



KEK



IMSS

物質構造科学研究所

KENS MLFのDAQシステムとストレージシステム

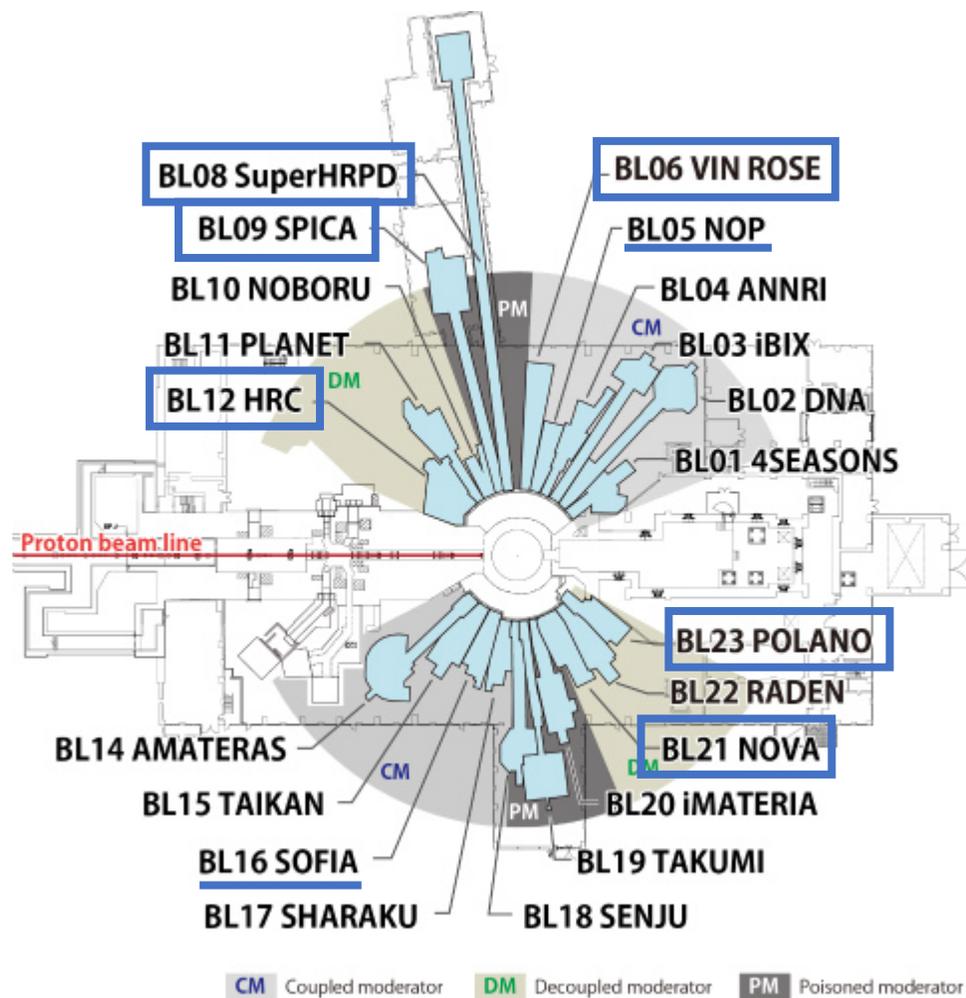
KEK 物質構造科学研究所 中性子科学研究系
瀬谷智洋

2022年11月17日

はじめに

- KENS (KEK Neutron Science Division) ではJ-PARC MLFで8本のビームラインを運用している。この内**DAQ-Middleware**を使用して**データ収集を行っている6本のビームライン**について、KENS-DAQグループにて実験用計算機システムの構築と運用を行っている。
- システムの構築と運用の効率化、導入コストの削減のため2017年よりKVM (Kernel-based Virtual Machine)による仮想化に取り組んでおり、現在では全てのビームラインにて**仮想化環境によるデータ収集**を行っている。これらは統一された構成となっており、OSやDAQ-MWのバージョンの他、ローカルネットワークの構成ルールなど、**様々な点で共通化**が図られている。また、**5年に1回ハードウェアのリプレイス**を実施しており、新OSへの対応などもその際に順次実施している。
- 2017年**以前は野放し**の状態ですerverが運用されており、ストレージ残量が無いのに実験を行いデータを取りこぼしたり、故障に気が付かずに実験を進めてしまうような状態だったが、**統合監視システムを導入**し、serverの健全性や**リソースの状態を常に監視する**体制を整えた。
- 共通化はソフトウェアに限らず、ハードウェアでも行われており、搭載しているメモリーモジュールやストレージなども**極力同一のもの**を選定するようにしている。
- **安定的な実験遂行のため、実験用計算機環境とストレージシステムシステムの整備を行った。**

MLF：物質生命科学研究所



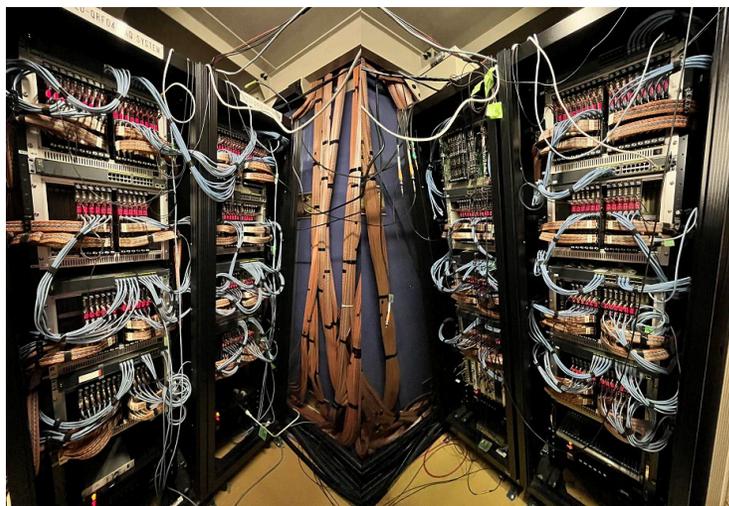
全21ビームライン中、
8ビームラインをKEKが運用している。

KEK ビームラインのDAQシステム



- NeuNETはVMEのPSD(Position Sensitive Detector)用に開発されたリードアウトモジュール。
- **SiTCP**でデータを読み出す。
- 1ボードあたり40MSPSのADCを16chを搭載し、**1枚で8本**のPSDの信号を処理する。
- MLFのおよそ8割の実験装置で使われる。

BL09 真空散乱槽

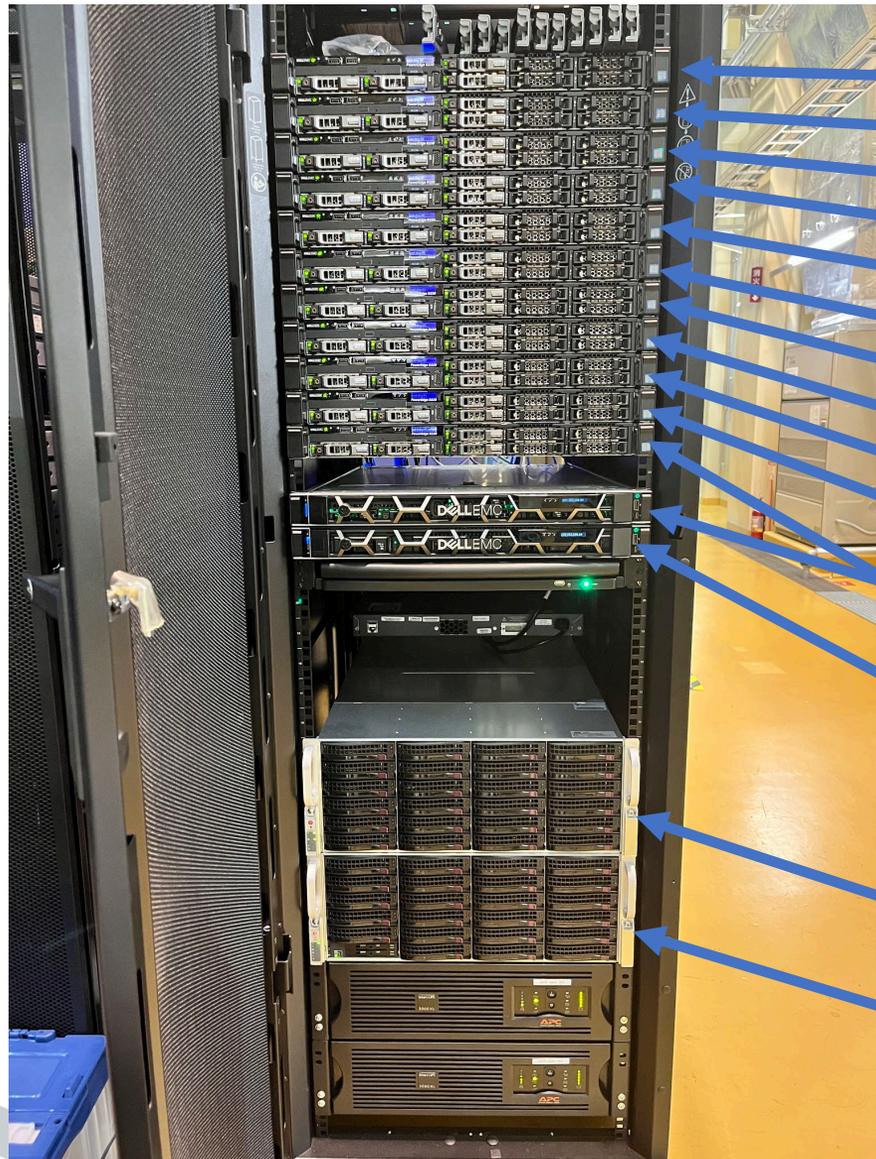


BL09のNeuNET

	PSD	NeuNET	ADC ch	DAQサーバー
BL06	-	-	-	2
BL08	704	88	1408	10
BL09	1592	199	3184	10
BL12	384	48	768	2
BL21	976	122	1952	7
BL23	352	44	704	3
合計	4008	501	8016	34

- 全体で**4008本**のPSD、**8016ch**の信号を処理する。
- 501枚のNeuNETボードと**34台のDAQサーバー**でDAQしている。
- DAQソフトは**DAQ-MW**で、MLF用のコンポーネントを作成してデータ収集している。

仮想化以前のDAQシステム



- daq-bs-l
- daq-bs-r
- daq-ha
- daq-qa-l
- daq-qa-r
- daq-la-l
- daq-la-r
- daq-sa
- daq-gem1
- daq-gem2
- edis-01vm
- iroha-01vm
- iroha-04vm

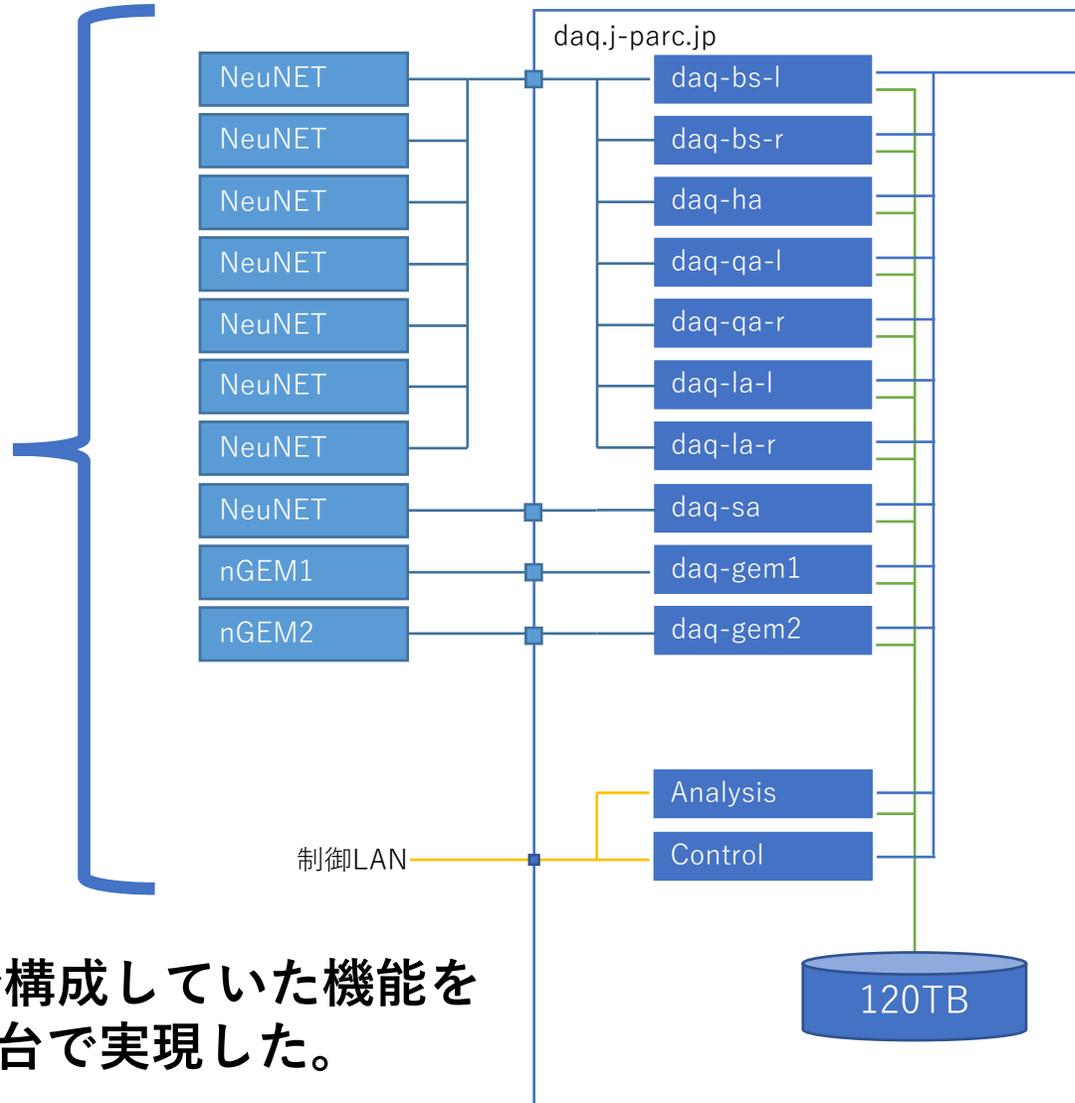
DAQ-MW

解析
実験制御

- nfs-01
240TB
 - JBOD
320TB
- ↑ min SAS

ストレージ

仮想化されたDAQシステム



- 仮想化によりOS、DAQ-MWのインストール作業の**大幅な作業量の低減**。
(KVMイメージをクローンすれば良い)
- **リプレイス作業の簡素化**
- バックアップも**KVMイメージをまるごとバックアップ**すれば良い。
- **6台の実験装置で共通の構成。**
- **サーバーの導入/管理コスト削減**
- ストレージ容量は減ったが、それでも1~2年分の容量があり、バックアップシステムと合わせて問題ない構成にした。
- 極力MLFに計算機を配置しない構成となり、**ハードウェアの故障やリプレイスの際の作業が簡素化**できる。

今まで15台の計算機で構成していた機能を
KVMの仮想化により1台で実現した。

サーバーの削減効果

BL08



BL08では19インチラックも撤去したためデスクトップタイプのT640を選定した。

BL21



上段:Dell PowerEdge R7515
下段:Dell PowerEdge R540

BL23



BL09



Dell PowerEdge R7515

	2017年以前	2022年
BL08	14	1
BL09	31	1
BL12	8	1
BL21	24	1
BL23	9	1
合計	86	5

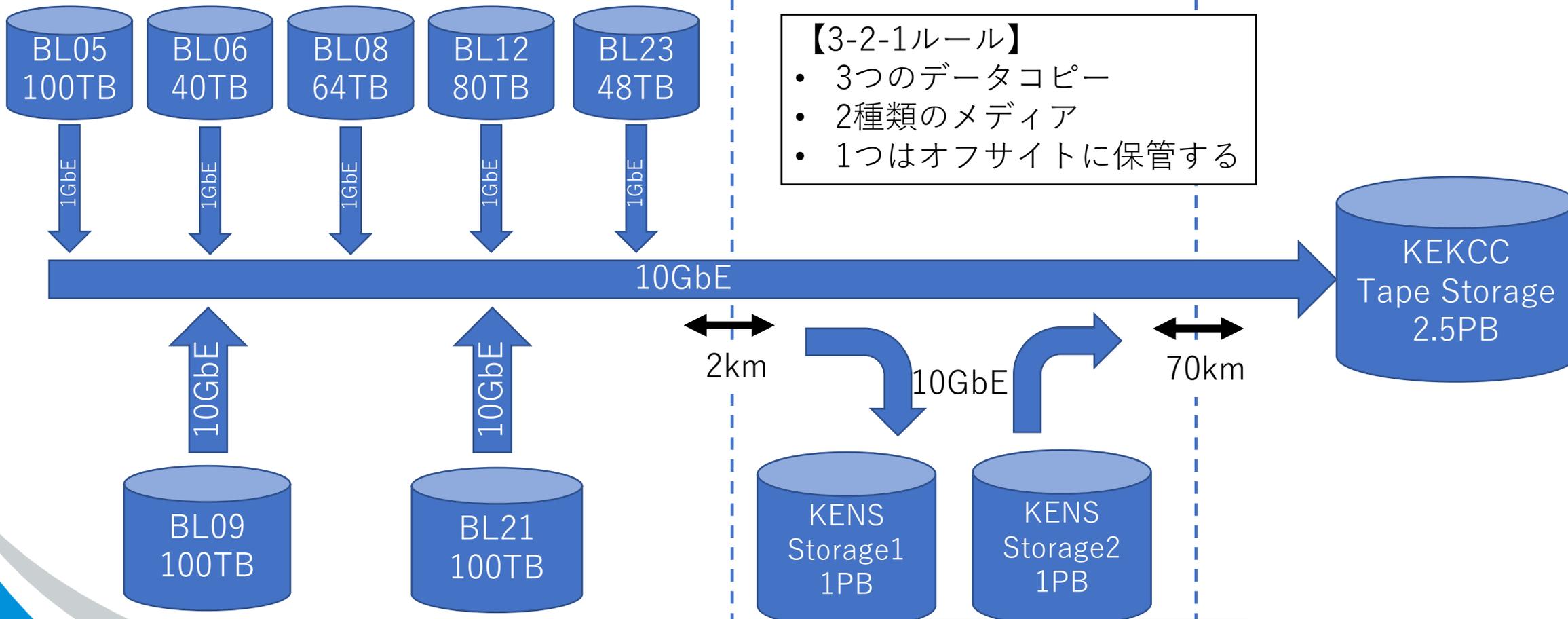
- KVMによる仮想化を実施する前は、ビームラインごとに10台~20台程度の物理サーバーを用意し実験環境を整備していたが、仮想化を導入してからは、大幅な物理サーバーの削減が実施できた。
- **導入コストは93%削減された。**

KENSのストレージシステム

MLF

KEK東海

KEKつくば

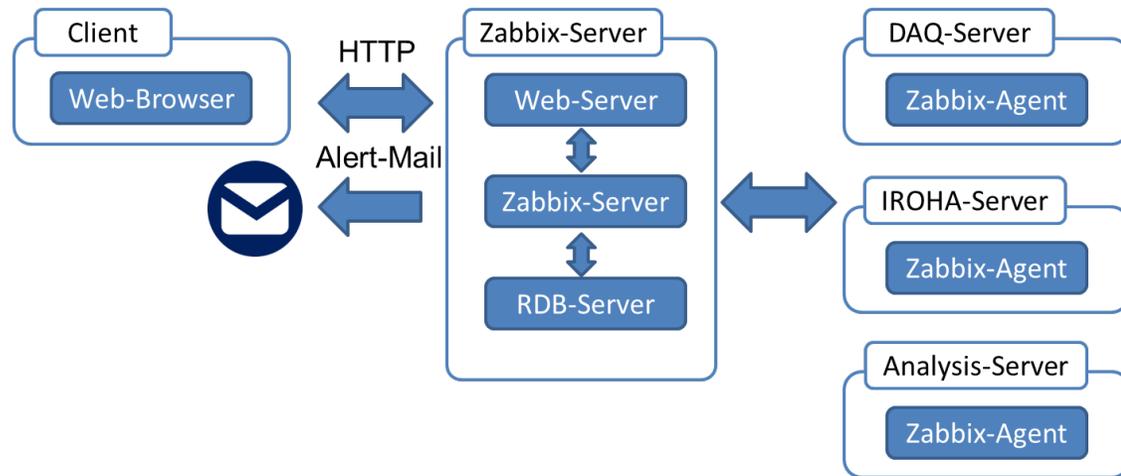


- 【3-2-1ルール】
- 3つのデータコピー
 - 2種類のメディア
 - 1つはオフサイトに保管する

毎週水曜日に自動的にバックアップが実行される。

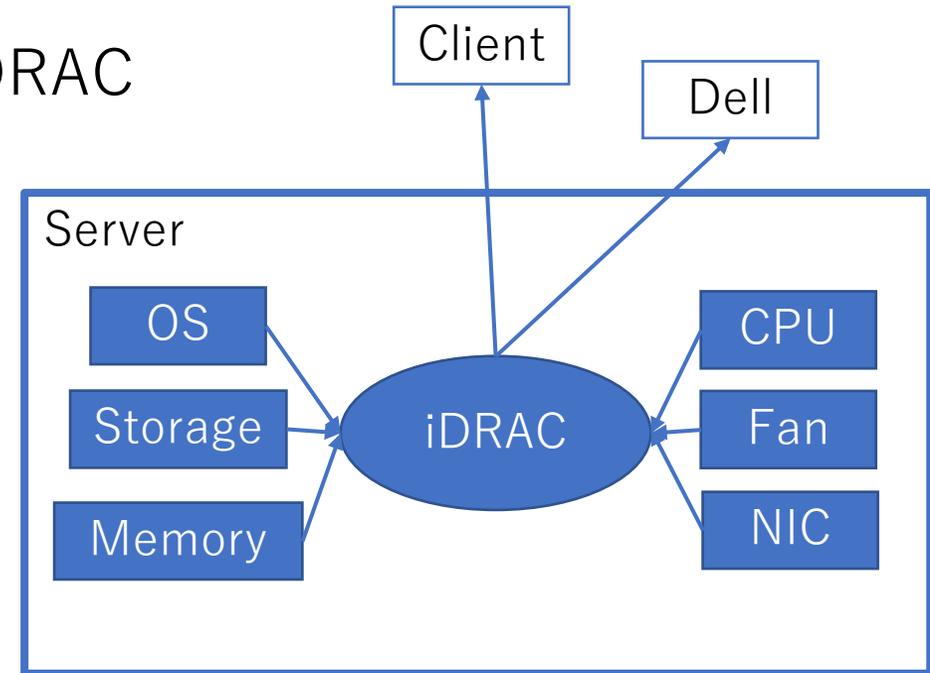
今後の開発監視システム

Zabbix



- Zabbixは、エージェントソフトをサーバーにインストールしてつかう**監視ソフト**。
 - 様々なOSに対応し、ネットワーク機器やUPSにも対応する。
 - サーバーのリソースを**常に監視**し、予め設定された閾値を超えると、アラートメールを送信する。
- >**不具合を見逃さない!**

iDRAC



- iDRAC(integrated Dell Remote Access Controller)は、マザーボード上に実装された、**サーバー管理**ツール。
- **OSに依存せずに独立**に動作する。
- サーバーの**健全性**を確認し、異常があればアラートメールを送信する。
- オプションで**リモート操作**が可能で、OSのインストール作業から完全リモート操作ができる。

今後の開発

●DAQ-MWのバージョンアップ

- RHEL8に対応した**DAQ-MW2.0.0**（プレリリース）のテスト

- DAQ-MW **MLF Component1.6.0**のテスト

- DAQ-MW開発当時は1台の計算機あたりの**リソースが少なかった**ので、並列分散処理はとても良いアイデアだった。しかし現代では1台あたりの計算機で64コア256GBメモリを搭載するようなマシンも入手できる環境である。であれば**マルチスレッド、マルチプロセス化**したDAQソフトでDAQできるのではないか？

- DAQ-MW MLF ComponentのGathererは**CORBA通信のオーバーヘッド**で本来の性能を引き出すことができていない。1回のコンポーネント間通信でやり取りされるデータが小さいため。1.6.0では**バッファリング機能**を追加して**オーバーヘッドを削減**した。もしかしたら1.6.0で1台のDAQ-CPUでDAQできるかも？

- 使い勝手が悪くとも、**シンプルで自分たちで責任で運用できる**DAQソフトの開発が必要。

●NeuNETの更新

- ハードウェアはSpartan-3EからArtix-7に載せ替えた。合わせてADCも1チップで8chのものに載せ替えた。

- HDLコードの刷新が必要。

DAQシステムの式年遷宮

式年遷宮：伊勢神宮の社殿を20年に1度建て替えること

- 現在のDAQシステムは2008年に開発された。
- DAQ-MW NeuNET共に開発者の退職もしくは退職が近い。
- 式年遷宮のように20年周期で総入れ替えというのは、継続的な開発という意味ではとても良い仕組み。
- 折しも、RHEL8のサポート期限はMLFファーストチームからちょうど20年。