

2024/07/12 新スーパーコンピュータXD2000の運用開始に向けた説明会

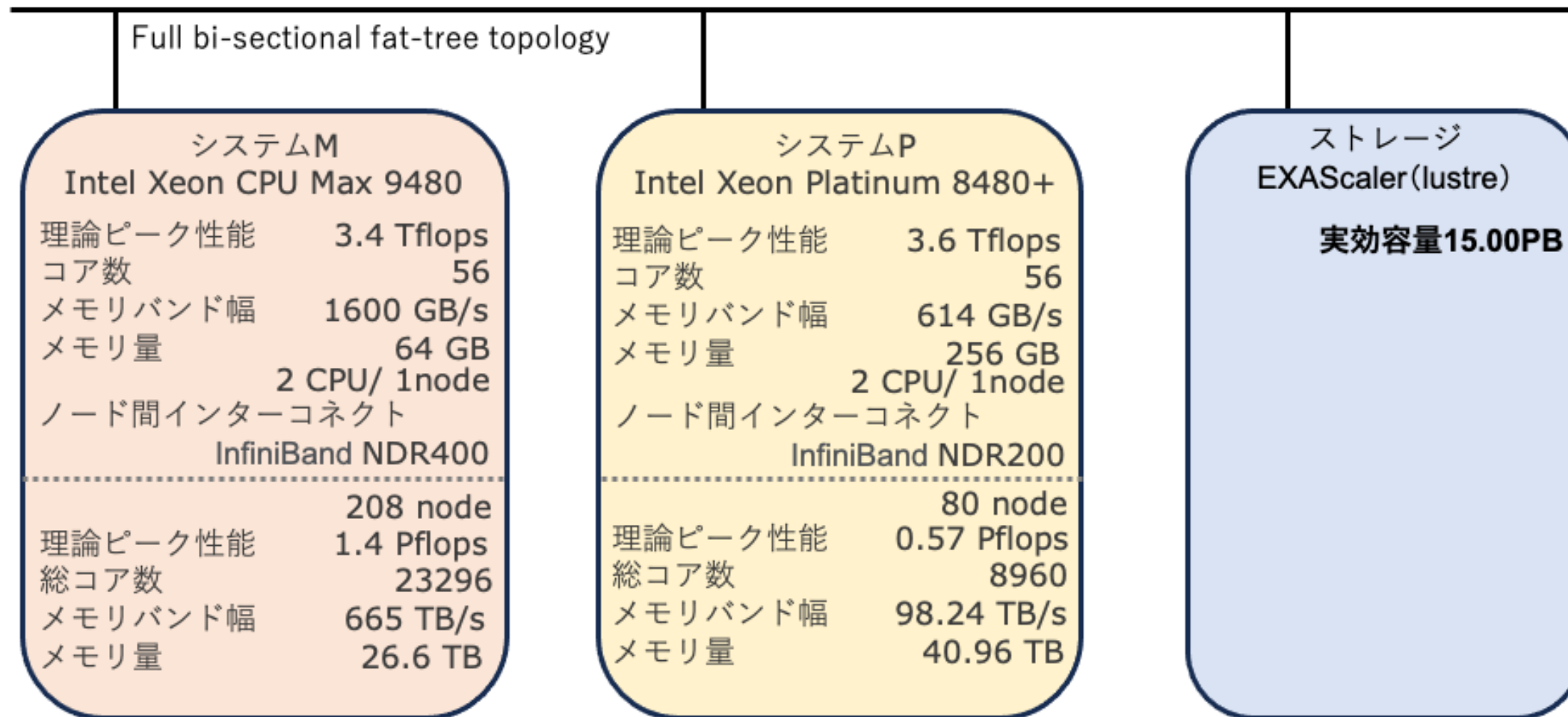
新スーパーコンピュータXD2000への 申請における注意点

Center for Computational Astrophysics

目次

1. XD2000の特徴
2. XDのカテゴリの変更について
3. 申請における注意点
4. XD2000以外の機材について

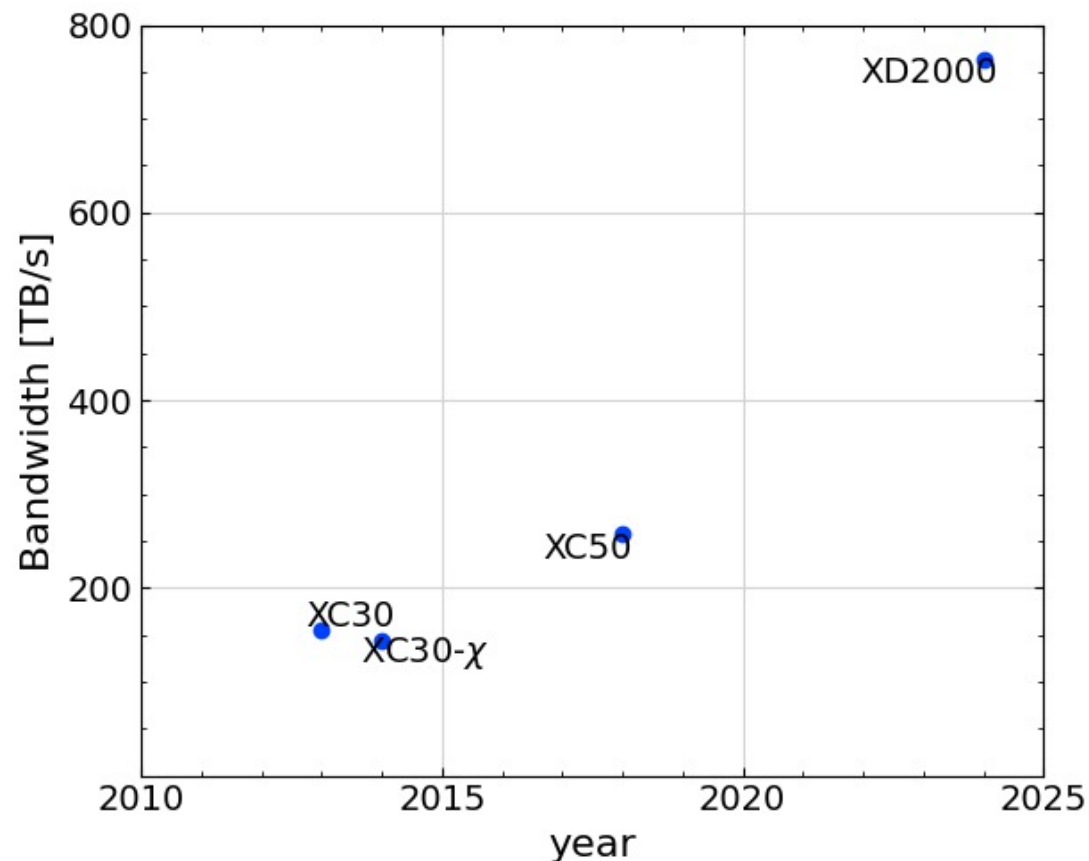
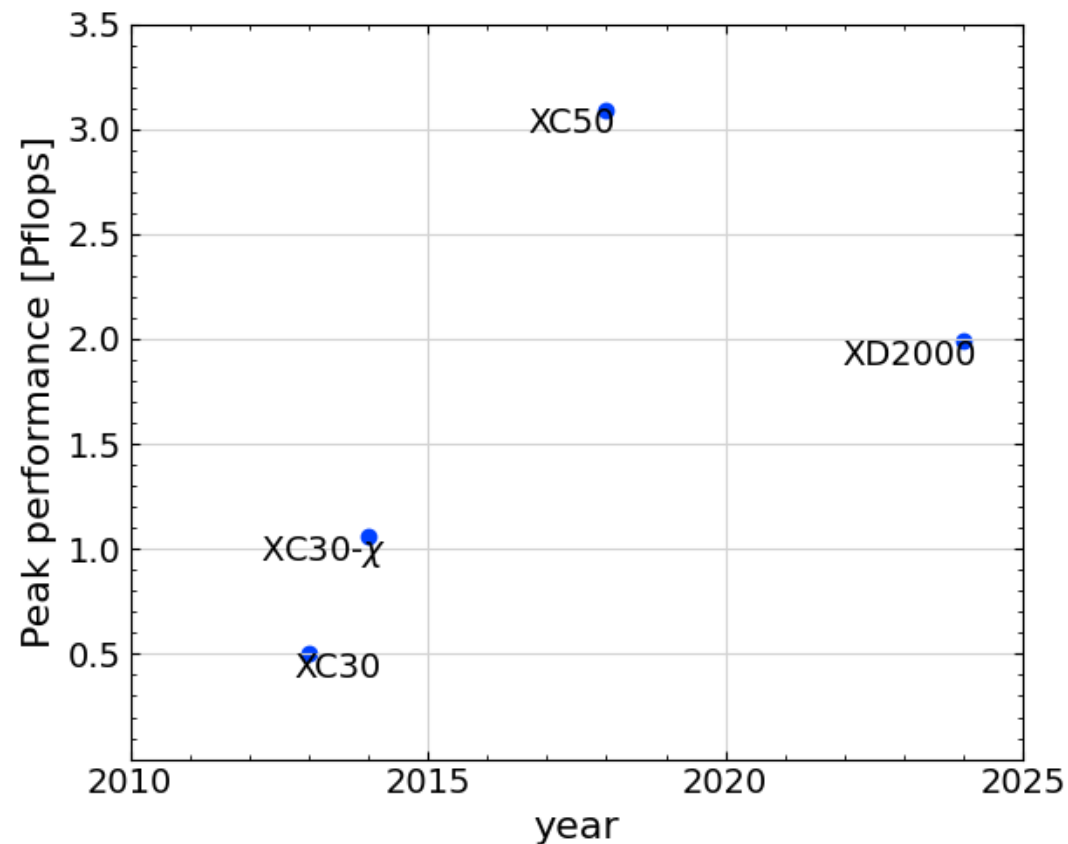
HPE Cray XD2000構成



メモリバンド幅特化

メモリ量確保

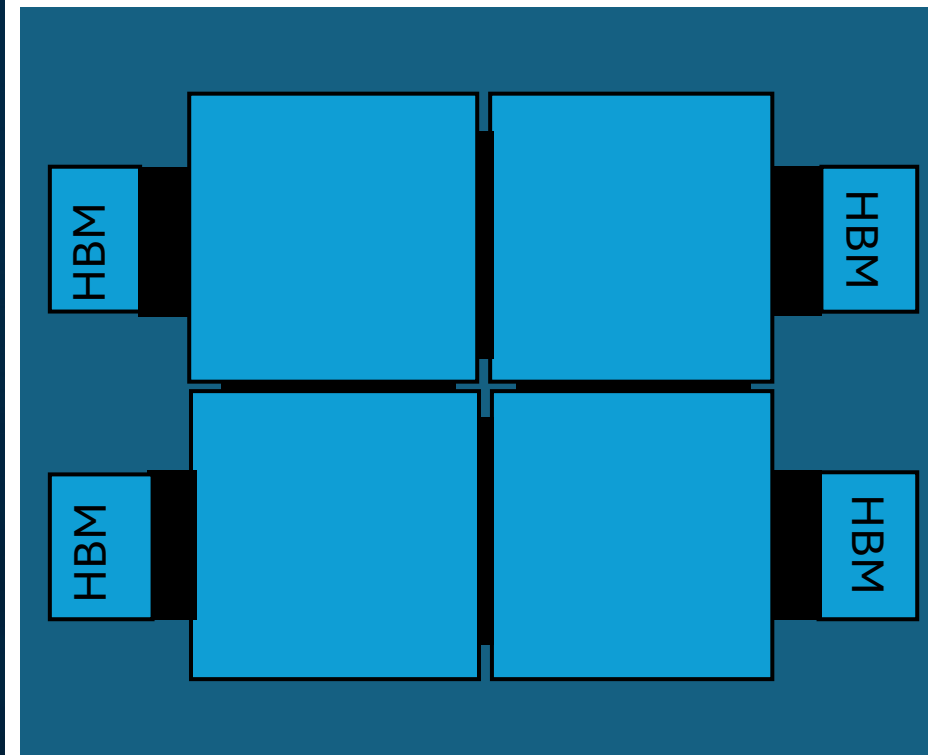
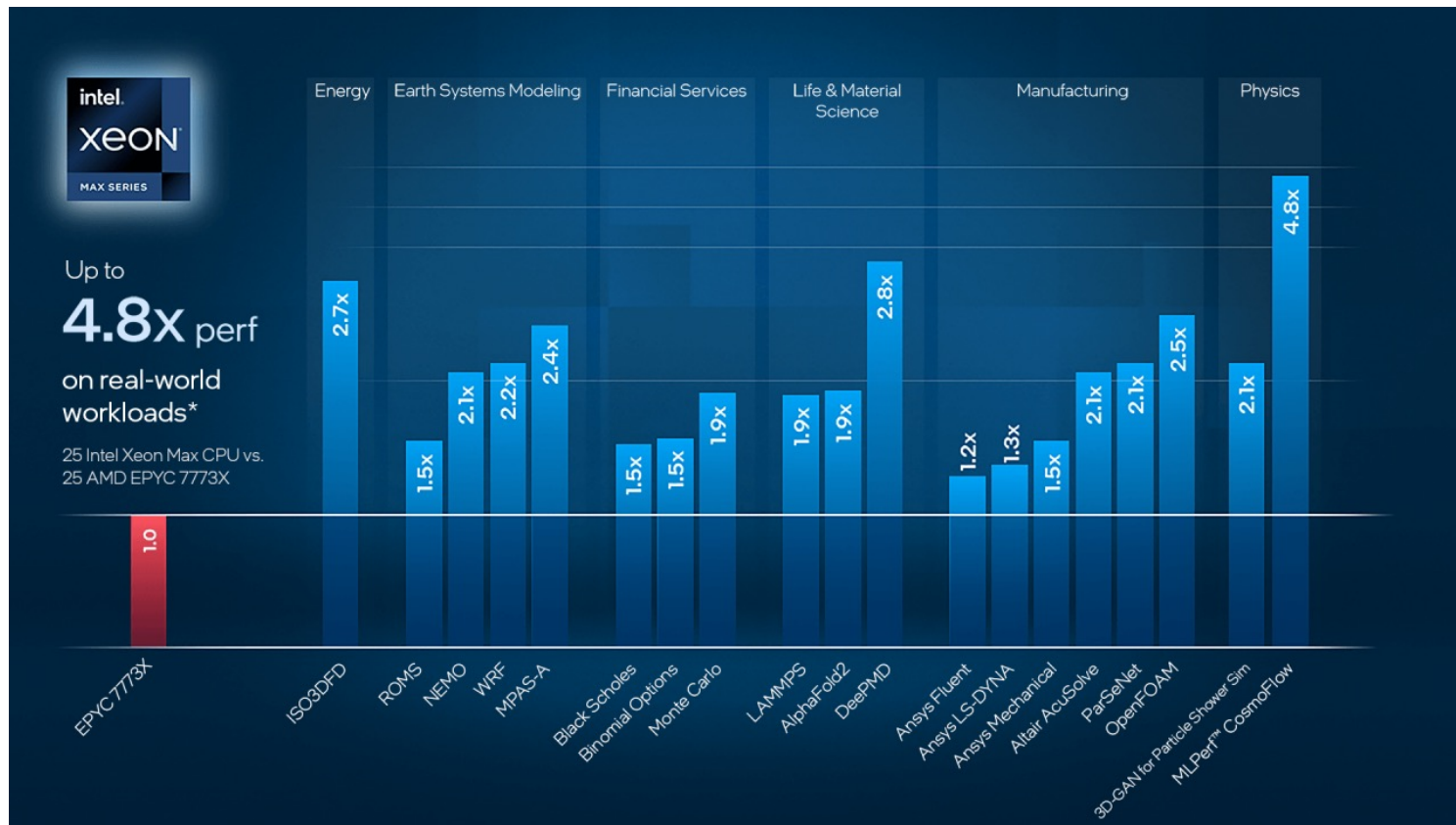
CfCAスパコンの性能の変化



今回は理論ピーク性能は下がり、メモリバンド幅は上がる

現在、多くのアプリで計算効率が悪く、ピーク性能は良い指標ではなくなっている。実際はメモリバンド幅が重要であり、多くのアプリで計算速度は上がることを期待している。

Maxの速度向上の期待



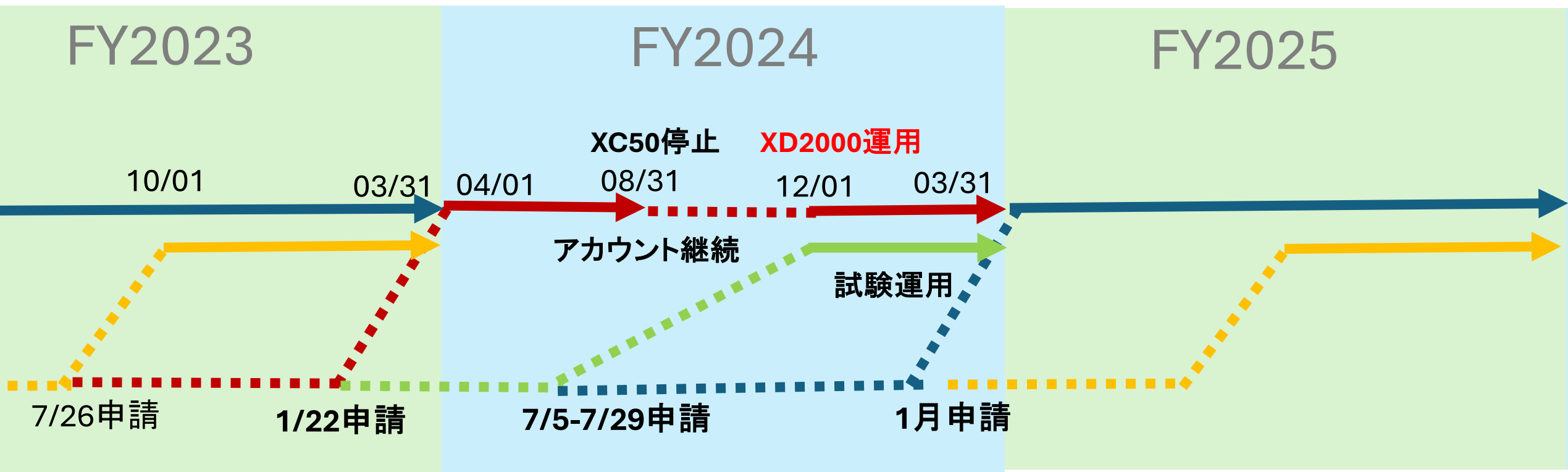
EPYC 7773Xに対して1.5-4倍程度のスピードアップ(intel 資料)

2次元の流体計算でSkylakeに対して9倍程度のスピードアップ(CfCA)

Maxは4つのチップを一つにまとめたようなデザインでそれぞれにHBMがついている。

1CPUに対して4以上MPI processを使うことで性能が最大限発揮される(1ノード8process)。

スケジュール



2024/1/22の申請でFY2024中はアカウント継続中。→

同じカテゴリで(XC-A, XC-B等)12/01運用開始のXD2000が利用できる。

7/5-7/29 募集中の申請は新規とカテゴリを変更したい方向け。→

12/01-03/31はキュー構成など頻繁に変わる試験運用。

FY2025の募集は通常通り1月に予定。→

カテゴリの変更

Sへ申請

XC-S

Aへ申請

XC-A

XC-B+

Bへ申請

XC-B

XC-MD

随時申請

XC-trial

随時申請

XD-S

Aへ申請

XD-A

XD-B+

Bへ申請

XD-B*

XD-B

XD-MD

随時申請

XD-trial

取扱変更

新設

XD-B*

読み方はエックスディービースター

これまで520コア、13ノード(XC-B最大)に相当するメモリ~5TBで計算していた方向け。

次期システムではXD-BではシステムPの6ノードが最大利用でき、使用できるメモリは~3TBとなる。このメモリでは研究の継続が難しいという方向けに、システムPを6ノード使用できるXD-B*を新設した。その場合メモリを6TB利用できる。

カテゴリBの申請時にチェックボタンを押すのを忘れないこと。

システムPにおいて、P-large-b*キューの利用を希望しますか？ / Do you want to use P-large-b* queue in System P? *

- ☐ はい
- ☐ いいえ

P-large-b*キューは、大きなメモリ量が必要なジョブ実行のためのキューで、P-large-bキューの2倍の単一ジョブ最大コア数をご利用いただけます(単一ジョブ最大コア数は896)。ただし同時実行可能コア数はP-large-bと同じです。

P-large-b*キューの利用可否はCfCA 時間割り当て委員会にて決定されます。

XD-S

これまで大規模実行はあまり行われてこなかった。

- 申請の仕方がわかりにくいのではないかな？
- 富岳ではなくCfCAで大規模実行をする意味が明瞭ではなくなってきたのではないかな？

こちらでやりかたを決めるのではなく、利用者からニーズを掘り起こしたい。12月1日運用開始後に**随時募集し、利用者と相談しながら実行のやり方を決める。**

- 申請がきたら採否はTACで判断する
- どのようなやり方をしたいのかは利用者が指定(システムMだけ利用 or 全系利用。全系を数日 or 半系を1週間)
- 申請書はXD-Aなどの申請を再利用して構わない(その場合サイエンスはレフェリーにより審査済み)。
- 申請から2ヶ月程度後の実行を想定。
- **通常の「問い合わせフォーム」から受け付ける**

キュー構成

- XC50は40core/node, 1005node => 40200 core
- XD2000 システムMは112core/node, 208node => 23296 core
- XD2000 システムPは112core/node, 80node => 8960 core

これまでのキュー構成にコア数を比をかけたものを初期のキュー構成とする。

詳細はWEBページを参照

largeキュー

このキューは各カテゴリごとに以下の設定値でジョブを実行できるキューです。

カテゴリ	XD-A	XD-B+	XD-B	XD-MD	XD-Trial
キュー名	M-large-a	M-large-bp	M-large-b	M-large-md	M-large-t
最大同時利用可能コア数(ノード数)	11648(104)	2128(19)	672(6)	448(4)	112(1)
単一ジョブ最大コア数(ノード数)	11648(104)	2128(19)	336(3)	224(2)	112(1)
最大同時投入数	無制限	無制限	無制限	無制限	無制限
最大同時実行数(括弧内は混雑した場合の値)	10(1)	10(1)	10(1)	10(1)	5(1)
継続時間	24 hour	24 hour	24 hour	24 hour	4 hour

申請書の変更点

- 申請にメモリ量の見積りが必須となった
- メモリ量の見積りはXCを利用して測っておく必要がある。やり方は後述。

申請書の変更点

- 申請にメモリ量の見積りが必須となった
申請フォーム

使うシステムを選択してください(複数選択可)/Select the system you want to use (multiple selection possible) *

☒ システムM

☒ システムP

システムM：ノード時間積 [k nodes hours] / The value of multiplying the number of nodes by the total calculation time in System M [k nodes hours] *

k nodes hours

システムMで使う予定のノード時間積(使用するノード数 × 計算時間 [h])の見積もりを 10^3 nodes hoursの単位で記入してください。/Please write down the estimated value of (number of nodes)*(time [h]).

システムM：要求メモリ量 [GB/node] / The requested memory size per node in system M [GB/node] *

GB/node

システムMで要求するノード当たりのメモリ量の見積もりをGB/nodeの単位で記入してください。/Please write down the requested memory size per node in system M [GB/node].

申請書の変更点

- 申請にメモリ量の見積りが必須となった
TeXの申請書、3-2requiredres.texに記述

```
% Describe your estimate of the amount of computing resources required
%
% Required for ALL the categories
% Language: Japanese or English

% For the XD-* categories, uncomment the following lines with appropriate values
%\def\Mnodetime{0}    % node*time [k nodes hours] for M
%\def\Mmempernode{0} % memory/node [GB/node]      for M
%\def\Pnodetime{0}    % node*time [k nodes hours] for P
%\def\Pmempernode{0} % memory/node [GB/node]      for P
```

申請書の変更点

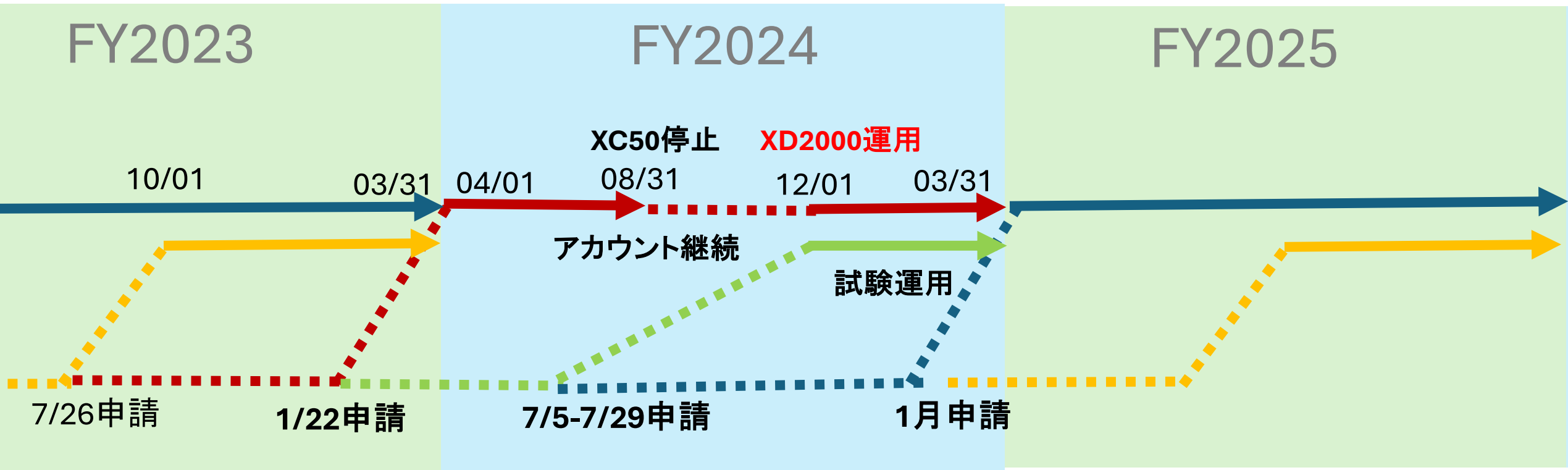
- 申請にメモリ量の見積りが必須となった

Wordの申請書

ここではメモリ使用量としてノード当たりの値 (単位は GB/node)を記入することに注意してください。システム M が搭載するノード当たりのメモリ量は 128 GB/node であり、システム P では 512 GB/node です。また、ここで見積もった数値は以下の表内にも記入してください。使用しないシステムについては 0 を入力します。なお表の位置と黒字で書かれた文字は変更しないでください。↵

↵	システム M↵	システム P↵
ノード時間積(k nodes hours)↵	8.4↵	33.6↵
ノード当たりのメモリ量(GB/node)↵	50↵	200↵

スケジュール



7月29日までに使用メモリを測定する。申請しない方も測っておくと良い。
8月31日までにxcのwork領域をコピーしておく。workはXD-2000に引き継がれない。一方でxcのhomeはひきつがれる。(とはいえバックアップを推奨)

XD2000以外の機材について

- 解析サーバ、ファイルサーバGPUは通常通り稼働（定期メンテナンス以外）
- 中規模サーバ、計算サーバは統合、OS更新などで長期のメンテナンス

中規模サーバと計算サーバの統合

- 現在
 - 中規模サーバ→全利用者が申請無しで利用可能
 - 計算サーバ→利用申請を行い、審査を通過後に利用可能
- XD2000導入の時期以降
 - 中規模サーバと計算サーバの機材が統合され、単一のシステムに
 - 利用申請を行い、審査を通過後に利用可能
- 日程の案
 - 2024年9月1日-11月末 システム構築作業 (運用停止)
 - 2024年12月- 運用開始 (早まる可能性もあり)
- その他
 - 作業領域(/cfca-workなど)にあるファイル群は保持されます
 - 日程や運用方法の詳細は今後の広報をお待ちください