

[Save to Evernote](#)

CHEP 2016

Updated Oct 27, 2016

Mon Oct 10

- Siegrist talk
 - https://indico.cern.ch/event/505613/contributions/2314340/attachments/135082/2039181/Siegrist-2016-10-10_CHEP_2016_v8.pdf
 - Need more resources than currently allocated to HEP for 2025
 - Mentions Art
 - Networking (medium term is 5 years)
 - LHC run 4 2026 is Exabyte era
 - ESNet is ASCR -> success is due to partnership of HEP with ASCR
 - ASCR is responsible in DOE to increase the computing capacity; HEP will need to collaborate with ASCR and port code to use those resources
 - ASCR / HEP collaboration: s17
 - Long term: Deep learning, Smart networking, Neuromorphic, Quantum computing National Strategic Computing Initiative (NSCI) - Exec order - Quantum Information Science (Q Computational Science, Q Entanglement sensors, Partnering with HEP).
- Future of Compute - Mark Seager
 - <https://indico.cern.ch/event/505613/contributions/2314341/attachments/135088/2039327/CEP2016-FutureOfCompute-SeagerV1.pdf>
 - Big Data Analytics - Government and research (Juman brain project); Commerce and industry (reduce RnD / time to market); New Users / Uses (Genomics, Clinical)
 - Project Moonshot for Cancer: Predictive oncology
 - Intelligent devices: 100x per year (moor's law 2x every 18 months)
 - Intel: form a PC company to a data center + IoT
 - Intel: Focussing on HPC and HTC
 - Future: integration of CPU, Memory - hot: fast storage (e.g. luster substitutes local buffers) - cold: large storage

runners) - cold, large storage

- Future of Distributed Data and Workload Management - Ian Fisk
 - https://indico.cern.ch/event/505613/contributions/2323233/attachments/135091/2039442/Outline_IMF_CHEP16_V1.pdf
 - In HL LHC era: assume 6-10x improvements / speed; 10x more data; 60x more processing needs
 - Medium term: Keep custodianship of the data and learn to deploy the data to where the compute cost is cheap.
 - Long term: Data reduction centers (consolidate selections between groups); investigate data transformations instead of “blind” event processing
 - (s22: good summary of future)

Infrastructure track

- ATLAS and LHC computing on CRAY - Francesco Giovanni Sciacca
 - <https://indico.cern.ch/event/505613/contributions/2227411/attachments/134846/2039429/Oral-v3-88.pdf>
- Integration of the Chinese HPC Grid in ATLAS Distributed Computing - Andrej Filipci
 - <https://indico.cern.ch/event/505613/contributions/2227415/attachments/134778/2033077/Oral-198.pdf>
- OCCAM: a flexible, multi-purpose and extendable HPC cluster - Stefano Bagnasco
 - <https://indico.cern.ch/event/505613/contributions/2227418/attachments/134785/2041350/Oral-232-v3.pdf>
- The OSG Open Facility: An on-ramp for opportunistic scientific computing - Bo Jayatilaka
 - <https://indico.cern.ch/event/505613/contributions/2227426/attachments/134517/2032830/Oral-277.pdf>
- Dynamic resource provisioning of the CMS online cluster using a cloud overlay
 - <https://indico.cern.ch/event/505613/contributions/2238070/attachments/134761/2039377/oral-412.pdf>
 - Can add up to 75% of the total CMS T1 pledges by using the online cluster for offline when not in use, using cloud technologies.

Data Handling track

<https://indico.cern.ch/event/505613/sessions/207450/#20161010>

- CERN's Ceph infrastructure: OpenStack, NFS, CVMFS, CASTOR, and more! - Xavier Espinal Curull
- CEPHFS: a new generation storage platform for Australian high energy physics - Goncalo Borges

- The US Exascale Project - Paul Messina
 - https://indico.cern.ch/event/505613/contributions/2314342/attachments/135214/2041460/Messina.ECP.CHEP_2016-10-10.final.pdf
 - Address application devel, software technology, hw technology, exascale system (integration)
 - 10 yrs projects for 17 DOE labs
 - Goals: 2 diverse architectures (vendors), US HPC Leadership, foster app development, ease of use

 - Diversity and Bias: Creating Inclusive Workplaces, See Bias, Block Bias - Caroline Simard
 - Block bias in Evaluations
 - https://indico.cern.ch/event/505613/contributions/2314343/attachments/135878/2039324/Simard-Creating_Inclusive_WorkplacesCHEP.pdf
 - discuss and agree to criteria in advance
 - notice higher bar and insist on universal application
 - notice when someone is given a “pass” and ask what criteria was used
 - discard unnecessary of narrow criteria
 - block undue criticism of personality
 - focus on top criteria
-

Tue Oct 11, 2016

- Networking - View from ESnet - Inder Monga
 - https://indico.cern.ch/event/505613/contributions/2323222/attachments/134805/2033650/CHEP_Networking_Plenary_Talk_Monga_v2.pdf
- Networking - View from HEP - Shawn McKee
 - <https://indico.cern.ch/event/505613/contributions/2323224/attachments/134770/2032909/Plenary-Oral-McKee.pdf>
 - SDN available in the HEP-related networks in the 3-7 years scale
- Simulation - Key Issues for the Coming Decade - Federico Carminati
 - <https://indico.cern.ch/event/505613/contributions/2323230/attachments/134738/2032244/Oral-Plenary-Carminati.pdf>
- High Performance Computing Solutions - Tom Raisor (DELL-EMC)
 - https://indico.cern.ch/event/505613/contributions/2323242/attachments/135284/2043034/CHEP_Fall_2016_-_DELLEMC_Preso.pdf

Distributed computing track

<https://indico.cern.ch/event/505613/sessions/207449/#20161011>

- Software and Experience with Managing Workflows for the Computing Operation of the CMS Experiment - Jean-Roch Vlimant.
- Stability and scalability of the CMS Global Pool: Pushing HTCondor and glideinWMS new limits - Antonio Perez-Calero Yzquierdo
- Advances in Grid Computing for the Fabric for Frontier Experiments Project at Fermilab - Ken Herner

- Multi-threaded Geant4 on Intel Many Integrated Core architectures
- Investigations of Future Computing Platforms for High Energy Physics

- **Provenance-aware optimization of workload for distributed data production** - Dzmitry Makatun
 - https://indico.cern.ch/event/505613/contributions/2230716/attachments/134677/2031005/Oral-246-Makatun_CHEP_2016_draft_1.pdf
 - Similar concept to what we are discussing with IIT for a PhD program
 - Simulated the placement of data at 51 site, 2.7 PB of data
 - Does not include elastic resources, although plans are made every 12 h, so they could be integrated.
 - Assume guaranteed resources, rather than simulating preemption

- Software - HEP Software Foundation Community White Paper Status - Peter Elmer
- Software Challenges and Commonalities - Panel

Wed Oct 12, 2016

- Clouds in HEP - Randall Sobie (IPP/Victoria)
 - <https://indico.cern.ch/event/505613/contributions/2323243/attachments/135301/2043464/Sobie-Cloud-CHEP.pdf>
- Azure: the cloud platform for digital transformation - Taylor Newill (MS)
 - 75% of the Azure workload is on Linux
 - millions of cores with exponential growth month over month
- CCF - Kevin Phelan

- GCE- Karan Bhatia
 - 10B photos per day
 - HW-dedicated to machine learning is 10x more cost effective as per watt.
 - Tensor Flow is open source
- CERN OpenStack Cloud - Tim Bell
 - Grow on-premises services
 - Example challenge: CC-based procurement for Clouds
 - high-value services are not standardized
- Panel with the people above
 - Ride Google or MS network where network is not good.
 - 60% of HEP Capital is on storage - can we keep custodianship in 10 yrs or we become anachronistic?
- Quantum Computing - John Martinis

Distributed Computing Track

<https://indico.cern.ch/event/505613/sessions/207449/#20161012>

- The Cherenkov Telescope Array production system for Monte Carlo simulations and analysis - Luisa Arrabito
- Evolution of user analysis on the Grid in ATLAS - Federica Legger
- Opportunistic data locality for end user data analysis - Max Fischer
- Geographically distributed Batch System as a Service: the INDIGO-DataCloud approach exploiting HTCondor - Sara Vallero
- The Cloud Area Padovana: from pilot to production - Lisa Zangrando
- On-demand provisioning of HEP compute resources on cloud sites and shared HPC centers - Thomas Hauth

Thu Oct 13, 2016

- Computing for HENP Experiments and Programs - Blurring Online and Offline - Pete Hristov
- Non-HEP: Data Management for LSST - Jacek Becla
- Non-HEP: Computing Challenges at Light Sources - Dula Parkinson

Distributed Computing session

<https://indico.cern.ch/event/505613/sessions/207449/#20161012>

<https://indico.cern.ch/event/505613/sessions/207449/#20161013>

- Experience in using commercial clouds in CMS - Maria Girone
- The HNSciCloud project - Helge Meinhard
 - Helix Nebula
 - s5 on procurement and legal challenges.
- Technical challenges of HNSciCloud - Martin Gasthuber
- From Physics to industry: EOS outside HEP - Xavier Espinal
- Volunteer Computing Experience with ATLAS@Home - David Cameron

Software Development

<https://indico.cern.ch/event/505613/sessions/207451/#20161013>

- “Big Data” in HEP: A comprehensive use case study - Oli Gutsche
- Big Data Analytics for the Future Circular Collider Reliability and Availability Studies - Volodimir Begy
- XRootD Popularity on Hadoop Clusters - Luca Menichetti
- Hadoop and friends - first experience at CERN with a new platform for high throughput analysis steps - Prasanth Kothuri
- Big Data Analytics Tools as Applied to ATLAS Event Data - Ilija Vukotic
 - Performance measurements on doing cuts on data miss comparison with root
- Developing and optimizing applications for the Hadoop environment - Prasanth Kothuri
 - Data ingestion, formats and processing framework are key aspects
 - Use these tools to build app: Kafka, Parquet, Spark

Plenary

- Machine Learning in Science and Beyond - Pierre Baldi
 - Transmuting Physicists in Data Scientists - Kathy Copic
 - <http://insightdatascience.com/> Look at the 7-week fellowship
 - How to learn data science tools: meet-ups on products; Podcast software engineering daily
 - SciPy for data science
-

Fri Oct 14

- Storage - Sony Everspan - Horst Schellong
 - Everspan storage system - Optical Archival Disc - Write once media - Enterprise evolution of BlueRay
 - Designed to compete with tapes..
 - 300 GB per media; 500 GB in 2018 -> 1 TB -> 2TB -> 4TB
 - Restore 1.3 PB/day
 - Read: MB/s: 314 Everspan peak; 280 - Everspan avg; 300 - LTO7
 - Write 140 MB/s with verification
 - 100 yr media life guaranteed -> no migration
 - 176 PB per library at 0.05kW/PB -> \$7.7k/yr (HDD 50 times higher)
 - Price comparable to LTO7



Add a save button to the internet
Get Web Clipper

[Terms of Service](#) | [Privacy Policy](#)