

データ共有を念頭においたGRID環境のテスト

長岡技術科学大学 佐々木幸次、高橋弘毅

新潟大学 大原謙一

大阪市立大学 神田展行

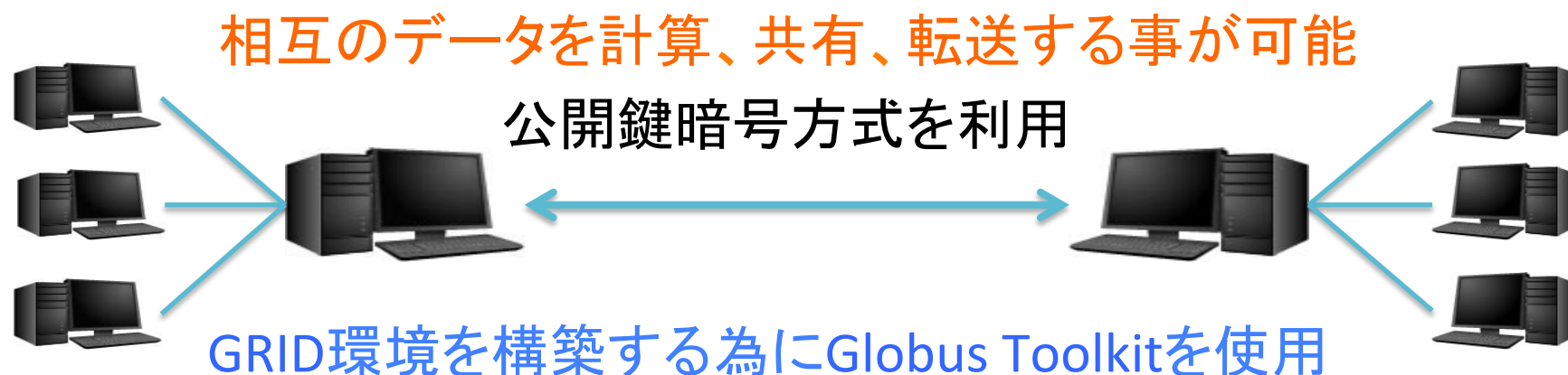
KAGRA Collaboration

GRID技術の特徴

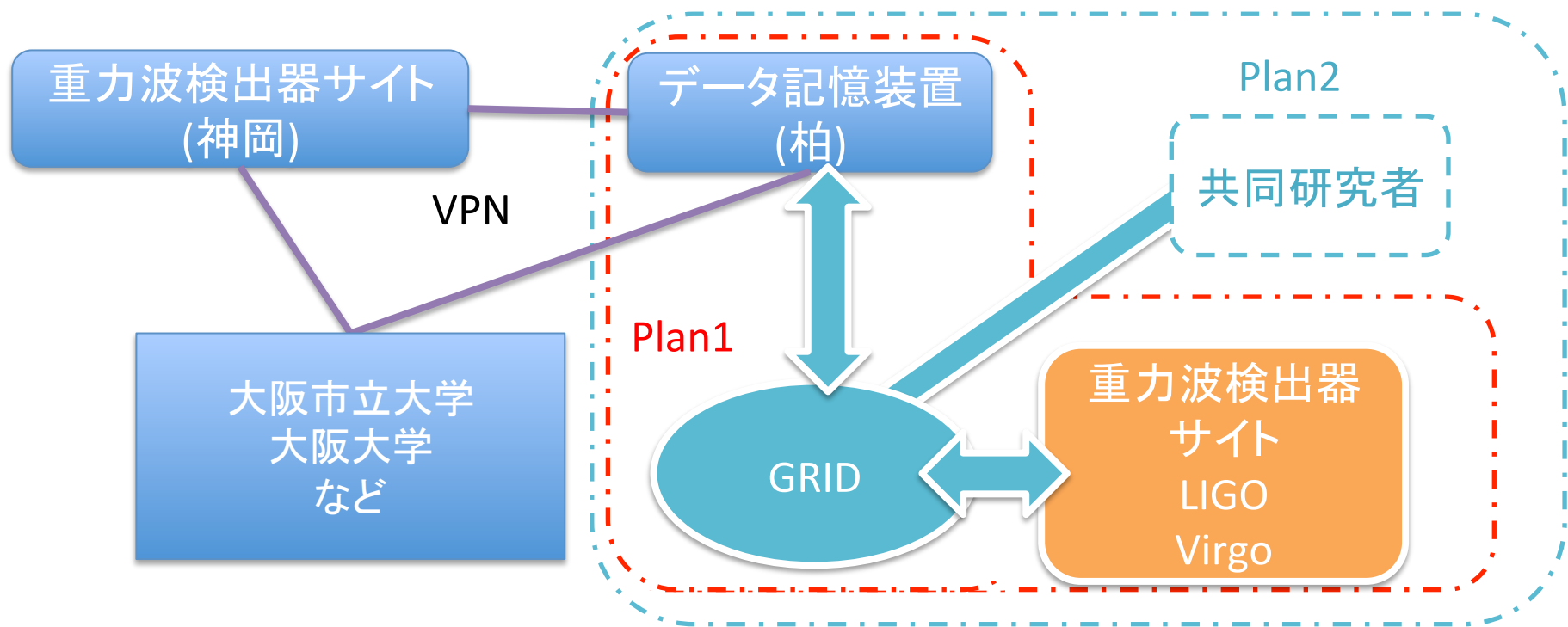
✓ KAGRAとLIGO、VIRGOがデータを共有する為にGRID技術の使用を検討

➤ GRID技術の特徴

- ✓ 安全に使用出来るセキュアな環境を提供
- ✓ 遠隔地にあるサーバーと意識せずに接続可能
- ✓ 複数台のサーバーと接続しても、一つとして表現が可能



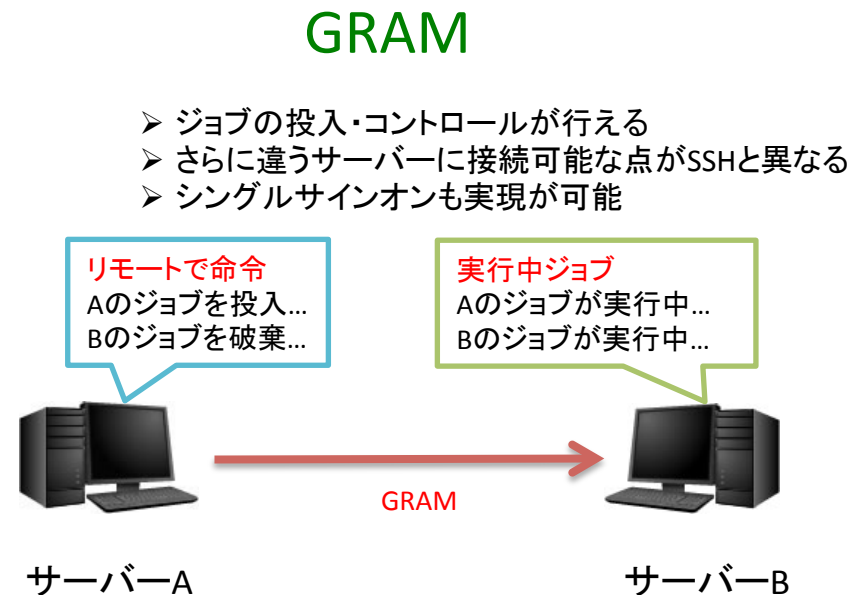
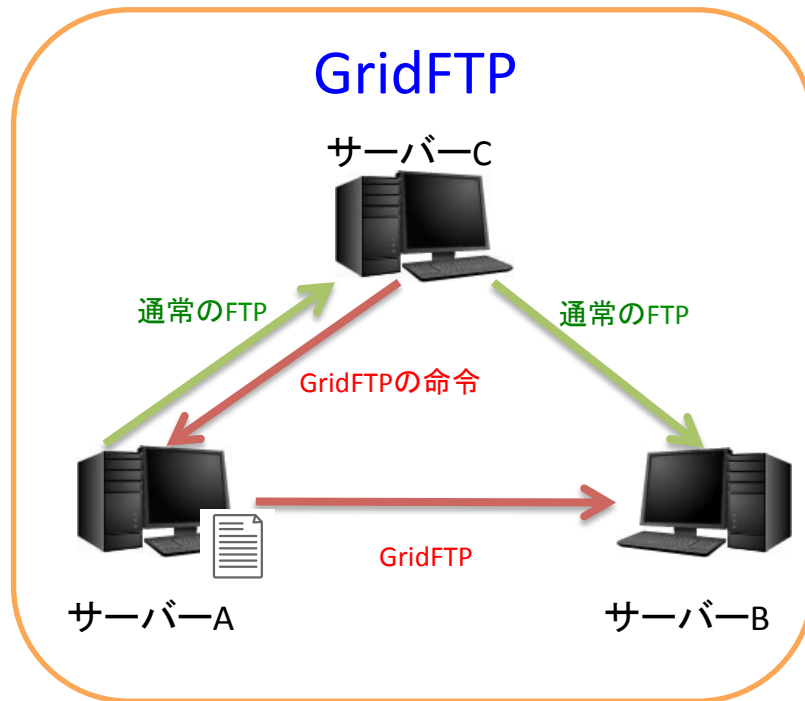
KAGRAにおけるGRIDの立ち位置



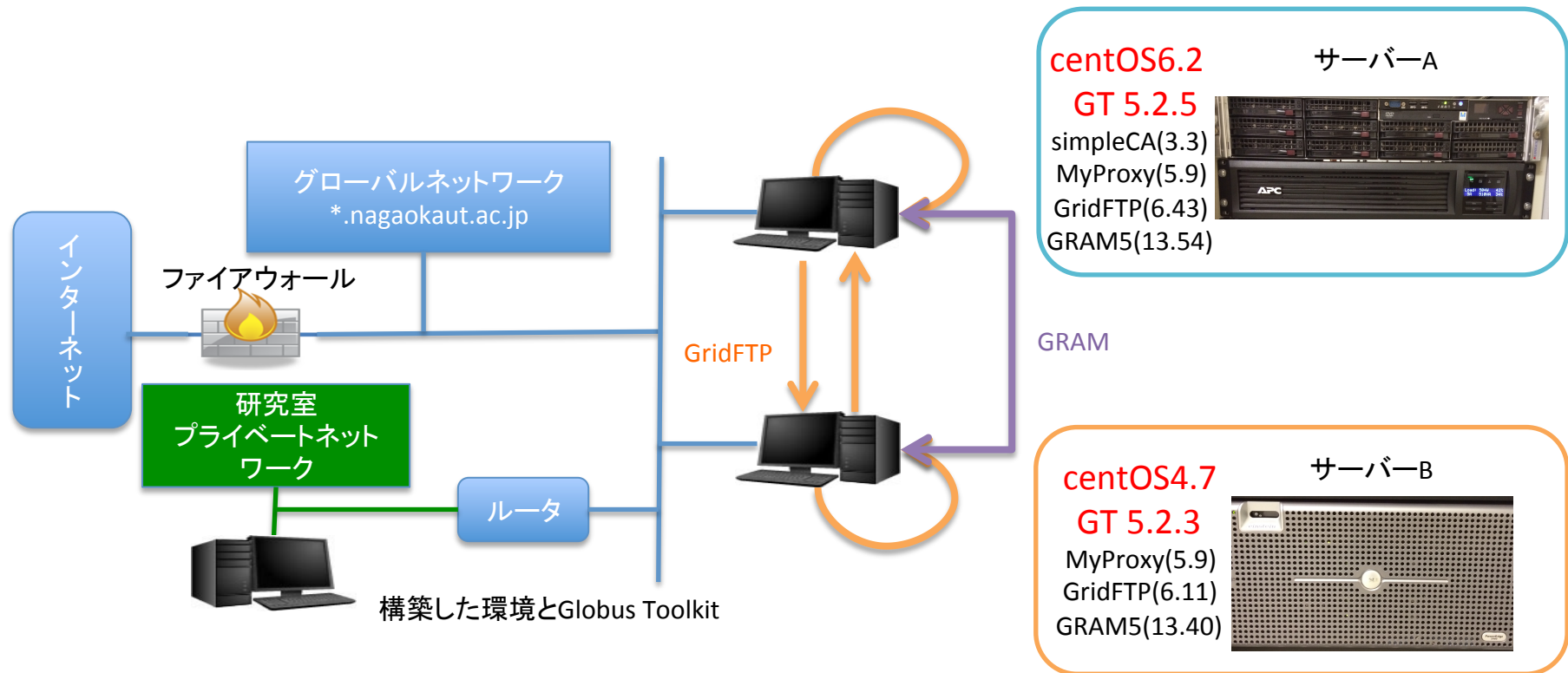
GRID環境の構築

Globus toolkit

- ✓ GRID環境を構築する一つのツール。機能は大まかに2種類存在する
 - ✓ GridFTPは第三者を中継せずにファイルの転送が可能
 - ✓ GRAMは遠隔でジョブを実行可能



構築したGRID環境



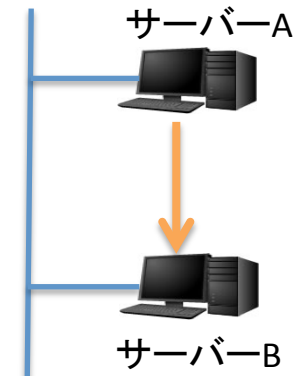
※学内の最大通信速度はファイアウォール以下で1Gbps
MyProxy : Proxy証明書の代理発行サーバー

SCPとGridFTPでのファイル転送速度の比較

- ✓ サーバーAからサーバーBのファイル転送

– コマンド:

```
globus-url-copy gsiftp://serverA.domain.jp/home/gtuser/datafile.zip \  
gsiftp://serverB.domain.jp/home/gtuser/datafile.zip
```



- ✓ 312MBのデータを間隔20秒、試行100回の平均を求めた

サーバーAからサーバーB間	平均速度
SCP	63.4±5.0 MB/s
GridFTP	99.2±4.1 MB/s
RCP	111.4MB/s

KAGRAデータフローの要求値 20MB/s

GridFTPでの転送速度の向上の考察

- GridFTPの転送速度向上の手段
 - ✓ バッファサイズをより大きくする事
 - ✓ 並列コネクションの数を増加させる事

バッファサイズとコネクションのオプションによる指定

```
$ globus-url-copy -p 1 -tcp-bs 1250000 gsiftp.....
```

推奨TCPバッファサイズ = 帯域幅 × RTT / 並列コネクション数

※RTT(Round Trip Time)とは: 信号を発してから応答が返って来るまでの時間
※オプションを指定しなければ自動でバッファサイズが調節される

- 並列コネクションの数を過度に増加させるとスループットが低下
- OSの上限値以上のバッファサイズを設定出来ない
- 実際の環境ではGridFTP以外のサービスが動作している
⇒過度にバッファサイズを設定すると他のサービスに影響を与える

※高瀬亘氏の 広域ネットワークにおけるGridFTPによる効率的なデータ転送 GridFTP Performance Analysis より引用

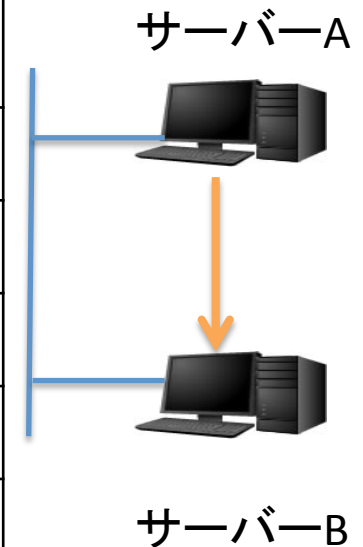
GridFTPでの転送速度の向上の考察

Pingで60回試行したときのRTT

rtt min/avg/max/mdev = 0.064/0.103/0.163/0.021 ms

平均RTTが0.1ms

設定バッファサイズ	転送速度MB/s コネクション数=1	転送速度MB/s コネクション数=2
12.5KB	91.6±4.0	101.6±5.9
125KB	102.5±4.9	101.3±6.2
1.25MB	102.0±5.5	101.2±7.1
2.5MB	101.6±5.9	101.7±6.2
6.25MB	101.8±6.4	101.6±5.9



帯域幅1Gbpsによる速度計測

まとめ

- GlobusToolkitを用いたGRID環境の構築に成功した
- SCPとGridFTPの速度転送の比較を行った結果、GridFTPがSCPの転送速度よりも速い事が分かった
- ただし、速度の差の原因が理解出来ていない為、今後も調査が必要

今後の課題

- ✓ VOMS (Virtual Organization)
 - ✓ European Data Gridプロジェクトで開発されたGRID技術
 - ✓ 日本のKAGRAがVOMS(KAGRA VO)をデータ転送系で利用
 - ✓ 仮想的な組織による計算機資源の制限
⇒メンバデータベースによって研究グループのメンバを判断する

Globus Toolkit



認証局を使用した通信でのデータ共有



KAGRA VOによるデータ(計算機資源)の利用制限



VOMS(KAGRA VO)

高エネルギー加速器研究機構がGRID環境サービスを提供

高エネルギー加速器研究機構のGRID環境との接続テストを行う