

「情報基盤全体からの視点： 徒然なる定年教官の視座」 given by 安浦先生

喜連川優
国立情報学研究所

喜連川のスパコンへの思い。結構頑張ってきました。

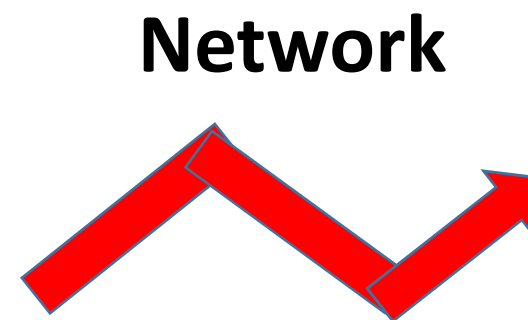
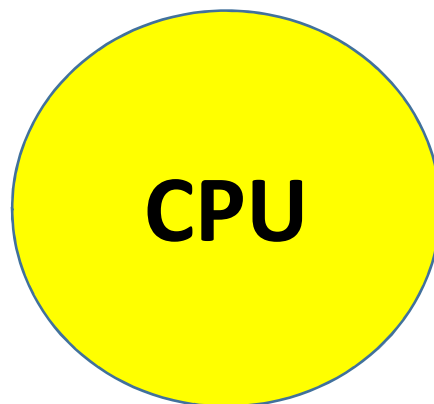
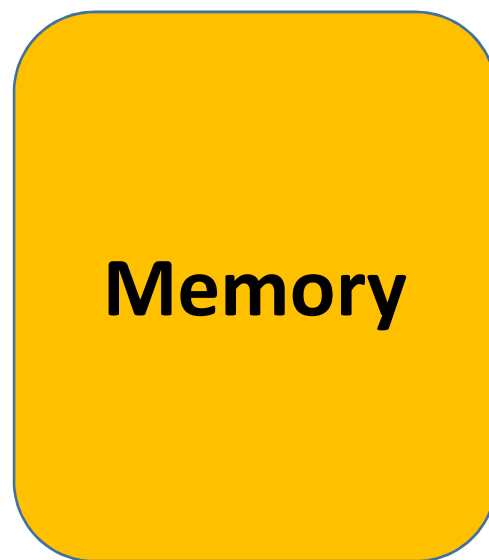
- 蓮舫発言「2番でいいのでは」に 反旗（情報処理学会副会長時代に蓮舫議員反対の学会声明を発出）
 - スパコンの産業は極めて小さく、理事のマジョリティはとても苦労しました。
- オープンを訴え続けて来ました
 - 京の開発当初は国家基幹技術で機密。出来てからしかISVはアーキテクチャを見れない！？ こんな常識はずれな状況はあり得ないと科学官として訴える
 - 富岳に関しては「可能な限りオープンにすべし！」を連呼

定年を迎える老兵の
次の世代への気持ちを纏めました

コロナ時代：
原点に戻って考えるチャン
スと捉えるのはどうで
しょうか？

大局的に見ると

ざっくり分けると3つの要素

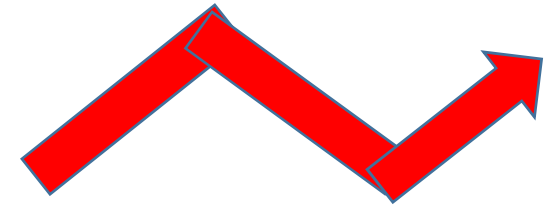


ざっくり分けると3つの要素

- スパコン 1976
- Internet 1984
- Data(4th paradigm) 2009

データ

ネットワーク



スパコン

距離を置いた見方

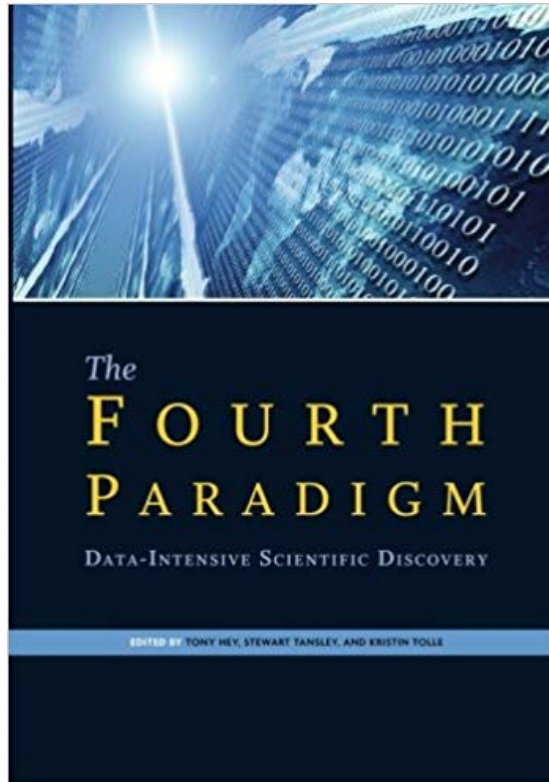
- スパコン競争：最も歴史がある分野（小生が学生の頃CRAYは存在した）
- Computational science、計算が世界を変えるという考えに昔から若い研究者は魅了され続けて来た。（良いことであった）
- ネットワークはTCP/IPのUNIXバンドリングが起爆剤（1984）
- データはProgram/Dataという意味では古くからあるが、データ駆動科学は**最近の潮流 2009**
 - **NIIは次期SINETでデータ基盤を構築予定。マスタプラン、ロードマップでご承認**

データ駆動サイエンス

2009年10月1日

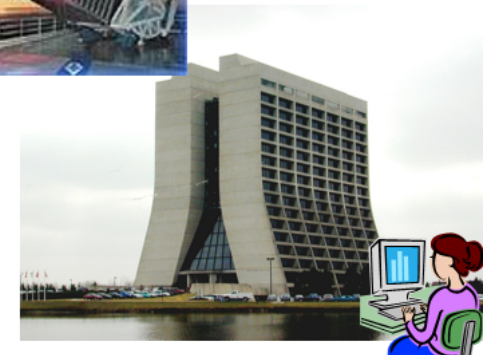
計算からデータへのシフト

支配方程式の無い世界への挑戦



The Evolution of Science

- **Observational Science**
 - Scientist gathers data by direct observation
 - Scientist analyzes data
- **Analytical Science**
 - Scientist builds analytical model
 - Makes predictions.
- **Computational Science**
 - Simulate analytical model
 - Validate model and makes predictions
- **Data Exploration Science**
 - Data captured by instruments
 - Or data generated by simulator
 - Processed by software
 - Placed in a database / files
 - Scientist analyzes database / files



第3の科学から第4の科学へ

スパコンの時代から データの時代へ

支配方程式がある世界から未だ方程式
を出しえない難しい科学の世界へ

（実は後者が大半）

3つのコンポーネントは、データの時代にどう対処すべきでしょうか？

- ネットワークは 均質にならざるを得ない特性
 - 一旦作ると変えられない(TCP/IP)
 - SINETは東阪のみ400G
- データは 現時点で注目株
 - 計算資源とは別のエンティティとして見据えるのが妥当
 - データの特性により非均質あり
 - これまでは計算の高速化だけに注目するだけで済んでいたが、日本の国家としての情報基盤を考える際には、**スパコンだけではなくデータ基盤を合わせてしっかり考えないとバランスが悪くなる**
- 計算資源（スパコン）詳細は後にするとして
 - **今のスパコンは大きくデータ駆動科学にシフトBDEC**
 - **計算のパワーをデータパワーに融合する時代**

IT全般を見た時、今日の所謂 グローバルプラットフォームの 勝つ戦略は見えてきましたよね。

- 基盤層はスケールを大きくすることで、コストを徹底的に下げることが重要(Google, FB, Amazon, ..)
- 昔に比べると ベンダーが減る。種類が減る。
- 同質性 ⇒ 低コスト化（製作コスト、維持コスト）
- 中国を除くとスパコンは殆ど同じ。昔はもっとバラエティがあって夢があった。
- 付加価値は上位層で出す
- 戦略：
 - 類似するものは集約する方が優位
 - 異質なもの（小林レポート）をそれぞれに先鋭化
 - しかし 色々な文化があること、リスクヘッジにも配慮

前頁を脳裏に置いて
「計算資源（スパコンの一般化）」を
考えますと？

視座 1 : そもそも計算資源は どれだけ要るのでしょうか？

- スパコンは要るか？ ⇒ 要るに決まっている！
絶対必要
- ポイントは、日本に「どの程度の計算資源」が必要か？の議論があまりない。
- 一つの方法は学術会議に問う？（ユーザが決めること）
マスタープラン リニアコライダも含め何にどの程度資金を投入するかの真摯な議論の場
- 計算に依存する学術は多いはず
- 今の計算資源量では足りないという結論だって十分にあり得るのではないのでしょうか。

視座2：どんな計算資源（スパコン） を作るべきか？ 小林委員会：ヘテロ

将来のHPCIのあり方に関する検討ワーキンググループ 報告書（概要）

R1.6 HPCI計画推進委員会
将来のHPCIの在り方に関する検討WG報告書

◆HPCIの現状と課題

- 文科省では、平成26年度からフラッグシップマシンとしてスーパーコンピュータ「富岳」の開発に着手。令和3～4年の共用開始を目指し、順調に進捗。
- 各大学の基盤センター等と連携し、HPCIも順調に運用されており、平成30年度末までの研究課題は1,342件（のべ11,685人）。
- 一方、計算科学分野や情報科学分野の動向として、①利用分野拡大およびユーザーニーズの多様化、②アーキテクチャの多様化（半導体微細化の限界等）、③データ量の指数増加、④情報分野における教育への期待の高まりが指摘されている。

◆将来のHPCIのあり方について

- 最先端の科学的成果を創出するためのHPCIを、引き続き整備・運用していくことは我が国の科学力、産業力等の維持・向上に必要不可欠。
- 将来のHPCIの姿を具体化していく上で、以下の4点に十分留意しながら検討すべき。

①アーキテクチャの多様性とプラットフォームの整備

- 将来的にはHPCIには異なるアーキテクチャを混在させ、多様なアプリケーションに適用させるべき。フラッグシップマシンの定義や役割の再検討、HPCIの全体構成の長期的、俯瞰的視点からの検討が必要。
- ユーザーの要望が多様化するため、多様なアプリケーション分野への対応、計算基盤利用プロセスの高効率化、高生産化を実現するための取組（ソフトウェア、人材育成、利用支援等）を行うことの重要性が増していく。

②データ活用基盤の構築

- HPCI内外から得られる大規模データ処理のインフラであるHPCIをそのニーズと整合性のとれたシステムとすべき。
- データが有効に活用されるための基盤構築を推進する必要。
- 低レイテンシで高バンド幅なネットワークの存在が重要であり、SINETの機能強化及びHPCI構成システムとの連携が必要。

③ボーダレス化の進展

- 世界最高性能のスパコンを1つの国の技術と予算で作り上げることは困難になりつつあり、解析されるデータも国境を越えて収集されることが想定されるため、HPCI分野における国際連携を推進すべき（国内外拠点が参画する計算・データ共有基盤構築、大規模データの共同解析プロジェクトの推進、HPCI分野の国内外機関間の共同研究の促進等）。
- 民間企業によるクラウド型サービスを研究者が利用することが常態化しつつあり、科学、産業を支える計算基盤として、国と民間それぞれの特徴を踏まえた上で、ユーザーからみて最適となるHPCIを実現すべき。

④人材育成

- アーキテクチャ、OS、コンパイラ、基本ライブラリ等のシステム系の技術を保持し続けるための人材育成の努力が不可欠。
- 高度な技能をもつ支援要員の育成に力をいれるとともに、キャリアパスを産学の関係者で検討すべき。
- ボーダレス化に対応し、国際連携を推進する観点からの人材育成も重要。

◆実現するための具体的な取組例

- 将来的にHPCIに導入される可能性がある斬新な未開拓技術およびそれらを用いた計算機構成方式、基盤ソフトウェア・アルゴリズムを研究する基礎研究プロジェクト
- 多様なアーキテクチャから構成されるHPCI全体の有効活用を実証するための研究開発プロジェクト（ソフトウェア等の利用環境整備含む）
- 海外の研究拠点との連携を推進するプロジェクト（国外拠点も参画する計算・データ共有基盤の構築、大規模データの国際共同解析プロジェクト等）
- 潜在的ユーザーがHPCIを活用するための支援するための体制強化、利用支援人材のキャリアパス構築をHPCI参画基盤が連携して実施するプロジェクト
- 科学的課題・社会的課題を見据えつつ、その解決に必要な計算科学・データ科学的アプローチと、その実現に必要なネットワークも含めたHPCI全体像を示すロードマップの策定
- 技術研究開発プロジェクトからHPCIの利用に関するプロジェクトに跨ったワークショップの開催

ヘテロをバランス良く！

第6期科学技術基本計画に向けた今後のHPCIの方向性に関する検討 論点まとめ

出典:R1.10.18第4回情報委員会資料2

■各種応用分野に強みを発揮する専用機的发展

- ・ 量子コンピュータが最適化問題や暗号計算などに強みを発揮する可能性が指摘されているように、それぞれの特定分野に強みを発揮する専用の計算資源がそれぞれの分野のHPCとして役割分担するような計算手法が増えてくることが想定される。実際には、専用のハードウェアアーキテクチャを用いる場合もあれば、汎用性の高いハードウェア上に仮想化技術で実装される場合も想定する必要がある。

■多様なシステムからなるバランスの取れたHPCI

- ・ 国の機関及び民間によって今後提供される先端的な計算資源の動向も踏まえつつ、我が国の科学技術・学術研究の発展に不可欠な計算インフラの全体像を俯瞰し、多様なシステムから構成されるバランスの取れたHPCIを検討していく必要がある。複数の先端的計算資源や大規模データベースを効率よく接続する高速かつ高信頼なネットワークとその上のデータ流通まで含めた国レベルの統合的なシステム構成と、技術の進歩を意識した継続的なシステムの更新や保守の体制を考えるべきである。

■ソフトウェア技術と人材の重要性

- ・ 応用分野別にコ・デザインを高度に適用・開発することで専用化された計算資源で構成されるHPCIをユーザーのニーズに応じて効果的・効率的に組み合わせて使いこなすための基盤的なソフトウェア技術の重要性が増してくる。利用者の立場から、利用できる計算資源を柔軟に組み合わせて利用することを可能とするコンパイラや仮想化技術を含む「基盤ゲートウェイソフトウェア」とでもいうような技術開発を推進していく必要がある。また、このような汎用的な基盤ソフトウェアの研究開発・普及を支える人材を長期的視点から継続的に育成、支援していくことが重要である。

視座3. どういう風に創るのが良いのでしょうか？ (調達が肝、コンソリの時代)

- 日本にはFHNと3つも必要？という議論を大学の先生はします
- しかし、大学も一杯あって同じ。そもそもスパコンをバラバラに調達しているのも同じかもしれないのではないのでしょうか？
- 色々な機関が色々なスパコンを所有している。
 - しかし、スパコンベンダは僅か。
 - **スパコンの種類（ベンダ数） << 調達機関**
- **調達力UPを狙うべきかな？
(ジャーナル問題と同じ)**
- そうすれば、かなり節約出来るはずです！
- **東大＋筑波大で既に実績があります。得るもの > 違和感**
- **MDX：調達した後に等分割する方式！実際やれば出来る！**

コンソリスパコンの時代

- どのみち、ほって置いたら、AWS HPCに負けてしまうのではないのでしょうか？ **これをとて**
も老兵としては心配しております。
- その前に出来る限り体力をつけることが必須。
- 「**自国に計算パワーを確保**」することが大切であることがこのコロナで判りました。

何の為に？調達力をあげ、調達の省力化を敢えてしないと思うかと言いますと：

DXの先導

- 基盤センター数は減らない（地域のサポート）
- 各基盤センタの予算も人員も減らない
- 物理マシン数は減り、クラウド型へシフト
 - 管理コスト減少
 - 調達の手間大幅減少
 - （可視化など一部はローカルに残る）
- **生まれた余剰パワーで**
- **教育のDX， 研究のDX、事務のDXを推進（現在、余力が無い！）**

デジタル化はそんなに必要ない
ということでありましたら、こ
の論述は意味を失います。

もしデジタル化が妥当と考えますと、
アホな喜連川としては、これが一つの
方向かと。

もっと他の道があるかもしれませんが、
それを否定するものではありません。

バラバラに同じマシンが沢山ある！？

国立研究開発法人 宇宙航空研究開発機構	宇宙航空研究開発機構	宇宙航空研究開発機構			
国立研究開発法人 宇宙航空研究開発機構	宇宙航空研究開発機構	宇宙航空研究開発機構			
国立研究開発法人 日本原子力研究開発機構	日本原子力研究開発機構	日本原子力研究開発機構			
国立研究開発法人 日本原子力研究開発機構	日本原子力研究開発機構	日本原子力研究開発機構			
国立研究開発法人 量子科学技術研究開発機構	量子科学技術研究開発機構	量子科学技術研究開発機構			
気象庁(NAPS10)	気象庁	気象庁			
気象庁気象研究所	気象庁	気象庁			

長期的展望において、
現状のようにスパコンをバラ
バラと多機関が維持するやり
方で本当に良いのか？

ちょっと違うけれどもたいして変わら
ないマシンが沢山

バーチャルクラウドの方が妥当では？
運用面ではその方が圧倒的に楽なはず

視座4 コストをどう考えるか？ (京、富岳に特化した視点)

- 1200億は実際 高価と言えるでしょう。
- 理研AIP (AIのプログラム) が40年分作れる。
CRESTならもっともっと沢山作れます
- 情報処理学会の研究会は約40個
- CSは弱いと言われる。10個程度のCS分野プロジェクトを作れば、もっとCSが強くなれるのでは。現状は富岳を情報のプロジェクトと見なされ、集中しすぎと言わざるを得ないのでは無いでしょうか。他の情報プロジェクトが立ちにくいことを憂慮します。

視座5 「元気の素」 と思えば京、富岳は高くは無いと思います 情報系のシンボルでしょう。

- ノーベル賞は我が国はそこそこあるものの、ITで「元気の素（世界で一番）」が日本にはあまりないのが実情です。
- 国民の栄養ドリンクと思えば 十分ペイするカンフル剤と言えましょう。
- ココアもバグ2回、東証もダウン、キャッシュレスの不具合 ⇒ 元気の出るIT素材が希薄
- しかし、それでも情報枠ではない点をはっきりと合意をとるべきだと思います。
- 我国ではHPC以外の情報分野に殆ど予算が回らないのはとても残念です。（IT研究全体を持ち上げることが必須）

視座 6 元気の素となったとしても、きちんとした整理は不可欠だと思います

- ・「京」でなければ得られなかった成果は何かという整理は今後を考えると必須でしょうか？
- ・基盤センターの複合体では無理であることを明確化することは重要ではないでしょうか。

視座7「よく考えるとやはりおかしい！
1200億円は情報分野のものでは無い」
スパコンは情報分野のモノと思われな
いようにすることが大切です。

- 情報分野がお世話しているが、計算資源は多様な科学者のもの（これを何度も言い続けることが重要）
- ドメイン学問から資金が拠出される構造が本来は妥当と
感じます。
- フォーマルに議論しているのが「学術会議マスタープラン」と言えます。全分野からエンドースを得ることが王道かもしれません。（6年）
- 情報分野が受益者のように見えるのが不幸の始まりです

視点8 情報分野としての価値もあるはずです。（富岳を活用出来ているのでしょうか？）

- もちろん価値は大きいはず（8百万個のプロセッサコアは大魅力）
- 今はC（大きなユーザ空間）を持ったものが勝つのであり、スパコンには多分沢山のユーザ付いていると想定されます。
- LISA的研究が出来ます。
 - Large Installation System Administration Conference
- しかし日本からそのような視点でのシステム研究は何故か希薄なのは残念です（もっとやってもいいのではないかと感じます **ROIが悪いので研究者が少ないのかもしれません**）
- 情報分野の価値に関する議論が「あいまい」と感じます
- **第6期 ⇐ 健全な「評価システム」が重要だと思います。**
- 情報分野システム研究としての拠出額をどう考えるのか？
- 情報分野としては何が欲しいのか？（いじれるシステム）

視座9 必須の工夫：次世代育成

- 学生や若手研究者がなるべく身近に使えるようにすることが一番大切でしょう。。
- 次の世代の育成は極めて重要
- MITリンカーンラボでのサザランドを思い出しましょう。（無茶苦茶な実験を可能にすることで無茶苦茶な発想が生まれます）京都賞受賞者

視座10　そもそもスパコンは文科省だけでやるのでしょうか？

- 不老の富士通の発表は判り易くて、とても面白かったです！
 - ARMのHPC実装はスクラッチからだったとのこと
- 富岳は日本の技術力の維持に貢献したと言って間違いないと感じました。
- でも、それは経済省のmatterかもしれない？（商情局）
- バイオ戦略　＞　IT（？）

雜感

要するに

- 小林委員会「1つのデカいスパコンは、やれば作れる」
- 一手先、二手先を見ると、既にゲームは変わっており、小林委員会は「多様なスパコン」の時代を示唆。ACM 50周年のパネルと同じだと思います。
- この難しい課題を考える事が全てではないか

何を「勝ち」と考えるか？

- 15年先を見据えたとき、従来のゲームを続けてレガシーの勝ちを狙い続けるのか？という問いがあります。日本は「ゲームの決まった闘い」をしている余裕があるのでしょうか？
- サムソンはなんと世界第二位のバイオ産業シェアです。
- 当面の金メダルではなく、その先のプラチナメダルを目指すべきかもしれません（多くの日本の産業が凋落しているeg.半導体）

何故日本のCSが強くないか？ 海外の日本人から見た日本の違和感

- 古井レポート
 - 要するにそもそも戦い方が中途半端
 - グローバルタレントの厳しい争奪
 - もっと集中、もっと厳しい評価機構
 - 大学の研究室単位でITインフラをメンテすることは一切無い。無駄が多すぎ
- 根底から考える必要があるでしょう。「**本来、スパコンの議論はまずはユーザがやるべきかと感じます**」もっとCSの抜本的グレードアップに関する大切な議論が不可欠ではないかと思います。

ご清聴ありがとうございました

- 大学の定年を迎えるものの役割として、大局から考えてみました。
- 「データなる潮流」・「DXの流れ」を上手に乗り越えて行くには、追加定員が無い中では、相当のパワーを捻出しなくてはならず、違和感を伴うかもしれないけれど、現在ある仕事を根底から見直すことも必要ではないかと感じなくもありません。
- 間違いも多々あるかと思います。ご指摘お願い致します
- 気持ちの整理には時間がかかると想定されます。
- これまで培った気持ちや文化も重要です。
- なお、頭が悪いので、この資料は何度も書き換えましたがまだ不明瞭やもしれず、失礼致します。