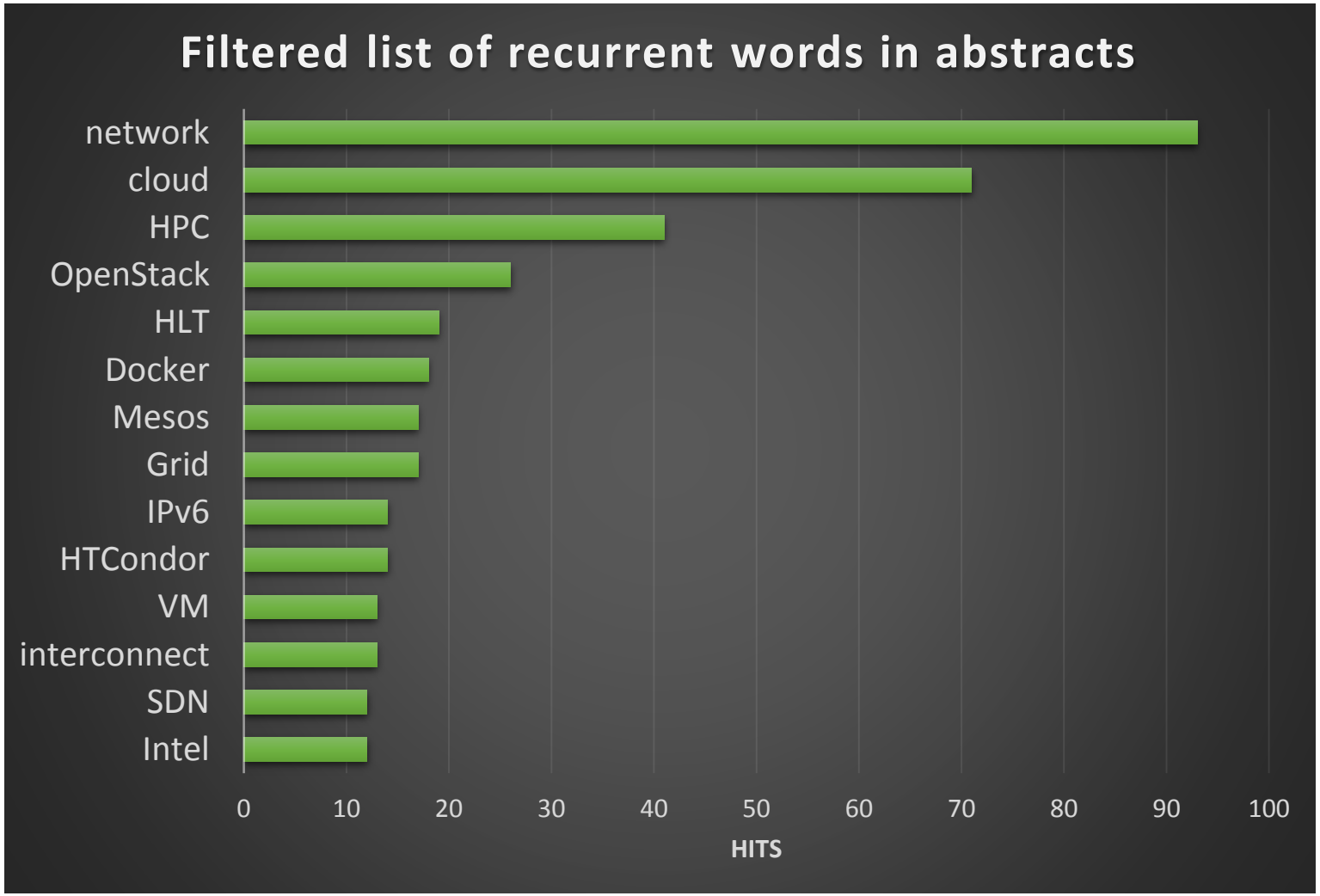
Track 6 - Infrastructure

Catherine Biscarat (LPSC Grenoble, IN2P3/CNRS)
Olof Bärring (CERN)
Margaret Votava (Fermi National Accelerator Laboratory)
Francesco Prelz (Università degli Studi e INFN Milano)
Oksana Shadura (CERN)

T6 in a nutshell



Orals	44
Posters	31
Total contributions	**75**

Topic	Sessions
Site reports	6.1 & 6.2
Networks and Interconnects	6.3 & 6.4
Methods and Tools	6.5 & 6.6
Scheduling and Provisioning	6.6 & 6.7

Site reports

- Ease opportunistic utilization of clusters
  - More scientific output from dedicated clusters (HPC, HLT)
    - Experiments get more resources and fund granters usually like the increased utilisation
  - Accommodate for the differences
    - Restrictions to login access and Linux installation
    - Assure immediate resource reclaim to priority user (e.g. HLT when beams are back)
  - Not unidirectional
    - Elastic HPC usage of non-HPC resources
  - The aggregate increments can represent a considerable part of annual production capacity

  →Don't forget compliance to (EGI) security policies..

Site reports

- Computing facilities
  - On- / off-premise dilemmas
    - Comprehensive cost comparison of public Cloud offerings and dedicated facilities
      - Fund granting may be subjected to genuine consideration of public Cloud alternatives
      - Procurement of Cloud capacity
    - Seamless and elastic extension to remote facilities
      - Remote (institutional) farms to extend T1 + cloud burst
    - Interconnected private Clouds
  - Hardware lease vs upfront purchase
  - Machine room aspects
    - Facility operation beyond IT equipment: meter and control power, temp., airflow,…
  - Lightweight site setup, site-in-a-box, virtual data centre
    - Enable smaller sites to contribute resources in an useful manner

Networks and interconnects

- IPv6
  - It’s approaching and it’s real.. Broad agreement on migration plan
    - IPv6 only on CPU nodes
    - Dual-stack central shared services and storage
    - Full IPv6 adoption by 2018?

Networks and interconnects

- IPv6
  - It's approaching and it's real.. Broad agreement on migration plan
    - IPv6 only on CPU nodes
    - Dual-stack central shared services and storage
    - Full IPv6 adoption by 2018?
- PerfSonar
  - Recurring CHEP topic since 2007(*)
  - Data storage in full-fledged production in 2015
  - Data is used, e.g. PUNDIT for automated analysis and escalation of real network problems

(*) https://indico.cern.ch/event/3580/abstract-book.pdf

Networks and interconnects

- High-bandwidth and/or low latency interconnects
  - Usually not for HEP, at least not offline processing
  - Potential drivers are rack level technology trends (NVMeoF, RSA, …)
  - Infiniband, Omnipath, RapidIO, APEnet and an interconnect developed in ExaNest project

Networks and interconnects

- High-bandwidth and/or low latency interconnects
  - Usually not for HEP, at least not offline processing
  - Potential drivers are rack level technology trends (NVMeoF, RSA, …)
  - Infiniband, Omnipath, RapidIO, APEnet and an interconnect developed in ExaNest project

- SDN
  - To optimise WAN transfers by allocating bandwidth and load balancing multiple large flows over diverse paths spanning multi-domain networks

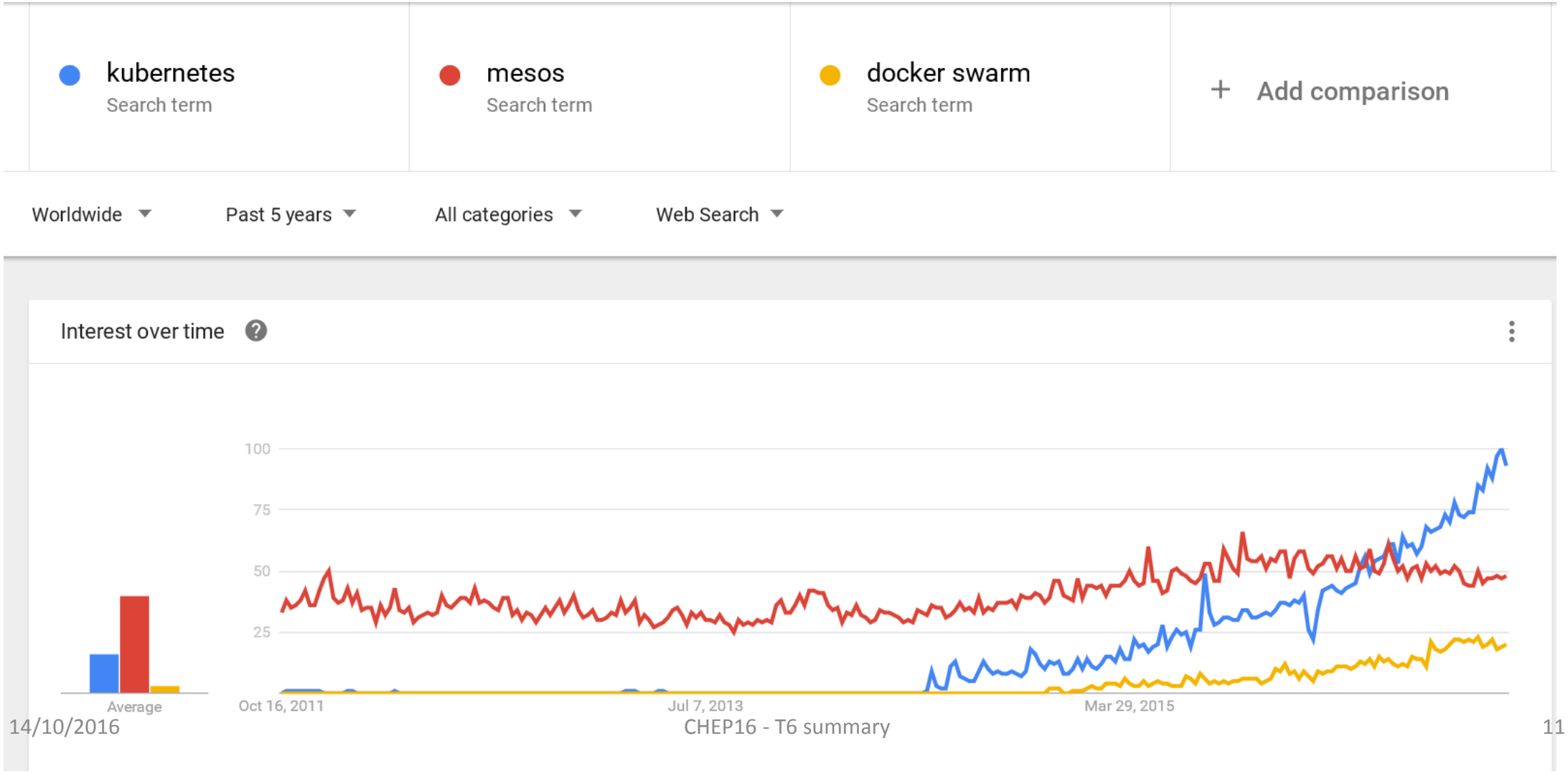
Methods and Tools

- Containers
  - Lightweight with bare-metal perf → attractive alternative to VMs
  - Simple to use tools to enable packaging of apps with their dependencies

Methods and Tools

- Containers
  - Lightweight with bare-metal perfs → attractive alternative to VMs
  - Simple to use tools to enable packaging of apps with their dependencies
- Container orchestration
  - hot topic…: Mesos, Kubernetes or Docker Swarm

Google trends



Methods and Tools

- Containers
  - Lightweight with bare-metal perfs → attractive alternative to VMs
  - Simple to use tools to enable packaging of apps with their dependencies
- Container orchestration
  - A hot topic…: Mesos, Kubernetes or Docker Swarm
  - Large number of contributions on this topic, not all in Track 6
    - Experiment specific services
    - Grid services
    - Web-based analysis
    - PaaS for web central hosting
  - SHIFTER ≈ DOCKER for HPC

Methods and Tools

- PaaS
  - Multi-discipline scientific communities data lifecycle management on federated data infrastructure
  - Data and computing platform for scientific communities provisioned over hybrid (private or public) e-infrastructures
  - Interconnected, secure private clouds with universal file access

Scheduling and Provisioning

- HTCondor
  - Broad adoption in HTC
  - Massively scalable
  - Perhaps less suitable for HPC
    - Scheduling of MPI jobs

Scheduling and Provisioning

- HTCondor
  - Broad adoption in HTC
  - Massively scalable
  - Perhaps less suitable for HPC
    - Scheduling of MPI jobs
- Other schedulers
  - HPC: SLURM, LSF and some PBS or Torque
  - SGE: still at DESY but phasing out

Scheduling and Provisioning

- Benchmarking
  - Need for lightweight benchmarks
    - Cloud procurement
    - Predict job run-time to optimise batch slot utilisation in pilot jobs
  - Features
    - Scale as the real workload
    - … and non-intrusive, self-contained and zero footprint
  - Grim reality
    - Matching is scaling difficult to achieve and sustain over time
    - Reproducibility is utterly uncertain

Summary of the summary

- Big sites – big challenges
- Small sites – …also big challenges
- Numerous undertakings – site-wide, national and international - for better resource utilisation with some common key objectives
  - Consolidation
    - seamless integration of various resources, sometimes with rather incompatible features
  - Serve multiple sciences
  - Enabling of opportunistic utilisation of virtually just any resource coming across
  - Automation
    - Machine assisted service mgmt. and system administration