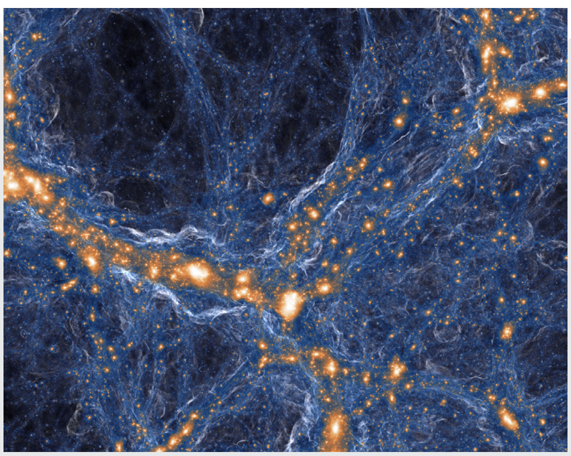
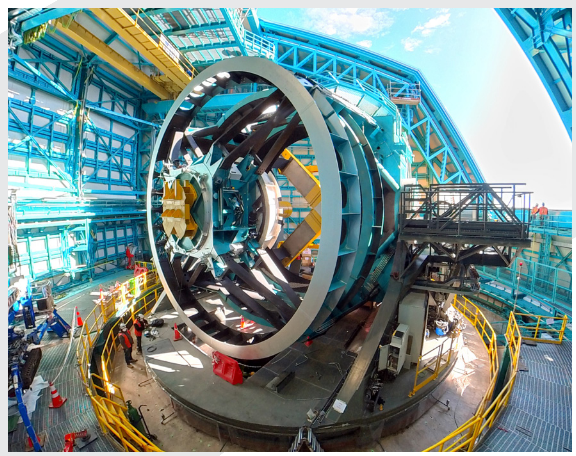
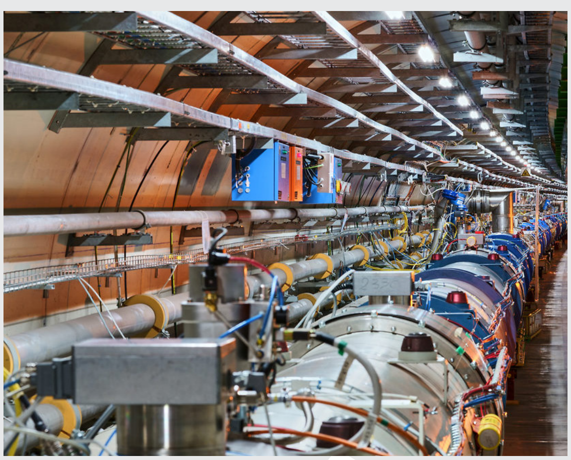
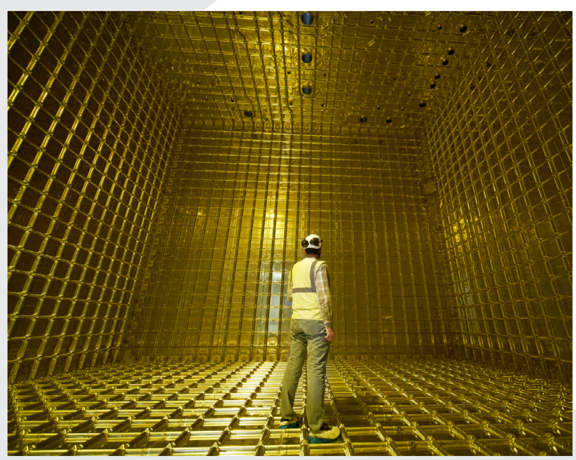
ESnet

ENERGY SCIENCES NETWORK

High Energy Physics Network Requirements Review: One-Year Update

April 2022

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY

Office of Science

High Energy Physics Network Requirements Review: One-Year Update

April 2022

Office of High Energy Physics, DOE Office of Science
Energy Sciences Network (ESnet)

ESnet is funded by the US Department of Energy, Office of Science, Office of Advanced Scientific Computing Research. Carol Hawk is the ESnet Program Manager.

ESnet is operated by Lawrence Berkeley National Laboratory (Berkeley Lab), which is operated by the University of California for the US Department of Energy under contract DE-AC02-05CH11231.

This work was supported by the Directors of the Office of Science, Office of Advanced Scientific Computing Research, Facilities Division, and the Office of High Energy Physics.

This is LBNL report number LBNL-2001492[1].

Cover Images:
(**Top left**) the new universe simulation model, dubbed Illustris, courtesy of Simons Foundation
(**Top right**) Wide view of the telescope mount inside the dome at the Rubin Observatory, courtesy of H. Stockebrand/Rubin Obs/NSF/AURA
(**Bottom left**) LHC tunnel image, courtesy of CERN
(**Bottom right**) protoDUNE detectors at CERN, courtesy of Max Brice/CERN) PHENIX image

[1] https://escholarship.org/uc/item/65h4t5zc

Participants and Contributors

Garhan Attebury, *University of Nebraska-Lincoln*

Nicole Avila, *University of Chicago*

Stephen Bailey, *Lawrence Berkeley National Laboratory*

Justas Balcas, *California Institute of Technology*

Amanda Bauer, *Rubin Observatory*

Lothar Bauerdick, *Fermi National Accelerator Laboratory*

Chris Bee, *Stony Brook University*

Doug Benjamin, *Argonne National Laboratory*

Kurt Biery, *Fermi National Accelerator Laboratory*

Kenneth Bloom, *University of Nebraska-Lincoln*

Bob Blum, *Rubin Observatory*

Andrey Bobyshev, *Fermi National Accelerator Laboratory*

Brian Bockelman, *University of Wisconsin-Madison*

Tim Bolton, *Kansas State University*

Vincent Bonafede, *Brookhaven National Laboratory*

Julian Borrill, *Lawrence Berkeley National Laboratory and University of California, Berkeley*

Tulika Bose, *University of Wisconsin-Madison*

Joseph Boudreau, *University of Pittsburgh*

Steve Brice, *Fermi National Accelerator Laboratory*

Benjamin Brown, *Department of Energy, Office of Science*

Paolo Calafiura, *Lawrence Berkeley National Laboratory*

Simone Campana, *European Organization for Nuclear Research*

Dale Carder, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

John Carlstrom, *University of Chicago and Argonne National Laboratory*

Eric Colby, *Department of Energy, Office of Science*

John Corlett, *Lawrence Berkeley National Laboratory*

Eli Dart, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Kaushik De, *University of Texas at Arlington*

Phil DeMar, *Fermi National Accelerator Laboratory*

Richard Dubois, *Stanford University*

Daniel Eisenstein, *Harvard University*

Johannes Elmsheuser, *Brookhaven National Laboratory*

Simon Fiorucci, *Lawrence Berkeley National Laboratory*

Mark Foster, *SLAC National Accelerator Laboratory*

Stuart Fuess, *Fermi National Accelerator Laboratory*

Robert Gardner, *University of Chicago*

Gil Gilchriese, *Lawrence Berkeley National Laboratory*

Heather Gray, *University of California Berkeley*

Chin Guok, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Oliver Gutsche, *Fermi National Accelerator Laboratory*

Julien Guy, *Lawrence Berkeley National Laboratory*

Salman Habib, *Argonne National Laboratory*

Damian Hazen, *Lawrence Berkeley National Laboratory and The National Energy Research Scientific Computing Center*

Katrin Heitmann, *Argonne National Laboratory*

Ken Herner, *Fermi National Accelerator Laboratory*

Saswata Hier-Majumder, *Department of Energy, Office of Science*

Michael Hildreth, *University of Notre Dame*

David Jaffe, *Brookhaven National Laboratory*

Jeff Kantor, *Rubin Observatory*

Wesley Ketchum, *Fermi National Accelerator Laboratory*

Mike Kirby, *Fermi National Accelerator Laboratory*

Alexei Klimentov, *Brookhaven National Laboratory*

Markus Klute, *Massachusetts Institute of Technology*

Robert Kutschke, *Fermi National Accelerator Laboratory*

Eric Lancon, *Brookhaven National Laboratory*

David Lange, *Princeton University*

Kevin Lannon, *University of Notre Dame*

Paul Laycock, *Brookhaven National Laboratory*

Tom Lehman, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

James Letts, *University of California San Diego*

Michael Levi *(Dark Energy Spectroscopic Instrument (DESI) Director), Lawrence Berkeley National Laboratory*

Mark Lukasczyk, *Brookhaven National Laboratory*

Adam Lyon, *Fermi National Accelerator Laboratory*

Krista Majewski, *Fermi National Accelerator Laboratory*

Dan Marlow, *Princeton University*

Phil Marshall, *SLAC National Accelerator Laboratory*

Edoardo Martelli, *European Organization for Nuclear Research*

David Mason, *Fermi National Accelerator Laboratory*

Shawn Mckee, *University of Michigan*

Andrew Melo, *Vanderbilt University*

Bogdan Mihaila, *National Science Foundation*

Ken Miller, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Bill Miller, *National Science Foundation*

Inder Monga, *Lawrence Berkeley National Laboratory and Energy Sciences Network (*

Maria Elena Monzani, *SLAC National Accelerator Laboratory*

Harvey Newman, *California Institute of Technology*

Will O'Mullane, *Rubin Observatory*

Nathalie Palanque-Delabrouie, *The French Alternative Energies and Atomic Energy Commission*

Ramon Pasetes, *Fermi National Accelerator Laboratory*

Abid Patwa, *Department of Energy, Office of Science*

Christoph Paus, *Massachusetts Institute of Technology*

Srini Rajagopalan, *Brookhaven National Laboratory*

Quentin Riffard, *Lawrence Berkeley National Laboratory*

Kate Robinson, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

David Schelgel, *Lawrence Berkeley National Laboratory*

Heidi Schellman, *Oregon State University*

Kate Scholberg, *Duke University*

Jennifer Schopf, *Indiana University*

Uros Seljak, *University of California, Berkeley*

Elizabeth Sexton-Kennedy, *Fermi National Accelerator Laboratory*

Richard Simon, *Lawrence Berkeley National Laboratory*

Eric Smith, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Maria Spiropulu, *California Institute of Technology*

Tavia Stone Gibbins, *Lawrence Berkeley National Laboratory and The National Energy Research Scientific Computing Center*

Rune Stromsness, *Lawrence Berkeley National Laboratory*

Matevz Tadel, *University of California San Diego*

Kevin Thompson, *National Science Foundation*

Steve Timm, *Fermi National Accelerator Laboratory*

Margaret Votava, *Fermi National Accelerator Laboratory*

Paul Wefel, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Torre Wenaus, *Brookhaven National Laboratory*

Andrew Wiedlea, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Linda Winkler, *Argonne National Laboratory*

Frank Wuerthwein, *University of California San Diego*

Wei Yang, *SLAC National Accelerator Laboratory*

Xi Yang, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Jim Yeck, *University of Wisconsin, Madison*

Alexandr Zaytsev, *Brookhaven National Laboratory*

Jason Zurawski, *Lawrence Berkeley National Laboratory and Energy Sciences Network*

Report Editors

Ben Brown, *Department of Energy, Office of Science:* Benjamin.Brown@science.doe.gov

Dale Carder, *ESnet:* dwcarder@es.net

Eric Colby, *Department of Energy, Office of Science:* eric.colby@science.doe.gov

Eli Dart, *ESnet:* dart@es.net

Carol Hawk, *Department of Energy, Office of Science:* Carol.Hawk@science.doe.gov

Saswata Hier-Majumder, *Department of Energy, Office of Science:* saswata.hier-majumder@science.doe.gov

Bill Miller, *Department of Energy, Office of Science:* bill.miller@science.doe.gov

Ken Miller, *ESnet:* ken@es.net

Abid Patwa, *Department of Energy, Office of Science:* abid.patwa@science.doe.gov

Kate Robinson, *ESnet:* katerobinson@es.net

Lauren Rotman, *ESnet:* lauren@es.net

Andrew Wiedlea, *ESnet:* awiedlea@es.net

Jason Zurawski, *ESnet:* zurawski@es.net

Table of Contents

1 Executive Summary

About ESnet

The Energy Sciences Network (ESnet) is the high-performance network user facility for the US Department of Energy (DOE) Office of Science (SC) and delivers highly reliable data transport capabilities optimized for the requirements of data-intensive science. In essence, ESnet is the circulatory system that enables the DOE science mission by connecting all its laboratories and facilities in the US and abroad. ESnet is funded and stewarded by the Advanced Scientific Computing Research (ASCR) program and managed and operated by the Scientific Networking Division at Lawrence Berkeley National Laboratory (LBNL). ESnet is widely regarded as a global leader in the research and education (R&E) networking community.

ESnet interconnects DOE national laboratories, user facilities, and major experiments so that scientists can use remote instruments and computing resources as well as share data with collaborators, transfer large data sets, and access distributed data repositories. ESnet is specifically built to provide a range of network services tailored to meet the unique requirements of the DOE's data-intensive science.

In short, ESnet's mission is to enable and accelerate scientific discovery by delivering unparalleled network infrastructure, capabilities, and tools. ESnet's vision is summarized by these three points:

1. Scientific progress will be completely unconstrained by the physical location of instruments, people, computational resources, or data.
2. Collaborations at every scale, in every domain, will have the information and tools they need to achieve maximum benefit from scientific facilities, global networks, and emerging network capabilities.
3. ESnet will foster the partnerships and pioneer the technologies necessary to ensure that these transformations occur.

Requirements Review Purpose and Process

ESnet and ASCR use requirements reviews to discuss and analyze current and planned science use cases and anticipated data output of a particular program, user facility, or project to inform ESnet's strategic planning, including network operations, capacity upgrades, and other service investments. A requirements review comprehensively surveys major science stakeholders' plans and processes in order to investigate data management requirements over the next 5–10 years. Questions crafted to explore this space include the following:

- How, and where, will new data be analyzed and used?
- How will the process of doing science change over the next 5–10 years?
- How will changes to the underlying hardware and software technologies influence scientific discovery?

Requirements reviews help ensure that key stakeholders have a common understanding of the issues and the actions that ESnet may need to undertake to offer solutions. The ESnet Science Engagement Team leads the effort and relies on collaboration from other ESnet teams: Software Engineering, Network Engineering, and Network Security. This team meets with each individual program office within the DOE SC every three years, with intermediate updates scheduled every off year. ESnet collaborates with the relevant program managers to identify the appropriate principal investigators, and their information technology partners, to participate in the review process. ESnet organizes, convenes, executes, and shares the outcomes of the review with all stakeholders.

This Review

In April 2022, ESnet and the Office of High Energy Physics (HEP) of the DOE SC organized an ESnet requirements review of HEP-supported activities. Preparation for the review included identification of key stakeholders: program and facility management, research groups, and technology providers. Each stakeholder group was asked to prepare formal case study documents about the group's relationship to the HEP program to build a complete understanding of the current, near-term, and long-term status, expectations, and processes that will support the science going forward. A series of pre-planning meetings better prepared case study authors for this task, along with guidance on how the review would proceed in a virtual fashion.

The HEP program's mission is to understand how the universe works at its most fundamental level by discovering the elementary constituents of matter and energy, probing the interactions between them, and exploring the basic nature of space and time. This research and development (R&D) inspires young minds, trains an expert workforce, and drives innovation that improves the nation's health, wealth, and security.

The scientific objectives and priorities for the field recommended by the HEP Advisory Panel are detailed in its recent long-range strategic plan, developed by the Particle Physics Project Prioritization Panel (P5). HEP research is inspired by some of the most fundamental questions about our universe. What is it made of? What forces govern it? How did it evolve to the way it is today? Finding these answers requires the combined efforts of some of the largest scientific collaborations in the world, using large arrays of the most sensitive detectors in the world, at some of the largest and most complex scientific machines in the world.

HEP supports US researchers who play leading roles in these international efforts and world-leading facilities at our national laboratories that make this science possible. HEP also develops new accelerator, detector, and computational tools to open new doors to discovery science, and through the Accelerator Stewardship program, works to make transformational accelerator technology widely available to science and industry.

This review includes case studies from the following HEP facilities, experiments, and joint collaborative efforts:

- Large Hadron Collider (LHC) experimentation and operation
  - ATLAS (A Toroidal LHC ApparatuS) experiment
  - Compact Muon Solenoid (CMS) experiment
  - LHC operations
  - High-Luminosity (HL) era of the LHC
- Neutrino experiments at Fermilab
  - Short-Baseline Neutrino Program (SBN)
  - Deep Underground Neutrino Experiment (DUNE)
- DESI
- Belle II experiment
  - Muon experimentation at Fermilab
  - Muon G minus two (g-2)
- Muon-to-electron-conversion experiment (Mu2e)
- Dark Energy Science Collaboration (DESC)
- The Vera C. Rubin Observatory (Rubin Observatory) and the Legacy Survey of Space and Time (LSST)
- Cosmic Microwave Background — Stage 4 (CMB-S4)

- Cosmological simulation research
- LZ (LUX-ZEPLIN) Dark Matter Experiment

Requirements reviews are a critical part of a process to understand and analyze current and planned science use cases across the DOE SC. This is done by eliciting and documenting the anticipated data outputs and workflows of a particular program, user facility, or project to better inform strategic planning activities. These include, but are not limited to, network operations, capacity upgrades, and other service investments for ESnet as well as a complete and holistic understanding of science drivers and requirements for the program offices.

We achieve these goals by review of the case study documents, discussions with authors, and general analysis of the materials. The resulting output is a set of review findings and recommendations that will guide future interactions between HEP, ASCR, and ESnet. These terms are defined as follows:

- **Findings**: key facts or observations gleaned from the entire review process that highlight specific challenges, particularly those shared among multiple case studies.
- **Recommendations**: potential strategic or tactical activities, investments, or opportunities recommended to be evaluated and potentially pursued to address the challenges laid out in the findings.

The review participants spanned the following roles:

- Subject-matter experts from the HEP activities listed previously.
- ESnet Site Coordinators Committee members from HEP activity host institutions, including the following DOE labs: Argonne National Laboratory (ANL), Brookhaven National Laboratory (BNL), the European Organization for Nuclear Research (CERN), Fermi National Accelerator Laboratory (Fermilab), LBNL, the National Energy Research Scientific Computing Center (NERSC), and SLAC National Accelerator Laboratory (SLAC).
- Networking and/or science engagement leads from the ASCR high-performance computing (HPC) facilities.
- DOE SC staff spanning both ASCR and HEP.
- ESnet staff supporting positions related to facility leadership, scientific engagement, networking, security, software development, and R&D.

The review produced several important findings from the case studies and subsequent virtual conversations:

- Pandemic delays are persistent and have caused certain delays in design, testing, construction, and experimental run-time.
- Increased data volumes were reported by several experiments, but not by significant margins beyond the predictions that were made during the 2020 case studies.
- R&D efforts conducted during the pandemic, in several projects, are investigating more efficient mechanisms to retain, process, and disseminate data.
- Leveraging "data challenges," synthetic exercises to evaluate the performance of hardware and software using previously collected, or simulated, data sets has increased.
- Case studies that primarily leverage grid computing models continue to explore ways to leverage resources at HPC facilities.
- Experiments that make data sets available to a wide set of collaborators, either through a data management system that intelligently distributes, or a portal system that can be used by anyone, have reported that some engineering work will be needed in the future to understand bottlenecks.

- Investigation into the use of machine learning (ML) and artificial intelligence (AI) approaches continues, and this leverages a mixture of commercial clouds as well as resources located at R&E HPC facilities.
- The use of commercial cloud resources to support components of experimental workflows continues to be an area of active investigation.

Lastly, ESnet will be following up with participants in the coming years on several recommendations that were identified:

- ESnet will continue to participate with community members in discussions surrounding best common practices for data mobility and management.
- ESnet will continue to discuss progress of the transatlantic capabilities with key HEP stakeholders and site upgrades affiliated with ESnet6 commissioning, and continue to maintain and augment cloud connectivity options.
- ESnet will continue to work with R&D efforts in the LHC community, along with the Large Hadron Collider Open Network Environment (LHCONE) and Large Hadron Collider Optical Private Network (LHCOPN) planning groups.
- ESnet will participate in DOE efforts such as the Infrastructure Architecture Blueprint Activity (IRI-ABA).
- ESnet will discuss some of the use cases that ESnet services can address related to ML and AI research.

2 Review Findings

The requirements review process helps to identify important facts and opportunities from the programs and facilities that are profiled. These points summarize important information gathered during the review discussions surrounding case studies and the HEP program in general:

- Pandemic delays are persistent and causing delays in design, testing, construction, and experimental run-time.
- Increased data volumes were reported by several experiments, but not by significant margins beyond the predictions that were made during the 2020 case studies. These increases are mostly related to minor changes in the process of analysis, as well as increased runtimes.
- R&D efforts conducted during the pandemic, in several projects, are investigating more efficient mechanisms to retain, process, and disseminate data. This has contributed to increases in volume, as well as use of network resources to reach distributed computational and storage facilities.
- Leveraging "data challenges," synthetic exercises to evaluate the performance of hardware and software using previously collected, or simulated, data sets has increased. Some projects (e.g., the LHC, DUNE, Rubin, and CMB-S4) have found these exercises to be valuable as they plan implementation, or changes, to their computational models.
- Case studies that primarily leverage grid computing models continue to explore ways to leverage resources at HPC facilities. These investigations are mostly related to converting some workflow components (e.g., simulation production), versus re-engineering more complex analysis approaches.
- Experiments that make data sets available to a wide set of collaborators, either through a data management system that intelligently distributes, or a portal system that can be used by anyone, have reported that some engineering work will be needed to understand bottlenecks:
  - Data platforms (e.g., based on the Modern Research Data Portal[1] (MRDP)) have noted a more "random access" pattern that can lead to delays in fetching data from storage. Incorporation of caching may assist for this use case.
  - Data management systems (e.g., Rucio and others) are being adopted to manage data dissemination more centrally, and have exposed storage and networking bottlenecks.
- Use of the Globus platform for data movement continues to vary depending on the use case. Case studies that leverage grid computing paradigms will gravitate toward data managers native to that framework (e.g., Open Science Grid, Rucio, etc.), while experiments that use HPC facilities will adopt approaches that are native and supported (e.g., data portals).
- Investigation into the use of ML and AI approaches continues, and this leverages a mixture of commercial clouds as well as resources located at R&E HPC facilities. In some cases, a "hybrid" approach of using both (e.g., central processing unit-based work at an HPC facility, and graphics processing unit-based work at a cloud) can be used successfully, but in doing so network connectivity must be carefully monitored and planned for.
- The use of commercial cloud resources to support components of experimental workflows continues to be an area of active investigation. Data volumes, computational responsiveness, and network performance are all important factors.

[1] https://doi.org/10.7717/peerj-cs.144

- ESnet continues to work with sites, facilities, and R&E partners on planning to upgrade network connectivity to meet the data volumes required for projected experimental loads.
- ESnet continues to plan for the expansion of transatlantic network capacity to support LHC, DUNE, and other use cases.

3 Review Recommendations

ESnet recorded a set of recommendations from the HEP and ESnet requirements review that extend ESnet's ongoing support of HEP-funded collaborations. Based on the key findings, the review identified several recommendations for HEP, ASCR, ESnet, and ASCR HPC facilities to jointly pursue:

- ESnet will continue to participate with community members in discussions surrounding best common practices for data mobility and management. These include but are not limited to DUNE, CMB-S4, Rubin, DESI, and the LHC projects.
- ESnet will continue to discuss progress of the transatlantic capabilities with key HEP stakeholders.
- ESnet will continue discussions regarding site upgrades affiliated with ESnet6 commissioning.
- ESnet will continue to work with R&D efforts in the LHC community.
- ESnet will continue to work with the LHCONE and LHCOPN planning groups.
- ESnet will participate in DOE efforts such as IRI-ABA.
- ESnet will discuss some of the use cases that ESnet services can address related to ML and AI research.
- ESnet will continue to maintain and augment cloud connectivity options.

4 Requirements Review Structure

The requirements review is designed to be an in-person event; however, the COVID-19 pandemic has changed the process to operate virtually and asynchronously for several aspects. The review is a highly conversational process through which all participants gain shared insight into the salient data management challenges of the subject program/facility/project. Requirements reviews help ensure that key stakeholders have a common understanding of the issues and the potential recommendations that can be implemented in the coming years.

4.1 Background

Through a case study methodology, the review provides ESnet with information about:

- Existing and planned data-intensive science experiments and/or user facilities, including the geographical locations of the experimental site(s), computing resource(s), data storage, and research collaborator(s).
- For each experiment/facility project, a description of the "process of science," including the goals of the project and how experiments are performed and/or how the facility is used. This description includes information on the systems and tools used to analyze, transfer, and store the data produced.
- Current and anticipated data output on near- and long-term time scales.
- Timeline(s) for building, operating, and decommissioning of experiments, to the degree these are known.
- Existing and planned network resources, usage, and "pain points" or bottlenecks in transferring or productively using the data produced by the science.

4.2 Case Study Methodology

The case study template and methodology are designed to provide stakeholders with the following information:

- Identification and analysis of any data management gaps and/or network bottlenecks that are barriers to achieving the scientific goals.
- A forecast of capacity/bandwidth needs by area of science, particularly in geographic regions where data production/consumption is anticipated to increase or decrease.
- A survey of the data management needs, challenges, and capability gaps that could inform strategic investments in solutions.

The case study format seeks a network-centric narrative describing the science, instruments, and facilities currently used or anticipated for future programs; the network services needed; and how the network will be used over three timescales: the near term (immediately and up to two years in the future); the medium term (two to five years in the future); and the long term (greater than five years in the future).

The case study template has the following sections:

- **Science Background:** a brief description of the scientific research performed or supported, the high-level context, goals, stakeholders, and outcomes. The section includes a brief overview of the data life cycle and how scientific components from the target use case are involved.
- **Collaborators:** aims to capture the breadth of the science collaborations involved in an experiment or facility focusing on geographic locations and how data sets are created, shared, computed, and stored.

- **Instruments and Facilities:** description of the instruments and facilities used, including any plans for major upgrades, new facilities, or similar changes. When applicable, descriptions of the instrument or facility's compute, storage, and network capabilities are included. An overview of the composition of the data sets produced by the instrument or facility (e.g., file size, number of files, number of directories, total data set size) is also included.
- **Process of Science:** documentation on the way in which the instruments and facilities are and will be used for knowledge discovery, emphasizing the role of networking in enabling the science (where applicable). This should include descriptions of the science workflows, methods for data analysis and data reduction, and the integration of experimental data with simulation data or other use cases.
- **Remote Science Activities:** use of any remote instruments or resources used in the process of science and how this work affects or may affect the network. This could include any connections to or between instruments, facilities, people, or data at different sites.
- **Software Infrastructure:** discussion of the tools that perform tasks, such as data source management (local and remote), data-sharing infrastructure, data-movement tools, processing pipelines, collaboration software, etc.
- **Network and Data Architecture:** what is the network architecture and bandwidth for the facility and/or laboratory and/or campus? The section includes detailed descriptions of the various network layers (local area network, metropolitan area network, and wide-area network) capabilities that connect the science experiment/facility/data source to external resources and collaborators.
- **Cloud Services:** if applicable, cloud services that are in use or planned for use in data analysis, storage, computing, or other purposes.
- **Data-Related Resource Constraints:** any current or anticipated future constraints that affect productivity, such as insufficient data transfer performance, insufficient storage system space or performance, difficulty finding or accessing data in community data repositories, or unmet computing needs.
- **Outstanding Issues:** an open-ended section where any relevant discussion on challenges, barriers, or concerns that are not discussed elsewhere in the case study can be addressed by ESnet.

5 Review Summary

The following facilities and projects were featured in the 2020 HEP requirements review, and were asked to provide updates for 2021:

- LHC experimentation and operation
  - ATLAS experiment
  - CMS experiment
  - LHC operations
  - HL era of the LHC
- Neutrino experiments at Fermilab
  - SBN
  - DUNE
- Dark Energy Spectroscopic Instrument (DESI)
- Belle II experiment
- Muon experimentation at Fermilab
  - Muon g-2
  - Mu2e
- DESC
- The Rubin Observatory and the LSST
- CMB-S4
- Cosmological simulation research
- LZ Dark Matter Experiment

5.1 LHC Experimentation and Operation (e.g., ATLAS, CMS, and Shared Operations)

The ATLAS and CMS experiments indicated that LHC physics operations could not resume in 2021 as a result of the pandemic, and the next season of LHC running will instead begin in 2022 with a one-year extension to now end in December 2025. This next Run 3 will continue to provide larger data volumes and, to accommodate completing the installation and commissioning of the HL-LHC accelerator and detector upgrades, also lead to Run 4 (e.g., HL operations era) beginning in early 2029, as shown in Figure 1.

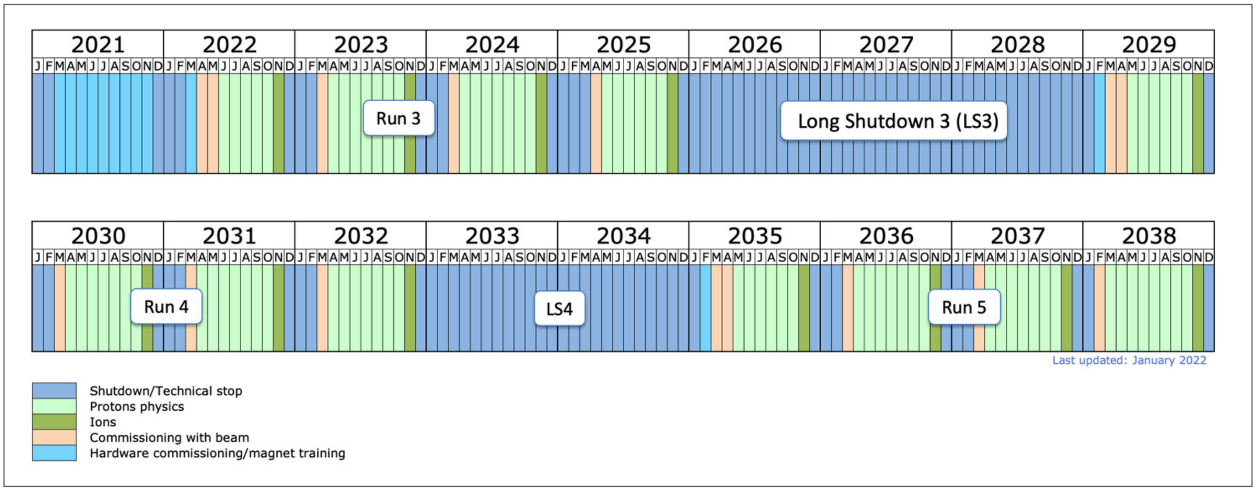


Figure 1: LHC and HL-LHC schedule

The CMS and ATLAS experiments report that they are preparing for the resulting shift in schedule in several important ways. Work on each experiments' hardware, computation, storage, and software infrastructure has resulted in some minor changes from the originally published case study.

R&D on operation of the ATLAS and CMS detectors has resulted in increased data rates and processing goals, namely through work to manage data coming off the detector during the "post trigger" phase of experimentation; previous experimentation may have not used as much of this data, but now it can be managed and will result in two times more volume for the archival storage component.

The CMS migration from Physics Experiment Data Export (PhEDEx) to Rucio has been completed, which has simplified operations for data management and also resulted in a reduction in transfer rates. The root cause is still being investigated, but early indications are that reduced physics operations are the primary reason, versus a fundamental change in the underlying networks or software packages. LHC Run 3 observations will help understand this phenomena further.

Great progress has been made towards developing a next generation system that coordinates the use of network and site resources in the framework of the Global Network Advancement (GNA) group's Data Intensive Sciences and the Software Defined Networking (SDN) for End-to-End Networked Science at the Exascale (SENSE) and AutoGOLE working groups[2]. A major milestone achieved in 2021-2 has been the establishment of two persistent testbeds with global expanse:

- A dynamic overlay network based on SENSE/AutoGOLE layer 2 overlays
- A fully programmable testbed at 30 sites using the GEANT Rare/freeRtr network operating system, as well as SONIC.

Other major milestones include:

- Progress in SENSE development, moving from research to a production-ready, multidomain system.
- Rapid progress towards the integration of SENSE with FTS and Rucio, in two efforts, for CMS and for ATLAS. For CMS, building on the SENSE virtual circuit foundation, flow steering is handled at layer 3 through the use of IPv6 subnets mapped onto each of the XRootD directors.

[2] https://www.dropbox.com/s/ainofw1lbxp5a91/SC22_NRE_GlobalPetascaleWorkflows_Update122022.docx?dl=0

- Migration towards 400Gbps network technologies extending to end sites, in order to develop the methods and systems needed between now and the start of the HL LHC era. 4 independent 100Gbps connections to the Caltech and UCSD/SDSC campuses supported by CENIC, Internet2 and ESnet, including virtual circuit extensions to StarLight and Fermilab, are one example. These are expected to migrate to 400Gbps connections and include additional 100Gbps connected Tier2 sites during 2023.
- Design of a system based on source-based routing (such as PolKA), where fully programmable networks and production networks can co-exist compatibly side-by-side, and where advanced network services can be moved progressively into production in a non-disruptive way.

ATLAS and CMS continue the pattern of annual reviews of projected resource needs. This was complicated during 2021 due to the shifting schedule (e.g., later start and longer run-time), but it has also allowed for a deeper analysis of the data-taking capabilities going forward. Both experiments have revisited "pileup" volume estimates for LHC data collection which are now expected to increase in 2035 during Run 5 operation, versus an earlier target of 2027 during Run 4. This is not attributed to the luminosity targets, but rather the changes to the science. ATLAS and CMS will address this via an exercise tentatively to be called "Pileup2000" and will feature some ramp-up of preparedness like the data challenges that the LHC community has attempted in previous years.

The Research Networking Technical Working Group (RNTWG), formed in response to the January 2020 LHCONE/LHCOPN joint meeting[3] with the WLCG experiments, has focused on enabling network flow visibility anywhere in the network. This is the first of three primary work areas for the RNTWG. The goal is to allow anyone to find the owner and purpose of any research and education flow anywhere in the network. To do this the RNTWG created a subgroup on packet marking. During 2021-2022 the focus was on developing packet and flow marking capabilities to provide the needed visibility for R&E network traffic. A technical specification document has been created for the flow and packet marking work[4]. In early 2022, a service was created which supports storage and transfer activities in marking netwwork flows or packets. The medium term goal is to have a significant amount of LHC traffic identified by owner and activity in time for the next WLCG Data Challenge, currently estimated to be in March of 2024.

ATLAS and CMS continue to prepare for computation, storage, and network upgrades around their ecosystem. They are actively engaged with ESnet on plans for transatlantic upgrades, as well as the upgrades of T2 facilities (campuses and regional networking infrastructure). ESnet continues to work with the LHC community on data challenges in 2022 and 2023. Data Challenge 2 (in 2023) may see a slight schedule shift as the LHC detector operations may shift.

5.2 Neutrino Experiments at Fermilab

Both SBN and DUNE experienced delays due to the pandemic in terms of supply chain issues and shifting schedules for construction and operation. Both case studies were originally presented using pre-pandemic timelines, and those are now adjusted for the reality of the delays.

5.2.1 SBN

SBN features three liquid Argonne detectors in its design. In terms of construction and operations of the SBN detectors, the largest has now completed commissioning work and will transition into data taking and operation. This will be an approximately year-long exercise. The new detector is still under construction, and is not expected to come online until 2023.

[3] https://indico.cern.ch/event/828520/
[4] https://docs.google.com/document/d/1x9JsZ7iTj44Ta06IHdkwpv5Q2u4U2QGLWnUeN2Zf5ts/edit#heading=h.2msfykqhodwc

SBN reports that MicroBooNE has now reached analysis mode after a one-year delay due to the pandemic: it will not take any more data, and will focus on the task of analyzing the results. The overall process of science has not experienced any change beyond timeline delays.

SBN computation and storage infrastructure has seen increases in capability: namely adding 2 PB more disk-based storage (dCache) to deliver data products at an increased rate. With more disk storage available, there is less need to rely on the slower tape archival system for reading raw (or derived) data during analysis. This could increase network requirements over time by a modest amount, but there is not enough experience to fully understand the impacts of this change.

R&D efforts to develop workflows that leverage HPC resources for simulation continue, as well as efforts to manage data streaming and reconstruction workflows. These are leveraging the Theta resource at the ALCF.

Overall, the progress of the software infrastructure is continuing, with a series of demos planned in 2022 and 2023. It is expected to enter production by 2023. The next one to two years will be critical for Icarus and MicroBooNE data processing. Current development will affect long-term data analysis functionality.

5.2.2 DUNE

While DUNE experienced delays and shifting schedules, progress was made in several key areas including the analysis framework, the model for processing, and approaches to data mobility. As DUNE prepares for operations, it has adjusted expectations for many aspects of the data chain and is exploring the adoption of technologies in use for other large-scale efforts like the LHC.

The current distributed processing model for DUNE will feature 10 sites that act as storage and 50 that are involved in computation: overall, there will be 20 PB of disk between all facilities. Rucio is expected to be the tool used for data mobility and placement between these facilities. Each compute facility will perform copy of data from a disk site during processing. All primary archival storage (e.g., tape) will occur at Fermilab, with a secondary archive being distributed among several European facilities. Reconstruction activities will leverage a model like the Worldwide LHC Computing Grid (WLCG) from the LHC, with the reconstructed copies being returned central to Fermilab and distributed to participating sites. It is anticipated there will be two copies of all data (worldwide) for a one- to two-year timeframe to facilitate distributed analysis. One area of consideration is a partnership with EU sites, and the impacts that data reads/writes from archival storage will have on wide-area data transfer rates.

The WLCG-like approach will leverage DOE, university, and international partner sites in the future, meaning that network resources should be considered carefully via data challenge exercises. DUNE is actively testing out data transfer capabilities for the overall workflow. One limitation is that the data format used (HDF5) does not facilitate streaming workflows, so transfers will be bulk during operations. The impacts of implementing a network-based caching solution are being considered to help alleviate some of the pressure on this part of the DUNE workflow, and discussion occurred during the presentation on ways this may be integrated.

DUNE presented some results of R&D on job operation. Experience has shown that the tooling is making the correct choices about data transfer locality (e.g., computing close to where data resides), and early indications are that 15–20 MBps network speeds can be seen during analysis jobs when fetching. Caching could certainly offer some speedup, and will be investigated.

DUNE expects first light by 2028 or 2029. The schedule may shift more. Data challenges will start prior to this; some will begin in June 2022. The expectation for this is to experience CERN to Fermilab flows and reach 40 Gbps rates over a five-day period. ProtoDUNE experience is ramping up slowly to full-DUNE data rates, but will not run for the longer time periods that DUNE expects to.

DUNE ended with some reminders on data volumes and expectations:

- By 2029, the DUNE supernova detection components are expected to produce 160 TB of data in around 100 seconds.
- DUNE will produce 100 TB a day during normal operations starting in fall of 2022.
- DUNE will produce around 30 PB of data a year throughout the ecosystem.

5.3 DESI

DESI operations have remained stable since the writing of the case study, with some minor friction that the project is experiencing related to the tools used to deliver data to the collaboration; in particular, several bottlenecks related to discovery and access of data were discovered as more users have started to use the analysis framework. Over the previous year, the project has released a public data set, and has found that the access patterns, which can be considered "random access," have caused many problems for the way the data is stored and retrieved. In some cases, this could manifest as a denial of service (DOS)-like form of cyberattack on the database systems. The team members continue to work on this as they prepare for more public data releases. The interface that is exposed to such data releases may need to be re-engineered to something that can handle the expected volume of requests, along with the semi-random way they may be accessed.

DESI continues to use rsync to replicate data sets between NOIRLab at the University of Arizona to NERSC over the network connectivity between the sites (Sun Corridor network, Interent2, and ESnet). In practice, this system has been found to meet expectations for replication speed and ease of use, except in cases where deep nesting of directory structures may take a long time to be discovered and synced between sites. The group has explored Globus, but found that symbolic links in the data could cause the Globus software some friction, resulting in partial synchronization.

DESI expects no changes to data volumes or process of science over the short to medium term.

5.4 Belle II Experiment

Belle II has implemented the multi-site model to handle archiving of data, as discussed in the case study. This new pattern of data distribution did not result in any significant change to the overall volume of data flowing from the instrument, or across networks to the collaborating sites. A second change to the data distribution mechanisms (e.g., the adoption of Rucio as used in other HEP experiments) has also been implemented, with BNL serving as the central coordination site for data movement; the adoption of the tool was without major complications.

The accelerator itself will be entering into a multi-month shutdown phase for upgrades starting in 2022, which will result in higher data rates in 2023/2024. A second shutdown is still planned for the 2026/2027 time frame.

5.5 Muon Experimentation at Fermilab

Scientists from Fermilab gave updates on the two major muon experiments: Muon g-2 and Mu2e. Both are located at Fermilab, and will follow a similar paradigm for operations, computing, storage, and networking.

5.5.1 Muon g-2

Muon g-2's initial case study could not definitively discuss operational experience since the report was written slightly before the experiment started run-time. The one-year update to the case study has seen Muon g-2 running for a little over a year, taking experimental data, and several papers are being published with the resulting observations. Due to pandemic-related delays, a modification to the schedule occurred, which was updated and appears in Figure 2.

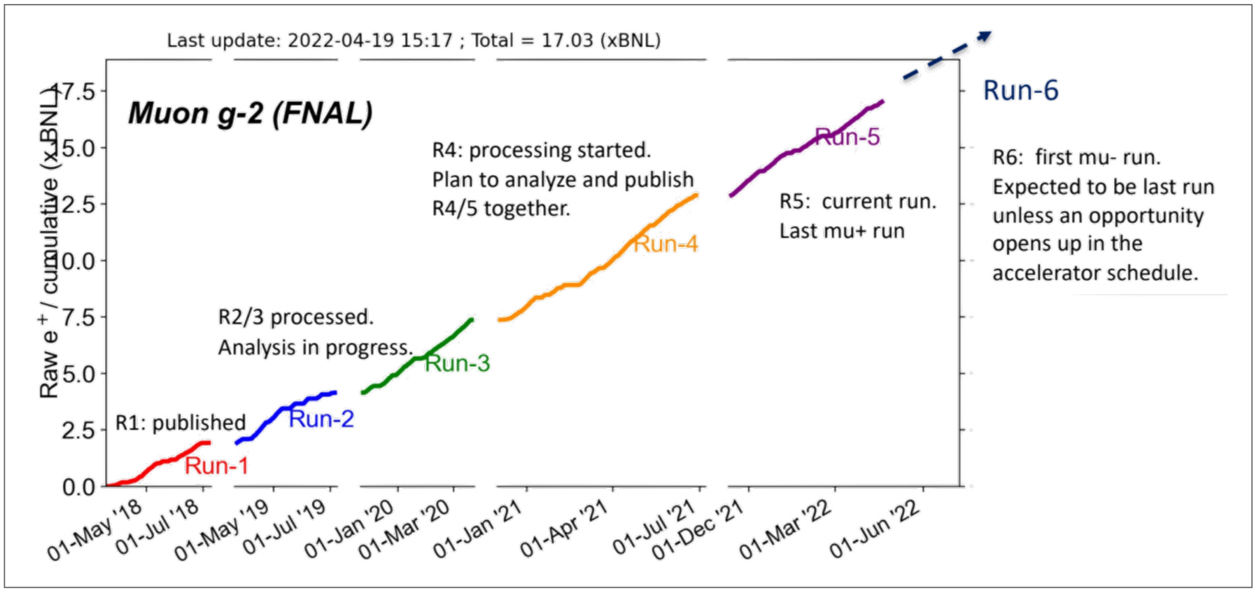


Figure 2: Muon g-2 schedule

Several factors went into the schedule modifications:

- Some aspects of the experimental timeline require shutdown and reconfiguration of the instrument (e.g., changing polarity of magnets). As an example, an experiment proposed by collaborators in Japan can operate with only one form of polarity; thus it will be set to run before changing anything to accommodate other experimentation.
- It is expected that production analysis will continue into fiscal year 2024, with published results in fiscal year 2025.

These changes to the schedule will not affect the models used for computation, which will remain steady throughout. The production software has been frozen, and the only potential disruption could be the use of additional grid computing hardware resources during run-time (local or remote), but it is not expected this will have a noticeable difference on outcomes.

One minor change that could leverage network resources would be leveraging computational resources at international partner sites (e.g., in Europe), in particular running reconstruction workloads and testing new algorithms. The new approaches are leveraging emerging work in ML, and use a different (nonproduction) set of software. To accomplish this testing, it is expected that raw data will be transferred from Fermilab to European resources, and then results would be retrieved. Exact data volumes are unknown, but are estimated to be on the order of hundreds of TB and may have a bottleneck due to reading from storage versus any known network problem.

G-2 continues to leverage HPC resources at NERSC, namely for simulation production. As discussed in the report, the input data is very small, but the large (e.g., TB) sized output data must be transferred back to Fermilab. This will continue during the length of the project run.

The aforementioned changes (and the longer run-time of the experiment) will not change the overall expected data volumes on an experiment, yearly, or project basis that were enumerated in the case study.

5.5.2 Mu2e

As noted in the case study, Mu2e uses the technology footprint that g-2 is currently using for operations, and will therefore not start until after g-2 completes. As mentioned in Section 2.6.1, this is now expected to be late 2024 or early 2025 (e.g., first beams), given the adjustments to the timelines due to the pandemic. Overall construction has been slowed, and current progress is around 70% of where it was originally expected to be based on pre-pandemic estimates.

Mu2e is not expecting changes to the overall data volume of the experiment:

- Initial calibration runs will produce hundreds of TBs with an estimated 1 PB being sent to European partners for processing. Large data volumes will not be returned to Fermilab.
- Full operation could extend to 2031, and will produce in excess of 7 PB on a yearly basis, which will leave Fermilab and be sent to partners in the global grid computing model.

Additional changes to the schedule are possible, and will depend on the progress and run-time of g-2.

5.6 The Rubin Observatory and the LSST and DESC

The Rubin Observatory, the LSST, and DESC all leverage shared instrumentation deployed in Chile and are working to construct an end-to-end workflow that will transport data from deployment site to SLAC in the United States, as well as a cloud-based platform operated by Google. This hybrid model is described in Figure 3, and the networking is described in Figure 4.

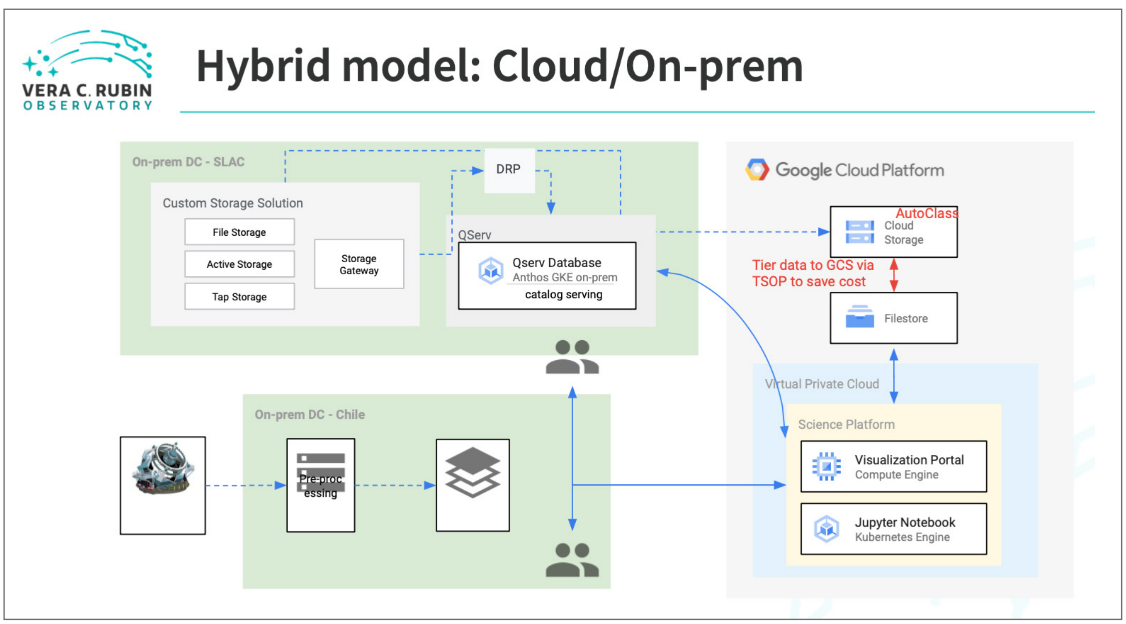


Figure 3: Rubin architecture

Details are emerging on the specifics of the hybrid computing model that will leverage Google Cloud, but the current planning and experimentation is investigating the following:

- The international data path will still leverage connections provided by the National Science Foundation International Research and education Network Connections (NSF IRNC) program via Florida International University / AMPATH.
- The cloud platform is being designed to serve as a "user analysis" solution for some collaborators, as well as the general public.
- It is not expected that full image data will be stored on the cloud platform due to size and complexity. At this time, it is hard to predict what the total volume of data storage in the cloud will grow to. See Figure 5 for expected data rates.
- Data served to DOE collaborators (e.g., DESC) will come from SLAC directly, and in most cases go to NERSC (where DESC operates).
- As much as 50% of the data processing could be done by international collaborators (e.g., located in Europe). Results of the processing will be sent from, and have to be returned to, the storage resources at SLAC.
- Data processing will not begin until at least December of 2024, but it is possible additional pandemic delays may delay deployment and data challenge activities.

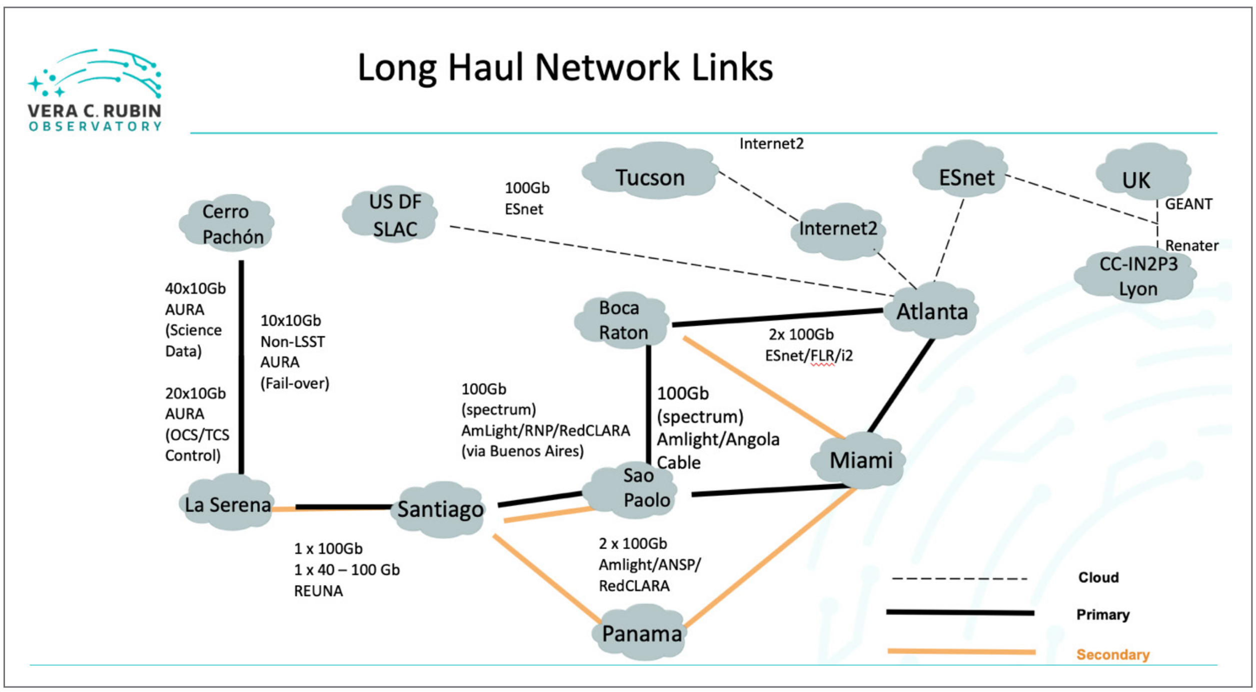


Figure 4: Rubin wide-area networking

Usage and access patterns for the data (from Google Cloud or SLAC) are still being investigated, and cannot be predicted. The largest data volumes are still expected to flow from the instrument to SLAC, and then from SLAC to other locations (e.g., DOE sites, Google Cloud, other collaborators). Production data-taking activities are expected to flow to SLAC directly.

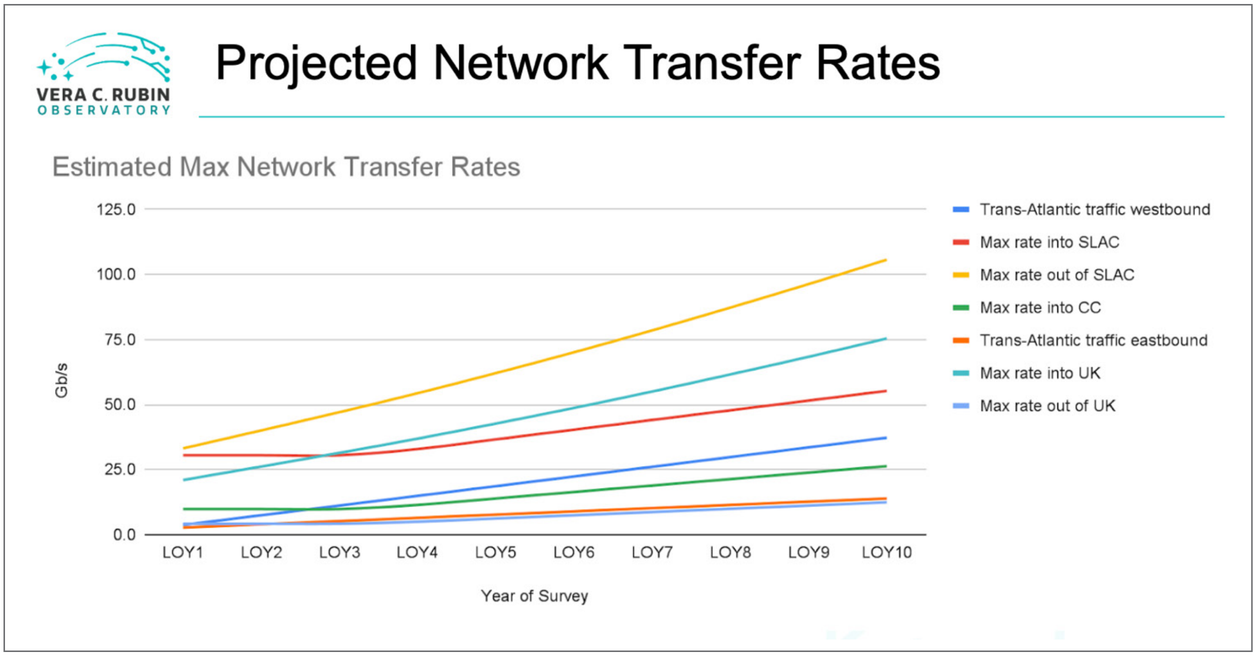


Figure 5: Rubin data rates

Additional updates on the DESC were not made available.

5.7 CMB-S4

A representative from CMB-S4 was not present, but members of the ESnet Science Engagement Team provided the following information based on prior discussions:

- The overall experimental schedule will shift due to COVID delays and has not been finalized.
- The CMB-S4 team members started a series of data challenges with their computing partners at NERSC. The goal of this early investigation is to fully exercise the storage and computational infrastructure, via the generation of simulation data. 1 PB of simulations are being generated and run through the analysis software and hardware infrastructure.
- CMB-S4 has had discussions about ways to disseminate the data from this instrumentation, and will work with NERSC to deploy a dedicated "portal" based on the concept of the MRDP that uses dedicated data-movement hardware and the Globus data transfer platform.
- The project is still investigating a data management and dissemination platform, and is exploring the use of Rucio. No choice has been made at the time of the update

5.8 Cosmological Simulation Research

An update on cosmological simulation research was not made available.

5.9 LZ Dark Matter Experiment

An update on the LZ Dark Matter Experiment was not made available.

6 Review Discussion

The discussion during the two webinars focused primarily on:

- Pandemic-related delays to the supply chain and staffing shortages.
- Data volume changes that result either from new approaches to the science, or improvement to the overall workflow.
- R&D efforts, and how these will transition to production.
- The ongoing data challenges used by scientific communities (planned, and planning).
- Data dissemination and sharing via different tools and platforms.
- Emerging use cases that leverage cloud resources.
- The status of ESnet upgrades that will increase capacity and reachability to locations around the world.

The following sections offer some of the content that was discussed related to these topics.

6.1 Pandemic Delays

A universal truism for HEP efforts is that the pandemic has increased timelines, caused shortages in staff and components, and generally disrupted plans that were discussed a year ago. Almost all facilities and experiments have made plans, or are planning, for schedule and resource changes as a result.

6.2 Data Volume Changes

Most of the original estimates given in the case studies for data volumes have held steady. A few case studies have made modest changes to their project total volumes due to the LHC and HL-LHC schedule changes mentioned in Section 5.1. The LHC experiments are planning to start later, and potentially run longer, which will increase the baseline data expectations.

The LHC experiments have reported that changes to their underlying triggering architecture, as well as certain changes in analysis, will result in some data reduction capabilities over time. This will be more noticeable during Run 3.

6.3 R&D Efforts

Despite any operations schedule updates, as mentioned in Section 5.1, many R&D efforts continue to move forward and make progress and are entering into production footing around the HEP community. For example, the LHC experiments are making progress in the use of dynamic networking, measurement and monitoring, and data caching as they prepare for Run 3 resuming in 2022. DUNE continues to experiment with analysis approaches on the software and hardware side. Rubin, and DUNE, continue to investigate the use of commercial clouds (see Section 5.6).

6.4 Data Challenges

LHC Run 4, also known as the start of the High-Luminosity LHC era or HL-LHC, is expected to produce approximately 10 times more data than LHC Run 3, which begins in the spring of 2022. Stable LHC Run 4 operations are currently projected to begin by the summer of 2029. In order to prepare for the factor of 10 increase in global performance, the LHC community is conducting four data challenges to ensure that the community is prepared for HL-LHC at the start of Run 4. Each data challenge runs like a dress rehearsal, with specific performance targets to be achieved by all participants. Data Challenge 1 was conducted in the first two weeks of October 2021. The first week was the Network Challenge, which involved CERN, Tier 1, and Tier 2 sites. The second week was the Tape Challenge, which involved sites with tape systems (in the US, this is primarily BNL for ATLAS and Fermilab for CMS). The performance target was 10% of HL-LHC needs (which is equivalent to 100% of Run 3 needs). Data Challenge 1 was a success, both from an ESnet perspective and from a community perspective.

Figure 6 shows the ESnet network experienced significant load during the data challenge. Data Challenge 2, in 2023, is expected to produce three times as much load at a minimum. ESnet has capacity enhancements for transatlantic connectivity in the pipeline, which are expected to exceed the near-term needs of both Run 3 and Data Challenge 2. In the months since Data Challenge 1, the ESnet6 circuit and router deployment projects have been completed, dramatically increasing the terrestrial bandwidth of the ESnet backbone network. Note that one of the transatlantic circuits is idle in the screenshot: the cable carrying the circuit experienced a fault and that circuit was down before, during, and after the data challenge.

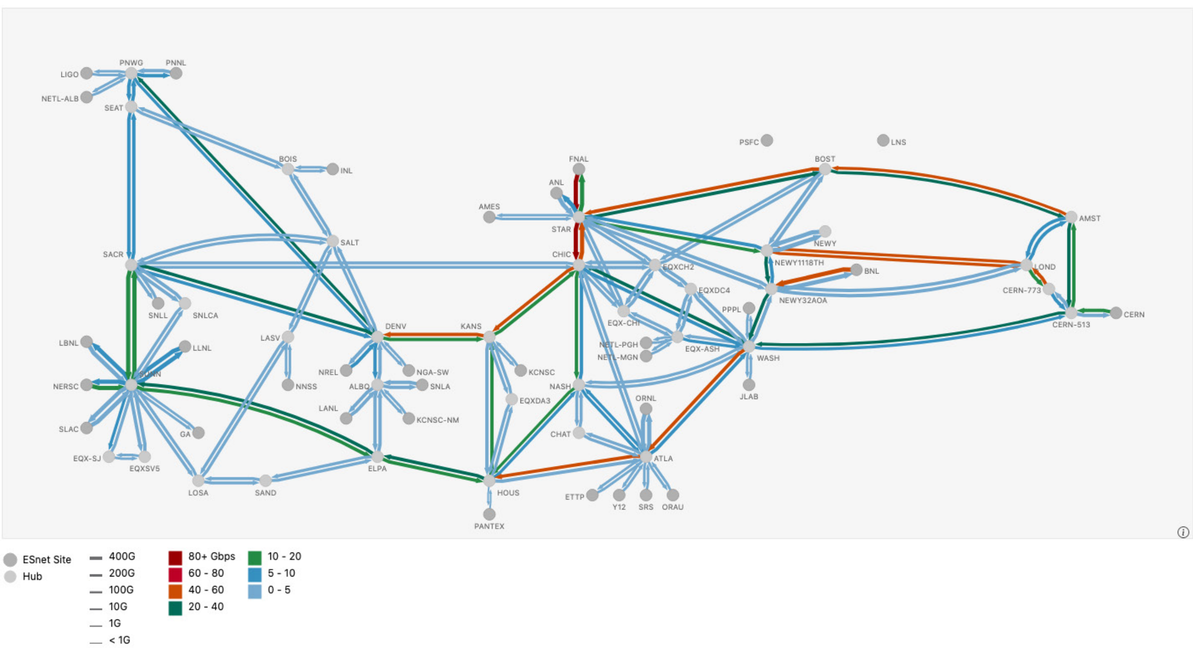


Figure 6: ESnet portal

During the LHC Blueprinting meetings before the data challenge, LHC community leaders brought up one of the goals of the data challenge from the perspective of the LHC experiments: the collection of network statistics to help with analysis after the data challenge. The ESnet Software and Engineering groups put together a statistics portal based on ESnet's new Stardust platform which provided a wealth of information. This was very successful, and the ideas will be further developed for Data Challenge 2, as shown in Figure 7.

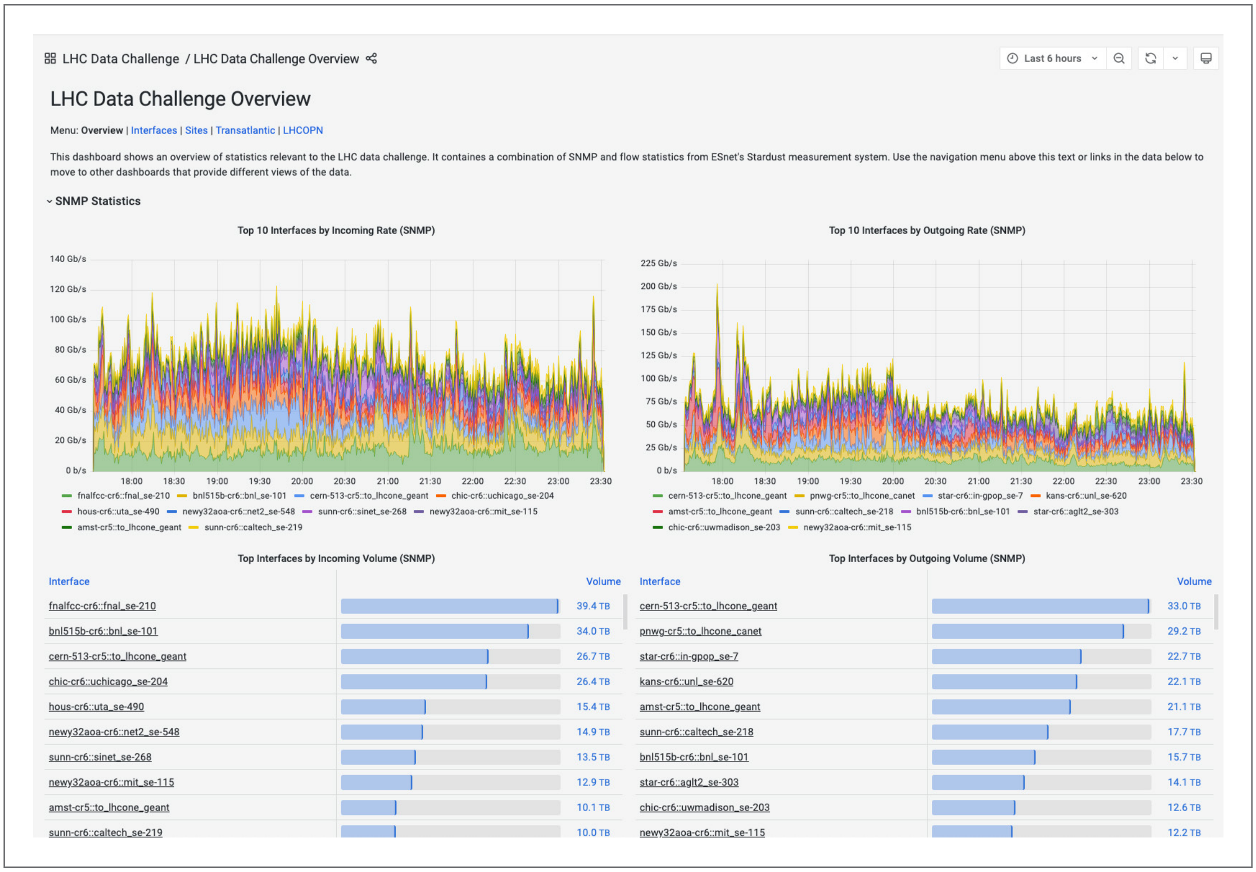


Figure 7: LHC data challenge via ESnet Stardust

As DUNE and Rubin get closer to operations, both have started structured activities to understand the capabilities of the wide-area networking paths between their affiliated facilities. These activities will increase in volume, and frequency, in the coming years.

CMB-S4 has started an effort to understand data constraints, and has been focused on storage and computation along with early considerations for wide-area data movement from the remote sites back to NERSC.

6.5 Data Dissemination and Sharing

During the pandemic, some experiments have modified the underlying infrastructure used to support their scientific workflows that handle data mobility. Rucio is now used by the LHC experiments and Belle II as the standardized way to manage data movement between participating facilities. Several other projects are investigating possible use of this tool as well.

Others are continuing to define what future data-movement tools may make sense for both internal use cases, as well as those that may face the general public in the form of portals. CMB-S4 continues to work with NERSC and ESnet, and has been exploring the use of the MRDP approach that leverages Globus.

6.6 Emerging Cloud Use Cases

Several HEP use cases leverage cloud resources for some of their prototyping and R&D efforts:

- Rubin is in the process of creating a cloud-based infrastructure to prepare for the future science needs of the project.
- DUNE has done similar work to experiment with the use of cloud resources.
- Several LHC-affiliated R&D efforts continue to explore how commercial clouds may offer some capabilities beyond traditional computing models.

6.7 ESnet Upgrades

An update on ESnet6 deployment was given by ESnet network engineering, including how this will enhance some of the connectivity that supports the HEP community. A plan to upgrade the LHC Tier 1 computing centers, connectivity to major exchange points that support LHC Tier 2 centers, and impacts to peering with international partners were discussed in some detail.

ESnet also briefed on the discussions surrounding the plans for transatlantic networking to support science in Europe, and offered a rough timeline for when these upgrades would be made to support ongoing HEP efforts.

List of Abbreviations

AI	artificial intelligence
ANL	Argonne National Laboratory
ASCR	Advanced Scientific Computing Research
ATLAS	A Toroidal LHC ApparatuS
BNL	Brookhaven National Laboratory
CMB-S4	Cosmic Microwave Background — Stage 4
CMS	Compact Muon Solenoid
DESC	Dark Energy Science Collaboration
DESI	Dark Energy Spectroscopic Instrument
DOE	Department of Energy
DUNE	Deep Underground Neutrino Experiment
Fermilab	Fermi National Accelerator Laboratory
FTS	File Transfer Service
Gbps	Gigabytes per second
HEP	High Energy Physics
HL	High-Luminosity
HPC	high-performance computing
IRI-ABA	Infrastructure Architecture Blueprint Activity
LBNL	Lawrence Berkeley National Laboratory
LHC	Large Hadron Collider
LHCONE	Large Hadron Collider Open Network Environment
LHCOPN	Large Hadron Collider Optical Private Network
LSST	Legacy Survey of Space and Time
LZ	LUD-ZEPELIN
MBps	Megabytes per second
ML	machine learning
MRDP	Modern Research Data Portal
NERSC	National Energy Research Scientific Computing Center
R&D	research and development
R&E	research and education
RNTWG	Research Networking Technical Working Group
SBN	Short-Baseline Neutrino
SC	DOE Office of Science
SDN	Software Defined Networking
TB	terabyte
WLCG	Worldwide LHC Computing Grid

22-SN-35039