



powered by
FUGAKU

資料3

2システム一体整備／運用の可能性について



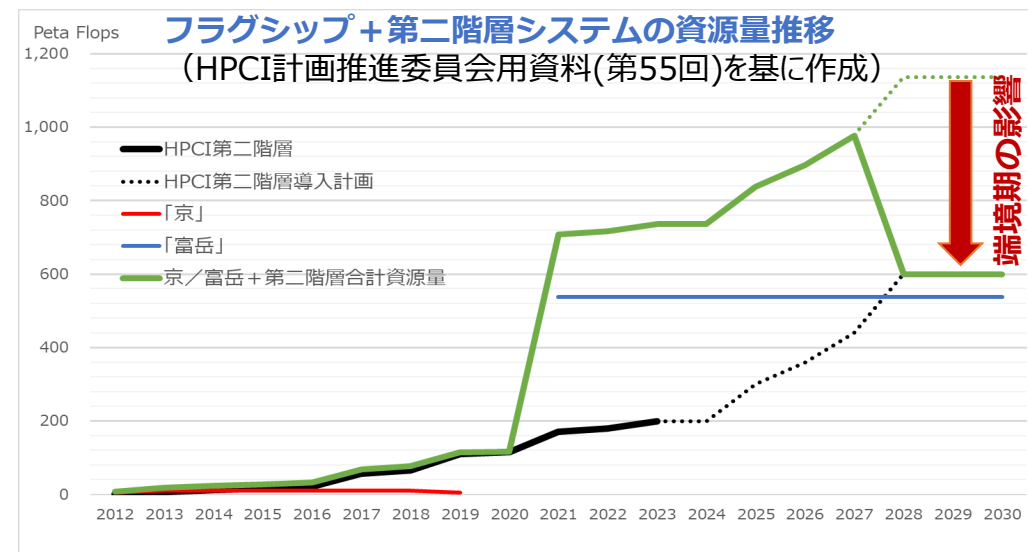
理化学研究所 計算科学研究センター (R-CCS)
センター長 松岡 聡

2024年3月19日



- 従来型の10年近いサイクルでの単一システム運用の課題

- 最新の計算手法等に適合したシステムを提供できない。
(例：深層学習手法や生成AIの高度な利用など)
- 運用サイクル後半では最新機種に対する性能的な魅力が薄れる他、電力効率も悪い。
- システム移行期に停止期間が生じ、計算リソース提供に端境期が生じる。



- 2 システム一体整備・運用のメリット

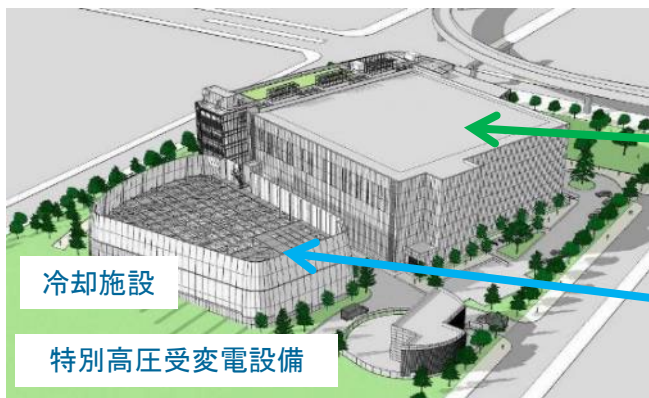
- 常に世界最高水準の計算性能と計算資源量を安定提供し、計算科学の研究開発サイクルを途切れなく回すことで成果創出を加速できる。
- 最新テクノロジーの採用・キャッチアップができる。
- リプレース時の空白期間をなくし継続的に計算資源を提供できるほか、ユーザーデータもシームレスに移行できる。
- 近接施設での一体整備・運用により、既存の運用体制・施設と共通化できる部分が多く効率化できるほか、最新設備へのアップグレードも容易となりエネルギー効率が高いデータセンタ運用が可能になる。

新システム導入に必要な設備環境

理化学研究所の国内拠点



R-CCS研究センター + データセンター



計算機室 50 m x 60 m = 3,000 m²
電力消費 最大 37 MW
 水冷システム

ガスタービンコージェネレーション設備
 5 MW x 2

- 老朽化に伴い既存の「富岳」の電気・冷却設備を継続して新システムにも利用するためには建屋やおよび設備の大規模な改修が必要なため、**長期の停止期間**が見込まれる。
 - 複数のシステムを同時に運用するためには、**既存の建屋および電気設備・冷却設備の容量が不足**するため増強する。
 - **カーボンニュートラル化**を実現するためにも、既存設備の抜本的改修を行う。
 - 新システム用の新たな建屋および電気・冷却設備の整備する。
 - **新たな建屋および電気・冷却設備を整備する場合、建屋整備に時間を要する。**
 - 一般的な公共工事の場合、調達行為を含め、計画から竣工まで**通常5年以上**が必要であり、さらなる長期化も予想されることから、DBO方式*等、民間への委託方式を活用することなど検討し、計画から竣工までの期間を短縮と事業継続性の向上を両立させる。
- *DBO方式とは、PFIに類似した事業方式の一つで、公共が資金調達を負担し、設計・建設、運営を民間に委託方式のこと。民間の提供するサービスに応じて公共が料金を支払う。民間が資金調達を行うのに比べ、資金調達コストが低いため、コスト縮減率≒VFMで有利になりやすいことが特徴（出展一般社団法人 日本水道運営管理協会）。
- 民間の一般データセンター運用技術を活用することで、**施設建設・運用コストの削減**するとともに、AIなどのインフラ設備運用に必須となる**HPCデータセンター運用技術を民間事業者**に**技術移転**することで事業の継続性を向上させる。
- **そのほかの課題**
 - 運用コストの大部分を占める**エネルギー(電力)コストの削減**がさら重要になるため、電力調達方法の柔軟化、および発生する莫大な**排熱の有効活用も民間との協業を検討**。
 - 隣接地を含めた土地利用の可能性の調査・検討（新たに特別高圧受変電設備を必要であり、最低**3年強**の期間が必要）。また、2つの建屋・電気・冷却設備運用のための**人員増強の必要あり**。

課題は克服可能

● 建屋の建設・運用コストの削減

- 低層階かつ簡易的な造りとし、過剰な建設コストを削減。
- 計算機の電力密度を向上させ、必要床面積を削減し建設コストを削減。
- 既存建屋と比較して低床のフリーアクセスフロアの仕様とし、計算機本体の重量制約を低減。
- 計算機のライフサイクルを踏まえ、新システムに特化せず、普遍的に利用可能に。
- 計算機と冷却設備を近傍に配置にし、設備の設計を簡略化。

● 冷却設備の整備・運用コストの削減

- 冷却塔を主体としたフリークーリング技術を導入することで、圧縮機主体の冷却設備から脱却し、冷却設備の導入・運用コストを削減。
- 温水冷却を主とする冷却手法とすることで、PUE=1.1以下（現在の「富岳」がPUE=1.3程度）を実現し、電力効率を「富岳」比2割向上。

「富岳」を含めた2システム一体運用

計算科学研究センター（R-CCS）運用技術部門によりスーパーコンピュータ「富岳」は維持・管理されている。運用技術部門のミッションとしては、「富岳」等のシステム稼働に必要な電気設備や機械設備、ネットワーク設備の管理・運用を行うとともに、システムおよび設備の効率的な運用に資する技術開発、およびシステム利用技術の高度化に資する研究開発、R-CCSで開発されたソフトウェアの普及や、Armアーキテクチャにおけるソフトウェアエコシステムの構築に資する取り組みを行っている。これらを推進するため、必要な専門技能を有したプロフェッショナルを集め、組織的に対応している。



庄司 文由 部門長



井口 裕次 副部門長



施設運転技術ユニット
三浦 信一 ユニットリーダー



システム運転技術ユニット
井口 裕次 ユニットリーダー



ソフトウェア開発技術ユニット
村井 均 ユニットリーダー



データ連携技術ユニット
甲斐 俊彦 ユニットリーダー



先端運用技術ユニット
山本 啓二 ユニットリーダー

2024年3月現在

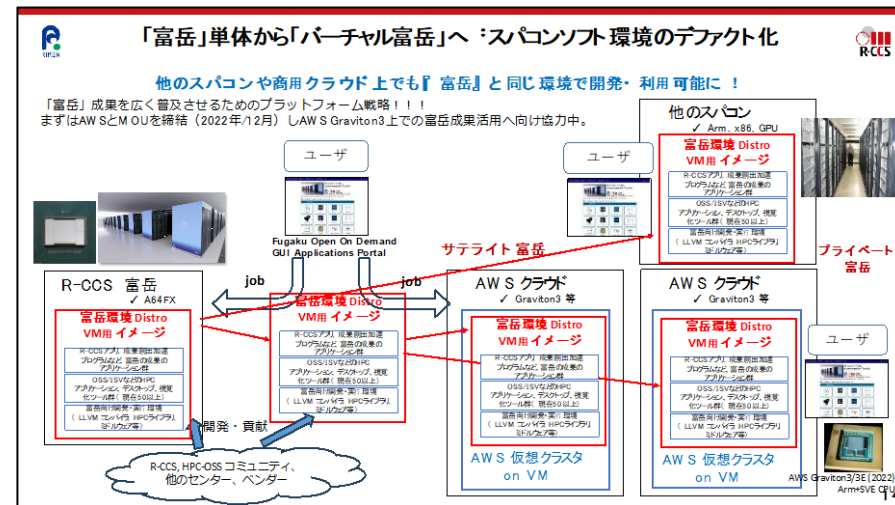
2 システム一体運用における課題と解決案

- **消費電力および冷却設備運転の調整**
 - ・ スーパーコンピュータ「京」、「富岳」における省電力運用の知見を活用しつつ、ジョブ履歴データの分析に基づく消費電力と熱負荷の予測を考慮したジョブスケジューリングにより対応可能。
- **システム間のソフトウェア環境の互換性の担保**
 - ・ バーチャル富岳の成果である、コンテナ等の仮想化技術およびSpack等のソフトウェアパッケージツールを活用することで、ソフトウェア利用環境の共通化が可能。
- **2つのシステムを透過的に利用できる環境の整備**
 - ・ 「富岳」運用の成果である、OpenOn Demandやワークフローツール等を連携させることで、異なるシステムを透過的に利用することが可能。
- **商用クラウドおよびHPCIシステム群等との連携**
 - ・ 上記の知見およびHPCI共用ストレージ運用の知見等も活用しつつ、商用クラウドサービス、HPCIを構成する計算資源および大型実験設備等との連携を促進。



課題は克服可能

バーチャル富岳の成果例



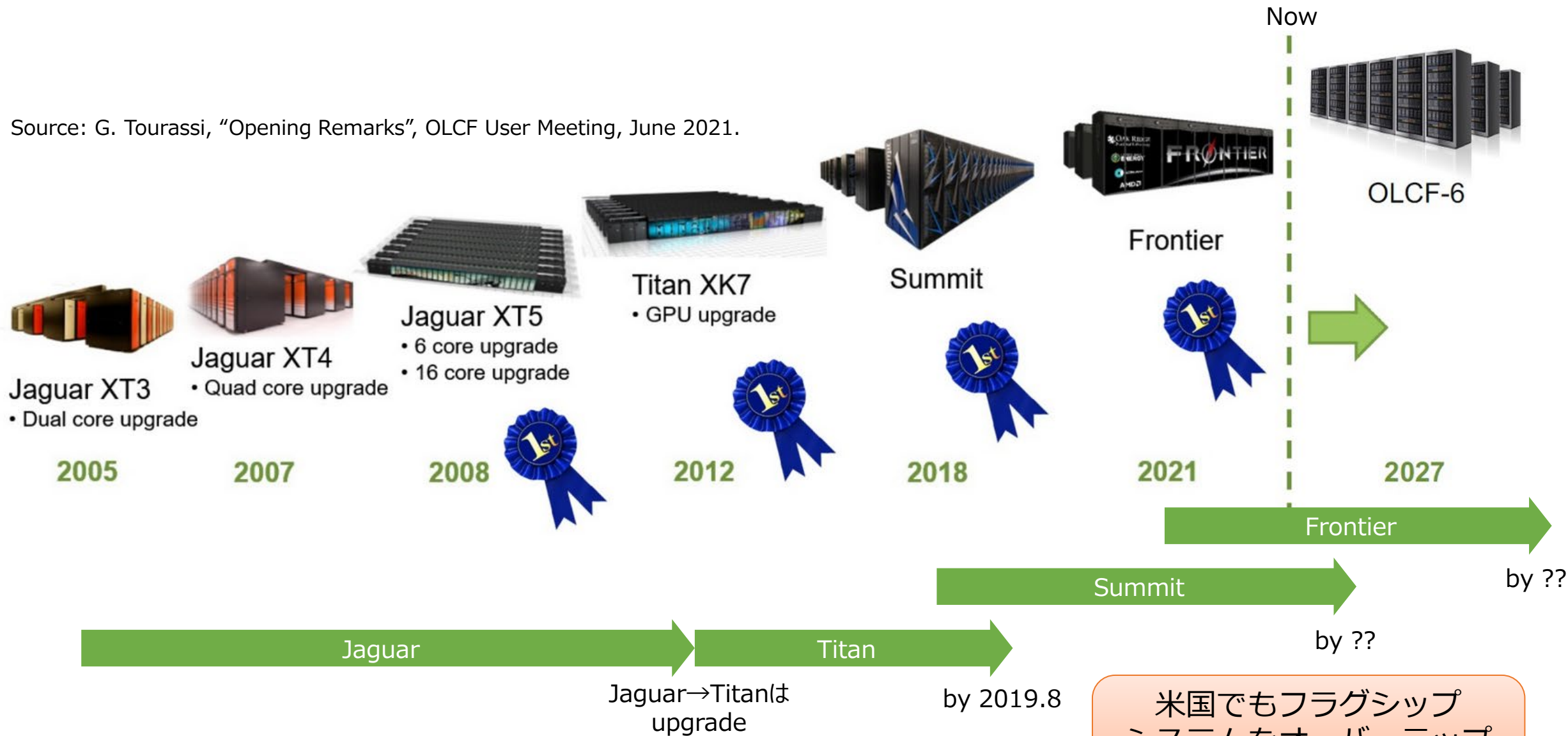
まとめ

- フラグシップ・スーパーコンピュータを空白期間なく整備・運用する意義は際めて高い。
- スーパーコンピュータ「富岳」を運用しつつ、次世代計算基盤を整備・運用（2システム整備／運用）することは可能。
- 「富岳」の近接地に整備することで、改修・増強は必要なものの、「富岳」の高度化・利用拡大で構築された資産を有効活用できる。

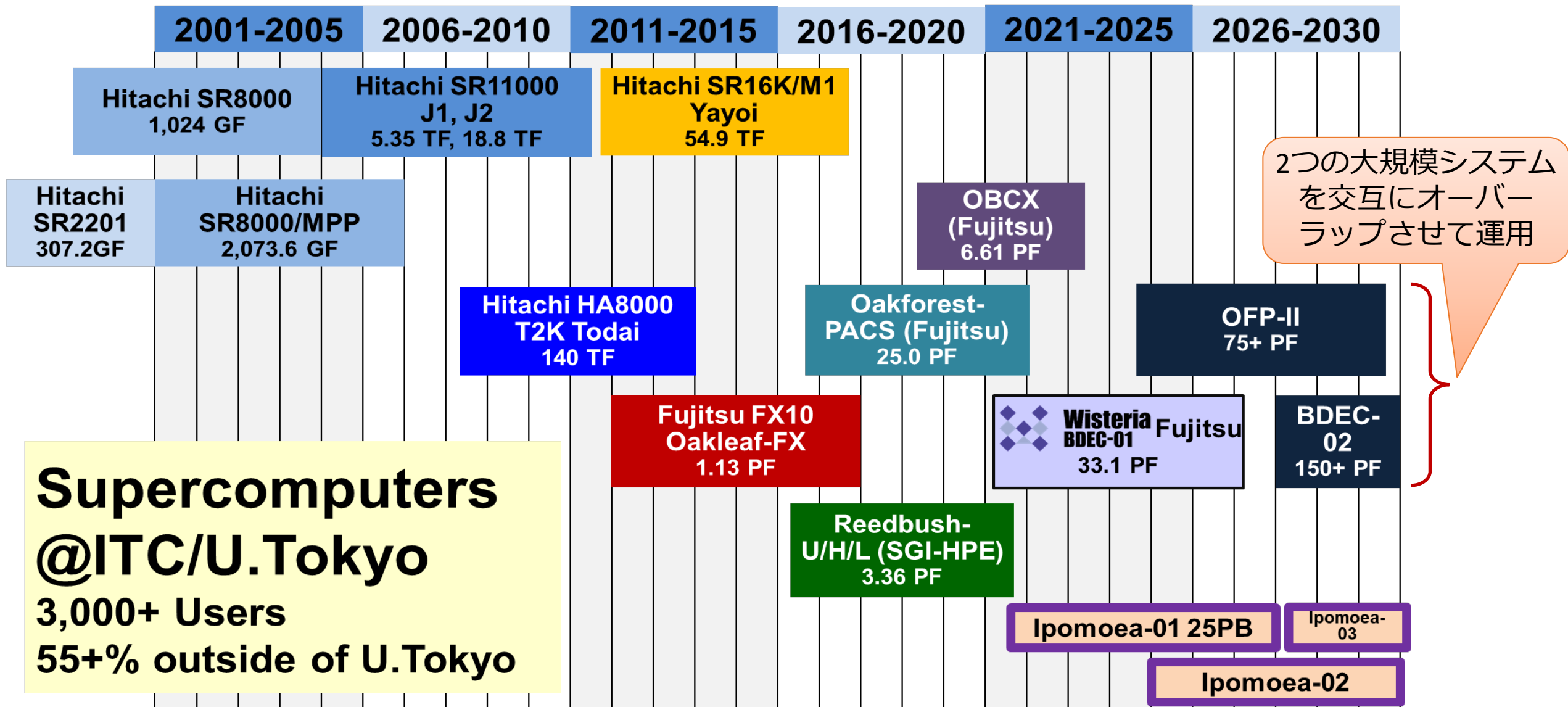
参考資料

米国ORNLのフラグシップシステムの運用状況

Source: G. Tourassi, "Opening Remarks", OLCF User Meeting, June 2021.



米国でもフラグシップシステムをオーバーラップさせて運用することが通常

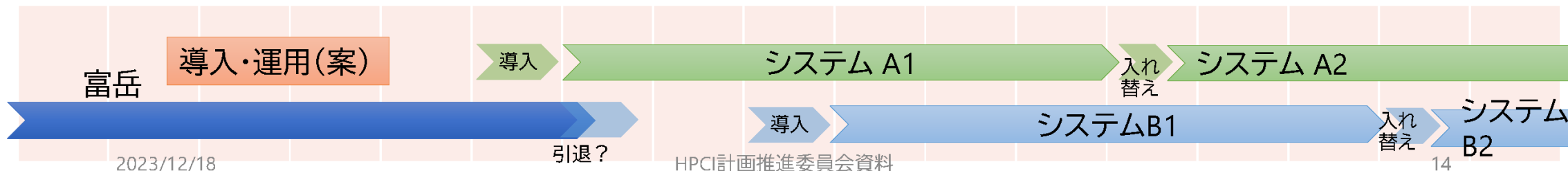


Supercomputers @ITC/U.Tokyo
 3,000+ Users
 55+% outside of U.Tokyo

次期(+その先の)フラグシップシステム整備に向けて


- これまでのフラグシップシステム開発の問題
 - 計画立案からシステムの運用終了までのサイクルが10年程度(京の場合)
 - 継続的な計画がなかったため、次期システムが決まるまで運用期間の方針も立たない
 - 富岳もいつまで運用するか現時点での見通しが無い
 - システム移行のタイミング: いわゆる**端境期** → ユーザの研究計画に多大な影響
- 理想的な形(**提案**)
 - 更新時期をずらし一体的な並行運用について検討
 - **研究開発の継続性も維持、技術動向の変化に即時対応できる**
 - 計算機システムの寿命は導入からせいぜい**6年**: コンポーネントの保証(End-Of-Life)の観点、これまでの運用経験から見て妥当な線
 - 製品のライフサイクルは、技術革新のサイクルの影響で延びてきているものの、6年目以降は厳しい
 - 7年目から保守費が急激に増加(IT機器は5年程度が多い)、機材によってはリプレースしないと継続運用できない
 - 利用者にとっては陳腐化し魅力が低下
 - 引き続き実現可能性を検討中

端境期をどう解消するのか?



次期(+その先の)フラグシップシステム整備:施設

一方で

- 次期システム向けに既存の富岳の設備の大規模な改修が必要
(理研R-CCS施設運転技術ユニットによる)
 - 冷却システムの主となっているコジェネレーションシステム(CGS)設備の寿命
 - 冷却配管の摩耗: 全交換が必要、並行配管には経路が不足
 - これらのリプレイス等には富岳停止後2年ほどの工事期間が必要
 - 富岳と新システムを並行設置するには、現行の計算機棟では、追加設置場所・電力設備容量の点で不足
- 
- 次期システムについて、計算ノードの検討だけでなく、建屋等の施設や冷却等の設備の拡大や増強についても、早急に検討を進めることが必要