

シンギュラリティ・サロン 第7回講演

「Changing the Game」

(次世代スパコン開発から、AGI開発へ)

2015年 7月25日

齊藤 元章

(株式会社PEZY Computing / 株式会社ExaScaler / UltraMemory株式会社)

自己紹介

1968年、新潟県長岡市生まれ(47歳)

幼少時より電子機器、無線等に親しむ(当時最年少の9歳でアマチュア無線技士免許取得)

1992年:新潟大学医学部を卒業し、東京大学附属病院放射線科医局に入局

1993年:研修医2年時、東大放射線科医局の最年少で北米放射線科学会に論文採択

1994年:東京大学医学系大学院に進学し、同時に研究開発法人を設立

1997年:世界初のリアルタイムCTを発表し、スタンフォード大学教授から渡米を招聘される

1997年:米国シリコン・バレーに医療画像システム法人を創業(累計売上1,000億円超)

2003年:米国インテル社のアンドリュー・グローブ会長(当時)とクレイグ・バレット社長兼CEO(当時)の推薦で、日本人で初のComputer World Honors(米国コンピュータ業界栄誉賞)を受賞(受賞部門はAward in the Medicine Category)

2004年:博士論文を提出して東京大学医学系大学院を卒業、医学博士

2007年:次世代歯科用3次元CT装置、次世代超音波装置開発のために2法人を創業

2010年:超並列メニーコアプロセッサ開発のPEZY Computing社を創業

2012年:東京大学に産学連携共同講座を開設

2013年:東京工業大学に産学連携共同講座を開設

2013年:超広帯域カスタムDRAMの独自開発のためにUltraMemory社を創業

2014年:新たな液浸冷却手法を考案し、同技術の展開のためにExaScaler社を創業

2014年:1,024コアのメニーコアプロセッサを開発し、初めてのスパコンを開発(Green500:2位)

2015年:プレ・シンギュラリティ期について考察した「エクサスケールの衝撃」を上梓

2015年:液浸冷却に特化した第2世代スパコンを開発、2台を新規設置(Shoubu/Suiren Blue)

講演内容

0. 講演タイトルについて
1. 医師(放射線科医)時代
2. 医療診断装置・画像システム開発期
3. プロセッサ専業開発期
4. スーパーコンピュータ開発期

5. 次世代スーパーコンピュータ開発へ
6. AGI(汎用人工知能)開発へ
7. シンギュラリティは？(質疑応答セッションにて)

“Changing the Game”

- 同じ土俵の上で、同じルールで闘う
(実力だけの勝負の世界であれば。機会平等でないことが少なくない)
- 同じ土俵の上で、異なるルールで闘う
(飛び道具的なものを持ち込んだり、圧倒的な物量や後ろ盾があれば)
- 異なる土俵で、全く異なるルールで闘う
(既存の力関係、実績、評価を関係なしに、新参の弱小グループでも、何とか互角以上の戦いを挑める唯一の方法)
- 異なる土俵で、そもそも闘いを不要にしたい
(基本的に闘いを好まない、平和主義者なもので、)
- 新しい試みを持ち込んだり、新しい展開を作り出したりする方が、圧倒的に楽しく遣り甲斐がある
- ゲームを変えること、変え続けていくことで、その先に予想だにしていなかった、全く新しい方向性と可能性が拓ける

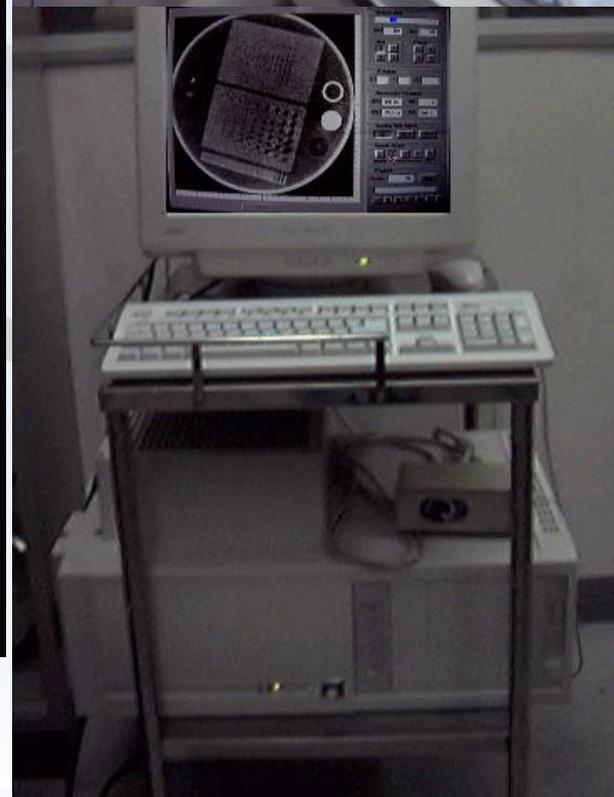
医師(放射線科医)時代

- 放射線科: 医学と工学の境界領域の専門科
(画像診断装置や放射線治療計画で、信号・画像処理、人工知能が最も求められる)
- 放射線科: 実は最も過酷な診療科(?)
(他科への画像検査・診断サービスと、自身の受持ち患者さんの診療で明け暮れる)
- とても遣り甲斐があり、有意義な経験を積ませて頂く
(人の生死に接し、そこに自身が直接関与することでしか得られない多くの物を得る)
- 一人の医師としての限界と、安定した環境への安住の懸念
(20万人の一人として生きるか、枠に嵌らない変異種として新たな道なき道を進むか)
- 研修医2年目に、世界初となる4次元描出法を開発して発表
(北米放射線科学会現地で、初めて海外医療、シリコン・バレーとの接点が出る)
- 最初の“Changing the Game”で、道を踏み出す(外す?)
(年間で数人、30年間で数百人ではなく、数百万人、数億人に寄与出来る道を選ぶ)

医療診断装置・画像システム開発期

- 革新的な診断装置や治療装置を開発して万人を救いたい
(20年前、1996年当時の診断装置や治療装置は、中身は旧態依然としていた)
- 最初はソフトウェア開発を試みたものの、、、
(ソフトウェアで実現出来ることに限界がすぐ見え、遣りたいことが殆ど出来なかった)
- 次にハードウェア開発に乗り出したものの、、、
(汎用コンポーネントを流用していたのでは、大手企業以上のことはなかなか出来ず)
- 結果として、新しいハードウェアを開発する必然性に直面
(小規模な開発グループでも、大企業に長期的に対抗できる戦略を思案)
- 次なる“Changing the Game”
(どうせ開発するなら、一番の基幹コンポーネントであるプロセッサから開発したい！)
- 更なる“Changing the Game”
(どうせなら、世界の中心であるシリコン・バレーに飛び込んで通用するか試したい！)

医療診断装置・画像システム開発期

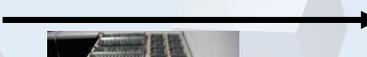


ハード、ソフトの連携開発 (Cross-Cascade Model)

Inventing new application idea from vast clinical field knowledge and experience



Precisely evaluate the need for the system performance and resources



Proprietary processor development



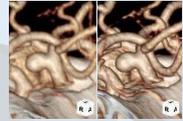
Realizing the application and clinical utilization



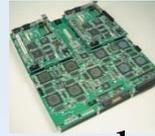
Necessary hardware system development



Inventing new application idea



Developing different type of hardware and original OS ...



Keep inventing new things ...



Next generation hardware needs



Much Larger scale and/or totally different application idea



Continuous development of various type of technologies



Continuously inventing new system, application, diagnostic devices



医療診断装置・画像システム開発期

創業者2名による合計13種類の大規模プロセッサ等開発履歴(現法人以前の11種)

極めて小規模な開発グループながらも、平均で1年半に1デザインの開発ペースで1995年から20年間に渡るプロセッサ等の開発を経験。

医療システム系法人内でのプロセッサ開発は、2005年以降の開発が鈍化。好不況の波が小さいとされる医療業界でも、収益悪化によって真っ先に開発費を削られるのがプロセッサ開発であった。過去3回の憂き目を見て、プロセッサ開発に完全に特化した事業体設立の必然性を痛感する。

| Processor | Year | Process | Die Size (mm) | Clock | Gates | Architecture | Core number | FLOPS | Power | Memory |
|--------------------|------|---------|---------------|--------|-------|--------------------|---------------|-------------|-------|--------|
| Version 1.0 | 1997 | 600nm | 8.0*8.0 | 50MHz | 1.2M | VLIW+SIMD | 1 Core/8 ALU | Fixed Point | 6W | SDR |
| Version 1.5 | 1999 | 350nm | 7.3*7.3 | 80MHz | 1.5M | VLIW+SIMD | 1 Core/8 ALU | Fixed Point | 3W | SDR |
| 3DVR Version 1.0 | 1999 | 350nm | 13.65*13.65 | 133MHz | 0.8M | Hardwired Pipeline | 2 Pipeline | - | 32W | DDR |
| Version 2.0 | 2001 | 250nm | 8.1*8.1 | 80MHz | 1.8M | VLIW+SIMD | 1 Core/8 ALU | 160M | 2W | SDR |
| 3DVR Version 2.0 | 2001 | 160nm | 15.6*15.6 | 250MHz | 3.2M | Hardwired Pipeline | 4 Pipeline | - | 20W | DDR |
| Version 2.0 shrink | 2003 | 180nm | 6.5*6.5 | 167MHz | 1.8M | VLIW+SIMD | 1 Core/8 ALU | 333M | 1W | SDR |
| Version 2.5 | 2003 | 180nm | 6.5*6.5 | 167MHz | 2M | VLIW+SIMD | 1 Core/8 ALU | 333M | 2W | DDR |
| DBF Version 1.0 | 2003 | 180nm | 11.5*9.6 | 40MHz | 2.5M | Hardwired Pipeline | - | - | 10W | - |
| Version 3.0 | 2005 | 130nm | 16.5*12.0 | 333MHz | 34M | RISC+VLIW+SIMD | 8 Core/40 ALU | 13.3G | 19W | DDR |
| Version 3.0 B | 2005 | 130nm | 9.5*12.0 | 250MHz | 20M | VLIW+SIMD | 1 Core/8 ALU | 8G | 6W | DDR |
| 3DVR Version 3.0 | 2008 | 130nm | 10.5*10.5 | 333MHz | 5.5M | Hardwired Pipeline | 2 Pipeline | - | 10W | DDR2 |

プロセッサ専門開発期

創業者による13種類の開発済みプロセッサのうちの現法人分と、今後の開発計画

2010年に設立したPEZY Computingでは、事業の状況に関わらずに、2年毎に1世代のプロセッサ開発の速度を維持したい。

開発中のPEZY-SC2では、一気に4倍となる4,096コアの集積を実現する計画。

PEZY-SC3以降では、集積度を更に上げ、高価なIPを再利用して、最先端プロセスを早期に利用可能とする目的から、Multi-Dieプロセッサ構成も計画。

今後も世代毎に搭載するコア数を2倍ずつ増やして、最終的には10kコアを大きく超える集積度を目指す。

その過程で、HMCやHBMでも不足することが確実なメモリ帯域を十分確保するため、独自のメモリIFと、低消費電力・超広帯域・大容量を実現する積層メモリを、グループ内で開発。

| Processor | Year | Process | Die Size (mm) | Clock | Gates | Architecture | Core number | FLOPS Double/Single | Power | Memory |
|---------------------------------|------|---------|---------------|---------|-------|-----------------|-------------|------------------------|-------|---|
| PEZY-1 | 2012 | 40nm | 21.0*16.8 | 533MHz | 220M | RISC+SMT (MIMD) | 512 Core | 166/333G | 35W | DDR3/Wide IO |
| PEZY-SC | 2014 | 28nm | 21.1 *19.5 | 733MHz | 580M | RISC+SMT (MIMD) | 1,024 Core | 1.5/3.0G | 70W | DDR4/Custom Ultra-Wide IO |
| PEZY-SC2 (under development) | 2016 | 14/16nm | TBD | 1GHz | 2.4G+ | RISC+SMT (MIMD) | 4,096 Core | 8.2/16.4T | 100W | HMC or HBM/ Custom TCI Stacked DRAM |
| PEZY-SC3 (in plan) | 2018 | 10nm | TBD | 1.25GHz | 5G+ | RISC+SMT (MIMD) | 8,192 Core | 20/40T | TBD | Custom TCI/TSW Stacked DRAM |
| PEZY-SC4 (in plan) | 2020 | 7nm | TBD | 1.5GHz | 10G+ | RISC+SMT (MIMD) | 16,384 Core | 50 /100T | TBD | Custom TCI/TSW Stacked DRAM |

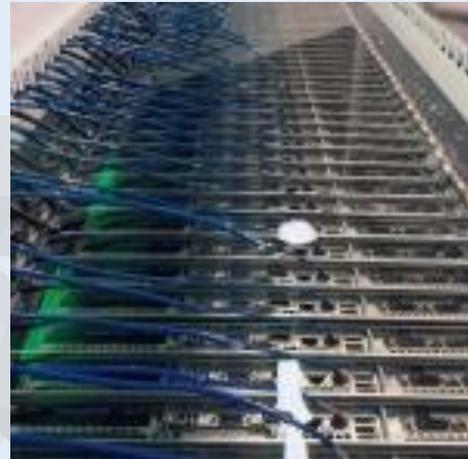
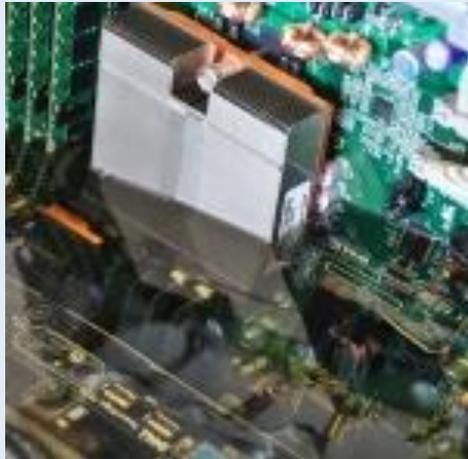
プロセッサ専業開発期

| Processor | | PEZY-1 | PEZY-SC |
|-----------|---------------------|--|--|
| 製造プロセス | | TSMC 40G (40nm) | TSMC 28HPM (28nm) |
| | ダイサイズ | 335mm ² | 412mm ² |
| コア性能 | 動作周波数 | 533MHz | <u>733MHz</u> |
| | キャッシュ | L1: 128KB, L2: 1MB, L3: 無し | L1: 1MB, L2: 4MB, <u>L3: 8MB</u> |
| 周辺回路 | 動作周波数 | 66MHz | 66MHz |
| IPs | 内蔵CPU (管理・デバッグ用) | ARM926 x 2 Cache L1: 16KB*2, L2: None | ARM926 x 2 Cache L1:32KB*2, <u>L2: 64KB</u> |
| | PCIe | PCIe Gen2 x 4Lane 6Port (2GB/s x 6 = 12GB/s) | PCIe Gen3 x 8Lane 4Port (8GB/s x 4 = <u>32GB/s</u>) |
| | DDR | DDR3 64bit 1,333MHz 4Port (10.6GB/s x 4 = 42.6GB/s) | DDR4 64bit 2,400MHz 8Port (19.2GB/s x 8 = <u>153.6GB/s</u>) |
| コア (PE) 数 | | 512 PE + 64 PE (redundancy) | <u>1,024 PE</u> |
| 演算性能 | | 533G Flops (単精度浮動小数点) 266G Flops (倍精度浮動小数点) | 3.0T Flops (単精度浮動小数点) <u>1.5T Flops</u> (倍精度浮動小数点) |
| 消費電力 | | 50W (Leak: 20W, Dynamic: 30W) | 60W (Leak: 10W, Dynamic: 50W) <u>46W@533MHz</u> (PEZY-1以下) |
| パッケージ | DDR版 | 40*40mm (1,517pin) | 47.5*47.5mm (2,112pin) |

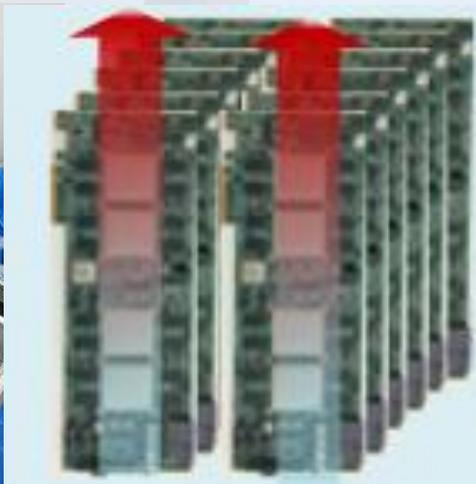
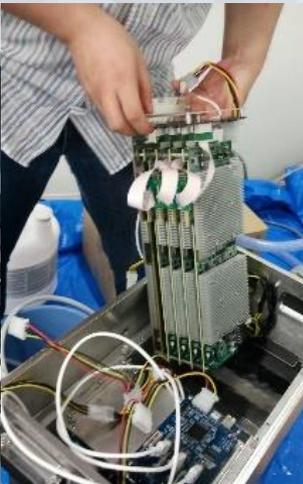
スーパーコンピュータ開発期

- PEZY-SC開発がほぼ完了し、Tape-out目前まで漕ぎ着ける
(巨大なプロセッサであったため、本当に一発で動作するかどうかは誰にも分からず)
- 2014年3月26日、我々の未来を変える運命的な分岐点
(東工大から理化学研究所に移動される直前の、牧野淳一郎先生に御面会)
- 全面的支援のお約束と、スパコン開発への道を明示される
(目指す方向性と技術的素性、開発能力を御評価頂いて、背中を強く押して頂く)
- とは言え、僅か10名で本当にスパコンを開発出来るのか？
(でも、もし本当に開発出来たとしたら、大きな可能性が拓け、未来が広がるはず…)
- 逆境を逆手に取っての、“Game Changer”を模索する
(全体でも2週間、正味は3日間で新しい液浸冷却手法を考案し、特急開発に着手)
- 7か月後の10月25日、液浸冷却スパコンが何とか形になる
(社内でも、誰一人として本当に完成し、稼働するとは思っていなかった…らしい)

スーパーコンピュータ開発期



米国の油浸冷却技術を参考に、多々存在する問題点の全てを抜本的に解決する手法を考案



高沸点フッ化炭素を冷媒とした、社内での基礎実験の様子(2014年4月から5月に掛けて)

スーパーコンピュータ開発期



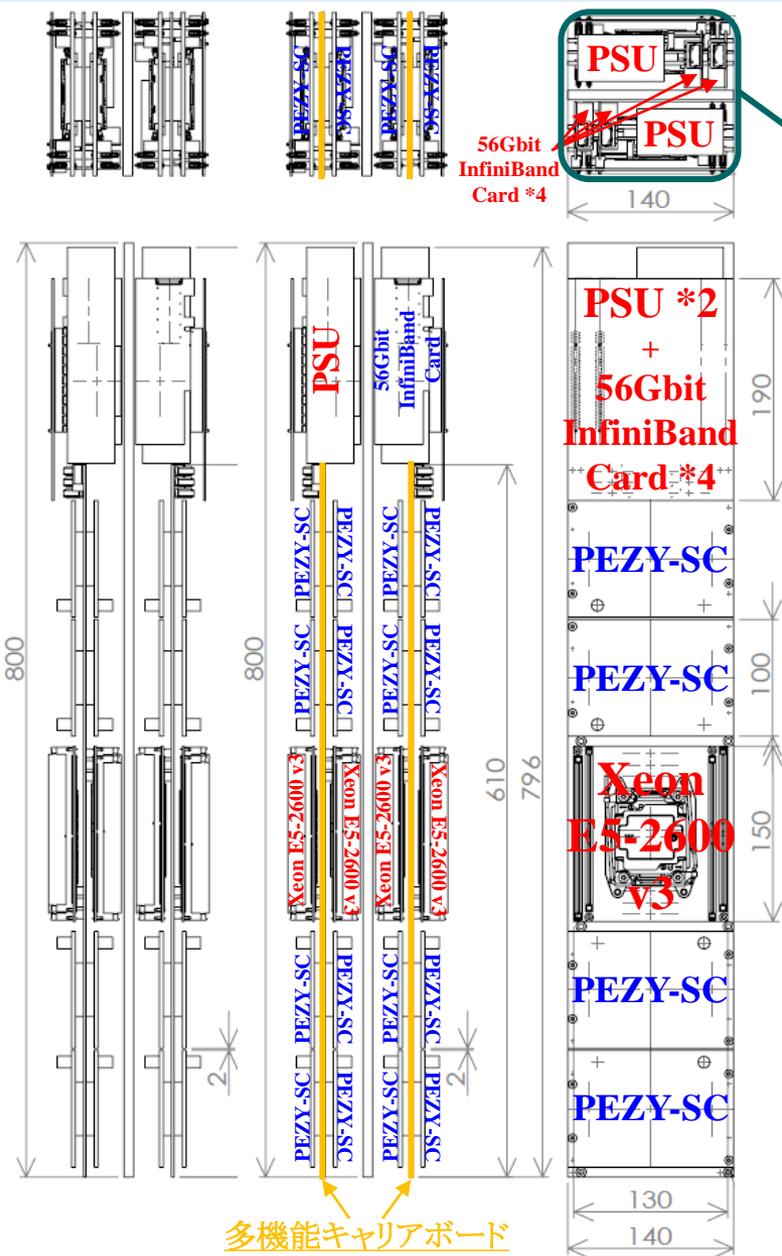
国の研究機関「高エネルギー加速器研究機構 (KEK)」に設置された初代スパコン「睡蓮」
(2014年10月)

スーパーコンピュータ開発期

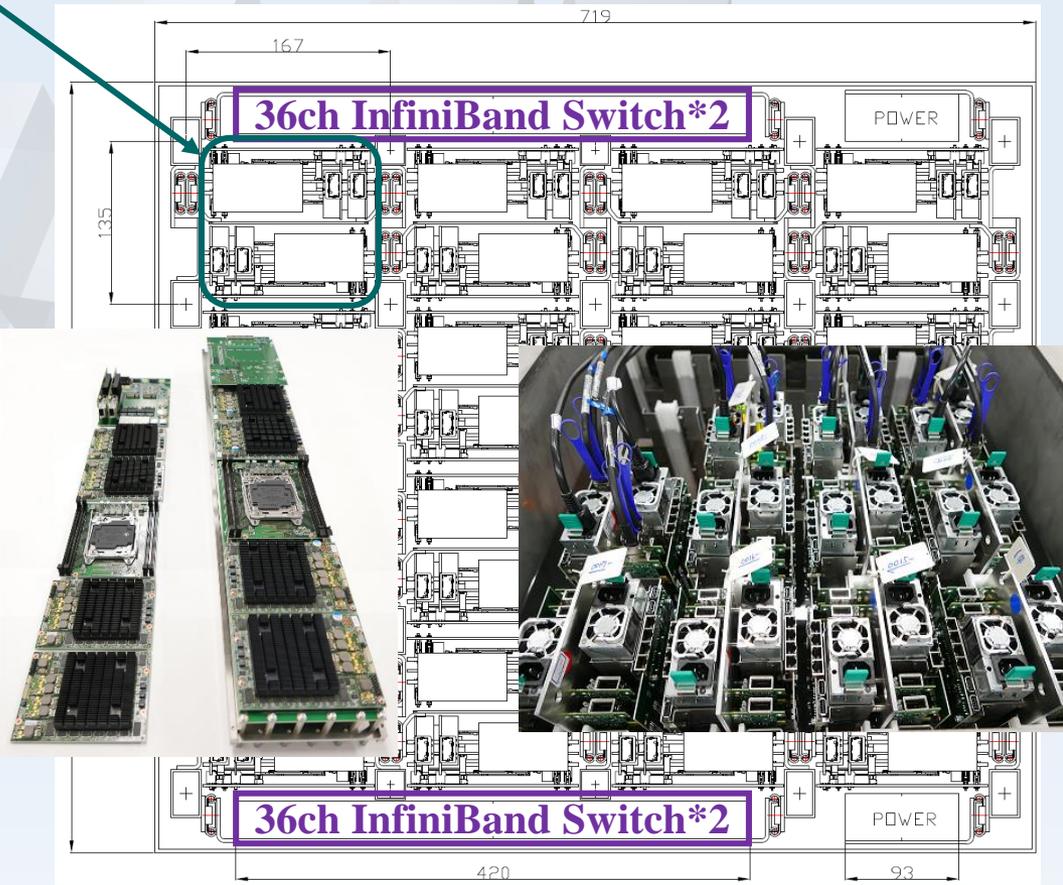
10月31日締切りの「Top500」では369位 (178.1TFLOPS)、
11月14日締切りの「Green500」では4.95GFLOPS/Wを計測して世界2位に



スーパーコンピュータ開発期

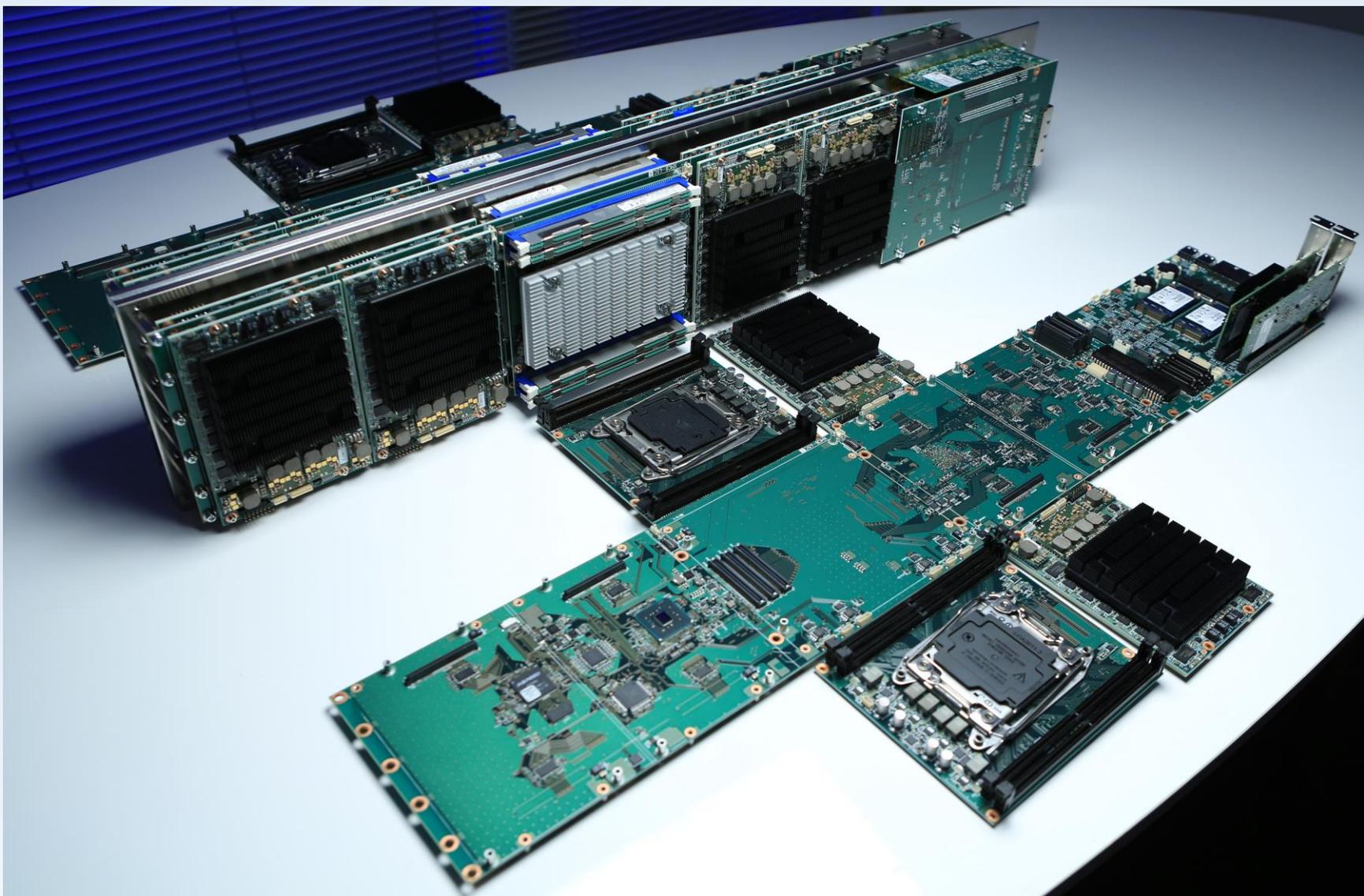


新型液浸槽「ESLiC-32」上面図
(Xeon*64, PEZY-SC*256, Switch*4)



← 新型液浸槽専用に新たに開発する、
モジュール+キャリアボード基板構成

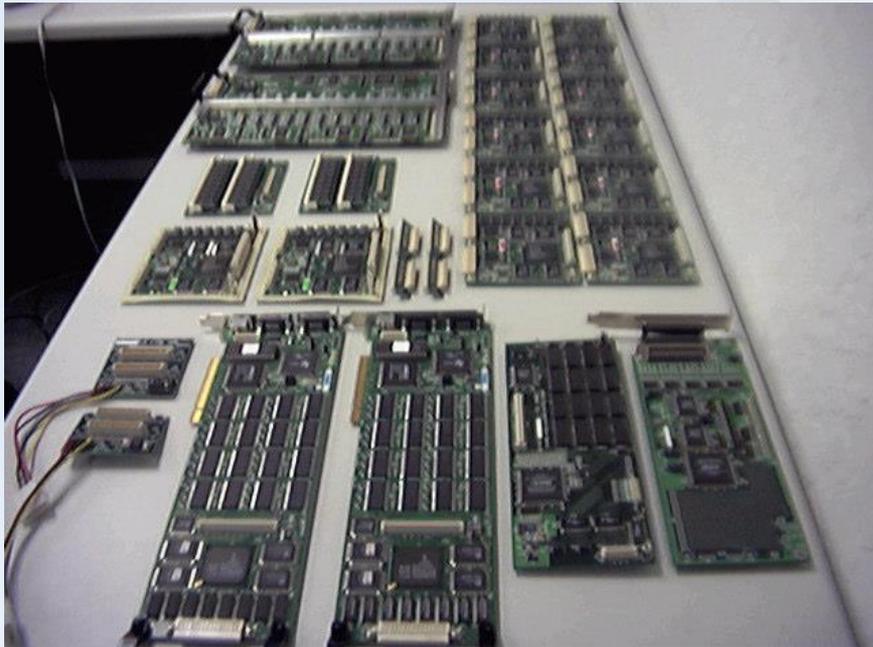
スーパーコンピュータ開発期



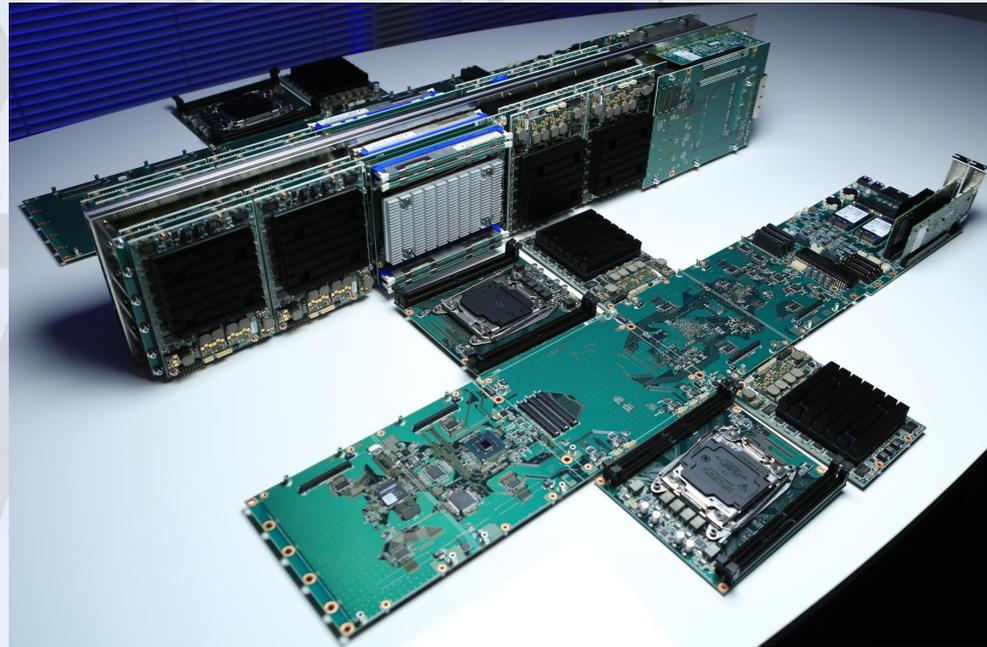
2015年2月から開発を開始して、液浸冷却に完全に特化した第2世代スパコンを、全てのコンポーネントを自社開発しながら、再び僅か5か月で完成させる。

スーパーコンピュータ開発期

1997年のリアルタイムCT再構成システム
のハードウェア構成



2015年(18年後)の第二世代液浸冷却スパコン
のハードウェア構成



やっていることは、一見余り変わっていない・・・
（「ブレていない」or「進歩していない」？）
（どうやら齊藤は、どうしても複雑なハードウェアを作りたいらしい、..）

スーパーコンピュータ開発期



RIKEN : Colors 知の旅人たち 全章 [Read More](#)

最新のExaScaler-1.4による5台構成の2 PetaFLOPS級スパコンが、「Shoubu (菖蒲)」として、理化学研究所の情報基盤センターに設置された。富士通の最新スパコンFX-100による、「HOKUSAI」に隣接して運用される。



研究論文 (STAP細胞) に関する取組み、情報等について (2015年4月9日)
4つの方針：研究不正の調査・科学的検証の実施・研究論文の取下げ・再発防止の取組みについて掲載しております。

- 一般の方 >
- 研究者・学生の方 >
- 企業の方 >
- 報道関係の方 >
- 理研関係者 >

プレスリリース

トピックス

2015年6月25日

重力によって移動方向が変わらないオーキシンを発見



理研らの国際共同研究グループは、植物ホルモン「オーキシン」の一種であるフェニル酢酸 (PAA) が、重力によって移動方向が変わらないユニークな特徴を持つことを発見しました。 [続きを見る...](#)

2015年6月25日

ExaScaler及びPEZY Computingが、理化学研究所と共同研究契約を締結し理研情報基盤センターに2 PetaFLOPS級の液浸冷却スーパーコンピュータ「Shoubu (菖蒲)」を設置



株式会社ExaScalerと株式会社PEZY Computingは理研と共同研究契約を締結し、理研情報基盤センターにExaScaler-1.xの2PetaFLOPS級の液浸冷却スーパーコンピュータ「Shoubu (菖蒲)」を設置します。 [続きを見る...](#)

スーパーコンピュータ開発期

最新(2015年6月末時点)の「Top500」結果

「Shoubu (菖蒲)」: 160位

・演算性能: 412.67TFLOPS

大規模システムの設置に時間を要し、全体の6割の構成で、500MHzという低い周波数での測定に留まる。

「Suiren (睡蓮)」: 366位

・演算性能: 206.57TFLOPS

基本的な性能改善作業は完了しており、次回に向けては現状を維持。

「Suiren Blue (青睡蓮)」: 393位

・演算性能: 193.91TFLOPS

Suirenと比較して、メモリ容量が半分で性能が低かったため、次回はこれを増強することで、Suirenを超える性能の実現を目指す。

スーパーコンピュータ開発期

最新(2015年7月時点)の「Green500」結果

(何故か、未だに発表されていません。)

「Shoubu (菖蒲)」: ___ 位 (おそらく1位ではないかと、)

・消費電力効率性能: 7.032 GFLOPS/W

「Suiren Blue (青睡蓮)」: ___ 位 (おそらく2位ではないかと、)

・消費電力効率性能: 6.842 GFLOPS/W

「Suiren (睡蓮)」: ___ 位 (おそらく3位ではないかと、)

・消費電力効率性能: 6.217 GFLOPS/W

(2014年11月の世界1位の値は、5.272GFLOPS/W)

スーパーコンピュータ開発期

ステルス

Ben R. Rich SKUNK WORKS

スカンク・ワークスの秘密

ベン・R・リッチ[著] 増田興司[訳]

戦闘機



齊藤の座右の書:「ステルス戦闘機(スカンク・ワークスの秘密)、1997年 Ben R. Rich 著」

製造段階で数千人規模の作業となる最先端の超高度戦闘機技術開発も、主要な開発は10人程度の少数精鋭による極小さなチームで短期間に、しかも矢継ぎ早に行われていた。

(書籍紹介文)

例のない創造型組織の責任者が明かす成功のノウハウ

数々の最新鋭軍用機を開発した知られざる
[創造型組織]の全て

スカンク・ワークス(ロッキード先進開発計画)はF-104・U-2・SR-71・F-117Aなどのハイテク機を短期間・低コストで開発

〈少数精鋭・独立・秘密〉を第一義にした例のない組織の責任者がその全貌を明かす



講談社

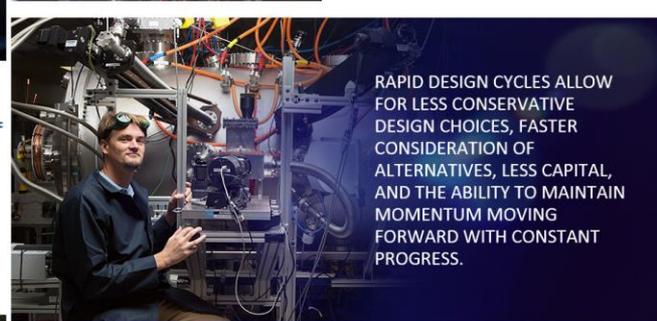
スカンク・ワークスの最新の開発成果

2014年10月に、「CFR:小型熱核融合炉」の開発成功を発表

“**The compact size is the reason** that we believe we will be able to **create technology quickly**. The smaller the size of the device, the easier it is to **build up momentum** and develop it faster. Instead of taking five years to design and build a concept, it takes **only a few months**.” (スカンク・ワークス的小規模チームが小型システムを開発することで、開発速度が更に飛躍的に増すことに。液浸冷却スパコンも同じコンセプト)



To mimic the energy created by the sun and control it here on earth, we're creating a concept that can be contained using a magnetic bottle. The bottle is able to handle extremely hot temperatures, reaching hundreds of millions of degrees. By containing this reaction, we can release it in a controlled fashion to create energy we can use.
The heat energy created using this compact fusion reactor will drive turbine generators by replacing the combustion chambers with simple heat exchangers. In turn, the turbines will then generate electricity or the propulsive power for a number of applications.



次世代スーパーコンピュータ開発へ

・「今回の“Game Changer”は？」

- 1) MIMD型プロセッサの圧倒的な超々メニーコア化
- 2) 低消費電力・大容量の積層DRAMを独自開発
- 3) プロセッサ・DRAM間の無線による超広帯域接続
- 4) スイッチチップを独自開発し、ファットツリーを1チップ化
- 5) インターコネクトを磁界結合で無線化、無ケーブル化
- 6) 3重合液浸冷却で超高集積化・小型化・低消費電力化
- 7) 上記全てを2年で開発し、2020年までに2世代分進化

・詳細については、配布の資料を御参照下さい。

・おそらくは、7項目のうちの1項目でも実現出来たとしたら、スパコン業界では相当な“Changing the Game”になると推測。(因みに、今回も社内では誰一人として実現を信じていないらしい?)

次世代スーパーコンピュータ開発へ

| Processor | | PEZY-SC | PEZY-SC2 |
|-----------|-------|---|--|
| 製造プロセス | | TSMC 28HPM (28nm) | TBD (14-16nm FinFET) |
| | ダイサイズ | 412mm ² | 400-500mm ² |
| コア性能 | 動作周波数 | 733MHz | <u>1GHz</u> |
| | キャッシュ | L1: 1MB, L2: 4MB, L3: 8MB | 50MB in total (TBD) |
| 周辺回路 | 動作周波数 | 66MHz | 66MHz |
| IPs | 内蔵CPU | ARM926 x 2 管理・デバッグ用 | 64bit CPU (TBD) 汎用演算用 |
| | PCIe | PCIe Gen3 x 8Lane 4Port (8GB/s x 4 = 32GB/s) | PCIe Gen3/4 x 8Lane 6Port (48-96GB/s) |
| | DRAM | DDR4 64bit 2,400MHz 8Port (19.2GB/s x 8 = 153.6GB/s) | Custom Stacked DRAM 8Port (500GB/s x 8 = 4.0TB/s) |
| コア (PE) 数 | | 1,024 PE | <u>4,096 PE</u> |
| 演算性能 | | 3.0T Flops (単精度浮動小数点) 1.5T Flops (倍精度浮動小数点) | 16.4T Flops (単精度浮動小数点) <u>8.2T Flops</u> (倍精度浮動小数点) |
| 消費電力 | | 60W (Leak: 10W, Dynamic: 50W) | 100W (Leak: 10W, Dynamic: 90W) |
| パッケージ | | 47.5*47.5mm (2,112pin) | Multi-Die Package (TBD) |

次世代スーパーコンピュータ開発へ

次世代システム「**ExaScaler-2.0**」開発 (試験稼働開始:2016年末予定)

- ・**4,096コア@1GHz、8.2TFLOPS**のメニーコアプロセッサ「**PEZY-SC2**」(**PEZY社**)
- ・独自IF接続(合計**4TB/s**)による**大容量・超広帯域DRAM**(**UltraMemory社**)
- ・高効率・省スペースの第二世代「**液浸冷却システム**」(**ExaScaler社**)

上記3つの要素技術が密接に統合されることによって構築される次世代スパコンシステム「ExaScaler2.0」は、15-20GFLOPS/Wの電力消費効率を達成し、タワー筐体1台の体積で5PetaFLOPSの性能の到達を目標。

次々世代システム「**ExaScaler-3.0**」開発 (試験稼働開始:2018年末予定)

- ・**8,192コア@1.25GHz、20TFLOPS**のメニーコアプロセッサ「**PEZY-SC3**」(**PEZY社**)
- ・独自IF接続(合計**10TB/s**)による**大容量・超広帯域DRAM**(**UltraMemory社**)
- ・高効率・省スペースの第三世代「**液浸冷却システム**」(**ExaScaler社**)
- ・TCI技術による**Multi-Die接続、超高速・多チャンネル・3次元積層Switch**(**TCI社**)

上記4つの要素技術を密接に統合させて構築する、次々世代スパコンシステム「ExaScaler3.0」は、25-30GFLOPS/Wの電力消費効率を達成して、タワー筐体1台の体積で15PetaFLOPS超の性能の到達を目標とする。67台分で1ExaFLOPSの性能を実現し、ExaScalerの名を具現化したい。

AGI (汎用人工知能) 開発へ

• 再びいつもながら、「今回の“Game Changer”は??」

• 現状についての認識

- 1) ディープラーニング系は、既に先行者多数
- 2) ソフトウェアによるアプローチは、既に先行者多数
- 3) 現在のコンピューティングの先に、AGI実現は困難
- 4) 現在のコンピューティングの先に、意識創発は困難
- 5) 特化したハードウェアによるアプローチは、極めて少数
- 6) 特化したハードウェアは、実際の脳構造と比較すると極めて小規模であり、依然として本質的とは思われない

• 結論：全く新しい脳機能ハードウェアを開発して、過去最大の“Game Change”を起こすことで、AGIを実現する！

大脳皮質アルゴリズムには専用アーキテクチャが必要 (但し、既存の最新CMOS半導体製造技術の範囲内での実現を提唱)



Mapping HTM Algorithms to Hardware

HTM algorithms requires hardware matched to the algorithm's needs

1. High Connectivity (more expensive than the computational requirements)
2. Local Memory and Parameter Storage
3. Simple, Low Precision Computation
4. Configurable / Adaptable
5. Sparse activity

Dr. Dan Hammerstrom, NICE 2014

Conventional processors are a poor match to cortical algorithms:

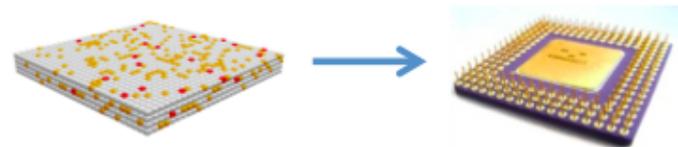
- Constrained: processor/memory partition, limited parallelism
- Excessive: high precision, tiered caches, complex instruction sets, pipelines, etc.



Conventional Solutions

Custom architectures can address HTM's requirements:

- High-risk exotic devices unnecessary
- Utilize conventional CMOS fabrication optimized for HTM architecture/computational model
- Can benefit from latest advances in CMOS.



Hierarchical Temporal Memory

Specialized HTM processor

残念ながら、この方法では集積度は殆ど上げられない (TSVと、極薄化しない半導体ウェハと、接着剤使用では早々に限界が)

3D Wafer Scale Integration of wafers

frontside wiring on Final wiring wafer (e.g.- GMP-to-GMP communication)

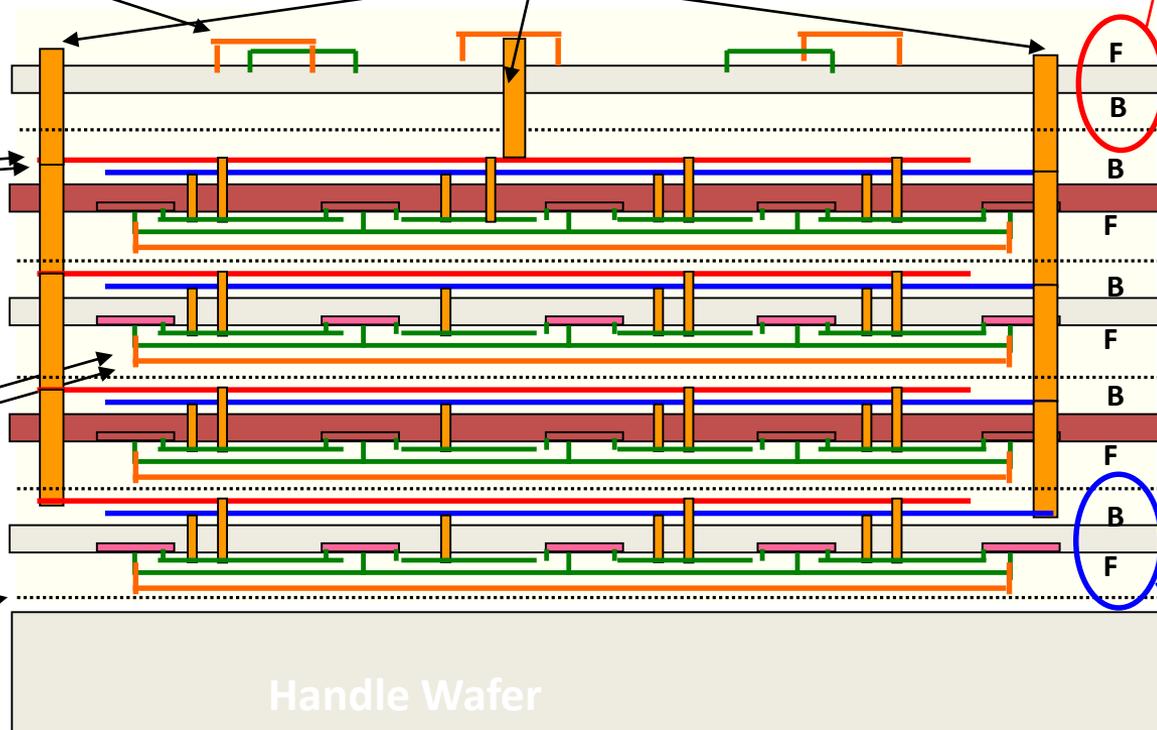
inter-strata TSV (e.g, communication from GMP level to SSP level, and between SSPs, e.g, for power & ground)

Preferred orientation, but not required

backside wiring on SSP wafer (e.g.- power & ground,)

frontside wiring on SSP wafer (e.g.- SSP-to-SSP communication)

bonding interface



Final wiring wafer

SSP Wafer #4

SSP Wafer #3

SSP Wafer #2

SSP Wafer #1

Shown here for 4 wafers

B=backside
F=frontside

AGI (汎用人工知能) 開発へ

- 現在のコンピューティングの方向性は、SIMD・メニーコア化で数千~数万コアに、数~数十レーンのインターコネク트에
- 脳では1,000億のニューロンに、100兆以上のシナプス結合 (1個のニューロンに1,000個以上のシナプス結合) で、全く逆
- この比率を保ちつつ超々高集積な、全く新しいプロセッサ (これまでのプロセッサと構造も概念も異なる) の開発が必要となる。この実現を可能とするのが、

磁界結合技術

+

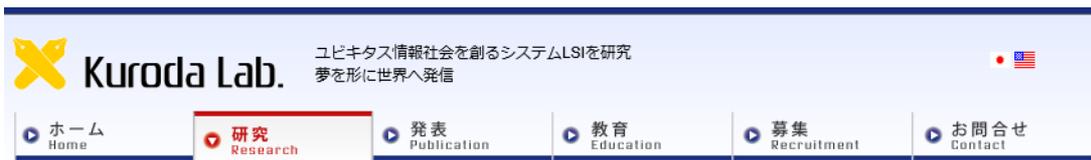
半導体ウェハ極薄化技術

+

フュージョン・ボンディング (ウェハ直接貼合わせ)

磁界結合技術の積極的な採用

慶応大学理工学部電子工学科黒田研究室の黒田忠広教授が2004年から研究されて来た、**超高速・低消費電力の伝送**を可能とする**磁界結合 (Inductive Coupling)**技術を、独自プロセッサと独自メモリそれぞれの2.5次元接続、3次元積層接続と、そのIF技術として積極的に採用します。



ホーム > 研究 > 研究メンバー > 黒田 忠広 (くろだ ただひろ)

研究メンバー



黒田 忠広 (くろだ ただひろ)

誘導結合を用いた近接場ワイヤレスチップ間通信や画像認識などを研究。60件の招待講演と21件の著書を含む200件以上の技術論文を発表。100件以上の特許を申請。

メールアドレス
kuroda@elec.keio.ac.jp

1982年東京大学工学部電気工学科卒。工学博士。同年東芝入社。CMOS SRAM、ゲートアレイ、スタンダードセルを開発。1988年から1990年までカリフォルニア大学バークレイ校にて客員研究員としてLSI CADを研究。1990年に東芝に戻り、BICMOS ASIC、ECLゲートアレイ、通信用高速CMOS LSI、携帯・マルチメディア用低電力LSI、無線通信用LSIの研究開発に従事。1995年に基板バイアスでトランジスタのしきい値電圧を制御する可変しきい値電圧技術(VTCMOS)を発明しチップに応用。1997年にDC-DCコンバータをチップに搭載して電源電圧を制御する可変電源電圧(VS)方式を提案し、世界で始めてマイクロプロセッサとDSPに応用。2000年に慶應義塾大学に移り、2002年より教授。広島大学とカリフォルニア大学バークレイ校の客員教授を兼任。誘導結合を用いた近接場ワイヤレスチップ間通信や画像認識などを研究。60件の招待講演と21件の著書を含む200件以上の技術論文を発表。100件以上の特許を申請。VLSI回路シンポジウムのシンポジウム委員長、ASP-DACの副委員長、A-SSCC、ICCAD、SSDMの分科委員会委員長、VLSI回路シンポジウム、CICC、DAC、ASP-DAC、ISLPEDなどのプログラム委員を歴任。2005 IEEE System LSI賞、2005 P&I Patent of the Year賞、2006 LSI IP Design賞、2006 IP/SoC Best Design Paper賞、2007 ASP-DAC Best Design賞、2008 電子情報通信学会業績賞などを受賞。IEEEフェロー、IEEE SSCS管理委員会メンバー、IEEE上級講師。



黒田研究室について

研究目標と研究プロジェクト



講義情報を確認する

日程と講義資料ダウンロード



黒田教授について

経歴と連絡



学部3年生へ

黒田研に興味のある方



発表論文

対外的な論文の紹介



お問い合わせ

メンバーへのご連絡



- 磁界結合を利用したチップ間無線通信**
SiP (System in Package) 積層実装されたチップ間を磁界結合で通信することを研究します。従来のCMOS高速リンクに比べて、3桁以上の低電力化を図ります。非接触通信の利点を生かした新しい応用も提案します。
- 近距離パルス無線通信**
数cmの距離をパルスで無線接続する通信技術を研究します。数Gb/sの高速通信を有線に比べて低コストで高信頼に実現し、可動性の利点を生かせる応用を提案します。また、低速通信では消費電力の大幅削減を追求します。
- 画像認識**
人の顔を検出してビデオカメラなどの焦点や露光や拡大率を自動制御する技術を研究します。画像検索、セキュリティー、物体追尾などへの応用も提案します。アルゴリズム (SW) とLSI (HW) の総合最適設計が重要です。
- センサネットワーク**
環境情報から人に役立つ情報を抽出し加工する技術を研究します。センサのデータから行動を分析し、知識労働の生産性の向上に役立てます。さらに人間の知覚や思考にまで踏み込んだ分析を目指します。
- コンバージョン**
無線通信用のADC/DACのデータ変換と、無線給電でのAC/DCやチップの電力制御用のDC/DCの電力変換を研究します。
- 高速シリアルリンク**
40Gb/s以上の高速チップ間シリアル接続技術を研究します。60GHz CMOS無線接続技術も研究します。

「磁界結合」技術普及の突破口

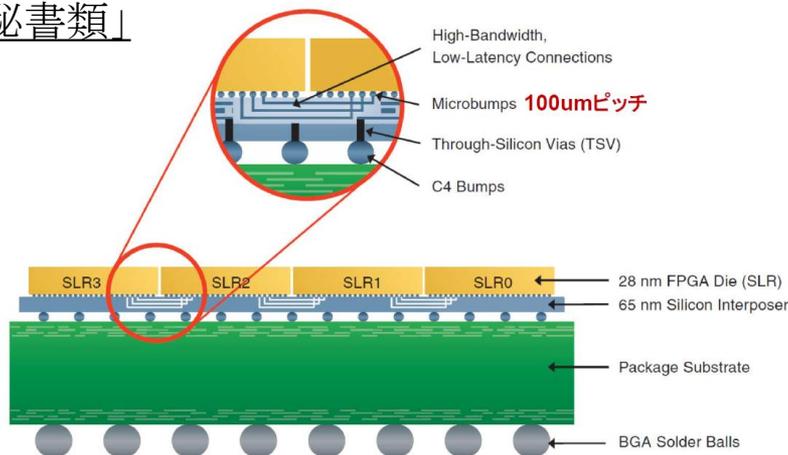
PEZY社が、2013年に東京工業大学大場研究室とディスコ社と共同開発した半導体ウェハ極薄化技術では、30cmウェハ全体を4umまで均質に極薄化して、完全高速動作と標準的な耐久性を確認。

半導体ウェハ極薄化によりアンテナ部を極小化(100分の1以下)して、伝送距離(磁界部と電送部の双方)を大幅に縮小。大幅な伝送速度向上と、劇的な転送電力の削減が可能に。

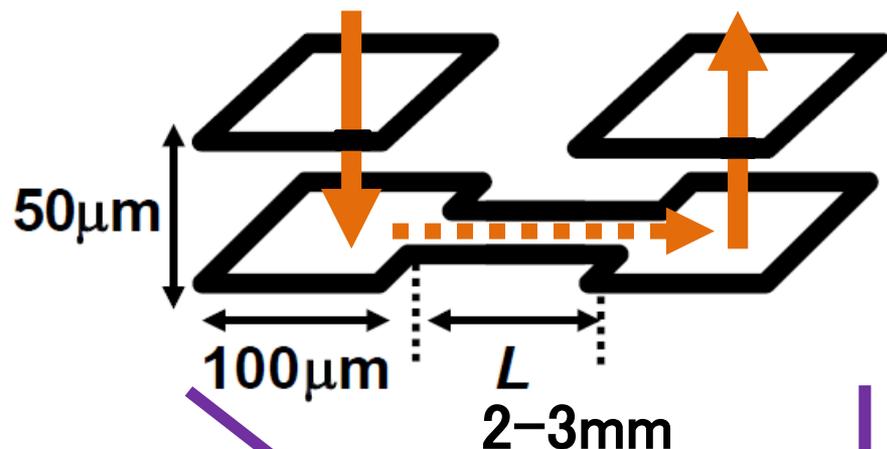
μバンプ(100μピッチ) + Siインタポータ

1.6Gb/s/link

「守秘書類」

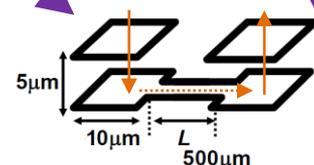


現在の一般的な2.5次元実装では、Siインタポータが必要で、伝送距離が長く、TSV工程も不可欠



2-3mm
アンテナ部面積が
100分の1になり、
伝送距離も
極小に

2.5次元インタポータを用いない2次元磁界結合IF接続についても考案済み。
(特許申請準備中)



NSPU: Neuro-Synaptic Processing Unit

- 1 μm の半導体ウェハ厚実現で、磁界結合アンテナ(コイル)は2 μm 角の小ささに。そして有線配線は不要で、アンテナの上下に論理回路の配置が可能。インターコネクットの集合体の中にコアを埋め込んでプロセッサを構成出来ることに
- 計算上は、800 cm^3 (人間の脳は1,300 cm^3)の大きさに100兆個のインターコネクットを実現しつつ、その中に1,000億コアを埋め込むことが可能
- 人間のシナプスと比べ、半導体で構成・製造されるNSPUは、電氣的接続は伝送速度(120m/秒対30万km/秒)で250万倍、信号発火頻度(周波数が数Hz対数GHz)で10億倍も高速
- 10億倍の速度を考慮すると、6Lの体積で73億人分の脳を実現することが可能に。恐らく、AGIとASI(人工超知能)への最短経路で、最有力の実現手法であると確信

NSPU: Neuro-Synaptic Processing Unit

- 2 μm の半導体ウェハ極薄化と、4 μm の磁界結合アンテナの集積については2018年にも実現を見込む。1 μm の極薄化と2 μm の磁界結合アンテナの集積は、2020-22年実現を予想
- 演算コア部については、2020年には7nmプロセスが、2022年には5nmプロセスが利用可能になるとされる。現在28nmで400 mm^2 の面積に1kコアの集積度だが、同じ面積に2020年には16kコア、2022年には32kコアの集積も可能に
- 10cmの厚さには1 μm のウェハであれば100k枚積層出来るので、10cmキューブ(1L)には25億コアを集積可能であるが、そもそも非ノイマン型であれば1,000億コアの集積も現実的
- 給電と抜熱については、別途、新しい手法が必要であるが、アイデアは既に幾つか検討中。実現出来ると信じて進むことで道は必ず拓け、新しい発想が次々に産まれることを確信

Keep Changing
The Game”