DOMA R&D for HL-LHC

Maria Girone CERN-IT

(Thanks to Ian Bird and Simone Campana, along with the members of the HSF CWP effort)
The LHC Upgrade Programme

- **RUN3**: LHCb and ALICE upgrades
- **RUN4**: ATLAS and CMS upgrades

Graph showing the progression of luminosity and trigger rates over different runs and years.
The events at the HL-LHC

• Increased complexity due to much higher pile-up and higher trigger rates (1kHz@Run2 – 7.5kHz@HL-LHC) will bring several challenges to reconstruction algorithms
Challenge of HL-LHC

- The next phase of the LHC will see the data rate and complexity both increase by a factor of 5-10
  - Estimates for the needed increase in resources and expected gap are 4-8

- The community is developing an R&D program to help evolve computing models, services and operations and shrink the resource gap
Planning and Preparation

- The HEP community has been dealing with large and increasing samples of data for decades. Currently we have approximately an Exabyte of data under management, 750K CPU cores distributed world-wide.

- One of our strengths is the ability to develop and plan coherently to cope with computing challenges.
  - The WLCG (Worldwide LHC Computing Grid) is the successful outcome of this approach; a common infrastructure for all the experiments, common protocols and services.

- The HEP Software Foundation (HSF) Community White Paper (CWP) is now defining a roadmap for HEP software and computing R&Ds for the 2020s.
  - For the field this is seen as crucial for preparation for HL-LHC.
  - The CWP consists of 13 work packages, each one defining a set of R&Ds.
  - Other research communities with computing needs at the level of WLCG will coexist on the same infrastructure.

- The concepts outlined in this talk come mostly from the “Data Organization Management and Access” section.

http://hepsoftwarefoundation.org
DOMA Strategy for HL-LHC

Three DOMA R&D areas relevant to this workshop

- **Improve the organization and access of the data**
  - Capitalize on improvements from industry in big data technologies
  - Have data management systems able to exploit different types of storage (including volatile storage)

- **Consolidation of Storage**
  - Data-lake(s) approach to concentrate data storage in fewer larger centers
  - Evolve from data placement to data serving

- **Optimize the distribution of data to users**
  - Implement more intelligent caching
  - Increase diversity of the storage types from archival+disk to solutions optimized for specific use cases (high performance, non-persistent, low cost/low reliability, archival, etc..)

16th November 2016  Maria Girone DOMA Workshop
Data Organization and Big Data Tools

• Study the impact of column-wise vs. row-wise organization of the data on the performance of each type of access
  • Reading entire events vs. reading single objects from many events
  • Efficiency of compression schemes and mappings onto hardware architectures

• Evaluation of industry tools for HEP data analysis (ex. CMS physics data reduction project with Intel)
  • Tools to index events and reuse selections (à la map-reduce)
  • Tools to parallelize data access like Spark

• Investigate specialized hardware for data transformations
  • Specialized hardware like FPGAs for tracking reconstruction (ex. EC funded DEEP-EST project at Jülich Supercomputer Centre)

More in J. Pivarski’s talk
Storage Consolidation – Data Lakes

- Many Funding agencies wish to consolidate storage in a region, to optimize operational costs, with in mind also other communities beyond HEP
  - Currently HEP operates with many independent sites (~200) and there are many replicas of the data. On average there are 2-3 replicas for every sample produced

- The consolidation might be physical: reduce the number centers offering a storage service
  - Reduces number of services and the effort needed to run storage

- Or logical: implement a distributed storage system across centers
  - Reduces the number of replicas

- This R&D should evaluate which storage solutions are suitable to implement a distributed storage system and build prototypes
  - Within WLCG will do a PoC with EOS (CERN developed parallel storage system). Interest from SKA to join

- The R&D should also acquire experience with data processing and data management for those prototypes

More in B. Bockelman’s talk
Possible Model for future HEP computing infrastructure

Simulation resources

Cloud users: Analysis

HEP Data lake
Storage and compute

1-10 Tbps

1-10 Tbps
Caching

• Consolidating storage implies breaking CPU/data co-location, further than now
  • Today between 10-20% of data serving is remote

• This R&D should evaluate the need for a storage caching layer for those sites outside the data-lakes and for which use cases
  • Improving the intelligence of the caching

• The evaluation should consider benefits in terms of CPU performance while accessing data and effort in operating the infrastructure
  • Does operating cache maintain high CPU efficiency and reduce effort in operations?
Outlook

• The HL-LHC program would require more computing than we could possibly afford unless we change how we work
  • This challenge is launching a lot of R&Ds
  • Looking for external communities to peer with
• Industry has been analyzing very large and very diverse datasets
  • We are investigating which ‘Big Data’ tools can be adopted
• Storage consolidation and data-lakes are a priority for WLCG
  • Maintaining accessibility to the data by a globally distributed community of thousands of users is a requirement, so data serving will be an active development area