

素粒子原子核研究所・計算機グループの現在の主な活動は Belle II コンピューティング・グループ（とくに Core Computing グループ）に参加し、実験におけるオンライン-オフライン間 RAW データ転送とその分散管理、さらに分散計算環境の改良およびその定常的運用である。またこれと並行して、国内外の研究機関と計算機ならびにソフトウェアに関する情報を共有し、世界的な動向を注視しつつ、KEK で行われている素粒子原子核実験や将来計画における持続可能な計算機環境を目指して活動を行っている。

1. オンライン-オフライン間 RAW データ転送の改良と実装・運用

2023 年 7 月の活動報告[1]に記載したとおり、2022 年 6 月から始まった Belle II 実験の長期運転停止期間 (LS1) 中に Belle II DAQ 側 (オンライン側) のデータ収集系の大幅な変更がおこなわれた。これに伴い、オフライン側への RAW データ転送系の改変が必要となった。この改変による利点を要約すると

- DAQ 側で Sequential ROOT (以後 SROOT と表す) と呼ばれる独自のフォーマットで記録していた RAW データを汎用的に使用できる ROOT 形式で保存することで、これまでオフライン側で行っていた SROOT→ROOT 形式変換が不要になる
- SROOT 形式は圧縮されておらず、データサイズは ROOT 形式の 2~3 倍の大きさがある。ROOT 形式の RAW データを転送することで、オンライン-オフライン間のデータ転送量を縮減でき、今後期待される高ルミノシティ環境への耐性が上がる

ことが挙げられる。これらの改変に加えて、より連続的な RAW データの転送も実装することとした。変更前は転送すべき RAW データがある程度オンライン側で蓄積されるのを待ち、後でまとめてオフライン側へ転送する、という方法をとっていたが、変更後は転送に必要な RAW データの情報をデータベース化し、新しい情報がデータベースに加えられることを切っ掛けとして RAW データ転送を開始するように改良した。これにより、これまでバンチで処理されていた RAW データ転送がラン毎に行われるようになり (図 1 参照)、より迅速に必要なオフライン処理を施せるようになった。

この改変に伴うオフライン側の RAW データ転送系の開発は、Belle II DAQ グループならびに Distributed Computing グループ協力の下、2022 年 11 月より Belle II Core Computing グループが中心となって行い、概念設計、各コンポーネントの開発、ユニットテストを繰り返し行ってきた。2023 年 11 月から 2024 年 1 月に断続的に行われた Belle II 検出器のテストオペレーション (Test Campaign I, II, III)、ならびに公式の Global Cosmic Ray Run のオペレーションを通して、システムの動作確認を行った。2024 年 2 月からは LS1 後初の Physics Run が再開され、大幅なシステム変更であったにもかかわらず、これまでの所、順調に運用されている。

このシステム改変に付随してもう一つ改良できる点がある。取った RAW データはトリガーの種類によって、検出器の Calibration のみに使うものと Calibration だけでなく物理解析にも使うものに大別できる (前者は prescale factor によってデータを間引くが、後者は間引かない)。Belle II 実験では Prompt Calibration と呼ばれる行程を約 2 週間毎に行うことにしており、そのタイミングで prescale factor を決定、その後 RAW データの選別を行っていた。しかし、prescale factor を必要としない後者のデータは、RAW データが使えるようになれば 2 週間も待つ必要はなくすぐに作成で

きる。そこでシステム変更後は SROOT→ROOT 形式変換がオフライン側で必要なくなったことから、これに使用されていた計算資源を利用することで、RAW データ転送から連続してラン毎に後者のデータを作成できるように改良中である。これにより Prompt Calibration の行程ならびにそれに続く物理解析用データの生成を効率化することを目指している。

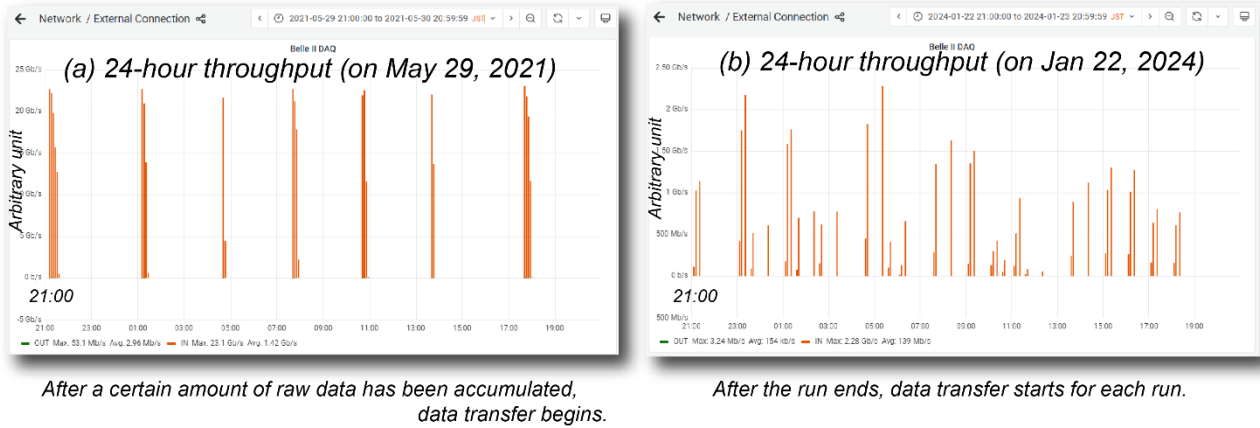


図 1：オンライン→オフライン RAW データ転送のある一日におけるネットワーク・スループット（縦軸は任意）、(a)システム変更前の状況。ある程度 RAW データがオンライン側で蓄積されたあと、オフライン側へ転送されているため、バンチ構造となっている。(b)システム変更後の状況。ランが終わるとほぼ同時にデータ転送が開始され、速やかに RAW データ転送が終了している。

2. DIRAC & Rucio 合同ワークショップの開催

分散計算環境を構築するため各計算機サイトで運用しているシステムは、小規模計算機クラスターから大規模バッチシステム、クラウド・コンピューティング、HPC (High Performance Computer) と多種多様である。これら異なる計算資源を相互運用出来るようにするため、LHCb を中心に DIRAC [2]と呼ばれるインターウェアが開発され、ジョブならびにデータの双方が管理できるようになった。また、多数の計算機サイトに分散したデータを统一的に管理、利用することを可能にする Rucio [3]と呼ばれる分散データ管理ツールも ATLAS を中心に開発された。Belle II 実験は現在両者を基本ツールとして利用している唯一の実験であることから、これまでヨーロッパを中心に独立して開かれていた DIRAC と Rucio ワークショップを 2023 年 10 月 16 日から 20 日にかけて初めて合同で KEK にて開催した。[4] アメリカ合衆国、スイス (CERN)、フランス、イギリス、イタリア、ロシア、メキシコ、中国、日本から合計 40 名ほどの専門家ならびにユーザーが集結し、それぞれが培ってきた技術や経験、また問題点を共有することで、よりよい分散計算環境を実現するため、活発な議論ならびに現地でのハッカソンを行った。これにより双方の開発者、さらにはそれを利用している実験の計算機担当者らとの面識を広げ、問題点や開発項目の共同認識に至るなど、有意義な会議を開催することが出来た (図 2)



図 2 : 2023 年 10 月に KEK で開催された DIRAC & Rucio 合同ワークショップのグループ写真

3. Data Challenge 2024 への参加

Belle II 実験では RAW データを実験のホスト機関である KEK に保管するほか、その複製を共同研究機関のデータ・センターに分散保管する。 これらを RAW データ・センターと呼び、BNL (アメリカ合衆国)、CNAF (イタリア)、DESY、KIT (ともにドイツ)、University of Victoria (カナダ)、CC-IN2P3 (フランス) の 6 機関に設置している。 Belle II 実験では今後のルミノシティ向上により 40TB/day の RAW データが期待され、これらを高速かつ継続的に KEK から各 RAW データ・センターに転送できる国際ネットワーク回線が生命線となる。 仮に転送に一時的に不具合があった場合にも速やかに回復できるためには、期待されるデータ取得レートの数倍の回線容量が必要となる。 一方、これらの RAW データ・センターの多くは LHC 実験でも主要なデータ・センターとして機能しており、将来的に Belle II 実験のデータ転送と(HL-)LHC 実験のデータ転送がネットワークの使用において競合する可能性がある。 WLCG (Worldwide LHC Computing Grid) [5] では、これらの状況を見越した計算資源やネットワーク・インフラの評価を進めており、その一環として 2024 年 2 月 12 日から 23 日にかけて “Data Challenge 2024” というネットワークだけでなくモニター系も含めたストレス・テストを計画し、Belle II 実験も計画段階から参加した。 図 3 に Data Challenge 2024 の期間中に観測したネットワーク・トラフィックを示す。 今回、Belle II 実験は前出の 40TB/day を基準値とし、達成目標を $5 \times 40\text{TB/day}$ (5 日分の Raw データを 1 日で送る量)、成功判断基準を $3 \times 40\text{TB/day}$ (3 日分の Raw データを 1 日で送る量) に設定した。 図 3 に示すように、KEK からの全体的なトラフィックは概ね成功判断基準を達成していることが分かる。

Belle II 実験で採用している分散計算環境は LHC 実験の分散計算環境のために開発されたツールを多く使用しており、また RAW データ・センターだけでなく参加している多くの計算機サイトは既に LHC 実験のため WLCG に加盟している。 そのため WLCG の活動や方向性とは切っても切り離

せない。また WLCG 側も LHC で行われている実験だけでなく、他実験や他分野への拡張も重要と考えている。2015 年より Belle II 実験は LHC 実験以外で初めて WLCG に参加することが許され、それ以降、情報共有を初めとした協力関係が続いている。

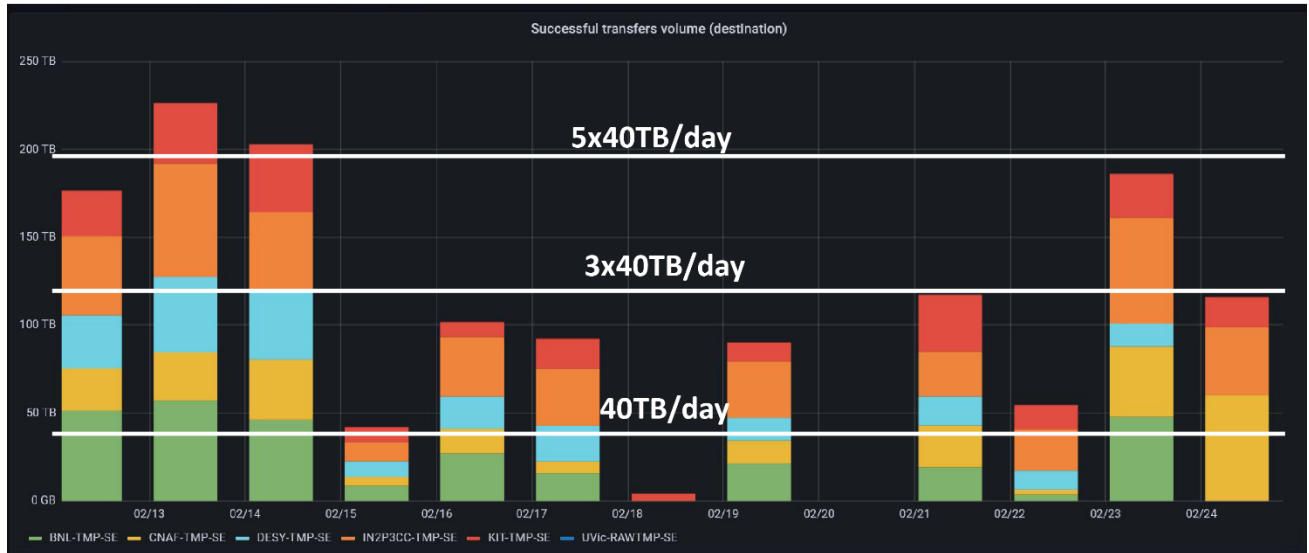


図 3 : Data Challenge 2024 (2024 年 2 月 12 日から 2 月 24 日) 期間中に観測した Belle II 実験のテストに起因する KEK から各 RAW データ・センターへのネットワーク・トラフィック。色は各 RAW データ・センターへの転送量を表している。なお、Data Challenge 2024 期間中、University of Victoria は整備中であったため転送量が少なかった。

4. 参考資料

- [1] : 2023 年 7 月の活動報告 : コンピューティンググループ
https://www2.kek.jp/ipns/ja/research_introduction/report/4854/
- [2] : DIRAC, the Interware, <http://diracgrid.org/>
- [3] : Rucio, Scientific Data Management, <https://rucio.cern.ch/>
- [4] : DIRAC & Rucio Workshop 2023, <https://indico.cern.ch/event/1252369/>
- [5] : "Worldwide LHC Computing Grid", <http://wlcg.web.cern.ch/>