貢献の一環として民間企業へ提供する事業を実施した。2009 年度は第 1 期に新規利用拡大 1 件の応募があり、採択された。第 2 期にトライアルユース 1 件の応募があり、採択された。また、2009 年 7 月に「第 3 回先端的大規模計算シミュレーションプログラム利用シンポジウム」を開催した。

## 4 企業ユーザー利用支援

2008 年度から、大規模高性能並列計算によるイノベーションと科学技術の発展に寄与することを目的として、HA8000 システムの資源のうち最大 10%(上記先端研究施設共用促進事業利用を含む)までを企業利用に有償で提供することとし、利用規程の変更を含む制度の改定を実施し、2008 年 10 月より企業利用支援の制度を開始した。2009 年度は、課題の公募を 3 回実施し、3 社の応募があり、3 件が採択された。

## 5 講習会・広報活動

基礎的な並列プログラミング教育を目的とした、国内に例を見ないユニークな取り組みとして「お試しアカウント付きスパコン利用講習会」がある。2009 年度は同講習会を 3 回実施したほか、2008 年度に引き続いて、HA8000 クラスタシステムを使用した「科学技術計算のためのマルチコアプログラミング入門」を開催した。

2009 年度は広報誌「スーパーコンピューティング」を 6 回発行した他、特集号として、「2008 年度公募型プロジェクト成果報告書(2010 年 2 月)」を 1 回発行した。

## 6 ファイルシステムの増強

HA8000 クラスタシステム導入当初から運用している HSFS(Hitachi Striping File System)ファイルシステムは、大規模ファイルの入出力に特化して最適化されていることから、小規模の入出力は苦手

としている。それは、HA8000 クラスタシステム利用者全員に実施したファイルシステムに関するアンケート結果(スーパーコンピューティングニュース Vol.11 No2,2009.3 を参照)にも如実に現れている。そこで、以下の増強を行った。

- NFS 領域の提供
- ノード、ファイルサーバ間ネットワークの増強
- Lustre ファイルシステムの導入

## 7 参考文献

### 報道関連

[報道 1] 朝日小学生新聞 2009 年 6 月 2 日 1 面:次世代スーパーコンピューター、2012 年に世界最高性能をめざす.

[報道 2] 日経産業新聞 2009 年 12 月 2 日 10 面:次世代スパコン計画継続を要望、東大研究者ら.

# スーパーコンピューティング業務

## システム管理係

## システム運用係

### 1 2009 年度のシステム整備状況

本部門で提供しているスーパーコンピューターシステムのサービスについて、2009 年度中に行ったシステム整備状況を以下に述べる。

SR11000/J2 については、従来からのサービスを継続して行った。HA8000 クラスタシステムについては、ファイルシステムの見直しを行うなどの利用者からの要望を踏まえた計算機環境の整備を行った。

また、大規模 (超並列) ジョブ実行サービスや教育利用への支援など、従来から継続して行ってきたサービスに加え、利用支援のため講習会や研究会を実施した。

#### 1.1 2009 年度のサービス変更点

2009 年 4 月以降に行った、主なサービス変更点は以下の通り。

- ・HA8000 クラスタシステムのファイルシステム見直し
  - ・高レスポンス領域の新設
  - ・大容量 NFS 領域の新設
  - ・並列ファイルシステムの導入 (Lustre ファイルシステムの導入)
- ・並列ファイルシステム導入にあたり、HA8000 クラスタシステム内ネットワーク構成変更
- ・HA8000 クラスタシステム向け並列処理対応可視化システム AVS/Express PST の導入
- ・SR11000/J2 の /batch 拡張

#### 1.2 教育利用

前年度に引き続き、教育利用のためのシステム提供 (試行、無料) を行い、並列プログラミング教育として 9 件の利用があった。今年度からは、HA8000 クラスタシステム上で行っている。

- ・東京大学工学部・工学系研究科 (前期)「スパコンプログラミング(1),(I)」
- ・東京大学工学部・工学系研究科 (後期)「スパコンプログラミング(1),(I)」
- ・東京大学工学系研究科「実践的シミュレーションソフトウェア開発演習」
- ・東京大学理学部「計算機特別演習」
- ・東京大学理学系研究科「並列計算プログラミング・先端計算機演習」
- ・東京大学情報理工学系研究科「実践コンピュータシステムⅡ」
- ・東京大学情報理工学系研究科「コンピュータ科学特別講義Ⅱ」
- ・東京大学教養学部 全学ゼミ「スパコンプログラミング研究ゼミ」
- ・電気通信大学電気通信学研究科「並列分散計算法特論」

## 2 2010 年度のシステム整備計画

利用者からの要望が大きかった、大規模ジョブ実行サービスを拡大し、月 1 回 512 Node 実行サービスを開始する予定である。また 2010 年度以降に、SR11000/J2 システムの更新が予定されていることから、仕様の策定等を実施する予定である。

## 3 専門委員会報告

本センター運営委員会のもとにスーパーコンピューティング専門委員会を設置し、全国共同利用スーパーコンピューターシステムおよびその運用に関する事項について、企画、立案及び審議を行っている。2009 年度に開催した同専門委員会の主な議事内容は以下のとおりである。

### 第 11 回スーパーコンピューティング専門委員会

日時：2009 年 7 月 21 日 (火) 13:00 ～

議題：・次期スーパーコンピュータ導入について

- ・HA8000 クラスタシステム 512 ノードサービスについて
- ・お試しアカウント付き講習会 (試行) への企業所属者の参加許可について
- ・東京大学情報基盤センタースーパーコンピューターシステム利用規程別表の改正について
- ・共同利用・共同研究拠点認定について
- ・HA8000 クラスタシステム導入と当部門の活動について
- ・T2K オープンスパコン (東大) 共同研究プロジェクトについて
- ・スーパーコンピュータ若手利用者推薦 (試行) 報告
- ・試行期間限定 T2K オープンスパコン (東大) HPC 特別プロジェクト
- ・スーパーコンピューターシステム教育利用 (試行) 報告
- ・日本応用数学会「2008 秋の学校」報告
- ・お試しアカウント付き並列プログラミング講習会 (試行) 報告
- ・先進スーパーコンピューティング環境研究会 (ASE 研究会) 報告
- ・クラスタシステム上での並列プログラミングコンテスト開催報告
- ・先端研究施設共用イノベーション創出事業／研究開発施設共用等促進費補助金 (先端研究施設共用促進事業) について
- ・HA8000 クラスタシステムの企業利用支援について
- ・HA8000 クラスタシステムのファイルシステム改善について
- ・平成 20 年度スーパーコンピューティング部門決算報告
- ・プログラム相談状況
- ・システム稼動状況報告
- ・その他

第 12 回スーパーコンピューティング専門委員会

日時：2010 年 2 月 15 日 (月) 10:00 ～

議題：・次期スーパーコンピュータシステムの調達方針について

- ・お試しアカウント付き講習会 (試行) への企業所属者への参加許可について
- ・HA8000 クラスタシステムにおける 512 ノードサービス実施について
- ・学際大規模情報基盤共同利用・共同研究拠点公募共同研究 (試行) 採択課題について
- ・次期スーパーコンピュータシステムの柏地区設置について
- ・研究開発施設共用等促進費補助金 (先端研究施設共用促進事業) について
- ・HA8000 クラスタシステムの企業利用支援について
- ・HA8000 クラスタシステムのファイルシステムについて
- ・HA8000 クラスタシステム Linpack 性能更新について
- ・システム稼動状況報告
- ・次世代スーパーコンピュータの状況
- ・その他

## 4 システム利用状況

2009 年度におけるジョブ統計は表 1、表 2 のとおりである。スーパーコンピューターの利用状況については、SR11000/J2 では、ノード稼動率が 70% を越える月が 10 ヶ月になるなど、今年度も多数の利用者に利用された。また、HA8000 クラスタシステムにおいてもノード稼動率が 70% を超える月が 9 ヶ月となるなど利用が増加している。

### 4.1 SR11000/J2 (AIX 5L) のジョブ処理状況

表 1. ジョブ処理状況

年月	登録者数	実利用者数	TSS 件数	NQS 件数	CPU 時間		ノード稼動率 (%)	/batch ファイル 使用量 (MB)
					sr11000-s (時間)	NQS (時間)		
0904	435	170	5,359	5,188	142.75	836,272.39	60.4	16,902,546
0905	444	164	4,019	4,069	98.22	731,565.34	65.4	17,440,916
0906	450	177	6,137	6,176	256.00	997,624.49	74.6	17,774,848
0907	455	169	6,602	5,833	255.62	1,048,315.38	85.7	18,935,323
0908	455	161	3,754	4,399	116.89	715,326.88	81.2	19,495,755
0909	470	159	5,282	6,216	205.25	1,009,711.69	82.9	19,696,566
0910	488	176	5,869	5,873	187.33	1,100,864.09	87.5	20,599,707
0911	479	179	6,463	4,976	239.07	1,147,867.52	88.9	20,555,041
0912	480	179	6,437	5,403	253.46	1,133,115.71	86.6	20,768,456
1001	483	177	7,001	7,077	553.46	1,080,649.52	83.1	14,192,044
1002	485	168	4,899	4,943	166.66	938,286.66	74.6	12,458,855
1003	493	168	4,546	8,239	310.70	999,764.77	74.8	11,047,233
合計	506	300	66,368	68,392	2,785.42	11,739,364.45		

(注) ノード稼動率とは 1 ノードが 100% 稼動したと仮定した場合のサービスノードに対する稼動率である。

## 4.2 HA8000 クラスシステムのジョブ処理状況

表 2. ジョブ処理状況

年月	登録者数	実利用者数	TSS 件数	NQS 件数	CPU 時間		ノード稼働率 (%)	/home ファイル 使用量 (MB)
					ログイン (時間)	NQS (時間)		
0904	487	223	7,601	19,372	1,866.10	3,799,181.42	64.8	66,347,273
0905	514	196	6,116	15,248	222.97	3,263,086.22	66.1	65,400,700
0906	536	195	8,292	27,167	272.60	4,612,945.74	70.3	68,179,392
0907	554	228	7,883	21,255	295.87	4,667,510.38	71.0	70,301,434
0908	600	185	2,710	8,928	82.87	1,754,884.58	69.0	70,703,556
0909	690	282	8,793	30,031	768.29	4,134,908.04	73.3	83,818,170
0910	794	252	7,889	31,469	3,686.13	4,015,565.26	76.6	88,663,879
0911	815	228	8,088	23,356	2,726.02	4,487,161.27	83.5	89,554,967
0912	846	253	9,016	19,524	546.81	5,155,018.75	85.5	95,627,546
1001	858	234	9,111	31,733	283.01	5,765,520.28	89.5	98,313,098
1002	864	239	6,891	21,657	339.53	5,156,265.85	88.7	111,090,619
1003	910	243	6,465	22,796	380.33	5,117,848.62	88.1	111,006,933
合計	970	617	88,855	272,536	11,470.53	51,929,896.41		

(注) ノード稼働率とは 1 ノードが 100% 稼働したと仮定した場合のサービスノードに対する稼働率である。

## 5 プログラム相談と刊行物

### 5.1 プログラム相談

サービスの一環として、利用者からのプログラム相談を受け付けているが、オープンフロアサービスを行っていないため、プログラム相談員を常設せず、事前予約により本センター教職員の面談による対応を行っている。また、以前と同様に E-Mail による相談も受け付けている。

表 3. 相談状況

	対面相談	E-Mail 相談	合計
4 月	0	37	37
5 月	0	25	25
6 月	1	20	21
7 月	0	23	23
8 月	0	6	6
9 月	0	16	16
10 月	0	21	21
11 月	0	16	16
12 月	0	13	13
1 月	0	7	7
2 月	0	13	13
3 月	2	7	9
小計	3	204	207



## 5.2 刊行物

2009 年度に刊行したものは以下のとおりである。

- ・ スーパーコンピューティングニュース Vol.11 No.3 ～ Vol.12 No.2 (2009 年度)
- ・ スーパーコンピューティングニュース Vol.12 No. Special Issue 1 2010.2  
「平成 20 年度公募型プロジェクト報告」

「スーパーコンピューティングニュース」は、スーパーコンピューティング部門の教員、職員が部門連絡会議等で討議のうえ掲載事項・内容を検討し、スーパーコンピューターの利用者に対して利用に関する適切な情報提供を行うとともに、センターへの要望・提案を受け付けている。スーパーコンピューター利用者へよりわかりやすく情報伝達を行うとともに、コンピューター利用サービスの改善に役立っている。さらに同じ内容を本センターの Web ページにも掲載している。2009 年度は昨年度掲載した「T2K オープンスパコン (東大) チューニング連載講座」の番外編、「これからの並列計算のための GPGPU 連載講座」など執筆している。また、Vol. 12 No.1 (2010 年 1 月) 号からは表紙を改装、掲載内容についても一層利用者の皆様に役立つものとするよう推進していく。

また従来から発行しているものには、以下のものがある。

- ・ HA8000 クラスタシステム 利用の手引き 第 2 版 (2009.3)
- ・ 東大センターにおけるスーパーコンピューター利用入門 - SR11000 編 - 第 4 版 (2009.3)
- ・ ベクトル並列型スーパーコンピューター SR11000 チューニングガイド (2005.5)

## 6 職員による研究開発活動

業務系の職員は、日々利用者サービスの向上やシステム運用の効率化・高度化を図りながら業務に携わっており、その成果を発表することで更なる動機付けに期待される。このような成果の発表の場として、全国共同利用情報基盤センター長会議のもとに研究開発連合発表会講演会を毎年開催している。

### 6.1 第 31 回全国共同利用情報基盤センター 研究開発連合発表講演会

国立情報学研究所の当番で、第 31 回全国共同利用情報基盤センター研究開発連合発表講演会が 2009 年 11 月 6 日 (金) に開催され、本センターからも職員が参加し発表を行った。

本講演会は日々急激に進展するスーパーコンピューターを中心とした計算機、ソフトウェア及びネットワークに関する最新の技術について知識・サービスの発表・討議を行い、情報基盤センター群相互の情報交換を図るものである。

＜本センターの研究開発論文＞  
(東京大学情報基盤センター)

- ・「HA8000 クラスタシステムの運用と整備について」 - 宮寄洋 (発表者) -

HA8000 クラスタシステムで行っている各種サービスの概要、サービス向上のための取り組みなど、導入当初から行っている取り組みを報告したものである。

詳しくは、「研究開発論文集 No.31 平成 21 年 11 月全国共同利用情報基盤センター」(P41 ～ P46)をご覧願いたい。

# 講習会

研究支援係

システム管理係

システム運用係

## 1 概要

2009年度は表1に示すような並列計算プログラミングに関連した講習会を開催した。スパコン(HA8000 クラスタシステム)の臨時アカウントを無料で発行しての講習会は国内でも他に例が無い先駆的な取り組みであり、学習効果の促進とともに利用者の拡大にも貢献している。受講者はノートPCを持参してもらい、無線LAN経由でスパコンにログインする。受講者が復習できるように、アカウントの有効期限は、講習会の日から1週間～2週間程度としている。受講者は、これまでは、スパコンの利用資格者(学生、教育機関・研究機関教職員)に限定してきたが、2009年10月より、社会貢献の一環として企業ユーザーもスパコン利用が可能となっており、企業ユーザーも講習会を受講できるようになった。

表1:2009年度に開催した講習会

名称	開催日	利用計算機	申込者数	備考
第5回お試しアカウント付き並列プログラミング講習会(試行)	2009年7月1日～2日	HA8000 クラスタ システム	26人	情報基盤センター主催
第6回お試しアカウント付き並列プログラミング講習会(試行)	2009年9月7日～8日		12人	
第7回お試しアカウント付き並列プログラミング講習会(試行):科学技術計算のためのマルチコアプログラミング入門:Hybrid 並列プログラミングモデルへの道	2009年9月14日～15日		38人	情報基盤センター主催
第8回お試しアカウント付き並列プログラミング講習会(試行)	2009年3月17日		17人	情報基盤センター主催
CCS HPC サマーセミナー2009	2009年7月23日～7月24日	—	7人	筑波大学主催
京都大学 並列プログラミング入門講習会	2009年9月3日～4日	—	5人	京都大学主催

## 2 お試しアカウント付き並列プログラミング講習会(試行)

2009年度は本センター主催による初心者向け並列プログラミングを表1に示すように3回実施した。以下に概要を示す。詳細はホームページを参照されたい<sup>1</sup>。

<sup>1</sup> <http://www.cc.u-tokyo.ac.jp/publication/kosyu/>

## 2.1 第5回講習会

HA8000 クラスタシステムを使用した、初心者向け並列プログラミング講習会を実施した。並列プログラミングの基礎、MPI、並列アプリケーション入門について講義・実習を実施した。

表2：第5回お試しアカウント付き並列プログラミング講習会（試行） スケジュール

開催日	時間	内容	講師
2009年 7月1日（水）	1000～1200	ノートパソコンの設定、テストプログラムの実行	片桐孝洋
	1300～1500	プログラミングの基礎（分割コンパイル）	鴨志田良和
	1500～1700	並列プログラミングの基本	片桐孝洋
7月2日（木）	1000～1200	MPIプログラミング実習	片桐孝洋
	1300～1700	MPIによる並列アプリケーション開発入門	中島研吾

## 2.2 第6回講習会

HA8000 クラスタシステムを使用した、初心者向け並列プログラミング講習会を実施した。並列プログラミングの基礎、MPI、並列アプリケーション入門について講義・実習を実施した。

表3：第6回お試しアカウント付き並列プログラミング講習会（試行） スケジュール

開催日	時間	内容	講師
2009年 9月7日（月）	1000～1200	ノートパソコンの設定、テストプログラムの実行	片桐孝洋
	1300～1500	プログラミングの基礎（分割コンパイル）	鴨志田良和
	1500～1700	並列プログラミングの基本	片桐孝洋
9月8日（火）	1000～1200	MPIプログラミング実習	片桐孝洋
	1300～1700	MPIによる並列アプリケーション開発入門	中島研吾

## 2.3 第7回講習会

近年マイクロプロセッサのマルチコア化が進み、様々なプログラミングモデルが提案されている。中でも OpenMP は指示行（ディレクティブ）を挿入するだけで手軽に「並列化」ができるため、広く使用されており、様々な解説書も出版されている。メモリへの書き込みと参照が同時に起こるような「データ依存性（data dependency）」が生じる場合に並列化を実施するには、適切なデータの並べ替えを施す必要があるが、このような対策は OpenMP 向けの解説書でも詳しく取り上げられることは余り無い。本講習会は、このような事情を考慮し、本センターと日本応用数理学会の共催で実施したものであり、「有限体積法から導かれる疎行列を対象とした ICCG 法」を題材として、科学技術計算のためのマルチコアプログラミングにおいて重要なデータ配置、reordering などのアルゴリズムについての講習、実習を実施した。

今回は、2008年6月より稼動を開始した、「T2K オープンスパコン（東大）（HA8000 クラスタシステム）」等最近のマルチコアプロセッサへの適用を念頭においた内容を実施した。本講習会のスケジュールを表4に示す。詳細はホームページを参照されたい。9月14日（月）の午前・午後、9月14日（火）午前の正味10時間弱の講義であるが、その中で目を引くのが、対象となるアプリケーション（有限体積法、ICCG法）に2時間、オーダリングのアルゴリズム、実装について3時間と合計半分程度を占めていることである。それに対して、OpenMP の文法に関わる部分はごく少ない。

表 4：科学技術計算のためのマルチコアプログラミング入門 スケジュール

開催日	時 間	内 容	講 師
2009 年 9 月 14 日 (月)	0930～1000	受付	中島研吾
	1000～1200	ICCG 法によるポアソン方程式ソルバー	
	1300～1400	OpenMP「超」入門	
	1415～1715	オーダリングについて	
	1715～1730	質疑	
9 月 15 日 (火)	0900～1100	並列化実装について	
	1100～1200	最新の研究事例の紹介、質疑、T2K オープンスパコンによる実習	

## 2.4 第 8 回講習会

HA8000 クラスタシステムを使用した、初心者向け並列プログラミング講習会を実施した。並列プログラミングの基礎、MPI について講義・実習を実施した。

表 5：第 8 回お試しアカウント付き並列プログラミング講習会（試行） スケジュール

開催日	時 間	内 容	講 師
2010 年 3 月 17 日 (月)	1000～1200	ノートパソコンの設定、テストプログラムの実行	片桐孝洋
	1300～1500	プログラミングの基礎（分割コンパイル）	鴨志田良和
	1515～1615	並列プログラミングの基本	片桐孝洋
	1530～1830	MPI プログラミング実習	

## 3 T2K 連携セミナー・講習会

筑波大学および京都大学が主催する T2K オープンスパコンに関連するセミナーおよび講習会を、連携の一環として、東京大学情報基盤センターの遠隔講義室を利用し公開講座を行った。

## シンポジウム・研究会

研究支援係

システム管理係

システム運用係

### 1 平成 20 年度公募型プロジェクト報告会「ペタ/エクサスケールコンピューティングへの道 2009」

2008 年度に実施した, 2 つの公募型プロジェクト:

- ・ 2008 年度「T2K オープンスパコン(東大)共同研究プロジェクト」(5 件)
- ・ 2008 年度「スーパーコンピューター若手利用者推薦(試行)」(後期) (4 件)

採択課題の報告会(平成 20 年度公募型プロジェクト報告会「ペタ/エクサスケールコンピューティングへの道 2009」)を 2009 年 5 月 11 日に小柴ホールにて実施した<sup>1</sup>。表 1 にプログラムを示す。51 名の出席があった。

表 1:平成 20 年度公募型プロジェクト報告会「ペタ/エクサスケールコンピューティングへの道 2009」  
(2009 年 5 月 11 日小柴ホール)概要

種 別	時 間	講演タイトル・研究代表者(所属(当時))
共同研究 プロジェクト報告	10:20-11:05	海洋大循環のマルチスケール連結階層モデリング 羽角博康 (東京大学気候システム研究センター)
	11:05-11:50	津波発生伝播の大規模 3 次元シミュレーション 古村孝志(東京大学大学院情報学環総合防災情報研究センター)
	13:00-13:45	超並列計算によるマルチスケール・マルチフィジックス心臓シミュレーション 久田俊明(東京大学大学院新領域創成科学研究科)
	13:45-14:30	電磁流体コードによる惑星磁気圏シミュレーション 深沢圭一郎(情報通信研究機構)
	14:30-15:15	地球ダイナモの新しいシミュレーションコード開発とその応用 陰山聡(神戸大学大学院工学研究科)
若手利用者推薦報告	15:30-16:00	実対称固有値問題に対する多分割の分割統治法の分散メモリ型並列計算機への実装 田村純一(埼玉大学大学院理工学研究科)
	16:00-16:30	大規模データから探る経済現象の数理 大西立顕(東京大学大学院法学政治学研究科)
	16:30-17:00	厳密な理論波形計算を用いた高解像度地球内部構造推定 竹内希(東京大学地震研究所)
	17:00-17:30	数値モデルによる海洋微小スケールプロセスの解明 松村義正(東京大学気候システム研究センター)

<sup>1</sup> <http://www.cc.u-tokyo.ac.jp/publication/sympo/03/>

## 2 T2K シンポジウム

筑波大学、京都大学との共催によるシンポジウム「T2K シンポジウム'09 in Tsukuba」を 2009 年 6 月 18 日に開催した<sup>2</sup>。

## 3 研究会

### 3.1 概要

2009 年度は表 2 に示すように 2 回の先進スーパーコンピューティング環境研究会 (Advanced Supercomputing Environment、ASE) (略称:ASE 研究会)を開催した。

表 2:2009 年度に開催した研究会

名称	開催日	参加者数
第 5 回 ASE 研究会	2009 年 8 月 21 日(金)	23 人
第 6 回 ASE 研究会	2010 年 2 月 12 日(金)	15 人

### 3.2 第 5 回 ASE 研究会

2009 年 8 月 21 日(金)14 時 00 分から 17 時 30 分まで、東京大学情報基盤センター大会議室にて、第 5 回 ASE 研究会を開催した。国内の大学、および企業からの参加者が 23 名あり、活発な議論がなされた。

第 5 回 ASE 研究会の招待講演として、電気通信大学大学院情報システム学研究科から大島聡史 博士を招待した。大島博士は、GPU とスーパーコンピュータを用いた高性能計算に関する講演を行った。また、東京大学情報基盤センターの中島研吾教授による、拡張階層型領域間境界分割に基づく並列前処理手法に関する講演、同センターから片桐孝洋特任准教授による、汎用な自動チューニングインターフェース OpenATLib を利用した疎行列反復解法ライブラリ Xabclib に関する講演を行った。

### 3.3 第 6 回 ASE 研究会

2010 年 2 月 12 日(金)14 時から 17 時まで、東京大学情報基盤センター4 階遠隔講義室において、第 6 回先進スーパーコンピューティング環境研究会 (ASE 研究会、Advanced Supercomputing Environment) が開催された。国内の大学、研究機関、企業から合計 15 名の出席者があり、活発な議論が行われた。

プログラムは表 3 に示すように、招待講演 1 件と情報基盤センタースーパーコンピューティング研究部門に所属する教員による 2 件の講演から構成されている。

招待講演として、ドイツの Hamburg-Harburg 工科大学から Jens-Peter Zemke 博士をお呼びし、線形反復解法において近年注目されている IDR(s)法に関する講演を行った。また、情報基盤センターの伊藤祥司 特任准教授による、反復解法の体系的性能評価に対する品質管理の講演を行った。また同センターの片桐孝洋 特任准教授により、文部科学省 e-サイエンス実現のためのシステム統合・連携ソフトウェアの研究開発、「シームレス高生産・高性能プログラミング環境」の支援のもとに、日立製作所中央研究所および東京大学情報基盤センターの共同研究で開発された、自動チューニング機能付き数値計算ライブラリ Xabclib における新機能である、数値計算ポリシー制御に関する講演を行った。

<sup>2</sup> <http://www.ccs.tsukuba.ac.jp/workshop/t2k-sympo2009/>



表 3 : 第 6 回 ASE 研究会 プログラム

講演者	講演タイトル
Jens-Peter Zemke 博士 (Hamburg-Harburg University of Technology, Germany) (招待講演)	IDR(s) and IDR(s) Eig in Parallel Computing
伊藤祥司 特任准教授 (東大・情報基盤センター)	Quality control approach for systematic performance evaluation of linear solvers
片桐孝洋 (東大・情報基盤センター)	Xabclib: An Iterative Solver with a General Auto-tuning Interface "OpenATLib" --- Development of Numerical Computation Policy Function

# 研究プロジェクト

研究支援係

システム管理係

システム運用係

## 1 2009 年度 T2K オープンスパコン(東大)共同研究プロジェクト

2009年4月～2010年3月の間、2009年度「T2K オープンスパコン(東大)共同研究プロジェクト」を実施した。共同研究プロジェクト期間中は64ノード専用キュー(1,024コア)を無料で利用できる。以下、共同プロジェクトの概要と2008年度採択課題について、簡単に紹介する。

### 1.1 共同研究プロジェクトの対象、目標

本共同研究プロジェクトでは「64ノード(1,024コア)」程度を使用する大規模計算を大量に行う研究を対象とする。この共同研究プロジェクトで採択された研究グループは、様々なシミュレーションのアルゴリズムの開発、プログラムの高速化に関する研究を本センターのスタッフと共同で実施する。研究成果は「T2K オープンスパコン(東大)」上でのその成果をライブラリ、HPCミドルウェア等のアプリケーション開発環境整備にフィードバックすることにより、利用環境の向上に資することを最終的な目標とする。

### 1.2 応募資格

申込者(代表者)は「国内の大学に所属する研究者、公共機関に所属する研究者」に限定する。研究グループのメンバーは「T2K」オープンスパコンの利用有資格者、すなわち：

- 国内の大学に所属する学生・研究者
- 公共機関に所属する研究者
- 企業に所属する研究者、技術者

でなければならない。

### 1.3 共同研究プロジェクトの要件

研究期間は2009年4月～2010年3月の間とする。本センターのスタッフとプログラムの並列化、高速化に関する共同研究を実施する。原則としてプログラムのソースコードを本センターの担当スタッフに対して提供する。本センターとして協力可能な分野は下記の通りである。

- スカラープロセッサ向けチューニング
- 線形ソルバー(密行列、疎行列)の高速化、チューニング
- 反復法前処理手法
- ハイブリッド並列化
- 並列適応格子法、動的負荷分散
- 細粒度タスク並列化、並列分散プログラミング言語
- ファイル転送効率化
- ユーザー所有クラスタや他のスパコンとの連携

共同研究期間中は、64ノード専用キューを無料で利用できる(但し、本共同研究プロジェクトで採択された他グループと共同で使用する)。

共同研究プロジェクトで得られた成果は当センター(スーパーコンピューティング部門)広報誌「スーパーコンピューティングニュース」に投稿していただく。また、当センターの主催するセミナー、ワークショップ等でご発表いただく場合がある。年度末に所定の様式に従って共同研究プロジェクト実績報告書(研究成果の概要、外部発表リスト等)を提出していただく。

研究結果の公表の際には、その論文、報告等に「T2K オープンスパコン(東大)」を利用したことを明記するとともに、当該論文、報告等の別刷又は写し 1 部を提出していただく。なお、本センター教職員の貢献度に応じて co-authorship を求める場合がある。性能データの誤解等に基づく誤謬防止のため投稿前に本センター教職員の助言を受けることが望まれる。

#### 1.4 2009 年度募集と採択課題について

2008 年度については 2009 年 3 月 10 日に募集を開始し、同 4 月 10 日に締め切った。合計 8 件の応募があったが、本センタースーパーコンピューティング研究部門教員による厳正なる審査の結果、以下に示す 7 課題を採択することにした。採否にあたっては、以下の点に特に留意した:

- 研究テーマの有用性、先進性
- T2K オープンスパコン(東大)を利用することの必要性
- 大規模計算に関する経験
- 共同研究プロジェクトとしての有用性
- T2K オープンスパコン(東大)利用環境向上への貢献

## 採択課題①：継続

課題名	地球ダイナモの新しいシミュレーションコード開発とその応用
代表者	陰山 聡
所 属	神戸大学大学院 工学研究科
分 野	地球電磁気学, 地球内部ダイナミクス
<p>地球ダイナモに関する計算機シミュレーション研究において、これまで我々は理想化されたモデルに基づく高解像度計算に注力してきた。今後は、地球外核の現実環境を可能な限り反映させた複合的なモデルに基づくシミュレーションの重要性が増すであろう。多数のスカラープロセッサが結合されたスーパーコンピュータはそのような複合シミュレーションに適している。本研究の目的は、T2K オープンスパコン（東大）に向けて最適化した地球ダイナモシミュレーションの新しいコードを開発することである。まずは地球シミュレータ向けに開発した従来の我々のコードを T2K（東大）向けに最適化し、そこで得られた知見を生かして、基本アルゴリズムから検討し直して新たなコードを開発する。単純モデルによる高解像度・短時間計算という従来のアプローチとは相補的な、複合モデルによる中解像度・長時間計算により、地球双極子磁場の成因とその逆転機構に関する新発見を目指す。</p>	

## 採択課題②：新規

課題名	市街地における風・温熱・光・音環境総合数値予測データベースの開発
代表者	坂本 雄三
所 属	東京大学大学院 工学系研究科建築学専攻
分 野	建築学, 都市工学, 数値流体力学
<p>本研究は市街地における街区スケールの風・温熱・光・音環境の数値予測を行ない、種々の省エネルギー対策、騒音対策といった改善策の効果に対するデータベース整備および Web インターフェイスを通じた公開を図るものである。最初に、数値地図データと航空機レーザスキャナデータから、建物の壁面や屋根の幾何形状を自動的かつ精度良く再現する。次に、幾何形状データから 2km 四方の街区を単位として、超並列計算に対応させたロード・バランス機能付き並列メッシュを用いて、風・温熱・音環境予測用格子を自動生成する。さらに、スカラー型並列計算機用にチューニングした音響解析機能付き対流・放射連成 CFD 解析コードおよび光環境予測コードを使用して、単位街区の風・温熱・光・音環境の数値解析を行なう。省エネルギー対策、騒音対策シナリオを数種類設定し、無対策の状態と対策ケースについてデータを蓄積し、数値解析結果を Web 地図インターフェイスを用いて一般に広く公開する。</p>	

採択課題③：継続

課題名	海洋大循環のマルチスケール連結階層モデリング
代表者	羽角 博康
所 属	東京大学 気候システム研究センター
分 野	海洋物理学
<p>海洋循環は、水平 1 km 以下のスケールで生じる対流や混合現象から、水平 100,000 km スケールに及ぶ全地球規模循環までの間で、様々なスケールを持つ現象のスケール間相互作用の結果として形成されている。海洋循環の形成や変動のシミュレーションは、数十年から場合によっては数千年という時間スケールを対象としており、全海洋を一様に 1 km 程度の解像度で表現しつつそれを実施することは 10 年程度の将来を見越しても不可能である。大循環の形成という視点から見た場合、小規模スケール現象が重要となる領域は限定されているため、適切な連結階層化によるシミュレーションが可能であり、また必要とされる。本研究は、ネステッドグリッド手法による連結階層海洋大循環モデルについて、高並列環境におけるチューニングを目的とする。</p>	

採択課題④：継続

課題名	超並列計算によるマルチスケール・マルチフィジックス心臓シミュレーション
代表者	久田 俊明
所 属	東京大学大学院 新領域創成科学研究科
分 野	バイオメカニクス, 非線形力学
<p>当研究チームでは、細胞イオンチャンネルや収縮タンパクの数理モデルから出発し有限要素法でモデル化された心室の収縮、血液の拍出に至る現象を一貫して再現できる国際的にも突出した心臓シミュレータ、<i>UT-Heart</i> の開発を行ってきた。一方、有限要素法に基づく心筋細胞も並行して開発し、両者を組み合わせることでシームレスなマルチスケール・マルチフィジックスシミュレータを開発中であり、最終的には次世代スパコン上での数十万個の細胞を用いたマルチスケールシミュレーションを実施することを目指している。昨年度 7 月の HPC 特別プロジェクトにおいては規模を縮小した心臓モデルと細胞モデルにより、最大 6144 コアを用いたマルチスケールシミュレーションを実施し、非常に良好なスケーラビリティを得た。また、さらなる高精度なシミュレーションを行なうことにより、基礎医学や創薬への貢献が可能であると考えている。本共同研究プロジェクトにおいても、次世代スパコン上での実行を見据えて、各モジュールのさらなる高速化・高機能化を実施する予定である。</p>	

## 採択課題⑤：継続

課題名	電磁流体コードによる惑星磁気圏シミュレーション
代表者	深沢 圭一郎
所属	九州大学 宙空環境研究センター
分野	宇宙プラズマ物理学
<p>宇宙空間はプラズマで満ちており、そのプラズマダイナミクスをシミュレーションにより研究することを目的とする。その宇宙空間において惑星磁気圏のようなグローバルな構造に注目する場合、電磁流体（MHD）近似が成り立ち、MHD 方程式によってその構造はよく表される。現在 MHD シミュレーションは磁気圏観測結果の理解だけでなく、観測に先立つ磁気圏ダイナミクスの予見にも役立っている。昨年度のプロジェクトにおいて、我々のコードをある程度東京大学 T2K スパコンに最適化できた。また、現在よりも 5 倍程度の高精細な空間解像度で計算が可能となった。これにより今まで我々が再現できていないプラズマ乱流構造、磁気圏がお椀型に変形するなどの観測結果を扱える可能性がある。本課題では、MHD コードを用いた惑星磁気圏シミュレーションを行い、特に土星磁気圏に焦点を当て、高解像度長時間発展下における惑星磁気圏の構造を明らかにすることを目的とする。シミュレーションコードは昨年度の結果から高効率の実績のある直行格子 2 次元領域分割を用いて、東京大学 T2K スパコンに更に最適化させていく。</p>	

## 採択課題⑥：継続

課題名	津波発生伝播の大規模 3 次元シミュレーション
代表者	古村 孝志
所属	東京大学大学院 情報学環総合防災情報研究センター
分野	地震学, 津波学
<p>将来発生が危惧される巨大地震による津波被害の軽減を目的に、大規模津波シミュレーションのための大規模計算コードを開発する。既に開発した 3 次元ナビエ・ストークス式の差分法計算に基づく津波計算コード (Saito and Furumura, 2008) を改良し、T2K オープンスパコン (東大) に代表される、スカラー型超並列計算機において高い並列化スケールビリティ (8192 コア以上) と実効性能 (ピーク性能比 8% 程度以上) を引き出すことのできる実用化コードを整備する。これにより、現在のモデルより計算時間を 10 倍以上、計算規模を 8 倍以上大きくした大規模津波計算を可能にする。さらに、デカルト座標系における津波計算を極座標系に改良し、遠地巨大地震による地球規模の津波計算を実現する。本津波計算コードを用いて、近年の 2007 年千島列島地震や 2004 年紀伊半島南東沖など、近年の津波地震の大規模・高精度津波シミュレーションを実施し、沖合ケーブル津波計で記録した津波データとの比較からモデルの有効性と精度を検証する。</p>	



採択課題⑦：新規

課題名	天体活動現象の輻射磁気流体シミュレーション
代表者	松元 亮治
所属	千葉大学大学院 理学研究科
分野	宇宙物理学
<p>さまざまな天体で観測される激しい時間変動，爆発的エネルギー解放，X線・ガンマ線放射，ジェット噴出などを伴う活動的現象を，輻射と磁気流体の相互作用を考慮した3次元シミュレーションによって解明する。このため，我々のグループで作成してきた磁気流体コードを輻射磁気流体コードに拡張し，ブラックホール降着円盤における状態遷移過程等をシミュレートする。まず，Flux Limited Diffusion (FLD) 近似に基づく輻射磁気流体コードと輻射輸送方程式の1次のモーメントまで使用するM1スキームに基づくコードを実装する。東京大学情報基盤センターのスタッフと共同で，マルチコアCPUにおけるスカラー効率の改善，輻射を含む項を陰的に解く場合の並列効率の改善等を実施する。次年度以後には空間3次元，光子方向2次元，振動数1次元の6次元輻射輸送方程式を効率的に解くスキームの実装に取り組み，直接計算によって輻射強度と輻射スペクトルの時間変動をシミュレートできるようにする。</p>	

### 1.5 2009年度報告会・2010年度募集

2009年度採択課題については、2010年5月21日(金)に報告会を実施する<sup>1</sup>。また、2010年度以降については本共同研究プロジェクトは学際大規模情報基盤共同利用・共同研究拠点<sup>2</sup>公募型研究プロジェクトの一部として発展的に解消する。

## 2 スーパーコンピュータ若手利用者推薦(試行)

当センターでは、概ね35歳以下の若手研究者(学生を含む)を対象とした利用者向けの、スーパーコンピュータ若手利用者推薦(試行)(以降、若手推薦(試行))による課題の公募を開始する。スーパーコンピューティング部門の教員により審査の上、採択された課題の計算機利用負担金(半年分)をセンターが負担する。

年2回公募し、年間で4件程度の優れた研究提案を採択する。継続申請と再審査の上で、最大で1年間の無料利用ができる。採択の方には、報告書の提出、研究成果の発表の際に若手推薦(試行)を利用したことの明記、およびセンターが発行する「スーパーコンピューティングニュース」誌の原稿執筆を採択の条件とする。

スーパーコンピュータを利用することで学術的にインパクトがある成果を創出できると期待される点、大規模計算、テーマの重要性、既発表文献の観点から以下の課題を採択した。

<sup>1</sup> <http://www.cc.u-tokyo.ac.jp/publication/sympo/04/>

<sup>2</sup> <http://jhpcn-kyoten.itc.u-tokyo.ac.jp/>

## 2.1 2009年(前期)採択課題

### 採択課題①

課題名	直接メタノール形燃料電池内の流動現象シミュレーション
代表者	高木 洋平
所属	静岡大学 工学部物質工学科

### 採択課題②

課題名	ランダムネス誘起二次相転移の数値的解析
代表者	村瀬 洋介
所属	東京大学大学院工学系研究科 物理工学専攻

### 採択課題③

課題名	高分子流動プロセスにおける混合・混練評価の数値的方法の構築
代表者	名嘉山 祥也
所属	九州大学大学院工学研究院 化学工学部門

### 採択課題④

課題名	大規模データから探る経済現象の数理
代表者	大西 立顕
所属	東京大学大学院法学政治学研究科 総合法政専攻

## 2.2 2009年(後期)採択課題

### 採択課題①

課題名	Study of QCD and effective models in the context of equilibrium and out of equilibrium quantum field theory with application to Heavy Ion Collisions, the Quark Gluon Plasma and Early Universe dynamics
代表者	ROTHKOPF ALEXANDER
所属	東京大学大学院理学系研究科 物理学専攻

### 採択課題②

課題名	量子 i.i.d.状態のシミュレーションとその理論的考察分散メモリ型並列計算機への実装
代表者	坂下 達哉
所属	電気通信大学大学院情報システム学研究科 ネットワークシステム学専攻

### 採択課題③

課題名	実対称固有値問題に対する多分割の分割統治法の分散メモリ型並列計算機への実装
代表者	田村 純一
所属	埼玉大学 大学院理工学研究科

### 採択課題④

課題名	直接メタノール形燃料電池内の流動現象シミュレーション
代表者	高木 洋平
所属	静岡大学 工学部物質工学科

### 3 参考文献

#### 広報誌

- [広報誌 1] スーパーコンピューティングニュース(東京大学情報基盤センター) 11-3, 2009.
- [広報誌 2] スーパーコンピューティングニュース(東京大学情報基盤センター) 11-4, 2009.
- [広報誌 3] スーパーコンピューティングニュース(東京大学情報基盤センター) 11-5, 2009.
- [広報誌 4] スーパーコンピューティングニュース(東京大学情報基盤センター) 11-6, 2009.
- [広報誌 5] スーパーコンピューティングニュース(東京大学情報基盤センター) 12-2, 2010.
- [広報誌 6] スーパーコンピューティングニュース 特集:平成 20 年度公募型プロジェクト報告(東京大学情報基盤センター) 12-S.1, 2010.

文部科学省 研究開発施設共用等促進費補助金(先端研究施設共用促進事業)  
**先端的大規模計算利用サービス**

研究支援係

**1 文部科学省 研究開発施設共用等促進費補助金(先端研究施設共用促進事業) 先端的大規模計算利用サービス**

**1.1 事業の趣旨と概要**

独立行政法人・大学等の研究機関等の保有する先端的な研究開発に係る施設及び設備(以下「先端研究施設」)の中には、広範な分野や多様な研究等に活用されることにより、様々な研究等に大きな成果をもたらす可能性のあるものが多数あるが、利用者支援等の共用のための体制が整備されていないことから十分に活用されていない。「研究開発施設共用等促進費補助金(先端研究施設共用促進事業)」は、それらの先端研究施設の共用を促進し、基礎研究からイノベーション創出に至るまでの科学技術活動全般の高度化を図るとともに国の研究開発投資の効率化を図るために平成 21 年度から文部科学省が開始した事業である。

本事業を通じて、産学官の研究者による戦略的かつ効率的な研究開発や、研究機関・研究分野を越えた横断的な研究開発活動を推進することにより、継続的に産学官の知の融合によるイノベーションを加速していくことを目指している。

全国共同利用施設であるスーパーコンピュータを有する7大学(北海道大学、東北大学、東京大学、名古屋大学、京都大学、大阪大学、九州大学)は、平成 19 年 7 月より平成 21 年 3 月まで文部科学省先端研究施設共用イノベーション創出事業の支援を受け「先端的大規模計算シミュレーションプログラム利用サービス」を通じ、社会貢献の一環として、大学で開発された応用ソフトウェアとスーパーコンピュータの民間企業への提供を実施してきた。

平成 21 年度からは、「先端研究施設共用促進事業」の支援を受け「先端的大規模計算利用サービス」として、より幅広い研究課題に対してスーパーコンピュータの提供を実施している。

本事業の特徴は、各大学が開発した応用ソフトウェア(シミュレーションソフトウェア等)と各大学が持つ国内有数のスーパーコンピュータを、それらの利用支援サービスと一体化して民間企業に提供することにある。このようにサービス、ソフトウェア、ハードウェアの 3 つを同時に支援する試みは、国内外を通じて初めてのものである。この事業の目的は、大学の持つ知と施設によって我が国の経済発展に貢献することである。

- 国立大学法人 東京工業大学において同プログラムからの支援を受けた「『みんなのスパコン』TSUBAME によるペタスケールへの飛翔」プログラムと連携し、民間企業へのスーパーコンピュータ提供事業を行っている。

**1.2 支援内容**

本事業は大きく 2 つの目的から成り立つ。

- 一企業では試すことが困難な大規模高性能並列計算支援  
文部科学省が進める次世代スーパーコンピュータプロジェクトでは、10 ペタフロップス規模のスパコンが実現され、産業界でも期待されている。一方で、企業では PC や PC を集めた PC

クラスタと呼ばれる コンピュータ上で、小規模な計算が行われているのが現状である。大規模並列計算を行ってみようと思っても、費用対効果も見込めないため、大規模計算資源を利用できていない。本支援は、このような企業に対して計算資源を提供する。

- **ASP 事業者支援**

計算シミュレーションによるもの作りの方法論が確立できていない企業の支援を目的とする、ASP 事業者を支援する。ASP 事業者により、今までスパコンを利用した「もの作り」を行ったことのない企業の利用促進が期待される。

募集課題については、「大規模計算利用(有償利用)」課題と「トライアルユース(無償試用)」課題の2つのカテゴリに分類し支援を行う。

平成 21 年度は委託事業から補助金事業への移行期間であり、第1期採択課題は平成 20 年度中の採択となったため、「先端的大規模計算シミュレーションプログラム利用サービス」で採択されていた「戦略分野利用推進」課題と「新規利用拡大」課題の2つのカテゴリについても支援を行った。第2期採択課題については、新しいカテゴリでの募集となり、旧カテゴリは継続課題のみの支援を行った。

### 1.2.1 大規模計算利用(有償利用)

従来の「戦略分野利用推進」と同様に、ライフサイエンス、もの作り技術、情報通信、環境エネルギー、社会基盤の対象課題分野を維持しつつ、他の分野における特に優れた課題にも対応できるようにするため、「大規模計算利用」と改称した。利用成果は従来どおり公開を原則とし、利用期間も従来どおり1年間とするが、継続回数は特に規定していない。大規模計算利用では、計算機利用料の一部が文部科学省の補助金によって負担されるので、通常の有償利用より安価で計算機を利用可能である。

### 1.2.2 トライアルユース(無償試用)

新しいニーズを掘り起こしイノベーション創出につながる利用課題も民間企業から募っている。

分野は特定していない。産業界の潜在的な高性能計算需要を掘り起こし、ソフトウェア、ハードウェアを有償で利用してもらえ下地を作ることが目的である。

### 1.2.3 戦略分野利用推進(無償利用・第2期は継続のみ)

ライフサイエンス、もの作り技術、情報通信、環境エネルギー、社会基盤分野を対象に、大学で開発された先端的シミュレーションソフトウェアとスーパーコンピュータの利用課題を民間企業から募り、企業によるイノベーション創出を支援する。民間企業が有するアプリケーションを利用した課題あるいは新規アプリケーションを開発する課題も募っていく。

- タンパク質機能シミュレータによる創薬開発
- ナノマテリアルシミュレータによる新材料開発
- 音・熱・力の解析シミュレータによるものづくり
- 機器・システムレベルのシミュレータ開発・利用実験による次世代情報通信技術開発
- 地盤と拡散物質・構造物解析シミュレーション技術の開発

### 1.2.4 新規利用拡大(無償利用・第2期は継続のみ)

新しいニーズを掘り起こしイノベーション創出につながる利用課題も民間企業から募っている。

分野は特定していない。産業界の潜在的な高性能計算需要を掘り起こし、大学と ASP (Application Service Provider) 事業との連携によるサービス、ソフトウェア、ハードウェアを有償で利用する下地を作ることが目的である。

### 1.2.5 利用可能なソフトウェア

- 革新的シミュレーションソフトウェアの研究開発 (RSS21) で開発されたソフトウェア群
- タンパク質構造解析ソフトウェア myPresto/AMOSS  
※AMOSSは日本電気株式会社の製品
- 分子動力学法に基づくソフトウェア OCTA
- 全電子第一原理計算ソフトウェア TOMBO
- 数値タービンシミュレーション
- 超臨界流体シミュレータ
- TAS-code (Tohoku University Aerodynamic Simulation Code)
- 時間領域差分法に基づく大規模電磁界解析ソフトウェア JetFDTD
- 可視化システム KVS
- ◇ 上記以外で企業がすでに利用しているアプリケーションをスパコンで実行することも可能。また、企業によるソフトウェアの新規開発も可能。

## 1.3 公募日程

利用課題の募集は、年2回(第1期、第2期)の定期募集を原則とするが、平成21年度は、予定していた課題数を採択できなかつたため、第3期公募を実施した。

### 1.3.1 平成21年度第1期公募

2月2日(月)	受付開始
3月6日(金)	受付締め切り
3月25日(水)	採択通知
4月1日(水)	利用手続き開始

東京大学情報基盤センターでは2月2日と2月19日に公募説明会を実施した。

### 1.3.2 平成21年度第2期公募

7月9日(木)	受付開始
9月4日(金)	受付締め切り
9月30日(水)	採択通知
10月1日(木)	利用手続き開始

東京大学情報基盤センターでは7月9日と8月3日に公募説明会を実施した。

### 1.3.3 平成21年度第3期公募(臨時)

10月1日(木)	受付開始
12月4日(金)	受付締め切り
12月24日(木)	採択通知
1月4日(月)	利用手続き開始

東京大学情報基盤センターでは10月13日と10月30日に公募説明会を実施した。

## 1.4 採択結果

### 1.4.1 平成21年度第1期公募

- 戦略分野利用推進
  - 申請3件(うち継続申請1件、東京大学利用希望は0件)



- 採択 3 件(うち継続課題 1 件、東京大学での採択は 0 件)
  - ◇ 株式会社豊田中央研究所 計算物理研究室「多階層複合材料の大規模シミュレーション」
  - ◇ 株式会社 Hyper-SIT 新技術・事業開発室「極限構造超微細新機能デバイスの開発」
  - ◇ 株式会社コベルコ科研 技術本部エンジニアリングメカニクス事業部 応用技術部 流熱加工技術室「エンジン燃焼室内の熱空力環境の予測技術の開発」
- 新規利用拡大
  - 申請 3 件(うち継続申請 1 件、東京大学利用希望は 1 件)
  - 採択 3 件(うち継続課題 1 件、東京大学での採択は 1 件)
    - ◇ HPC システムズ株式会社 新規事業企画室「量子化学計算 ASP におけるアプリケーションレベルの有効な資源配分の研究」
    - ◇ 北興化学工業株式会社 化成品研究所 ファインセラミックス研究室「自動車排ガス触媒の開発」
    - ◇ 株式会社日立東日本ソリューションズ 計算科学ソリューション部「FMO 法による電子状態計算と超並列機による性能評価」

#### 1.4.2 平成 21 年度第 2 期公募

- 大規模計算利用
  - 申請 1 件(東京大学利用希望は 0 件)
  - 採択 1 件(東京大学での採択は 0 件)
    - ◇ パナソニック株式会社 本社 R&D 部門 PE 技術開発室「大規模 EMC シミュレーション」
- トライアルユース
  - 申請 7 件(東京大学利用希望は 1 件)
  - 採択 7 件(東京大学での採択は 1 件)
    - ◇ ホンダエンジニアリング株式会社 車体研究開発部「塗装微粒化シミュレーション実現可能性の検討」
    - ◇ 株式会社東芝 京浜事業所「蒸気タービン高圧段翼列設計への数値タービンの実用」
    - ◇ 株式会社ベストシステムズ ソリューション事業部「大規模計算機システムにおける量子化学ソフトウェアの比較検証」
    - ◇ ソニーイーエムシーエス株式会社 稲沢テック 技術部「はんだ熱疲労長寿命化に向けた解析精度向上および時間短縮化の検討」
    - ◇ 株式会社コスミックシステム開発研究所「確率的交通量予測モデルの大規模計算機利用の有用性検討」
    - ◇ 株式会社デンソー システム開発部「車載電子機器の自動車搭載状態における電磁環境解析」
    - ◇ バンドー化学株式会社 R&D センター 要素技術開発部「有機分子におけるアモルファスバルク中の電子状態計算」
- 戦略分野利用推進(継続のみ)
  - 申請 2 件(東京大学利用希望は 0 件)
  - 採択 2 件(東京大学での採択は 0 件)
    - ◇ 三菱航空機株式会社 本社 技術部「国産小型民間航空機空力設計への大規模空力解析技術の適用」
    - ◇ TDK 株式会社 テクノロジーグループ SQ 研究所「熱アシスト磁気記録の記録磁化過程シミュレーション」

- 新規利用拡大(継続のみ)
  - 申請 2 件(東京大学利用希望は 0 件)
  - 採択 2 件(東京大学での採択は 0 件)
    - ◇ 北興化学工業株式会社 化成品研究所 ファインセラミックス研究室「自動車排ガス触媒の開発」
    - ◇ 株式会社日立東日本ソリューションズ 計算科学ソリューション部「FMO 法による電子状態計算と超並列機による性能評価」

#### 1.4.3 平成 21 年度第 3 期公募

- トライアルユース
  - 申請 1 件(東京大学利用希望は 0 件)
  - 採択 1 件(東京大学での採択は 0 件)
    - ◇ 大塚化学株式会社 総合研究所「電気二重層キャパシタの材料評価」

### 1.5 広報活動

新聞・雑誌に公募広告を掲載した。また 4 ページの事業リーフレットと 8 ページの 7 センター紹介パンフレットを 1900 部作成し、各種セミナー、展示会などで配布を行い、本事業の認知度を高める努力を行った。関係機関のメーリングリストへ公募説明会情報などの送信を依頼し配布した。共用ナビへの公募情報、シンポジウム情報を掲載した。

東京工業大学と連携して広報活動を行い、雑誌広告などは重複しないように担当を分けて掲載を行った。

#### 1.5.1 新聞広告

- 半導体産業新聞(8 月 5 日、8 月 19 日)
- 原子力産業新聞(8 月 6 日、8 月 20 日)
- 電気新聞(8 月 7 日、8 月 21 日)

#### 1.5.2 雑誌広告

- 日本化学会「化学と工業 11 月号～2 月号」(11 月 1 日～2 月 1 日)
- 電気学会「電気学会誌 11 月号～2 月号」(11 月 1 日～2 月 1 日)
- 日本機械学会「日本機械学会誌 11 月号～2 月号」(11 月 5 日～2 月 5 日)
- 日本計算工学会学会誌「2010 年第 1 号」(1 月 31 日)
- 日刊工業新聞「機械設計 3 月号」(2 月 10 日)

#### 1.5.3 展示会へのブース出展

- 第 8 回産学官連携推進会議(国立京都国際会館、6 月 20～21 日)
- イノベーション・ジャパン 2009 - 大学見本市(東京国際フォーラム、9 月 16～18 日)

### 1.6 シンポジウム開催

7 月 3 日に本プログラムの現状報告、採択企業からの本年度の成果報告などを実施するシンポジウムを行った。参加者は 112 名であった。

#### 第 3 回先端的大規模計算シミュレーションプログラム利用シンポジウム

開催日時： 2009 年 7 月 3 日 (金) 13:30～18:30

開催場所： 丸ビルホール

定員： 200 名

主催： 北海道大学情報基盤センター、

東北大学サイバーサイエンスセンター、  
東京大学情報基盤センター、  
名古屋大学情報基盤センター、  
京都大学学術情報メディアセンター、  
大阪大学サイバーメディアセンター、  
九州大学情報基盤研究開発センター

**趣旨：**

「先端的大規模計算シミュレーションプログラム利用サービス」は、文部科学省先端研究施設共用イノベーション創出事業の支援を受けて、平成 19 年 7 月より平成 21 年 3 月まで、全国共同利用施設であるスーパーコンピュータを有する 7 大学が、社会貢献の一環として、大学で開発された応用ソフトウェアとスーパーコンピュータの民間企業への提供を行ってきたプログラムです。平成 21 年度からは、研究開発施設共用等促進費補助金（先端研究施設共用促進事業）の支援を受け、「先端的大規模計算利用サービス」として、より幅広い研究課題に対して、スーパーコンピュータの提供を実施していきます。本シンポジウムでは、本取り組みによる課題の成果報告と、現在の採択課題の中間報告を口頭発表、ポスター発表にて行います。

**プログラム：**

- 13:00～ 受付開始
- 13:30～14:00 主催者挨拶  
美濃 導彦（京都大学 学術情報メディアセンター センター長）  
文部科学省挨拶  
田口 康（文部科学省研究振興局 研究環境・産業連携課 課長）
- 14:00～15:00 採択企業からの報告（口頭発表 1）  
戦略分野利用推進課題の成果報告
- 「熱アシスト磁気記録の記録磁化過程シミュレーション」水野友人（TDK 株式会社テクノロジーグループ SQ 研究所）
  - 「国産小型民間航空機空力設計への大規模空力解析技術の適用」真保雄一（三菱航空機株式会社 機体設計部 空力・性能グループ）
  - 「大規模並列化手法による高周波電磁界解析技術開発」村山敏夫（ソニー株式会社 モノ造り技術センター 技術開発室）
- 15:00～15:20 休憩
- 15:20～16:20 採択企業からの報告（口頭発表 2）  
新規利用拡大課題の成果報告と有償利用移行への課題について
- 「量子化学計算 ASP における有効な資源配分の研究」英憲悦（HPC システムズ株式会社 新規事業企画室）
  - 「大規模 HPC 向け ASP サービスに関する研究成果と課題」松井周一（NEC ソフト株式会社 ニューソリューション事業部 エンジニアリング SI グループ）
  - 「HA8000 システムによる大規模流体解析」山出吉伸（みずほ情報総研株式会社 サイエンスソリューション部）
- 16:30～17:30 ポスターセッション（採択企業から 6 社）
- 「Large Eddy Simulation によるエマルジョン燃焼シミュレーションの開発」庄子正和（株式会社エコ・ウェーブ名取研究所）
  - 「大規模 EMC シミュレーション」崎山一幸（パナソニック株式会社 本社 R&D 部門 PE 技術開発室）

- 「蛋白質と薬物の結合自由エネルギーの予測と生理活性物質探索－ACE阻害剤探索」 緑川淳（株式会社ワールドフュージョン）
- 「遷移金属酸化物における酸素欠陥に関する研究」 稲男健（株式会社村田製作所 野洲事業所 故障解析センター故障物理解析課）
- 「大規模計算結果の遠隔可視化システムに関する技術開発と実証試験」 宮地英生（株式会社ケイ・ジー・ティー ビジュアリゼーション事業部 技術部）
- 「タンパク質機能シミュレータによる創薬開発」 志水隆一（特定非営利活動法人バイオグリッドセンター関西）

順不同

17:30～18:30

懇親会

# HA8000 クラスタシステムの企業利用支援

## 研究支援係

### 1 HA8000 クラスタシステムの企業利用支援

#### 1.1 企業利用趣旨

工学系研究者、産業界における大規模計算シミュレーションに対する期待が高まっている一方で、その利用が進まず、また、企業におけるスパコン導入の動きは必ずしも旺盛ではない。その理由として、PC レベルでの小規模計算シミュレーションを大量に処理したいというユーザが多いこと、小規模計算シミュレーションを行っているユーザが大規模計算シミュレーションを行うような機会がないために、費用対効果を検討できないということもあげられる。さらに多くの現場では、計算シミュレーションの方法論が浸透していないということもあげられる。

一方、文部科学省は、世界が研究開発にしのぎを削る分野で日本の国際競争力を高めようと、大学や独立行政法人等の公的研究機関がもつ先端研究装置の企業への開放を拡大する方針を打ち出している。

このような状況を背景として、情報基盤センターは、次世代スーパーコンピュータにつながるユーザ育成も視野に入れつつ、大規模高性能並列計算を必要とする企業に対して計算資源を提供する。企業利用については、企業における単なる計算需要の負荷を肩代りするのではなく、以下の項目に合致するテーマを支援するものである。

- (1) 将来の科学技術発展に寄与する。
- (2) 大規模高性能並列計算分野の発展に寄与する。
- (3) 大規模高性能並列計算によるイノベーションに寄与する。

#### 1.2 支援内容

企業利用へ提供するものは、HA8000 クラスタシステム全資源の 10%以内であり、この 10%には先端研究施設共用促進事業の企業利用も含まれる。

以下の 2 つの利用者カテゴリで課題を公募し、審査委員会での審査を行う。企業利用では、HA8000 クラスタシステムの専用キューコースかノード固定コースの利用を申し込むことができる。これらのコースは 8 ノードを 1 単位とし、企業が必要なノード数を決定する。ただし、限られた資源を有効に利用するために、採択可能な課題に対して優先順位を付け、提供可能な計算資源を勘案し、申込内容の調整を行い、採択課題を決定する。

- アプリケーション開発者あるいは利用者

企業の開発現場において既に確立されている計算シミュレーションによるものづくりに使われるのではなく、産業利用として先端性を有する計算シミュレーションおよび応用分野の課題を支援する。このような先端性を有する計算シミュレーションでは、大規模な並列計算が必要であると考えられる。

- ASP(Application Service Provider) 事業者

計算シミュレーションによるものづくりを行ってきていない企業に対し計算シミュレーションによる高度ものづくりを支援するためには、計算シミュレーションソフトウェアの性質を理解し、入力データの作成、シミュレーション結果の解析など、きめ細かい支援が必要である。産業界に対して、このような広範な支援を大学が行うのは難しく、このような支援を行っていく企業との連携が不可欠である。このために、計算シミュレーションによるものづくりを従来行っていない企業への支援体制が整っている ASP 事業者を支援する。

大学が提供する計算資源は限られており、また、本来、企業は、自前でスパコンを確保されるべきものである。企業においては、大規模計算シミュレーションによるものづくりの可能性が得られた場合に、支援を終了する。ASP 事業者においては、ASP 事業者が自前でスパコンを確保できる環境が整い次第、支援を終了する。

### 1.3 平成 21 年度企業利用

平成 21 年度は 4 月、8 月、12 月利用開始の 3 回の課題公募を実施した。4 月利用開始の公募では、平成 20 年度に利用を開始した 2 社が継続して応募し、新規の応募はなかった。8 月利用開始の公募では、新たに 1 社の応募があった。12 月利用開始の公募では応募はなかった。

8 月利用開始の公募に応募してきた 1 社は、共用イノベーション創出事業での 1 年間の無償利用によって、ASP 事業を進めるにあたっての問題点を検討してきており、その後、共同研究にて事業化のめどが立ったので有償利用に移行した。

- 平成 21 年度第 1 回公募(4 月以降利用開始)
  - 申請 2 件(うち継続申請 2 件)
  - 採択 2 件(うち継続課題 2 件)
    - ◇ みずほ情報総研株式会社 サイエンスソリューション部  
「大規模流体解析サービス“MIZUHO FLOW STATION”」
    - ◇ 富士通アドバンステクノロジー株式会社 HPC 適用推進センター・富士通株式会社  
「情報機器における大規模電磁特性および流体解析」
- 平成 21 年度第 2 回公募(8 月以降利用開始)
  - 申請 1 件
  - 採択 1 件
    - ◇ NEC ソフト株式会社 ニューソリューション事業部 エンジニアリング SI グループ  
「科学技術計算 ASP サービス」
- 平成 21 年度第 3 回公募(12 月以降利用開始)
  - 申請 0 件



# スーパーコンピュータ利用による研究成果報告(2009年)

## スーパーコンピューティング部門

### 1 ベクトル並列型スーパーコンピュータ SR11000

2009年(2009年1月～2009年12月)における、ベクトル並列型スーパーコンピュータ SR11000の利用による研究成果報告については以下のとおりである。

#### 1.1 論文

##### 【物理学】

1. Fumio Nishimura, Toru Takahashi, Kazuyuki Watanabe, Takahiro Yamamoto: Bending Robustness of Thermal Conductance of Carbon Nanotubes: Nonequilibrium Molecular Dynamics Simulation: Applied Physics Express, The Japan Society of Applied Physics, 2, 035003
2. 青木秀夫、黒木和彦、有田亮太郎: Pnictogen height as a possible switch between high-Tc nodeless and low-Tc nodal pairings in the iron based superconductors: Physical Review B, アメリカ物理学会(APS), 79, 224511
3. K. Takano, H. Suzuki and K. Hida: Exact spin-cluster ground states in a mixed diamond chain: Phys. Rev., American Physical Society, B80, 104410
4. K. Hida, K. Takano and H. Suzuki: Finite Temperature Properties of the Mixed Diamond Chain with Spins 1 and 1/2: J. Phys. Soc. Jpn., Physical Society of Japan, 78, 084716

##### 【地球惑星科学】

5. Moon, J.-H., N. Hirose, J.-H. Yoon, and I.-C. Pang: Effect of the along-strait wind on the volume transport through the Tsushima/Korea Strait in September: Journal of Oceanography, 65, 17-29
6. Endoh, Takahiro and Toshiyuki Hibiya: Interaction between the trigger meander of the Kuroshio and the abyssal anticyclone over Koshu Seamount as seen in the reanalysis data: Geophysical Research Letters, 36, L18604, doi:10.1029/2009GL039389
7. Masaru Inatsu and Masahide Kimoto: A scale interaction study on East Asian cyclogenesis using a general circulation model with an interactively nested regional model.: Monthly Weather Review, American Meteorological Society, 137, 2851-2868
8. Ambe, Daisuke, Takahiro Endoh, Toshiyuki Hibiya, and Shiro Imawaki: Transition to the large meander path of the Kuroshio as observed by satellite altimetry: La Mer, 47, 1-2, 19-27
9. Tozuka, T., T. Qu, Y. Masumoto, and T. Yamagata: Impacts of the South China Sea throughflow on seasonal and interannual variations of the Indonesian Throughflow: Dynamics of Atmospheres and Oceans, Vol. 47, 73-85
10. M. Nishikawa, T. Hiyama, K. Tsuboki, Y. Fukushima: Numerical simulations of local circulation and cumulus generation over the Loess Plateau: Journal of Applied Meteorology and Climatology, American Meteorological Society, Vol. 48, No. 4, pp. 849-862
11. 広瀬 直毅・小林 亮祐・高山 勝巳: 対馬暖流分枝説の検証—データ同化の結果—, 海と空, 85, 2, 1-11
12. Takashi Yamamoto: Hybrid Kelvin - Helmholtz/Rayleigh - Taylor instability in the plasma sheet: Journal of geophysical research, American Geophysical Union, 114, A06207

##### 【応用物理学・工学基礎】

13. Ikutaro Hamada, Minoru Otani, Osamu Sugino, and Yoshitada Morikawa: Green's function method for elimination of the spurious multipole interaction in the surface/interface slab model: Phys. Rev. B, The American Physical Society, 80, 165411
14. Nakajima, K.: Strategies for Preconditioning Methods of Parallel Iterative Solvers in Finite-Element Applications on Geophysics: Advances in Geocomputing, Lecture Notes in Earth Science 119, Springer, 119, 65-118

##### 【機械工学】

15. 姜東赫, 米澤宏一, 上田達也, 山西伸宏, 加藤千幸, 辻本良信: 流量変動に対するインデューサ

の応答の LES 解析: ターボ機械, 日本工業出版, Vol. 37, No. 2, pp.94-103

【土木工学】

16. Sasaki, J., Sato, Y., Rasmeemasuang, T. and Shibayama, T.: On the mechanism of organic-rich sediment accumulation at the head of Tokyo Bay : Proc. of 5th Int. Conf. on Asian and Pacific Coasts, 67-74

【総合工学】

17. R. Yano, K. Suzuki, H. Kuroda: Thermally relativistic flows induced by gravitational-force-free particle motion in curved spacetime: Physical Review D, APS, 80, 123506, 10 pages  
 18. R. Yano, K. Suzuki, H. Kuroda: Formulation and numerical analysis of vibrationally coupled recombination of monatomic molecules using Boltzmann kinetic equation: Physics of Fluids, AIP, 21, 12, 127101  
 19. R. Yano, K. Suzuki, H. Kuroda: Analytical and numerical study on the nonequilibrium relaxation by the simplified Fokker-Planck equation: Physics of Fluids, AIP, 21, 4, 047104

## 1.2 口頭・ポスター発表

【物理学】

20. 観山正道、佐々真一: 液-固相転移点近傍における応力と剪断率の関係: 日本物理学会 2009 年秋季大会, 日本物理学会講演概要集 (第 64 巻第 2 号第 2 分冊), pp.198 (講演番号: 26pQL-8)

【地球惑星科学】

21. 遠藤 貴洋、市川 香: 台湾北東沖で観測された黒潮の数十日スケールの変動に関する数値実験: 日本海洋学会春季大会  
 22. 川崎高雄、羽角博康: 千島列島周囲に局在化した鉛直混合が引き起こす太平洋熱塩循環: 日本海洋学会 2009 年秋季大会, 日本海洋学会 2009 年秋季大会 予稿集  
 23. 東塚知己、横井孝暁、山形俊男: セーシェルドームの経年変動: 2009 年度日本海洋学会春季大会  
 24. Tozuka, T., T. Yamagata: Prediction of Climate Variations and Its Application in the Southern African Region: Southwestern Indian Ocean Workshop  
 25. 長船哲史、安田一郎: 潮汐 18.6 年振動に伴う北太平洋亜寒帯域における表層塩分の 20 年変動: 日本海洋学会春季大会  
 26. Satoshi Osafune, Ichiro Yasuda: Numerical study on bidecadal water mass variations in the subarctic North Pacific related to the 18.6-year tidal cycle: PICES Annual Meeting  
 27. 浅井丈昭, 見延庄士郎, 稲津将: 東シナ海の黒潮による梅雨への影響: 降水集中化の可能性: 気象学会  
 28. Y. Niwa, P. K. Patra, Y. Sawa, D. Belikov, M. Ikegami, R. Imasu, T. Maki, T. Machida, S. Maksyutov, H. Matsueda, T. Nakazawa, T. Oda, M. Satoh: CONTRAIL-Transport Model Intercomparison (TMI): Analysis of CO<sub>2</sub> Vertical Profiles: 8th International Carbon Dioxide Conference, ICDC8 Abstracts  
 29. Y. Niwa, R. Imasu, H. Matsueda, Y. Sawa, T. Machida, M. Satoh: Seasonal and interannual variation of CO<sub>2</sub> concentration over the western Pacific: model analysis with inversed fluxes: 8th International Carbon Dioxide Conference, ICDC8 Abstracts  
 30. Hirose, Naoki: Inverse estimation of empirical parameters in a circulation model for the East Asian marginal seas: 5th WMO Symposium on Data Assimilation  
 31. 広瀬 直毅, Sergey Varlamov: 日本海モデルによる急潮の予測: 水産学会シンポジウム「急潮の発生・伝播機構と定置網の被害防除」, 75, 891-892  
 32. 森岡優志、東塚知己、山形俊男: 南インド洋の亜熱帯ダイポールモード現象がアフリカ南部の降水量に及ぼす影響: 2009 年度日本海洋学会春季大会, 講演要旨集, pp16  
 33. 森岡優志、東塚知己、山形俊男: 南インド洋亜熱帯ダイポールモード現象に伴う混合層水温偏差の形成機構: 2009 年度日本海洋学会秋季大会, 講演要旨集, pp14  
 34. Morioka, Y., T. Tozuka, and T. Yamagata: Climate Variability in the southern Indian Ocean as Revealed by Self-Organizing Maps: 2010 AGU Fall Meeting  
 35. 佐藤陽祐、中島映至、鈴木健太郎、井口孝享: ビン型雲モデルを用いた対流性降水の降水粒子形成メカニズムに関する数値実験: 2009 年日本気象学会秋期大会, 日本気象学会 2009 年度秋季大会講演予稿集, pp.125  
 36. 佐藤陽祐、中島映至、鈴木健太郎、井口孝享: ビン型雲モデルを用いたビデオゾンデ観測との比較と考察: 第 11 回非静力学モデルに関するワークショップ, 第 1 1 回非静力学モデルに関するワークショップ講演予稿集, F-3  
 37. Tozuka, T., T. Doi, T. Kataoka, T. Miyasaka, and T. Yamagata: Climate Simulations in the University of

- Tokyo Coupled Model: JAMSTEC アプリケーションラボ 国際シンポジウム-南アフリカにおける気候変動と亜熱帯海洋の役割-
38. 東塚知己: 気候変動予測とアフリカ南部における応用: 2009年度日本海洋学会秋季大会シンポジウム「陸域と海洋の相互作用-海から陸・陸から海へ」
  39. 土井威志、東塚知己、山形俊男: 大西洋南北モードとギニアドームの相互関係について: 2009年度日本海洋学会春季大会
  40. S. Arakane, M. Satoh, and W. Yanase: Numerical study on the rapid development of the deep convection to the north of Typhoon Bebinca: The Fourth Japan-China-Korea Joint Conference on Meteorology
- 【基礎化学】
41. 佐藤浩太、戸島玲: 量子化学に基づくⅢ族化合物のSiC(100)面への吸着過程の研究: 第3回分子科学討論会, 第3回分子科学討論会, 1P130
- 【応用物理学・工学基礎】
42. 田村純一、坪谷怜、桑島豊、重原孝臣: 実対称固有値問題に対する多分割の分割統治法の共有メモリ型並列計算機における有効性: 2009年ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2009, HPCS2009 論文集, pp.97-104
- 【機械工学】
43. 山崎健哉、竹田真、松原幸治、櫻井篤: 空間的に発達する二次元曲がりチャンネル乱流への半径比の影響: 第23回数値流体力学シンポジウム, 第二十三回数値流体力学シンポジウム 講演要旨集, pp.3

### 1.3 受賞情報

44. Satoshi Osafune: Numerical study on bidecadal water mass variations in the subarctic North Pacific related to the 18.6-year tidal cycle: POC Committee Best Presentation Award: North Pacific Marine Science Organization (PICES):

## 2 超並列型スーパーコンピュータ HA8000 クラスタシステム

2009年(2009年1月～2009年12月)における、超並列型スーパーコンピュータ HA8000 クラスタシステムの利用による研究成果報告については以下のとおりである。

### 2.1 論文

#### 【情報学】

45. 石川裕、片桐孝洋、吉廣保: T2K オープンスーパーコンピュータと共用イノベーション: 電気学会誌, 電気学会, 129, 1, pp.28-31
46. 石川裕、片桐孝洋、佐藤三久、朴泰祐、中島浩: 高生産・高性能計算機環境実現のためのシステムソフトウェア: HPCS2009論文集(ポスター論文), 情報処理学会, pp.53
47. 鴨志田良和、田浦健次朗: 並列アプリケーションの性能を損なわないポーリング型のモニタリング: 情報処理学会研究報告2009-HPC-121, 情報処理学会, No. 8
48. 緑川博子、齋藤和広、佐藤三久、朴泰祐: クラスタをメモリ資源として利用するためのMPIによる高速大容量メモリ: 情報処理学会論文誌: コンピューティングシステム, 情報処理学会, Vol.2, No.4, pp.15-36
49. 片桐孝洋: ペタスケール計算を目指したMRRR法を用いた固有値ソルバの開発: 計算工学会講演論文集, pp.185-188
50. 山田進、今村俊幸、町田昌彦: マルチコアクラスタのネットワーク構造を考慮した並列密度行列繰り返し込み群法の通信手法: 日本計算工学会論文集, 2009, 20090015
51. Okada J, Hisada T: Study on Compressibility Control of Hyperelastic Material for Homogenization Method Using Mixed Finite Element Analysis: JCST, The Japan Society of Mechanical Engineers, 3, 1, pp. 89-100
52. Okada J, Washio T, Hisada T: Nonlinear Homogenization Algorithms with Low Computational Cost: JCST, The Japan Society of Mechanical Engineers, 3, 1, pp. 101-114

53. Sugiura S, Katayama S, Umetani N, Hisada T: Simulation study of aortic valve function using the fluid-structure interaction finite element method.; Advances in understanding aortic diseases, Springer Japan, pp. 53-60
54. Okada J, Washio T, Hisada T: Study of Efficient Homogenization Algorithms for Nonlinear Problems: Comput Mech, Springer Berlin / Heidelberg, Published online: 31 October 2009
55. Washio T, Hisada T: Convergence Analysis of Inexact LU-type, Preconditioners for Indefinite Problems in Incompressible Continuum Analysis: JJIAM, apan Journal of Industrial and Applied Mathematics, in print  
【ナノ・マイクロ科学】
56. 飛松啓司、寺澤麻子、山本貴博、多田朋史、渡邊聡: 多端子電気伝導シミュレータの開発とカーボンナノチューブへの応用: ナノ学会会報, ナノ学会, 第 8 巻, 第 1 号, pp.21-26  
【ゲノム科学】
57. Shin Sasaki, Cecilia C. Mello, Atsuko Shimada, Yoichiro Nakatani, Shin-ichi Hashimoto, Masako Ogawa, Kouji Matsushima, Sam Guoping Gu, Masahiro Kasahara, Budrul Ahsan, Atsushi Sasaki, Taro Saito, Yutaka Suzuki, Sumio Sugano, Yuji Kohara, Hiroyuki Takeda, Andrew Fire, Shinichi Morishita: Chromatin-Associated Periodicity in Genetic Variation Downstream of Transcriptional Start Sites.: Science, AAAS, Vol. 323, No. 5912, pp. 401 - 404
58. Taro L. Saito, Jun Yoshimura, Shin Sasaki, Budrul Ahsan, Atsushi Sasaki, Reginaldo Kuroshu and Shinichi Morishita.: UTGB Toolkit for Personalized Genome Browsers.: Bioinformatics, Oxford University Press, 25, 15, 1856-61
59. Wei Qu, Shin-ichi Hashimoto, Shinichi Morishita.: Efficient frequency-based de novo short read clustering for error trimming in next-generation sequencing.: Genome Research, Cold Springer Harbor Press, 19, 7, 1309-1315  
【数学】
60. 村上 弘: マルチコア CPU システムおよび小規模 SMP 並列システム上での Tall Skinny 型 QR 分解法の実験: 情報処理学会論文誌 コンピューティングシステム (ACS26), 情報処理学会, Vol.2, No.2, pp.19-29  
【物理学】
61. S.Aoki, K.I.Ishikawa, N.Ishizuka, T.Izubuchi, D.Kadoh, K.Kanaya, Y.Kuramashi, K.Murano, Y.Namekawa, M.Okawa, Y.Taniguchi, A.Ukawa, N.Ukita and T.Yoshie: Precise determination of the strong coupling constant in  $N_f=2+1$  lattice QCD with the Schroedinger functional scheme: Journal of High Energy Physics, SISSA International School for Advanced Studies, 10, 053, 1-20
62. T. Ikeshoji, E. Tsuchida, K. Ikeda, M. Matsuo, H-W. Li, Y. Kawazoe, and S. Orimo: Diffuse and doubly split atom occupation in hexagonal LiBH<sub>4</sub>: Applied Physics Letters, American Institute of Physics, 95, 22, 221901
63. Seiji Miyashita, Shu Tanaka, Hans De Raedt, and Bernard Barbara: Quantum response to time-dependent external field: Journal of Physics: Conference Series, IOPscience, 143, 012005-1-10  
【地球惑星科学】
64. Nozomu Takeuchi: A low - velocity conduit throughout the mantle in the robust component of a tomography model: Geophysical Research Letters, American Geophysical Union, 36, L07306, doi:10.1029/2009GL037590
65. Yuta Mitsui, Kazuro Hirahara: Coseismic thermal pressurization can notably prolong earthquake recurrence intervals on weak rate and state friction faults: Numerical experiments using different constitutive equations: JOURNAL OF GEOPHYSICAL RESEARCH -SOLID EARTH, 114, B09304
66. Yuta Mitsui, Kazuro Hirahara: Interseismic pore compaction suppresses earthquake occurrence and causes faster apparent fault loading: GEOPHYSICAL RESEARCH LETTERS, L20305
67. Kazuro Hirahara: Toward advanced earthquake cycle simulation: Journal of Disaster Research, 富士技術出版株式会社, Vol.4, No.2, pp.99-105
68. Y. Kosaka, H. Nakamura, M. Watanabe and M. Kimoto: Analysis on the Dynamics of a Wave-like Teleconnection Pattern along the Summertime Asian Jet Based on a Reanalysis Dataset and Climate Model Simulations: Journal of the Meteorological Society of Japan, 学術図書印刷, Vol. 87, No. 3, pp. 561-580
69. M. TAKIGAWA, M. Niwano, H. Akimoto, M. Takahashi, and K. Kobayashi: Projection of surface ozone over East Asia in 2020: Journal of Agricultural Meteorology, 日本農業気象学会, Vol. 65, No. 2, pp. 161-166  
【基礎化学】
70. 岩倉いずみ、小林孝嘉、藪下篤史: Direct observation of molecular structural change during intersystem



crossing by real-time spectroscopy with a few optical cycle laser: *Inorganic Chemistry*, the American Chemical Society, 48, 3523-3528

【応用物理学・工学基礎】

71. 中島研吾: マルチコアクラスタにおける有限要素法アプリケーションのための階層型領域間境界分割に基づく並列前処理手法: 情報処理学会研究報告, 情報処理学会, HPC-119, 103-108
72. 大西立顕, 高安秀樹, 高安美佐子: ページランクとオーソリティ・ハブによる企業間取引の有向ネットワーク解析: 統計数理研究所共同研究レポート, 226, pp.23-31
73. Takaaki Ohnishi, Hideki Takayasu, Misako Takayasu: Hubs and authorities on Japanese inter-firm network: Characterization of nodes in very large directed networks: *Progress of Theoretical Physics Supplement*, the Yukawa Institute for Theoretical Physics and the Physical Society of Japan, 179, pp.157-166
74. 中島研吾, 片桐孝洋: マルチコアプロセッサにおけるリオーダーリング付き非構造格子向け前処理付反復法の性能: 情報処理学会研究報告, 情報処理学会, HPC-120
75. 中島研吾: 「並列反復法と自動チューニング—マルチコア時代の並列プログラミングモデル—」, 特集: 科学技術計算におけるソフトウェア自動チューニング: 情報処理, 情報処理学会, 50, 6, 517-522
76. Nakajima, K.: Parallel Multistage Preconditioners by Extended Hierarchical Interface Decomposition for Ill-Conditioned Problems: *Proceedings of International Conference on Parallel Computing (ParCo2009)*, (in press)
77. Nakajima, K.: Flat MPI vs. Hybrid: Evaluation of Parallel Programming Models for Preconditioned Iterative Solvers on “T2K Open Supercomputer”: *IEEE Proceedings of the 38th International Conference on Parallel Processing (ICPP-09) (Second International Workshop on Parallel Programming Models and Systems Software for High-End Computing (P2S2))*, IEEE, 73-80

【機械工学】

78. 宮村倫司, 大崎純, 小檜山雅之, 恩田邦藏, 磯部大吾郎秋葉博, 堀宗朗, 梶原浩一, 井根達比古: ソリッド要素を用いた鋼構造 4 層骨組の高精度有限要素解析: 日本機械学会第 22 回計算力学講演会講演論文集, 日本機械学会, pp.218-219
79. 宮村倫司, 大崎純, 小檜山雅之, 磯部大吾郎, 秋葉博, 堀宗朗, 梶原浩一, 井根達比古: 数値震動台プロトタイプによる鋼構造超高層骨組および 4 層骨組の仮想震動実験: 日本建築学会 2008 年度大会(東北)学術講演梗概集, 日本建築学会, pp.399-400
80. H. Akiba, M. Ohsaki, T. Miyamura, M. Kohiyama, D. Isobe, M. Hori, K. Kajiwara, T Ine: Large Scale Parallel Structural Analysis System and Its Application to Nonlinear Seismic Analysis of High-Rise Building Frames: *Proceedings Chinese Congress of Theoretical and Applied Mechanics 2009*, Chinese Society of Theoretical and Applied Mechanics, pp.280-292
81. 大崎純, 宮村倫司, 小檜山雅之, 磯部大吾郎, 秋葉博, 堀宗朗, 梶原浩一, 井根達比古: 鋼構造建築骨組の弾塑性地震応答の高精度有限要素解析: 日本計算工学会講演論文集, 日本計算工学会, 14, pp.829-832
82. 横山博史, 加藤千幸: Fluid-acoustic interactions in self-sustained oscillations in turbulent cavity flows. I. Fluid-dynamic oscillations: *Physics of Fluids*, American Institute of physics, Vol.21, No.10, 105103, pp.1-8
83. 横山博史, 加藤千幸: 乱流境界層内のキャビティ音発生におけるフィードバック機構(第 1 報, 流体力学的振動): 日本機械学会論文集 B 編, 日本機械学会, Vol.73, No.732, pp.2369-2378
84. Nonaka, Jorji and Ono, Kenji and Miyachi, Hideo: A Decomposition Approach for Optimizing Large-Scale Parallel Image Composition on Multi-Core MPP Systems: *Eurographics Symposium on Parallel Graphics and Visualization*, Eurographics, 71-78
85. Nonaka, Jorji and Ono, Kenji and Miyachi, Hideo: Performance Evaluation of Large-Scale Parallel Image Compositing on a T2K Open Supercomputer: *IPJS Transactions on Advanced Computing Systems*, 情報処理学会, 2, 2, 122-130
86. Nonaka, Jorji and Ono, Kenji and Miyachi, Hideo: Performance Evaluation of Large-Scale Parallel Image Compositing on a T2K Open Supercomputer: ハイパフォーマンスコンピューティングと計算科学シンポジウム, HPCS2009, 情報処理学会, 73-80
87. Atsushi Sekimoto, Alfredo Pinelli, Markus Uhlmann, Genta Kawahara: The effect of coherent structures on the secondary flow in a square duct: 『*Advances in Turbulence XII, Proc.12th EUROMECH Eur. Turb. Conf.*』, Springer Proceedings in Physics 132, Springer, pp.329-332
88. Atsushi Sekimoto, Kazuhide Sekiyama, Genta Kawahara, Markus Uhlmann, Alfredo Pinelli: Buoyancy effects on low-Reynolds-number turbulent flow in a horizontal square duct: 『*Proc. Of the Sixth Interna-*

tional Symposium On Turbulence, Heat and Mass Transfer.』

【電気電子工学】

89. 古賀裕明: Effect of As preadsorption on InAs nanowire heteroepitaxy on Si(111): A first-principles study: Physical Review B, American Physical Society, 80, 24, pp.245302/1-6

【建築学】

90. 今野雅、西村 彩子、荻野 佳、坂本雄三: 大規模低層建物と大型屋根に囲まれた屋外広場内の風環境評価: 第 23 回数値流体力学シンポジウム, 日本流体力学会, G4-3

【総合工学】

91. Y. Idomura, H. Urano, N. Aiba, and S. Tokuda: Study of ion turbulent transport and profile formations using global gyrokinetic full-f Vlasov simulation: Nuclear Fusion, IOP, IAEA, 49, 065029

## 2.2 口頭・ポスター発表

【情報学】

92. 片桐孝洋: 次世代計算機環境における固有値解法と自動チューニング機能の開発: 計算科学研究センター主催、スーパーコンピュータワークショップ 2009、「次世代理論化学の新展開と超並列計算への挑戦」
93. 片桐孝洋: Towards Sparse Iterative Solver with Auto-tuning Facility on Petascale Computing Era: 第 14 回 NEXT(数値トカマク)研究会
94. 片桐孝洋: 次世代スパコンに向けた固有値解法と自動チューニング機能の開発: 大阪大学 蛋白質研究所 主催、蛋白研セミナー、「蛋白質のバイオスーパーコンピューティング」
95. Takahiro Katagiri: Auto-tuning facility for peta-scale computing: International Workshop on Peta-Scale Computing Programming Environment, Languages and Tools
96. 片桐孝洋、櫻井隆雄、黒田久泰、直野健、中島 研吾: 汎用的自動チューニングインターフェースをもつ疎行列反復解法ライブラリ: SACSIS 2009, SACSIS 2009 予稿集, pp.116-117
97. 片桐孝洋、齊藤竜彦、古村孝志、中島研吾: 3次元津波伝搬シミュレーションにおけるコード最適化一手法: 情報処理学会HPC研究会, 情報処理学会研究報告 2009-HPC-120
98. 片桐孝洋、櫻井隆雄、黒田久泰、直野健、中島研吾: OpenATLib:汎用的な自動チューニングインターフェースの設計と実装: 2009年並列/分散/協調処理に関する『仙台』サマー・ワークショップ (SWOPP 2009), 情報処理学会研究報告 2009-HPC-121
99. 櫻井隆雄、直野健、片桐孝洋、中島研吾、黒田久泰: OpenATLib を利用した疎行列ライブラリの開発と評価: 2009年並列/分散/協調処理に関する『仙台』サマー・ワークショップ (SWOPP 2009), 情報処理学会研究報告 2009-HPC-121
100. 片桐孝洋、黒田久泰: マルチコア・超並列計算機時代の自動チューニング機能付き疎行列反復解法ソルバ: 日本応用数学会 2009 年度年会
101. 片桐孝洋、黒田久泰: マルチコア環境における密および疎行列ソルバの自動チューニング機構の評価: 2009年並列/分散/協調処理に関する『仙台』サマー・ワークショップ (SWOPP 2009)
102. 櫻井隆雄、直野健、片桐孝洋、中島研吾、黒田久泰: クリロフ部分空間法に対する自動チューニングについて: 加速法フォーラム
103. Takahiro Katagiri, Takao Sakurai, Hisayasu Kuroda, Ken Naono and Kengo Nakajima: Xablib: An Iterative Solver with a General Auto-tuning Interface “OpenATLib”: International Workshop on Peta-Scale Computing Programming Environment, Languages and Tools (WPSE 2010)
104. 片桐孝洋: ソフトウェア自動チューニング技術の最新動向 ～マルチコア、ヘテロジニアス、10万並列な環境に向けた新しい最適化技術～: 第9回 ANS 研究会
105. 片桐孝洋: ソフトウェア自動チューニングの最新動向 ～数値計算ライブラリへの適用を例にして～: 京都大学グローバルCOE、知識循環社会のための情報学教育研究拠点
106. 片桐孝洋: ソフトウェア自動チューニング: パソコンからスパコンまでの先進最適化技術 ～数値計算ライブラリを中心に～: 情報処理学会東北支部、第350回研究講演会
107. 嶋志田良和、田浦健次朗: 並列アプリケーションの性能を損なわないポーリング型のモニタリング: 並列/分散/協調処理に関する『仙台』サマー・ワークショップ(SWoPP 仙台 2009), 情報処理学会研究報告 2009-HPC-121
108. 嶋志田良和: クラスタ監視ソフト VGXP の開発・運用を通しての評価と課題: 第9回 PC クラスタシンポジウム



109. 西田晃: 大規模科学計算向け汎用数値ソフトウェア基盤 SSI の大規模 SMP クラスタ環境への移植と性能評価: 先駆的科学計算に関するフォーラム 2009 ～先駆的計算科学研究プロジェクト成果報告会&新システム紹介～:先駆的科学計算に関するフォーラム 2009 予稿集, 資料 D
  110. 片桐孝洋: ソフトウェア自動チューニング入門 ～ペタフロップス、10 万並列を達成するための実装方式から数値アルゴリズムの自動最適化技術～: 第 38 回数値解析シンポジウム-NAS2009-, NAS2009 チュートリアル資料集
  111. Takahiro Katagiri: Xablib: A Sparse Iterative Solver with a Generalized Auto-tuning Interface, and Overview of Auto-tuning Studies in Japan: 米国エネルギー省 (DOE) 主催、Center for Scalable Application Development Software (CScADS) Summer Workshops, Workshop on Libraries and Autotuning for Petascale Applications
  112. 片桐孝洋: マルチコア・超並列時代の並列固有値ソルバ実装法 ～ブロック化、マルチキャストを中心に: 日本応用数理学会 3 部会連携 「応用数理セミナー」, 3 部会連携 「応用数理セミナー」資料集
  113. 山田進, 今村俊幸、奥村雅彦、町田昌彦: マルチコアクラスタ向け通信手法を用いた密度行列繰り込み群法の並列化: 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ (SWoPP 仙台 2009), 情報処理学会研究報告 HPC-121
  114. H. Midorikawa, K.Saito, M.Sato, T.Boku: Using a Cluster as a Memory Resource: A Fast and Large Virtual Memory on MPI: IEEE International Conference of Cluster Computing, proc of IEEE International Conference of Cluster Computing 2009 (Digital Object Identifier 10.1109/CLUSTER.2009.5289180 ), pp.1-10
  115. Washio T, Okada J, Hosoi A, Hisada T: Multiscale finite element approach for heart simulation: 原子力機構第 14 回 NEXT 研究会
  116. Sugiura S, Washio T, Okada J, Watanabe H, Hisada T: Multi-scale, multi-physics heart simulator of the University of Tokyo: 2009 Systems Biology speaker Series at the NIH
  117. 久田俊明: 実用化を目指すマルチスケール・マルチフィジックス心臓シミュレータ: SS 研 HPC フォーラム 2009
  118. 久田俊明: マルチスケール・マルチフィジックス心臓シミュレータ UT-Heart: VINAS Users Conference
  119. Watanabe H, Sugiura S, Hisada T: The looped heart does not save energy by maintaining the momentum of blood flowing in the ventricle: FEF09
  120. Iwamura T, Okada J, Washio T, Kadooka Y, Hisada T: An analysis of cardiac mechanoelectric feedback with multi-physics heart simulator: FEF09
  121. Washio T, Okada J, Hosoi A, Hisada T: Efficient parallel solution techniques of homogenization method for nonlinear problems and its application to multi-scale heart simulations: FEF09
  122. 鷺尾 巧, 岡田純一, 久田俊明: 心臓冠循環シミュレーションにおける並列前処理法について: 第 14 回計算工学会, 計算工学講演会論文集, pp.241～244
  123. 細井 聡, 鷺尾 巧, 門岡良昌, 久田俊明: 非線形問題における前処理付き反復法の収束性改善法について: 第 14 回計算工学会, 計算工学講演会論文集, pp.215-218
  124. Sugiura S, Hatano A, Washio T, Okada J, Watanabe H, Hisada T: Multi-scale approach for the understanding of cardiac function: The 36th Congress of the International Union of Physiological Sciences
  125. Okada J, Washio T, Sugiura S, Hisada T: UT-Heart - a Multiscale Multiphysics Heart Simulator: The 36th Congress of the International Union of Physiological Sciences
  126. Washio T, Okada J, Sugiura S, Hisada T: Large-Scale Integrated Model Is Useful for Understanding Heart Mechanisms and Developments in Medical Therapy: The 31st Annual International IEEE EMBS Conference, Program, Paper FrA09.1
  127. 加藤英樹, 竹内郁雄: Zen のクラスタ並列化: 第 14 回 ゲーム・プログラミング ワークショップ 2009, 第 14 回 ゲーム・プログラミング ワークショップ 2009 論文集, pp. 22-26
  128. 浜崎佑樹, 岡本泰英, 大石岳史, 池内克史: 大規模距離画像の並列統合手法の効率化とその評価: 情報処理学会 コンピュータビジョンとイメージメディア研究会, 情報処理学会 コンピュータビジョンとイメージメディア研究会
- 【ナノ・マイクロ科学】
129. 飛松啓司, 寺澤麻子, 山本貴博, 多田朋史, 渡邊聡: ナノスケールにおける多端子電気伝導シミュレータの開発: 次世代スーパーコンピューティングシンポジウム 2009, 次世代スーパーコンピューティングシンポジウム 2009 資料集, pp.49-50

## 【数学】

130. 村上弘: 小規模マルチコア SMP システム上での T-S 型算法による QR 分解、対称行列のブロック三重対角化: 2009 年ハイパフォーマンスコンピューティングと計算科学シンポジウム, HPCS2009 講演論文集(ポスター発表 P2-2), p.32
131. Hiroshi Murakami: A multi-staged algorithm of Householder-type orthogonal transformation for distributed parallel processing: SIAM Conference on Computational Science and Engineering(March 5th,2009) 口頭発表予定
132. 村上弘: ブロック鏡映変換を用いたハウスホルダ型ブロック三重対角化法の、マルチコア及び小規模 SMP システム上での性能: 第 4 回 ASE セミナー、口頭発表予定
133. 村上弘: マルチコア CPU 及び小規模マルチコア SMP システム上での OpenMP 並列化によるハウスホルダ三重対角化の性能実験: 先進的計算基盤システムシンポジウム SACSIS2009, SACSIS2009 論文集, pp.273-282
134. 村上弘: マルチコア CPU 及び小規模マルチコア SMP システム上でのハウスホルダ型三重対角化の実験: 日本応用数理学会 2009 年度年会, 日本応用数理学会 2009 年度年会予稿集, pp.185-196

## 【物理学】

135. T. Kawanai, S. Sasaki and T. Hatsuda: Charmonium-Nucleon Interaction from Quenched Lattice QCD with Relativistic Heavy Quark Action: Third Joint Meeting of the Nuclear Physics Divisions of the APS and JPS
136. 細井亮、吉岡大二郎: DMRG による分数量子ホール効果のエッジ状態の研究: 日本物理学会第 64 回年次大会

## 【地球惑星科学】

137. 三井 雄太、平原 和朗: 速度・状態依存摩擦則に支配された断層上でクラックとアスペリティについて考えながら自発的震源核生成を調べる: 地球惑星科学連合大会 2009
138. 三井 雄太、平原 和朗: 摩擦の時間的回復がゆっくりすべり現象の発生に及ぼす影響 -脱水流体のより本質的な効果を追究して-: 地球惑星科学連合大会 2009
139. Yuta Mitsui, Kazuro Hirahara: Parameter dependencies of spontaneous nucleation on circular homogeneous rate and state faults: Considering the difference of state evolution laws: French-Japanese Workshop on Earthquake Source
140. 光井能麻、堀高峰、宮崎真一、加藤尚之: 2次元平面断層上における地震発生サイクルモデルのデータ同化手法開発: 日本地震学会 2009 年秋季大会, 日本地震学会講演予稿集
141. Hok, S. and Fukuyama, E.: Dynamic rupture of shallow 3D dipping fault earthquakes using a new BIEM technique for half-space medium: American Geophysical Union 2009 Fall Meeting, EOS, S31A-1695
142. Y. Kosaka, H Nakamura, M. Watanabe and M. Kimoto: Dynamics of a Wave-like Teleconnection Pattern Along the Summertime Asian Jet: Fourth Japan-China-Korea Joint Conference on Meteorology
143. Y. Kosaka, H Nakamura, M. Watanabe and M. Kimoto: Dynamics of a Wave-like Teleconnection Pattern along the Summertime Asian Jet: MOCA-09: IAMAS-IAPSO-IACS 2009 Joint Assembly, M06.25/21302
144. 滝川雅之、高橋正明、秋元肇、小林和彦: 領域化学輸送モデルを用いたオゾンによる小麦収穫量推定: 日本農業気象学会, 日本農業気象学会 2009 年大会予稿集
145. M. Takigawa, M. Takhashi, H. Akimoto, and K. Kobayashi: Estimation of ozone-induced crop production loss using an online regional chemical transport model over East Asia: 日中韓国際気象学会, 日中韓国際気象学会予稿集
146. M. Takigawa and I. Uno: Comparison of the distribution of mineral dust calculated by WRF/Chem with the Mie-scattering Lidar observations in East Asia: International workshop on air-quality forecasting, Proceedings of the International workshop on air-quality forecasting
147. M. Takigawa, M. Takhashi, H. Akimoto, and K. Kobayashi: Future projection of surface ozone and its impact on crop yield loss over East Asia in 2020: WRF workshop 2009, Proceedings of WRF workshop 2009
148. 山本勝、広瀬直毅: 海洋データ同化 SST を用いた日本海域の気象シミュレーション: 日本気象学会, 日本気象学会 2009 年春季大会予稿集, pp.350
149. 山本勝、広瀬直毅: 日本海 SST が北西太平洋の気象に与える影響: 日本気象学会, 日本気象学会 2009 年秋季大会予稿集, pp.263
150. 山本勝: 金星大気低安定度層の 1 km スケール擾乱の数値実験: 日本気象学会, 日本気象学会 2009 年秋季大会予稿集, pp.227
151. 深沢圭一郎、梅田隆行、荻野瀧樹、田中高史: High Efficiency Computing with MHD and Vlasov Simulations of Global Magnetosphere: The 9th International School for Space Simulations
152. 深沢圭一郎、梅田隆行、荻野瀧樹、田中高史: スカラー型計算機における電磁流体シミュレーションの性能測定: STE シミュレーション研究会・宇宙プラズマ波動研究会

153. 梅田隆行, 深沢圭一郎, 梶野健太郎, 松本洋介, 中村琢磨, 篠原育: ブラソフモデルによる次々世代ジオスペースシミュレーション: 次世代スーパーコンピューティング・シンポジウム 2009, 次世代スーパーコンピューティング・シンポジウム 2009 資料集, pp.45-46
- 【基礎化学】
154. 関山秀雄, 伊藤彰, 高橋知裕, 辻正高, 濱毅之: CH<sub>3</sub>OCH<sub>3</sub>/Ar, CH<sub>3</sub>OCF<sub>3</sub>/Ar, CF<sub>3</sub>OCF<sub>3</sub>/Ar の構造に関する量子化学的研究: 第 32 回情報化学討論会, 第 32 回情報化学討論会 講演要旨集, pp.40-41
- 【複合化学】
155. 日野 理, 小林 将人, 長谷川 浩司, 中野 達也: TrpCage を対象とした FMO-DFT 計算に基づく密度汎関数の比較: 第 12 回理論化学討論会
156. 小林 将人, 日野 理, 長谷川 浩司, 中野 達也: Acceleration of MP2 calculation with fragment molecular orbital(FMO) scheme by resolution of identity(RI) approximation: 第 12 回理論化学討論会
157. 日野 理, 小林 将人, 長谷川 浩司, 中野 達也: FMO-DFT 計算における不安定性の改善: 第 3 回分子科学討論会 2009 名古屋
- 【応用物理学・工学基礎】
158. 大西立頭, 高安秀樹, 伊藤隆敏, 橋本優子, 渡辺努, 高安美佐子: 外国為替市場における価格変動の連検定: 平成 20 年度統数研研究会「経済物理学とその周辺」第 2 回研究会
159. Takaaki Ohnishi, Misako Takayasu, Hideki Takayasu: Network Motifs in Inter-firm Network: Applications of Physics in Financial Analysis 7th International Conference and Tokyo Tech - Hitotsubashi Interdisciplinary Conference
160. 大西立頭, 高安秀樹, 高安美佐子: 企業間取引ネットワーク上のランダムウォーク: 電子情報通信学会 2009 年総合大会
161. Nakajima, K., Furumura T., Ichimura, T., Nagashima, T., Okuda, H. and Saito, T.: Coupled Simulations in Integrated Predictive Simulation System for Earthquake and Tsunami Disaster: SIAM Conference on Computational Science and Engineering (CSE09), MS107: Multiphysics Modeling: Frameworks and Applications Part-I
162. Nakajima, K.: Robust and Efficient Parallel Preconditioning Methods with Hierarchical Interface Decomposition for Multicore Architectures: SIAM Conference on Computational Science and Engineering (CSE09), MS115: Current Auto-tuning Challenges: Multicore Architecture and Crucial Algorithms Part-II
163. Nakajima, K.: Optimization of Preconditioned Parallel Iterative Solvers for Finite-Element Applications using Hybrid Parallel Programming Models on “T2K Open Supercomputer (Tokyo)”: The 11th International Specialist Meeting on Next Generation Models on Climate Change and Sustainability for High Performance Computing Facilities
164. 中島研吾: Parallel Programming Support for Applications with Unstructured Meshes: Expectations for “Local View” of XscalableMP: International Workshop on Peta-Scale Computing Programming Environment, Languages and Tools (WPSE 2009)
165. 大西立頭, 高安秀樹, 高安美佐子: 企業間取引のネットワーク構造にみられる統計的性質: 日本物理学会第 64 回年次大会
166. 大西立頭, 高安秀樹, 伊藤隆敏, 橋本優子, 渡辺努, 高安美佐子: 連検定からみた金融危機前後の市場特性: 京都大学基礎物理学研究所 2009 年度前期研究会 経済物理学 2009
167. 大西立頭, 鮫島伸二郎, 高安秀樹, 高安美佐子: 企業間ネットワークのパーコレーションと臨界特性: 日本物理学会 2009 年秋季大会
168. 中島研吾: 並列前処理手法と領域分割, マルチコア時代の戦略: 分野横断型研究会「アルゴリズムによる計算科学の融合と発展」
169. Nakajima, K.: Parallel Multistage Preconditioners by Hierarchical Interface Decomposition on “T2K Open Super Computer (Todai Combined Cluster)” with Hybrid Parallel Programming Models: The 15th International Conference on Finite Elements in Flow Problems (FEF09)
170. Nakajima, K.: Evaluation of Hybrid Parallel Programming Models for Finite-Element Applications on T2K Open Supercomputer (Todai Combined Cluster): Third French-Japanese Workshop Petascale Applications, Algorithms and Programming (PAAP)
171. 中島研吾: T2K オープンスパコン (東大) におけるハイブリッド並列プログラミングモデルの最適化: 日本計算工学会第 14 回計算工学講演会
172. 中島研吾: 拡張階層型領域間境界分割に基づく並列前処理手法: 日本応用数学会「行列・固有値問題の解法とその応用」研究部会 (MEPA), 2009 年並列/分散/協調処理に関する『仙台』サマー・ワークショップ (SWoPP 仙台 2009)
173. 中島研吾: マルチコア時代の前処理付き反復法と性能チューニング: 日本応用数学会 2009 年会

174. Nakajima, K.: Flat MPI vs. Hybrid: Evaluation of Parallel Programming Models for Preconditioned Iterative Solvers on Multicore Clusters: The Fourth International Workshop on Automatic Performance Tuning (iWAPT2009)
175. JST-CREST 「原子力発電プラントの地震耐力予測シミュレーション」グループ: 原子力発電プラントの地震耐力予測シミュレーション: JST-CREST 「マルチスケール・マルチフィジックス現象の統合シミュレーション」研究領域 平成 21 年度公開シンポジウム  
【機械工学】
176. K. Matsuzaki, D. Kanno, H. Teshima, N. Shikazono, N. Kasagi: Three-Dimensional Numerical Simulation of Ni-YSZ Anode Polarization Using Reconstructed Microstructure from FIB-SEM Images: Solid Oxide Fuel Cells 11, pp.1829-1836
177. H. Iwai, N. Shikazono, T. Matsui, H. Teshima, M. Kishimoto, R. Kishida, D. Hayashi, K. Matsuzaki, D. Kanno, M. Saito, H. Muroyama, K. Eguchi, N. Kasagi, H. Yoshida: Quantification of Ni-YSZ Anode Microstructure Based on Dual Beam FIB-SEM Technique: Solid Oxide Fuel Cells 11, pp.1819-1828
178. 松崎勝久, 菅野大輔, 鹿園直毅, 笠木伸英: 混合導電性多孔質空気極の 3 次元数値解析: 日本機械学会 2009 年度年次大会, 日本機械学会 2009 年度年次大会講演論文集 (7), pp.115-116
179. 松崎勝久, 鹿園直毅, 笠木伸英: LSCF 空気極過電圧の三次元数値解析: 第 18 回 SOFC 研究発表会, 第 18 回 SOFC 研究発表会講演要旨集, pp.142-143
180. Atsushi Sekimoto, Alfredo Pinelli, Markus Uhlmann, Genta Kawahara: The effect of coherent structures on the secondary flow in a square duct: 12th EUROMECH European Turbulence Conference, 『Advances in Turbulence XII, Proc.12th EUROMECH Eur. Turb. Conf.』 Springer Proceedings in Physics 132, pages 329-332
181. Atsushi Sekimoto, Kazuhide Sekiyama, Genta Kawahara, Markus Uhlmann, Alfredo Pinelli: Buoyancy effects on low-Reynolds-number turbulent flow in a horizontal square duct: Turbulence Heat and Mass Transfer 6, Proc. Of the Sixth International Symposium On Turbulence, Heat and Mass Transfer
182. 関本敦, 関山和英, 河原源太, ウルマン・マルクス, ピネリ・アルフレド: 低レイノルズ数における加熱水平正方形ダクトの二次流れと乱流構造: 日本流体力学会年会 2009, 日本流体力学会年会 2009 講演論文集
183. 関本敦, 河原源太, ウルマン・マルクス, ピネリ・アルフレド: 正方形ダクト乱流におけるパフ構造: 日本機械学会流体工学部門講演会 2009, 日本機械学会流体工学部門講演会 2009 講演論文集
184. 菅野大輔, 松崎勝久, 手島久典, 鹿園直毅, 笠木伸英: 局所平衡に基づく固体酸化物形燃料電池燃料極の 3 次元数値シミュレーション手法の開発: 第 14 回動力・エネルギー技術シンポジウム, 第 14 回動力・エネルギー技術シンポジウム講演論文集, pp.185-188  
【電気電子工学】
185. Ken Shimizu, Takuya Saraya, and Toshiro Hiramoto: Physical Understandings of Si (110) B Hole Mobility in Ultra-Thin Body pFETs by <110> and <111> Uniaxial Compressive Strain: IEEE International Electron Device Meeting (IEDM), IEDM Technology Digest, session 19-6
186. 古賀裕明: Si(111)を基板とする InAs ナノワイヤ気相成長: As 吸着表面の第一原理的研究: 第 70 回応用物理学会学術講演会, 第 70 回応用物理学会学術講演会講演予稿集 No.1, p.397  
【建築学】
187. 今野雅, 西村 彩子, 荻野 佳, 坂本雄三: 大規模低層建物と大型屋根に囲まれた屋外広場内の風環境評価: 第 23 回数値流体力学シンポジウム, 第 23 回数値流体力学シンポジウム講演要旨集, G4-3
188. Masashi Imano, Hideaki Hoshino, Yoshihiko Akamine, Motoyasu Kamata, Ayako Nishimura, Yuzo Sakamoto: Validation of CFD Analysis of Cross-Ventilation in Detached House: Ventilation2009, the 9th International Conference on Industrial Ventilation, Proceedings of the 9th International Conference on Industrial Ventilation, CD-ROM
189. Masashi Imano, Hideaki Hoshino, Yoshihiko Akamine, Motoyasu Kamata, Ayako Nishimura, Yuzo Sakamoto: Validation of CFD Analysis of Cross-Ventilation in Detached House: ROOMVENT 2009, 11th International Conference on Air Distribution in Rooms, Proceedings of ROOMVENT 2009, 11th International Conference on Air Distribution in Rooms, CD-ROM
190. Ayako Nishimura, Masashi Imano, Kei Ogino, and Yuzo Sakamoto: CFD Analysis of Cross-Ventilation in an Existing Old Residential Building: Ventilation2009, the 9th International Conference on Industrial Ventilation, Proceedings of the 9th International Conference on Industrial Ventilation, CD-ROM  
【プロセス工学】
191. 竹田英司, 名嘉山祥也, 梶原稔尚, 重石高志, 富山秀樹: リード付ニーディングディスクを用いた二軸スクリュ押出機の熔融混練: 第 20 回 プラスチック成形加工学会 年次大会



192. 名嘉山祥也: 高分子流動プロセスにおける混合・混練評価と数値シミュレーション: 2009 年度粉体操作に伴う諸現象に関する勉強会
  193. Y. Nakayama and T.Kajiwara: Numerical Methods of Quantifying Mixing in Polymer Processing: 6th International Discussion Meeting on Relaxations in Complex Systems
  194. 竹田英司,名嘉山祥也,梶原稔尚,重石高志,富山秀樹: 二軸スクリュ押出機の溶融混練におけるリード付ニーディングディスクの特性評価: 化学工学会 第41回秋季大会
  195. 竹田英司,名嘉山祥也,梶原稔尚,重石高志,富山秀樹: 二軸スクリュ押出機の溶融混練におけるリード付ニーディングディスクの特性評価: 第57回レオロジー討論会
  196. 竹田英司,名嘉山祥也,梶原稔尚,重石高志,富山秀樹: 二軸スクリュ押出機の溶融混練におけるリード付ニーディングディスクの特性評価: 第17回プラスチック成形加工学会秋季大会(成形加工シンポジウム'09)
  197. Y. Takagi, D. Ando, Y. Okano, M. Sudoh, K. Furukawa, Y. Muramatsu and S. Adachi: Development of numerical simulation model for carbon dioxide effect on direct methanol fuel cell performance: The 5th Joint China/Japan Chemical Engineering Symposium
- 【総合工学】
198. Y. Idomura: Gyrokinetic Full-f Vlasov Simulation of Source Driven ITG Turbulence: The 4th IAEA-TM on the Theory of Plasma Instabilities, Proceedings of the 4th IAEA-TM on the Theory of Plasma Instabilities, pp.18
  199. Y. Idomura: Properties of avalanches and momentum transport in driven ITG turbulence: 7th General Scientific Assembly of the Asia Plasma and Fusion Association and Asia-Pacific Plasmas Theory Conference
  200. Y. Idomura, S. Jolliet: Properties of avalanches and momentum transport in driven ITG turbulence: 19th International Toki Conference on Advanced Physics in Plasma and Fusion Research
  201. S. Jolliet, Y. Idomura: Parallel filtering in global Vlasov simulations: 7th General Scientific Assembly of the Asia Plasma and Fusion Association and Asia-Pacific Plasmas Theory Conference
- 【薬学】
202. 平野 秀典, 沖本 憲明, 泰岡 顕治, 末松 誠, 安井 正人: 分子動力学的手法を用いたアクアポリンに関する研究: 第23回分子シミュレーション討論会, 第23回分子シミュレーション討論会講演要旨集, pp. 252-253
  203. 香川 璃奈, 平野 秀典, 泰岡 顕治, 安井 正人: 分子動力学的手法を用いた電解質水溶液中における水と脂質2重膜に関する研究: 第23回分子シミュレーション討論会, 第23回分子シミュレーション討論会 講演要旨集, pp. 134-135

## 2.3 著書

204. Scalable Software Infrastructure for Scientific Computing Project: Lis 1.2.6 User's Manual: Scalable Software Infrastructure for Scientific Computing Project:

## 2.4 受賞情報

205. 原健太朗: クラスタシステム上の並列プログラミングコンテスト 非数値計算部門: 第1位: PC クラスタコンソーシアム
206. 原健太朗: 第2回クラスタシステム上の並列プログラミングコンテスト 数値計算部門: 第2位: PC クラスタコンソーシアム
207. 飛松啓司: 多端子電気伝導シミュレータの開発とカーボンナノチューブへの応用: 若手優秀発表賞: ナノ学会

PART 5

**その他**

委員会委員等

講習会・セミナー

報道関係一覧





## 平成21年度委員会委員等

教員名	委員会委員等名	任 期
柴山 悦哉	社団法人情報処理学会 理事	20. 4. 1～21. 5.29
	東京大学 キャンパス計画委員会 柏地区部会委員	20. 4. 1～22. 3.31
	東京工業大学 連携教授	21. 4. 1～21. 9.30
	東京大学教養学部 非常勤講師	21. 4. 1～21. 9.30
	独立行政法人産業技術総合研究所 招へい研究員	21. 4. 1～22. 3.31
	日本ソフトウェア科学会 理事	21. 4. 1～23. 3.31
	社団法人情報処理学会 調査研究運営委員会委員	21. 5.29～22. 6. 1
	社団法人情報処理学会 トッププロ棋士に勝つためのコンピュータ将棋委員会委員	21. 5.29～22.10.31
	社団法人情報処理学会 論文誌ジャーナル/JIP 編集委員会 ITWG 顧問	21. 6. 1～23. 5.31
	社団法人情報処理学会 刊行物オンライン化委員会幹事	21. 6. 1～23. 5.31
	社団法人情報処理学会 記念論文選考委員会委員	21. 6. 1～23. 5.31
	財団法人りそな中小企業振興財団 中小企業優秀新技術・新製品賞 専門審査委員会委員	21. 9. 1～22. 8.31
	情報セキュリティ大学院大学 客員教授	21.10. 1～22. 3.31
田中 哲朗	社団法人情報処理学会 会誌編集委員会委員	20. 4. 1～22. 3.31
	社団法人情報処理学会 論文誌ジャーナル編集委員会副査	20. 6. 1～21. 5.31
	社団法人情報処理学会 論文誌ジャーナル編集委員会委員	20. 6. 1～22. 5.31
	全国共同利用情報基盤センター長会議 災害対策検討委員会委員	20.12.10～21.12. 9
	東京大学教養学部 非常勤講師	21. 4. 1～22. 3.31
	社団法人情報処理学会 ゲーム情報学研究運営委員会主査	21. 4. 1～23. 3.31
	社団法人情報処理学会 フロンティア領域委員会委員	21. 4. 1～23. 3.31
	社団法人情報処理学会 論文誌ジャーナル編集委員会主査	21. 6. 1～22. 5.31
関谷 貴之	東京学芸大学 非常勤講師	21. 4. 8～22. 3.19
丸山 一貴	早稲田大学 非常勤講師	21. 4. 1～22. 3.31
中川 裕志	東京大学 キャンパス計画委員会 本郷地区部会委員	20. 4. 1～22. 3.31
	東京大学駒場図書館 運営委員会委員	20. 4. 1～22. 3.31
	京都大学学術情報メディアセンター 全国共同利用運営委員会委員	20. 4. 1～22. 3.31
	社団法人人工知能学会 評議員	20. 7.17～22. 6.30
	北海道大学情報基盤センター 全国共同利用委員会委員	21. 4. 1～23. 3.31
	東京大学柏図書館 運営委員会委員	21. 4. 1～24. 3.31
二宮 崇	言語処理学会 年次大会プログラム委員	20.10. 3～21. 6.30
	国土交通省総合政策局 総合評価検討委員会委員	21. 3. 5～21. 8.31
清田 陽司	株式会社リッテル 上席研究員	20. 8. 1～22. 3.31
	株式会社三菱総合研究所 有職者検討会委員	21. 9. 9～21.10.28
若原 恭	独立行政法人情報処理推進機構 情報処理技術者試験試験委員会委員	20. 4. 1～22. 3.31
	名古屋大学情報連携基盤センター 運営委員会委員	20. 4. 1～22. 3.31
	名古屋大学情報基盤センター 共同利用・共同研究運営委員会委員	21. 4. 1～23. 3.31

委員会委員等

	東京大学評価支援室 室員	21. 4. 1～22. 3. 31
	社団法人電子情報通信学会 ネットワークソフトウェア時限研究専門委員会委員	21. 4. 1～23. 3. 31
	社団法人電子情報通信学会 変革を続けるネットワークソフトウェア技術特集号編集委員	21. 6. 15～22. 4. 1
	東京大学バリアフリー支援室 室員	21. 4. 1～22. 3. 31
	東京大学バリアフリー支援室 支援促進ワーキンググループメンバー	21. 7. 1～22. 3. 31
中山 雅哉	独立行政法人情報通信研究機構 短時間研究員	21. 4. 1～22. 3. 31
	独立行政法人宇宙航空研究開発機構 宇宙利用ミッション本部衛星利用推進センター主幹開発員	21. 4. 1～22. 3. 31
	社団法人電子情報通信学会 インターネットアーキテクチャ研究専門委員会副委員長	19. 5. 25～21. 5. 26
	社団法人電子情報通信学会 インターネットアーキテクチャ研究専門委員会顧問	21. 5. 23～22. 5. 22
	情報・システム研究機構 学術情報ネットワーク運営・連携本部ネットワーク作業部会委員	21. 11. 18～22. 3. 31
	社団法人電子情報通信学会 「Special Section on Architectures, Protocols and Applications for the Future Internet」編集委員会編集委員	22. 1. 22～22. 11. 1
	総務省関東総合通信局 関東 JGN2plus 懇話会会員	20. 9. 30～23. 3. 31
小川 剛史	社団法人情報処理学会 グループウェアとネットワークサービス研究運営委員会運営委員	20. 4. 1～22. 3. 31
	日本バーチャルリアリティ学会 サイバースペースと仮想都市研究委員会幹事	21. 1. 1～23. 12. 31
	東京大学文学部 非常勤講師	21. 4. 1～22. 3. 31
	社団法人電子情報通信学会 ネットワークシステム研究専門委員会専門委員	21. 5. 23～23. 5. 22
	社団法人情報処理学会 論文誌ジャーナル/JIP 編集委員会査読委員	21. 6. 1～24. 5. 31
関谷 勇司	社団法人電子情報通信学会 インターネットアーキテクチャ研究専門委員会専門委員	20. 5. 27～21. 5. 26
	社団法人電子情報通信学会 情報ネットワーク研究専門委員会専門委員	20. 5. 27～21. 5. 26
	社団法人電子情報通信学会 英文論文誌 B 編集委員会編集委員	20. 5. 27～23. 5. 26
	慶應義塾大学 SFC 研究所 上席所員 (訪問)	21. 4. 1～22. 3. 31
中村 文隆	法政大学 非常勤講師	21. 4. 1～21. 9. 30
中島 研吾	独立行政法人海洋研究開発機構 招聘技術研究主幹	21. 4. 1～22. 3. 31
	東京大学大学院理学系研究科 非常勤講師	21. 4. 1～21. 9. 30
	筑波大学計算科学研究センター 並列プログラミング言語検討委員会委員	21. 4. 1～22. 3. 31
	内閣府日本学術会議 小委員会委員	21. 7. 8～22. 7. 7
	東京大学フューチャー推進機構 運営委員会委員	22. 2. 1～23. 3. 31

金田 康正	東京大学物性研究所附属物質設計評価施設 スーパーコンピュータ 共同利用委員会委員	20. 4. 1～22. 3. 31
	東京大学物性研究所附属物質設計評価施設 スーパーコンピュータ 共同利用課題審査委員会委員	20. 4. 1～22. 3. 31
	東京大学気候システム研究センター 研究協議会委員	20. 4. 1～23. 3. 31
	大阪大学サイバーサイエンスセンター 運営委員会委員	20. 4. 1～22. 3. 31
	鶴岡工業高等専門学校 非常勤講師	21. 4. 8～21. 9. 30
	大阪大学サイバーサイエンスセンター 外部評価委員	21. 5. 15～21. 7. 31
	独立行政法人国立環境研究所 地球環境研究センタースーパーコン ピュータ研究利用専門委員会委員	21. 6. 4～22. 3. 31
	内閣府行政刷新会議 行政刷新会議分科会（ワーキンググループ） 評価者	21. 11. 11～23. 11. 10
石川 裕	全国共同利用情報基盤センター長会議 スーパーコンピュータ研究 会委員	18. 12. 1～21. 11. 30
佐藤 周行	社団法人情報処理学会 情報規格調査会委員	20. 3. 17～24. 3. 16
	社団法人情報処理学会 ハイパフォーマンスコンピューティング研 究運営委員会運営委員	20. 4. 1～22. 3. 31
	日本ソフトウェア科学会 理事	21. 4. 1～23. 3. 31
	国立情報学研究所 学術情報ネットワーク運営・連携本部認証作業 部会委員	21. 4. 1～22. 3. 31
	財団法人日本情報処理開発協会 電子認証等の民間制度・基盤の確 立に関する委員会委員	21. 11. 11～23. 11. 10
片桐 孝洋	社団法人情報処理学会 ハイパフォーマンスコンピューティング研 究運営委員会幹事	20. 4. 1～22. 3. 31
	日本応用数理学会 評議員	21. 4. 1～22. 3. 31
	日本応用数理学会 「行列・固有値問題とその解法」研究部会主査	21. 4. 1～22. 3. 31
	日本応用数理学会 学会誌編集委員会委員	21. 4. 1～22. 3. 31
	日本応用数理学会 JSIAM Letters 編集委員会委員	21. 4. 1～22. 3. 31
	東京大学教養学部 非常勤講師	21. 4. 1～21. 9. 30
	独立行政法人理化学研究所 客員研究員	21. 8. 1～22. 3. 31
伊藤 祥司	社団法人情報処理学会 論文誌コンピューティングシステム編集委 員会委員	22. 2. 1～22. 3. 31
	社団法人情報処理学会 ハイパフォーマンスコンピューティング研 究会運営委員会委員	22. 2. 1～22. 3. 31
	日本応用数理学会 「行列・固有値問題とその解決」研究部会運営 委員会委員	22. 2. 1～22. 3. 31
	日本応用数理学会 ネットワーク委員会委員	22. 2. 1～22. 3. 31
	東京電機大学理工学部 非常勤講師	22. 2. 1～22. 3. 31
	独立行政法人理化学研究所 客員研究員	22. 2. 1～22. 3. 31

## 平成21年度講習会・セミナー

### 1. センター共通

講習会・セミナー名	開催日	参照頁
・コンピュータ・ネットワーク利用セミナー		
(第69回)最近のセキュリティ事情	8/5	p.313
(第70回)ChemBioOffice活用法セミナー	11/10	p.314
(第71回)「平成21年度コンピュータネットワーク研修(東京大学技術職員研修)」講義	11/24-26	p.314

### 2. 情報メディア教育部門によるもの

講習会・セミナー名	開催日	参照頁
・新規利用者向け講習会		
本郷キャンパス	4/1, 4/10, 4/14, 4/16, 4/17, 4/21, 4/23, 4/27, 5/7, 5/13, 5/18, 5/28, 6/4, 6/12, 6/16, 6/22, 7/3, 7/9, 7/27, 8/4, 9/2, 9/24, 10/2, 10/7, 10/13, 10/22, 11/4, 11/17, 11/30, 12/10, 12/16, 1/12, 1/25, 2/5, 2/17, 3/4, 3/16	
駒場キャンパス	4/3, 4/6, 4/7, 4/8, 4/9, 4/10, 4/13, 4/15, 4/17, 4/21, 4/23, 4/30, 5/8, 5/13, 5/18, 5/26, 6/3, 6/11, 6/19, 6/29, 7/7, 7/22, 8/6, 8/18, 9/7, 9/25, 10/1, 10/6, 10/16, 10/21, 10/26, 11/5, 11/10, 11/16, 11/26, 12/4, 12/9, 12/15, 1/8, 1/14, 1/20, 1/26, 2/1, 2/12, 2/16, 2/22, 3/3, 3/11	
柏キャンパス	4/15, 4/24, 5/12, 10/14	p.272
その他	4/2, 4/3, 10/6	
・システム利用説明会		
本郷キャンパス	9/25, 3/24	
駒場キャンパス	9/24, 3/23	

### 3. 図書館電子化部門によるもの

講習会・セミナー名	開催日	参照頁
・情報探索ガイダンス(テーマ別ガイダンス)		
日本の論文を探すには?	5/18, 5/26, 7/10, 10/9, 11/4, 12/4, 1/21, 2/5	p.297
電子ジャーナルを利用するには? (11月から「電子ジャーナルで論文を入手」に改称)	5/19, 5/29, 7/15, 10/23, 11/19, 12/8, 1/20, 2/18, 3/3	
自宅から検索するには?	9/14, 10/7, 11/10, 12/9, 1/29, 2/10, 3/4	

文献検索早わかり	5/8, 5/21, 7/24, 9/16, 9/29, 10/21, 11/25, 12/10, 1/26, 2/12	
RefWorksを使うには	5/22, 7/3, 7/28, 9/15, 9/30, 10/13, 10/29, 11/13, 11/24, 12/16, 1/15, 2/24, 3/12	
国内の新聞記事を探すには？	2/4	
海外の新聞記事を探すには？	3/5	
・情報探索ガイダンス（データベースユーザトレーニング）		p.298
Web of Science + EndNote Web	10/28, 11/27, 2/2	
PubMed + RefWorks	2/9	
Web of Scienceを使うには？	3/9	
EndNote Webを使うには？	3/10	
・情報探索ガイダンス（分野別データベースユーザトレーニング）		p.298
Eighteenth Century Collections Online +GVRL	6/1	
Early English Books Online	6/1	
JapanKnowledge+	6/2, 6/11	
Littel Navigator	6/2	
ヨミダス歴史館	6/3, 6/11	
PIO/PAO	6/3	
eol	6/4	
InfoTrac Custom	6/5	
LRC with MLA+GVRL	6/5	
聞蔵Ⅱビジュアル	6/8, 7/1	
Engineering Village	6/9	
毎日Newsパック	6/10, 7/1	
ASSIA/LLBA	6/10	
Lexis.com	6/12	
LexisNexis Academic	6/12	
MathSciNet	6/15, 6/16	
SciFinder Scholar	6/17, 6/25	
Business Source Elite/SocINDEX	6/18	
Web of Science（人文科学系・自然科学系）	6/19, 6/26, 6/30	
JSTOR	6/22	
PsycINFO/PsycARTICLES	6/23	
CNKI	6/24	
BIOSIS Previews	6/26, 6/30	
EndNote Web	6/26, 6/30	
AGRICOLA Plus Text	7/30	
Westlaw International/Westlaw Japan	9/29	
RefWorks	3/31	
・総合図書館オリエンテーション		p.300
OPAC入門	4/8, 4/10, 4/13, 4/14, 4/15, 4/16, 4/17	



講習会・セミナー

OPAC入門（英語編）	4/13, 4/15	
留学生ガイドランス（日本語）	4/9	
留学生ガイドランス（英語）	4/14, 5/25	
秋季留学生ガイドランス（日本語）	10/8	
秋季留学生ガイドランス（英語編）	10/15, 10/19, 10/29	
・留学生向け情報探索ガイドランス		p.300
韓国語コース	5/14	
中国語コース	11/6	
・事務職員向け情報探索ガイドランス	8/5, 8/11	p.300
・秘書さんのための論文の探し方講習会	11/12	p.300
・出張講習会		p.300
学際情報学府，教養学部，工学部／工学系， 情報理工学系，理学部，文学部／人文社会科学系， 新領域創成科学研究科，千葉演習林， 社会科学研究所，教育学部，駒場Ⅱ，医学系	4/10, 4/17, 4/23, 5/12, 5/18, 5/20, 5/27, 5/28, 7/2, 8/7, 8/14, 9/17, 10/1, 10/9, 10/16, 10/20, 10/22, 10/27, 10/28, 11/2, 11/5, 11/11, 11/20, 11/30, 12/7, 12/18	
・セミナー・説明会		p.303
セミナー「学術雑誌の評価指標を考える」	5/13	
「RefWorks」説明会	5/14, 5/15, 5/20	
「Business Source Premier」説明会	12/10	

4. ネットワーク部門によるもの

講習会・セミナー名	開催日	参照頁
・UTnet Meeting		p.314
第7回	10/30	

5. スーパーコンピューティング部門によるもの

講習会・セミナー名	開催日	参照頁
・お試しアカウント付き並列プログラミング講習会（試行）		p.339
第5回	7/1-2	
第6回	9/7-8	
第7回	9/14-15	
第8回	3/17	
・先進スーパーコンピューティング環境研究会（ASE研究会）		p.343
第5回	8/21	
第6回	2/12	
・その他		p.339
CCS HPCサマーセミナー 東京会場（遠隔）	7/23-24	
並列プログラミング入門講習会 東京会場（遠隔）	9/3-4	

## 平成21年度報道関係一覧

### 情報メディア教育研究部門

- [報道1] 読売新聞 2009年5月6日朝刊:「GPS将棋」が初優勝.
- [報道2] 週刊将棋 2009年5月13日号:革命的結果勢力図一変 - GPS将棋が初優勝.
- [報道3] 朝日新聞 2009年5月13日夕刊:世界コンピュータ将棋選手権 - 「GPS将棋」が初優勝.
- [報道4] NHK BS2、囲碁将棋ジャーナル 2009年5月16日:世界コンピュータ将棋選手権速報.
- [報道5] NHK BS2、囲碁将棋ジャーナル 2009年5月30日:世界コンピュータ将棋選手権特集.
- [報道6] 読売新聞 2009年6月2日夕刊:「GPS」が乱戦制し初優勝.
- [報道7] 産経新聞 2009年7月26日:訪れるか将棋界「Xデー」.

### 学術情報研究部門

- [報道1] 日本経済新聞 2009年(平成21年)5月13日13面:東大発VBのリッテル、商品検索簡単に.

### スーパーコンピューティング研究部門

- [報道1] 片桐孝洋:東京大学新聞,2009年9月22日(火)発刊,第2480号,「10分でゲノム解読する時代に—ゲノム解読をスパコンで,人材育成が急務,情報悪用を防げ」,2009年9月.
- [報道2] 片桐孝洋,黒田久泰,中島研吾:科学技術政策研究所発行,「科学技術動向」11月号特集,「数値シミュレーションにおけるソフトウェア研究開発の動向—並列分散型のハードウェアとソフトウェア自動チューニング—」,2009年11月.

### スーパーコンピューティング部門

- [報道1] 朝日小学生新聞 2009年6月2日1面:次世代スーパーコンピューター、2012年に世界最高性能をめざす.
- [報道2] 日経産業新聞 2009年12月2日10面:次世代スパコン計画継続を要望、東大研究者ら.



# 東京大学情報基盤センター一年報

2009 年度（第 11 号）

## 編 集

東京大学情報基盤センター一年報編集委員会

編集委員長 田中 哲朗

編集委員 二宮 崇、小川 剛史、片桐 孝洋、佐藤 周行、大塚 浩一、  
早野 裕士、本多 玄、平野 光敏、柏 芳美、米山 浩、  
岩崎 直樹、宮地 亨、永岡 陽香

## 発 行

東京大学情報基盤センター

〒113-8658 東京都文京区弥生 2-11-16

電話 03-5841-2710

2010 年 6 月発行



