

J-PARC MLF NOVAにおけるDAQ用計算環境構築

KEK IMSS 瀬谷智洋

2017年10月2日

計測システム研究会@函館アリーナ

目次

- J-PARC MLFのDAQシステムの現状
- DAQ用サーバーの監視体制
- BL21 NOVAのDAQシステム
- 新しいNFSシステムの提案
- NFSの更新
- NFSのベンチマーク

J-PARC 施設
(日本原子力研究開発機構と
高エネルギー加速器研究機構の
共同事業)
2009年7月16日鳥瞰図

北
4

リニアック

3GeVシンクロトロン

ニュートリノビーム
(神岡へ)

物質・生命科学実験施設

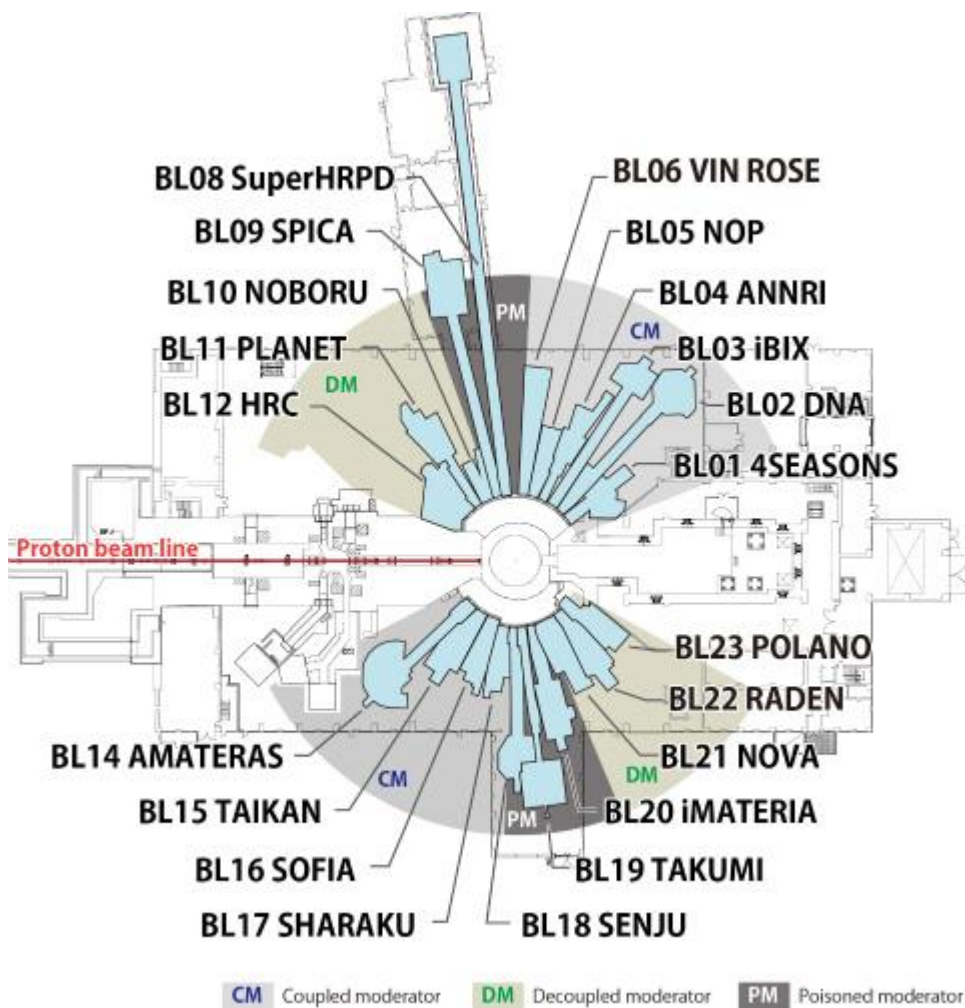
50GeVシンクロトロン

ハドロン実験施設

- 平成19年度ビーム
- 平成20年度ビーム
- 平成21年度ビーム



MLF: 物質生命科学研究施設



全21ビームライン中、
8ビームラインをKEKが運用している。

KEK 各ビームラインのDAQシステム

	NeuNET	DAQサーバー	NFS	解析サーバー
BL08	88	10	1(26TB)	2
BL09	199	16	1(146TB)	2
BL12	48	2	3(86TB)	1
BL21	122	7	2(250TB)	4
BL23	44	3	2(100TB)	1
合計	501	38	9(680TB)	10

- 全体で4008本のPSD、8016chの信号を、
- 501枚のNeuNETボードと38台のDAQサーバーでDAQしている。
- NFS,解析サーバーをあわせて、全部で80台のサーバーを実験に使用している。
- 安定的な実験の遂行にはサーバーの監視、管理が必要不可欠。



BL08

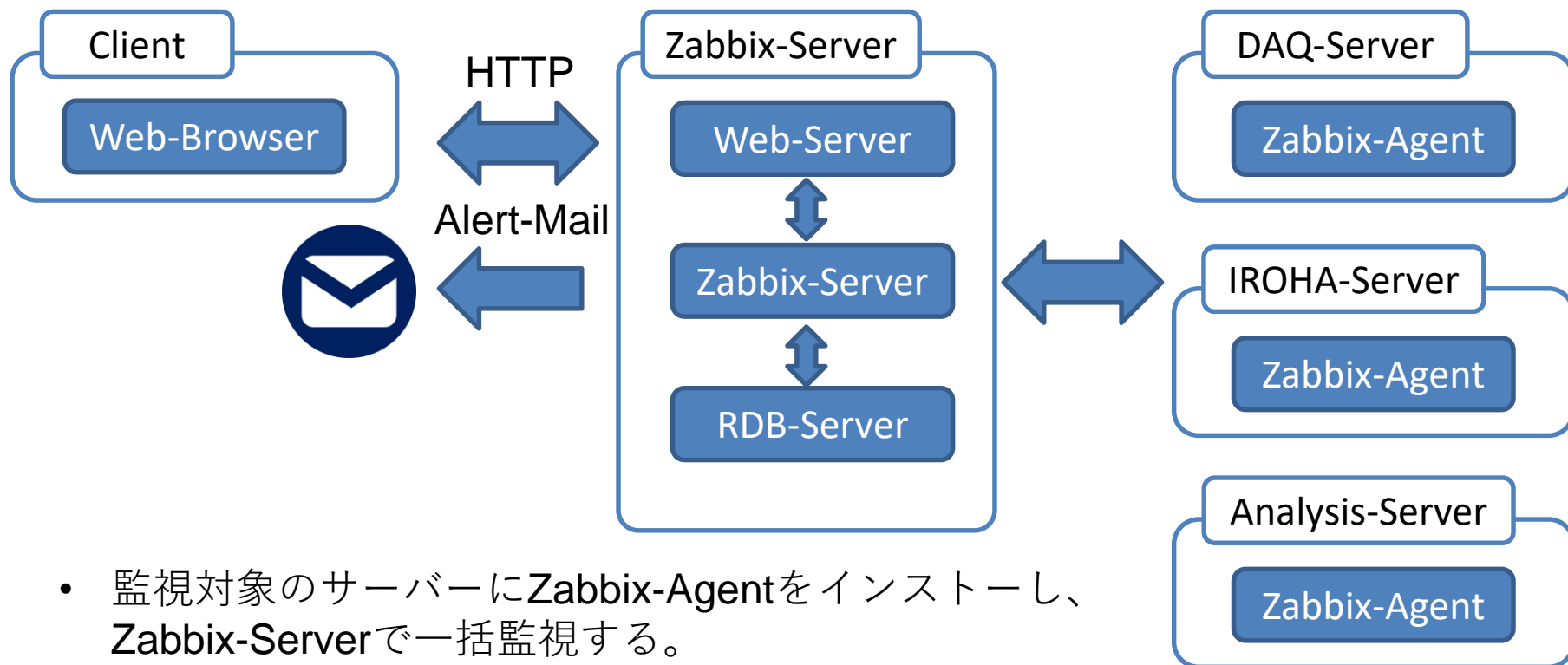


BL21



BL23

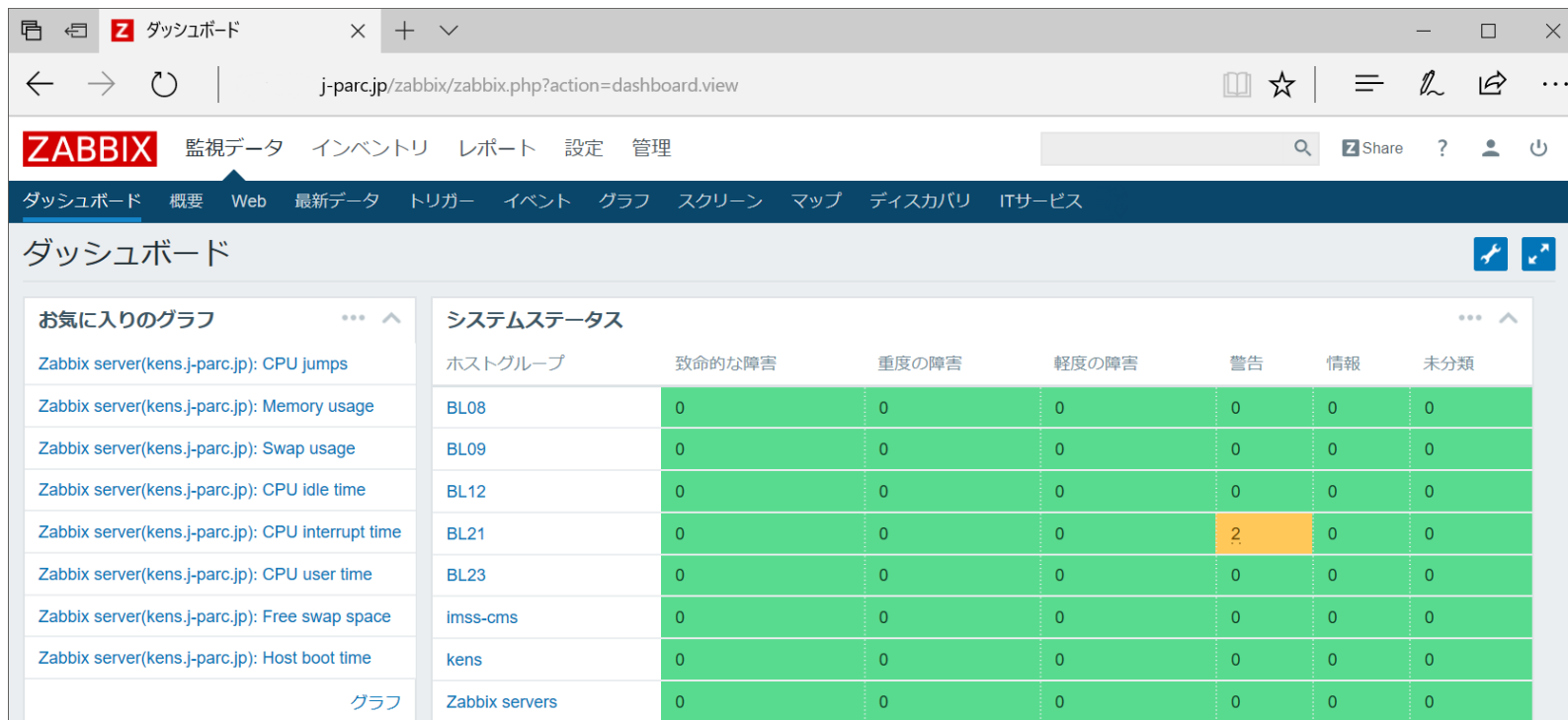
Zabbixによるサーバー監視体制の構築



- 監視対象のサーバーに**Zabbix-Agent**をインストールし、**Zabbix-Server**で一括監視する。
- 監視状況は、**Web**ブラウザで確認できる。
- 異常があれば、アラートメールが送信される。

Zabbixの運用状況

Zabbixの監視設定、状態確認はすべてWebブラウザから可能



The screenshot shows the Zabbix web interface. The top navigation bar includes 'ZABBIX', '監視データ', 'インベントリ', 'レポート', '設定', and '管理'. The main content area is titled 'ダッシュボード' and features two primary sections: 'お気に入りのグラフ' (Favorite Graphs) and 'システムステータス' (System Status).

The 'システムステータス' section displays a table of system status for various hosts. The table has columns for '致命的な障害' (Critical Issues), '重度の障害' (Severe Issues), '軽度の障害' (Minor Issues), '警告' (Warnings), '情報' (Information), and '未分類' (Unclassified). The '警告' column for host 'BL21' shows a value of '2', indicating two warnings.

ホストグループ	致命的な障害	重度の障害	軽度の障害	警告	情報	未分類
BL08	0	0	0	0	0	0
BL09	0	0	0	0	0	0
BL12	0	0	0	0	0	0
BL21	0	0	0	2	0	0
BL23	0	0	0	0	0	0
imss-cms	0	0	0	0	0	0
kens	0	0	0	0	0	0
Zabbix servers	0	0	0	0	0	0

サーバーを複数台運用しているビームラインに導入済み。
現在、BL08,09,12,21,23のビームラインで合計65台のサーバーを監視している。

DEEPI: DEEPI NETWORK TRAFFIC

BL21: Memory Usage

BL 21: Multi-purpose LAN Network Traffic

nova-ana1

Free disk space is less than 20% on volume /

2016/04/12 10:09:56

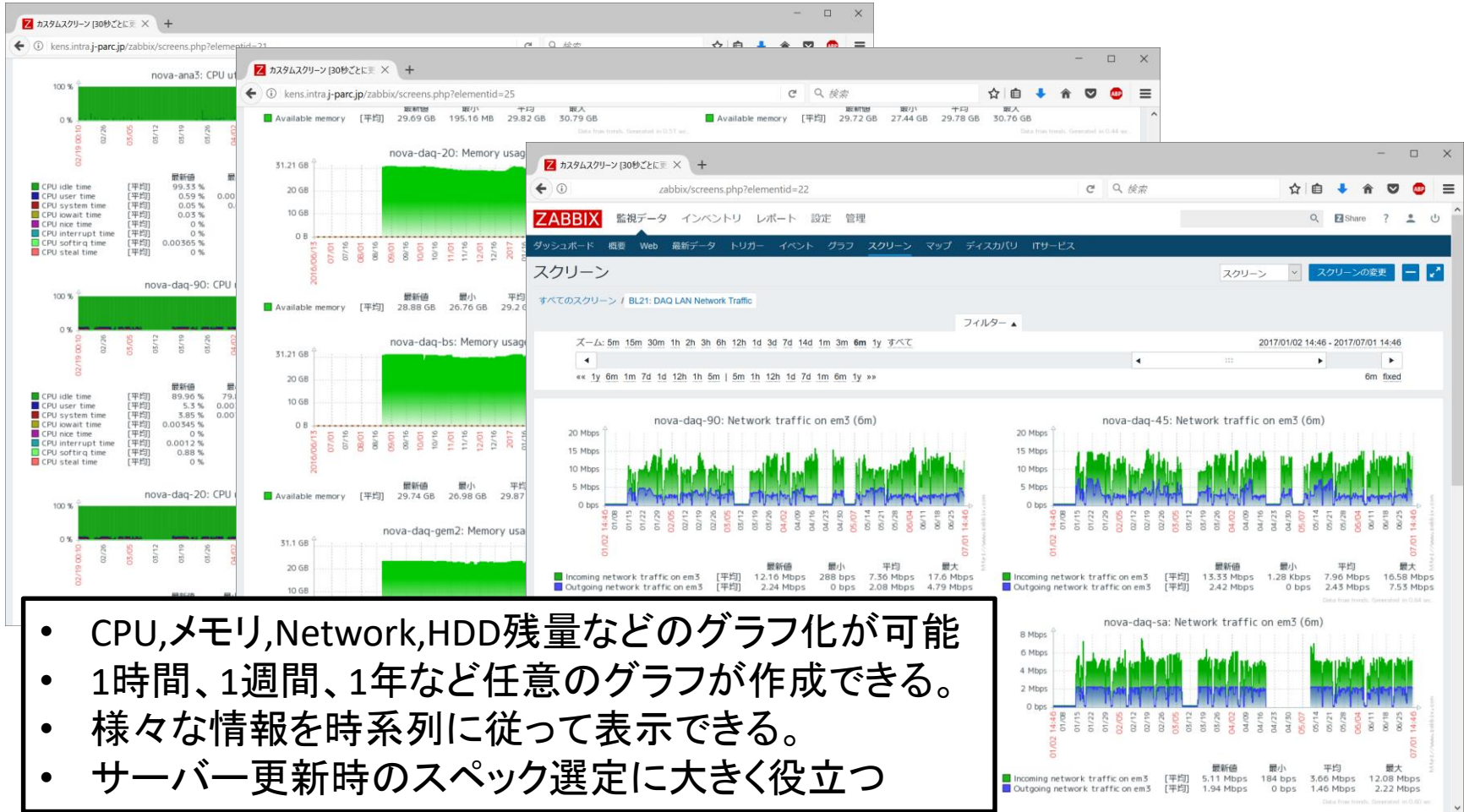
1y 4m 11d

いいえ

1

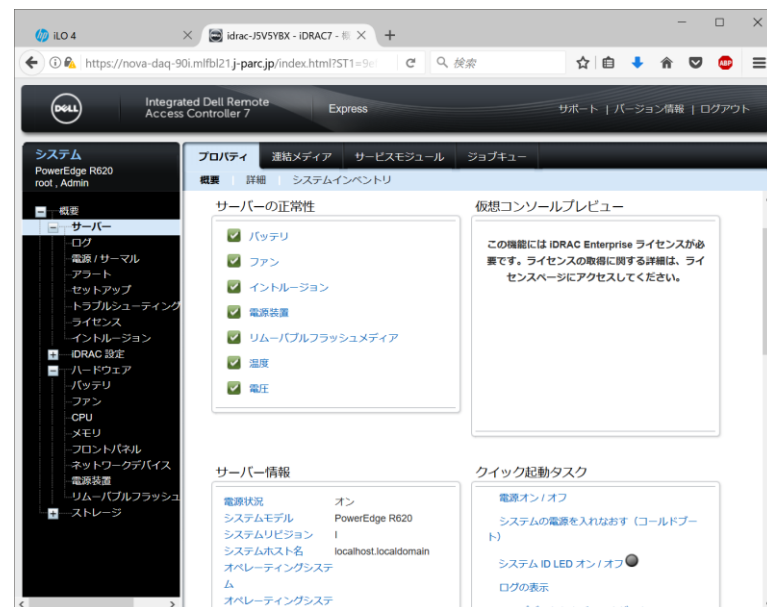
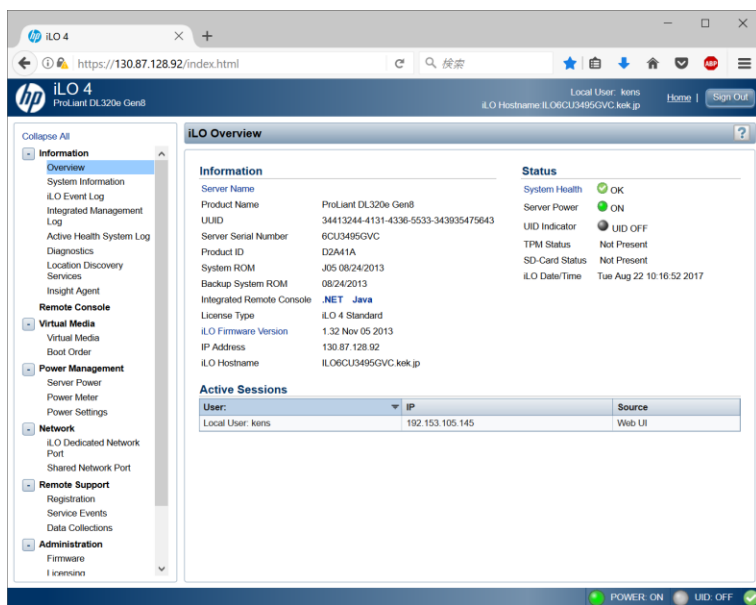
2件中2件の障害が表示されています 更新時刻: 10:14:16

Zabbixの監視機能

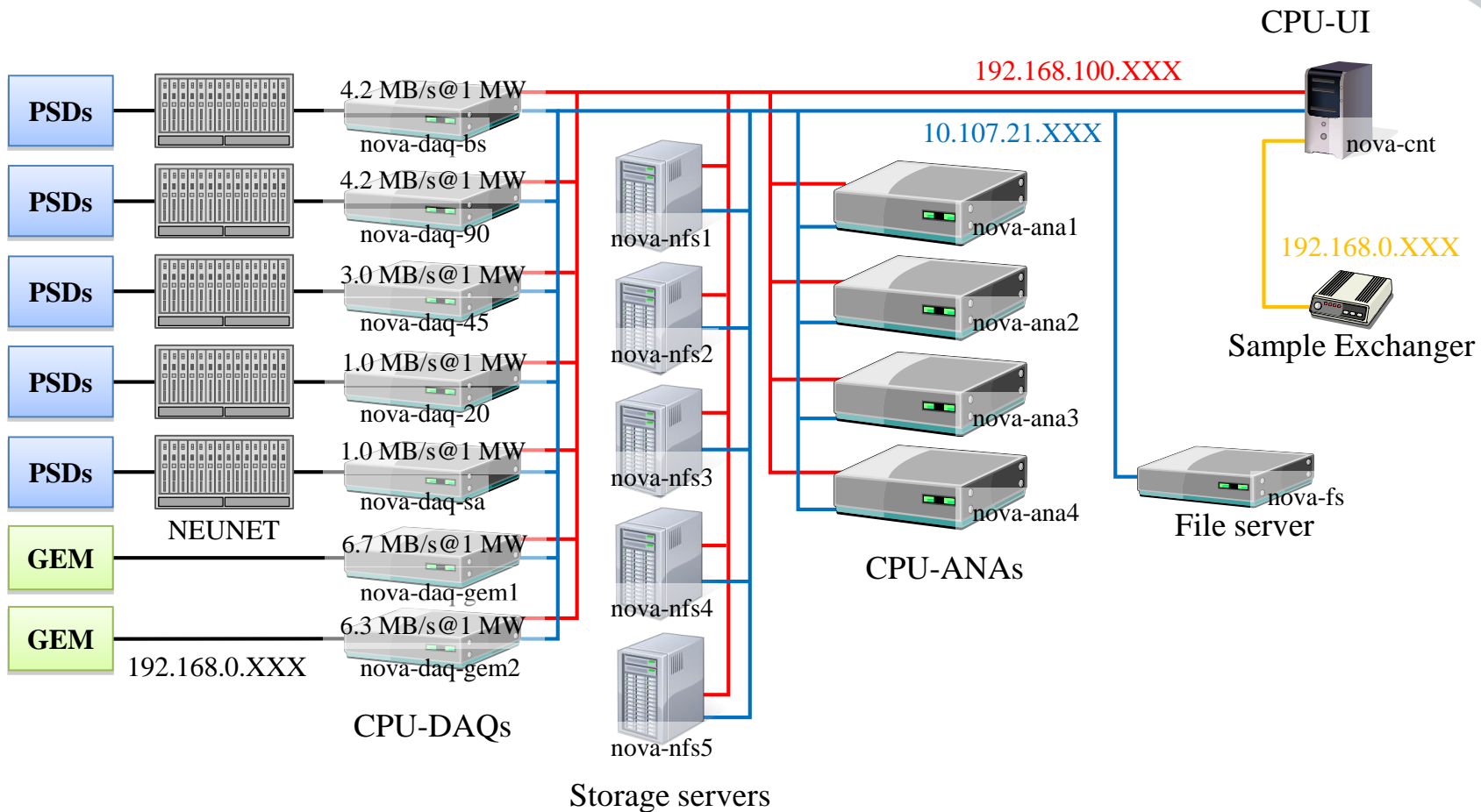


iDRAC iLOによるハードウェア監視

- iDRACはDELL、iLOはHPの組み込み型のハードウェア管理ツール
- マザーボードに実装され、OSの起動に関係なく、電源とネットワーク接続があればアクセス可能。
- Webブラウザからアクセス可能で、単独でアラートメール機能を持つ。
- Zabbixがサポートしない、HDD、ファンなどの故障を検知する。

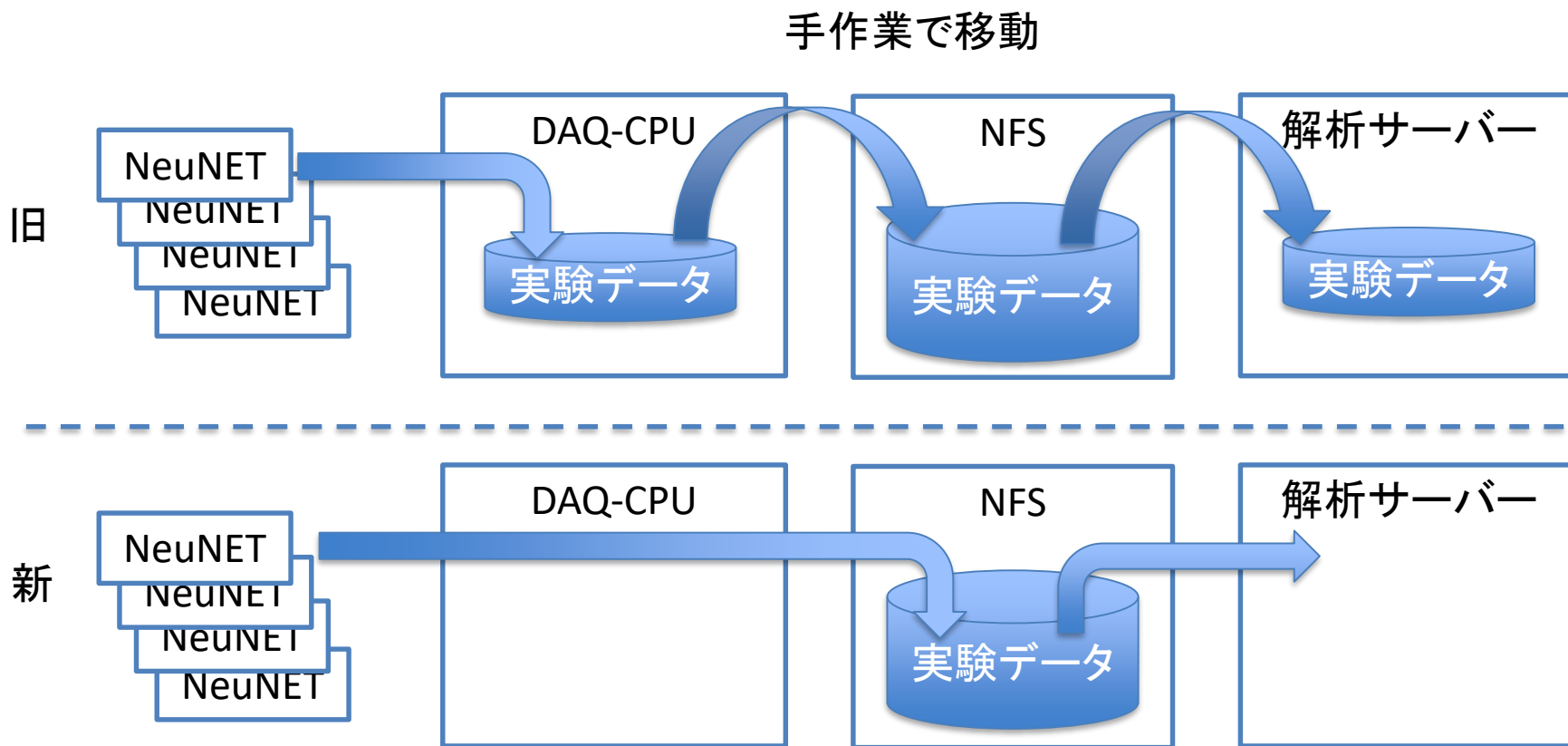


NOVAのDAQシステム



Total data size is approximately 26.4 MB/s (=2.3 TB/day) @ 1 MW.

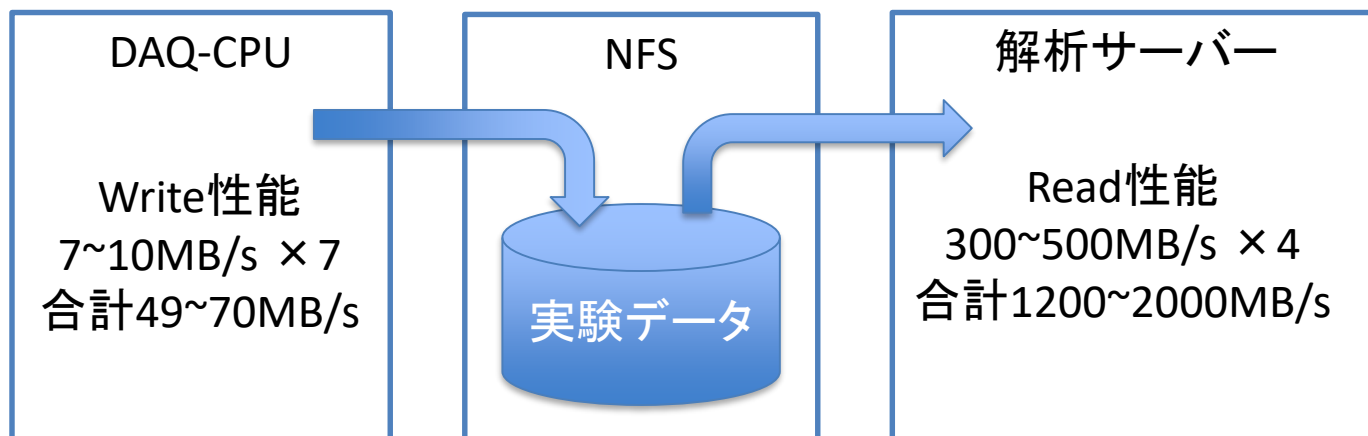
新ストレージシステムの構成



旧システムでは、各サーバーに実験データを置いていたが、
新システムでは、データはNFSだけに保存、各サーバーはNFSマウントして使う。

NFSの要件

- 冗長性を持たせるため、2台のNFSを購入する。
- 1台あたり100TB、2台で200TB以上の実効容量を持つこと。
- RAID 6に対応すること。
- ホットスワップに対応すること。
- 障害発生時には、スペアディスクを用いて元のRAIDを再構成する、オートリビルド機能を持つこと。
- Network File System(NFS)通信プロトコルに対応すること。
- 10GBASE-Tのポートを2ポート以上持つこと。このポートはリンクアグリゲーションに対応すること。



NFSの選定

メーカー	QNAP	ニューテック
製品名	TS-EC1280U-RP-R2	CloudyIII lx
OS	QTS 4	CentOS 7
HDD搭載容量	10TB×12 120TB	4TB×35 140TB
RAID6での実効容量	100TB	132TB
シーケンシャル・スループット	Read:3702MB/s Write:2685MB/s	Read:4052MB/s Write:3110MB/s
1台あたりの単価	—	—
総額(2台分の費用+保守)	—	—
テラバイト単価	—	—

NFSの設置

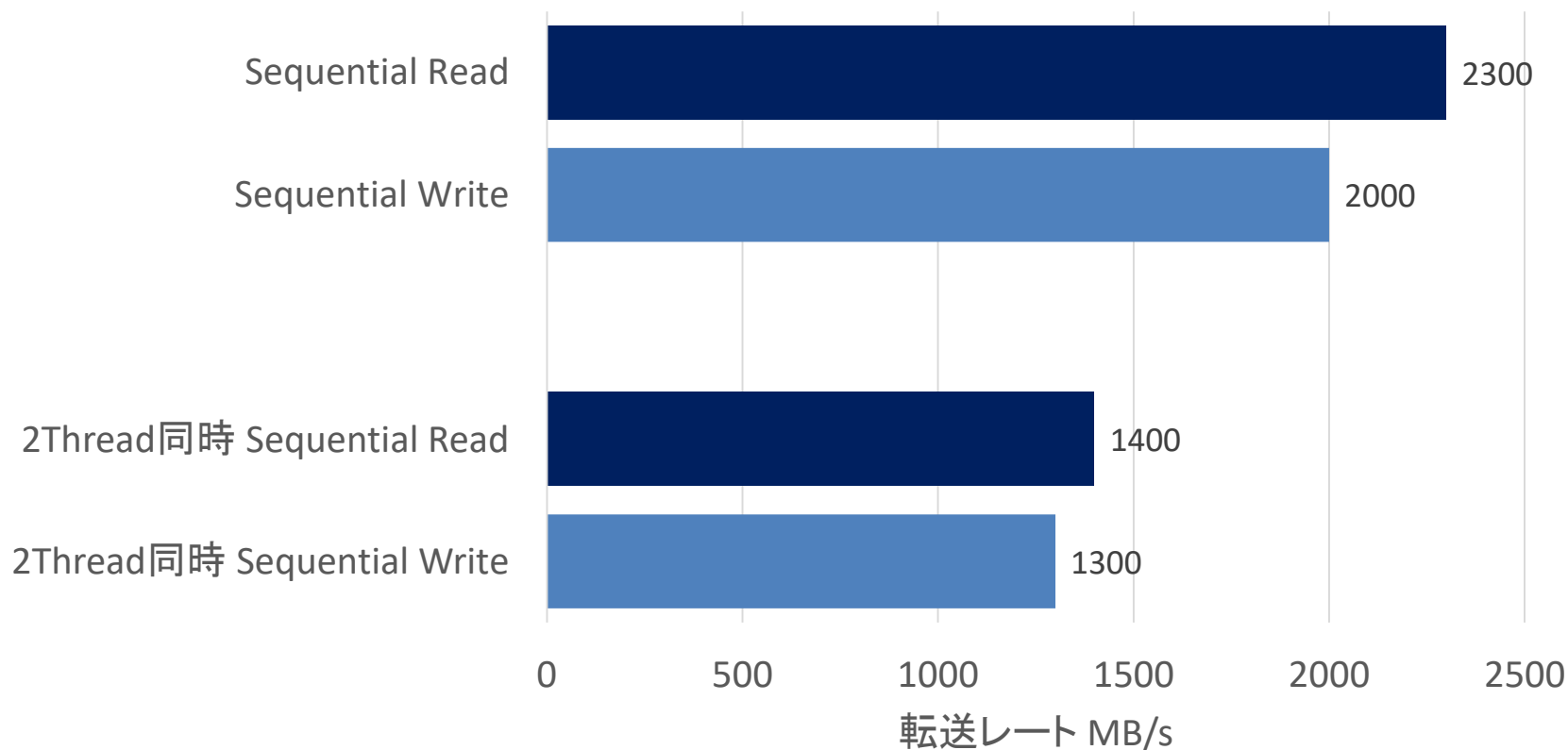


新NFS 正面



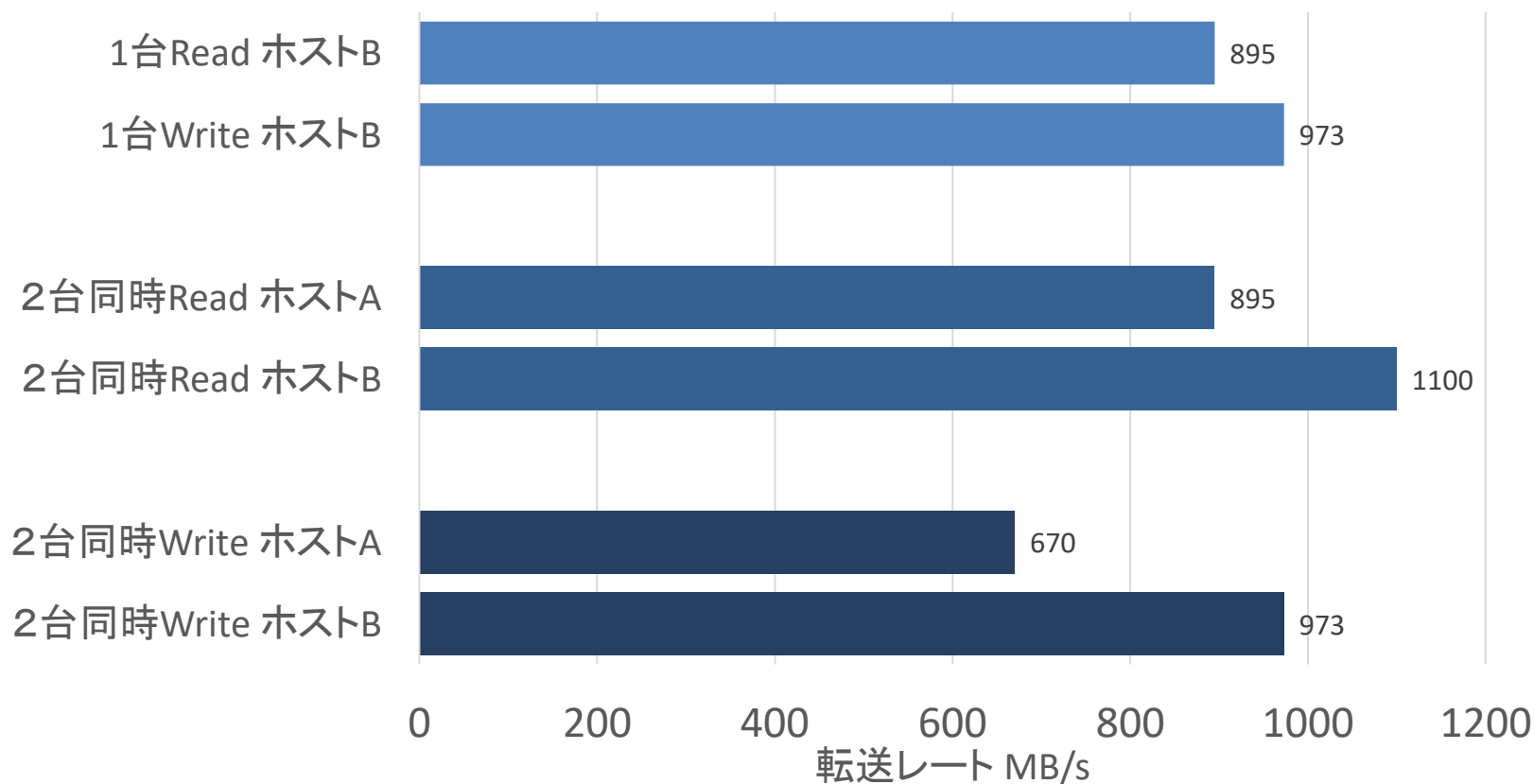
新NFS 背面

HDD単体のベンチマーク結果



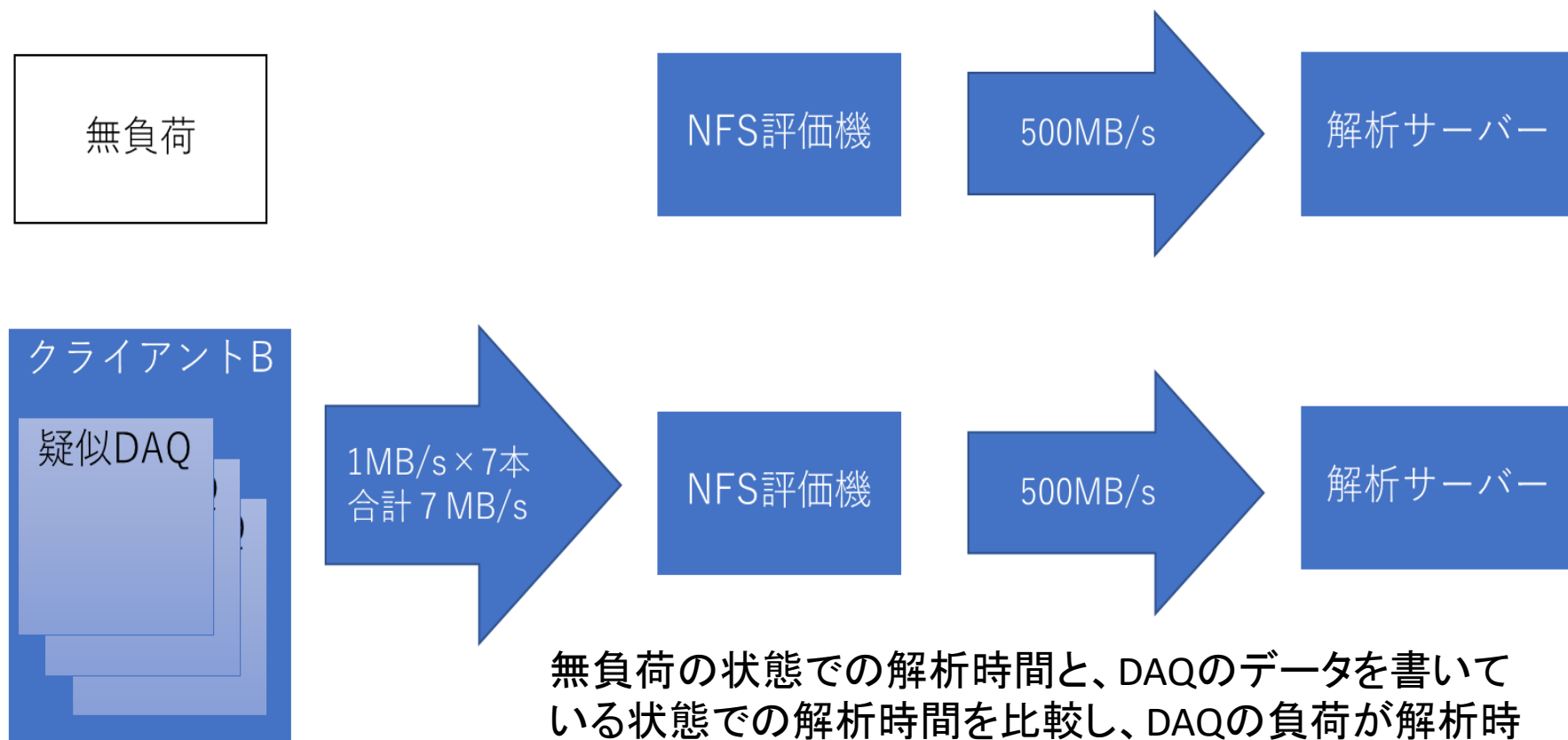
ddコマンドによる、HDDのシーケンシャルRead/Write性能試験

NFSのベンチマーク結果



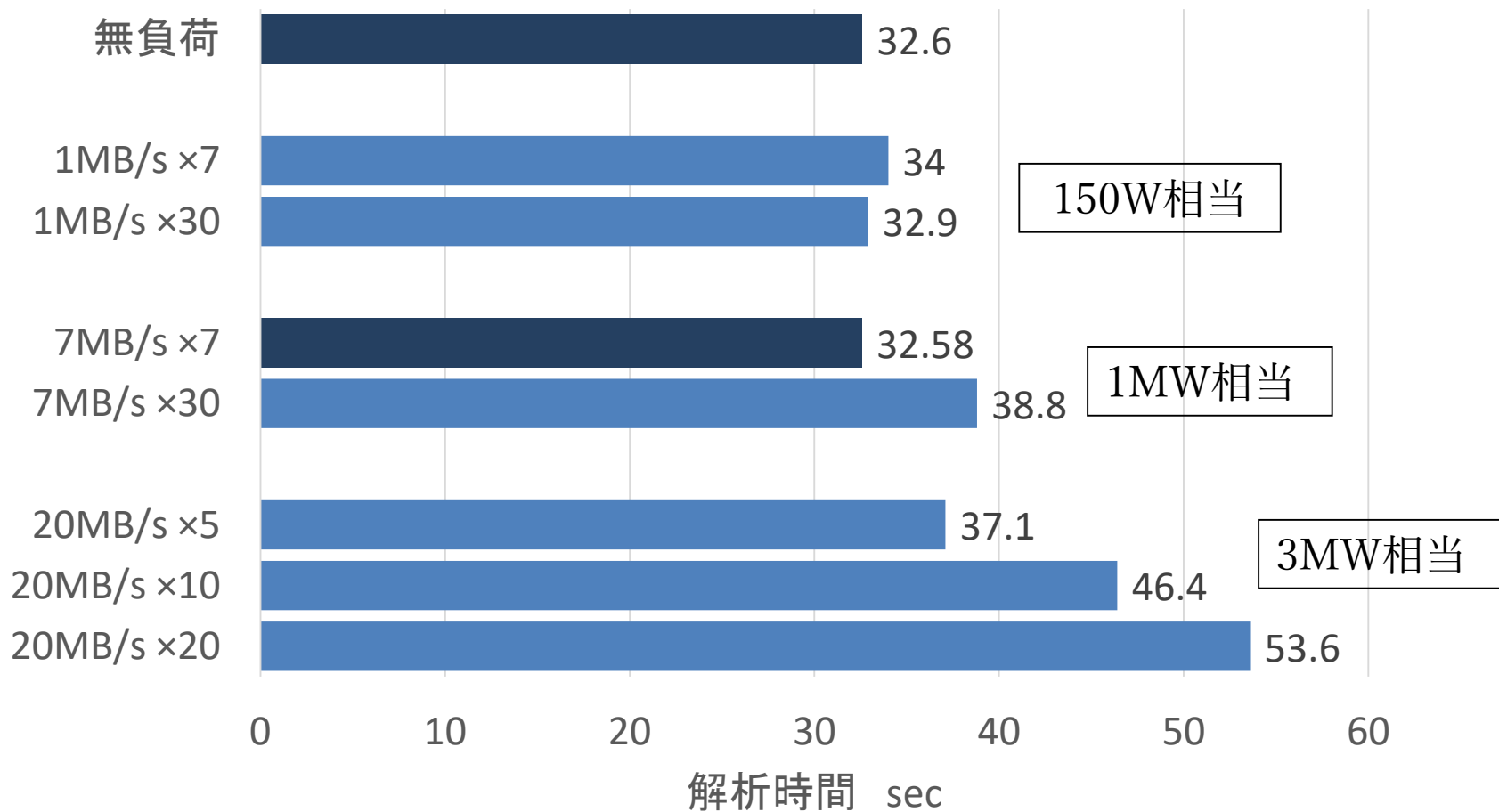
NFSマウントをした状態での、ddコマンドによるベンチマーク

実験環境に合わせたベンチマークの構成



無負荷の状態での解析時間と、DAQのデータを書いている状態での解析時間を比較し、DAQの負荷が解析時間にどの程度影響を与えるか比較した。

ベンチマーク結果

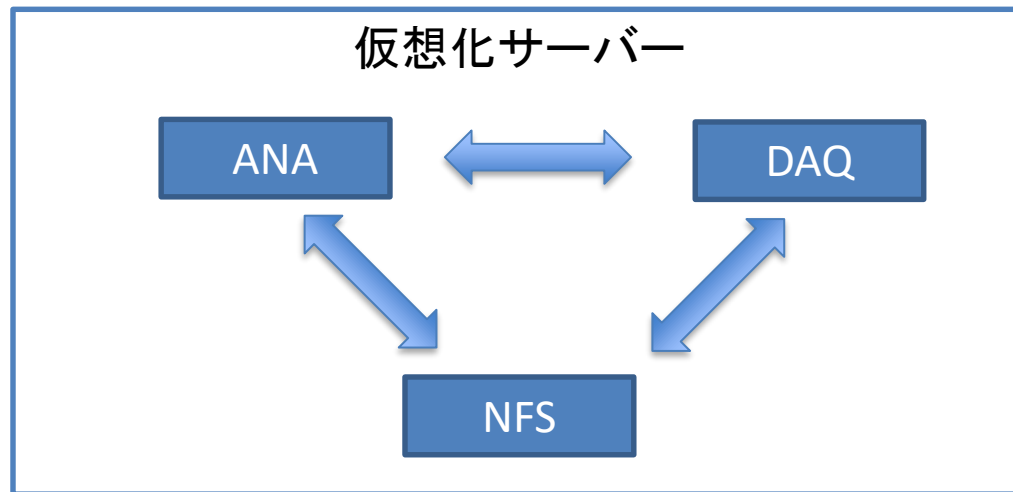
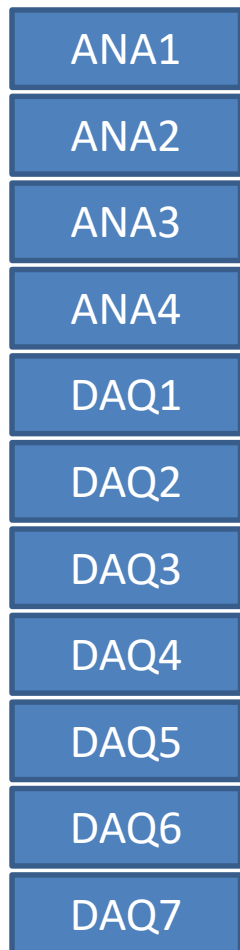


まとめ

1. 安定的な実験の遂行のために、Zabbix iDRAC iLO等を使った、DAQ用サーバーの監視体制を構築
2. Zabbix監視結果をもとに、新しいDAQサーバーの構成を提案。
3. J-PARC MLF NOVAにおいて、NFSの更新を実施。
4. 今後、MLF KEKの各ビームラインのDAQ用サーバーの更新を進めていく。

次に考えていること

DAQ,解析用サーバーの仮想化



- サーバーの管理が容易
- ストレージ次第では1台のマシンでDAQ可能
- 各仮想サーバーは、非常に高速なネットワークで結ばれる。
- サーバーの更新、不具合発生時は、仮想化イメージを移動するだけで対応可能。