[2007]九州大学情報統括本部年報 : 2007年度

https://doi.org/10.15017/1470731

出版情報：九州大学情報統括本部年報. 2007, 2008. 九州大学情報統括本部
バージョン：
権利関係：

第5章　イベント紹介

5.1　線形計算フォーラム～線形計算ライブラリへの取組み・概要・性能評価～

2007年9月25日（火）に，九州大学情報基盤研究開発センターにハイパフォーマンスコンピュータを提供している富士通，IBM，日立各社から大規模疎行列に対する連立1次方程式に代表される線形ライブラリを中心に概要・性能を紹介していただき，あわせて今後の取組みを伺いくとともに，線形計算に関する最先端の話題を理論面および応用面から紹介していただきました．

参加者は27名で，技術スタッフ，センター運用関係者も出席し，熱心な質疑応答が交わされました．

（フォーラムプログラム）

- 10:30～10:40 開会の挨拶
- 10:40～12:00 三上 次郎（富士通 (株)・ソフトウェア事業本部）
  「大型スパース行列の並列解法パッケージ MUMPS について」

  《概要》
  大型スパース線形方程式の並列解法パッケージ MUMPS について，その概要と利用方法，および数値事例を紹介する．
  MUMPS（a MUltifrontal Massively Parallel sparse direct Solver)は，欧州の共同プロジェクト PARSOL(1996-1999) で開発されたパブリックドメインソフトである．開発には欧州の研究機関 CERFACS, ENSEEIHT-IRIT, RAL の研究者が参画し，現在も精力的に改版が重ねられ，ENSEEIHT-IRIT から配布されている（http://mumps.enseeiht.fr/avail.html)．フォーラムでは，カリフォルニア大学から配布される実アプリケーションのデータに基づく数値事例も紹介する．

（昼食）

- 13:00～14:20 寒川 光（日本アイ・ビー・エム (株)・東京基礎研究所）
  「疎行列係数の連立1次方程式の直接解法と WSMP」

  《概要》
  はじめに大規模な疎行列を係数行列とする連立1次方程式を直接解法で解くことのメリットを，反復解法と比較して述べる．次に，係数行列が対称な場合について，順序付け (Ordering) と計算量の関係を示し，並列計算のアプローチを考える．これらの考察をもとに，WSMP（Watson Sparse Matrix Package）の手法と機能を紹介する．

（休憩）

- 14:30～15:50 直野 健（(株) 日立製作所・中央研究所），猪貝 光祥（(株) 日立超 LSI システムズ）
  「日立 SR11000 向け疎行列計算ライブラリ MATRIX/MPP/SSS と自動チューニングに向けた取り組み」

  《概要》
  SR11000 向け疎行列ライブラリ MATRIX/MPP/SSS の概要を紹介します．特に，反復法向け前処理（RCM, ND, MD）と反復解法（GMRES, BiCGStab）および直接法スパースソルバについてその特長を解説致します．また，研究活動として，疎行列ライブラリの自動チューニング化に向けた取り組みについても御報告します．

（休憩）

- 16:00～16:45 仙波 和樹（日本総研ソリューションズ），山田 隆（日本総研ソリューションズ），寒川 光（日本アイ・ビー・エム (株)）
  「疎行列ベクトル積プログラムの SMP/MPI 並列性能の比較」

  《概要》
  有限要素法による電磁界解析で用いられる共役勾配法のカーネルである，疎行列とベクトルの積を計算するプログラムにおいて，OpenMP を使用した SMP 並列と MPI を使用した分散メモリ並列性能の比較を，複数のプラットフォームで比較する．プロセッサの計算性能，キャッシュと物理メモリ間のデータ転送性能等，ハードウェア性能を考慮した分析を行うことで，ハードウェアに適した並列手法を検討する．

- 16:45～17:30 森 眞一郎（福井大），山崎 勇輔（福井大），依藤 逸（京都大），野田 裕介（京都大），粂 直人 (京都大), 富田 眞治 (京都大)
  「操作の連続性を考慮した手術シミュレーションの高速化」

  《概要》
  手術手技をシミュレートする手術シミュレーションでは，術者の動作に応じて変形する人体を有限要素モデル化した剛性方程式の求解問題をリアルタイムに処理しなければならない．このとき，手技には連続性が仮定できるため，連続する時系列において解くべき剛性方程式の間にも高い類似性が期待できる．本講演では，この性質を利用した高速化手法について検討した結果を紹介する．

○閉会の挨拶

＊ 敬称略.

＊ 講演時間は質疑応答時間を含みます.

5.2 線形計算フォーラム～4倍精度・多倍長精度演算への取り組みおよび実装技術の現状と今後～

2007年10月22日（月）に九州大学情報基盤研究開発センターにおいて開催しました．悪条件の問題に代表される丸め誤差の影響を受けやすい数値計算環境においては，多倍長演算を用いることが解の妥当性・収束性を得るための有効な手段であることが知られています．今回のフォーラムでは，最新の4倍精度・多倍長精度演算の実装技術について講演していただき，今後の方向性を討論しました．招待講演では，4倍精度版を実装した反復解法ライブラリの機能を性能評価とともに紹介していただきました．

参加者は37名で，技術スタッフ，センター運用関係者も出席し，熱心な質疑応答が交わされました．

（フォーラムプログラム）

- 13:00～13:30 緒方 隆盛（NEC HPC 販売推進本部）
  「高速4倍精度演算パッケージ ASLQUAD の開発とその有用性」

- 13:30～14:00 濱口 信行（(株) 日立製作所 ソフトウェア事業部ネットワークソフトウェア本部 第3ネットワークソフト設計部）
  「多倍長ライブラリによる精度評価と改善に関する考察」

（休憩）

- 14:15～14:45 大澤 暁（日本アイ・ビー・エム株式会社 公共クライアント IT 推進）
  「POWER アーキテクチャにおける4倍精度演算」

- 14:45～15:15 畑崎 隆雄（日本ヒューレットパッカード株式会社シェアードサービス本部 HPC ソリューション）
  「標準サーバでのプログラミングテクニックについて― Intel，AMD プロセッサを搭載したサーバの有効活用法―」

（休憩）

- 15:30～17:00 招待講演 小武守 恒（JST/東京大学）
  「反復解法ライブラリー Lis の紹介」

＊ 敬称略．

＊ 講演時間は質疑応答時間を含みます．

5.3　サイエンス・パートナーシップ・プロジェクト

大分工業高等専門学校制御情報工学科が企画したサイエンス・パートナーシップ・プロジェクト「スーパーコンピュータを用いた計算機実験」に，センターの教職員が協力しました．

サイエンス・パートナーシップ・プロジェクトは，「次代を担う人材への理数教育の拡充」施策の一環として，平成 18 年度より独立行政法人科学技術振興機構において実施されている事業で，学校と大学・科学館等の連携により，児童生徒の科学技術，理科・数学（算数）に関する興味・関心と知的探求心等を育成させることを目的としています．

実施項目は以下の通りでした．

実施計画（案）

- 7 月 27 日
  大分工業高等専門学校制御情報工学科の教員とセンター職員で実施内容や事前に学生に教えておくことなどについて事前打ち合わせ．

- 9 月 19 日
  大分工業高等専門学校にて事前講習会．UNIX システムの基礎とスーパーコンピュータで解く問題についての説明．

- 10 月 2 日
  九州大学情報基盤研究開発センターを学生と指導教員 42 名が訪問．スーパーコンピュータの講義を受講．また，研究施設や技術者の仕事などを見学．

- 10 月 17 日
  センター職員 2 名が大分工業高等専門学校を訪問．特別活動の時間に最新のコンピュータ技術についての講演会を実施．対象の学生は，3 年生 4 クラス（160 名）．

【参考資料】

「サイエンス・パートナーシップ・プロジェクト」報告書, スーパーコンピュータを用いた計算機実験, 大分工業高等専門学校 制御情報工学科, 2008 年 3 月．

5.4 国際会議 SC07 Exhibition における研究開発展示

SC07

SC は毎年米国で開催されている国際会議で，米国 ACM (American Computer and Machinary)学会が主催しています。参加者は 7,000～9,000 人と非常に多く，また参加する研究機関および企業も世界中から集まる，情報関連分野では世界最大規模のイベントです。SC07 は，米国ネヴァダ州にあるリノ(Reno)という町で開催されました。表 1 に SC07 の会議場や会期，および九州大学ブースの情報を記載します。

表 1 ：SC07 概要および九州大学ブース情報

SC07 Web Site	http://sc07.supercomputing.org/
場所	Reno-Sparks Convention Center (Reno, Nevada, USA)
SC07 期間	2007 年 11 月 10 日（土）～16 日（金）
Exhibition 期間	2007 年 11 月 12 日（月）～15 日（木）
九大ブース	No. 3119 (20x20 フィート)

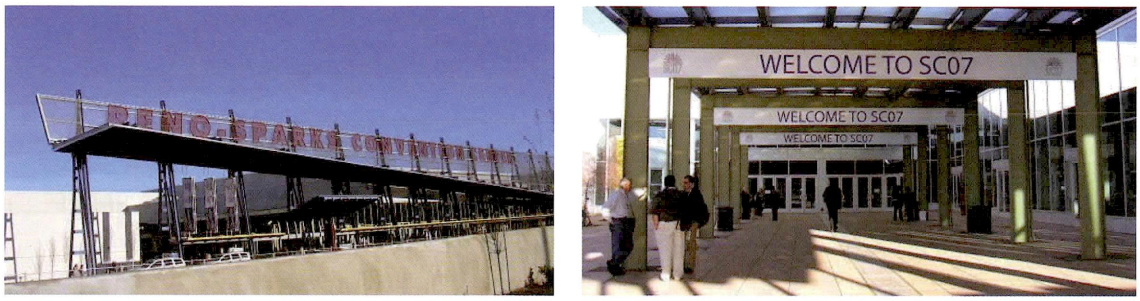


図 1 ：Reno-Sparks Convention Center

SC の正式名称は現在，「The International Conference for High Performance Computing, Network, Storage and Analysis」となっており，この正式名称が示すように，会議トピックは大規模計算や高性能の計算機，ネットワークやストレージと多岐にわたっています。しかし，SC07 Web サイトの URL (http://sc07.supercomputing.org/) が示すように，以前の名称である「Super Computing」が最大のトピックであることは変わりありません。

SC には主な行事として，研究論文の発表会(Technical Program)，最新技術に関する講習会(Tutorial)，および展示会(Exhibition)があります。他に招待講演や，パネルディスカッション，ネットワークやストレージの実験イベントなどが行われます。展示会では各センターのサービス紹介や，学術機関の研究やプロジェクトの紹介，さらに関連企業の新製品が展示されています。

当センターは，改組前の情報基盤センターであった 2003 年から展示を行っています。2006 年は申込遅れと展示希望者多数のためブースを予約できませんでしたが，2007 年に展示を再開しました。2007 年は情報基盤研究開発センターになって初めての展示でした。

図２：展示会(Exhivision)の様子

展示ポスター

九州大学ブースでは，下記の６種の展示を行いました。

(1) Pioneering the Computing & Communication Services for Academic Studies in Japan

(2) NGN Campus Network

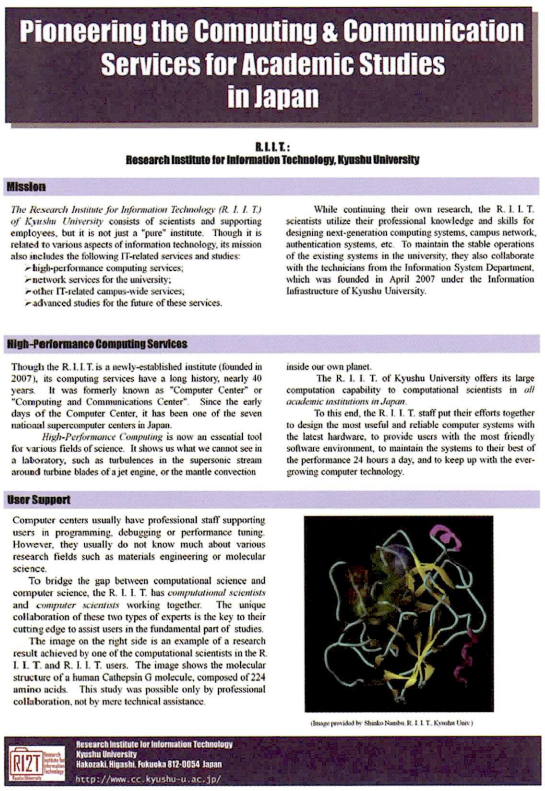
(3) User Authentication and Authorization in University

(4) Peta-Scale System Interconnect (PSI) Project

(5) Contribution to NAREGI Project (Work Package 6)

(6) Large Scale Reconfigurable Data Path Processor using Single Flux Quantum circuit

(1) Pioneering the Computing & Communication Services for Academic Studies in Japan



Pioneering the Computing & Communication Services for Academic Studies in Japan

R.I.I.T.:
Research Institute for Information Technology, Kyushu University

Mission

The Research Institute for Information Technology (R. I. I. T.) of Kyushu University consists of scientists and supporting employees, but it is not just a "pure" institute. Though it is related to various aspects of information technology, its mission also includes the following IT-related services and studies:

- high-performance computing services;
- network services for the university;
- other IT-related campus-wide services;
- advanced studies for the future of these services.

While continuing their own research, the R. I. I. T. scientists utilize their professional knowledge and skills for designing next-generation computing systems, campus network, authentication systems, etc. To maintain the stable operations of the existing systems in the university, they also collaborate with the technicians from the Information System Department, which was founded in April 2007 under the Information Infrastructure of Kyushu University.

High-Performance Computing Services

Though the R.I.I.T. is a newly-established institute (founded in 2007), its computing services have a long history, nearly 40 years. It was formerly known as "Computer Center" or "Computing and Communications Center". Since the early days of the Computer Center, it has been one of the seven national supercomputer centers in Japan.

High-Performance Computing is now an essential tool for various fields of science. It shows us what we cannot see in a laboratory, such as turbulences in the supersonic stream around turbine blades of a jet engine, or the mantle convection inside our own planet.

The R. I. I. T. of Kyushu University offers its large computation capability to computational scientists in *all academic institutions in Japan*.

To this end, the R. I. I. T. staff put their efforts together to design the most useful and reliable computer systems with the latest hardware, to provide users with the most friendly software environment, to maintain the systems to their best of the performance 24 hours a day, and to keep up with the ever-growing computer technology.

User Support

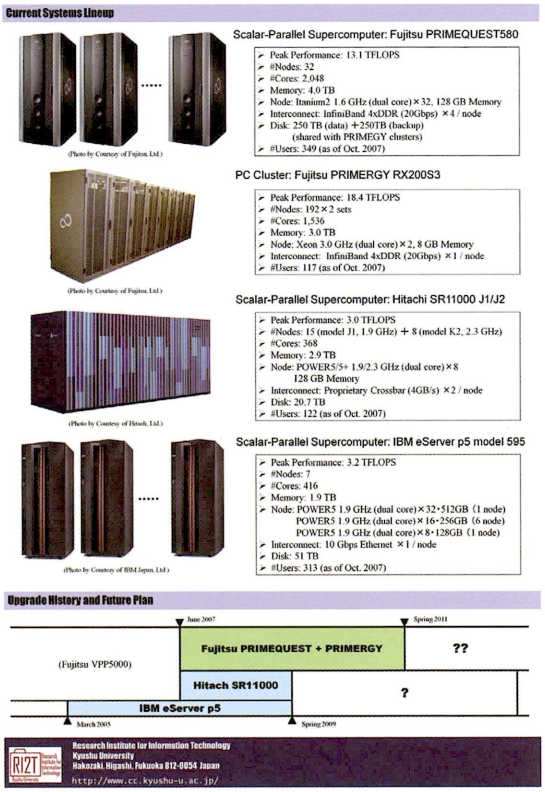
Computer centers usually have professional staff supporting users in programming, debugging or performance tuning. However, they usually do not know much about various research fields such as materials engineering or molecular science.

To bridge the gap between computational science and computer science, the R. I. I. T. has *computational scientists* and *computer scientists* working together. The unique collaboration of these two types of experts is the key to their cutting edge to assist users in the fundamental part of studies.

The image on the right side is an example of a research result achieved by one of the computational scientists in the R. I. I. T. and R. I. I. T. users. The image shows the molecular structure of a human Cathepsin G molecule, composed of 224 amino acids. This study was possible only by professional collaboration, not by mere technical assistance.

(Image provided by Shinko Nambu, R. I. I. T., Kyushu Univ.)

Research Institute for Information Technology
Kyushu University
Hakozaki, Higashi, Fukuoka 812-0054 Japan
http://www.cc.kyushu-u.ac.jp/



Current Systems Lineup

Scalar-Parallel Supercomputer: Fujitsu PRIMEQUEST580

- Peak Performance: 13.1 TFLOPS
- #Nodes: 32
- #Cores: 2,048
- Memory: 4.0 TB
- Node: Itanium2 1.6 GHz (dual core)×32, 128 GB Memory
- Interconnect: InfiniBand 4xDDR (20Gbps) ×4 / node
- Disk: 250 TB (data) +250TB (backup)
  (shared with PRIMEGY clusters)
- #Users: 349 (as of Oct. 2007)

(Photo by Courtesy of Fujitsu, Ltd.)

PC Cluster: Fujitsu PRIMERGY RX200S3

- Peak Performance: 18.4 TFLOPS
- #Nodes: 192×2 sets
- #Cores: 1,536
- Memory: 3.0 TB
- Node: Xeon 3.0 GHz (dual core)×2, 8 GB Memory
- Interconnect: InfiniBand 4xDDR (20Gbps) ×1 / node
- #Users: 117 (as of Oct. 2007)

(Photo by Courtesy of Fujitsu, Ltd.)

Scalar-Parallel Supercomputer: Hitachi SR11000 J1/J2

- Peak Performance: 3.0 TFLOPS
- #Nodes: 15 (model J1, 1.9 GHz) + 8 (model K2, 2.3 GHz)
- #Cores: 368
- Memory: 2.9 TB
- Node: POWER5/5+ 1.9/2.3 GHz (dual core)×8
  128 GB Memory
- Interconnect: Proprietary Crossbar (4GB/s) ×2 / node
- Disk: 20.7 TB
- #Users: 122 (as of Oct. 2007)

(Photo by Courtesy of Hitachi, Ltd.)

Scalar-Parallel Supercomputer: IBM eServer p5 model 595

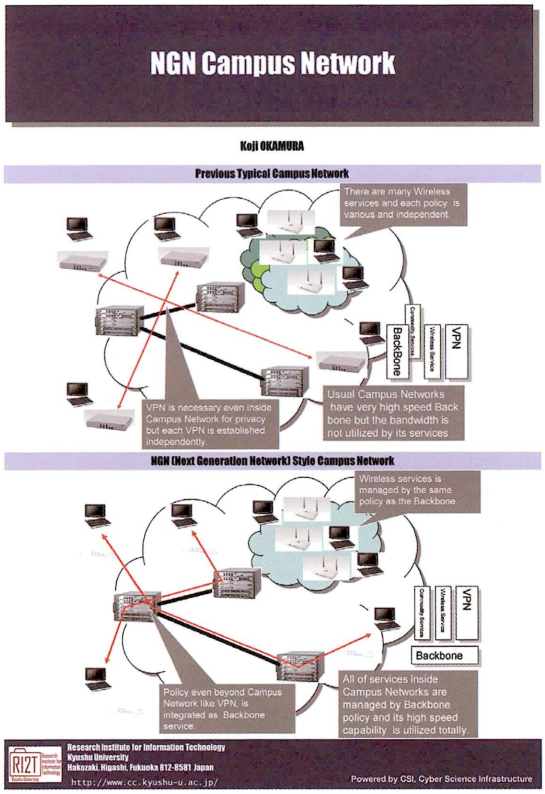
- Peak Performance: 3.2 TFLOPS
- #Nodes: 7
- #Cores: 416
- Memory: 1.9 TB
- Node: POWER5 1.9 GHz (dual core)×32・512GB (1 node)
  POWER5 1.9 GHz (dual core)×16・256GB (6 node)
  POWER5 1.9 GHz (dual core)×8・128GB (1 node)
- Interconnect: 10 Gbps Ethernet ×1 / node
- Disk: 51 TB
- #Users: 313 (as of Oct. 2007)

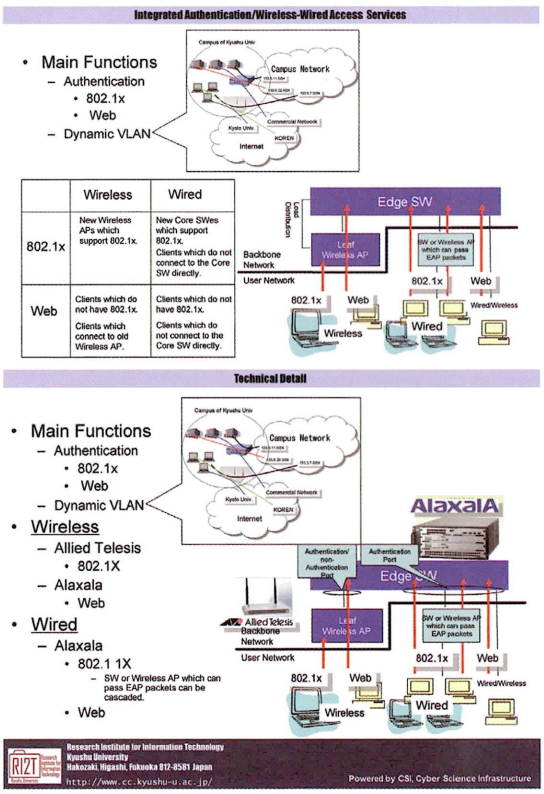
(Photo by Courtesy of IBM Japan, Ltd.)

Upgrade History and Future Plan

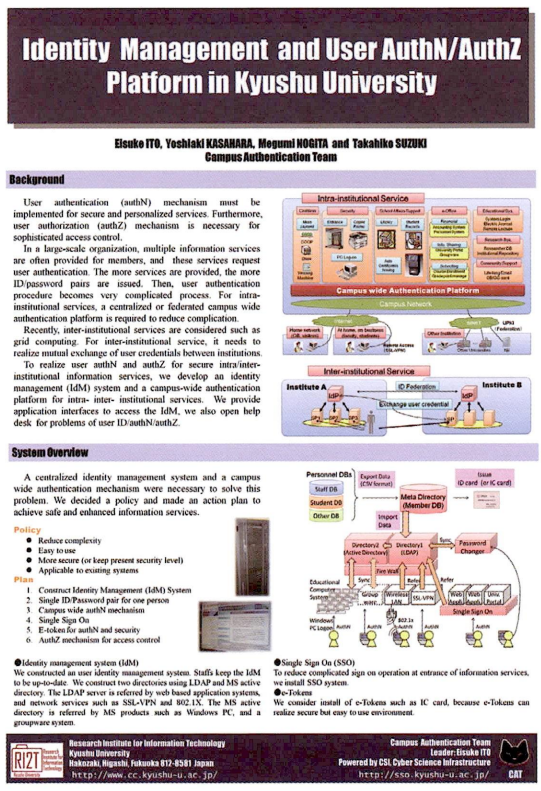
Research Institute for Information Technology
Kyushu University
Hakozaki, Higashi, Fukuoka 812-0054 Japan
http://www.cc.kyushu-u.ac.jp/

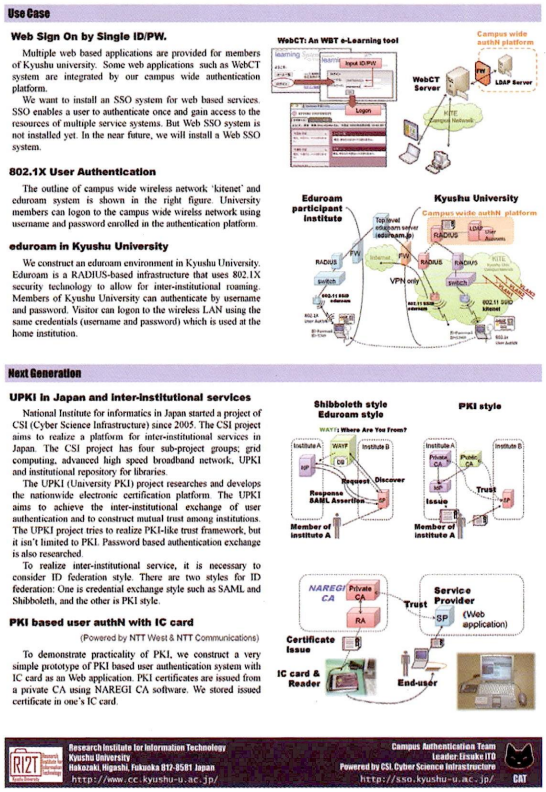
(2) NGN Campus Network





(3) Identity Management and User AuthN/AuthZ Platform in Kyushu University





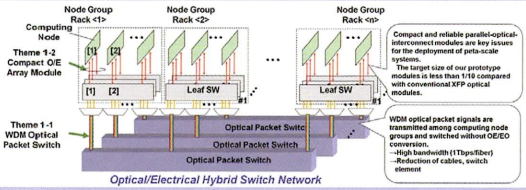
(4) Cutting the Edge of a Petascale Computing World (Peta-Scale System Interconnect Project)

Cutting the Edge of a Petascale Computing World

1: Ultra-High Capacity Interconnection by Photonic Switching Technology

Kyushu Univ., Fujitsu Limited, Fukuoka IST and ISIT Kyushu

Optical/Electrical Hybrid Switch Network for Peta-scale Interconnection



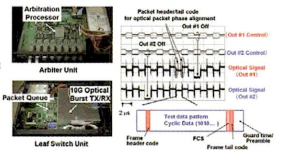
Optical/Electrical Hybrid Switch Network

Sub Theme 1-1 Development of Optical Packet Interconnect Switch System

2 x 2 optical packet interconnect system was developed and 1.2ms-length optical packet signal was successfully switched. The prototype optical packet system has the following functions;

(1) Arbitration unit for controlling the optical packet switch and Leaf switch to convert the electrical packet form computing node group into optical packet were developed. Optical packet signal scheduling and time-slot synchronization for optical packet forwarding is realized by using these functions.

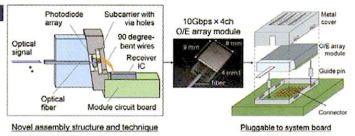
(2) Broadcast and select optical switch architecture using SOA (Semiconductor Optical Amplifier) was developed for the optical packet switch fabric. The switching speed of less than 10 nano-second, which is enough for 1.2ms-length optical packet switching, is obtained.



Sub Theme 1-2 Development of Compact O/E Array Module Technology

10Gbps ×4 channel O/E module

- **Small size (Low profile):**
  - ≦ 10 mm□(≦ 5 mm thick)
  - ⇒ High density packaging for Tbps class board-to-board optical interconnection
- **Low cost design**
  - ·Direct coupling to optical fiber
  - ·Bent wiring with subcarrier



Optical Components Unit
Fujitsu Limited
4-1-1 Kami Kodanaka, Nakahara-ku, Kawasaki 211-8588 Japan
http://www.fujitsu.com/ E-mail:photonic-ic@ml.labs.fujitsu.com

Peta-Scale System Interconnect Project
Leader: Kazuaki Murakami (Kyushu Univ.)
http://psi-project.jp

Cutting the Edge of a Petascale Computing World

2: Hardware & Software for High-Speed MPI

Kyushu Univ., Fujitsu Limited, Fukuoka IST and ISIT Kyushu

Background

Communication cost is an important factor on the performance of Peta-FLOPS computers, which consist of 10,000 to 100,000 nodes. Theme 2 works on research and development of technologies for implementing high-speed MPI (Message Passing Interface). MPI is a standard communication library for parallel programming. Especially, in theme 2, collective communications of MPI are the major targets and studied intensively.

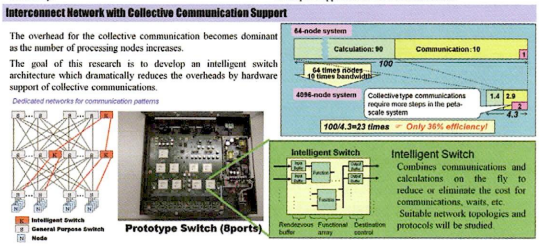
Collective communication is a type of communication such as Barrier, Broadcast, Gather, Scatter and Reduction, which involves a group of nodes. The effect of cost for collective communications becomes more significant as the number of node increases. Theme 2 is to reduce this cost by following technologies:

- Intelligent switch with collective communication support
- Adaptive approach for MPI Communication

Interconnect Network with Collective Communication Support

The overhead for the collective communication becomes dominant as the number of processing nodes increases.

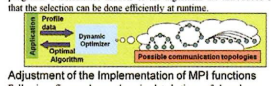
The goal of this research is to develop an intelligent switch architecture which dramatically reduces the overheads by hardware support of collective communications.



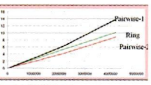
Adaptive Approach for MPI Communication

Runtime Selection of Collective Communication Algorithms

Choose the optimal algorithm of collective communication for the given situation of the parallel machine at runtime, such as the load-balance on each node or network collisions by other parallel programs. Performance models of the algorithm are introduced so that the selection can be done efficiently at runtime.

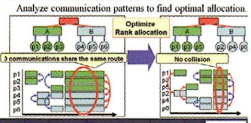


A new algorithm for MPI_Allgather: by considering the communication difference between inter- and intra-node



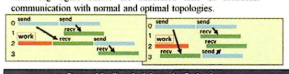
Rank Allocation to Avoid Collisions

Analyze communication patterns to find optimal allocation.



Adjustment of the Implementation of MPI functions

Following figure shows the simulated time of broadcast communication with normal and optimal topologies.



Research Institute for Information Technology
Kyushu University
Hakozaki, Higashi, Fukuoka 812-8581 Japan
http://www.cc.kyushu-u.ac.jp/

Peta-Scale System Interconnect Project
Leader: Kazuaki Murakami (Kyushu Univ.)
http://psi-project.jp

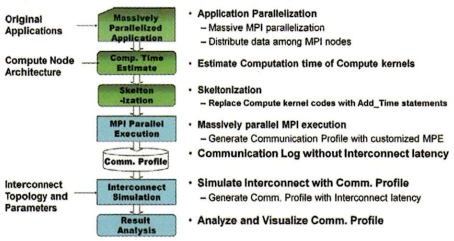
Cutting the Edge of a Petascale Computing World

3: Petascale System Evaluation Methodology

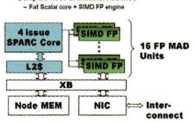
Kyushu Univ., Fujitsu Limited, Fukuoka IST and ISIT Kyushu

Peta-Scale System Evaluation Methodology for Interconnect Design

- **Divide & Conquer Approach**
  - –Separate Node compute time and Interconnect latency



- **Application Parallelization**
  - – Two dimensional parallelization
  - – Distributed Memory allocation
  - – HPL: original code is parallelized
  - – PHASE (MO for Solids)
    - Parallelize with wave functions and reciprocal lattice numbers
    - Over 50K nodes for 10,000 atoms
  - – OpenFMO (Fragment MO for Proteins)
    - Parallelize with electron structure calculations of fragments
    - Use MPI_GET, MPI_PUT for distributed data access
- **Skeltonization**
  - – For 100x or more faster execution speed
    - Replace Computer kernel code with Add_Time statement
    - Add_Time advances application "clock" as if Kernel code is executed
  - – Reduce Memory usage to 1/100 or less
    - Allows to run 100 or more node executions on each server node with installed DRAM memory
    - Future 100K node Peta-Scale system can be simulated with 1024 node server
- **Computation time Estimation**
  - – For Existing Systems
    - Run single node application and measure
  - – For Future Peta-Scale Systems
    - Compute node architecture defined
    - Fat Scalar core + SIMD FP engine



- Tools and Library functions developed
  - – Architecture Simulator
  - – SIMD Compiler
  - – dgemm, dgemm, fft, etc.
- Simulate Compute Kernel for Computation time estimate

Research Institute for Information Technology
Kyushu University
Hakozaki, Higashi, Fukuoka 812-8581 Japan
http://www.cc.kyushu-u.ac.jp/

Peta-Scale System Interconnect Project
Leader: Kazuaki Murakami (Kyushu Univ.)
http://psi-project.jp

Cutting the Edge of a Petascale Computing World

3: Petascale System Evaluation Methodology

Kyushu Univ., Fujitsu Limited, Fukuoka IST and ISIT Kyushu

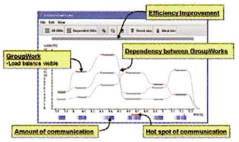
Peta-Scale System Evaluation Methodology for Interconnect Design

- **Communication Profile Generation**
  - – Execute parallel application on real machine
    - Real execution
    - Skeleton execution
  - – Generate communication profile
    - Suppose execution result on ideal network without interconnect latency
    - Include information about virtual elapsed times, send/receive ranks, etc.
  - – Toward Peta-scale system
    - Support a large number of processes (1,000 ~ >10,000)
    - Run a large number of processes on a small number of nodes
- **Analysis and Visualization**
  - – Tuning support
    - Concept of GroupWork

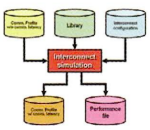
- Advice on program optimization:

- Discover network congestion
- Overview of communication traffic

  - – Efficient visualization tool

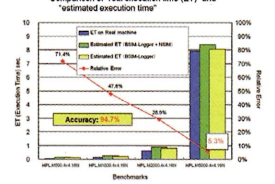


- **Interconnect Simulation**
  - – Generate new profile with interconnect latency



- Inputs - comm. profile without interconnect latency
  - interconnect configuration file
- Outputs - comm. profile with interconnect latency
  - performance file (estimated execution time, send/receive ranks, message size, hops, and collision frequency, etc)
  - – Ready for Peta-scale interconnect
    - Complete simulation in reasonable time
  - – Scalable and flexible evaluation
    - Various simulations based on interconnect specifications
- **Performance Estimation of Benchmarks**
  - – Profile generation + Interconnect simulation

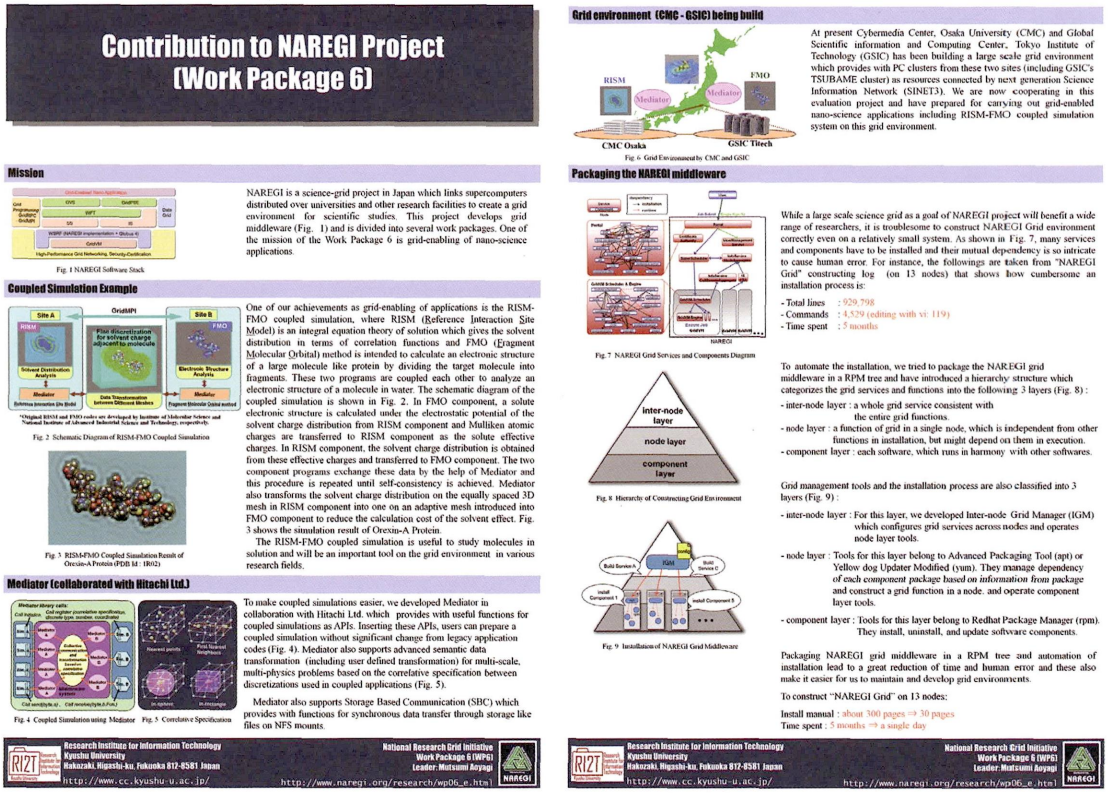
Comparison of "real execution time (ET)" and "estimated execution time"



Research Institute for Information Technology
Kyushu University
Hakozaki, Higashi, Fukuoka 812-8581 Japan
http://www.cc.kyushu-u.ac.jp/

Peta-Scale System Interconnect Project
Leader: Kazuaki Murakami (Kyushu Univ.)
http://psi-project.jp

(5) Contribution to NAREGI Project (Work Package 6)



Contribution to NAREGI Project (Work Package 6)

Mission

NAREGI is a science-grid project in Japan which links supercomputers distributed over universities and other research facilities to create a grid environment for scientific studies. This project develops grid middleware (Fig. 1) and is divided into several work packages. One of the mission of the Work Package 6 is grid-enabling of nano-science applications.

Fig. 1 NAREGI Software Stack

Coupled Simulation Example

One of our achievements as grid-enabling of applications is the RISM-FMO coupled simulation, where RISM (Reference Interaction Site Model) is an integral equation theory of solution which gives the solvent distribution in terms of correlation functions and FMO (Fragment Molecular Orbital) method is intended to calculate an electronic structure of a large molecule like protein by dividing the target molecule into fragments. These two programs are coupled each other to analyze an electronic structure of a molecule in water. The schematic diagram of the coupled simulation is shown in Fig. 2. In FMO component, a solute electronic structure is calculated under the electrostatic potential of the solvent charge distribution from RISM component and Mulliken atomic charges are transferred to RISM component as the solute effective charges. In RISM component, the solvent charge distribution is obtained from these effective charges and transferred to FMO component. The two component programs exchange these data by the help of Mediator and this procedure is repeated until self-consistency is achieved. Mediator also transforms the solvent charge distribution on the equally spaced 3D mesh in RISM component into one on an adaptive mesh introduced into FMO component to reduce the calculation cost of the solvent effect. Fig. 3 shows the simulation result of Orexin-A Protein.

The RISM-FMO coupled simulation is useful to study molecules in solution and will be an important tool on the grid environment in various research fields.

Fig. 2 Schematic Diagram of RISM-FMO Coupled Simulation

Fig. 3 RISM-FMO Coupled Simulation Result of Orexin-A Protein (PDB Id : 1R02)

Mediator (collaborated with Hitachi Ltd.)

To make coupled simulations easier, we developed Mediator in collaboration with Hitachi Ltd. which provides with useful functions for coupled simulations as APIs. Inserting these APIs, users can prepare a coupled simulation without significant change from legacy application codes (Fig. 4). Mediator also supports advanced semantic data transformation (including user defined transformation) for multi-scale, multi-physics problems based on the correlative specification between discretizations used in coupled applications (Fig. 5).

Mediator also supports Storage Based Communication (SBC) which provides with functions for synchronous data transfer through storage like files on NFS mounts.

Fig. 4 Coupled Simulation using Mediator

Fig. 5 Correlative Specification

Research Institute for Information Technology, Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812-8581 Japan
http://www.cc.kyushu-u.ac.jp/

National Research Grid Initiative, Work Package 6 (WP6), Leader: Matsumi Aoyagi
http://www.naregi.org/research/wp06_e.html

Grid environment (CMC - GSIC) being build

Fig. 6 Grid Environment by CMC and GSIC

At present Cybermedia Center, Osaka University (CMC) and Global Scientific information and Computing Center, Tokyo Institute of Technology (GSIC) has been building a large scale grid environment which provides with PC clusters from these two sites (including GSIC's TSUBAME cluster) as resources connected by next generation Science Information Network (SINET3). We are now cooperating in this evaluation project and have prepared for carrying out grid-enabled nano-science applications including RISM-FMO coupled simulation system on this grid environment.

Packaging the NAREGI middleware

Fig. 7 NAREGI Grid Services and Components Diagram

While a large scale science grid as a goal of NAREGI project will benefit a wide range of researchers, it is troublesome to construct NAREGI Grid environment correctly even on a relatively small system. As shown in Fig. 7, many services and components have to be installed and their mutual dependency is so intricate to cause human error. For instance, the followings are taken from "NAREGI Grid" constructing log (on 13 nodes) that shows how cumbersome an installation process is:

- Total lines : 929,798
- Commands : 4,529 (editing with vi: 119)
- Time spent : 5 months

Fig. 8 Hierarchy of Constructing Grid Environment

To automate the installation, we tried to package the NAREGI grid middleware in a RPM tree and have introduced a hierarchy structure which categorizes the grid services and functions into the following 3 layers (Fig. 8) :

- inter-node layer : a whole grid service consistent with the entire grid functions.
- node layer : a function of grid in a single node, which is independent from other functions in installation, but might depend on them in execution.
- component layer : each software, which runs in harmony with other softwares.

Fig. 9 Installation of NAREGI Grid Middleware

Grid management tools and the installation process are also classified into 3 layers (Fig. 9) :

- inter-node layer : For this layer, we developed Inter-node Grid Manager (IGM) which configures grid services across nodes and operates node layer tools.
- node layer : Tools for this layer belong to Advanced Packaging Tool (apt) or Yellow dog Updater Modified (yum). They manage dependency of each component package based on information from package and construct a grid function in a node. and operate component layer tools.
- component layer : Tools for this layer belong to Redhat Package Manager (rpm). They install, uninstall, and update software components.

Packaging NAREGI grid middleware in a RPM tree and automation of installation lead to a great reduction of time and human error and these also make it easier for us to maintain and develop grid environments.

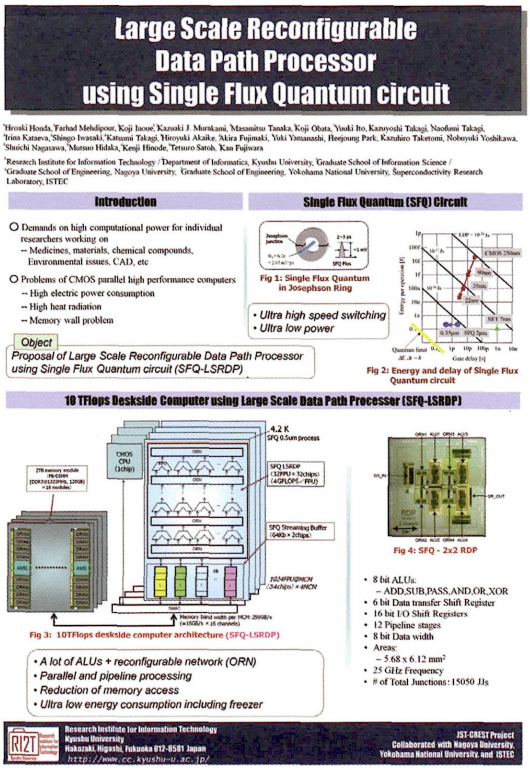
To construct "NAREGI Grid" on 13 nodes:

Install manual : about 300 pages ⇒ 30 pages
Time spent : 5 months ⇒ a single day

Research Institute for Information Technology, Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812-8581 Japan
http://www.cc.kyushu-u.ac.jp/

National Research Grid Initiative, Work Package 6 (WP6), Leader: Matsumi Aoyagi
http://www.naregi.org/research/wp06_e.html

(6) Large Scale Reconfigurable Data Path Processor using Single Flux Quantum circuit



Large Scale Reconfigurable Data Path Processor using Single Flux Quantum circuit

[1]Hiroaki Honda, [1]Farhad Mehdipour, [1]Koji Inoue, [1]Kazuaki J. Murakami, [1]Masanitsu Tanaka, [1]Koji Obata, [1]Yuuki Ito, [1]Kanyouhi Takagi, [1]Naofumi Takagi, [2]Irina Kataeva, [2]Shingo Iwasaki, [2]Katsumi Takagi, [2]Hiroyuki Akaike, [2]Akira Fujimaki, [3]Yuki Yamanashi, [3]Jeejoung Park, [3]Kazuhiro Taketomi, [3]Nobuyuki Yoshikawa, [3]Shuichi Nagasawa, [4]Mutsuo Hidaka, [4]Kenji Hinode, [4]Tetsuro Satoh, [4]Kan Fujiwara

[1]Research Institute for Information Technology / [2]Department of Informatics, Kyushu University, [3]Graduate School of Information Science / [4]Graduate School of Engineering, Nagoya University, [3]Graduate School of Engineering, Yokohama National University, [3]Superconductivity Research Laboratory, ISTEC

Introduction

- Demands on high computational power for individual researchers working on
  - -- Medicines, materials, chemical compounds, Environmental issues, CAD, etc
- Problems of CMOS parallel high performance computers
  - -- High electric power consumption
  - -- High heat radiation
  - -- Memory wall problem

Object

Proposal of Large Scale Reconfigurable Data Path Processor using Single Flux Quantum circuit (SFQ-LSRDP)

Single Flux Quantum (SFQ) Circuit

Fig 1: Single Flux Quantum in Josephson Ring

- *Ultra high speed switching*
- *Ultra low power*

Fig 2: Energy and delay of Single Flux Quantum circuit

10 TFlops Deskside Computer using Large Scale Data Path Processer (SFQ-LSRDP)

Fig 3: 10TFlops deskside computer architecture (SFQ-LSRDP)

Fig 4: SFQ - 2x2 RDP

- 8 bit ALUs
  - -- ADD,SUB,PASS,AND,OR,XOR
- 6 bit Data transfer Shift Register
- 16 bit I/O Shift Registers
- 12 Pipeline stages
- 8 bit Data width
- Areas:
  - -- 5.68 x 6.12 mm²
- 25 GHz Frequency
- # of Total Junctions: 15050 JJs

- *A lot of ALUs + reconfigurable network (ORN)*
- *Parallel and pipeline processing*
- *Reduction of memory access*
- *Ultra low energy consumption including freezer*

Research Institute for Information Technology, Kyushu University, Hakozaki, Higashi, Fukuoka 812-8581 Japan
http://www.cc.kyushu-u.ac.jp/

JST-CREST Project, Collaborated with Nagoya University, Yokohama National University, and ISTEC

おわりに

SC07 展示会では九大ブースへ多数の訪問がありました。参加正確な訪問者数は把握していないものの，日本から用意した 150 部の資料は全部配布しているので，150 名以上の方がブースへ訪問していただいたと考えております。また，当センターの SC07 参加者も，他のブースから様々な最新の研究開発動向および企業の最新製品を知ることができ，今後の研究開発およびセンターサービスのための知見を得ていると思います。SC の展示会参加はセンター活動のまとめや，デモシステム作成の契機になるなど，良い影響を与える機会になっていると思います。

次の SC08 は，米国テキサス州オースチンにある「Austin Convention Center」にて開催予定です。SC08 でも九州大学ブースとして展示予定です。

表 2 ：SC08 概要および展示予定ブース

SC08 Web Site	http://sc08.supercomputing.org/
場所	Austin Convention center (Austin, Texas, USA)
SC08 期間	2008 年 11 月 15 日（土）～21 日（金）
Exhibition 期間	2008 年 11 月 17 日（月）～20 日（木）
九大ブース	No. 2815 (20x20 フィート)