Track 3 Highlights
– Middleware and Distributed Computing

Track 3

- 41 oral talks in 7 sessions and 4 posters
  - Thanks to all presenters and audiences for very interesting sessions

~ 40 participants

- Track conveners and session chairs
  - Catherine Biscarat (L2I Toulouse, IN2P3/CNRS)
  - Tomoe Kishimoto (Univ. of Tokyo)
  - James Letts (Univ. of California San Diego)
  - Stefan Roiser (CERN)

Topics

Middleware and Distributed Computing

Topics

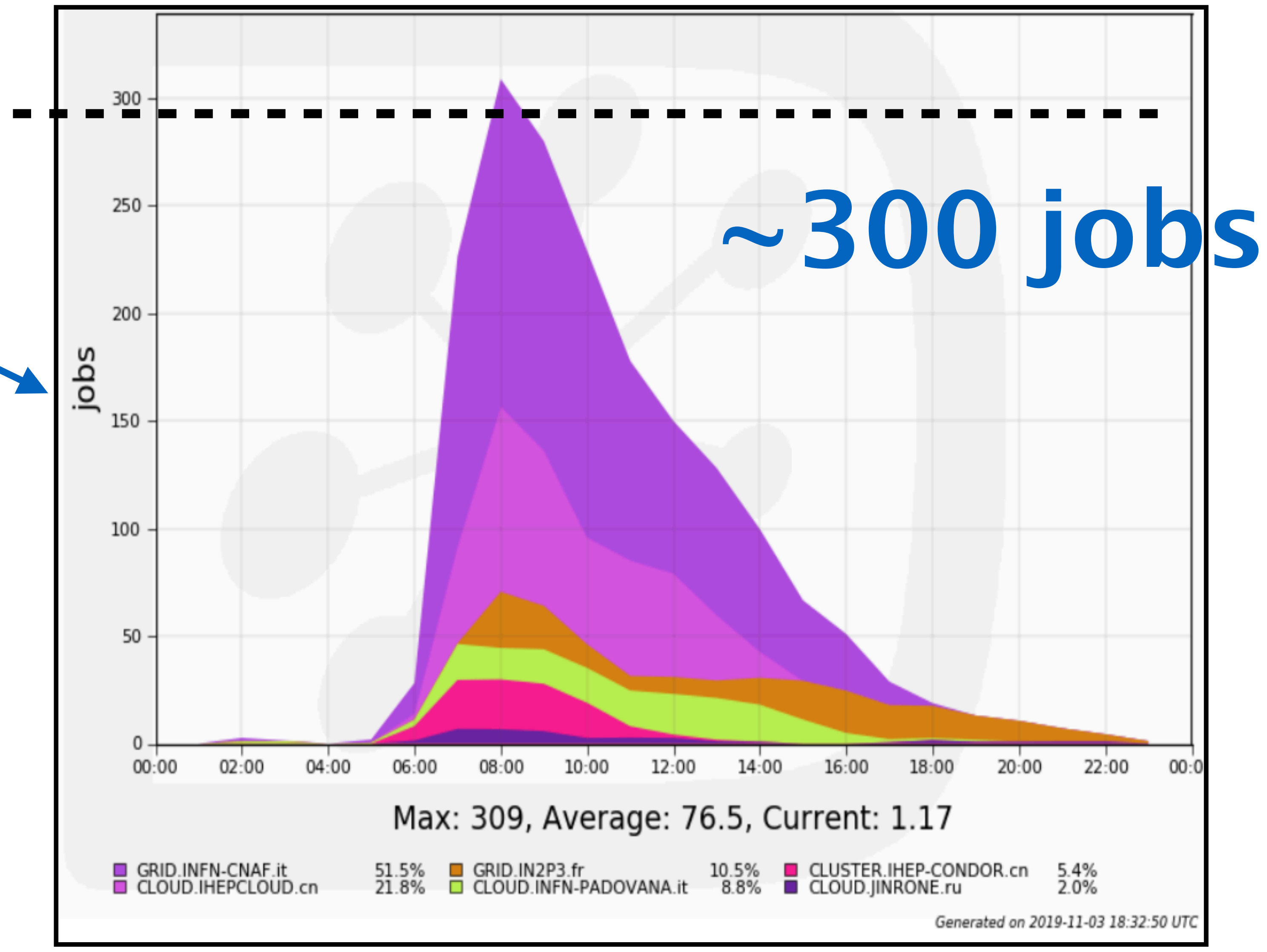
- Workload management
- HPC
- Lightweight site federation
- Operations
- Monitoring
- Identity
- Information System
- Cost evaluation

Disclaimer: Selected topics are biased. Apologies in advance if any highlights are missing due to lack of my understanding

Workload management

- Distributed computing software initially developed for LHC experiments are used/evaluated by multiple experiments and communities outside of LHC
  - e.g.) JUNO, Belle II, ILC use DIRAC to manage workloads

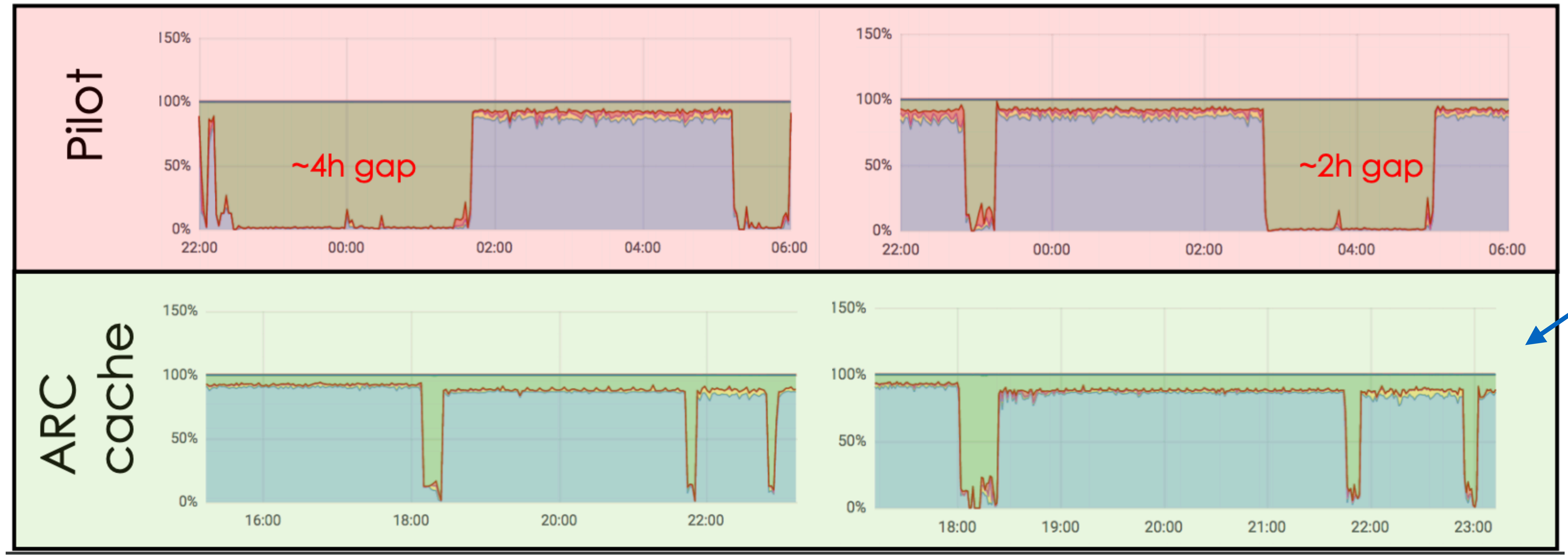
JUNO production system was tested with distributed computing resources



  - Reuse for midscale collaborations
  - Keeping multi-domain can improve the software and reduce duplication of work
- Small experiments sometimes do not have automated tools

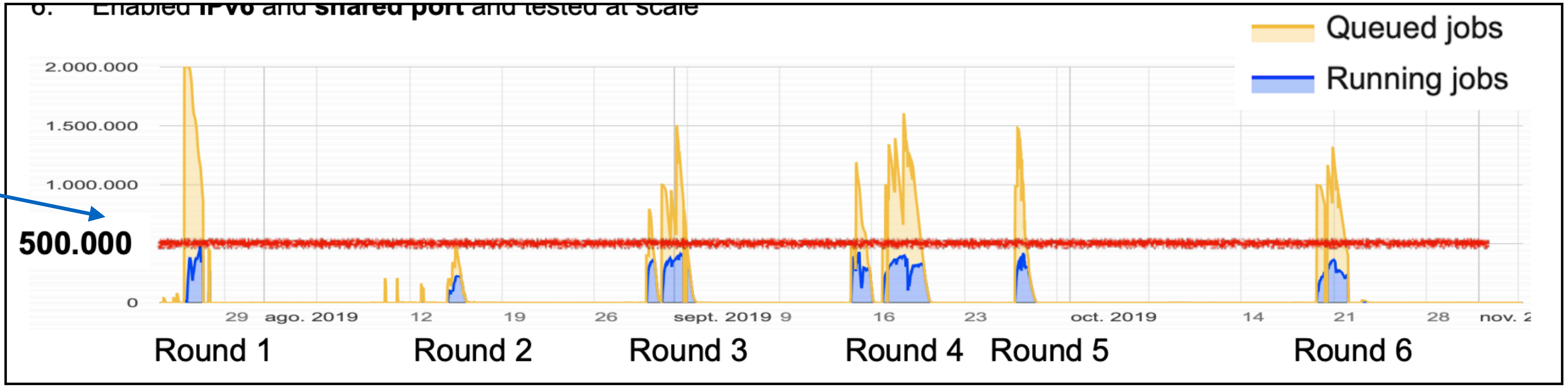
Workload management

- Workload management system and middleware continue to evolve..



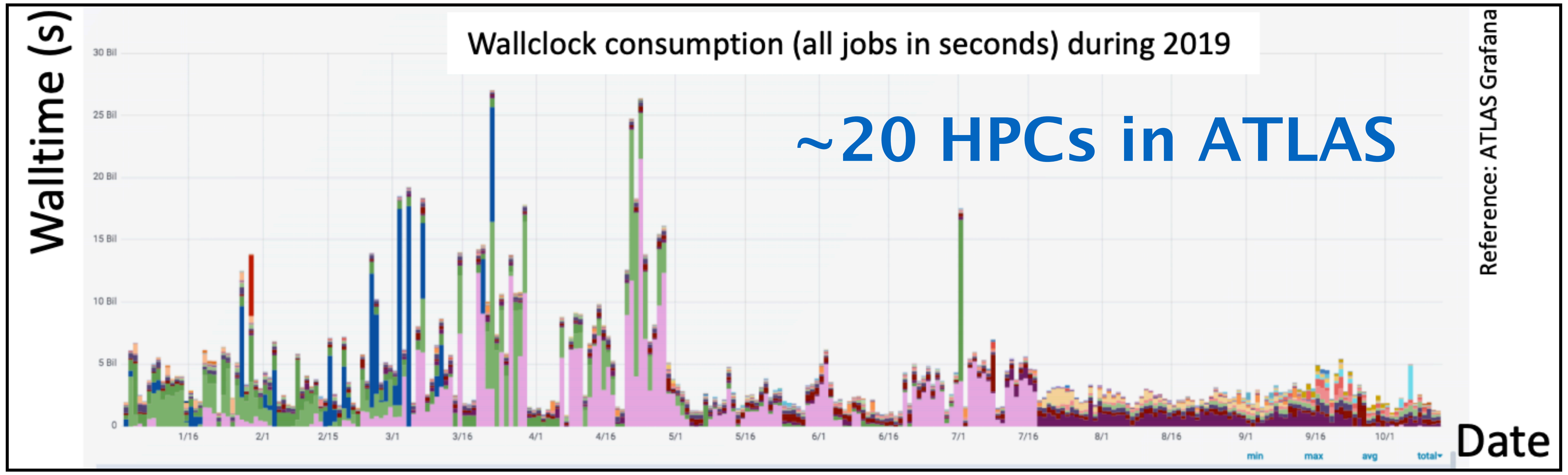
ARC stage-in/cache can improve CPU efficiencies for HPC and cloud resources

CMS tests scalability of submission interface for HL-LHC
450k running jobs



HPC

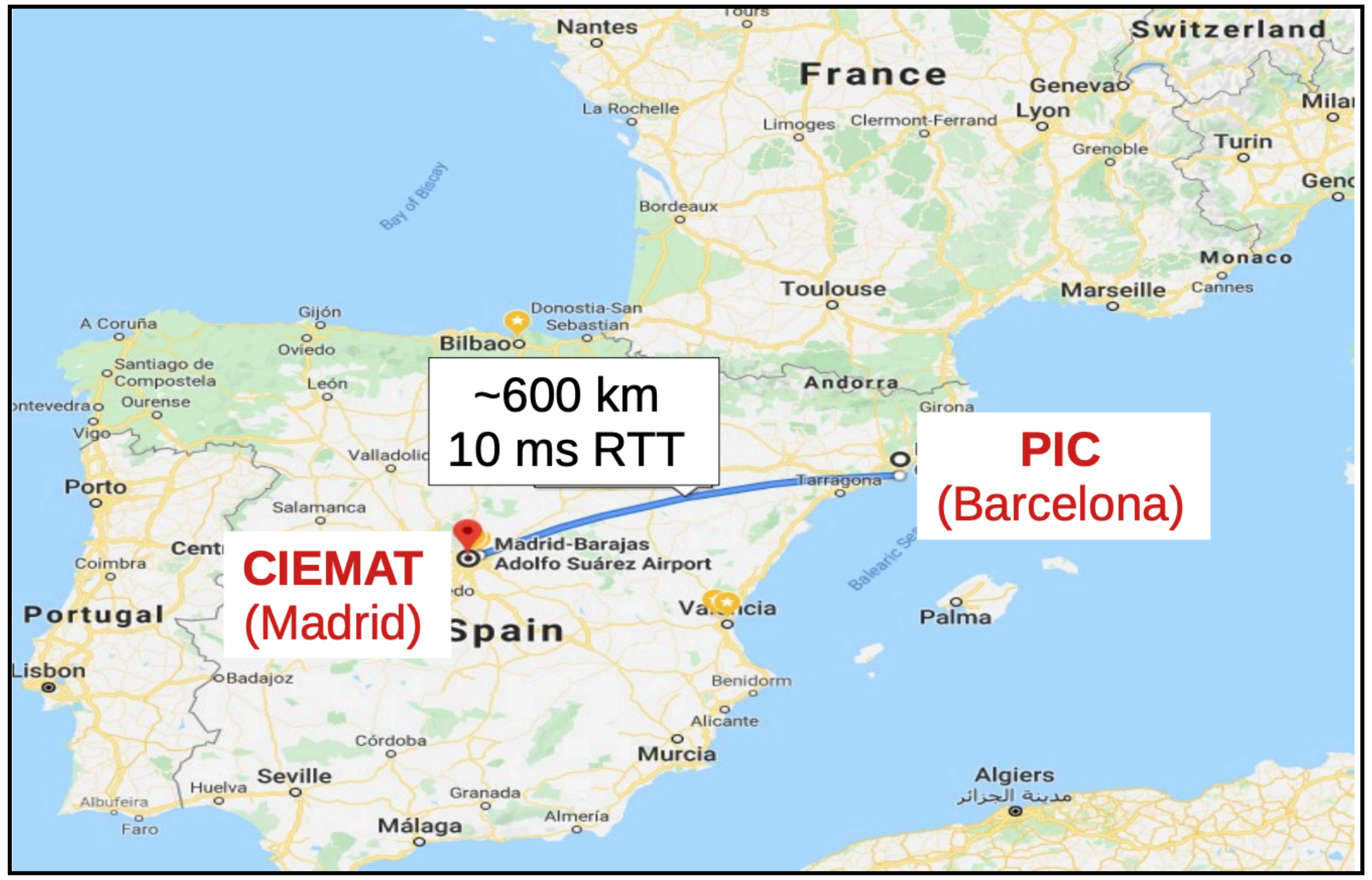
- Integrating HPC resources to experiment workflow is a hot topic in track3
  - LHC experiments, IceCube, SKA
  - Common challenges are different operational policies and architectures for each HPC → Container, Harvester, Pilot improvement, etc
- HPCs are used in production already

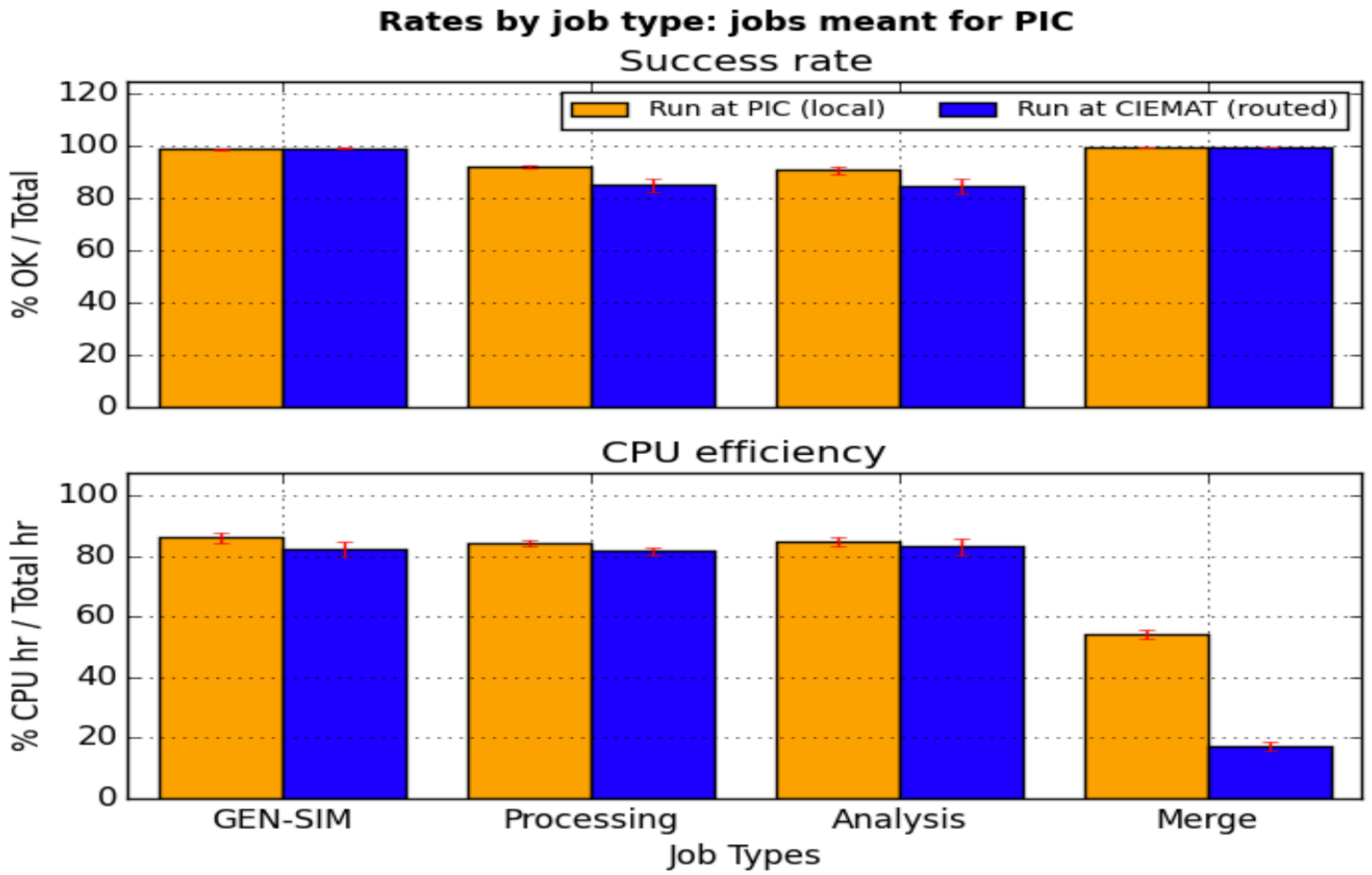


High scalability is confirmed in SKA1-LOW simulation

Lightweight site federation

- WLCG trend towards the federation of resources

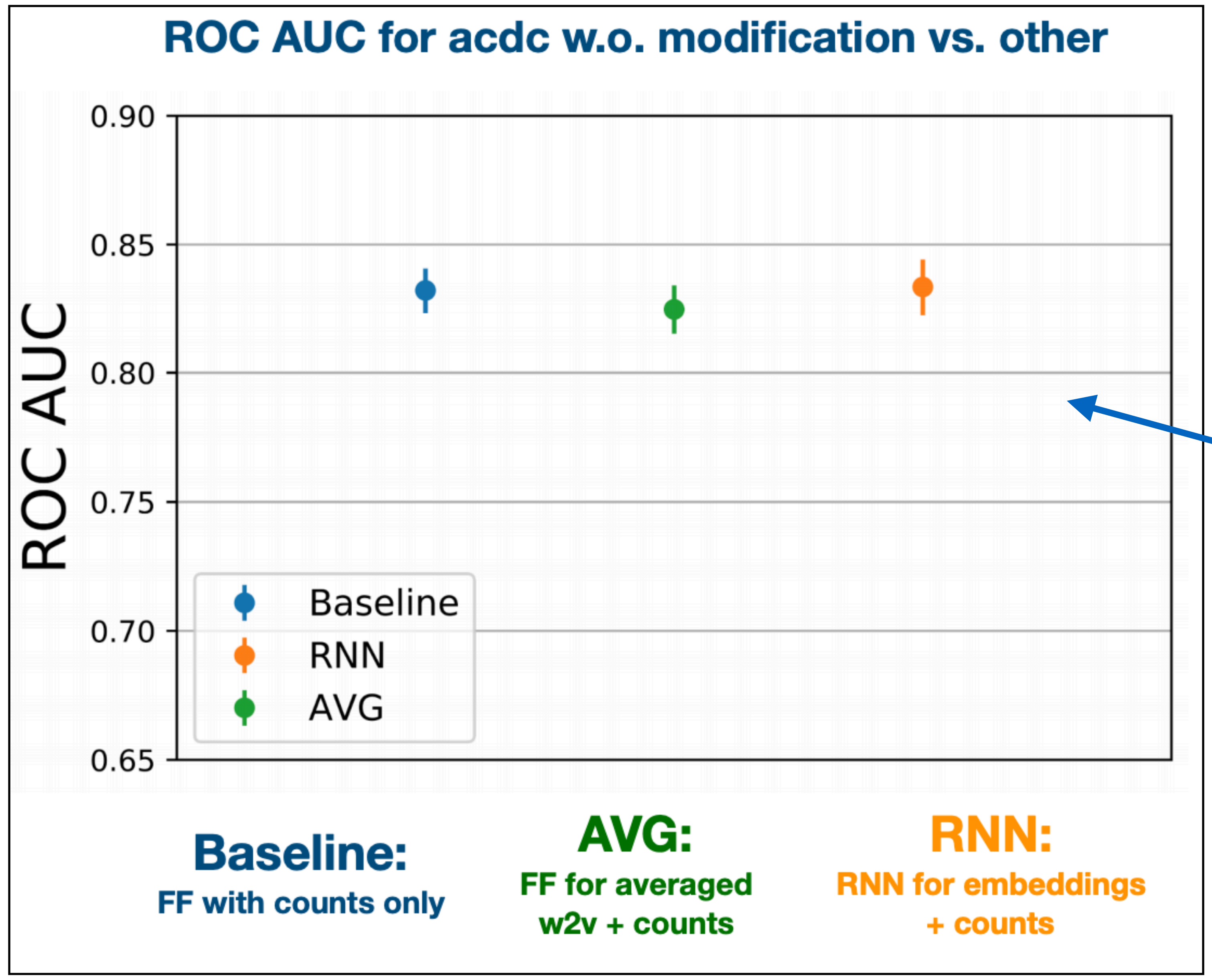




- The case of PIC and CIEMAT
  - No significant success rate or CPU efficiency degradation
    → The deployed lightweight federation has proven to work as expected

Operations

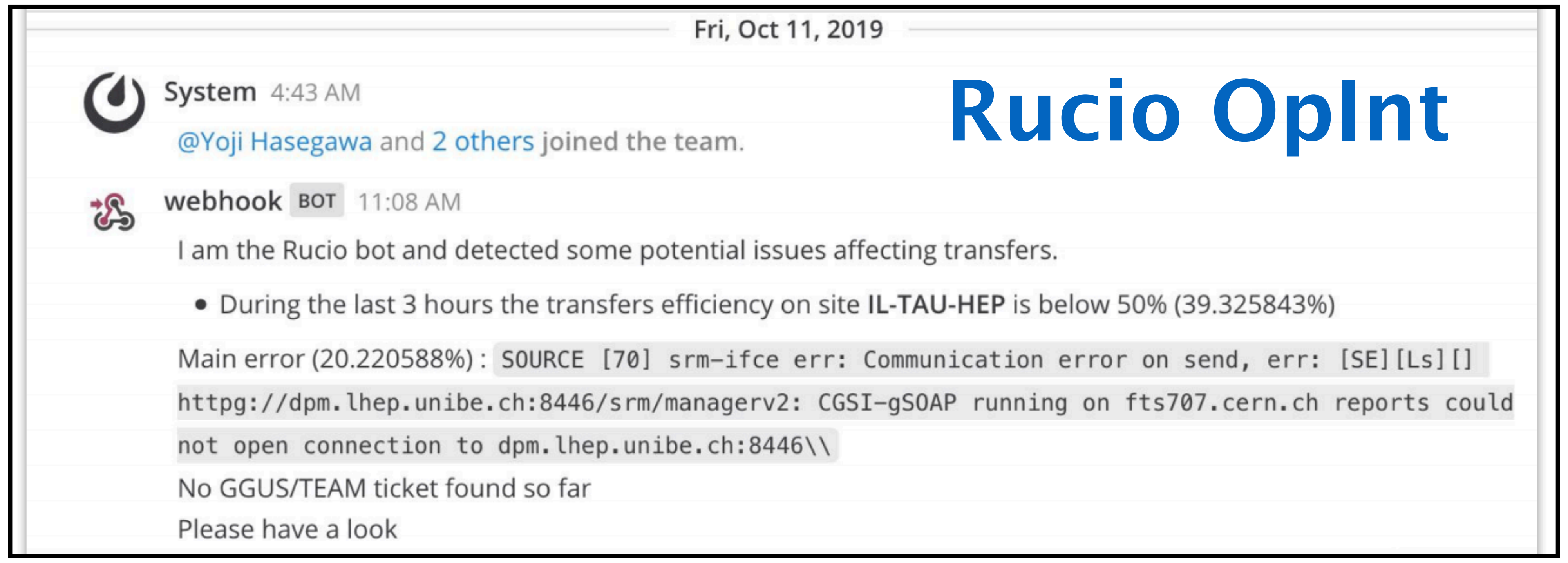
- ATLAS/CMS: 100+ people involved in computing operations



- Automation is a key for future operation

CMS creates a Machine Leaning model, which predict operator's action from error information

- Operation Intelligence (OI) is a cross-experiment project to improve computing operations



Monitoring

- ATLAS/CMS monitoring are migrating to CERN MONIT infrastructure, which is based on widely available open-source technologies

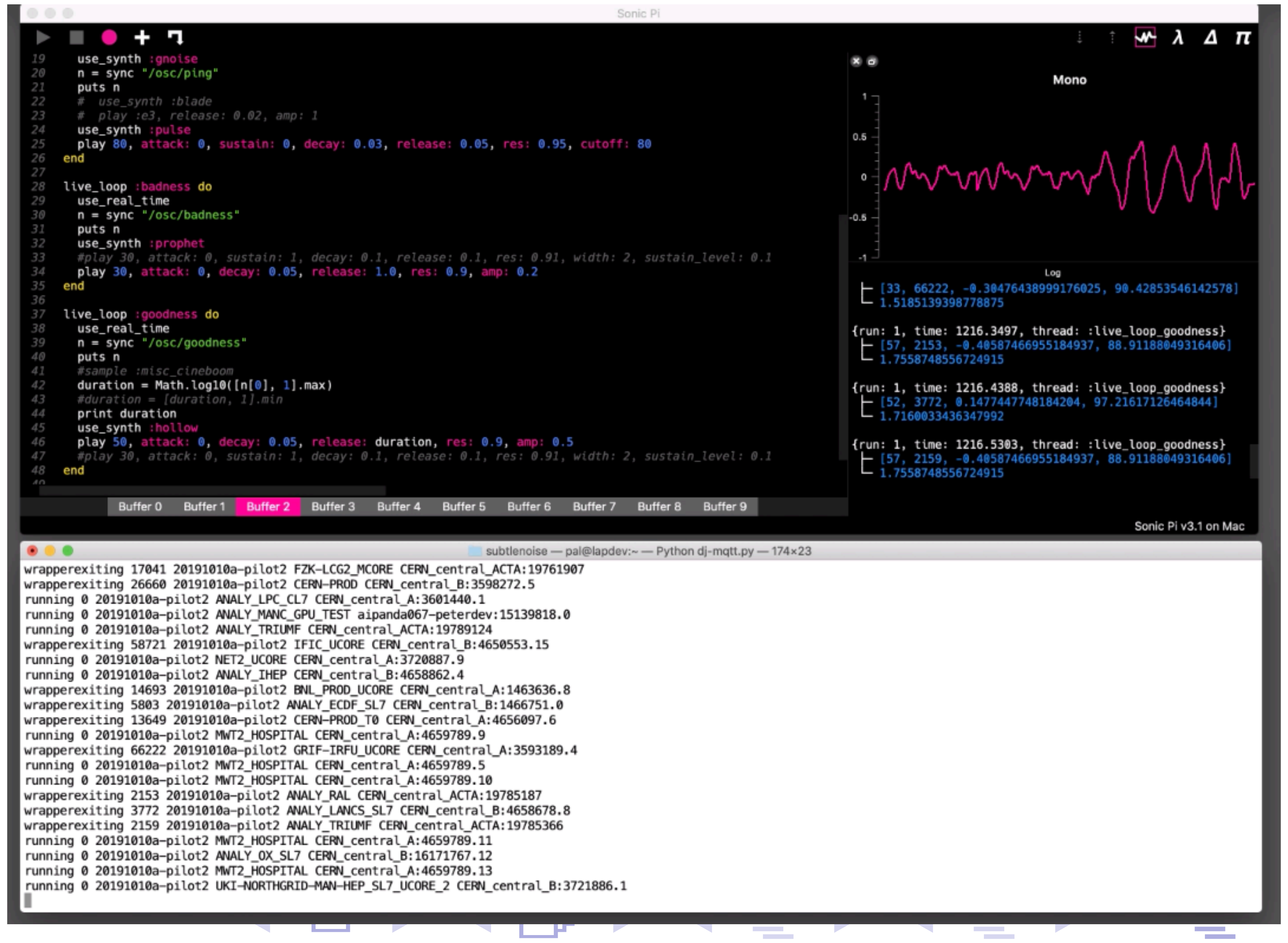
- Subtlenoise: Truly real-time monitoring with “noises”

Demonstration was very interesting ([link](#) to the contribution)

ATLAS data transfer matrix

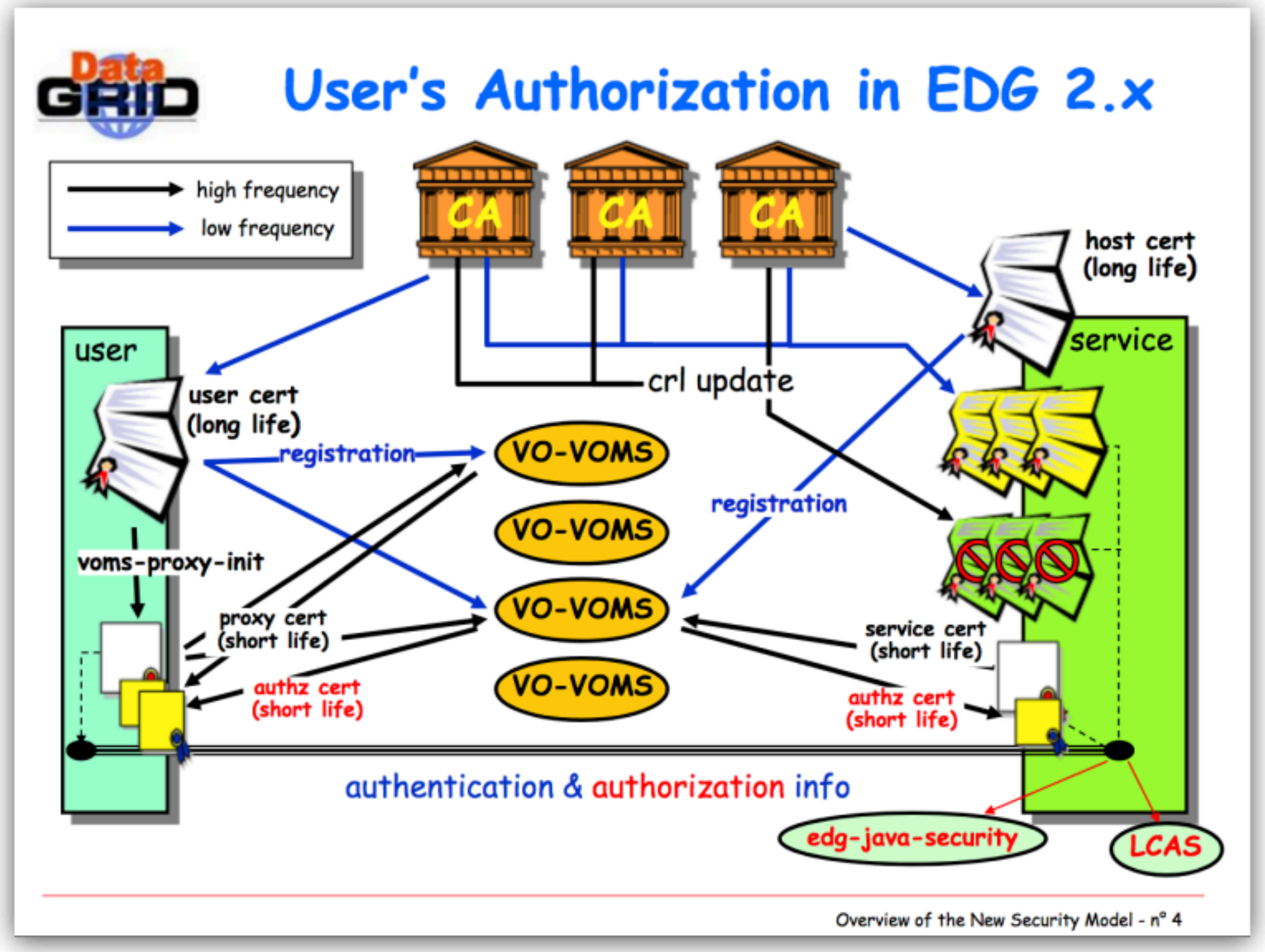
Efficiency

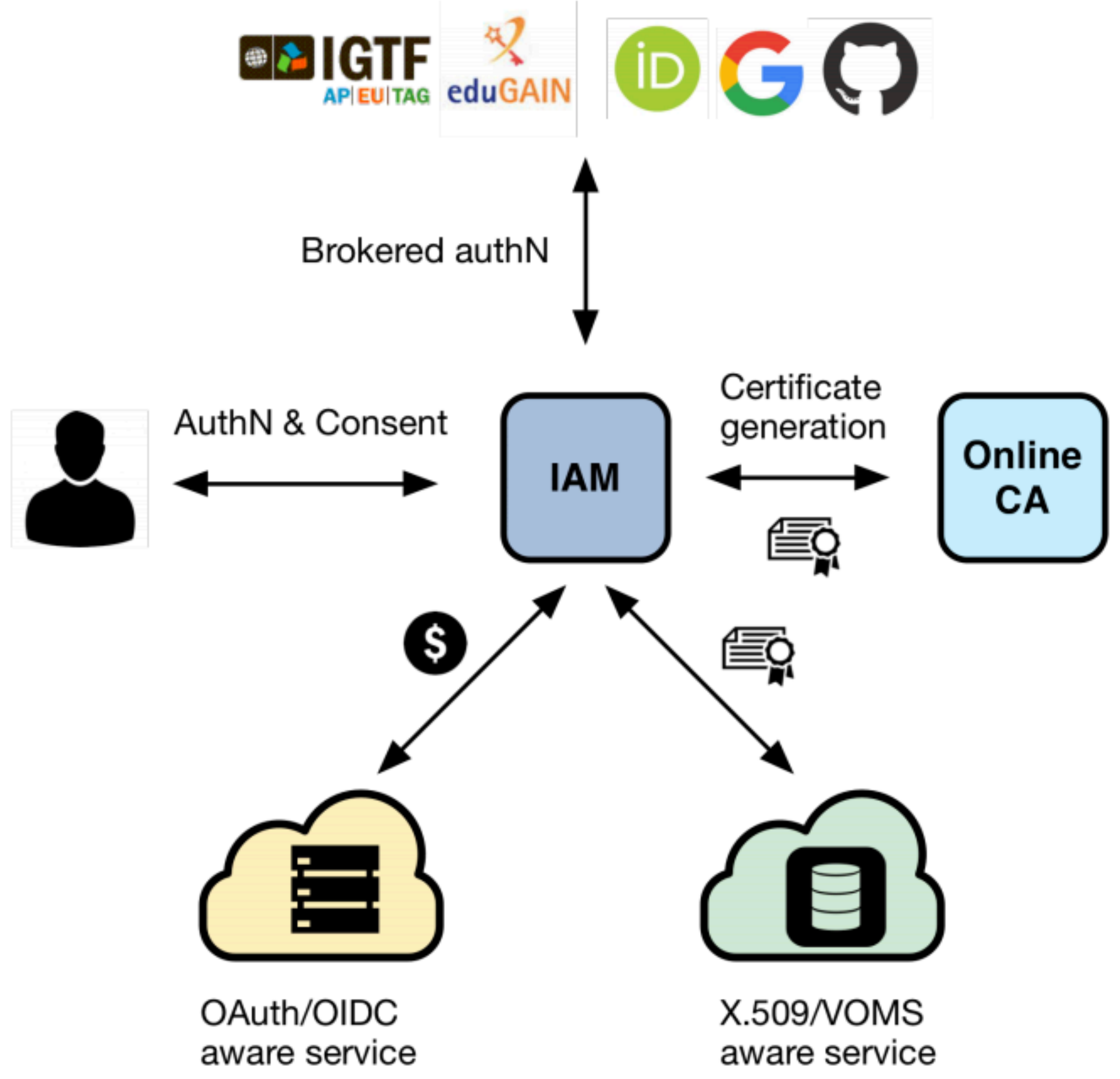
	CA	CERN	DE	ES	FR	IT	ND	NL	RU	TW	UK	US
CA	92%	97%	97%	65%	94%	93%	89%	91%	82%	72%	95%	84%
CERN	86%	96%	97%	96%	97%	96%	96%	93%	90%	85%	96%	87%
DE	76%	98%	91%	79%	87%	89%	96%	85%	82%	94%	93%	88%
ES	96%	98%	94%	99%	97%	93%	97%	89%	84%	99%	98%	90%
FR	95%	95%	98%	97%	96%	96%	96%	88%	92%	91%	95%	90%
IT	81%	93%	92%	85%	85%	97%	97%	86%	95%	97%	96%	83%
ND	89%	93%	95%	87%	74%	97%	98%	89%	92%	99%	98%	90%
NL	93%	43%	98%	98%	97%	97%	97%	97%	99%	99%	95%	90%
RU	85%	99%	99%	98%	95%	99%	98%	83%	100%	99%	98%	93%
TW	84%	100%	98%	98%	98%	95%	98%	71%	100%	-	97%	94%
UK	44%	74%	63%	28%	33%	30%	29%	19%	81%	77%	71%	74%
US	52%	88%	70%	53%	54%	62%	73%	77%	63%	88%	76%	80%



Identity

- WLCG authorization from x509 to Tokens is a hot topic

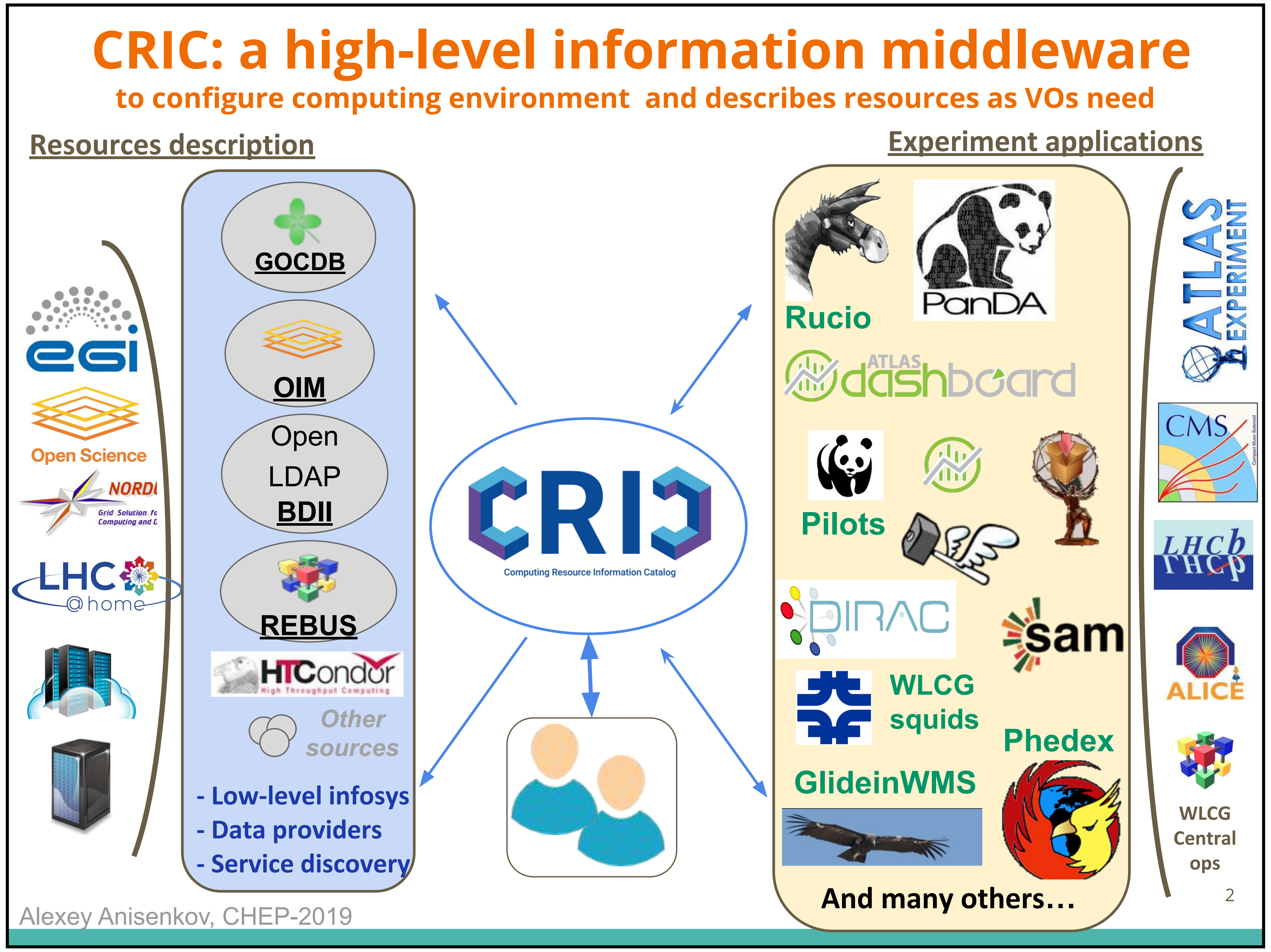


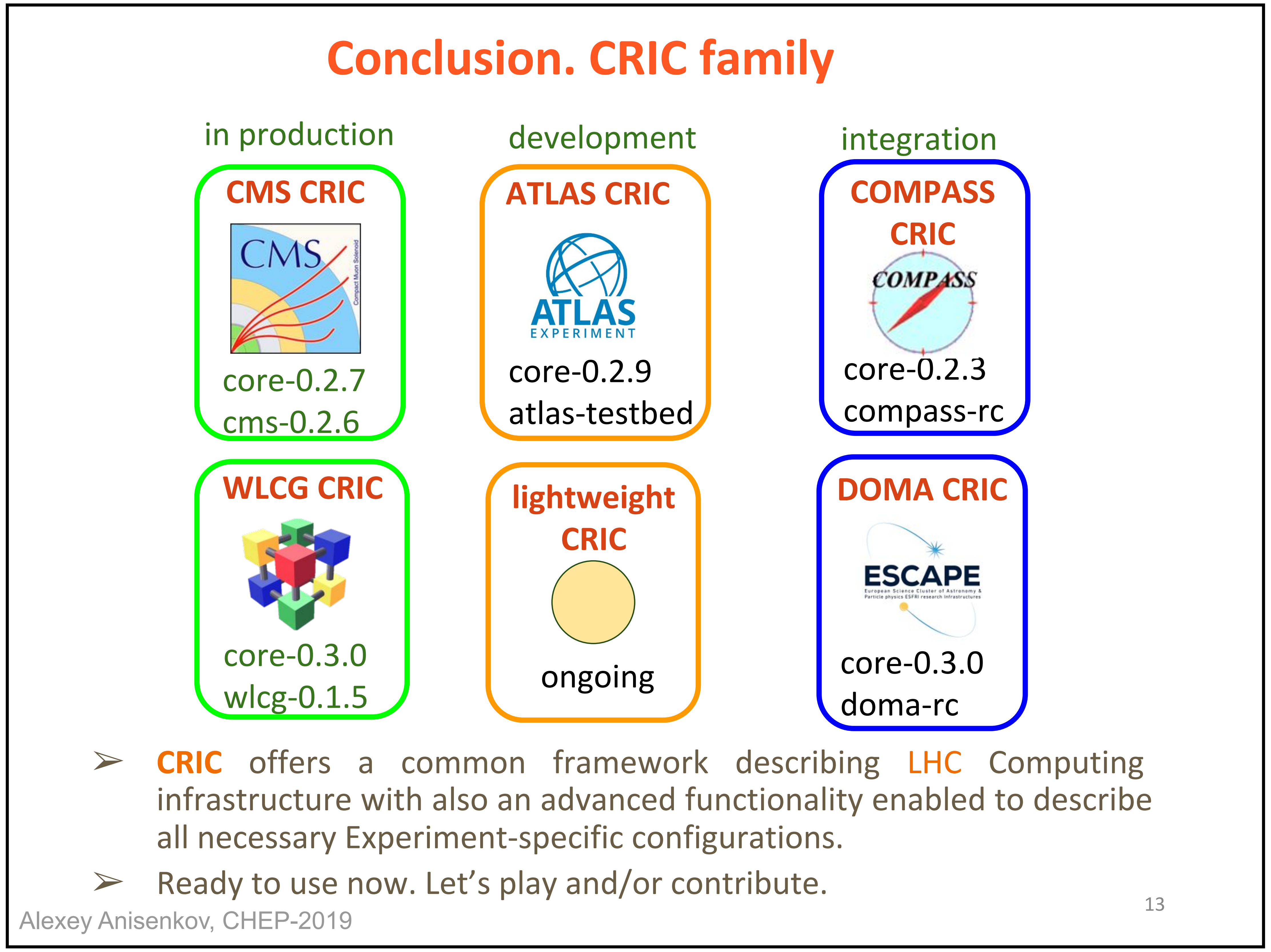


  - WLCG IAM instance is available to start the integration work
    - ✓ Successfully integrated with dCache, StoRM, XRootD (HTTP), FTS, RUCIO, HTCondor

Information system

- CRIC: a high level information middleware
  - Ready to use! CMS and WLCG in production



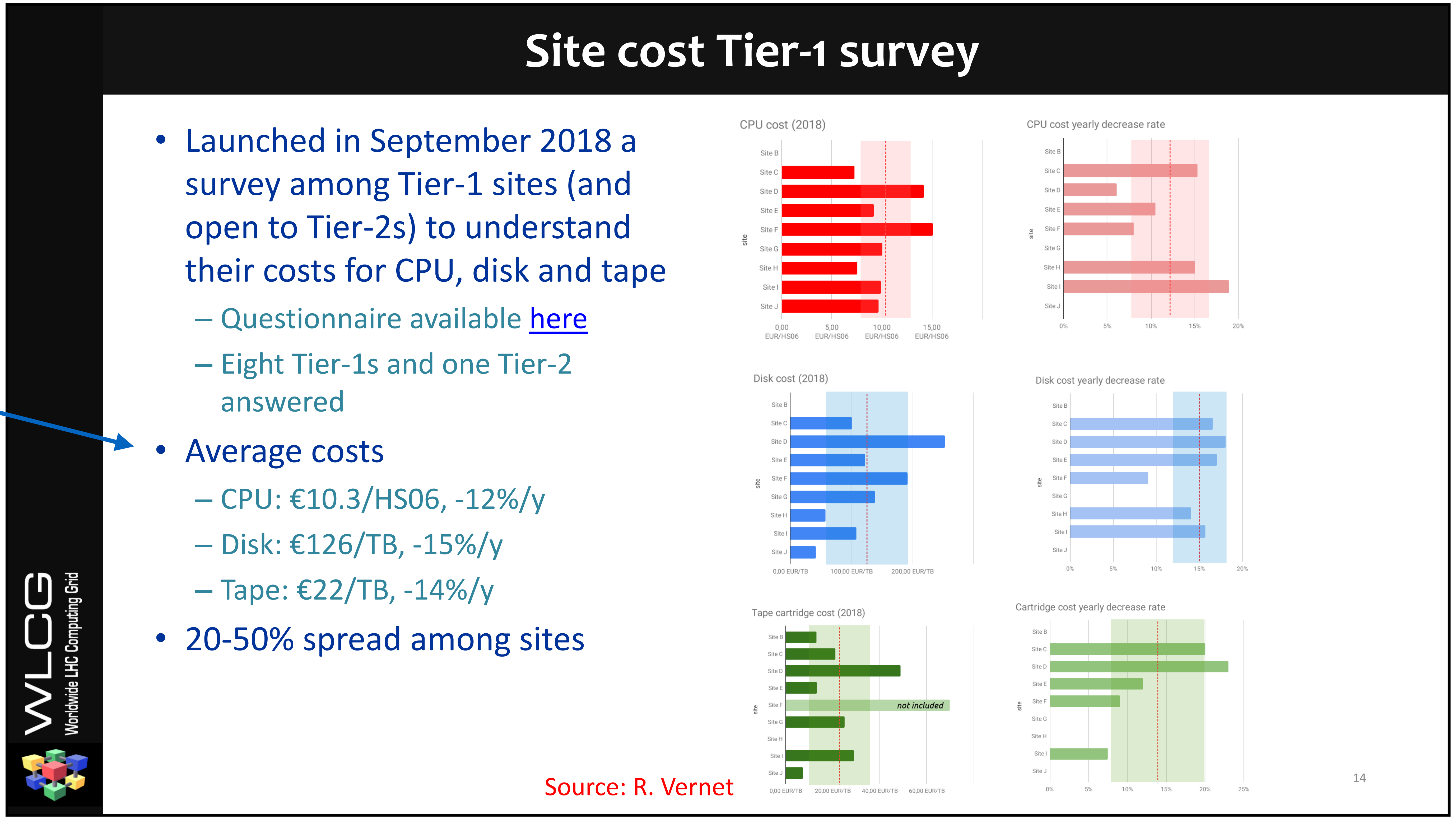


Cost evaluation

- Performance evaluation and cost modeling are important to understand the HL-LHC model
  - The working group was established in HSF/WLCG and showed important results

    Results of site cost survey

  - It is a key to include HPC resources (POWER, ARM, GPU, FPGA...) in the evaluation



Summary

- LHC distributed computing softwares are used in other experiments/ communities
- Integration of multiple resources is an active development item
  - Grid, HPC, cloud, GPUs, BOINC..
- Automation is a key for future distributed computing operations
- Performance and cost evaluation are also important for HL-LHC
  - Many "small" improvements can stack to provide significant gains!

Thank you!