WLCG Challenges

S. Campana (CERN)

Introduction

For HEP software and computing the time horizon of future challenges is the next 15 years

The main contributor to those challenges is HL-LHC, both in terms of volume and complexity. The largest needs come from ATLAS and CMS

The LHC computing resources are provided by the WLCG infrastructure. Other HEP experiments will share a large part of such an infrastructure. Other sciences will use many of the same facilities

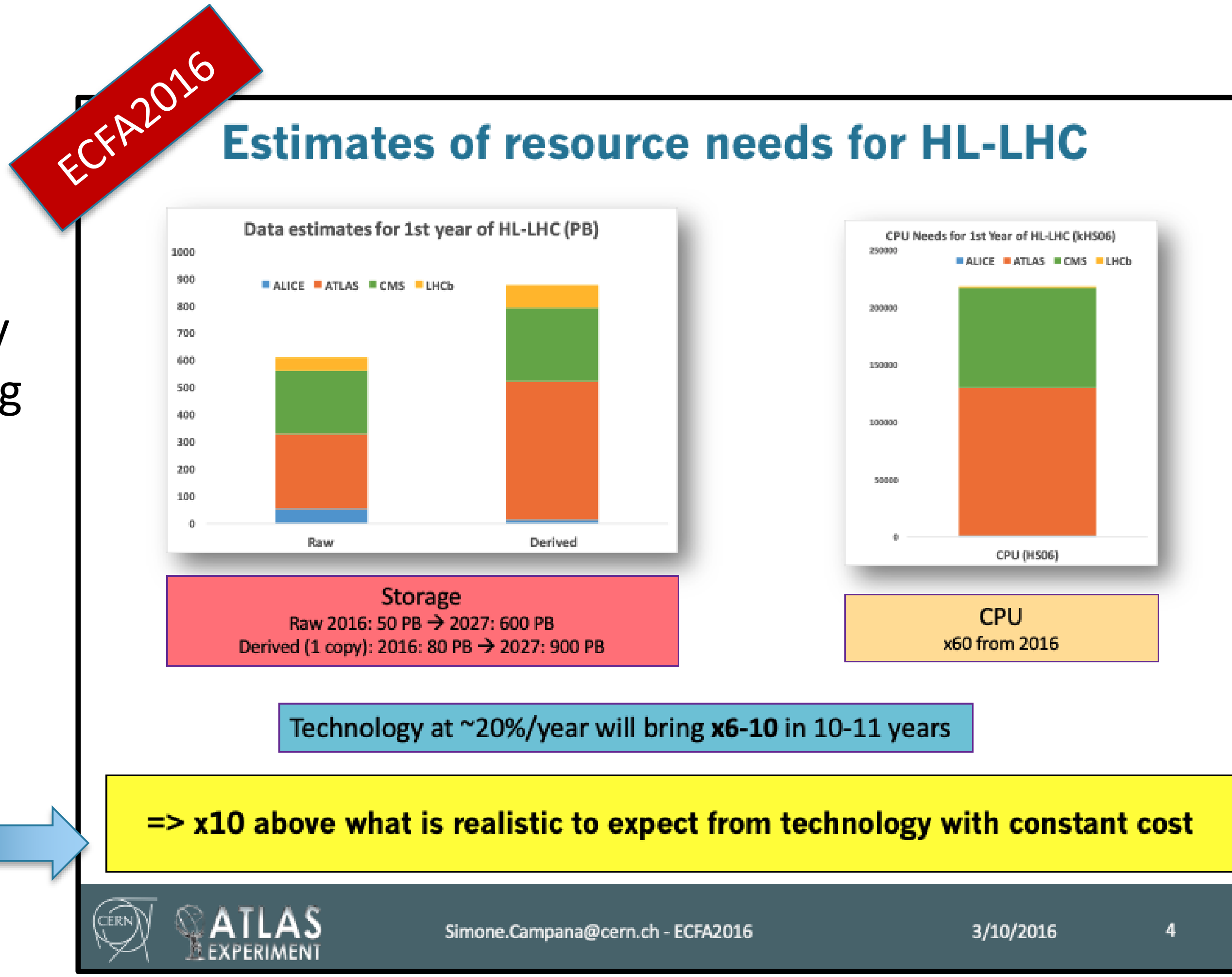
The HL-LHC challenge statement at ECFA 2016

In a nutshell:

- Projections assume constant funding every year for LHC computing

- Technology improvements will bring ~20% more resources every year

And this was the initial conclusion



The HL-LHC computing roadmap process

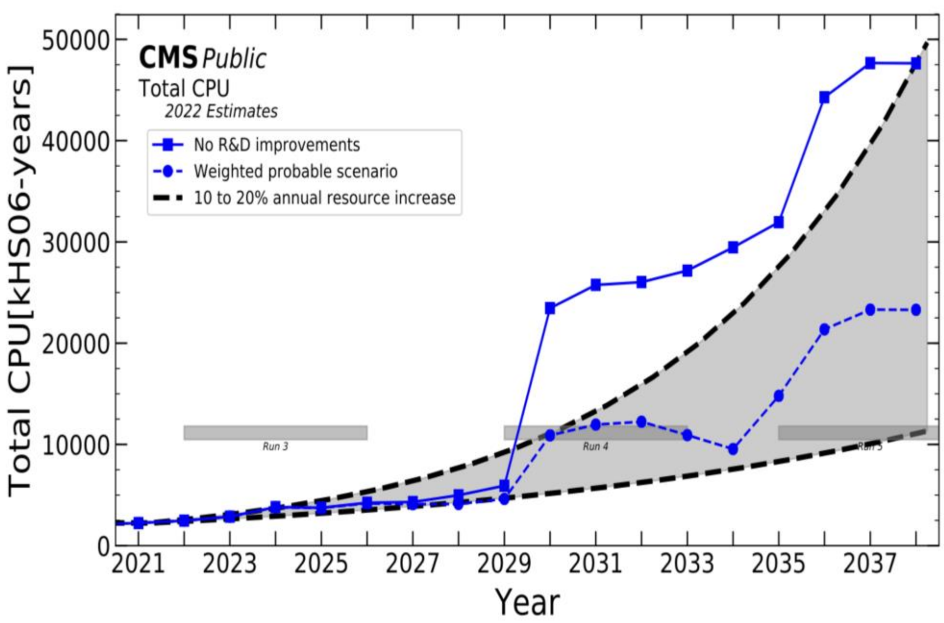
HEP Software Foundation Community Whitepaper: a bottom-up exercise. Identify the areas of work to address the HEP challenges of the 2020s

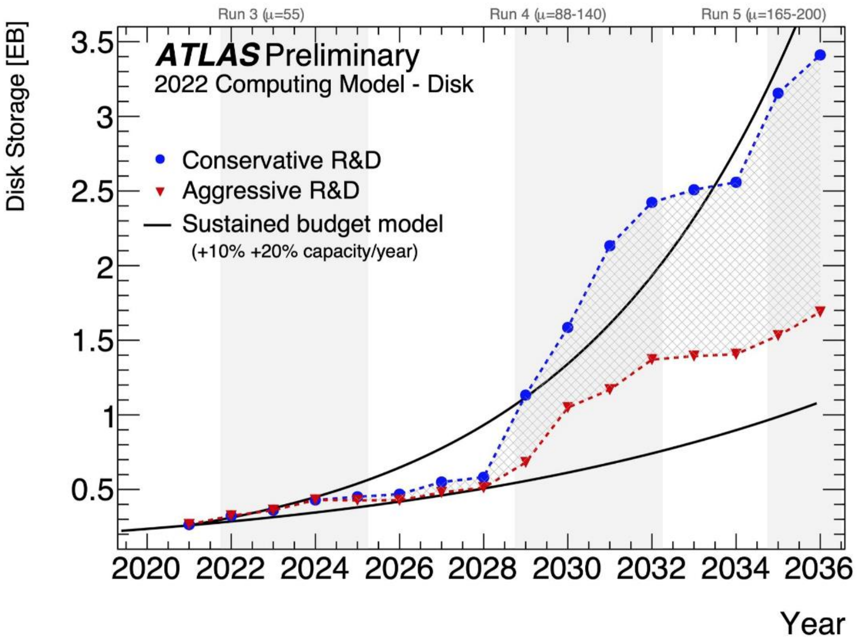
The first WLCG strategy toward HL-LHC document: a top-down high-level prioritization of the whitepaper, for the LHC needs

The LHCC review series of HL-LHC computing: a multistep process tracking the progress towards HL-LHC

- May 2020: review of ATLAS and CMS plans, Data Management (DOMA), offline software, the WLCG collaboration and infrastructure. Documents
- November 2021: update from ATLAS and CMS, common software activities (generators, simulation, foundation software, analysis, DOMA). Report

ATLAS and CMS needs for HL-LHC





- The gap between available and needed resources is filling up, assuming the main R&D activities are successful
- Investing in further (identified) R&D activities would fill this gap further. Need more effort
- There are still large uncertainties

Networks' role in HL-LHC

Networking will play a central role in HL-LHC as enabler for HEP computing

- Support the core functions of WLCG (data acquisition/archival/processing)
- Provide more flexibility to the computing models, allowing to optimise

WLCG continues engaging with Funding Agencies and NRENs to ensure that enough capacity is made available and the LHC traffic does not get segregated below a critical level.

Several R&Ds were launched to study how to better leverage the network resources in the data and processing infrastructures for HL-LHC

- Regularly discussed at the LHCONE/LHCOPN meetings

The LHCC sees the strategic role of networks and asked for regular updates. Tony Cass was asked to prepare a contribution for the next meeting (Nov 29th)

WLCG network data challenges

The data challenges are an incremental **process** to prepare for the HL-LHC network needs, through a regular dialog between the network providers, the experiments and the facilities.

We identified the main use cases at HL-LHC in terms of network use (RAW data export and reprocessing), for the 4 LHC experiments

We estimated the network needs including contingency and considering different scenarios

We set metrics and intermediate targets to be progressively challenged

The challenges offer the possibility to bring in production many network R&D activities

Network Data Challenges

from the WLCG Network Challenges document

T1	LHC Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Flexible Scenario in 2027	Data Challenge target 2027 (Gbps)	Data Challenge target 2025 (Gbps)	Data Challenge target 2023 (Gbps)	Data Challenge target 2021 (Gbps)
CA-TRIUMF	200	400	100	60	30	10
DE-KIT	600	1200	300	180	90	30
ES-PIC	200	400	100	60	30	10
FR-CCIN2P3	570	1140	290	170	90	30
IT-INFN-CNAF	690	1380	350	210	100	30
KR-KISTI-GSDC	50	100	30	20	10	0
NDGF	140	280	70	40	20	10
NL-T1	180	360	90	50	30	10
NRC-KI-T1	120	240	60	40	20	10
UK-T1-RAL	610	1220	310	180	90	30
RU-JINR-T1	200	400	100	60	30	10
US-T1-BNL	450	900	230	140	70	20
US-FNAL-CMS	800	1600	400	240	120	40
(atlantic link)	1250	2500	630	380	190	60
Sum	4810	9620	2430	1450	730	240

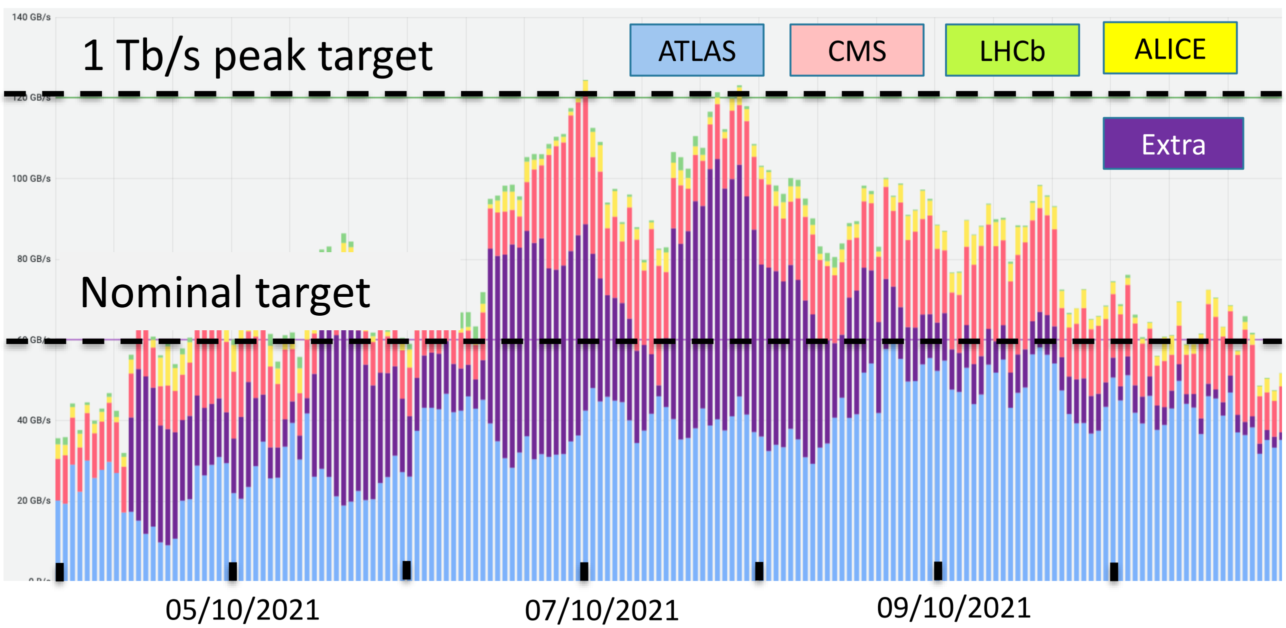
T1	%ATLAS	%CMS	% Alice	% LHCb	ATLAS+CMS Network Needs (Gbps) Minimal Scenario in 2027	Alice Network Needs (Gbps) Minimal Scenario in 2027	LHCb Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Flexible Scenario in 2027
CA-TRIUMF	10	0	0	0	200	0	0	200	400
DE-KIT	12	10	21	17	450	80	70	600	1200
ES-PIC	4	5	0	4	180	0	20	200	400
FR-CCIN2P3	13	10	14	15	450	60	60	570	1140
IT-INFN-CNAF	9	15	26	24	480	110	100	690	1380
KR-KISTI-GSDC	0	0	12	0	0	50	0	50	100
NDGF	6	0	8	0	110	30	0	140	280
NL-T1	7	0	3	8	140	10	30	180	360
NRC-KI-T1	3	0	13	5	50	50	20	120	240
UK-T1-RAL	15	10	3	27	490	10	110	610	1220
RU-JINR-T1	0	10	0	0	200	0	0	200	400
US-T1-BNL	23	0	0	0	450	0	0	450	900
US-FNAL-CMS	0	40	0	0	800	0	0	800	1600
(atlantic link)					1250	0	0	1250	2500
Sum	100	100	100	100	4000	400	410	4810	9620

- Numbers referring to ingress and egress (disjoint)
- Total ingress+egress
- Minimal: hierarchical model T0-T1-T2 traffic
- Flexible: chaotic model currently more realistic

2021 network data challenges

We executed data challenges testing the WLCG network and archive storages in preparation for Run-3 (Oct 2021)

The data challenges are an end-to-end test including storage, protocols, networks, data management services (e.g. Rucio, FTS)



Targets were met:

- the nominal transfer rate was sustained
- the peak transfer rate were reached

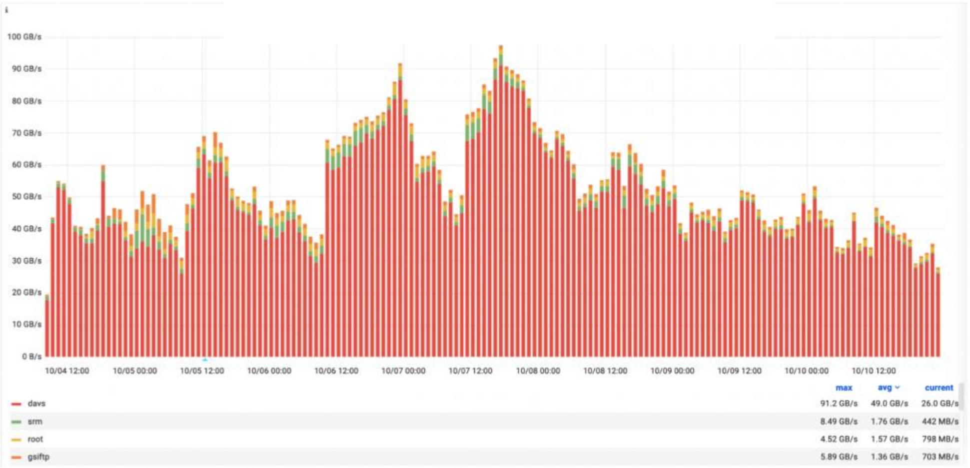
2021 network data challenges

The data challenges are not just about throughput but also functionality

In 2021 they provided an opportunity to commission new features that are now in use during Run-3.

- For example the HTTP protocol (replacing gridFTP) for asynchronous transfers

Transfer Rate per protocol



Traffic mostly through HTTP (**RED**)

CERN-PIC NOTED

NOTED is a Software Defined Network R&D project to share network traffic between different paths

Enabled during the data challenge between CERN and PIC

When the 6 Gbps LHCOPN link saturated, NOTED added the LHCONE link to complement it.

10Gbps target reached

The "2023" data challenges

The "2023" data challenges will likely happen in 2024 as it adapts better to the LHC schedule. Mario will give a dedicated talk later in this event

The target rates are known from the original document but what do we want to test should be discussed now

From the network perspective, take the opportunity to discuss at this meeting. What are the R&Ds that could be demonstrated in early 2024? **Set targets and define a plan.**

For example:

- Further demonstration of the NOTED capabilities?
- Can we monitor the traffic through packet marking for a few segments?
- Demonstration of the SENSE technologies at a large scale than the current testbed?

Collaboration with other HEP experiments

WLCG presented a joint paper with DUNE and Belle-2 to the Snowmass 2021 process

The paper presents the strategic directions to address the computing challenges of the experiments in the next decade. It complements the WLCG contribution to the European Strategy for Particle Physics in 2019

Snowmass 2021

- Consolidation of the WLCG scientific computing infrastructure
- Evolution of such an infrastructure to integrate modern technologies and facilities
- **Broadening the scope of the WLCG collaboration to create partnership with other HEP experiments**

Today DUNE,Belle-2 and JUNO are WLCG "observers" and share many services with WLCG (including some LHCOPN/LHCONE networks)

Physics > Computational Physics

[Submitted on 14 Mar 2022]

HEP computing collaborations for the challenges of the next decade

Simone Campana, Alessandro Di Girolamo, Paul Laycock, Zach Marshall, Heidi Schellman, Graeme A Stewart

Large High Energy Physics (HEP) experiments adopted a distributed computing model more than a decade ago. WLCG, the global computing infrastructure for LHC, in partnership with the US Open Science Grid, has achieved data management at the many-hundred-Petabyte scale, and provides access to the entire community in a manner that is largely transparent to the end users. The main computing challenge of the next decade for the LHC experiments is presented by the HL-LHC program. Other large HEP experiments, such as DUNE and Belle II, have large-scale computing needs and afford opportunities for collaboration on the same timescale. Many of the computing facilities supporting HEP experiments are shared and face common challenges, and the same is true for software libraries and services. The LHC experiments and their WLCG- partners, DUNE and Belle II, are now collaborating to evolve the computing infrastructure and services for their future needs, facilitated by the WLCG organization, OSG, the HEP Software Foundation and development projects such as HEP-CCE, IRIS-HEP and SWIFT-HEP. In this paper we outline the strategy by which the international HEP computing infrastructure, software and services should evolve through the collaboration of large and smaller scale HEP experiments, while respecting the specific needs of each community. We also highlight how the same infrastructure would be a benefit for other sciences, sharing similar needs with HEP. This proposal is in line with the OSG/WLCG strategy for addressing computing for HL-LHC and is aligned with European and other international strategies in computing for large scale science. The European Strategy for Particle Physics in 2020 agreed to the principles laid out above, in its final report.

Comments: contribution to Snowmass 2021
Subjects: **Computational Physics (physics.comp-ph)**
Cite as: arXiv:2203.07237 **[physics.comp-ph]**
(or arXiv:2203.07237v1 **[physics.comp-ph]** for this version)
https://doi.org/10.48550/arXiv.2203.07237

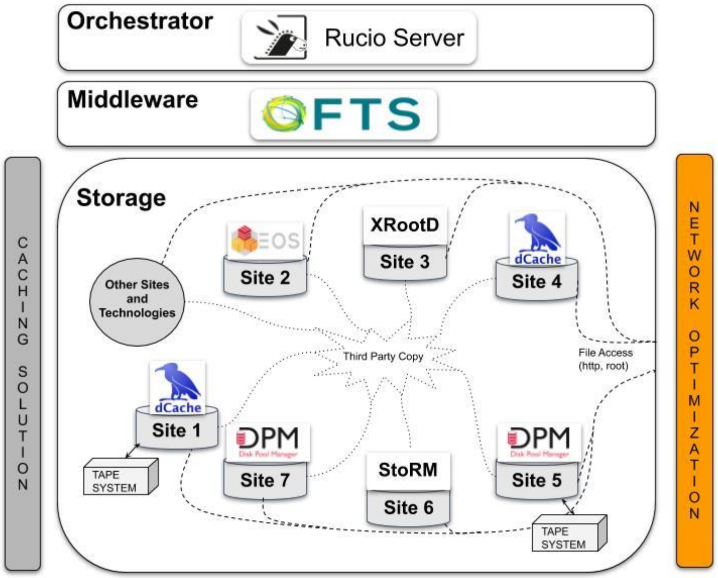
Collaboration with other sciences

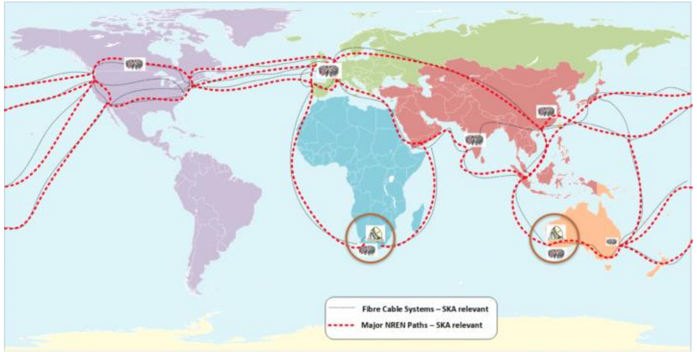
The ESCAPE project implemented a prototyped a data infrastructure prototype across Europe

- based on many of the WLCG building blocks and on top of many WLCG facilities

Examples of prototyped applications:

- SKA: data delivery from Perth and Cape Town to Europe and access through the data lake services





Experiments such as SKA will in future need a similar bandwidth as HL-LHC, sharing many of the network paths

It is now the opportunity to discuss how LHC, other HEP experiments and other sciences will cohabit

Conclusions

In the next 10 years WLCG will be faced with two major network challenges:

1) dealing with the HL-LHC data volumes and complexity
2) the cohabitation with other experiments and sciences on the same infrastructure

We should turn the two challenges into opportunities and the network community can and should play a leading role:

- Modernize the network services progressing with the ongoing R&D activities and bringing early prototypes in production

- Engage with other experiments and sciences to drive the evolution of the R&E networks and define a model for the future, based on the success story of LHCONE/LHCOPN and its community