

lio_transfer お忘れなく

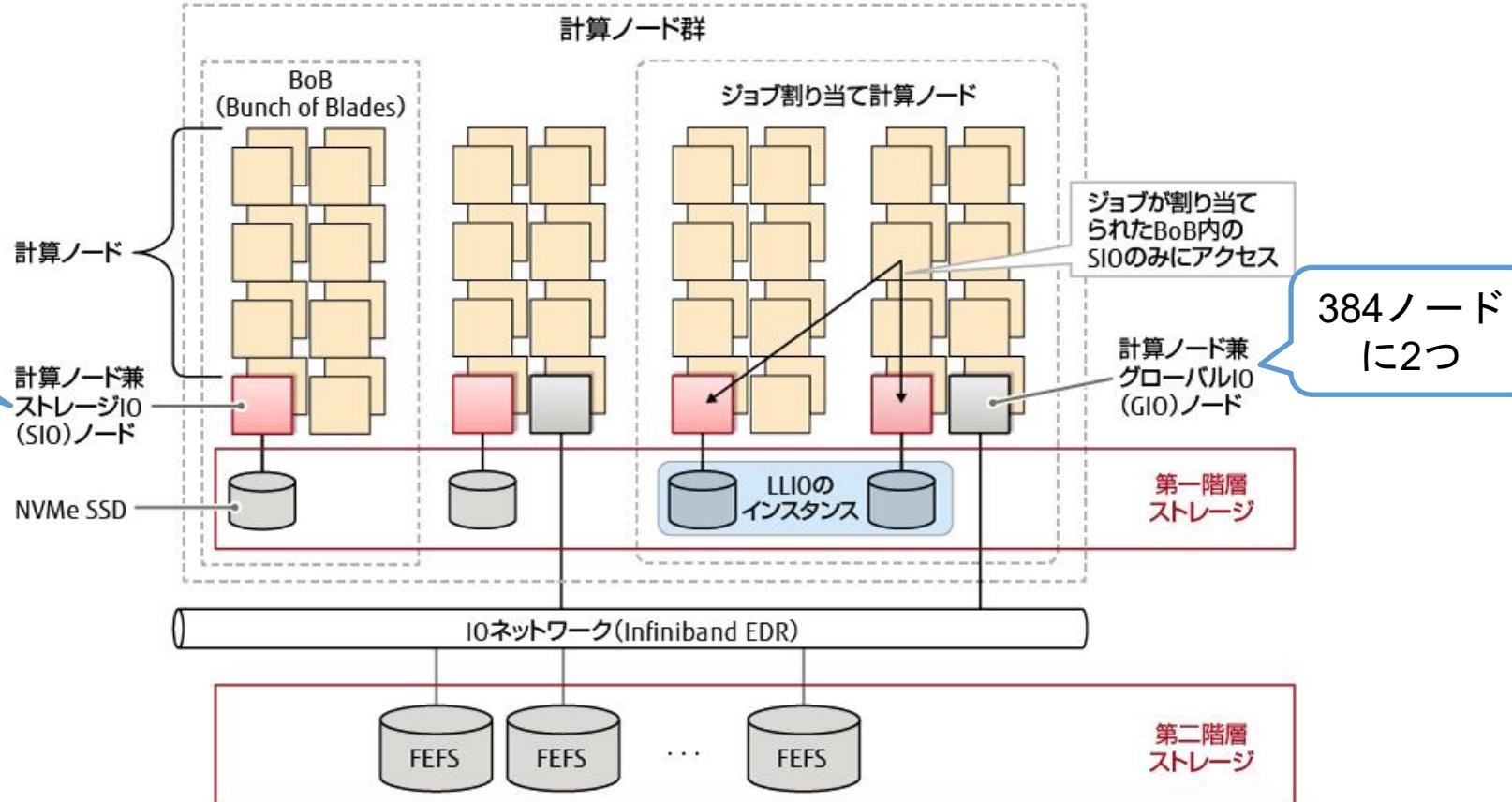
「富岳」ユーザーブリーフィング
トピックス(1) LLIOご紹介

2025年7月

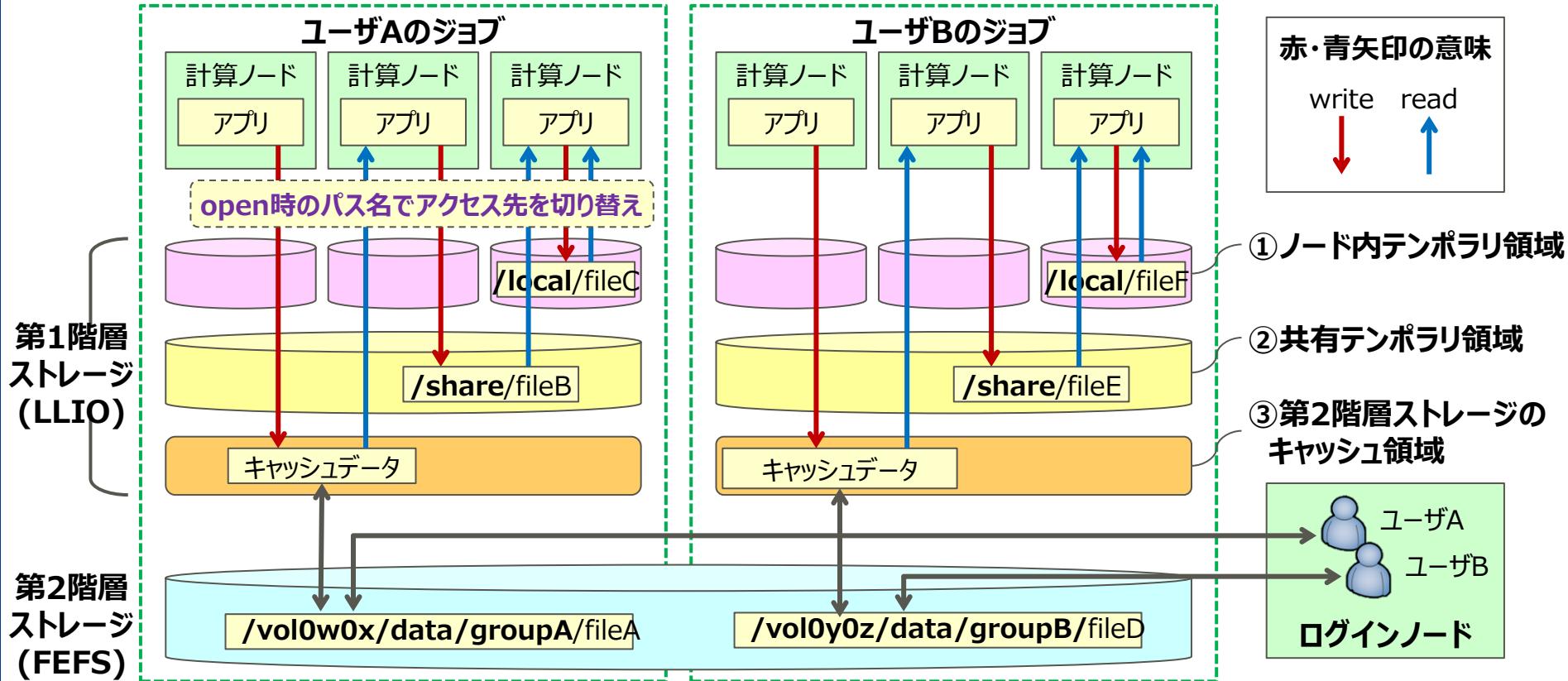
国立研究開発法人理化学研究所
計算科学研究中心 運用技術部門

おさらい - 「富岳」のIO構成

- <https://www.fujitsu.com/jp/about/resources/publications/technicalreview/2020-03/article05.html>

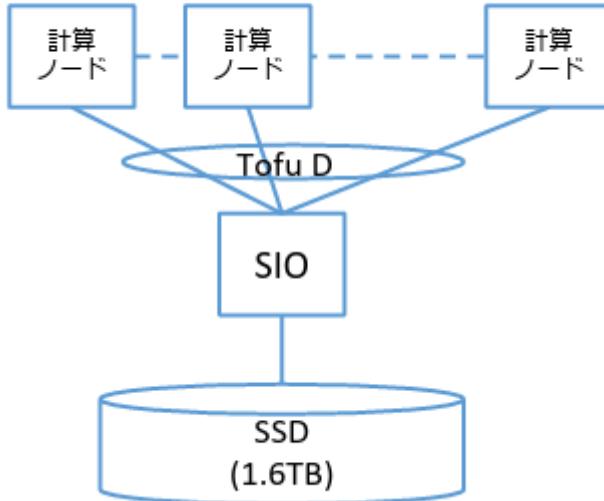


おさらい - 階層ファイルの概要 (LLIO+FEFS)



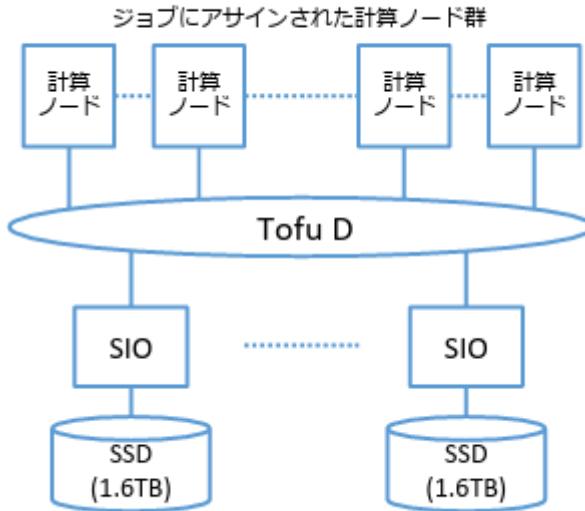
LLIO利用時は制限事項に注意してください。特に、ノード数が多い場合、対策を行わないとノードがダウンする場合があります。詳しくは手引書等を参照してください。

ノード内テンポラリ領域



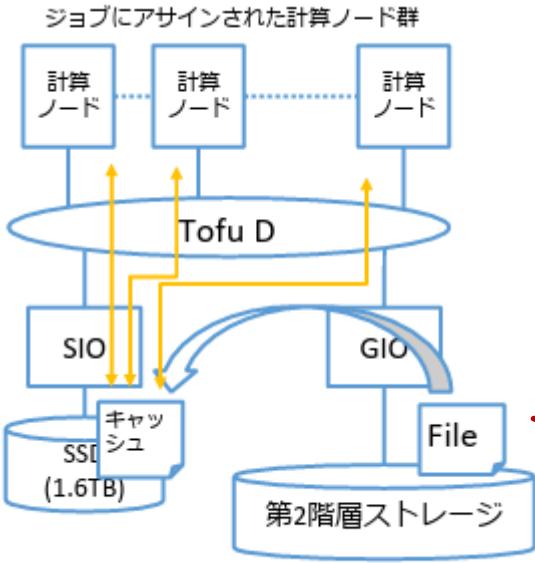
- ジョブに割り当てられたそれぞれの計算ノードで利用できる一時領域
 - ジョブ開始時に作成、終了時に削除
- SIOに接続されたSSDを16台の計算ノードで共有
 - 最大 87GiB / ノード
- SIOがメタデータサーバ

共有テンポラリ領域



- ジョブにアサインされた計算ノード間で共有できるテンポラリ領域
 - ジョブ開始時に作成、終了時に削除
 - 最大容量 87GiB × ノード数
 - SSD上にのみ存在
 - 第2階層へは書き出されない
- ジョブで使用できるSIOの数はジョブが使用するノード数でに比例
 - 規模の大きいジョブは、SIOを専有するようにノードを割り当てる
- LLIO用のMDSを用意

第二階層ストレージのキャッシュ領域



- 第2階層ストレージのキャッシュ機能
 - 同期/非同期クローズ
- ジョブにアサインされた全ての計算ノードから参照可能
 - ジョブ開始時に作成、終了時に削除
- 共有読み込みファイルの一斉転送機能

- 通常アクセスでは、一つのSIO上にキャッシュが作られるため、多数ノードからのアクセス時にアクセス性能が低下
- llio_transferコマンド
 - 全てのSIOにレプリカを作成し、高速なファイルアクセスを実現

LLIO制限超過 😳

- 1ファイルを多数ノードからアクセス → SIOがダウン
 - プログラムが扱うファイルだけでなくプログラムやライブラリが格納されたファイルに注意
 - 7,000ノードを超えるジョブ
 - 28,000プロセス以上からの同一ファイルアクセス
- llio_transferコマンドで全SIOにレプリカ作成

```
$ llio_transfer ./a.out
```
- 参考
 - [LLIO利用制限超過の通知がきた](#)
 - [LLIOのIOスローダウンを防ぐ（LLIO利用制限）](#)