

HPCI共用ストレージ東拠点の機材更新 に伴うデータ移行について

2017/12/22

中 誠一郎(東京大学)

機材更新の内容

HPCI共用ストレージ:H29更新イメージ

東大またはAICSによる単独運用可能な高いサービス継続性

遠隔地間を含む
メタデータフェイルオーバー
による高可用性

クラウドストレージの活用
によるアーカイブサービス

Cloud Storage

Meta Data
Storage
2~4TB

Meta Data
Storage
2~4TB

DataStorage
30PB~50PB

SINET5による広域・高帯
域データ通信

Meta Data
Storage
2~4TB

Meta Data
Storage
2~4TB

DataStorage
30PB~50PB

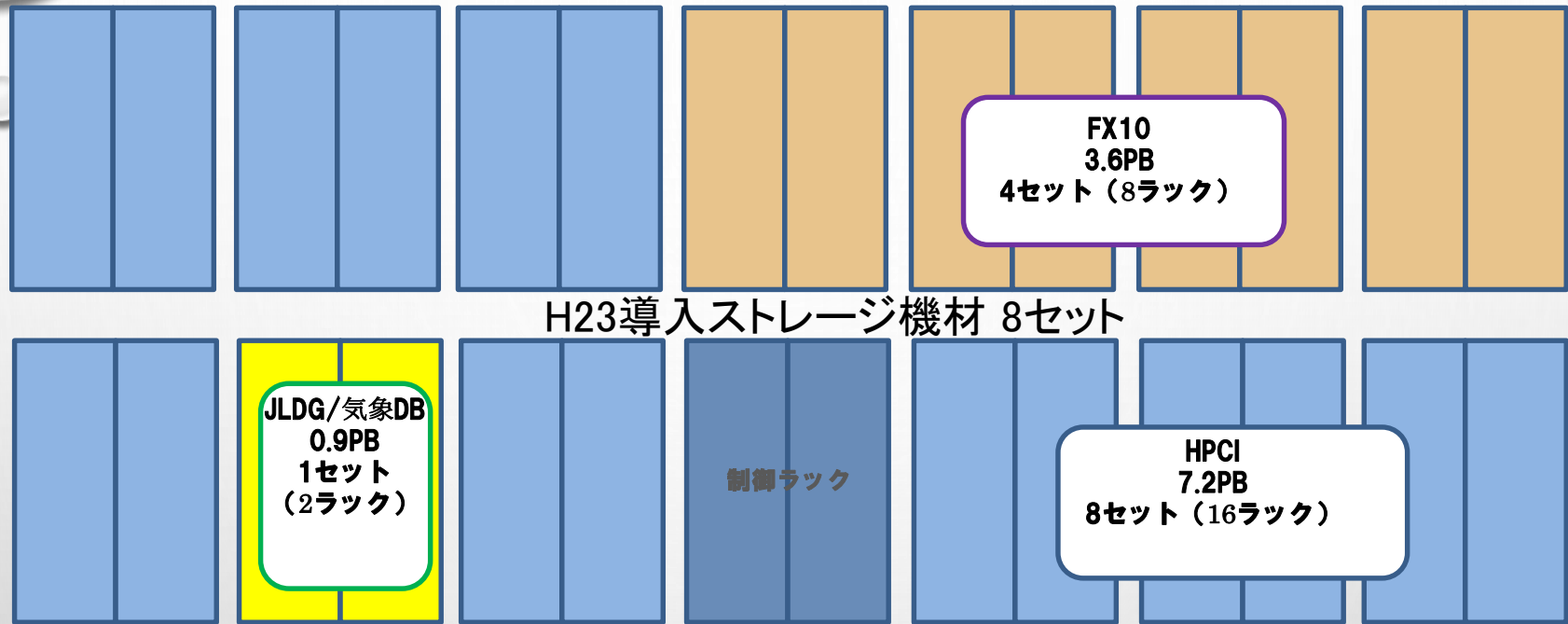
ラック
10~16台
のスペース

完全二重化による
高いデータ安全性

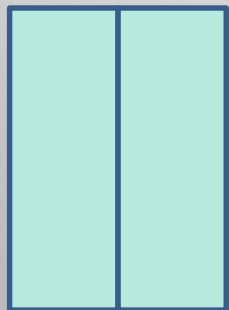
理研計算科学研究機構

東京大学情報基盤センター

Before 機材更新 (2016/7時点)



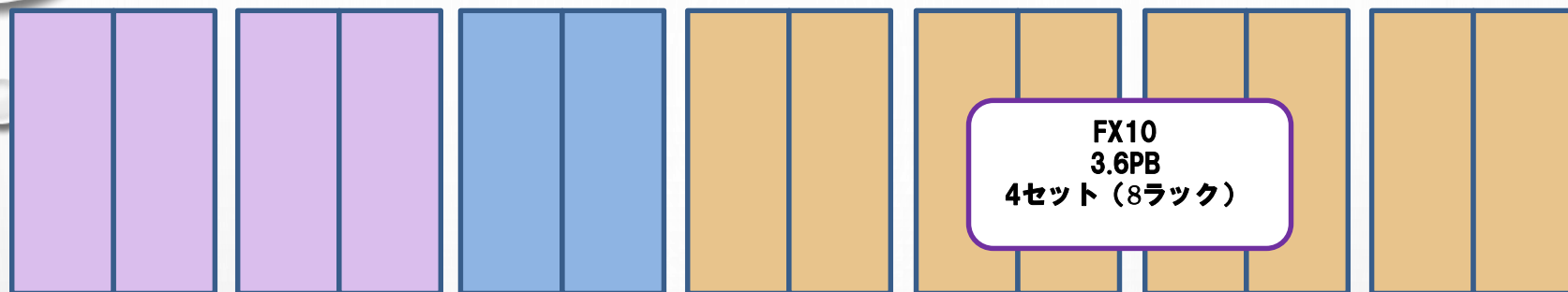
パイロットストレージ



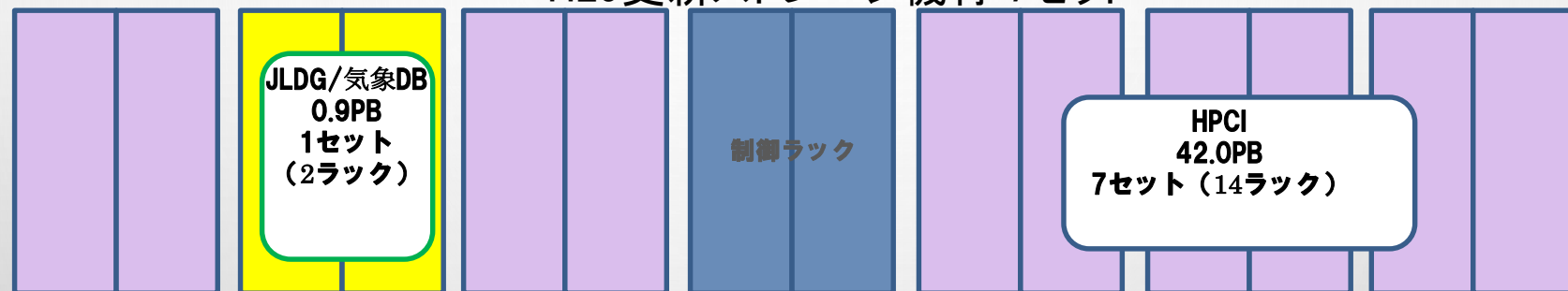
H24導入ストレージ機材 4セット



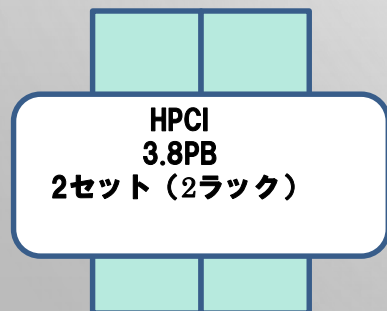
After 機材更新 (2017/10時点)



H29更新ストレージ機材 7セット



H28移行用ストレージ
(データ保存用ストレージ)



SFA14KXE 7セット
容量 42.0PB

データ移行方式

データ移行手順

データ移行の終了した既存ストレージ機材から順次撤去、その場所に新規ストレージ機材を導入・設置、その新規ストレージに残りの既存ストレージ機材からデータ移行するという、スライディングデータ移行方式を採用

そのためには、既存ストレージ機材からの最初のデータ移行先が必要で、H28/10にパイロットストレージを撤去、移行用ストレージ(データ保存用ストレージ)を設置済み

データ移行方式

【基本方針】

複製数1のファイルが多数存在しており、データ保全のため、データ移行時には移行元にファイルを必ず残す

□ gfprep方式(ファイル複製の移動コマンド)

- -mmオプション: 転送元ファイルを残す指定
- 余剰レプリカ自動削除機能: disable

□ replicainfo方式(複製自動配置機能)

- 余剰レプリカ自動削除機能: disable

□ 移行終了後に、移行元ファイルサーバのgfsdを停止し、replica_checkの周回でno available replicaエラーが出ないことを確認

□ 移行先ファイルサーバの安定稼働を確認した後に、移行元ファイルサーバをGfarmから削除

replicainfo方式

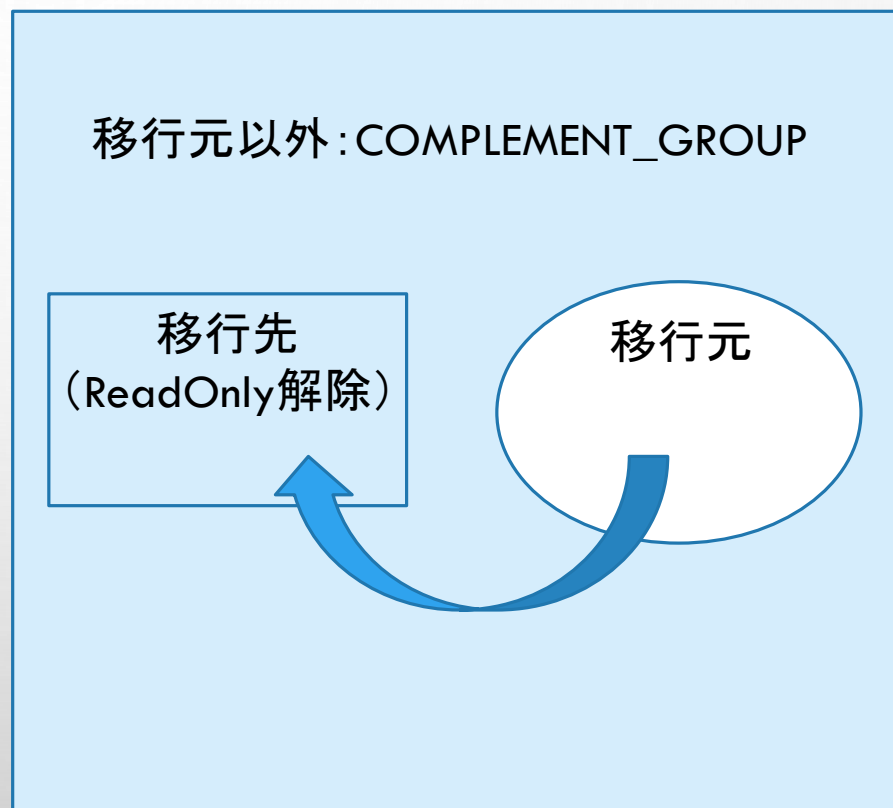
ホストグループ・replicainfo設定に基づいた、replica checkの自動複製配置機能を用いたデータ移行方式

- ・移行元サーバ以外の全ファイルサーバ（補集合サーバ群）を、ホストグループ（COMPLEMENT_GROUPと命名）に入れる

- ・gfarm:/// 及び NCOPY数明示設定箇所にてreplicainfo拡張属性を設定し、COMPLEMENT_GROUPに縮退（移行元サーバを外す）しても全ファイルのレプリカ数が維持されるように設定する

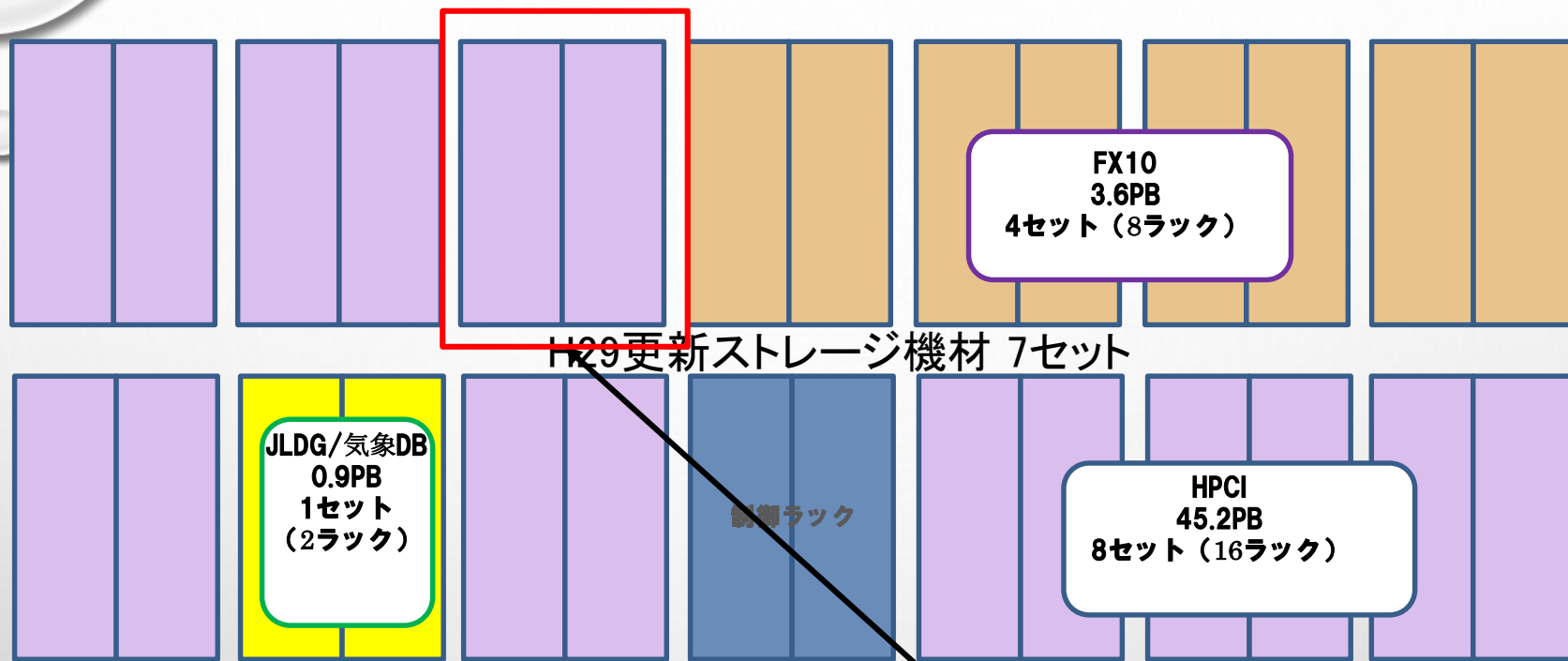
- ・COMPLEMENT_GROUPの中で移行先サーバのみ、ReadOnlyを解除する

- ・その結果、replicainfo設定に基づきreplica check周回時に、移行元サーバから移行先サーバへの自動複製配置（＝データ移行）が行われる。

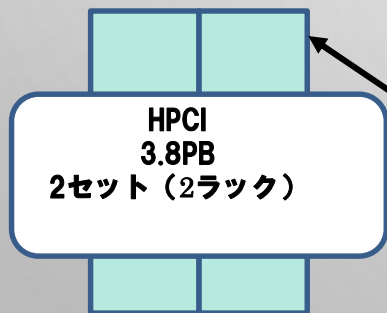


今後の予定

2018年度の予定



移行用ストレージ



2018年度上期:
データ入出力庫用途
ログインノード転用予定

2018/3月
8台目SFA14KXE増設:
容量 3.2PB