

喜連川・豊田・根本・吉永・合田
研究室

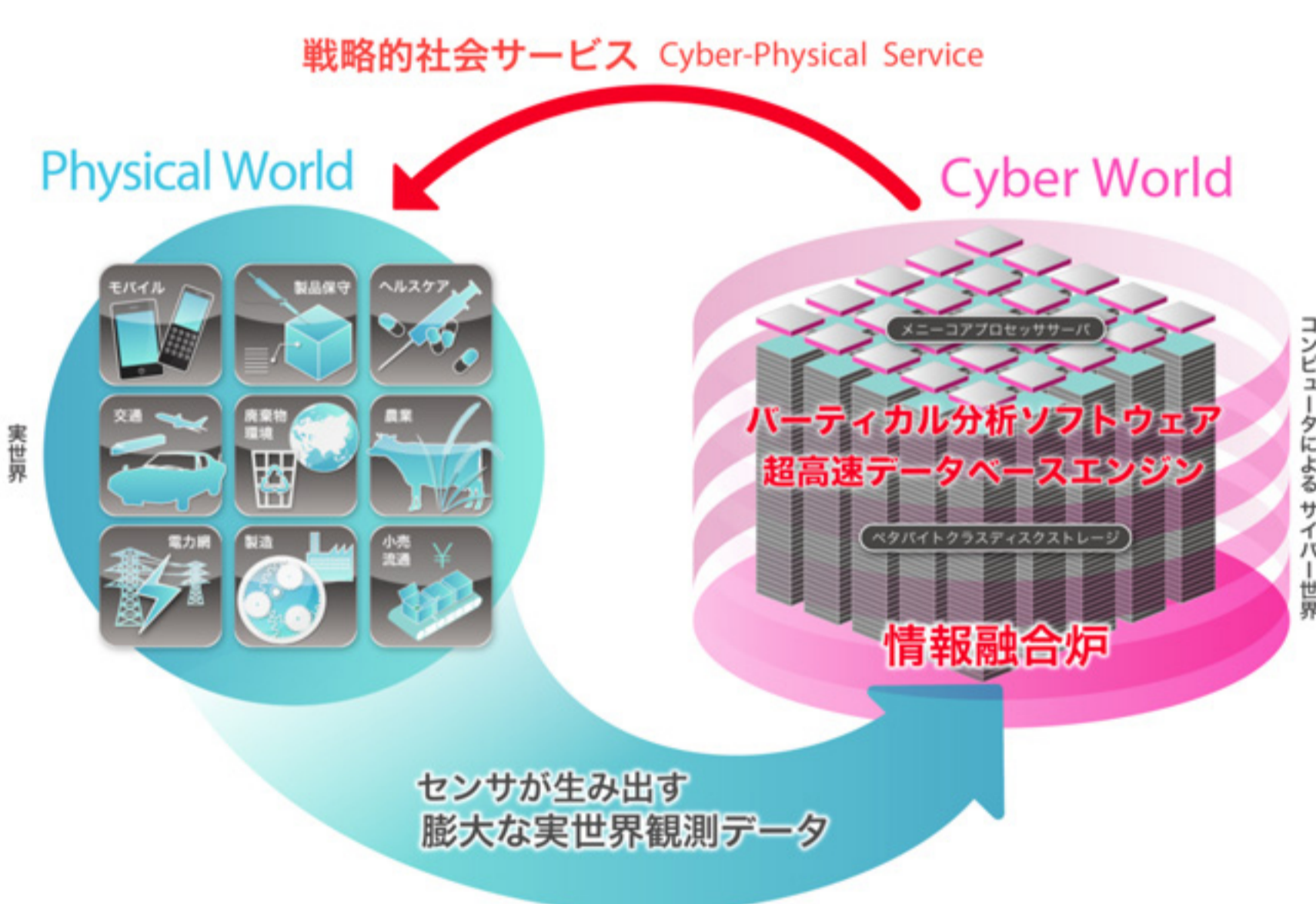
Kitsuregawa, Toyoda,

URL: <http://www.tkl.iis.u-tokyo.ac.jp/>



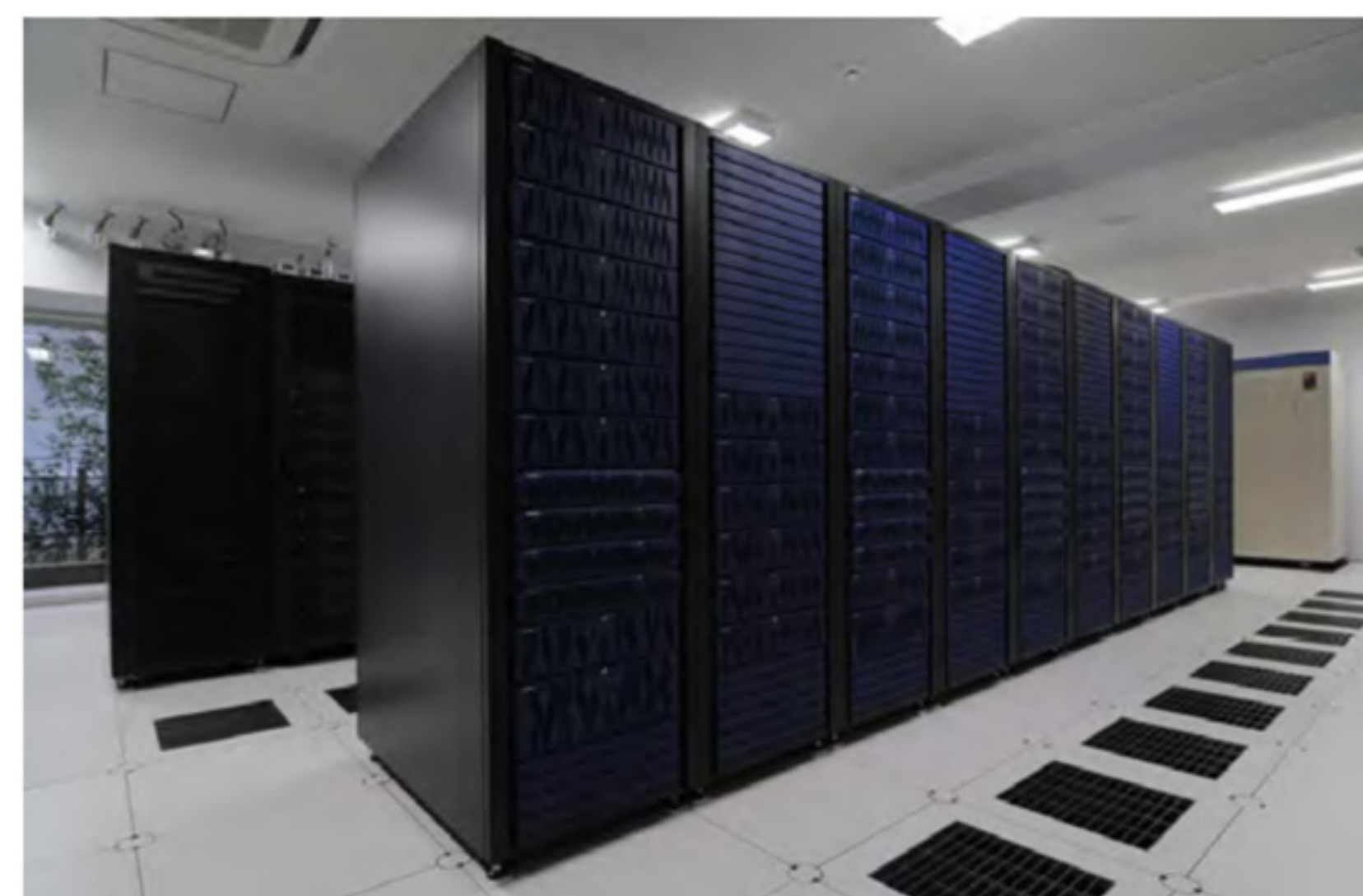
超巨大データベース時代に向けた最先端データベースエンジンの開発

「非順序型実行原理」なる新しいソフトウェア実行原理に基づく超高速データベースエンジン（大規模情報を扱うための基盤ソフトウェア）の開発を進めている。ビッグデータ時代のクエリ処理において、現行システム比で飛躍的な高速化を実現しつつある。また、当該データベースエンジンを核として、今までにない先進的なビッグデータ・IoT応用を創出するべく、実証システムの構築に取り組んでいる。加えて、新たな方向性として、データベースエンジンの省エネルギー化を進めている。



ペタバイト級地球環境情報融合システム

各種地球環境情報のデジタルライブラリ化と国際的ポータル構築を進めている。また、大規模データアーカイブ、メタデータ管理等、近年注目されている大規模データ処理のシステム技術に関する研究を進めている。様々な地球観測データを統合・解析することにより、科学的・社会的に有用な情報に変換できるシステムの実現を目指している。

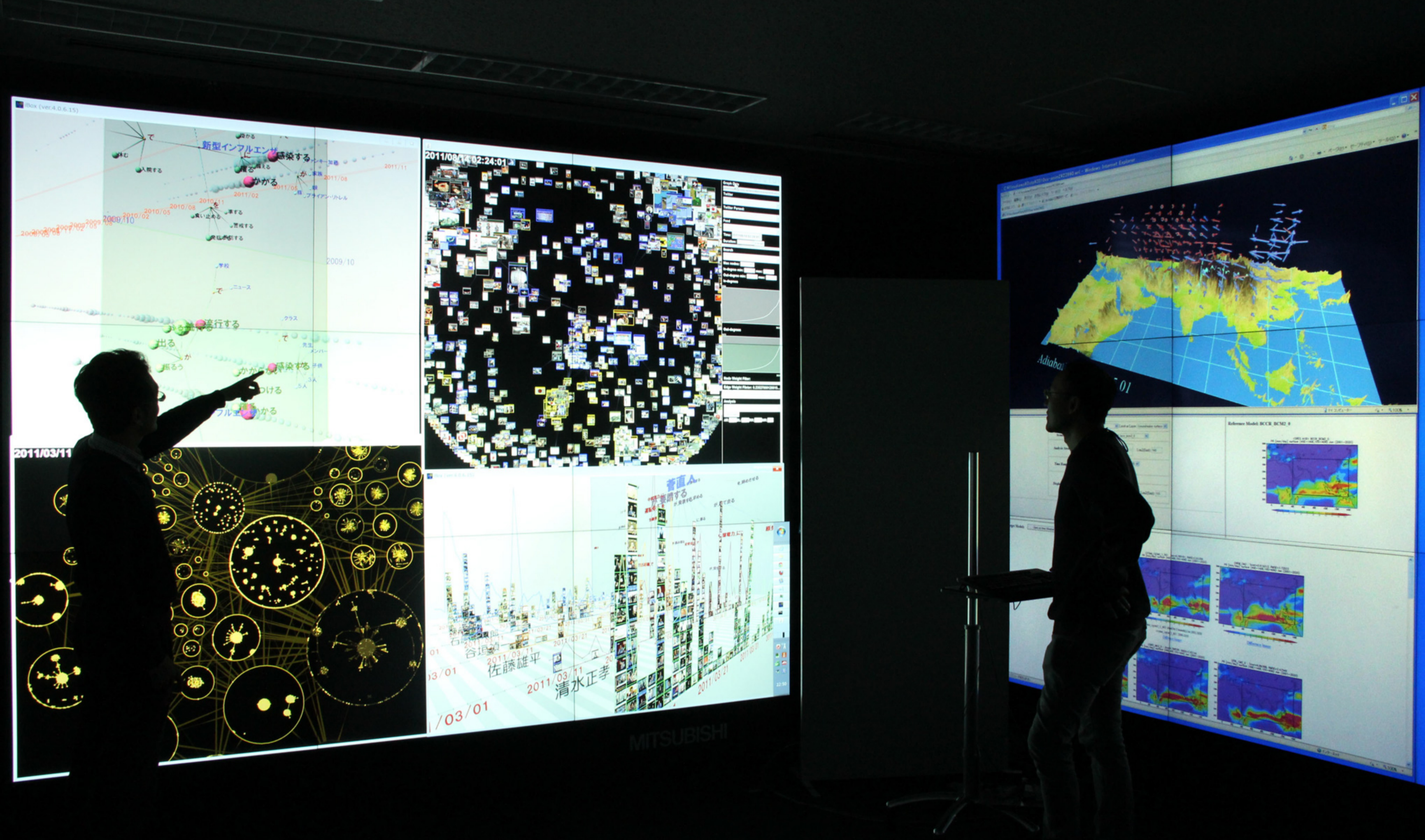


超大規模データ処理

ビッグデータを価値へ転換する情報エネルギー生成基盤の構築を目指し、最先端大規模ハードウェアを活用できるシステムソフトウェアから先進アプリケーションまでを広範に研究しています。モットーは「良く学び、良く遊ぶ」

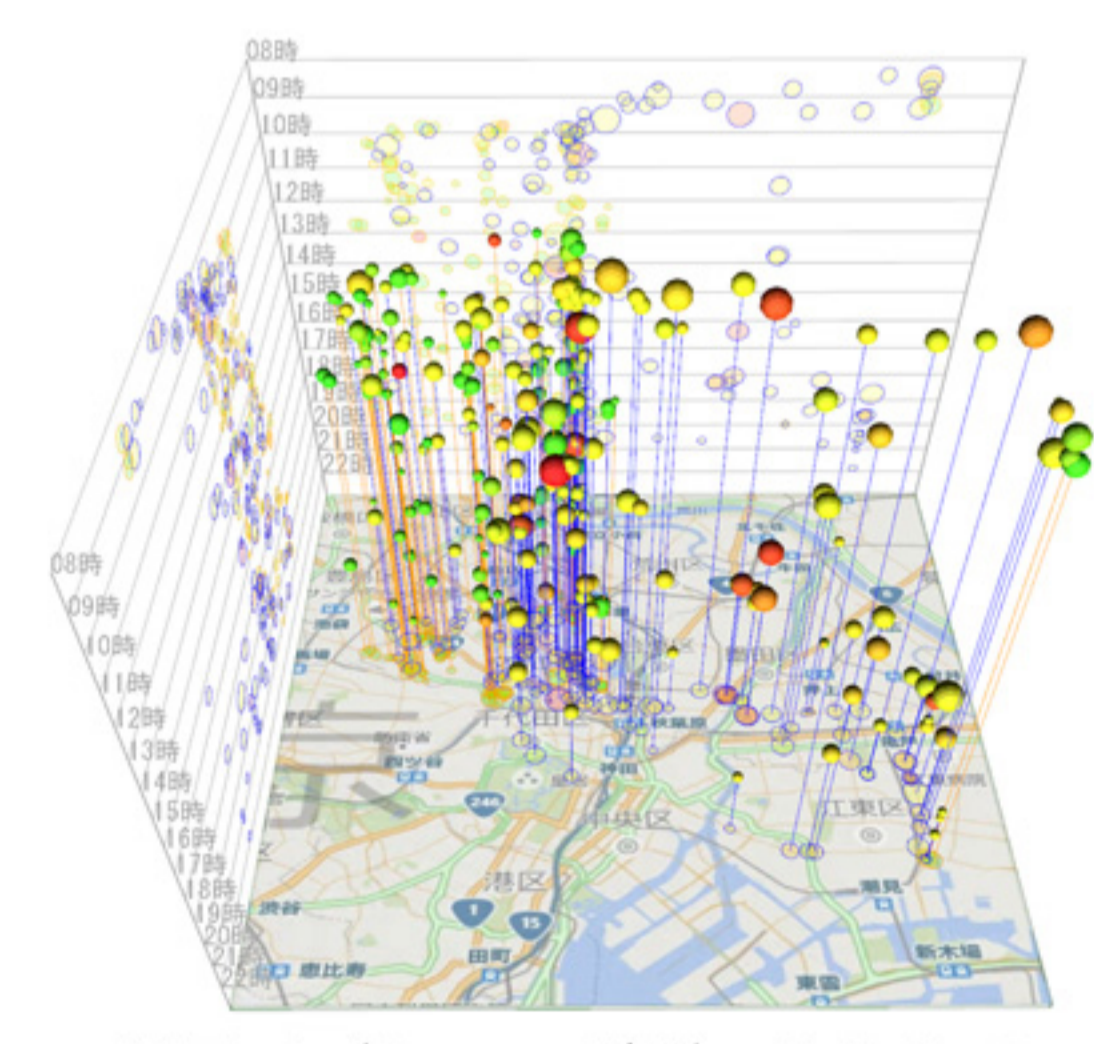
生産技術研究所Ee-505 IIS, Ee-505

本研究室では、データベース技術を基盤とし、非順序実行原理なる新方式に基づく超高速最先端データベースエンジン、ペタバイト級超大規模地球環境デジタルライブラリ、超大規模サイバー空間・実世界データのインタラクティブ解析システム、ソーシャル・ビッグデータ利活用のための言語処理基盤を始めとし、データ工学における諸課題に対し、『誰も見たことが無い、ドキッとする』システムソフト、先進アプリ、ハード、アルゴリズムの研究を進めている。どこかで見たことがある研究はしない。

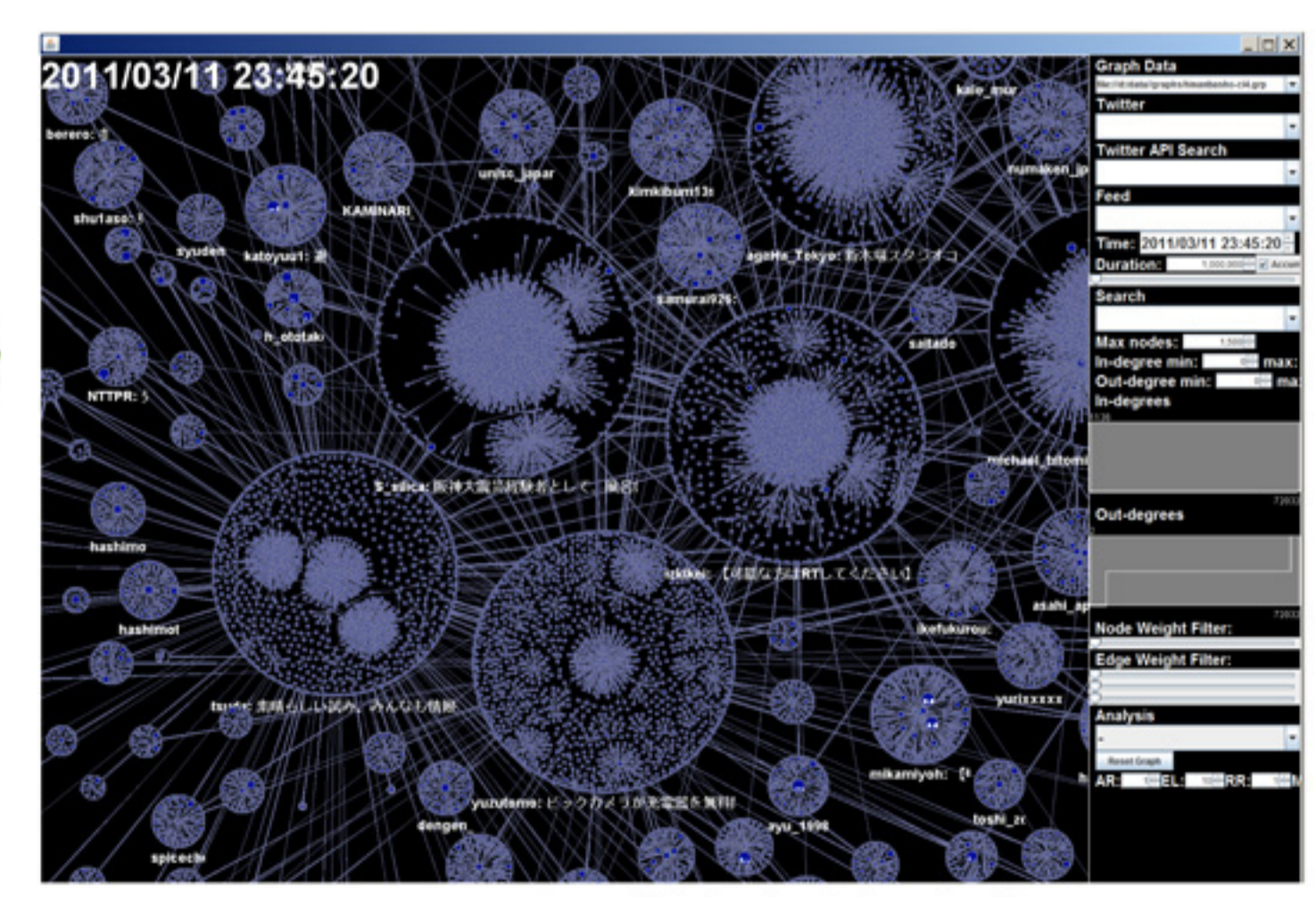


超大規模サイバー空間・実世界データのインタラクティブ解析システム

ウェブ・ソーシャルメディア等のサイバー空間と実世界は密接に連動しており、サイバー空間と実世界センサデータの融合解析による社会課題解決を目標とした研究を推進している。1999年から継続的に日本語ウェブページを大規模収集し、数百億URL、数十億ブログ記事、Twitterの数百億つぶやき等を含むウェブアーカイブを構築するとともに、ドライブレコーダデータ、交通トラフィックデータ、気象データ等の実世界データの収集・蓄積を行い、その構造、内容、時間変化等を解析するシステムを開発中である。膨大なサイバー空間・実世界データを、データマイニング、機械学習、リンク解析、自然言語処理、画像処理等を用いて解析し、様々な切り口で探索可能な可視化システムを大規模ディスプレイウォール上に実装している。



ドライブレコーダデータからのヒヤリハット地図生成



Twitter上の情報伝播可視化

超大規模かつ多様なソーシャル・ビッグデータ利活用のための言語処理基盤

マイクロブログ等を通して人々が発信する莫大で多様なソーシャル・ビッグデータを人間の代わりに読み解き、社会に役立つ知見を得るため、実際的な自然言語処理技術の研究を進めている。具体的に、超高速・高精度で適応的な基礎言語解析技術の開発、実世界テキストから得られた知識を矛盾なく集積するための方法論の提唱を行い、言語情報を基に非言語情報を整理して構造化するフレームワークの構築に取り組んでいる。一方で、超大規模データを最大限活用して、機械翻訳や対話システムなど自然言語を介した知的アプリケーションの高度化も進めている。

ソーシャル・ビッグデータ

- 7000万 ユーザ
- 3億 ユーザ
- 128万 ユーザ
- Wikipedia 2001-ウェブ 1993-
- Twitter 2006-
- ブログ 2002-

社会における個人の意見・体験

- 災害の報告 (竜巻・洪水)
- 不正の指摘 (STAP・五輪ロコ)
- 要望・クレーム (品切れ・異物混入)
- 評判 (新商品・政策)
- 実況 (イベント・デモ)

Real-Time Parsing / tweet+network

時間	ツイート数	単語
2011/03/11 14:00 - 16:00	36907	1824 地震 750 余震 514 津波 120 飲酒 81 火事 51 水災 39 通電止
2011/03/11 16:00 - 18:00	31028	986 余震 873 地震 512 津波 142 飲酒 70 二次災害 53 水災 40
2011/03/11 18:00 - 20:00	20857	702 地震 578 余震 214 津波 93 飲酒 71 火災 59 二次災害 46 水災 38
2011/03/11 20:00 - 22:00	19288	816 余震 762 地震 243 津波 84 飲酒 68 飲酒 45 水災 38 停電 30
2011/03/11 22:00 - 00:00	26359	931 余震 682 地震 511 津波 450 外出 118 飲酒 75 水災 49 火災 1
2011/03/12 00:00 - 02:00	25726	1114 余震 470 地震 291 外出 262 津波 80 運転手さん 60 飲酒 36 水災 1
2011/03/12 02:00 - 04:00	10980	381 余震 211 地震 199 外出 171 外出 96 津波 50 水災 41 火災 1
2011/03/12 04:00 - 06:00	19261	428 余震 360 余震 338 地震 252 外出 73 緊急地震速報 69 津波 37 地震 1
2011/03/12 06:00 - 08:00	8313	190 余震 190 地震 114 水災 93 津波 62 停電 60 外出 38 緊急地震速報

東日本大震災発生日における恐怖の対象の推移



塙研究室(Prof. Toshihiro Hanawa) Hanawa Laboratory

URL: <http://www.cspp.cc.u-tokyo.ac.jp/hanawa/>

情報基盤センター(柏・浅野)
Information Technology Center (Kashiwa/Asano)

演算加速機構と通信の融合・協調による次世代スパコン技術

- 動的再構成可能デバイスを用いて、演算加速装置間の高速ノード間通信と同時に、通信中のデータに対して軽量の演算処理を行うことで、システム全体での性能向上を目指す
- ポストムーア時代のスパコン要素技術につながる研究開発

密結合演算加速機構 (Tightly Coupled Accelerators: TCA) アーキテクチャ

GPU に代表される演算加速機構 (アクセラレータ) は演算性能が極めて高く、比較的消費電力が少ないため、科学技術計算において広く使われています。

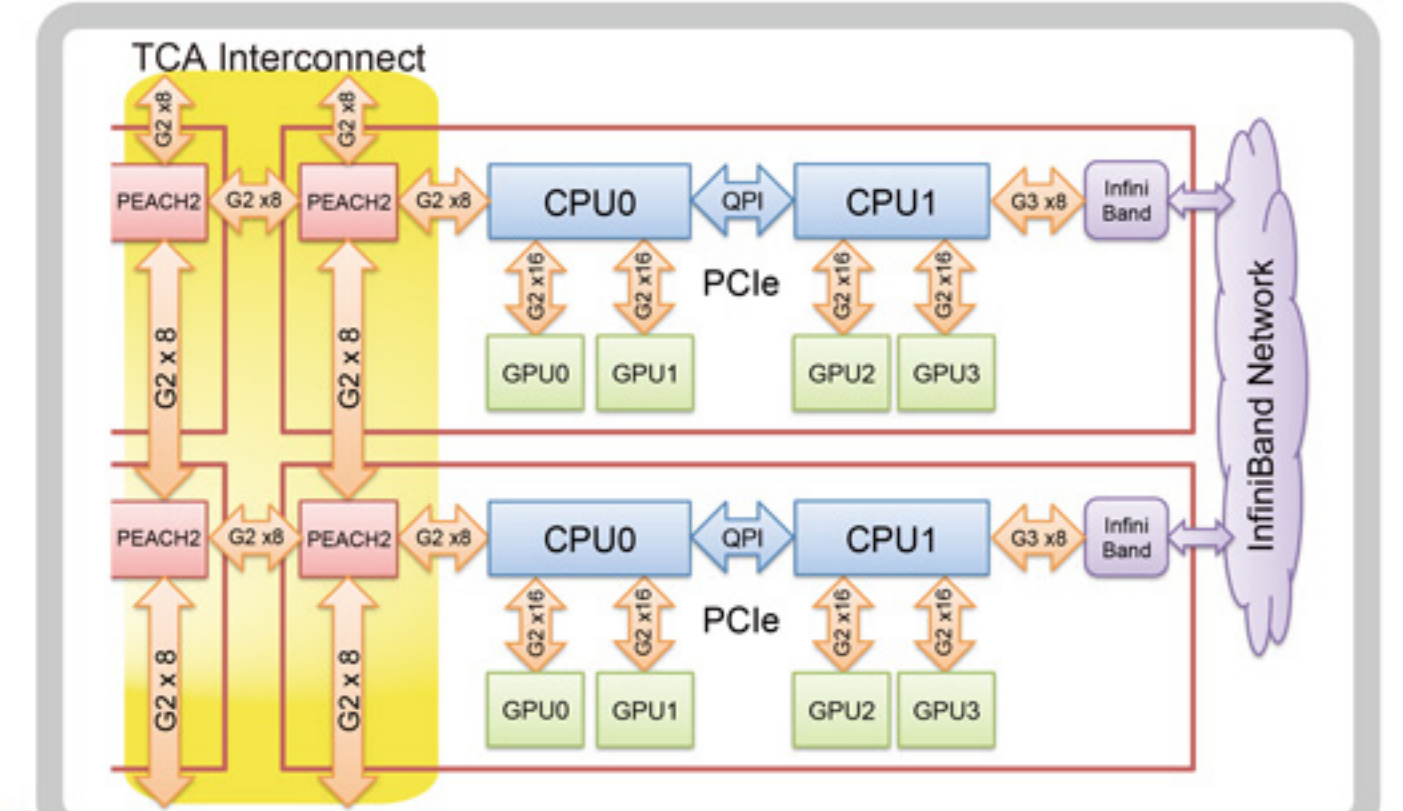
しかし演算加速機構は、ホストに PCI Express などの IO バスで接続されて使われるため IO バンド幅の制限が性能に与える影響が大きくなります。特に、ノードを超えた演算加速機構間の通信の際には、ホストとの間でのメモリコピーが通信遅延の増加を引き起こします。その問題を解決するため、計算ノードをまたぐアクセラレータ間の直接通信を実現する **TCA (Tightly Coupled Accelerators)** を提案しています。

PEACH2 (PCI Express Adaptive Communication Hub ver. 2) チップは、TCA を実現するための通信チップです。設計の柔軟性・開発コストなどの点から FPGA (Field Programmable Gate Array) を用いています。

PEACH2 による TCA には以下のような特徴があります

- ノードを超えた GPU 間直接通信を実現 (PCIe Gen2 x8 = 40Gbps)
 - オーバーヘッドの削減
- 複数ノード間で PCI Express アドレス空間の共有が可能
 - プログラミングが容易

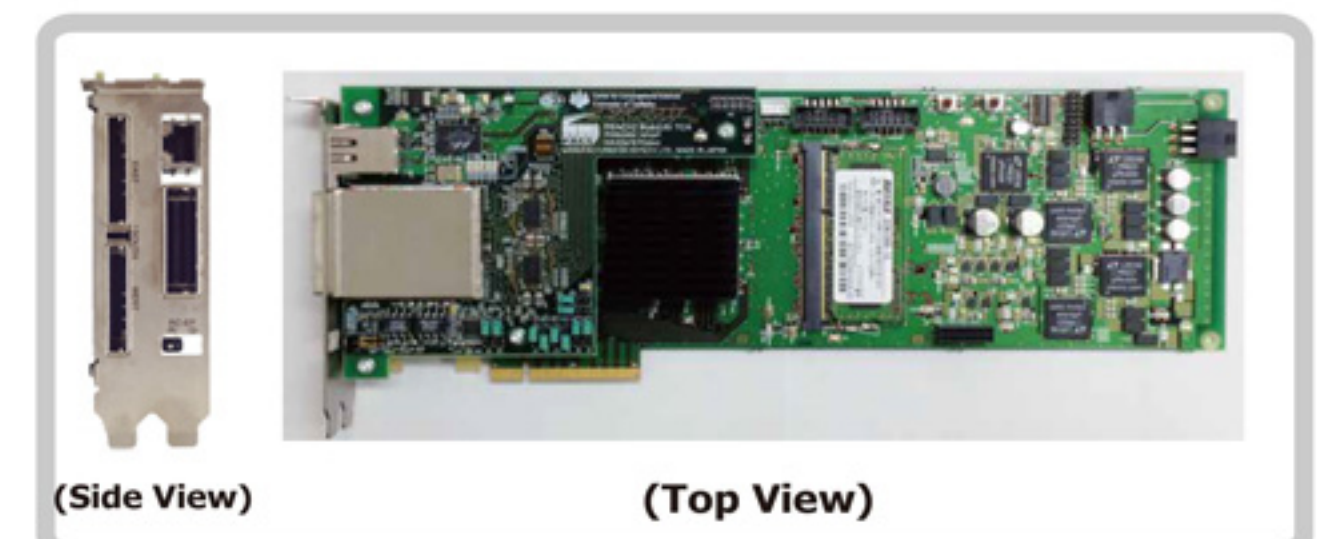
HA-PACS/TCA は TCA アーキテクチャに基づく実証 GPU クラスタとして筑波大学に設置されています。**TCA インタコネクト**だけでなく **InfiniBand QDR** を備えているため、性能比較も容易です。



HA-PACS/TCA 計算ノードのブロック図



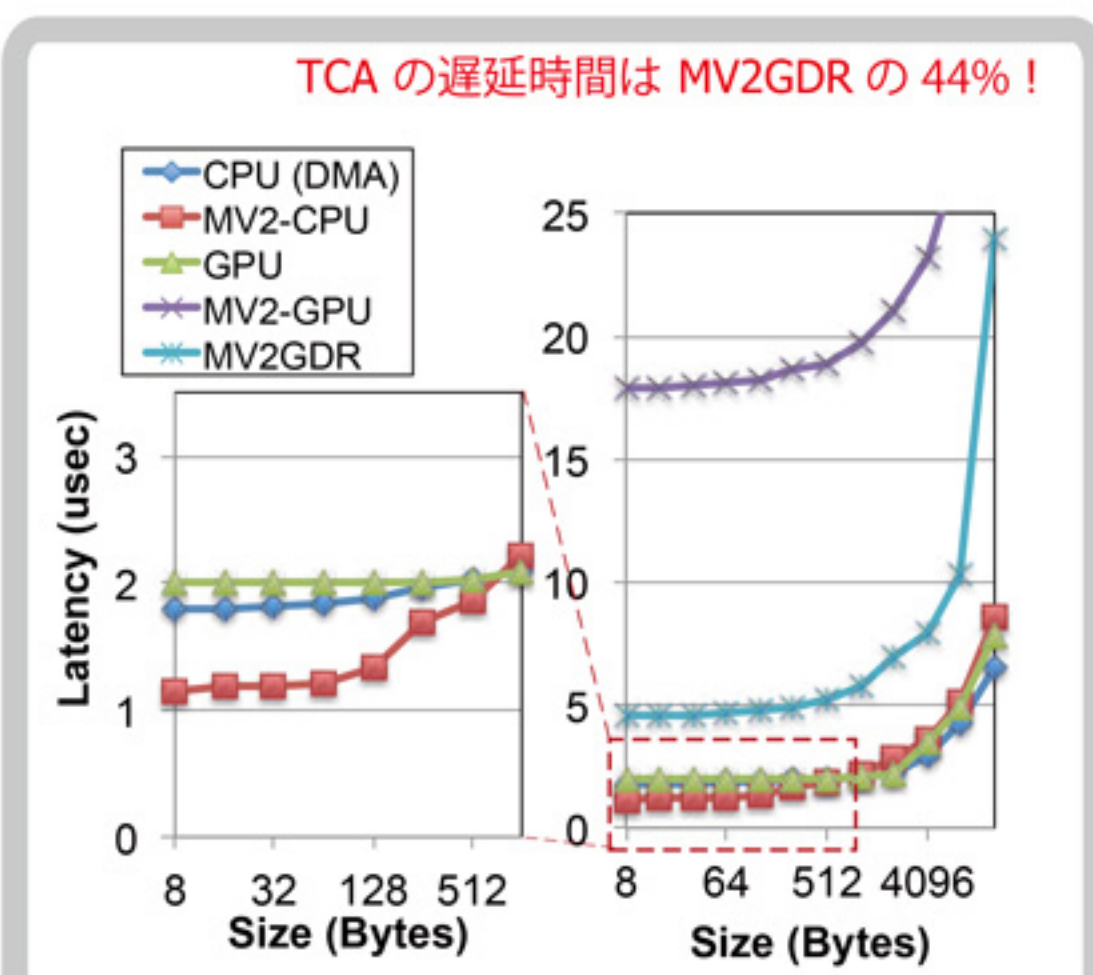
HA-PACS/TCA の写真
(手前 5x2 ラック、筑波大学)



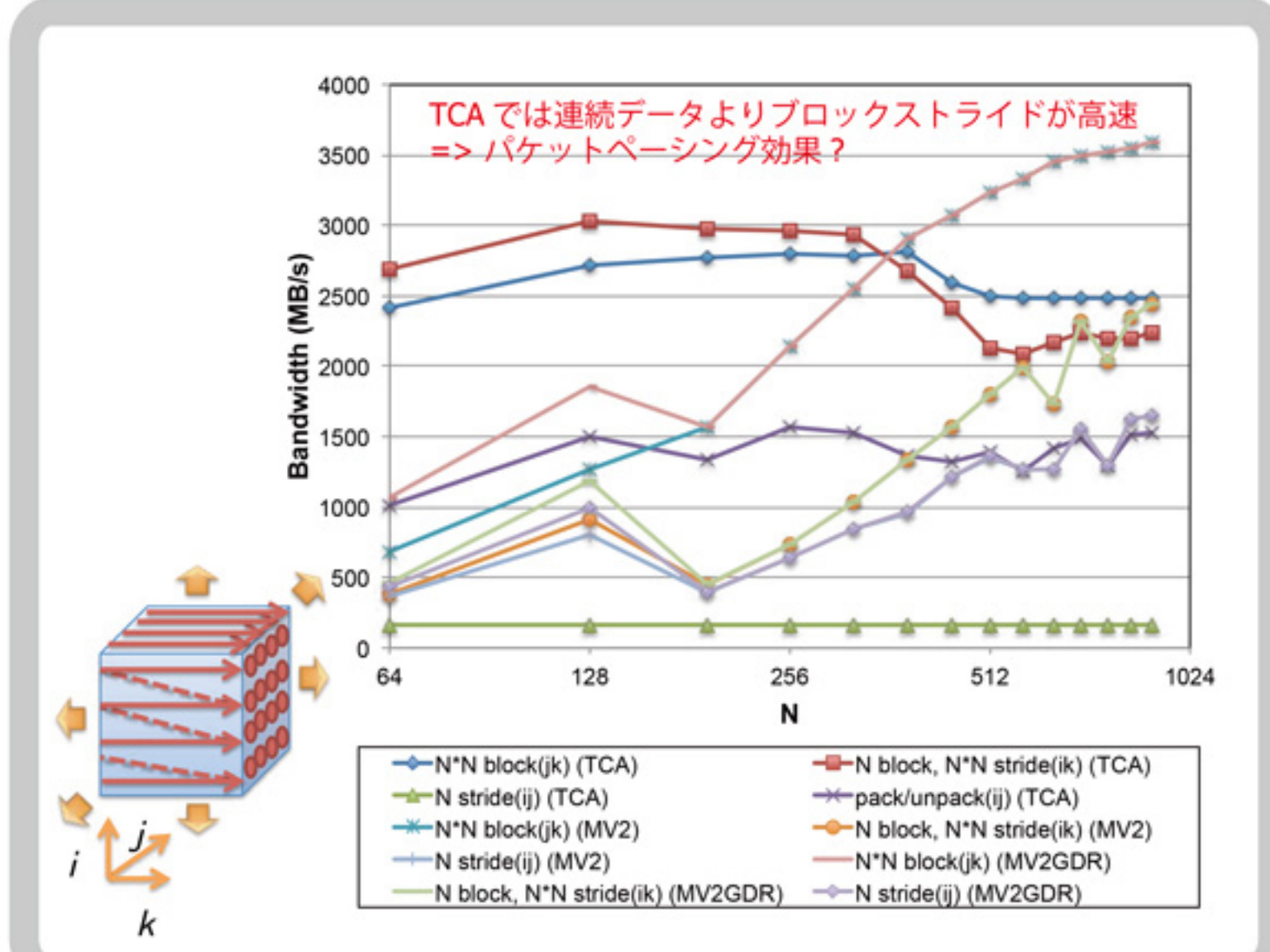
TCA 通信ボード
(PCIe CEM Spec., double height)

TCA による通信性能

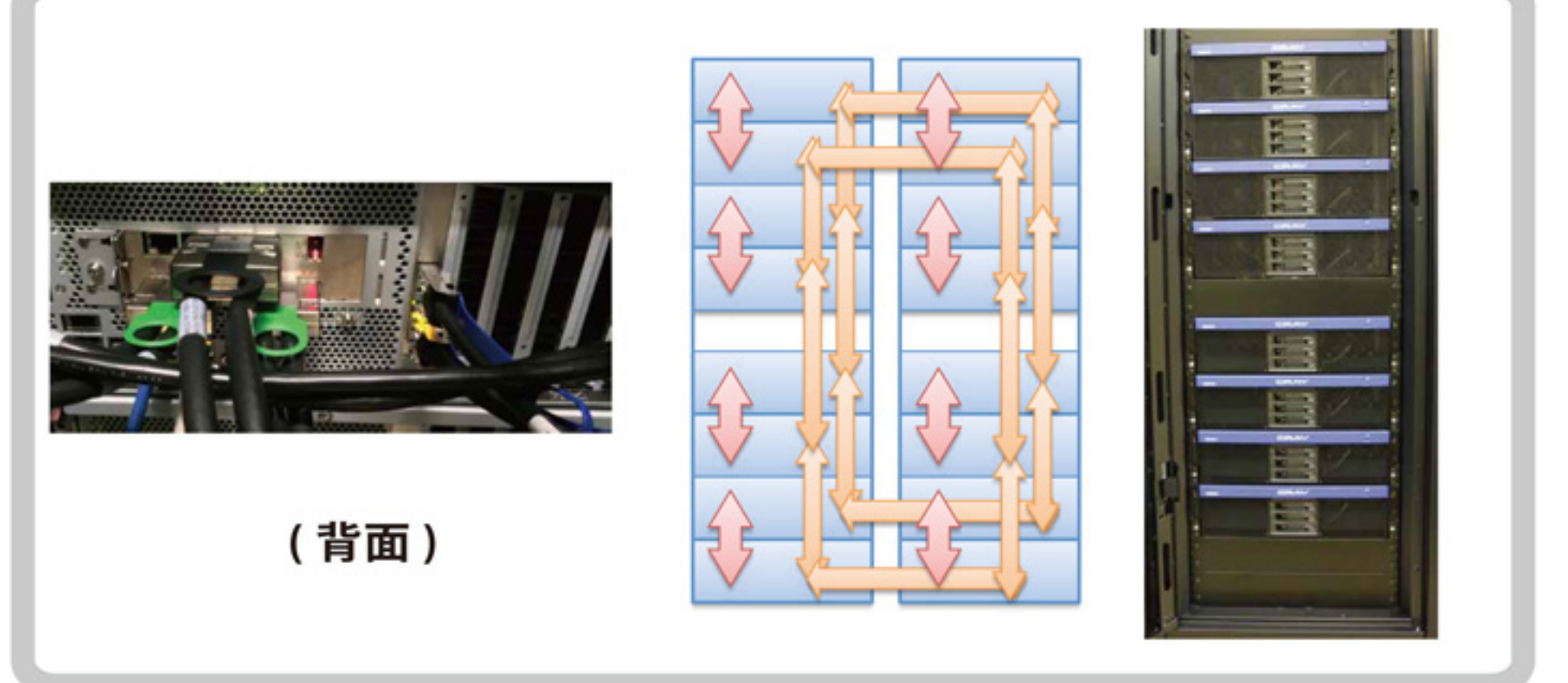
CPU: CPU 間隣接通信, GPU: GPU 間隣接通信
MV2: MVAICH2, MV2GDR: with GDR



DMA における Ping-pong 通信遅延



ステンシル計算における各平面における袖領域交換の DMA 通信性能



(背面)

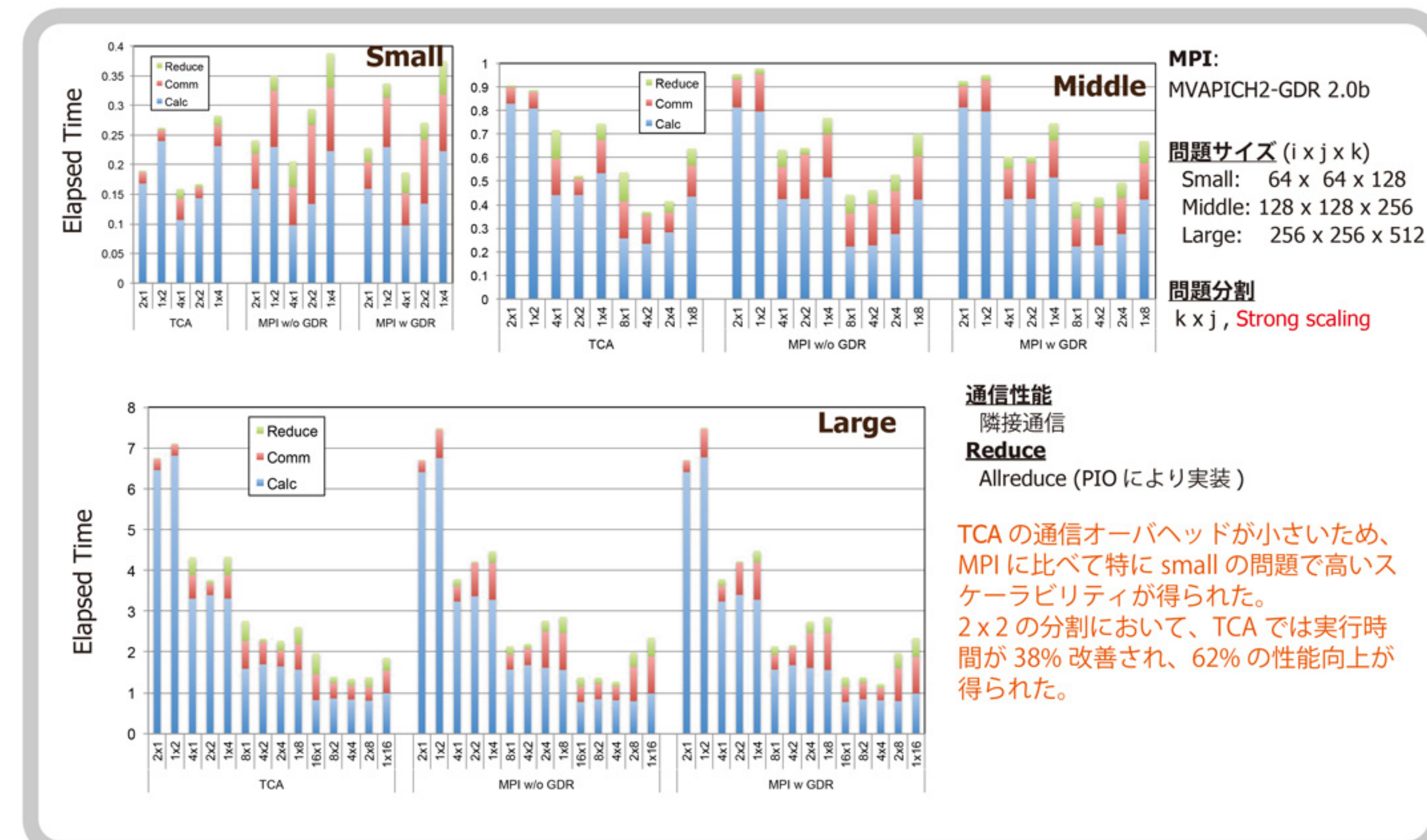
HA-PACS/TCA 計算ノードの写真,
TCA サブクラスタの構成 (16 nodes/group)

TCA における通信

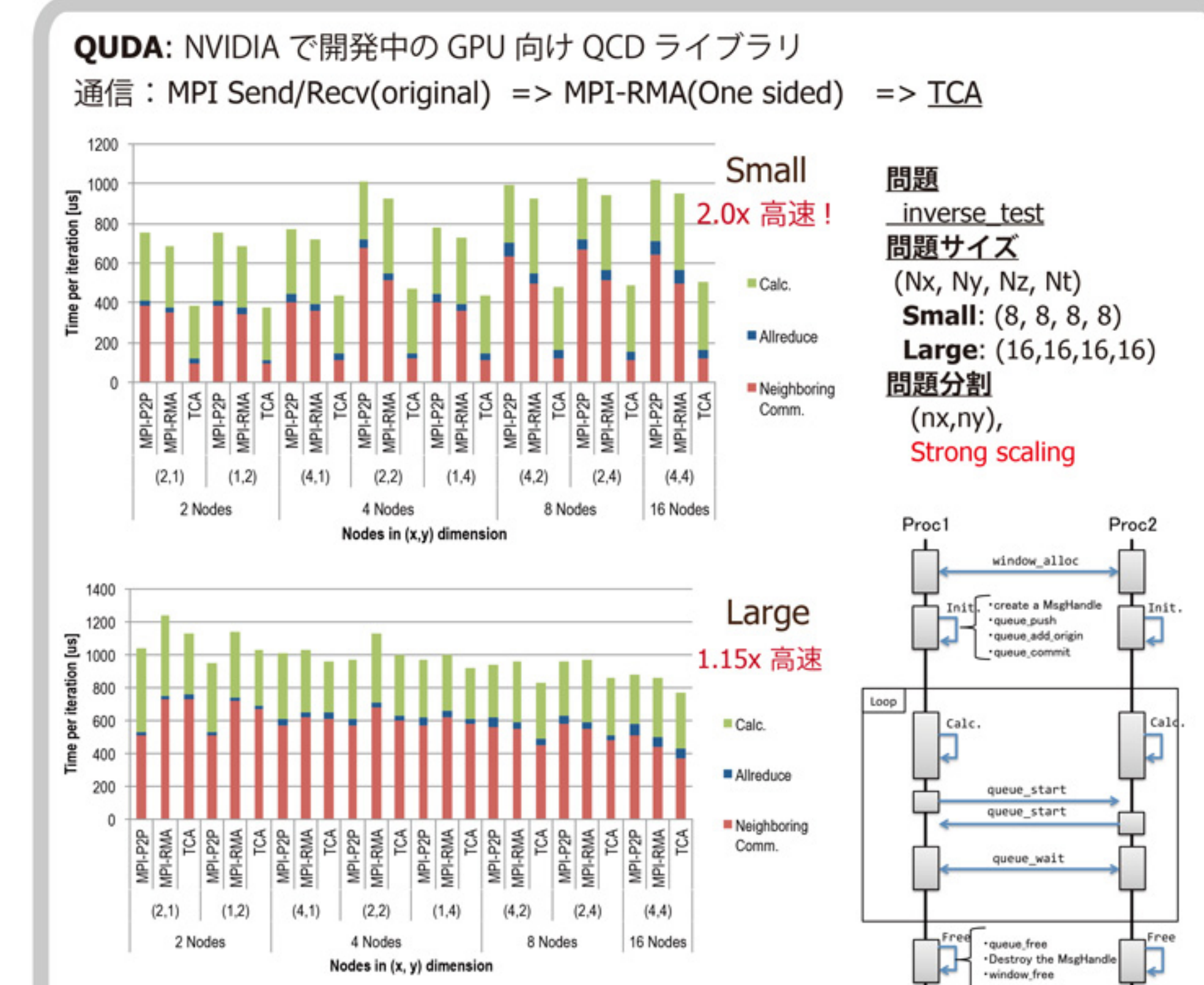
PIO (プログラム IO): CPU の Store 命令により直接リモート書込み
▶ 短メッセージに有利、超低遅延 (0.8usec)

DMA (ダイレクトメモリアクセス): 4ch, チェーン機能、ブロックストライド機能
レジスタモード: 16 個までの DMA 要求をレジスタに登録
内蔵メモリモード: PEACH2 内蔵メモリから DMA 要求を取得
ホストメモリモード: ホストメモリから DMA 要求を取得

姫野ベンチマークへの適用



TCA 通信機構を用いた QUDA



次世代スパコンを創造する!

既存のスパコンを使ったアプリケーションの最適化や分析などにより、問題点を明らかにすると同時に、独自ハードウェアによる専用ネットワークおよび通信ライブラリの開発などを行っています。FPGAを用いたアプリケーションのオフローディングと通信の融合、プログラミングとの統合が目下のテーマです。



長谷川研究室 Hasegawa Laboratory

URL: <http://sites.google.com/view/yoshihikohasegawa/>

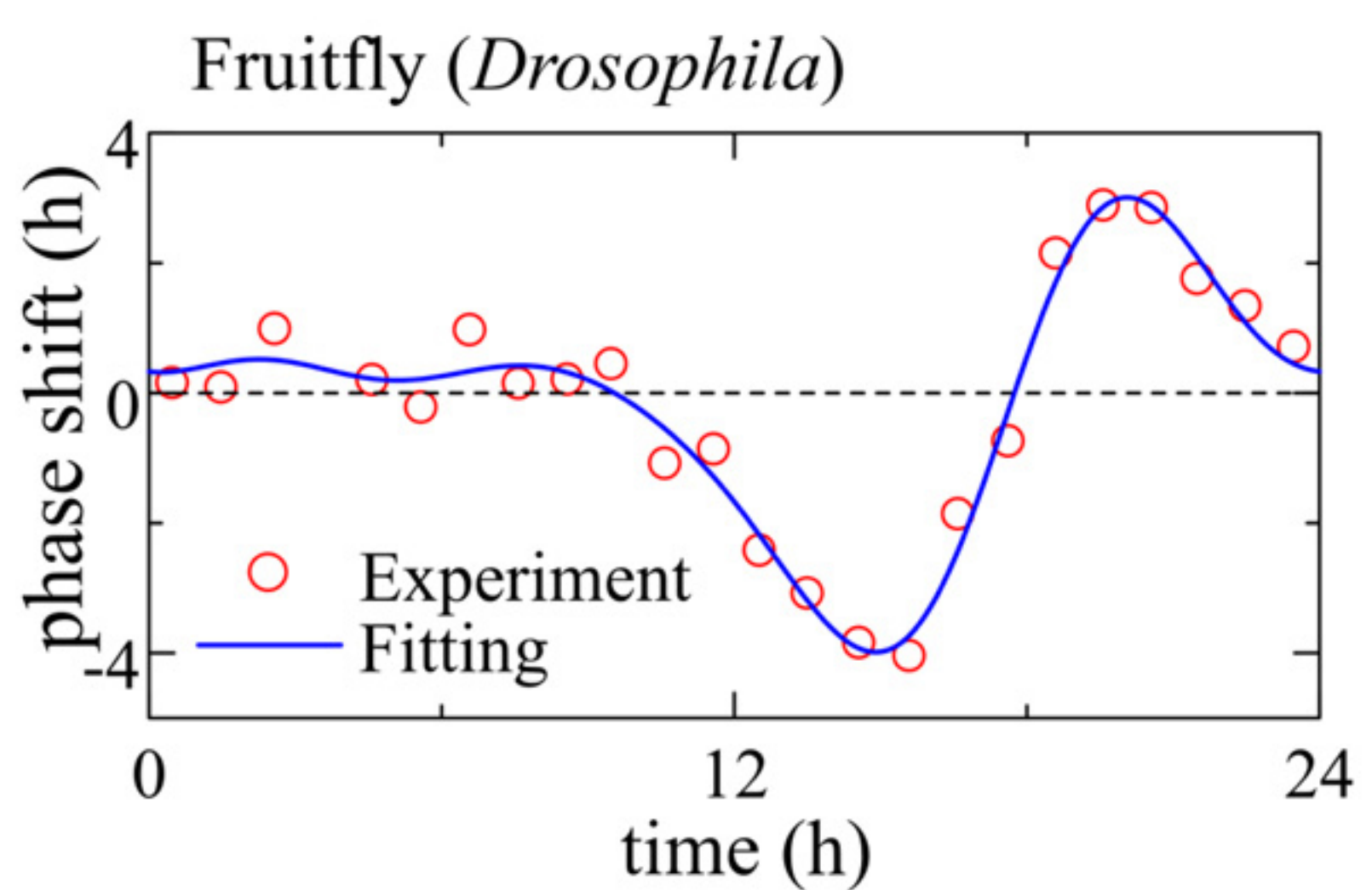
工学部2号館12F122A3
Bldg. Eng-2 12F Room 122A3

数理生物学・確率過程・力学系

21世紀は情報と生物の時代と言われており、近年生命現象の理解は多くの分野で最先端の研究課題となっています。従来、生物学は実験による解析が主でしたが、観測技術の発展により定量的なデータが得られるにつれ、生命現象を司るメカニズムの理解が進んで来ました。それに伴い、情報科学や物理学などの理論的枠組みによって非線形現象を解析する「数理生物学」や「システム生物学」が急速に発展しています。本研究室は情報科学や物理学などの手法を用い、理論的に生物を理解することを目指しています。より具体的にはDynamical Systems(力学系・動的システム)やStochastic Process(確率過程)といった理論を用います。

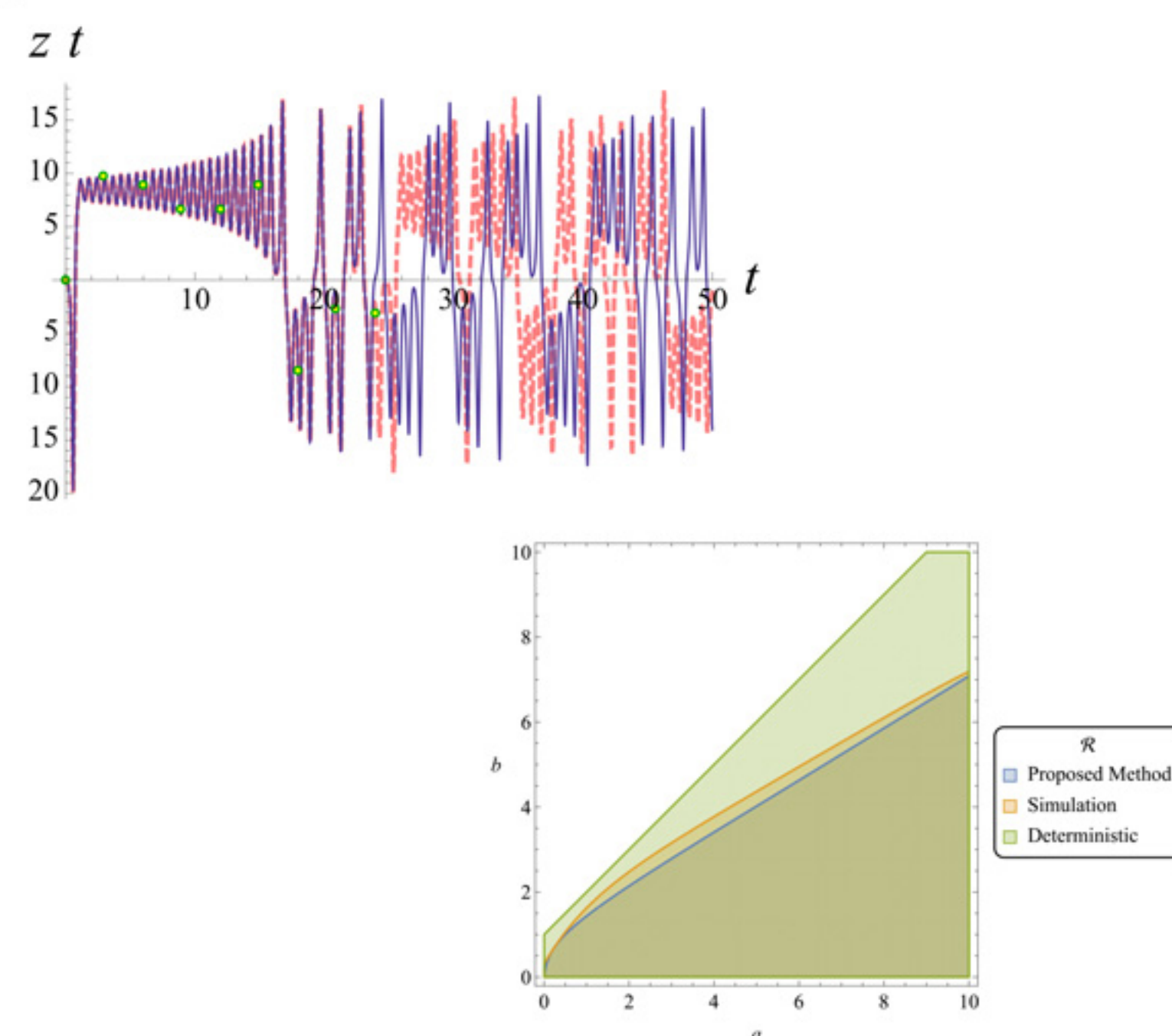
生体振動子

生体は24周期の「時計」を持っています。一見すると、工学的な時計とは異なるように見えますが、電気回路で知られるファンデルポール振動子と多くの共通点を持っています。そのため、微分方程式として生体振動子記述することで、その性質の理解・予測を行うことが可能です。



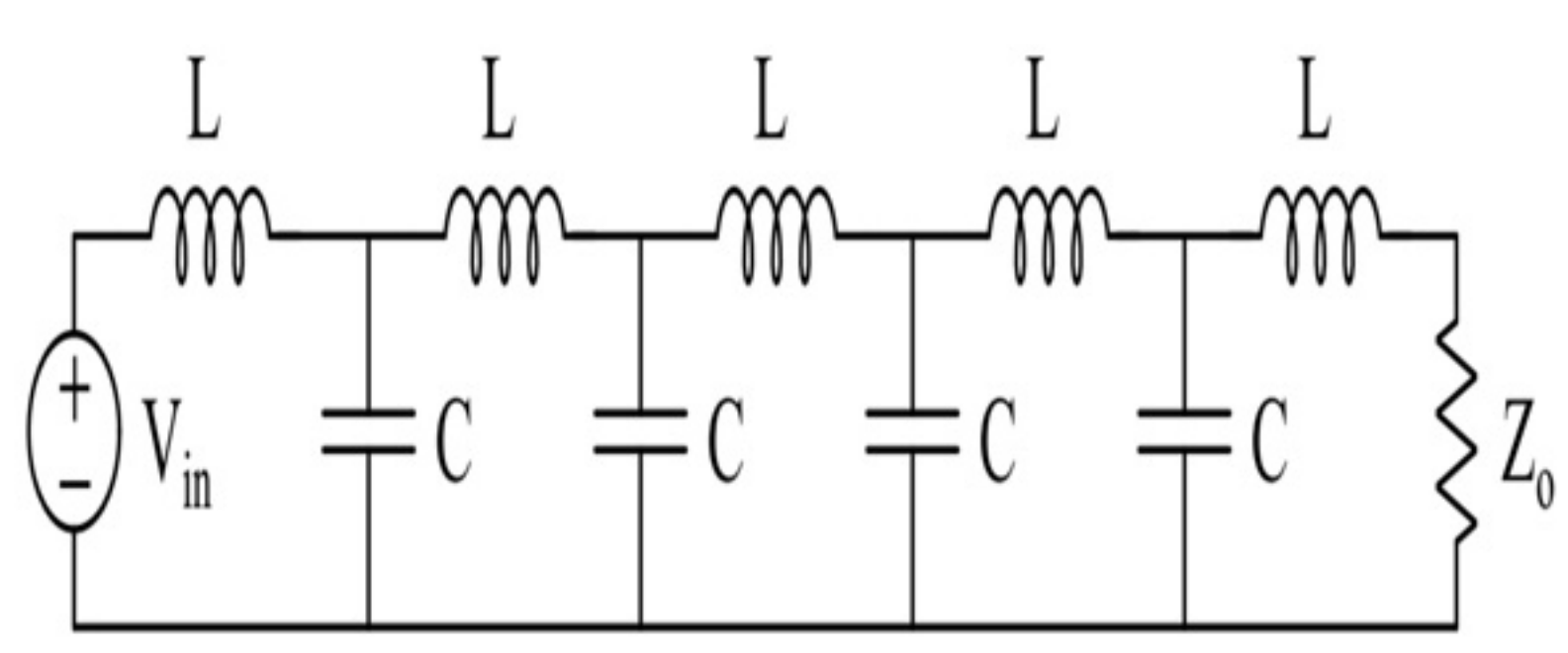
代数生物

生命現象においては非線形性が本質的に重要な役割を担いますが、線形代数では表せない非線形性を多項式によって解析する手法が注目を集めています。生命現象の理解に代数的手法を適用する研究を行っています。



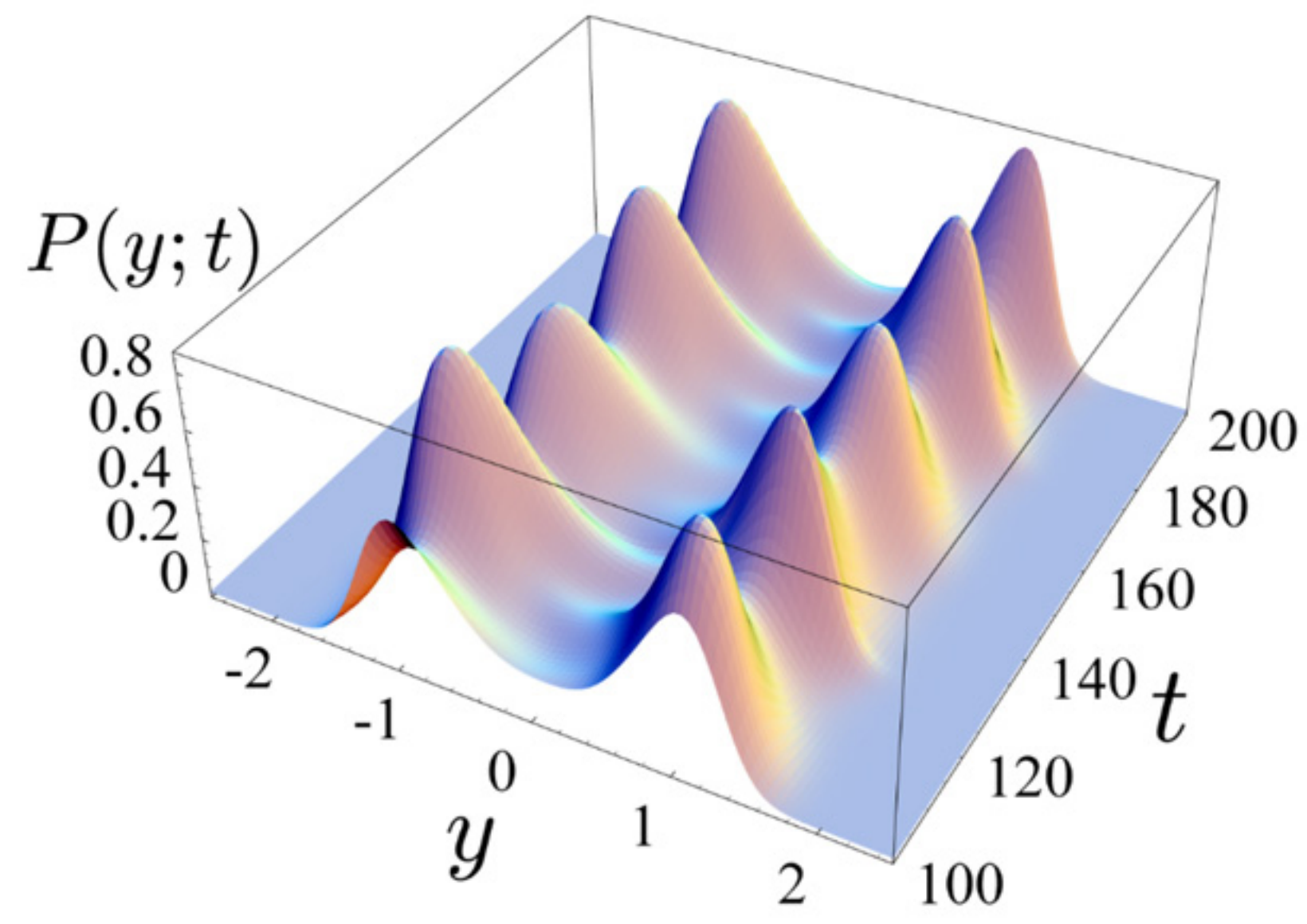
生物と電気工学

生化学反応は微分方程式によって記述されますが、多くの場合、その反応には遅延要素を含みます。しかし、遅延微分方程式は解析が難しいという難点があります。本研究室では電気回路のアイデアを用いた遅延微分方程式の近似方法を開発しています。



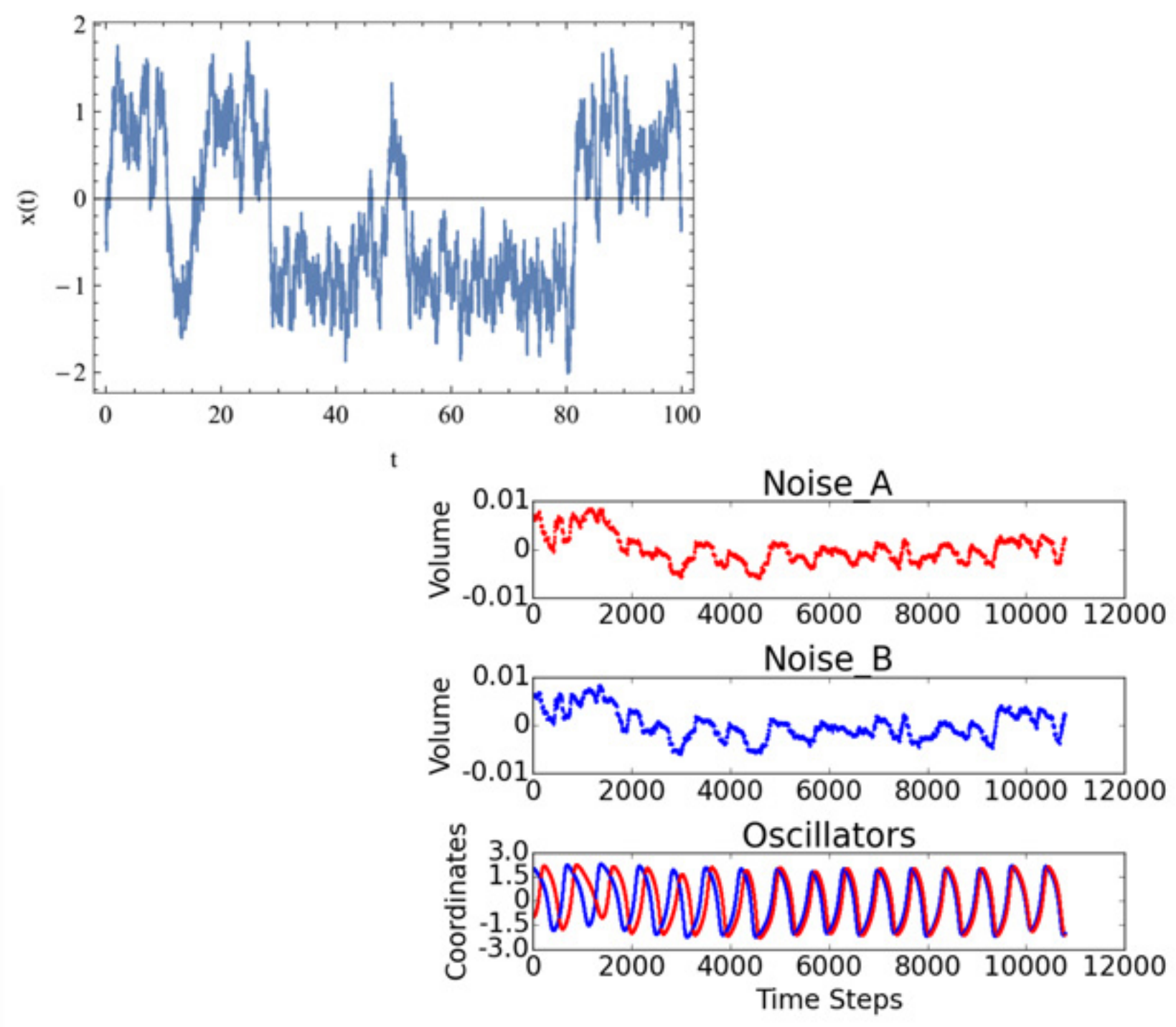
確率微分方程式

確率微分方程式の確率密度関数の時間変化は、簡単な場合を除き求めることが非常に難しいことが知られています。確率密度関数の時間変化を、解くことが簡単な関数の組み合わせによって近似する手法の研究を行っています。



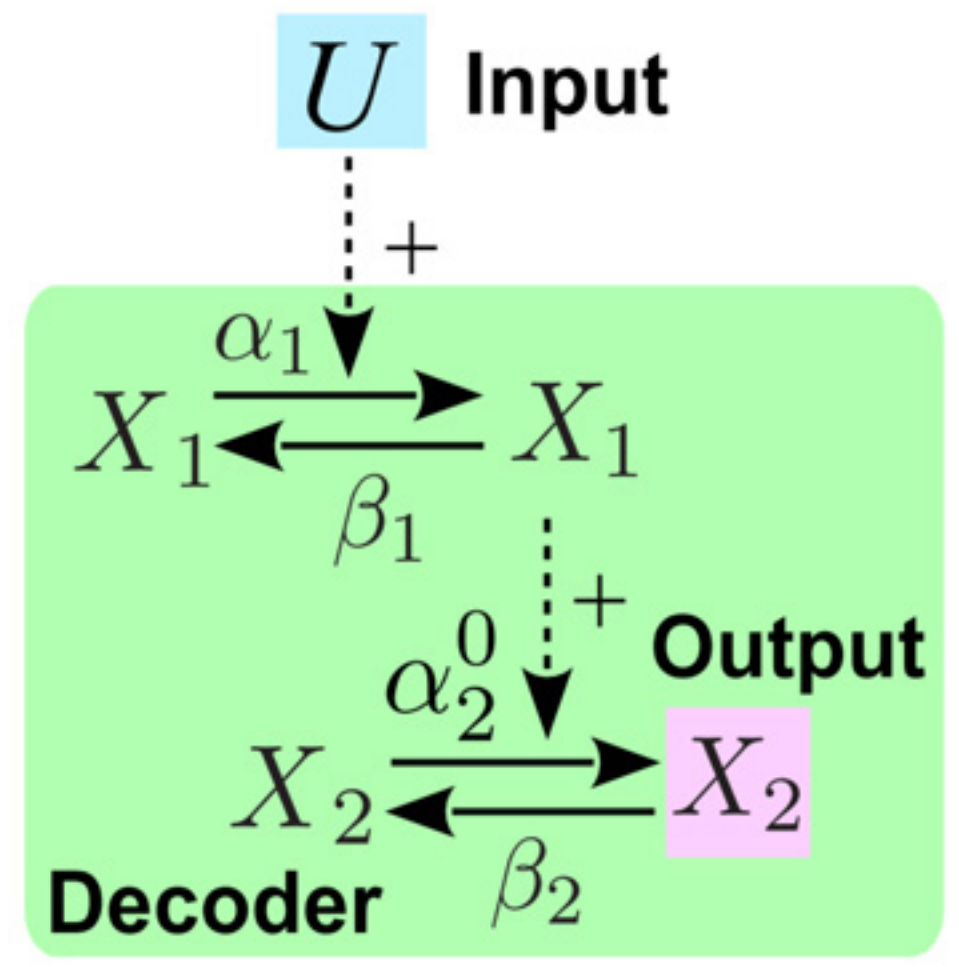
同期現象

概日時計をはじめとした生体振動子は決定的に振る舞うのではなく、常に外乱(ノイズ)の影響を受けています。本研究室では、ノイズ下にある振動子のダイナミクス解析や、ノイズ下の振動子の同期現象の工学的な利用方法について研究しています。



生命現象の熱力学

生命現象は微小な細胞内で行われるため、その効率とエネルギー消費量には密接な関係があることが知られています。そのため、生命現象は省エネルギーになるように進化的に変化してきました。シグナル伝達系や生体振動現象とエネルギー消費の関係について熱力学などを用いて解析しています。



数理で紐解く生命現象

当研究室では生物を対象として扱いますが、数理最適化や非線形物理学における手法を適用し、生命の全容解明を目指しています。

橋田研究室(Prof. Hasida) Hasida Laboratory

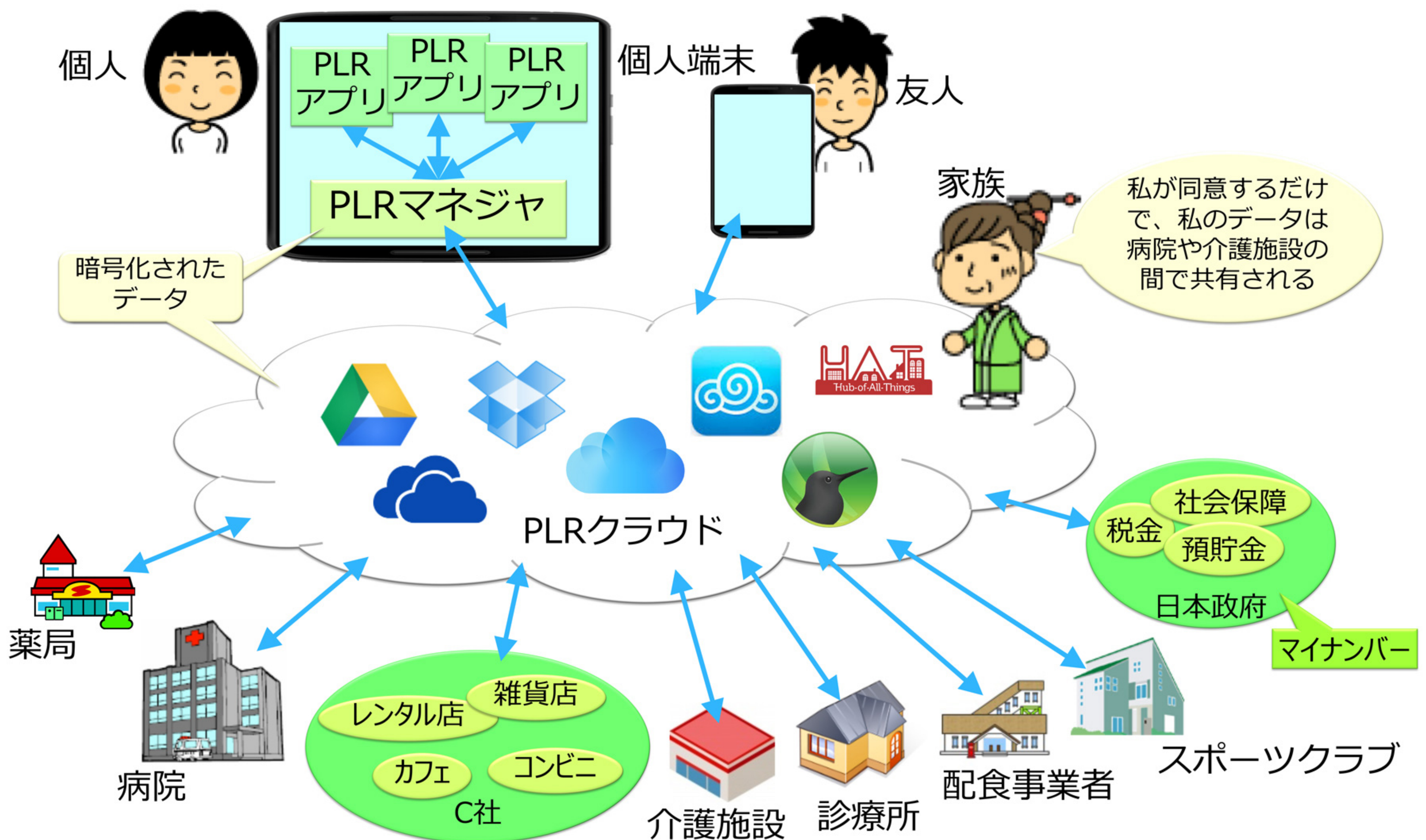
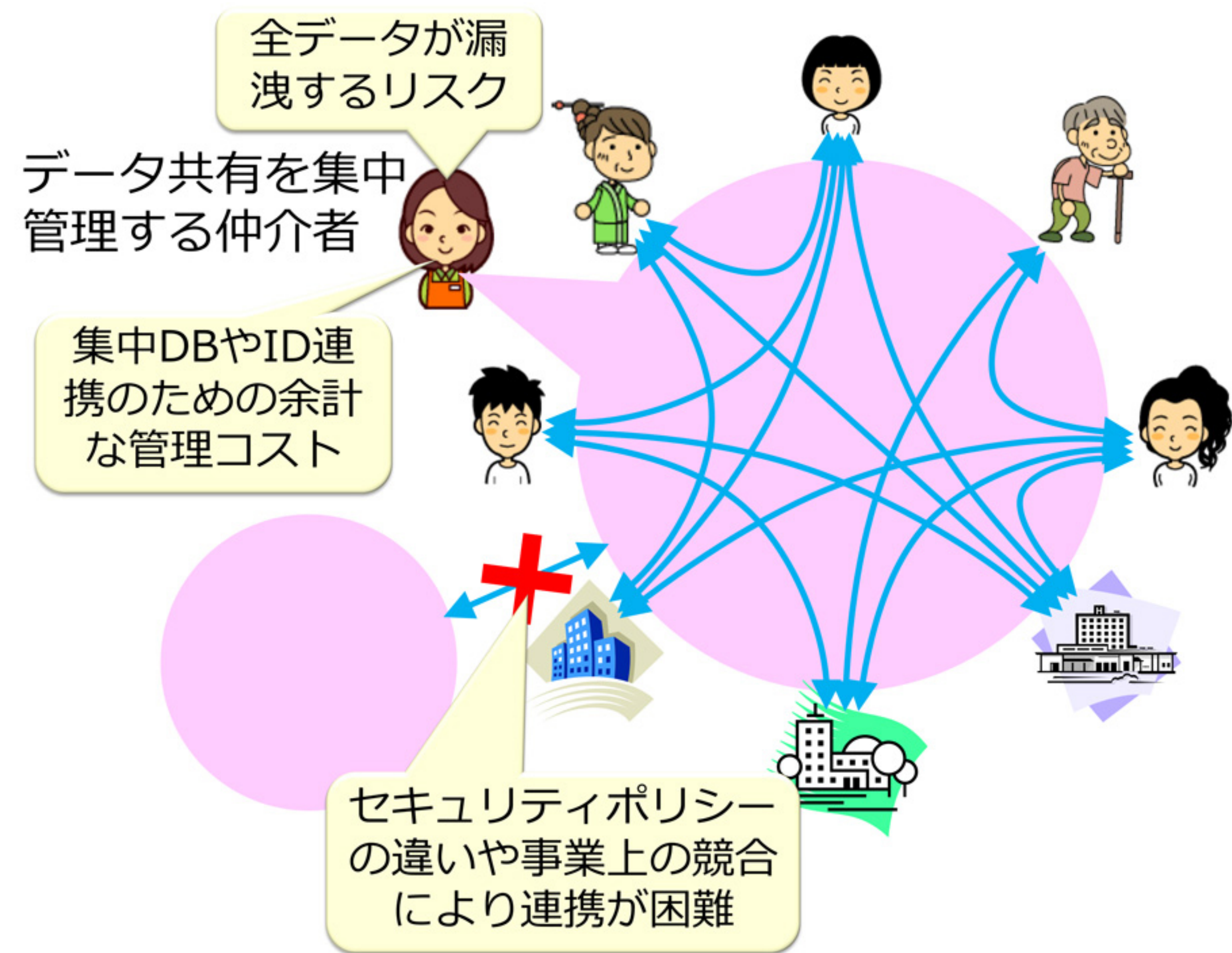
URL: <http://www.sict.i.u-tokyo.ac.jp/members/hasida/>

工学部2号館9F 92D4
Bldg. Eng-2 9F Room 92D4

パーソナルデータエコシステム

たとえばA病院の診療記録をB病院に開示して安全で効果的な治療を受けるなど、パーソナルデータを複数の事業者が共有することでサービスの価値が高まります。しかし、右図のようにデータの本来の所有者と利用者以外の仲介者がデータ共有を集中管理するこれまでの方法は、わざわざコストをかけて大量のデータが漏洩するリスクを生んでしまいます。そこで、下図のようにPLR (個人生活録) という仕組みによって個人(代理人)が仲介者を経ずに本人のデータの流通と活用をコントロールする社会を実現し、サービスの本来の受益者である個人にとっての価値を最大化するような競争とイノベーションを促進します。

集中管理によるデータ共有



国が推進している医療制度改革や地域包括ケアのため、ヘルスケア事業者は明確に役割を分担しながら患者や被介護者のデータを共有して協働することが必須になります。また、銀行等のサービスがオープンAPI化されることにより、決済やPOSデータの活用も個人が中心になるでしょう。PLRは従来の集中型のデータ共有技術よりもはるかに安価で安全で利便性が高いので、これらの改革に大きく貢献できると期待されます。当研究室では、政府や産業界と連携してPLRによるパーソナルデータエコシステムの普及を目指しています。

PLR

個人が特定の事業者に依存せずに自ら本人のデータを蓄積・管理し、データの種類や相手を指定してデータを安全に共有できるようにするスマートフォン等のアプリ。これにより、いわゆるサービスを含むさまざまな協同作業を自律分散協調的に実現し、産業や文化の発展を促すことができると考えられます。



入江研究室(Asso. Prof. Hidetsugu Irie)
Sakai-Irie Lab

URL:<http://www.mtl.t.u-tokyo.ac.jp/>

学部 電子情報工学科 本郷
大学院 情報理工・電子情報学専攻

工学部2号館11階 112B2
Bldg. Eng-2 11F Room 112B2

コンピュータと人間の幸福な共存を目指して

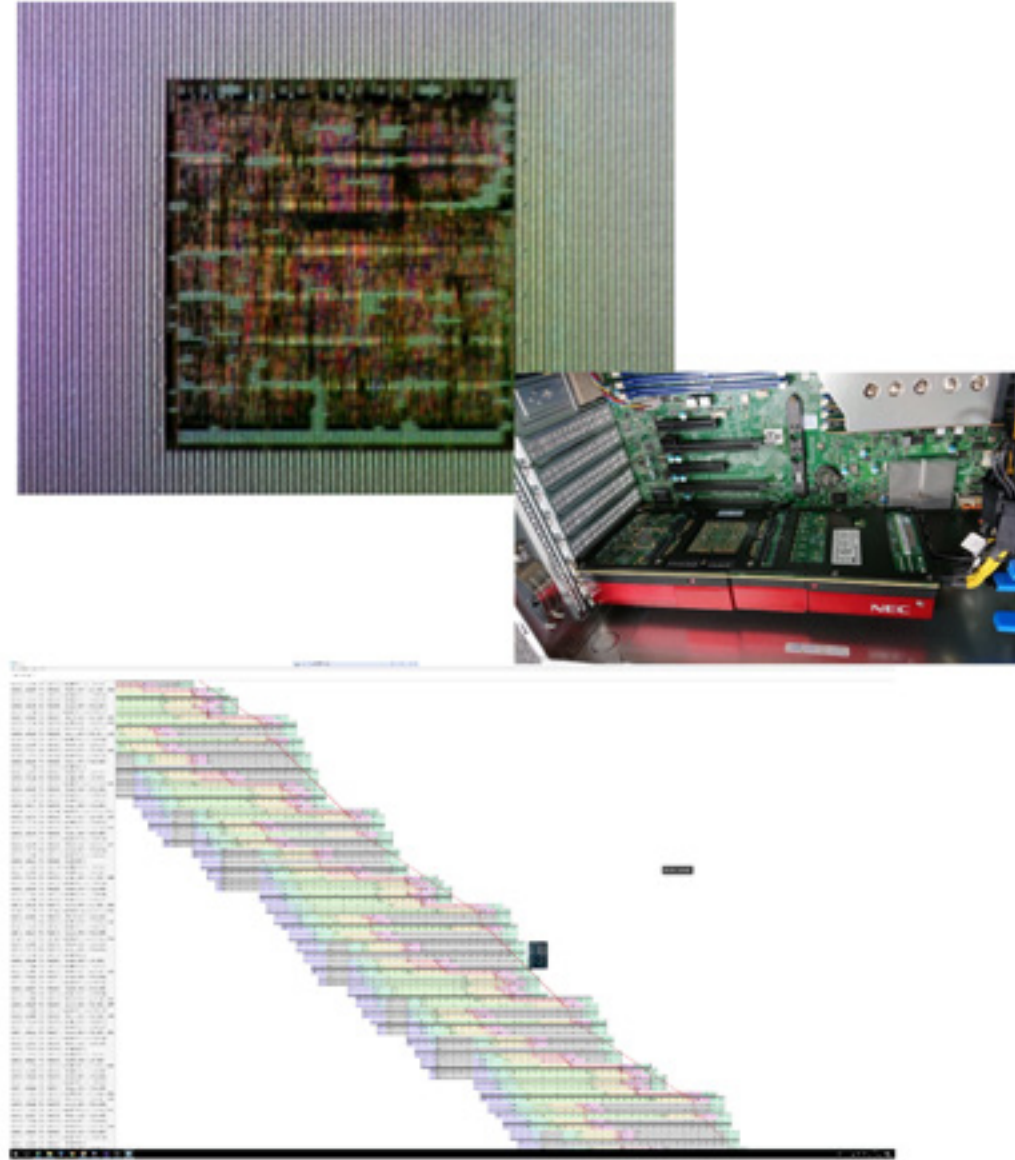
本研究室ではコンピュータと人間の幸福な共存を大目標とし、コンピューティングシステムの高性能化・高信頼化・セキュリティ向上を目指して、プロセッサアーキテクチャを軸に、言語・コンパイラ・OSからアプリケーションまでを視野に入れた、幅広い研究を行っています。

マイクロプロセッサ

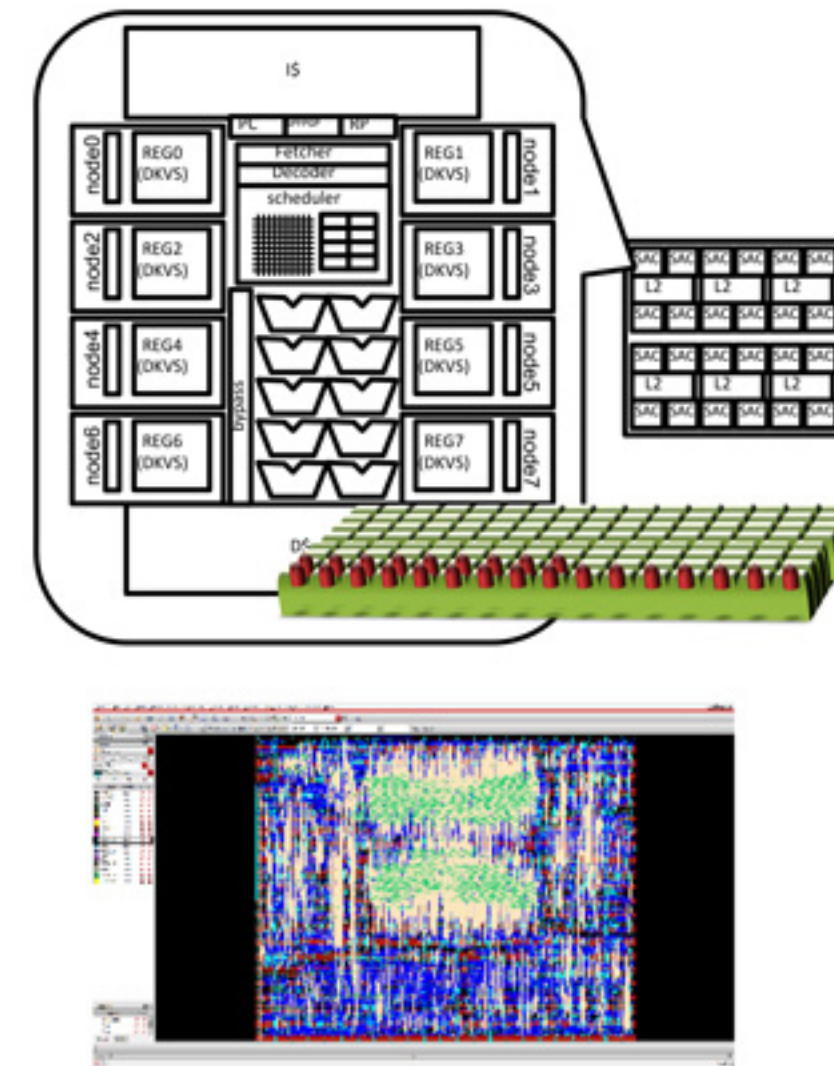
情報社会の趨勢を決める中核の一つは、コンピュータの中核部、プロセッサです。コンピュータ・アーキテクチャはこのグランドデザインを決める重要な技術で、同じチップ面積/電力条件でもアーキテクチャの善し悪しで性能は全く変わります。

このグループではクラスタ計算機や大型FPGAボードといったハードウェア資源、ハードウェア設計用CADやハウスメイドの高性能シミュレータおよびソフトコアなどのソフトウェア資源を駆使して、コンピュータの成長可能性を更に広げる研究を進めています。

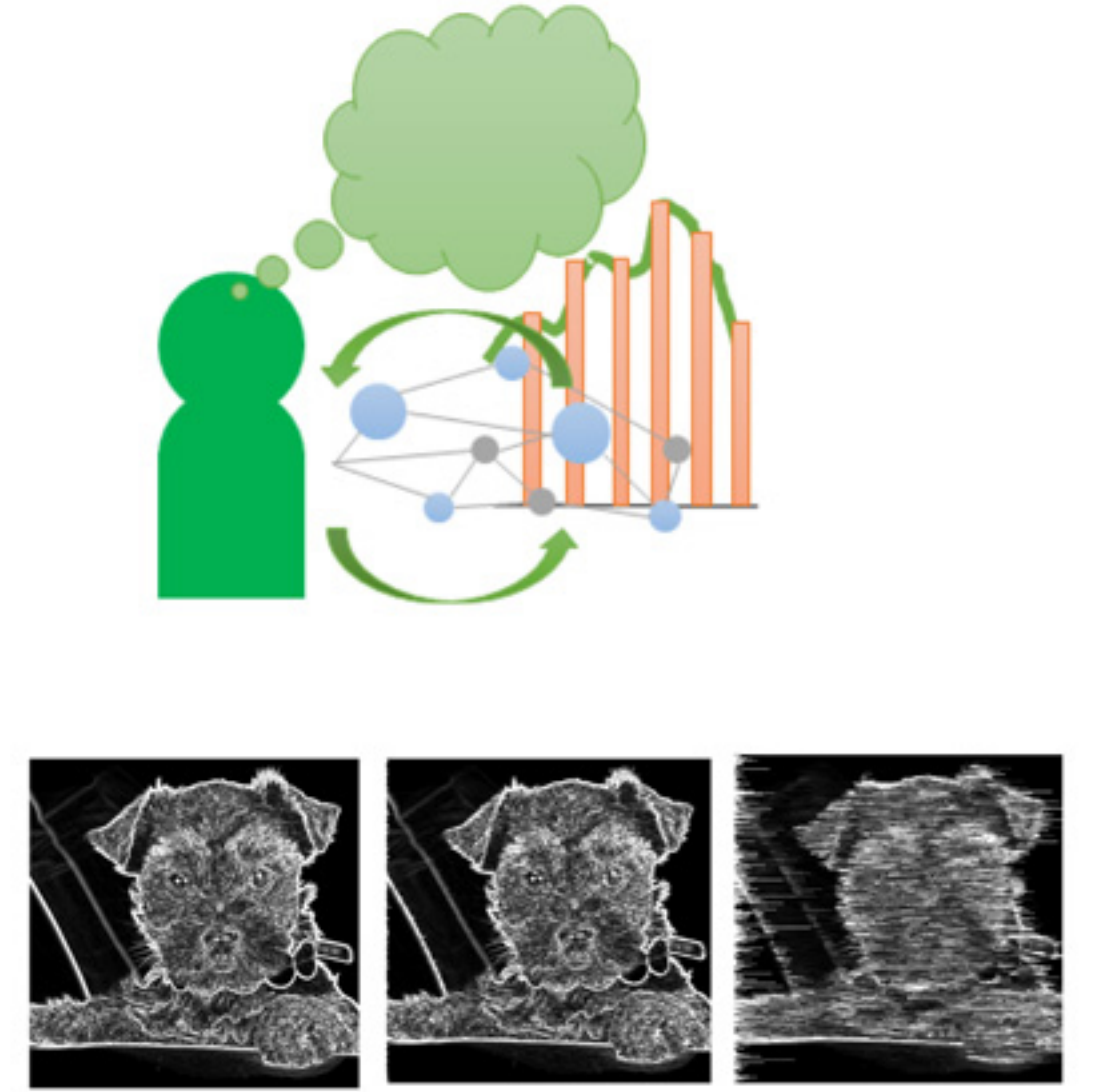
マイクロアーキテクチャ



新しいILPアーキテクチャとそのコンパイラ最適化



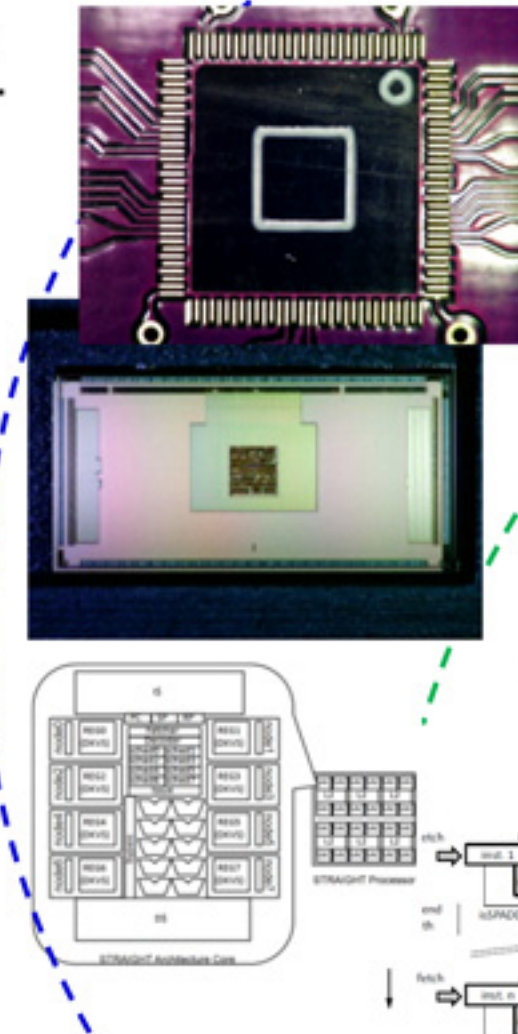
オンデマンド近似計算



コンピュータ・アーキテクチャ

- CPUアーキテクチャ
- SiP, チップ試作
- ハードウェア
- FPGA
- アクセラレータ
- メモリシステム
- 近似計算

より強力なコンピュータ



ユーザを察するコンピュータ



ヒューマン・コンピュータ・インタラクション

- NUI
- HMD
- お供ロボット・ドローン
- エッジコンピューティング
- エンタテインメント
- リアルタイム画像処理

人間とコンピュータの幸福な共存

コンピュータ・セキュリティ

- セキュアプロセッサ
- VM・HV
- ブロックチェーン・分散処理
- 情報フロー追跡
- ヒューマンエラー対策

安心して使えるコンピュータ

HCI (ヒューマン・コンピュータ・インタラクション)

このグループでは高効率高性能計算技術を軸として、ウェアラブル・車載・IoT・ドローンといったユーザを取り巻くコンピュータが、その計算力でユーザの意志や状態を素早く察して対話やサービスを行う技術を開発しています。特注のドローン、透過型/非透過型HMD、各種先進センサ、エッジコンピュータなどのハードウェアや企業が持つ実地ログなど様々な学習用データを活用して、ユーザによりそのコンピュータの在り方を提案しています。

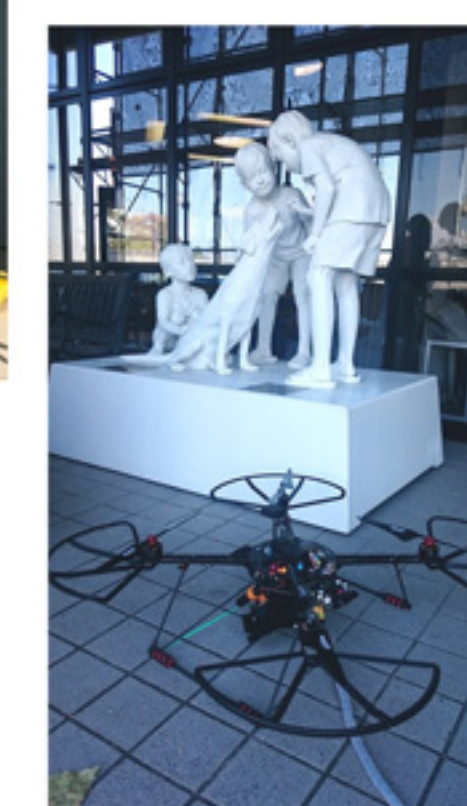
光学透過型HMDのための直観的インタフェース



エッジコンピュータによる車両・歩行者支援



コミュニケーション支援・インタラクション技術



研究を一言で表現

より強力なコンピュータ、あるいはより軽量で柔軟なコンピュータを実現する研究を行っています。また、それらの新しいコンピュータが可能とするユーザ支援サービスについて研究しています。



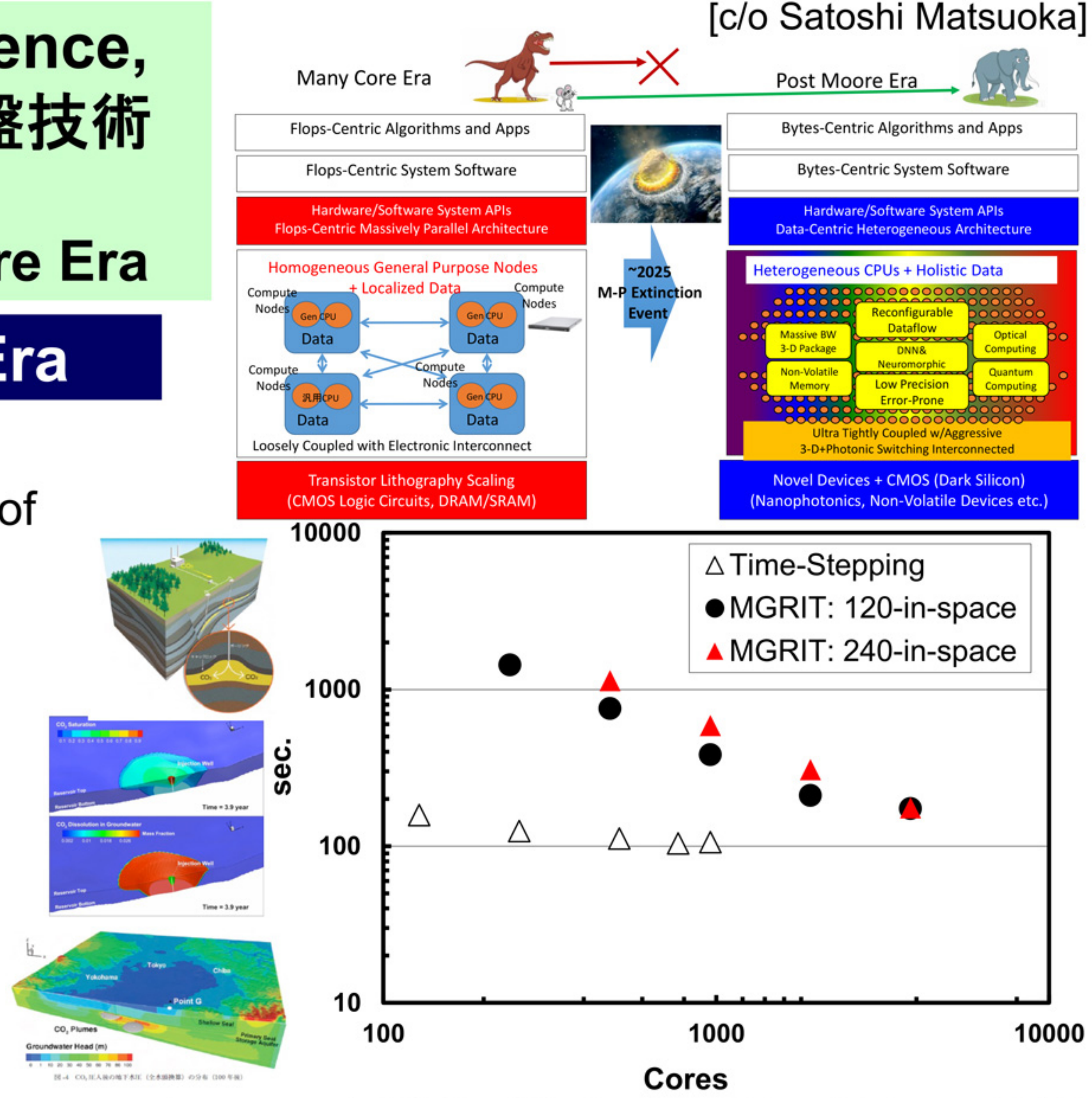
第3の科学「計算科学」(The Third Pillar of Science, Computational Science)の発展に貢献する基盤技術
ポストムーア時代のアプリケーション&アルゴリズムへ向けて
Towards Applications & Algorithms in Post Moore Era

Applications & Algorithms in Post Moore Era

- Assumptions & Expectations after 2025 ...
 - High bandwidth in memory and network, Large capacity of memory & cache
 - Memory: 3D Stacked Memory etc.
 - Network: Optical Communication etc.
 - Large & heterogeneous latency due to hierarchy in memory & network
 - Utilization of FPGA
 - High concurrency with $O(10^3)$ threads on each node

Applications & Algorithms in Post Moore Era

- Compute Intensity \Rightarrow Data Movement Intensity**
 - Integration of CSE & Data Analysis/Machine Learning
 - Data Driven Approach
- Hierarchical Methods for Hiding Latency
 - Parallel in Space/Time (PiST): Parallelization in *time* direction for time-dependent problems, suitable for supercomputer systems with large latency & with network hierarchy in the Post Moore Era.
 - hCGA in Multigrid Method: Hierarchical Coarse Grid Aggregation, suitable for large systems, and hierarchical networks
 - H Matrix Solver
- Communication/Synchronization Avoiding/Reducing Algorithms
 - Pipelined Algorithm [Ghysels et al. 2014]: hiding collective communications by using asynchronous collective comm. (e.g. `MPI_Iallreduce`) in MPI-3
 - Utilization of Dynamic Loop Scheduling
- Power-aware Methods
 - Approximate Computing, Power Management, FPGA



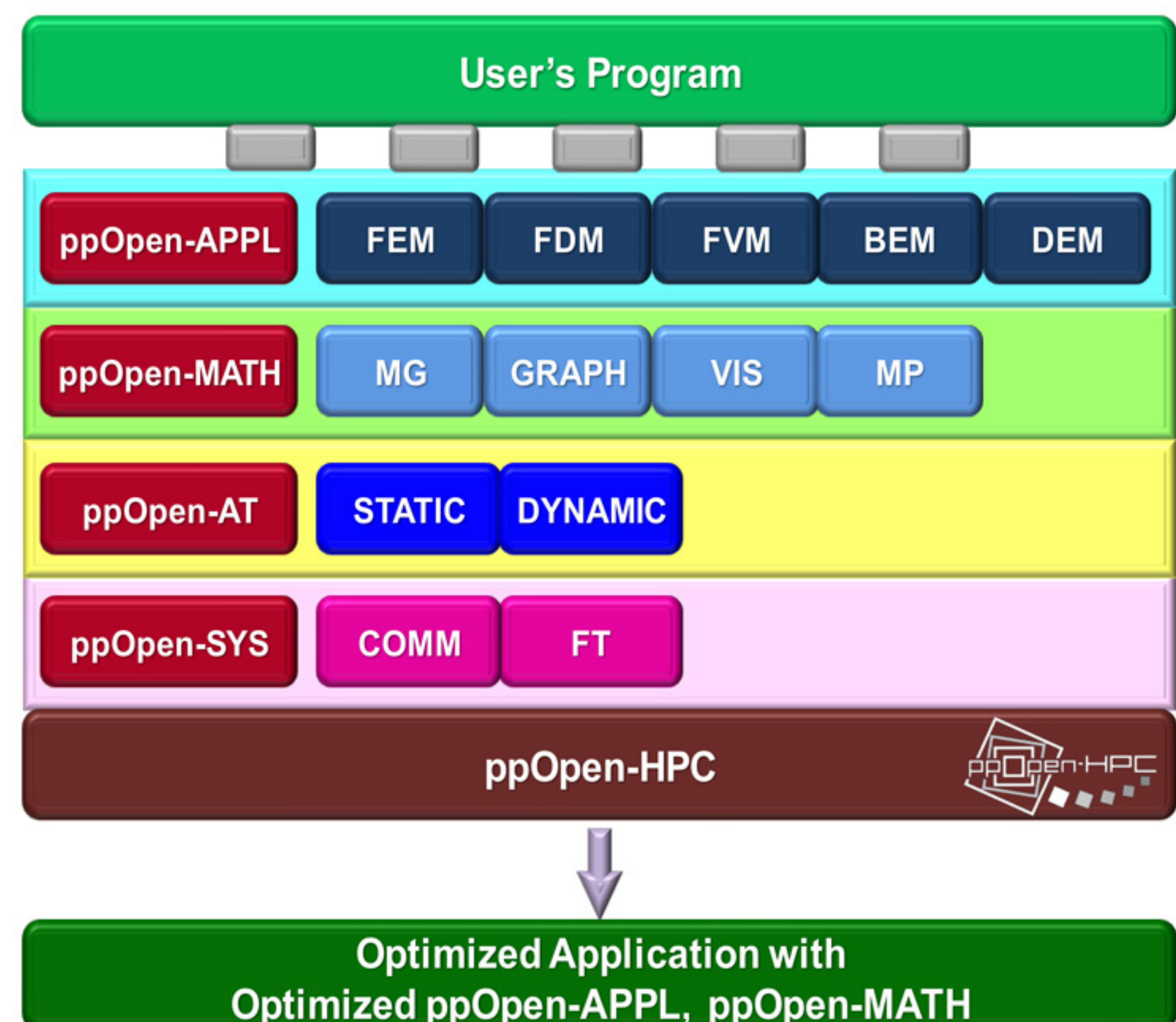
Goal: Integration to ppOpen-HPC

- Open source infrastructure for development & execution of optimized and reliable simulation code on post-peta-scale (pp) systems (JST/CREST since 2011)
 - ppOpen-APPL: Framework for Application Development
 - ppOpen-MATH: Math Libraries
 - ppOpen-AT: Automatic Tuning
 - ppOpen-SYS: System Software



Future Development of ppOpen-HPC

- Post-K Era (Memory Wall): Block-Structured AMR, SELL-C- σ
- Post Moore Era: Post-Pre/Post Processor for Parallel-in-Space/Time (PiST)
 - Nonlinear Algorithm
 - AMR, Visualization, Coupler for Multiphysics



スパコンで第三の科学に貢献!

スパコンを駆使した大規模シミュレーションによる計算科学は理論・実験に続く「第3の科学」と呼ばれています。当研究室では、スパコン上での大規模シミュレーションを支える数理的基盤の研究開発をScience-Modeling-Algorithm-Software-Hardware (SMASH)の幅広い観点から実施し、「第3の科学」の発展に貢献します。



坂井研究室(Prof. Shuichi Sakai) Sakai-Irie Lab

URL:<http://www.mtl.t.u-tokyo.ac.jp/>

学部 電子情報工学科 本郷
大学院 情報理工・電子情報学専攻

工学部2号館11階 112B2
Bldg. Eng-2 11F Room 112B2

コンピュータと人間の幸福な共存を目指して

本研究室ではコンピュータと人間の幸福な共存を大目標とし、コンピューティングシステムの高性能化・高信頼化・セキュリティ向上を目指して、プロセッサアーキテクチャを軸に、言語・コンパイラ・OSからアプリケーションまでを視野に入れた、幅広い研究を行っています。

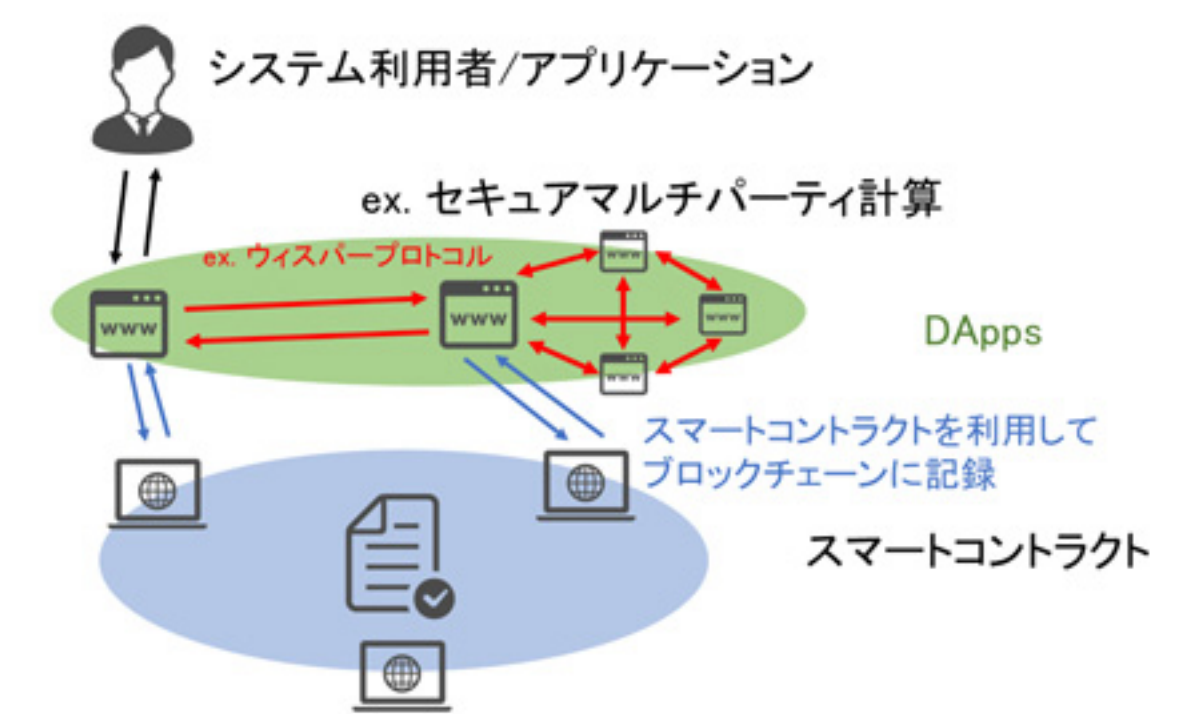
コンピュータセキュリティ

コンピュータシステムの発展とともに、多くのユーザが利便性のため大切なデータや処理をクラウドに委ねたり、他者からアクセスできる様々なセンサを日常空間に配置したりするようになってきました。一方で、クラウドオペレータによる情報漏洩、IT家電のクラッキングや誤作動によるプライバシー漏洩、ウェブサイトやブラウザへの注入攻撃、Meltdown/Spectreに代表されるようなマイクロアーキテクチャレベル脆弱性、不正ハードウェア埋め込みなど、攻撃の複雑さ高度さも増えています。このグループでは、ユーザの想定外の事態を引き起こさない、賢く用心深いコンピュータシステムを目指して研究を進めています。

セキュアハイパーバイザによる 情報フロー追跡



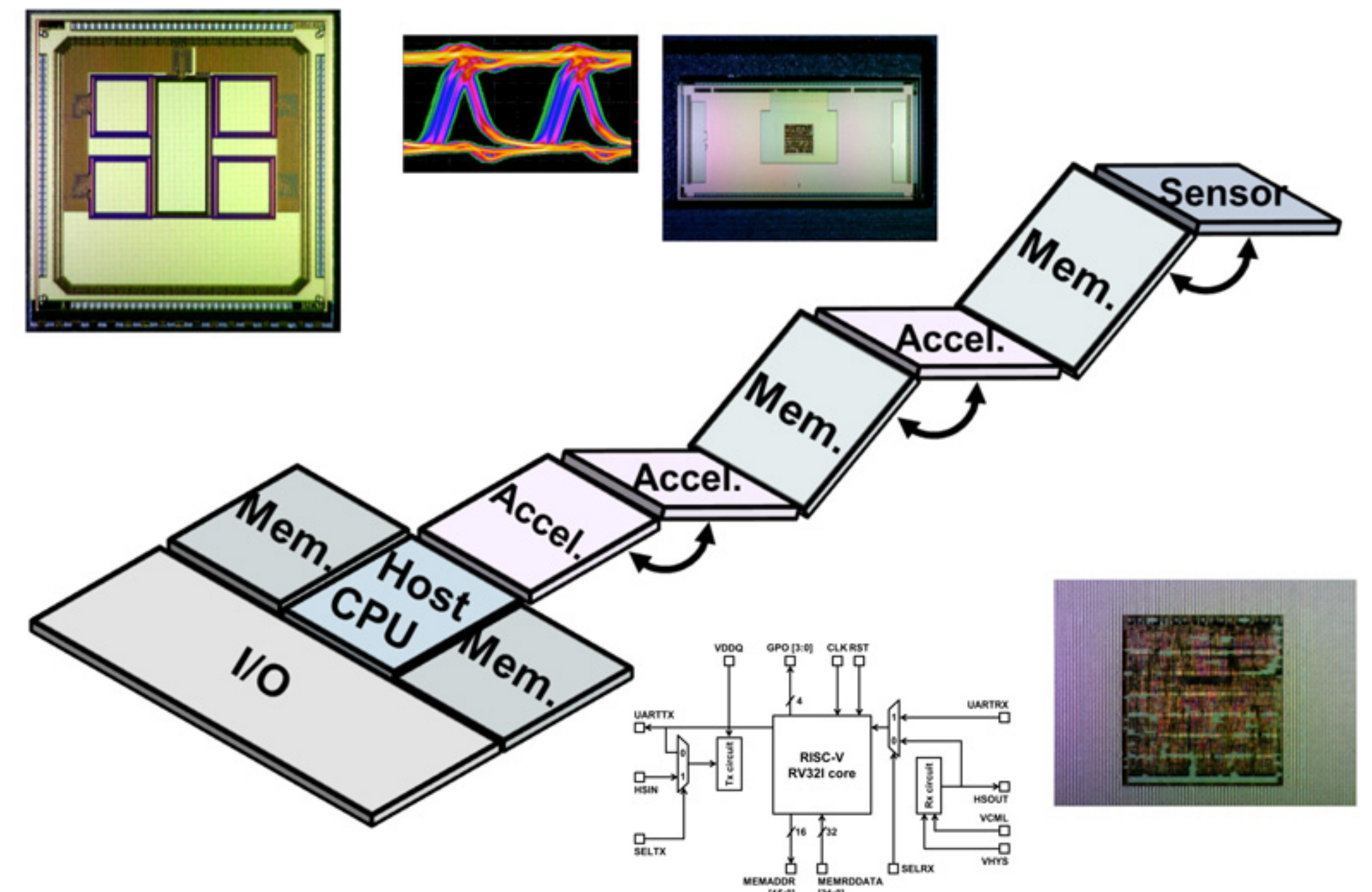
ブロックチェーンと 分散セキュアフレームワーク



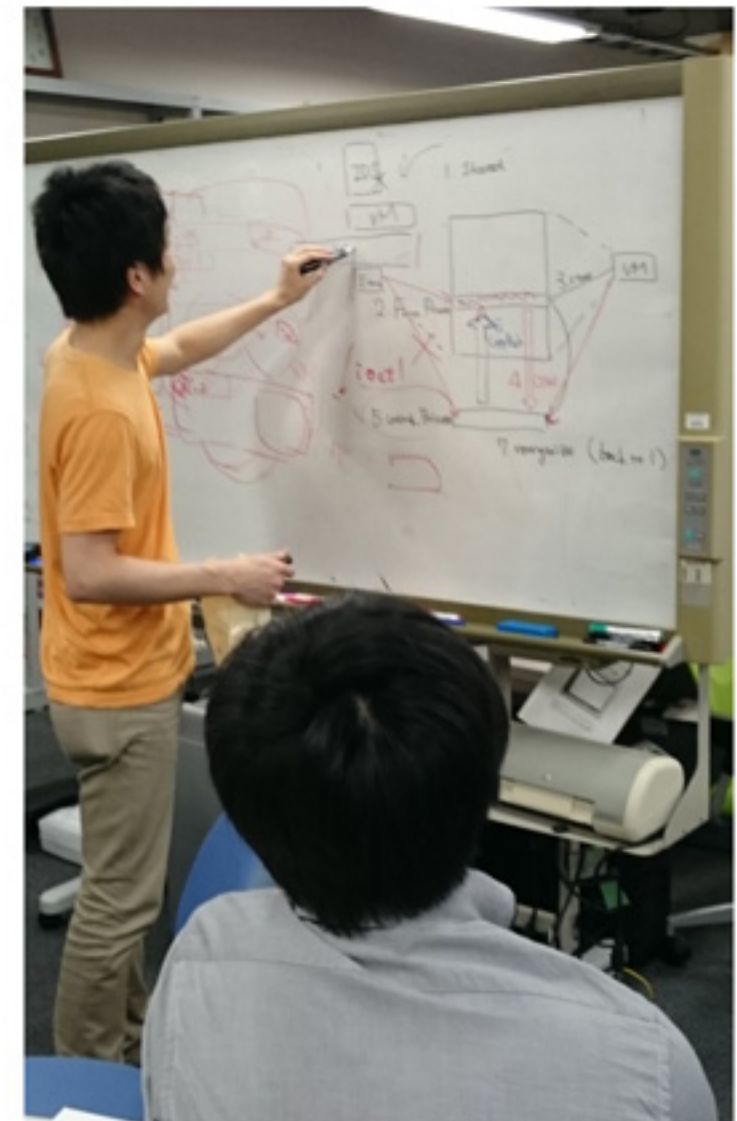
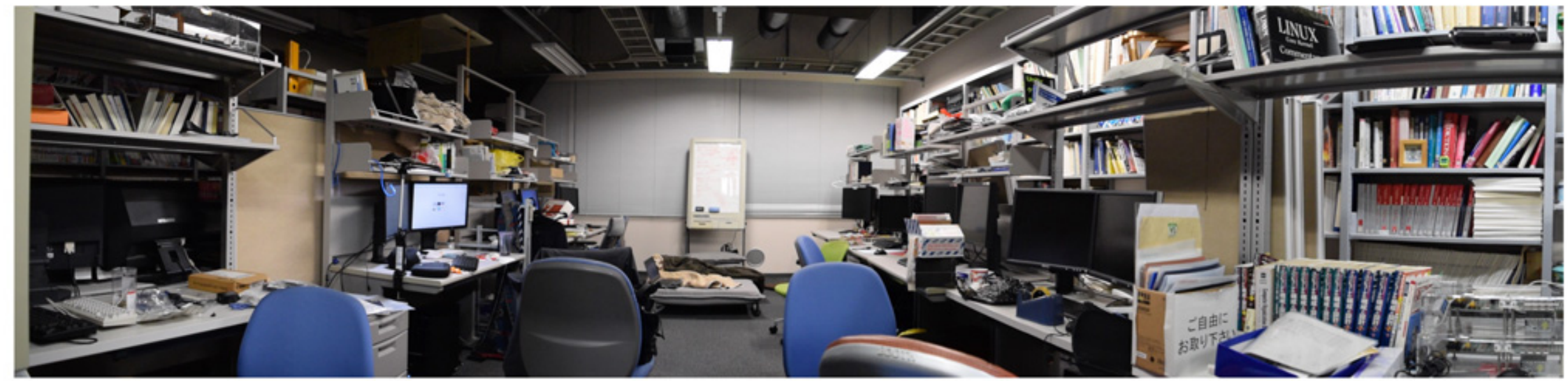
形状自在コンピュータ

コンピュータのダウンサイジングはウェアラブルコンピューティングやワイヤレスセンサネットワークなど新しい使い方を可能としてきました。本研究は、さらに微細な場所や可動部分に入り込めるコンピュータの実現を目指し、形状自在コンピュータを提案しています。これは、横並びにした微細チップのそれぞれの外縁に通信用のコイルを配置し、接続したチップ同士が誘導結合によって通信を行い、全体として1つのコンピュータとして機能するシステムです。一つ一つのチップは微少で、また、隣接するチップ辺での折れ曲がりやずれが可能であるため、微少で柔軟なコンピュータの実現が可能となります。

このシステムの実現にむけ、チップ間プロトコルやクロック・電力供給、アプリケーションなどの開発を、実際にチップを製造しながら行っています。

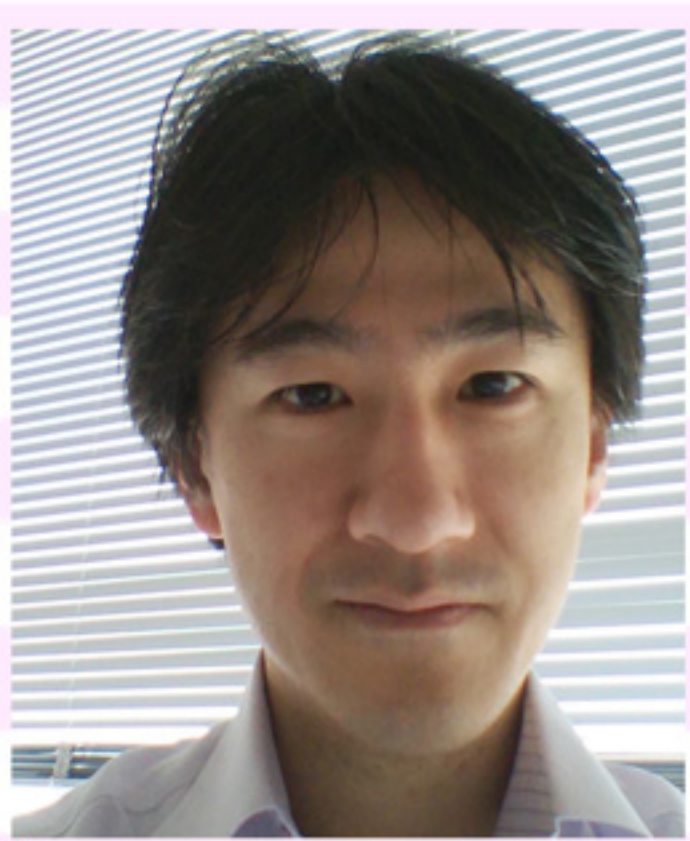


研究室の様子



研究を一言で表現

本研究室ではコンピュータと人間の幸福な共存を大目標とし、コンピューティングシステムの高性能化・高信頼化・セキュリティ向上を目指して、プロセッサアーキテクチャを軸に、言語・コンパイラ・OSからアプリケーションまでを視野に入れた、幅広い研究を行っています。



田浦研究室 Taura Laboratory

学部 電子情報工学科 本郷
大学院 情報理工学系・電子情報学専攻

工学部2号館10F101C2
Bldg. Eng-2 10F Room 101C2

About Us

<http://www.eidos.ic.i.u-tokyo.ac.jp>

Research Theme

Our central research goal is to make **high performance computing more productive and accessible** to everyone. To this end, we are working on a range of topics of various layers, including programming language design, compilers and interpreters, runtime systems, operating systems, file systems, databases and data processing systems, end-user tools for parallel processing, and new parallel algorithms.

A Message to Prospective Students

Achieving high performance on today's computers is no easy task, whether it is a distributed cloud, supercomputer, or a multicore server; our research topics broadly share the spirit of making high performance programming easy. We especially welcome students who are motivated/excited/attracted by this goal and have a gut of changing the state of the art in the field.

Recent Topics

Runtime Systems

■ MassiveThreads

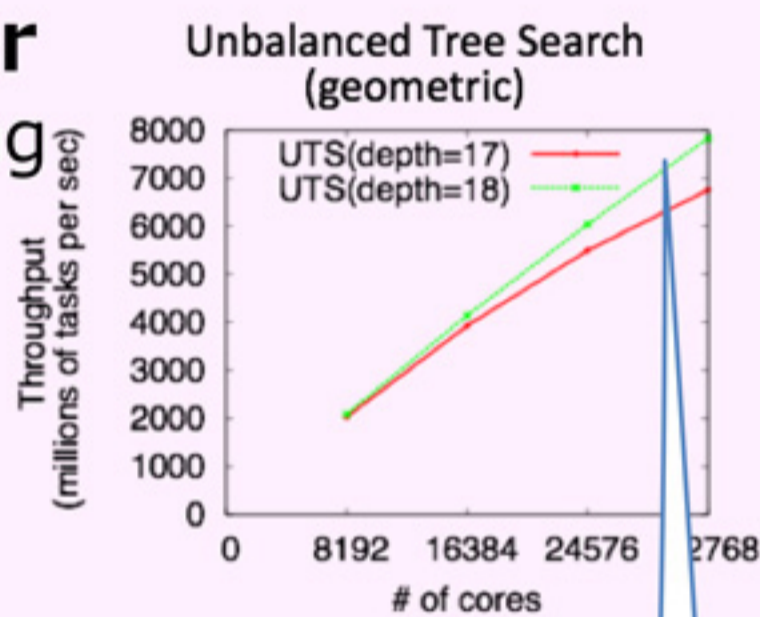
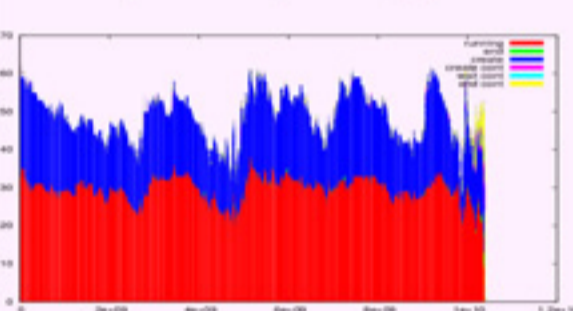
- A **task parallel system** (♠) for multicore/manycore servers
- You create millions of tasks and MassiveThreads automatically distributes them to multiple cores within a single server

■ MassiveThreads/DM and /GAS

- A task parallel system across multiple nodes (clusters/supercomputers)
- Shared memory-like programming on distributed memory machines

■ Performance analyzer and visualizer

- Capture a complete history of load balancing and visualize it as a directed acyclic graph
- Tell you who is to blame (system or you programmer?)



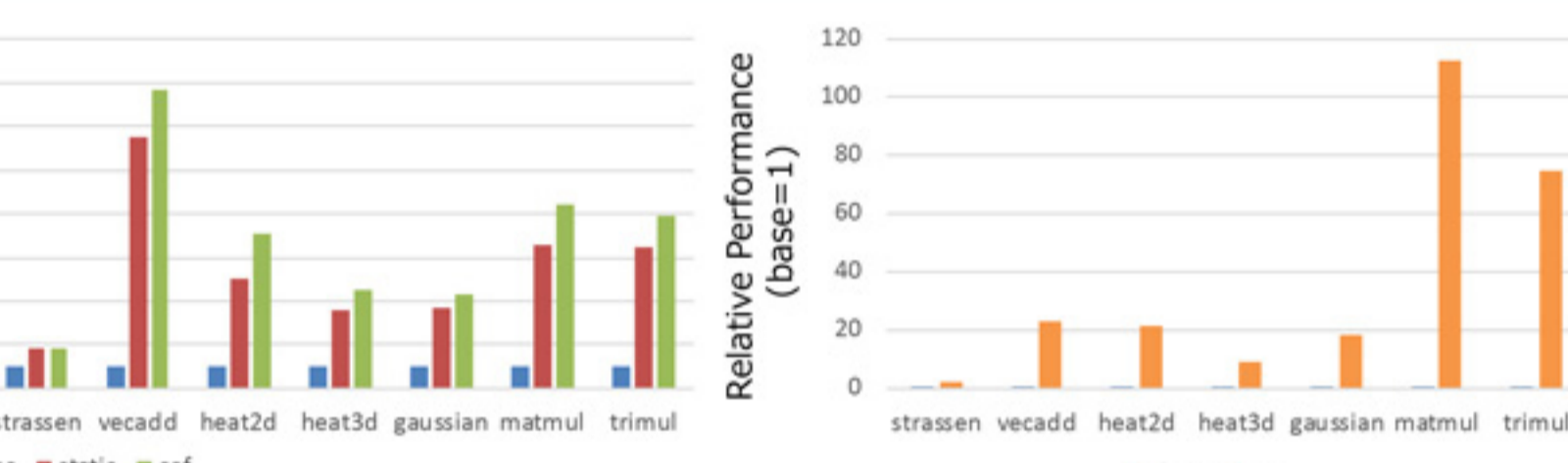
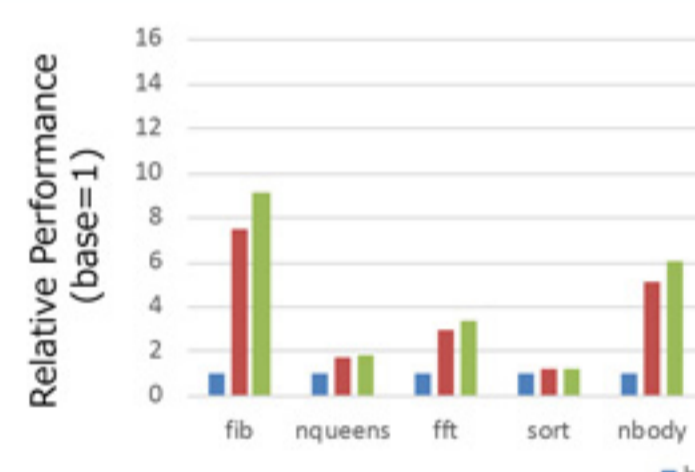
MassiveThreads/DM with >30,000 processors (> 7 billion tasks/sec)



Compiler

■ Automatic cut-off and vectorization of dynamic parallelism

- You express parallelism as dynamically created tasks
- The compiler automatically adjusts task granularity to reduce tasking overhead
- Furthermore, it compiles them to exploit **SIMD instructions** (♣)



Glossary

■ Task parallel systems (♠)

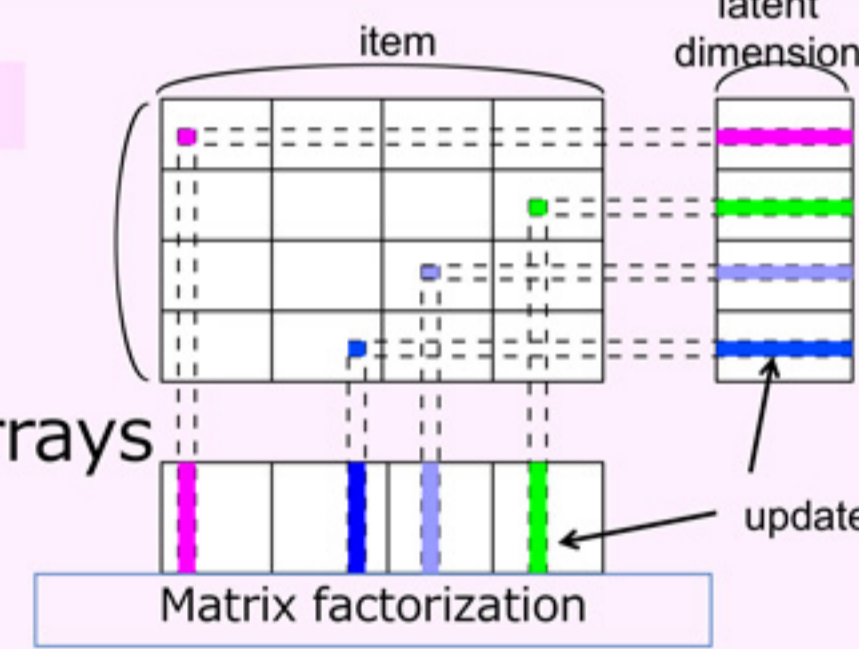
- Systems supporting dynamic creation of parallelism (tasks) at arbitrary program points of execution and their automatic load-balancing

■ SIMD Instructions (♣)

- SIMD = acronym of Single Instruction Multiple Data
- Instructions that operate on multiple data

Algorithms

- SIMD & parallel set intersection
- Parallel matrix factorization for recommendation systems
- SIMD & parallel sorting for structure of arrays
- Parallel Burrows Wheeler Transform
- Parallel Fast Multipole Method



Environment

■ HPC servers we have

- Xeon Phi (60 cores/240 threads)
- 36 cores (72 HWT) Intel Haswell-EP
- 32 cores (64 HWT) AMD Piledriver

■ Supercomputers we use

- FX10 supercomputers @ U-tokyo (76800 processors)
- Tsubame 2.5 @ Titech



Big Data Processing

- **Fast and expressive graph databases**
 - We are working toward a parallel graph databases supporting powerful regular path queries (RPQ)
- **Transactional memory maps for emerging persistent memory**
 - We are extending Linux operating system to *atomically* make changes to a file persistent
 - With emerging non-volatile memories, we expect the technology to help *atomic writes* to a persistent storage much more efficient

■ Background : why programming for high performance is difficult?

- Today's high performance computers are **parallel** machines
- Parallelism is in several layers; peak performance is the **product** of them; you ignore any and the performance is limited to a fraction of the ideal performance
 - **SIMD** (instructions that work on multiple data)
 - **Multicores** (CPUs in a single server box)
 - **Multiple nodes** (servers connected by a network)
- Today, you need to master a distinct programming skill for each layer and combine them
- Programming languages (compilers and runtime systems) should play a major role in relieving the burden, by presenting a **uniform** view of parallelism
 - You extract parallelism and the system maps them on the machine
- Automating the mapping is challenging as we must
 - Balance loads among processors
 - Minimize communication among them



Our Goal

社会を変える個人認証技術の確立を目指して

Next generation techniques for personal authentication

Our Model

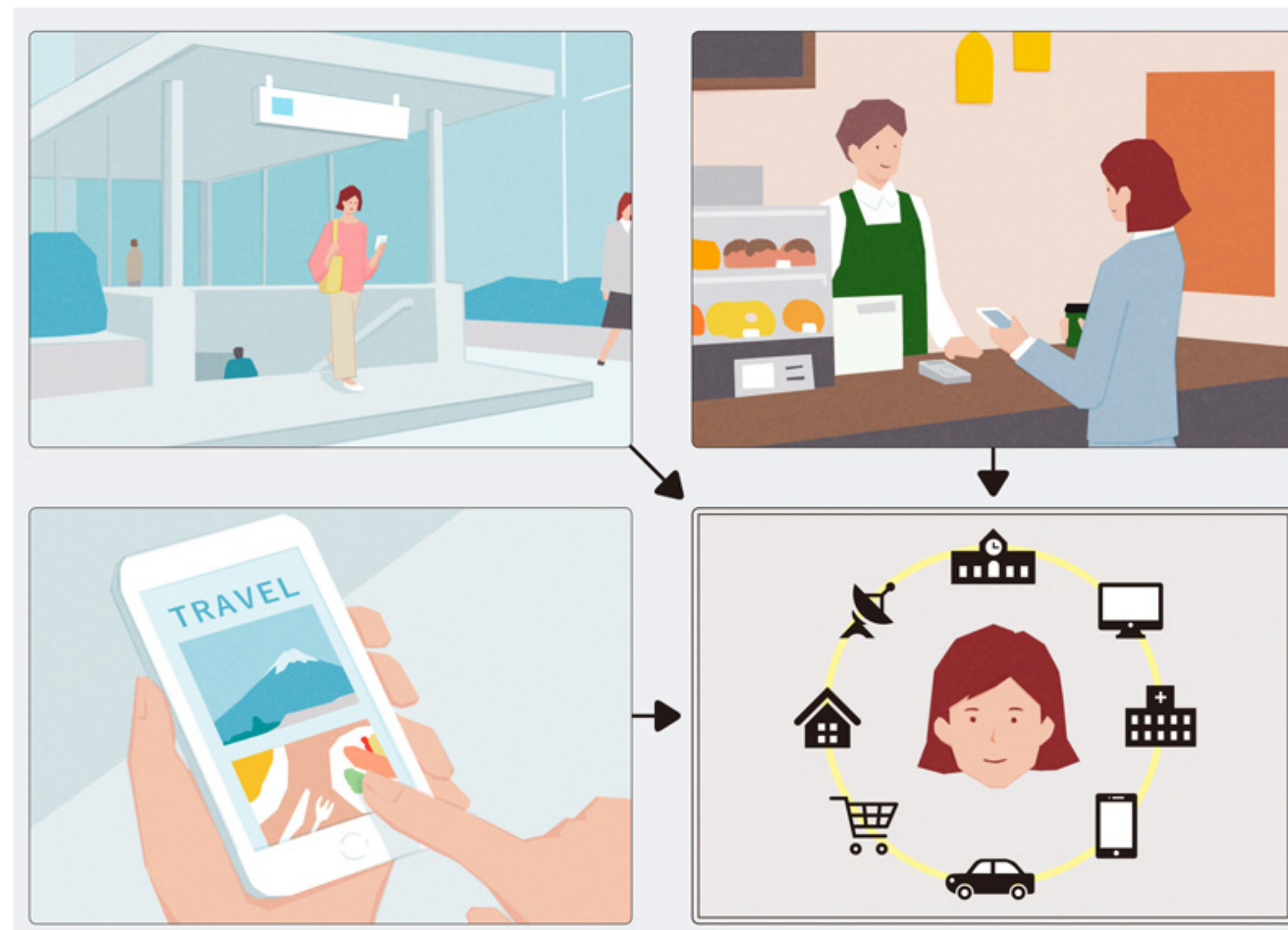
Personal Authentications and Sensor Information

- IDとパスワードから変更が急務
The recent prevalence in cases of leaked IDs and passwords mean that changes are urgently required.

- センサーや実データの解析をすることで、新たな認証技術を目指す
We will apply some techniques to user authentication so that the server can identify who the user is based on the sensor data collected by sensory input device.

Privacy Protection Techniques by Real Data

- リアルデータを利用して、どの程度の情報が集まるかによって、プライバシーのリスクが高まるかを検討
We focus on analyzing real data collected by sensory input devices, and discuss the risks that would identify how much information is necessary to complete tasks versus how much causes privacy risks.



ライフスタイル認証: 普段と同じ行動をしているかどうか、ライフログデータを解析して、個人を認証する技術。
例えば、よく使う駅、毎日夜コーヒー店、アプリの利用履歴等のデータを活用します。



You can choose any research topic in the field of information security and I would be very happy to encourage you to pursue new challenges.

次世代認証技術の確立

インターネット上における電子決済をはじめ様々な社会インフラには、詐欺や情報の詐取など様々な問題があります。実際の問題に応じた安全安心な技術の提案を行うことによって、色々な技術を利用して、次世代に活躍できる技術を探していきます。社会を変えられるような提案、待っています。