

SINET6の活用事例紹介②

— LHCアトラス実験 —

澤田 龍・田中 純一・齊藤 真彦・松井 長隆

東京大学 素粒子物理国際研究センター

2024年6月11日



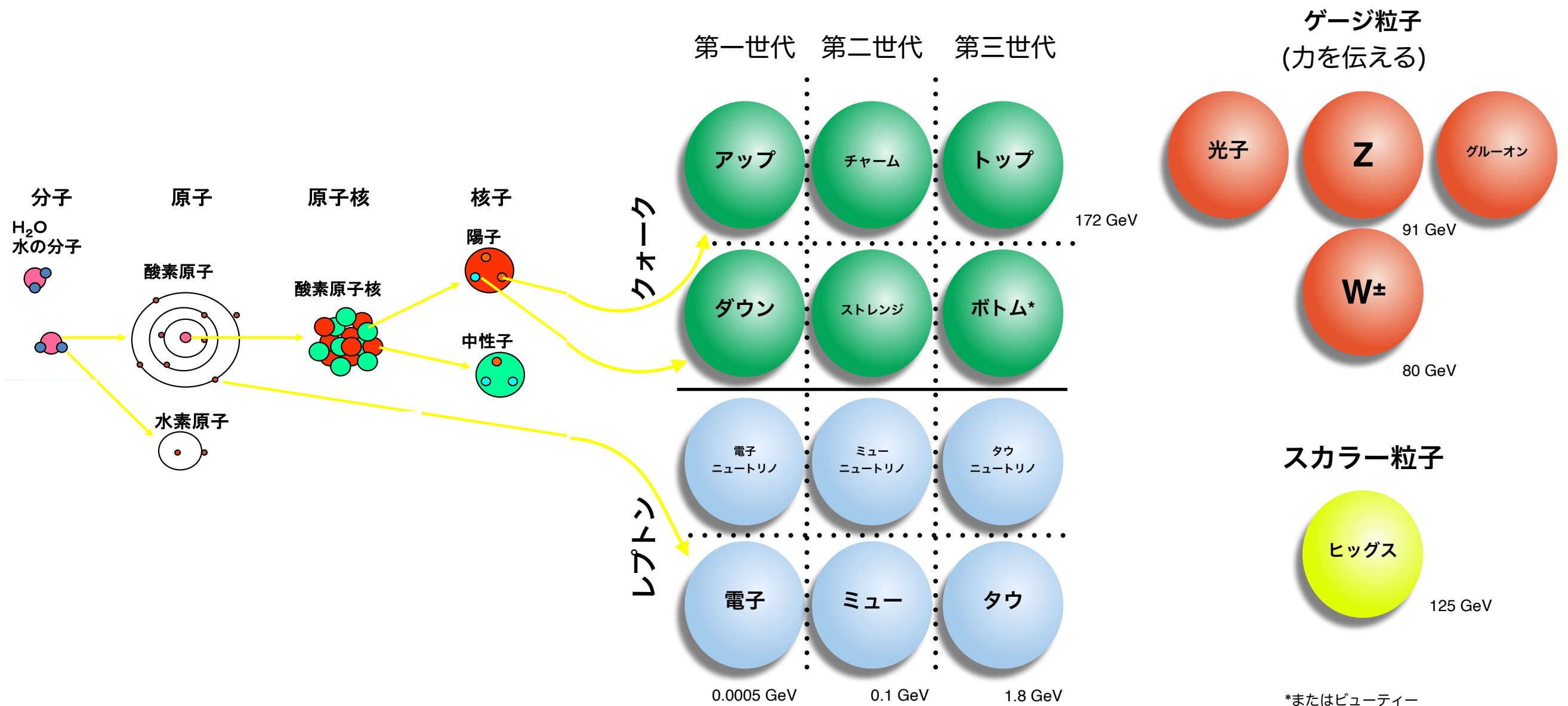
東京大学
素粒子物理国際研究センター
International Center for Elementary Particle Physics
The University of Tokyo



ATLAS
EXPERIMENT

素粒子物理

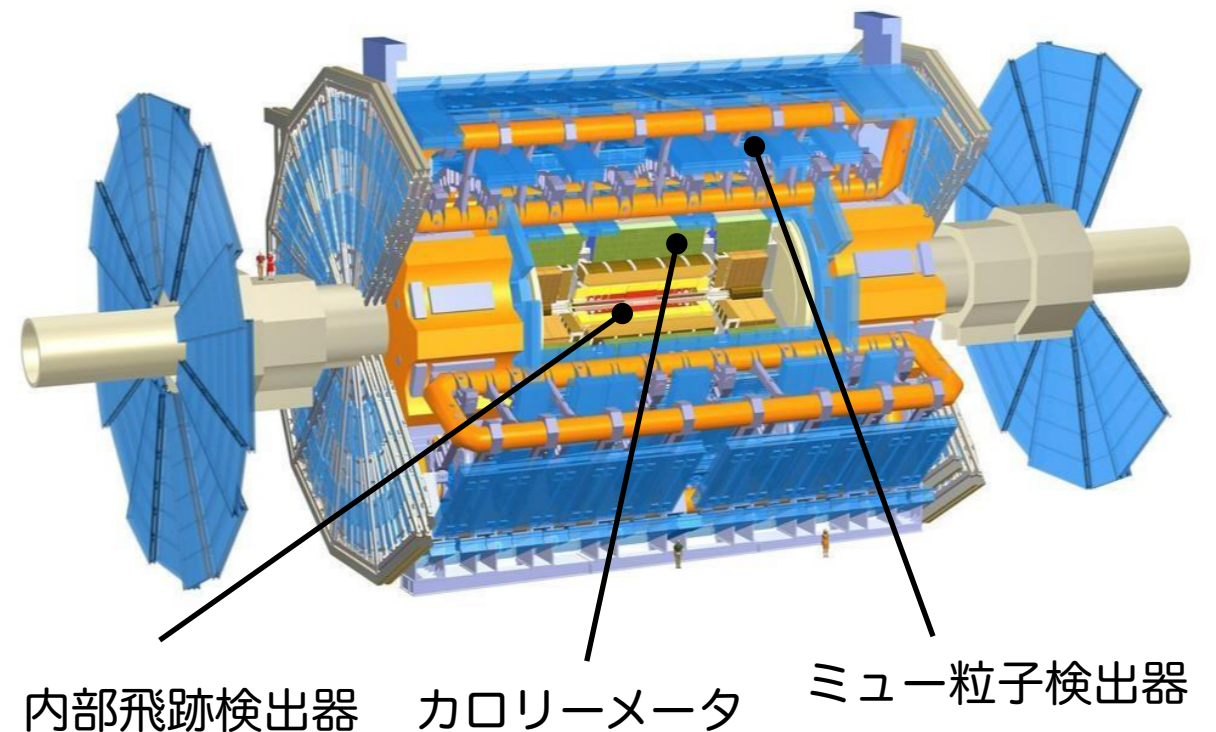
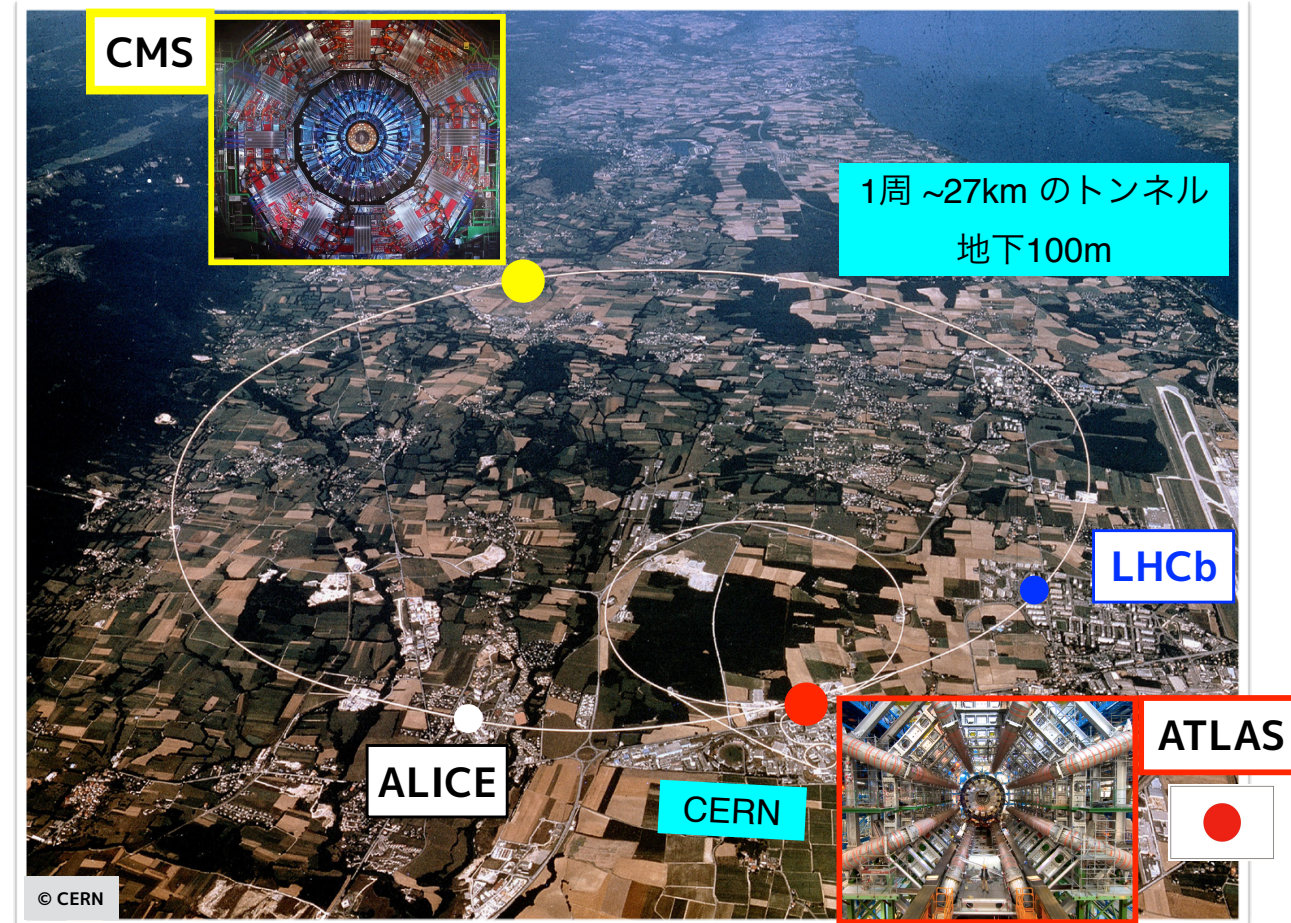
素粒子



- 電子の発見からはじまり、理論と実験の両輪で「標準理論」を作り上げてきた
- 2012年のHiggs粒子の発見によってひとまず「完成」

LHC ATLAS実験

- ・ スイスとフランスをまたぐ実験施設 (Large Hadron Collider)
- ・ 世界最高エネルギー (現在13.6 TeV)での陽子ビーム同士の衝突
- ・ ATLAS検出機:
 - ・ 直径25m, 長さ46m, 7000トン
- ・ センサーの読み出しチャンネル数: ~ 1億
- ・ 約40カ国から3000人の科学者が参加、日本からも約150人
- ・ 素粒子物理研究
 - ・ ヒッグス粒子の発見 (2012年)
 - ・ 精密測定、稀な崩壊の探索
 - ・ 標準模型研究:
 - ・ トップ粒子、電弱ボソン、フレーバー...
 - ・ 新粒子探索:
 - ・ 超対称性粒子、暗黒物質候補、重いヒッグス粒子...



ヒッグス粒子

粒子に質量を与えるヒッグス場

- ・ もしヒッグス場が0なら全ての粒子は質量0で光の速さになる

2012年にLHC (ATLASとCMS実験)で発見



宇宙のはじまりと未来を知るためのヒッグス研究

- ・ なぜ宇宙は物質で満たされているのか？
 - ・ なぜ反物質と対消滅せずに残っているのか
- ・ 将来にわたって今のような時空間 (真空) は安定なのか？

→これらの問いに答えるために、さらなるヒッグス研究が必要

ヒッグス粒子

粒子に質量を与えるヒッグス場

- ・ もしヒッグス場が0なら全ての粒子は質量0で光の速さになる

2012年にLHC (ATLASとCMS実験)で発見



宇宙のはじまりと未来を知るためのヒッグス研究

- ・ なぜ宇宙は物質で満たされているのか？
 - ・ なぜ反物質と対消滅せずに残っているのか
- ・ 将来にわたって今のような時空間 (真空) は安定なのか？

→これらの問いに答えるために、さらなるヒッグス研究が必要

ヒッグス粒子

粒子に質量を与えるヒッグス場

- ・ もしヒッグス場が0なら全ての粒子は質量0で光の速さになる

2012年にLHC (ATLASとCMS実験)で発見



宇宙のはじまりと未来を知るためのヒッグス研究

- ・ なぜ宇宙は物質で満たされているのか？
 - ・ なぜ反物質と対消滅せずに残っているのか
- ・ 将来にわたって今のような時空間 (真空) は安定なのか？

→これらの問いに答えるために、さらなるヒッグス研究が必要

ヒッグス粒子

粒子に質量を与えるヒッグス場

- ・ もしヒッグス場が0なら全ての粒子は質量0で光の速さになる

2012年にLHC (ATLASとCMS実験)で発見



宇宙のはじまりと未来を知るためのヒッグス研究

- ・ なぜ宇宙は物質で満たされているのか？
 - ・ なぜ反物質と対消滅せずに残っているのか
- ・ 将来にわたって今のような時空間 (真空) は安定なのか？

→これらの問いに答えるために、さらなるヒッグス研究が必要

ヒッグス粒子

粒子に質量を与えるヒッグス場

- ・ もしヒッグス場が0なら全ての粒子は質量0で光の速さになる

2012年にLHC (ATLASとCMS実験)で発見

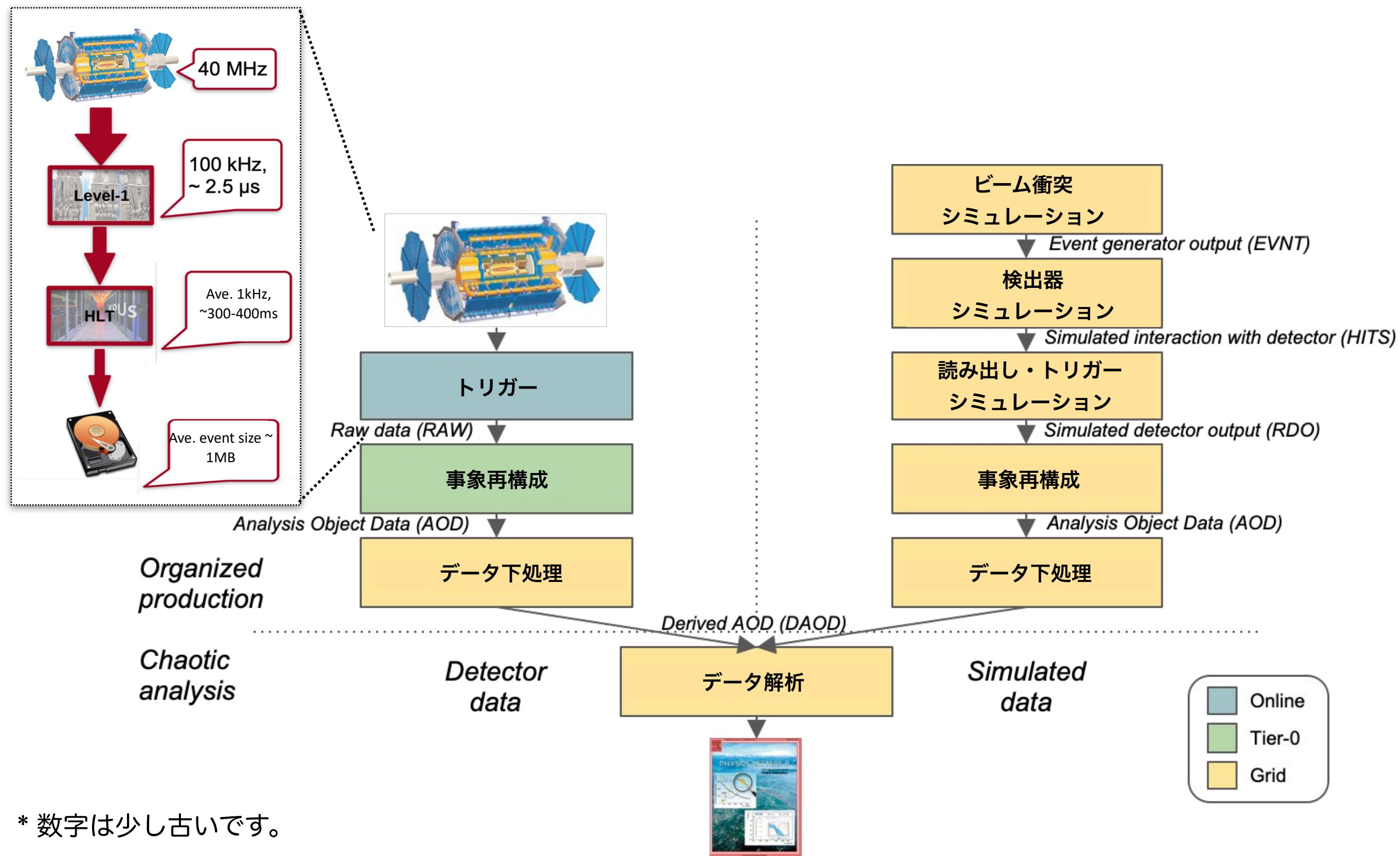


宇宙のはじまりと未来を知るためのヒッグス研究

- ・ なぜ宇宙は物質で満たされているのか？
 - ・ なぜ反物質と対消滅せずに残っているのか
- ・ 将来にわたって今のような時空間 (真空) は安定なのか？

→これらの問いに答えるために、さらなるヒッグス研究が必要

ATLAS実験データの流れ



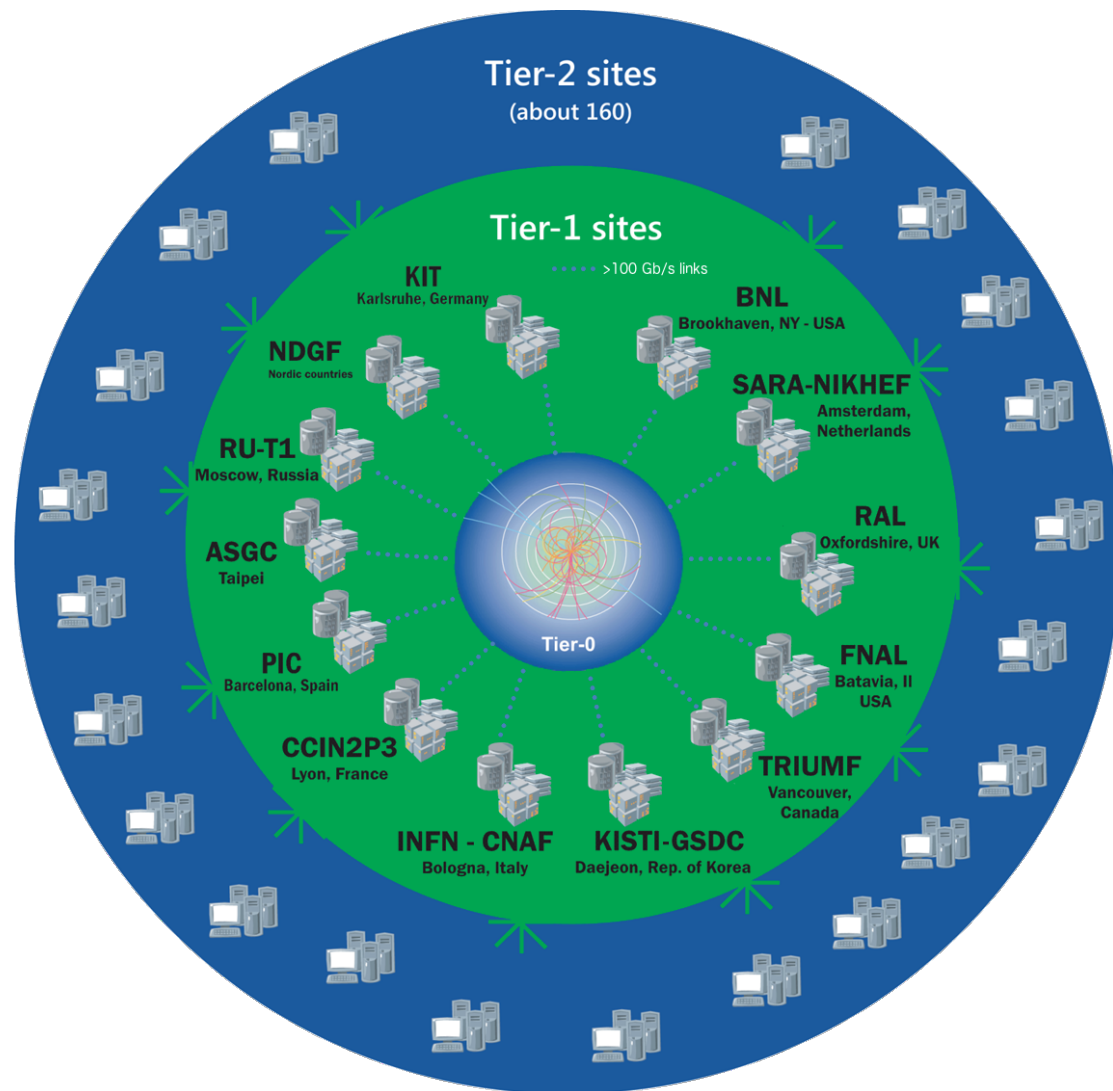
* 数字は少し古いです。

LHC実験と計算機

- LHCでは $O(100\text{PB})/\text{年}$ のデータが生成される
 - ・ 実験データ、シミュレーションデータ、中間ファイル、最終的に解析に使用されるファイル
 - ・ 一つのデータセンターでは扱えない量なので、世界各国で共同して管理
 - ・ WLCG という枠組みで連携
- 実験期間やビーム強度の増強によって増えるデータ量に応じた計算資源 (CPU)が必要になる。
 - ・ データは一度解析をしたら終わりではなく、解析テクニック (アルゴリズム) を改良しつつ、再解析する。

WLCGグリッド

<https://wlcg-public.web.cern.ch/tier-centres>



- 世界各所の計算機システム(サイト)を高速ネットワークで接続。
- 大量のデータを独立した大量のプロセスで処理する High Throughput Computing (HTC) — 多くのサイトでは商用サーバーにLinuxを搭載したもの。
- (ある程度の要求はあるものの) それぞれの計算機システムで使っている計算機、ストレージ、ミドルウェアは様々。
- 最近では、HPCや商用クラウドも利用。

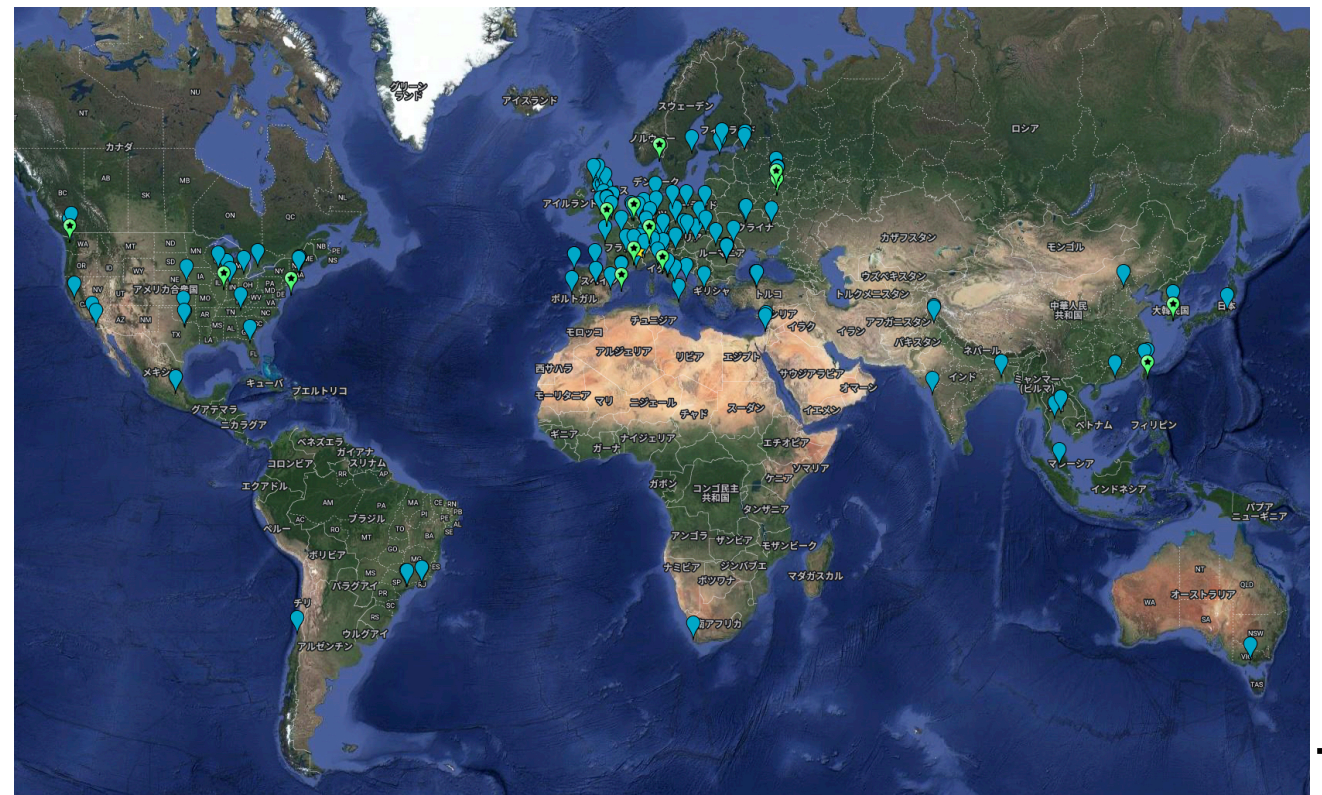
Tier-0: ATLASではCERN

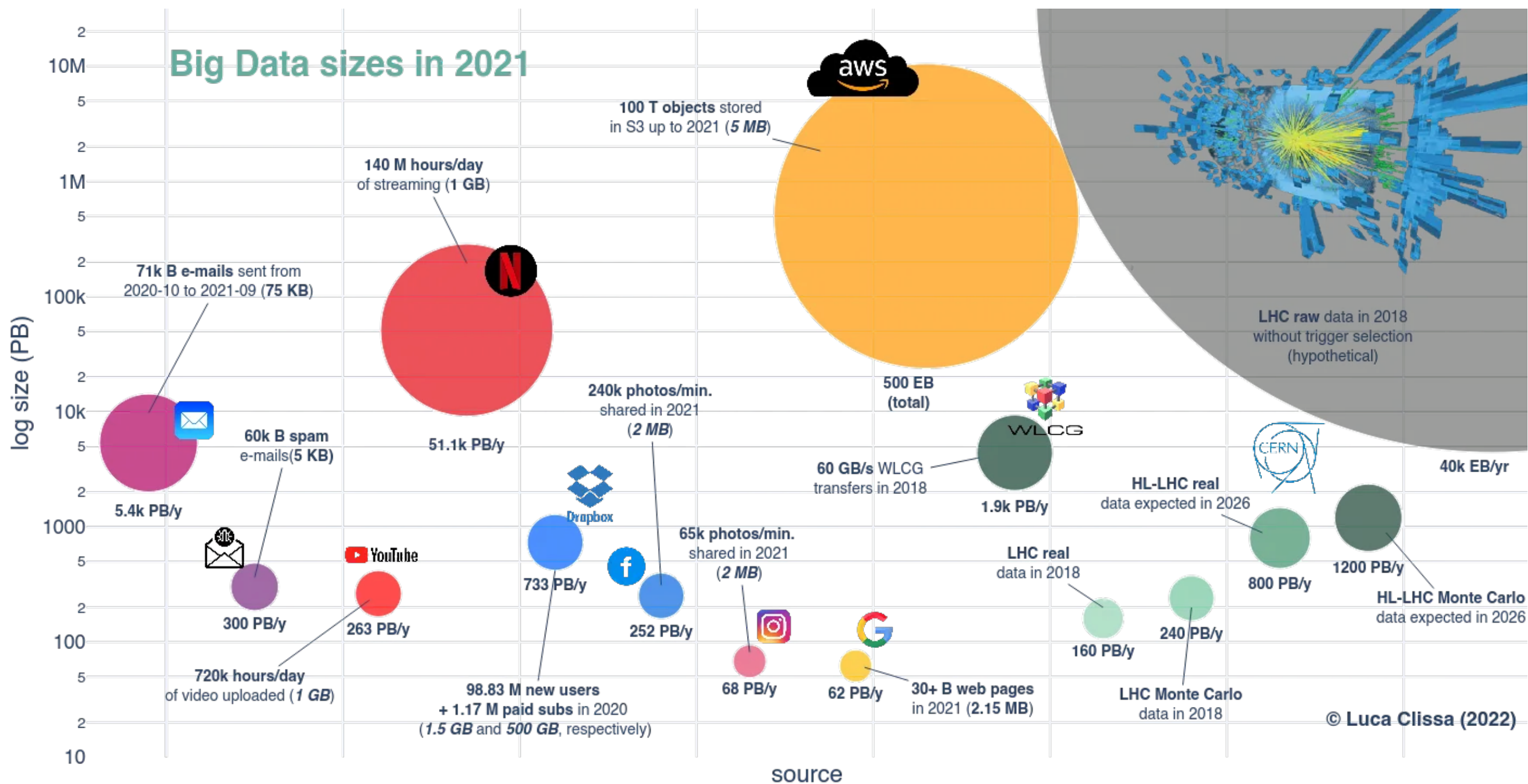
データの保存、
データを他のサイトに配布
実験データのプロセス、シミュレーション...

Tier-1: データの長期保存 (テープ)

実験データのリプロセス、シミュレーション...

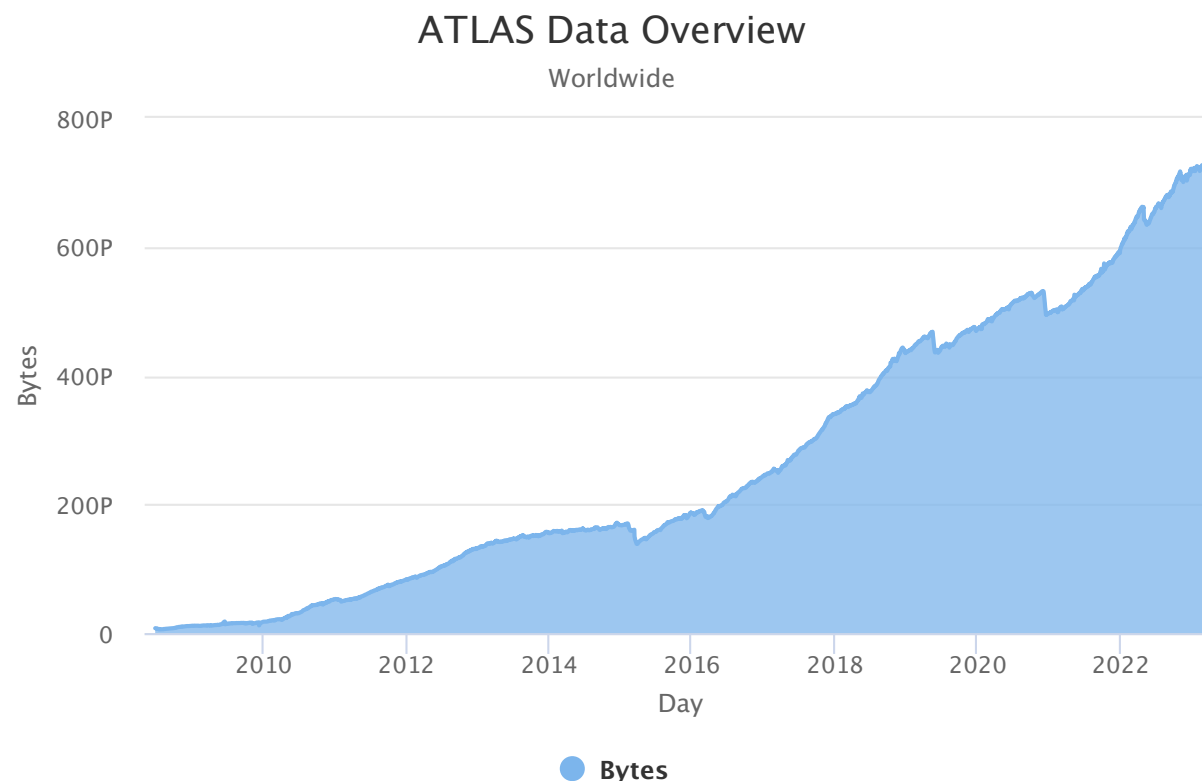
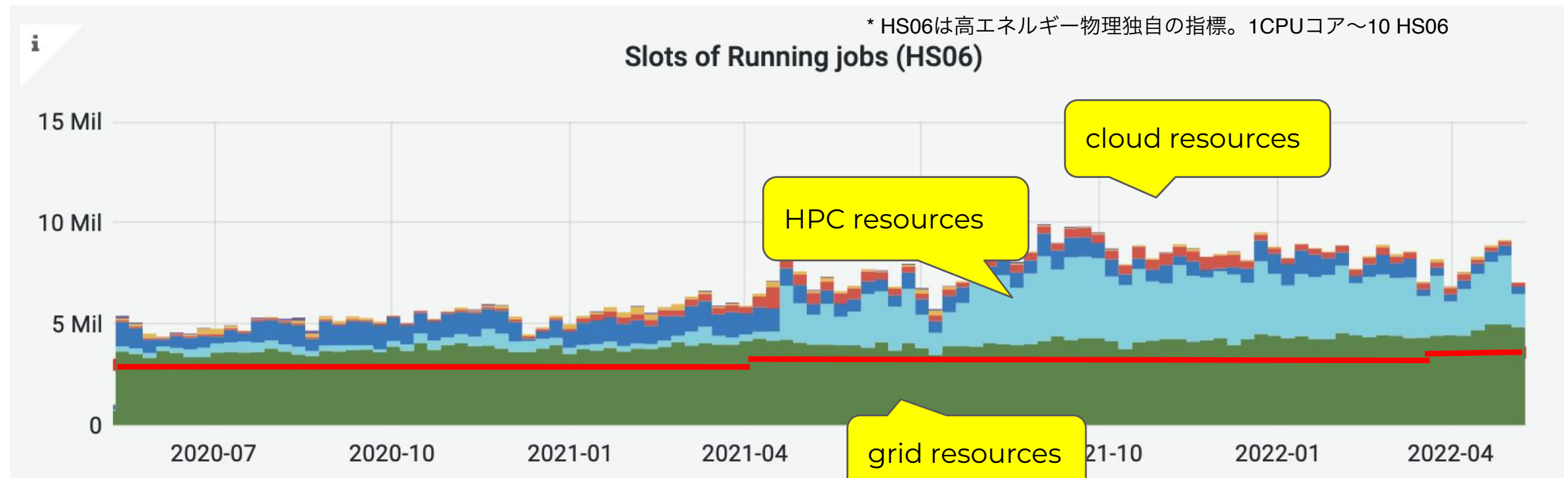
Tier-2: シミュレーション、ユーザー解析...





<https://datapane.com/u/lclissa/reports/dkjK28A/big-data-2021/>

ATLAS実験の計算機資源



- [Gridサイト](#)
(大学や研究所のオンプレミス) が安定して”プレッジ”以上のリソース提供。
- [HPC](#) (米国、欧州)も強力
- [Cloud](#)は主にトリガー用計算機の利用
- [ATLAS@home](#) (ボランティア)
- 常時~70万 CPU コアでジョブ実行
(ピークでは100万以上)

ATLAS日本グループの コンピューティング

- LHCアトラス実験用ICEPP地域解析センター計算機システム (本郷)
 - ① Tier-2 (Tokyo-LCG2): WLCGのグリッドサイトの機能
 - ② Tier-3: ATLAS日本グループメンバー用



地域解析センター計算機システム

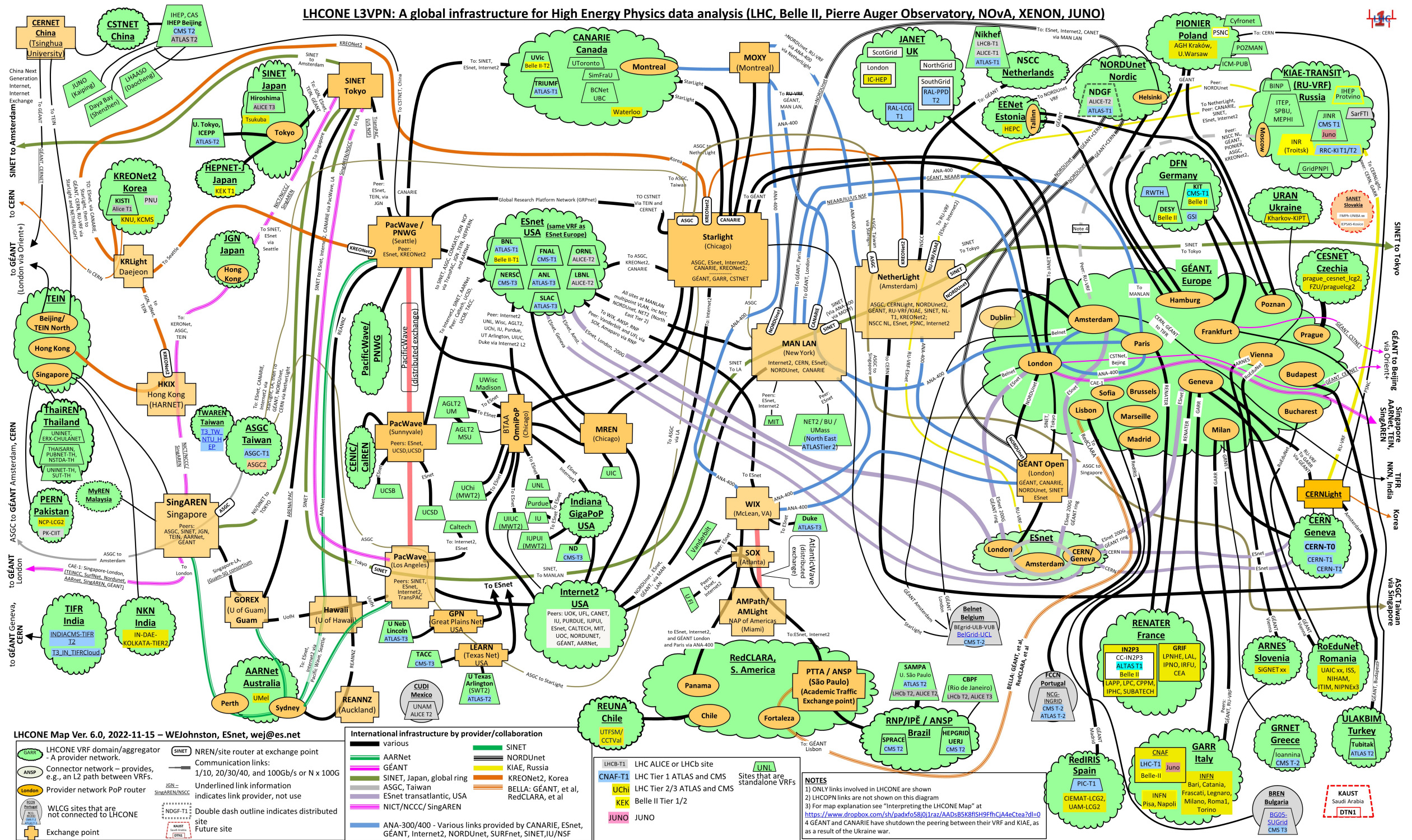
2006年からサービス開始、3年ごとにシステム更新。

主な機器はレンタル。第6期システム (2022-2024) が稼働中。

- 15000 CPU コア (DELL C6520 304台, それぞれXeon Gold 5320 × 2),
- 22 PB ディスクストレージ (Infortrend DS3024 72台)
- SINET6 に接続。100Gbpsに変更
- WLCGが標準としているAlma Linux 9を使用。
- ATLAS実験グループに参加している日本全国の大学で共同利用
- アクティブなユーザー数:
 - ・ 海外: 約1000人
 - ・ 国内: 約120人

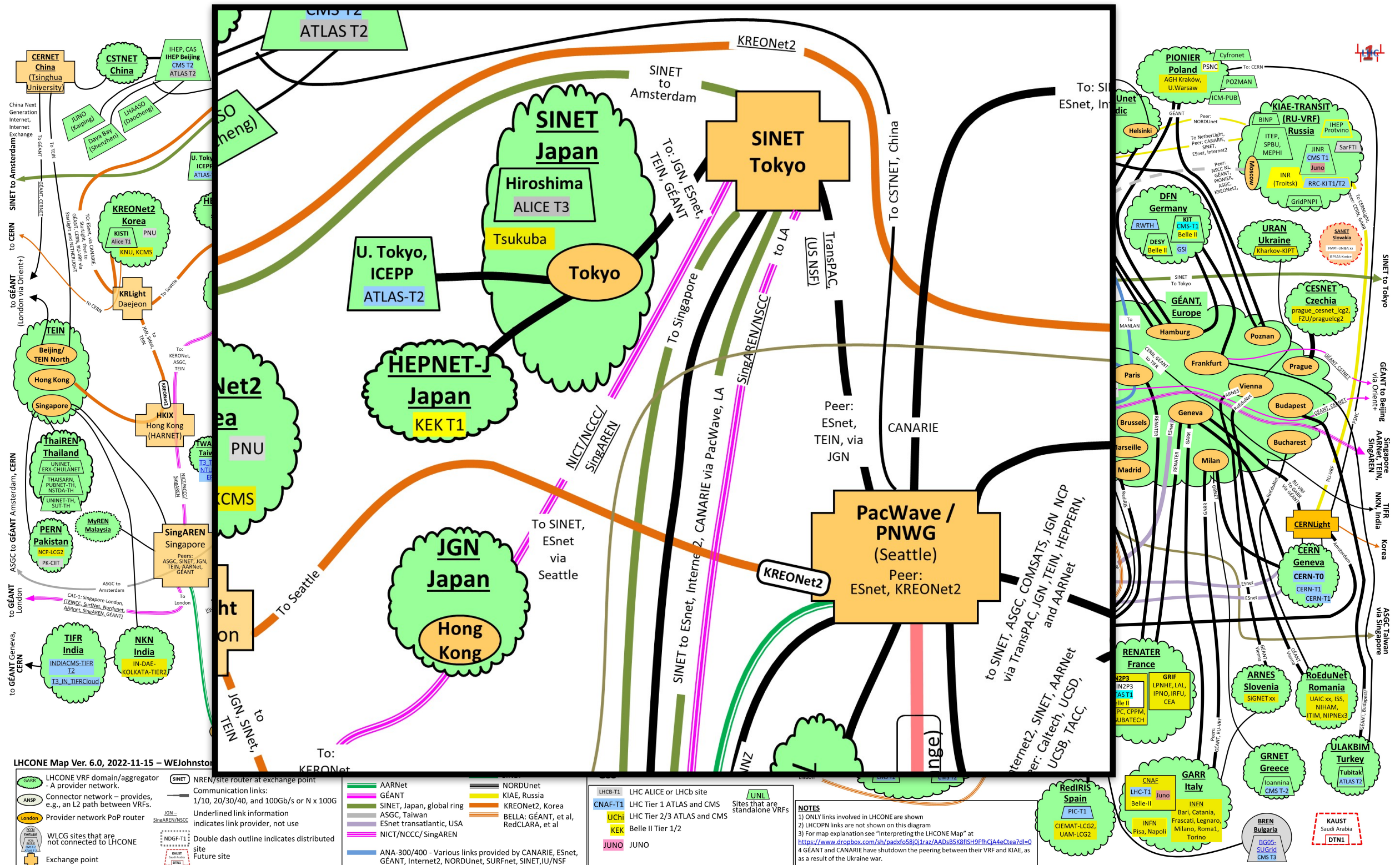
国際ネットワーク: LHCONE

LHCONE L3VPN: A global infrastructure for High Energy Physics data analysis (LHC, Belle II, Pierre Auger Observatory, NOvA, XENON, JUNO)



国際ネットワーク: LHCONE

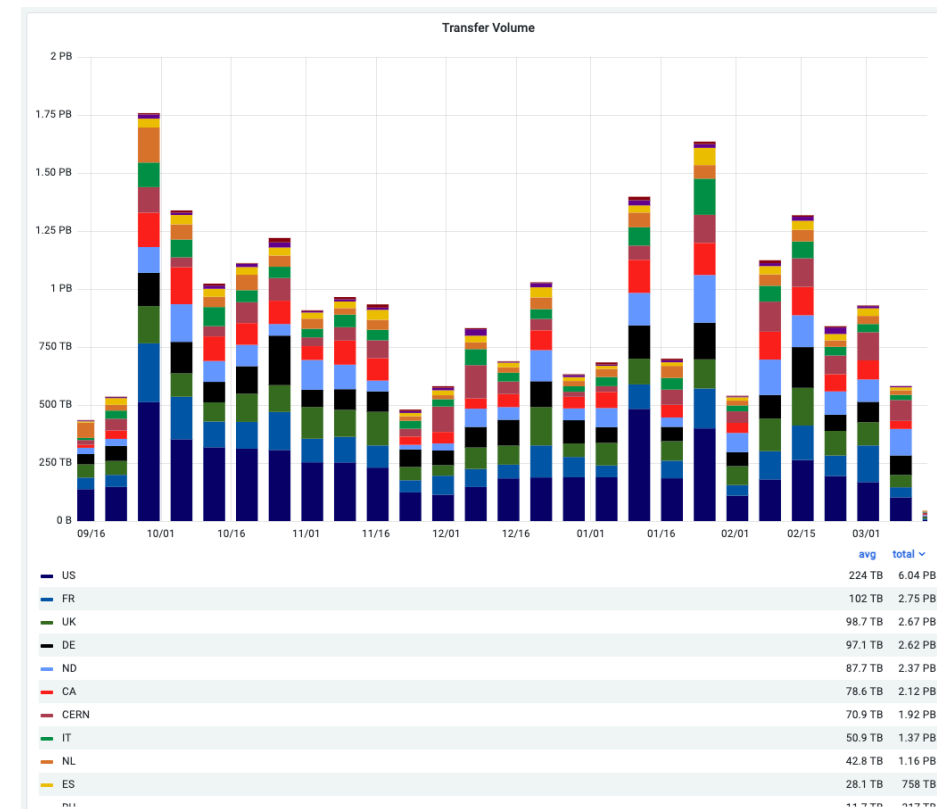
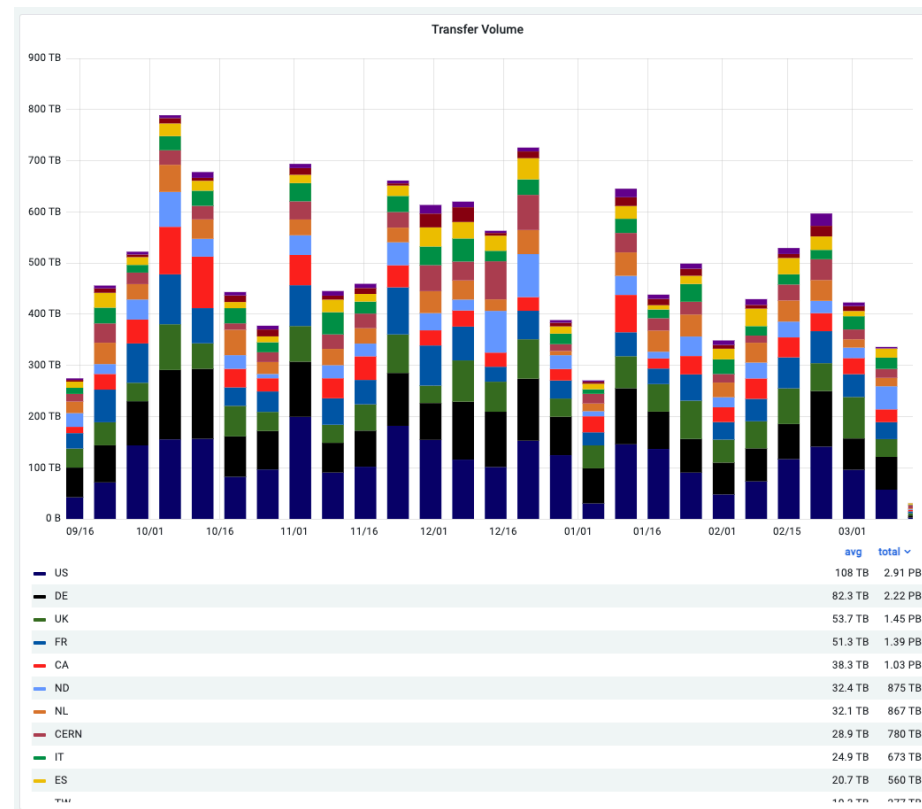
国内では、SINET、東大情報基盤センター、KEKと連携して運用



対外接続

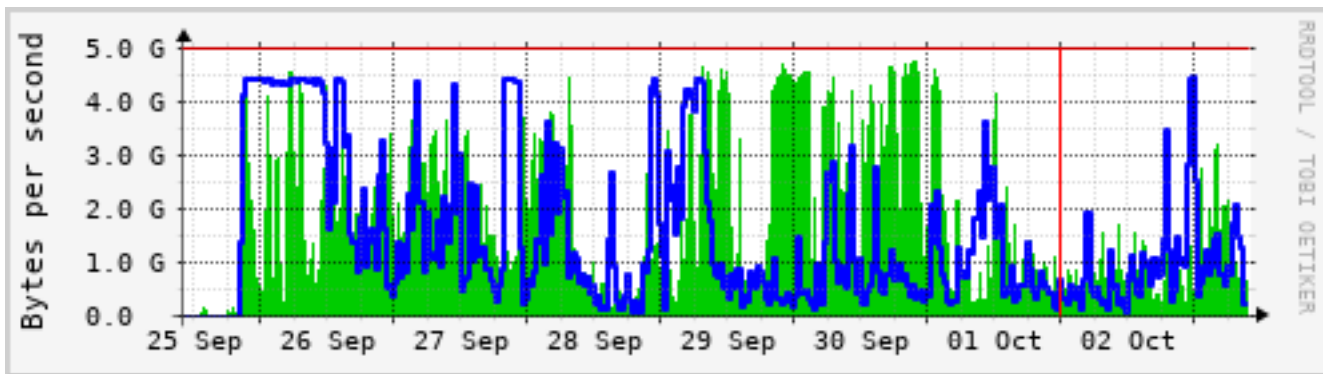
- 25GbE 接続のファイルサーバー 24 台が主要な対外接続ホスト
 - ・ 対外接続 100Gb を活用。
- 将来的には計算ノードが外部ファイルサーバーへアクセスするモデルも考えられる。 海外ではすでに、ディスクを持たないサイトとディスクを持つサイトのグループとしての運用 (データレイク) がなされている。

世界各国と通信

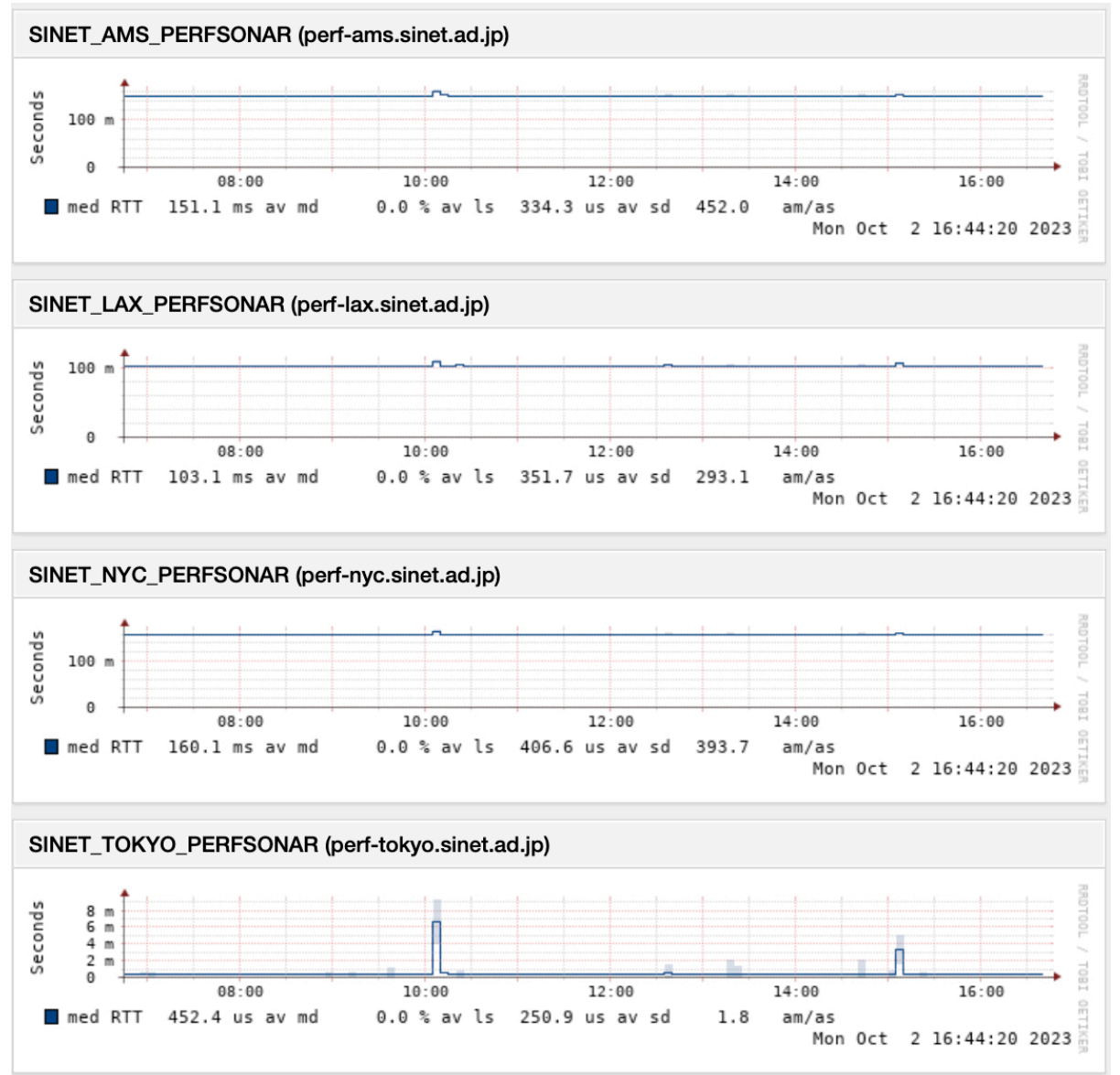


ネットワーク利用

- SINETを通して、海外と通信
 - RTT > 100ms
- 主にwebdav, xrootdのTCPで送受信
- 米国、欧州、アジアと世界各地と通信
 - 相手先が一つではないので、特定の相手向けにチューニングができない。パケットロスが通信速度に大きく影響。



RTT on IPv6

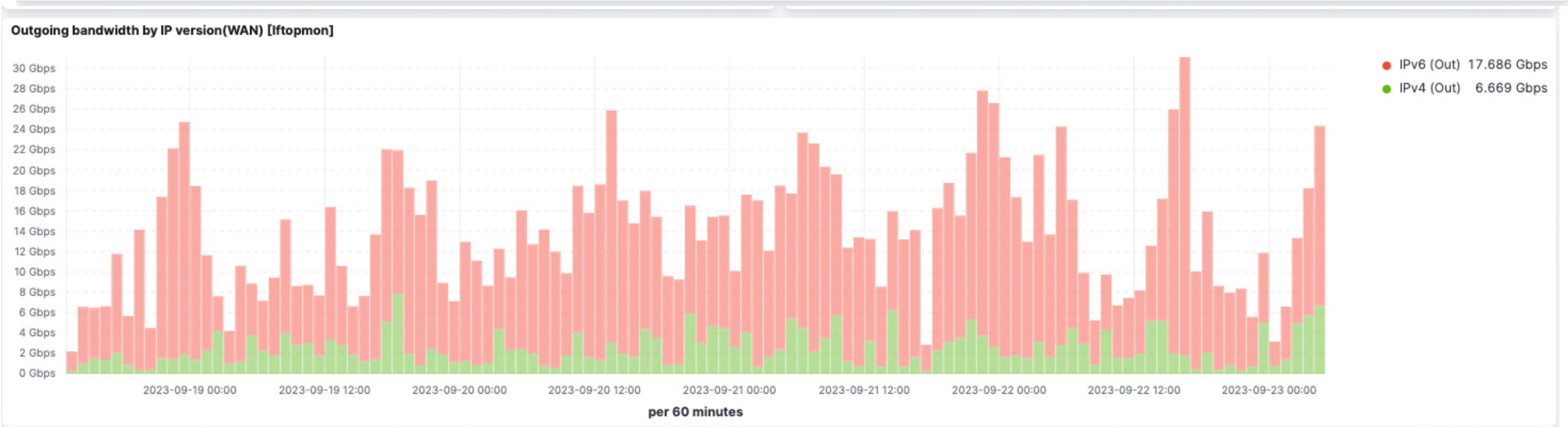
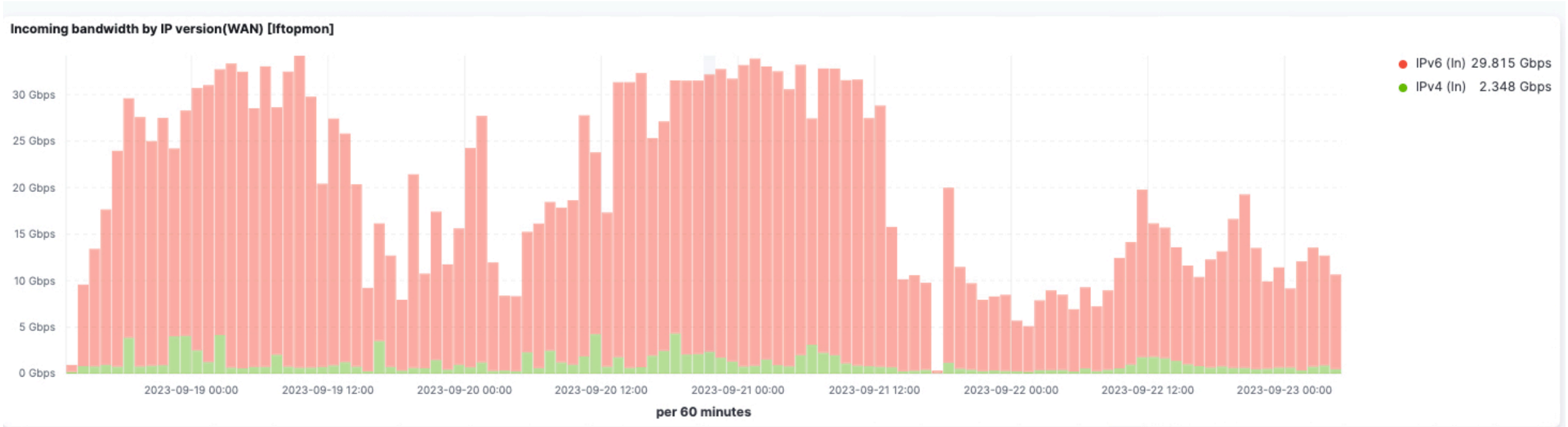


WLCGとSINETのperfsonarを活用

IP version

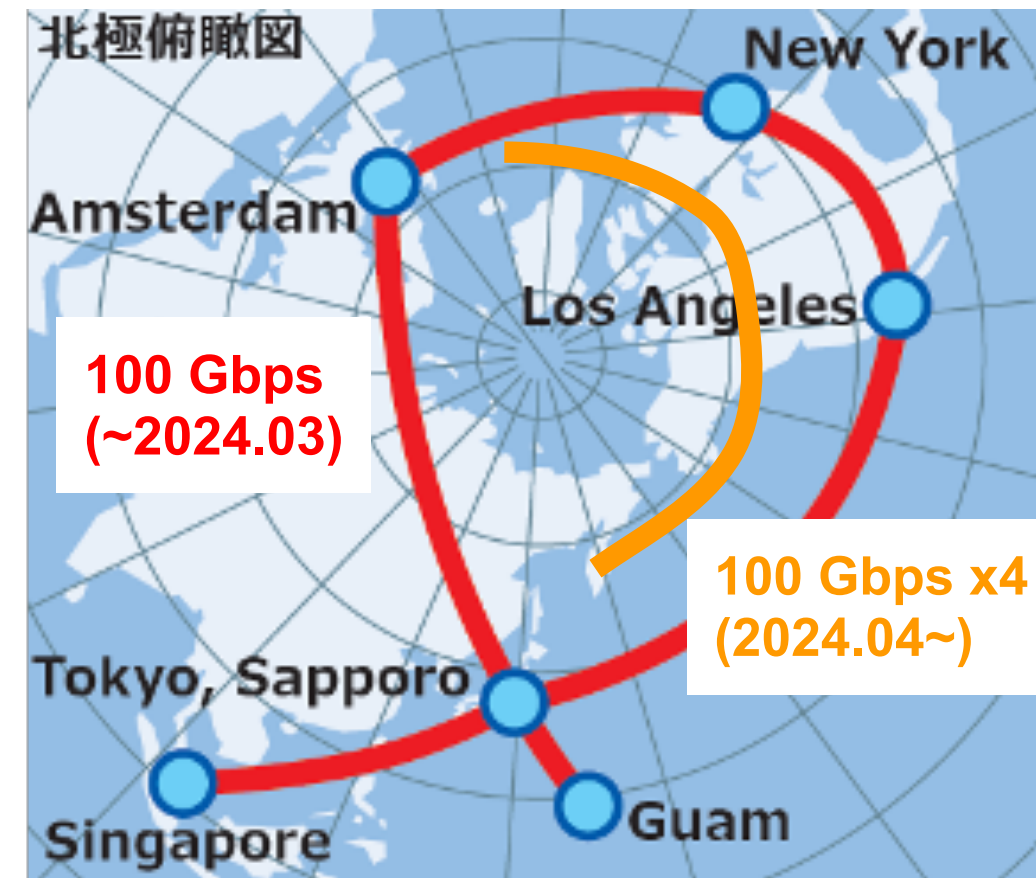
IPv6通信がメイン

SINET IPv6アドレスから東大のPIアドレスへ移行済み



欧州線(東京-アムステルダム線)の切り替え

- 東京 - アムステルダム間はユーラシア大陸に敷設されたケーブルが使われていた。
- 2024年4月に北米経由に切り替え
- バンド幅は増加: 100 Gbps → 100 Gbps x 4
- レイテンシも増加: 150 ms → 200 ~ 250 ms
 - productionへの影響はみられていません



切り替え前後のネットワーク転送量

Data Challenge

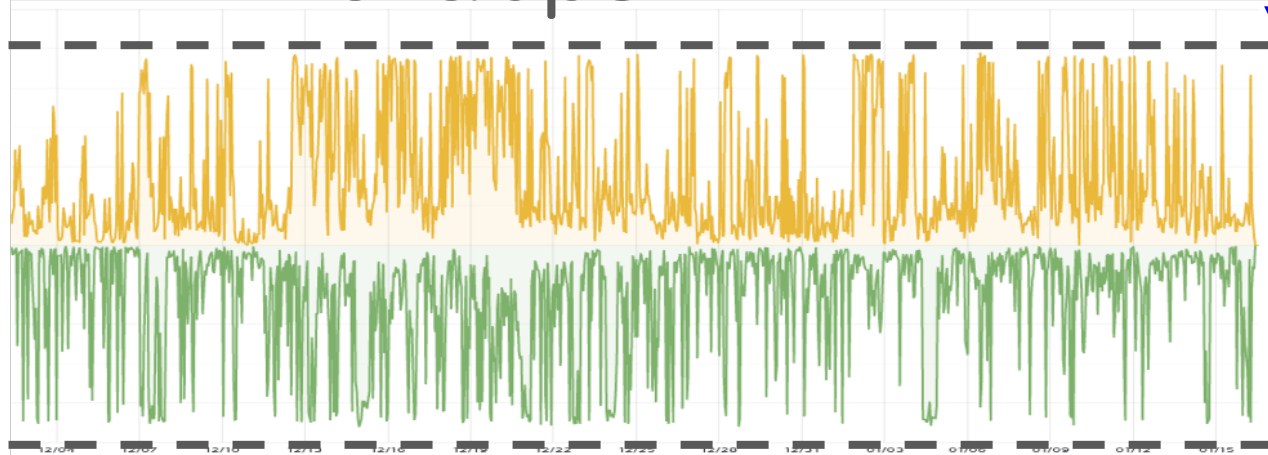
2024

(HL-LHCに向けた
Network負荷テスト)

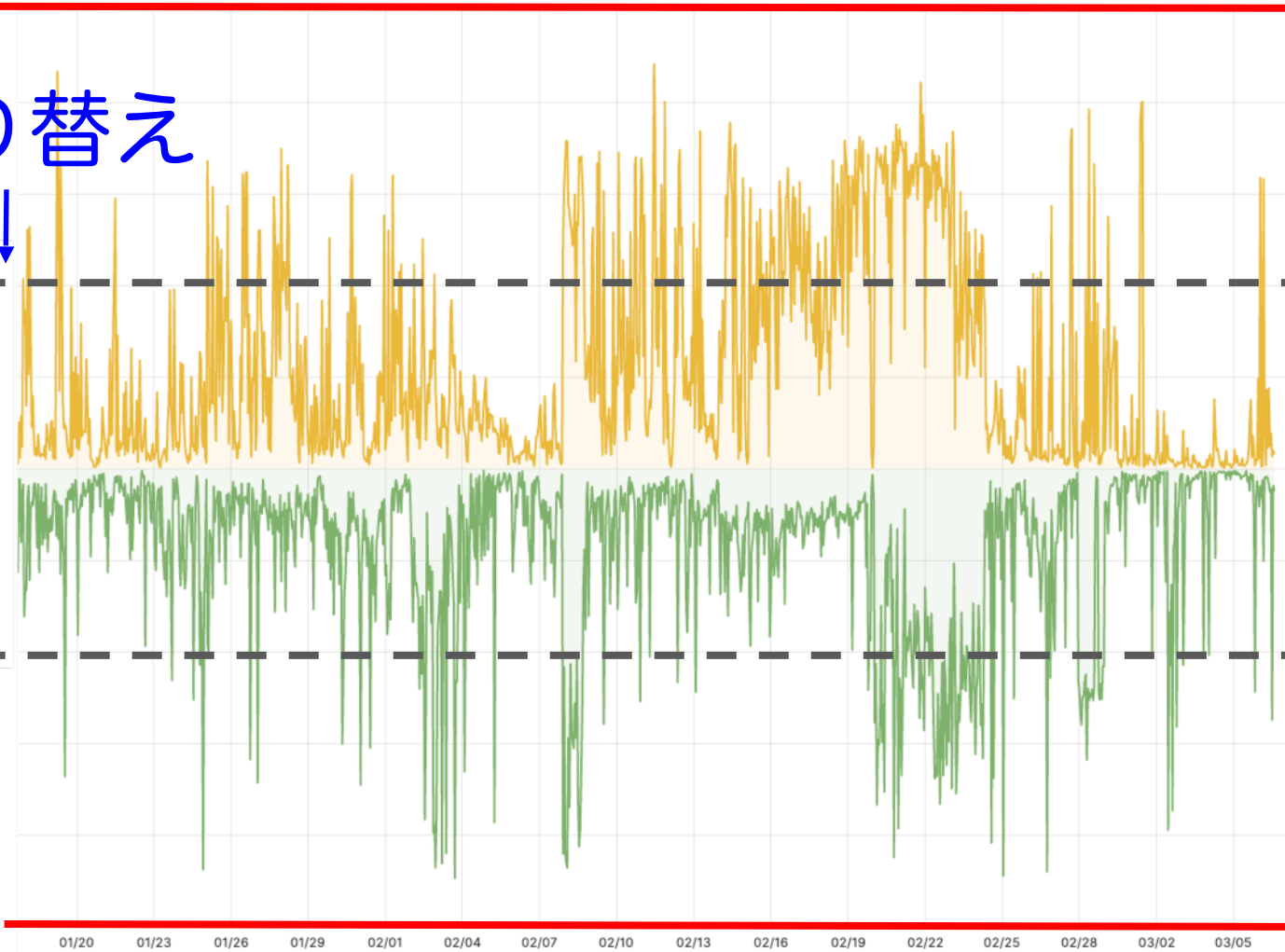
100 Gbps

外から内 40 Gbps

切り替え

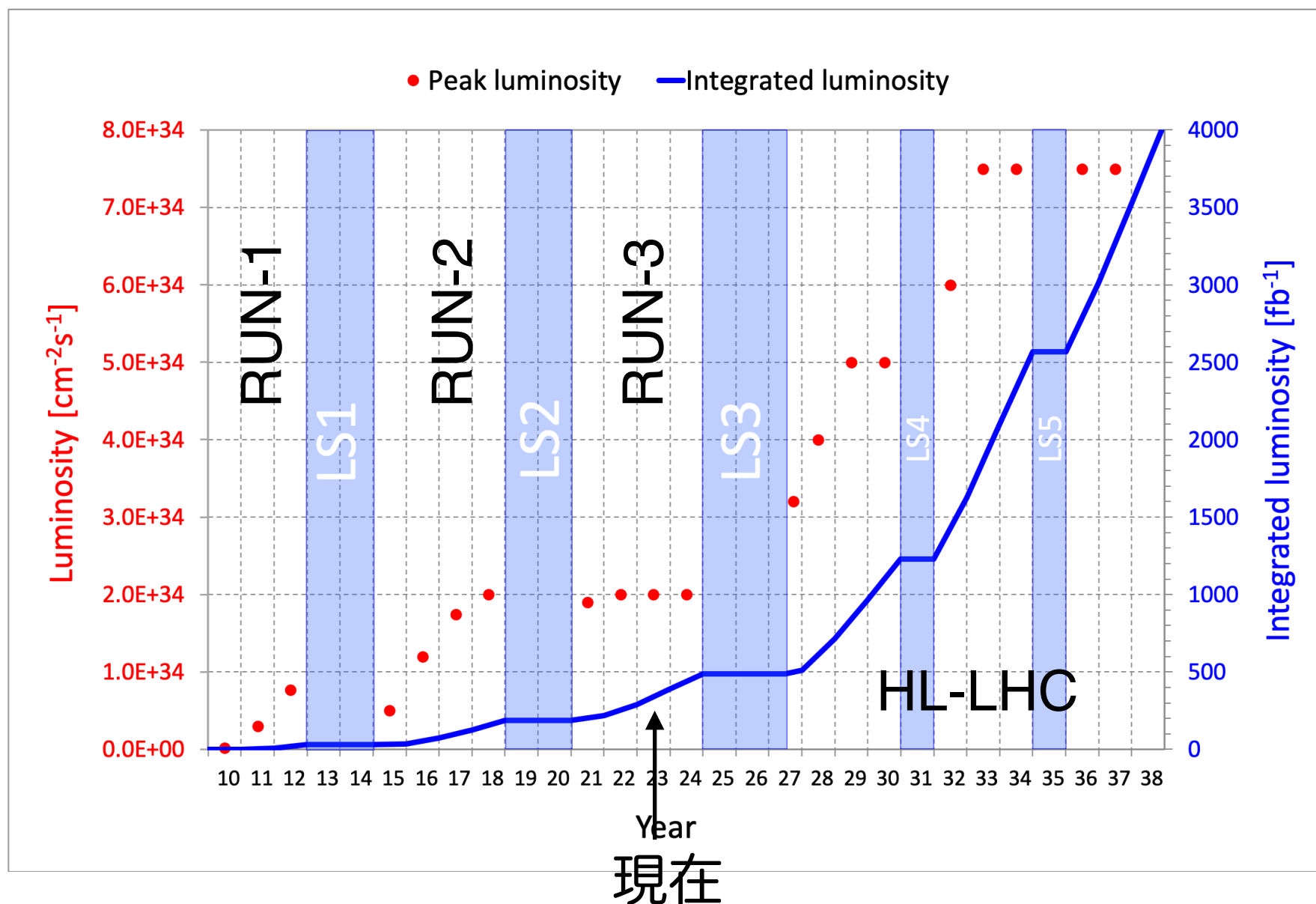


内から外



- 以前の上限(40Gbps)を超える転送が頻繁に起きていた
- 100Gbps近い転送も処理できている
 - アクセスパターンによってはファイルサーバーのIOが限界に近いため、ファイルサーバーのメモリの増強を検討

LHC 今後の計画

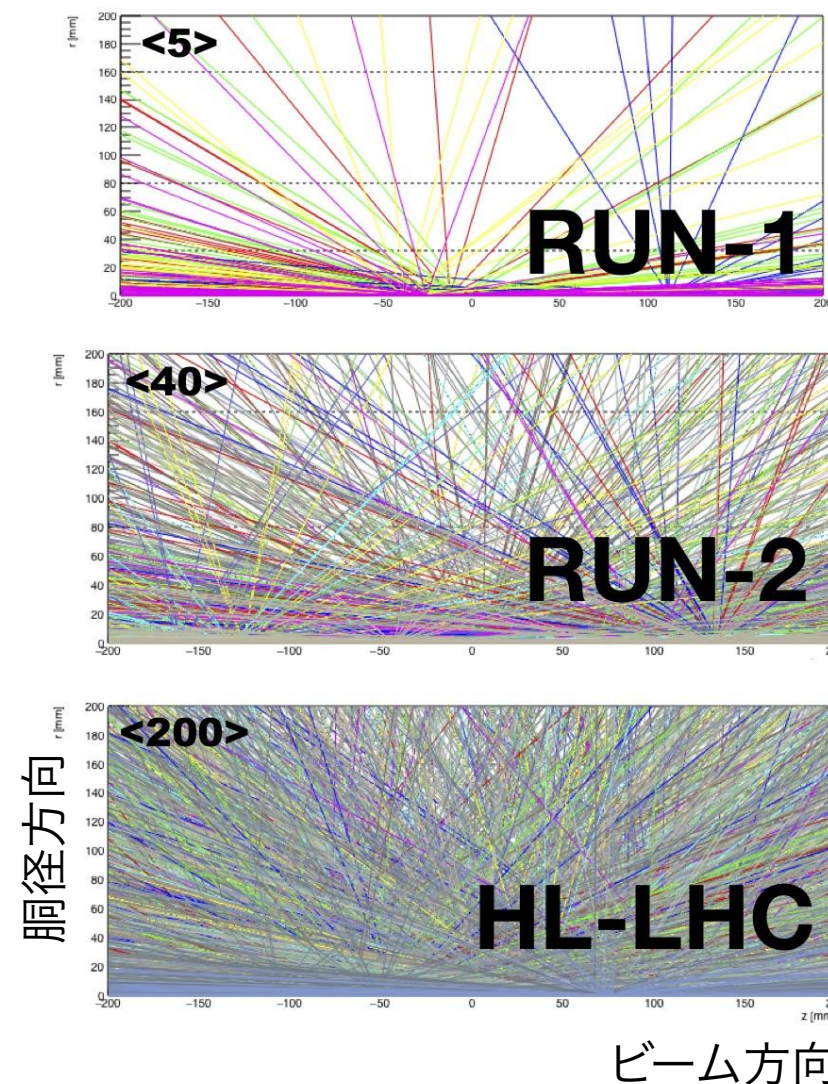


<https://lhc-commissioning.web.cern.ch/schedule/images/LHC-ultimate-lumi-projection.png>

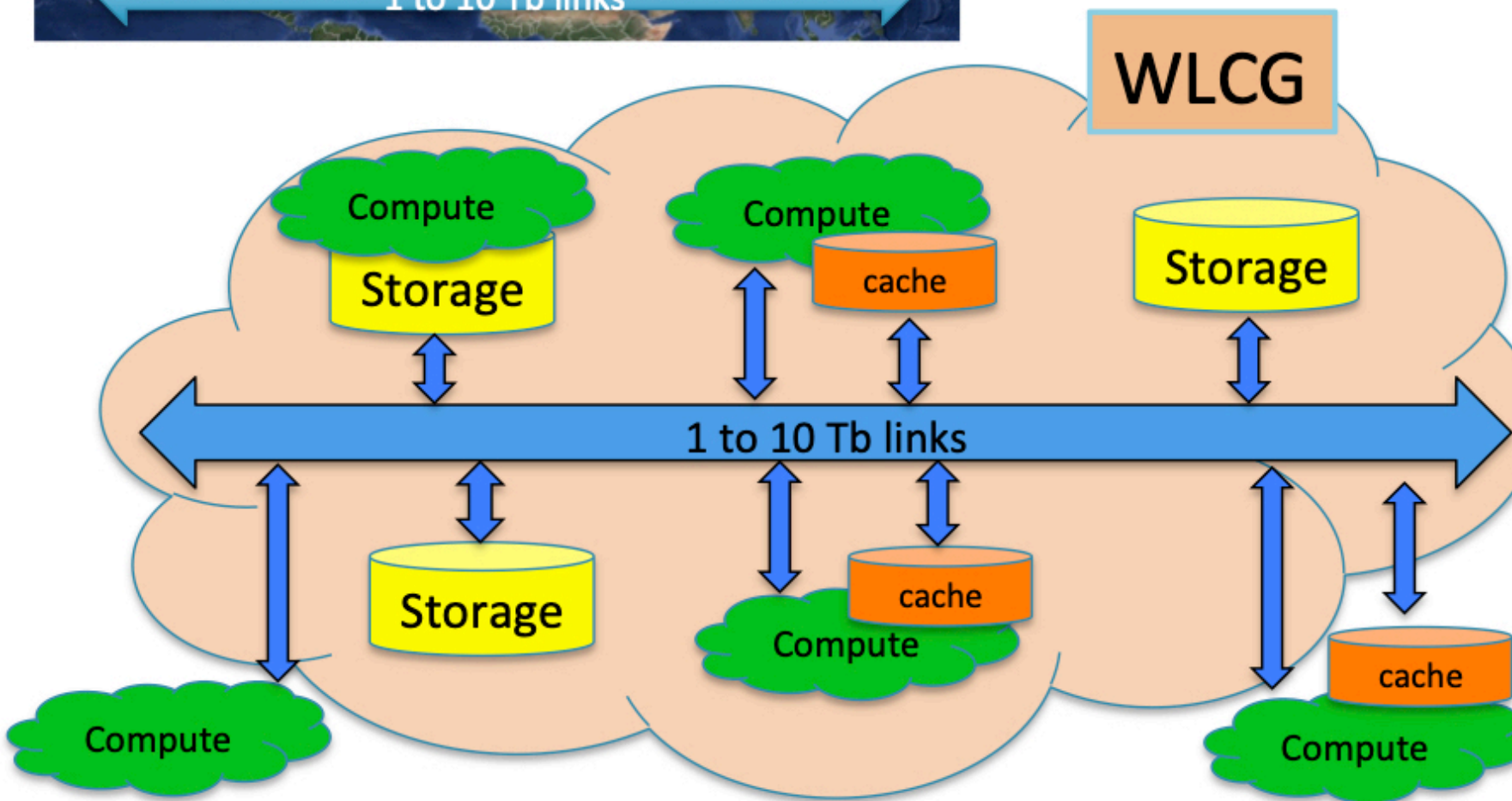
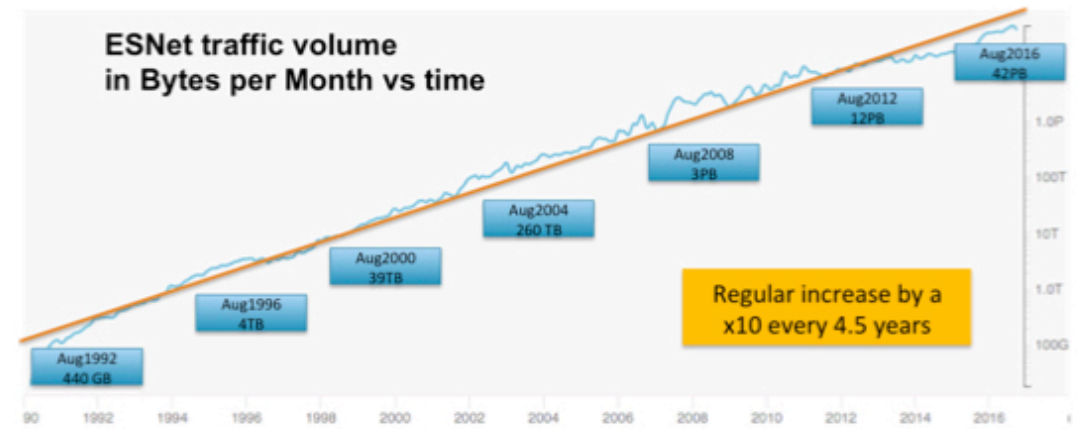
2040年ごろまでに、今までの~20倍のデータを貯める計画。
一回の陽子束衝突で起こる陽子衝突数も数倍になる。

今後より多くのリソースが必要。

ビーム交差あたり反応数



将来のコンピューティングモデル



- ・ サイト同士を高速ネットワークで接続
 - ・ データを少数サイトに集約して管理コストの削減
- ・ Gridだけではなく、HPCやクラウドリソースも利用。
 - ・ 東大地域解析センターではコールドストレージとしてAWSを検討：SINET-AWSの利用も？

Tier2は将来400Gbpsが想定されている。

まとめ

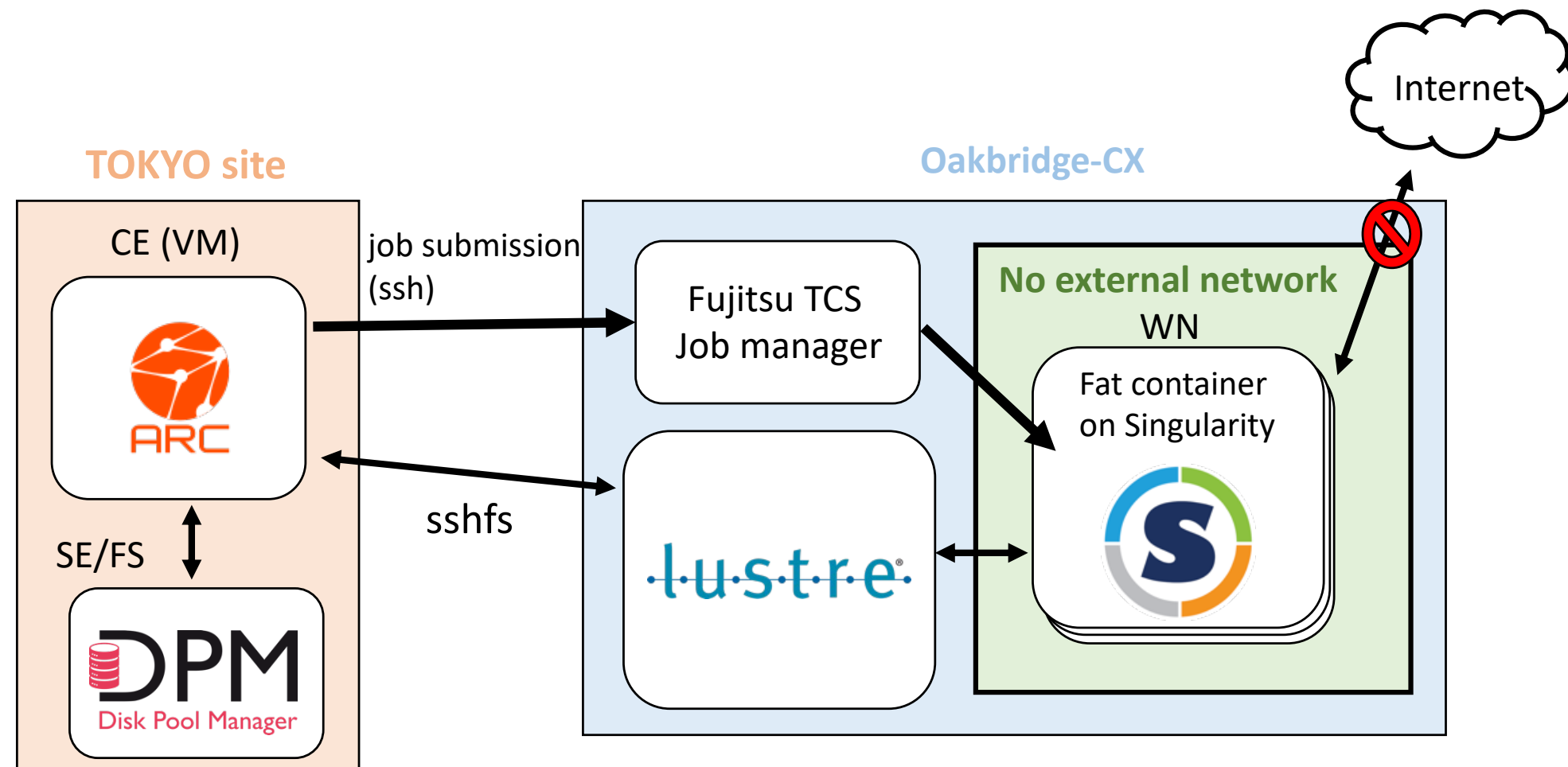
- LHC ATLAS実験は、世界の根源的な問いに答えるためのO(100PB/年)のデータを取得し、2040年代まで解析を行う。
 - ・ 実験を行う基盤技術は加速器、測定器、コンピューティング
- 世界中のデータセンターを接続したGridを基礎として、HPCやクラウドも利用。
 - ・ 高速なネットワークが必要不可欠
 - ・ ビーム強度の増強にともない、必要なリソースも増大する
 - ・ ネットワークのさらなる高速化によって、より効率的にリソースを利用するコンピューティングモデルが実現できる。

バックアップ

HPC 利用

東京大学のHPCを利用してR&D

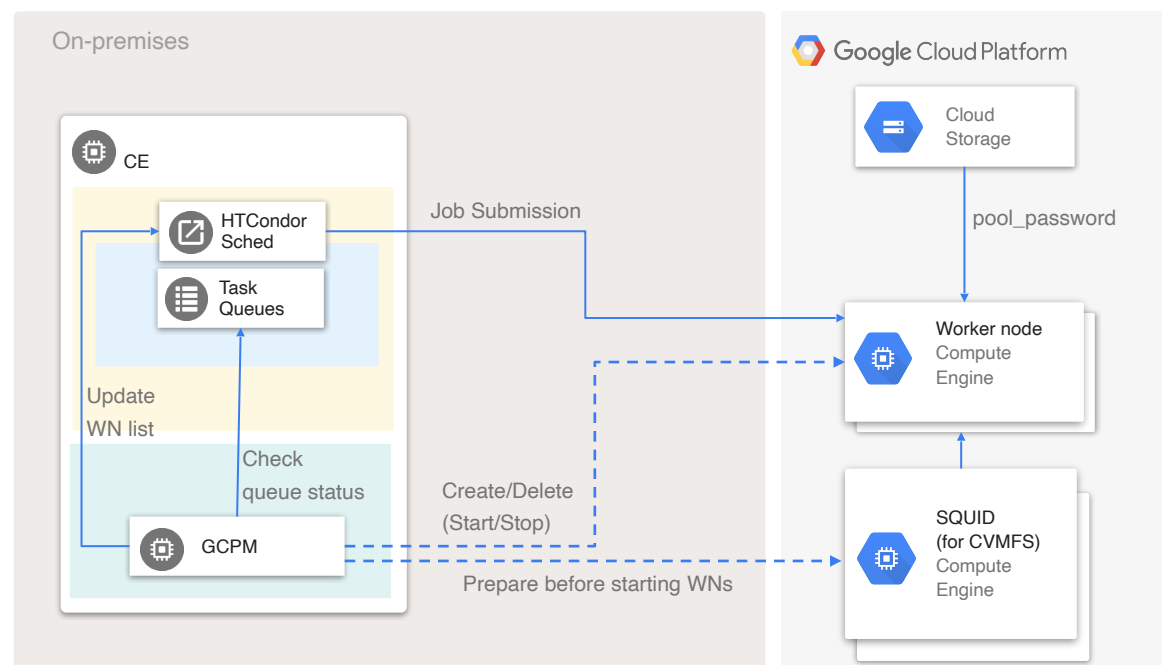
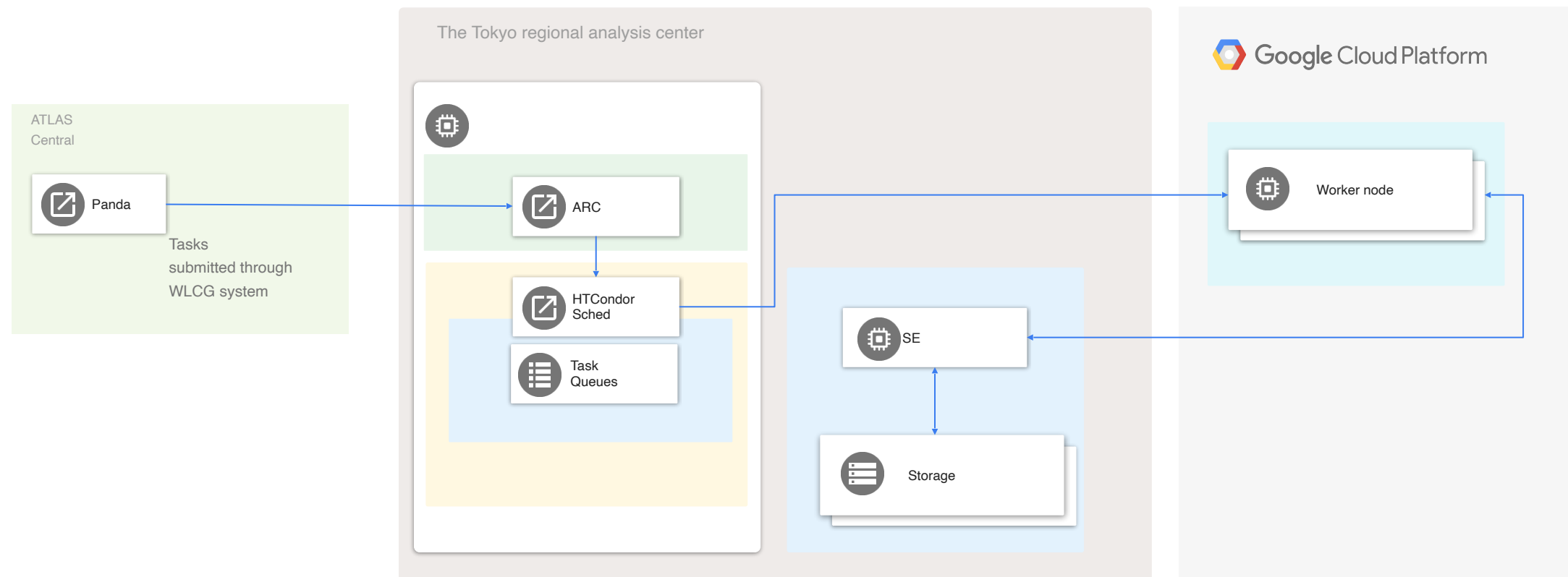
- Oakbridge-CX
 - Xeon Platinum 8280, 1,368ノード、6.61 PFLOPS
- Wisteria/BDEC01
 - A64FX, 7680ノード、33.1 PFLOPS
 - 現状ではATLASソフトウェア未対応。将来的には利用可能に



WNから外につながらないため、専用の”大きな”コンテナで対応。 23

クラウド(Google)利用

データ保存はオンプレミス(東大ICEPP)を利用、
一時的なキャッシュデータをGCPに作成してCPUを利用



HTCondorからGCPを利用するパッケージを開発
Google Cloud Condor Pool Manager → <https://github.com/mickaneda/gcpm>