



24th International Conference on Computing in High Energy & Nuclear Physics

4-8 November 2019, Adelaide, Australia

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**

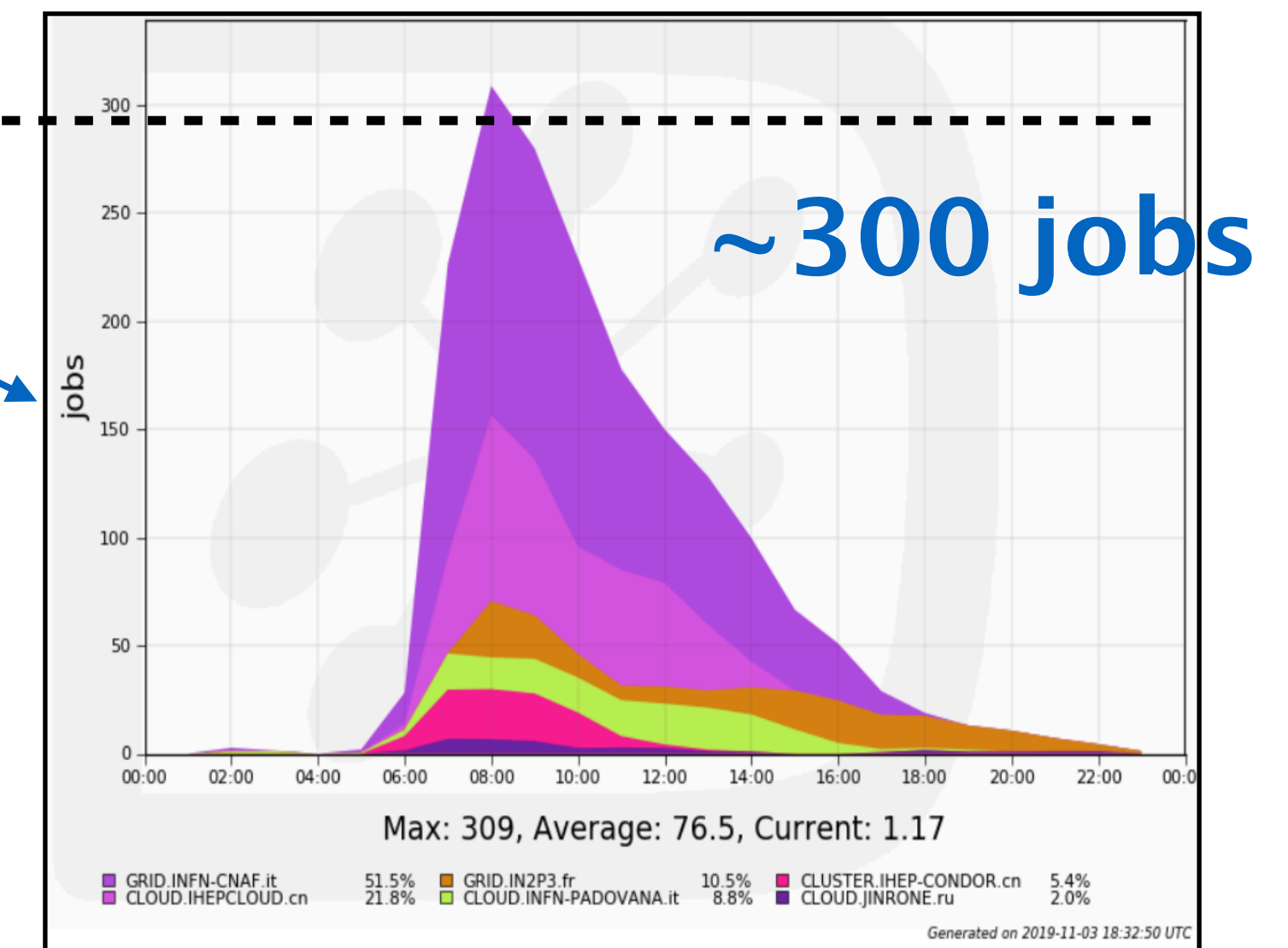
Disclaimer: Selected topics are biased. Apologies in advance if any highlights are missing due to lack of my understanding

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

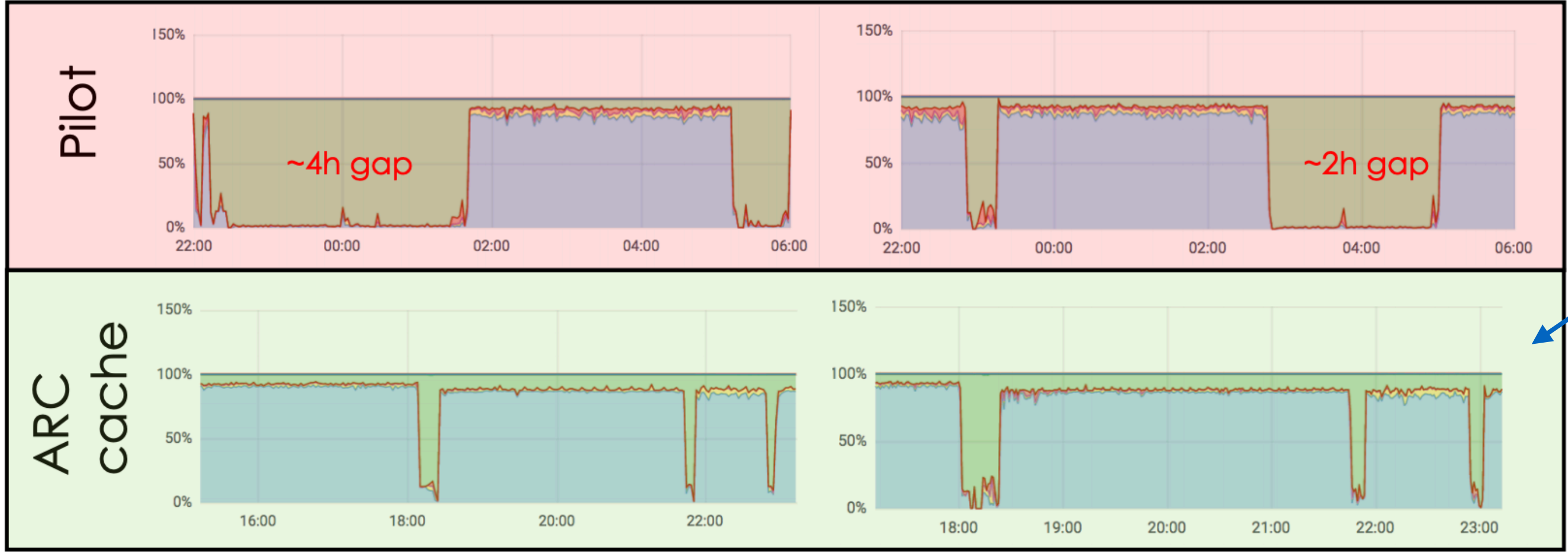
- Reuse for midscale collaborations
- Keeping multi-domain can improve the software and reduce duplication of work



- ▶ Small experiments sometimes do not have automated tools

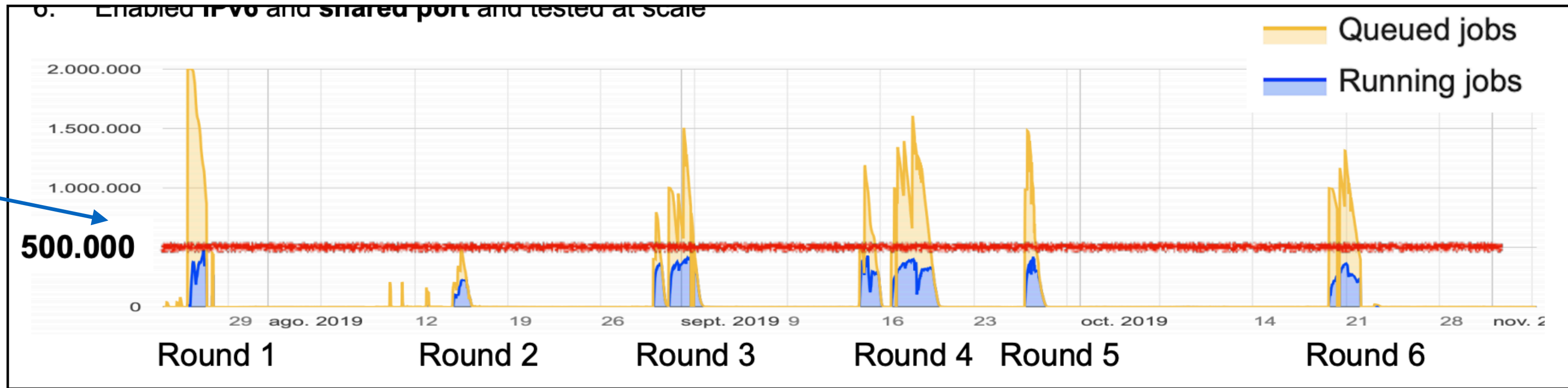
Workload management

- ▶ Workload management system and middleware continue to evolve..



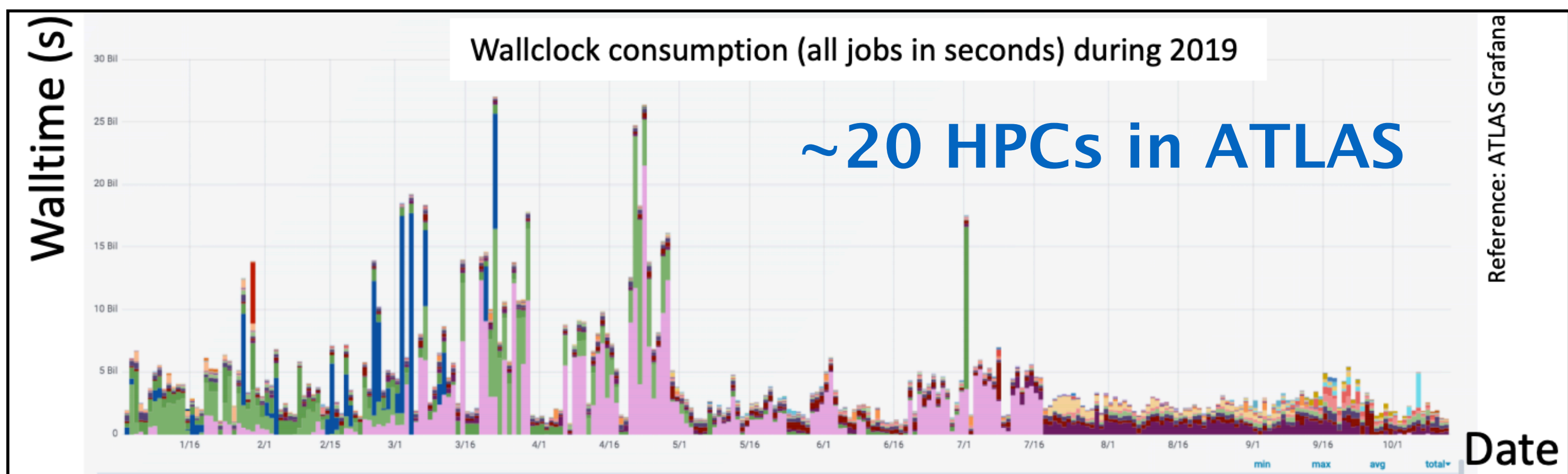
ARC stage-in/cache can improve CPU efficiencies for HPC and cloud resources

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

Latest News

ICRAR

OAK RIDGE National Laboratory

中国科学院上海天文台 SHAO Shanghai Astronomical Observatory CAS

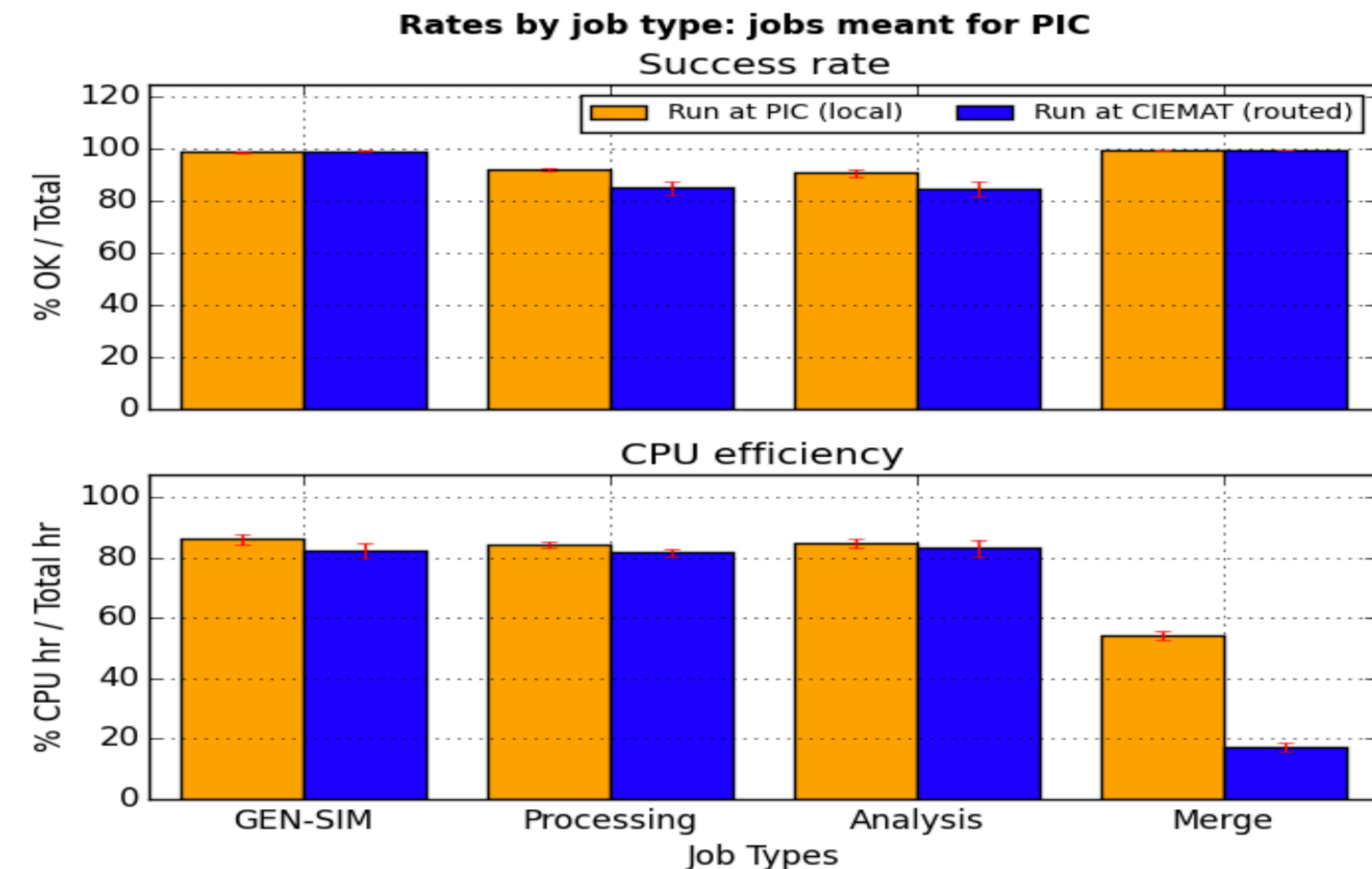
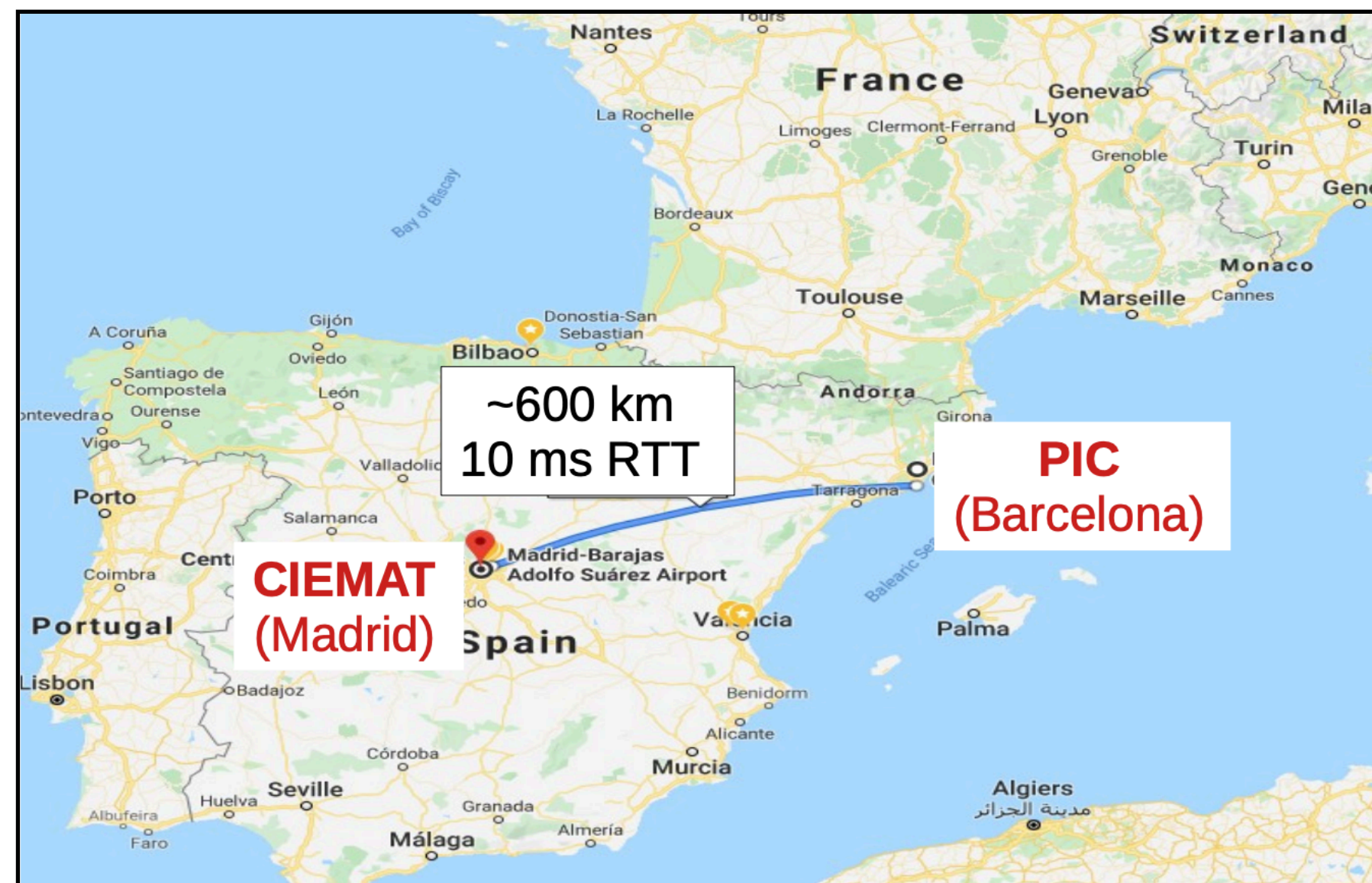
We had been running a SKA1-LOW simulation on SUMMIT using

4,561 nodes
27,360 V100 GPUs

essentially the whole machine available at the time of launching the job...

Lightweight site federation

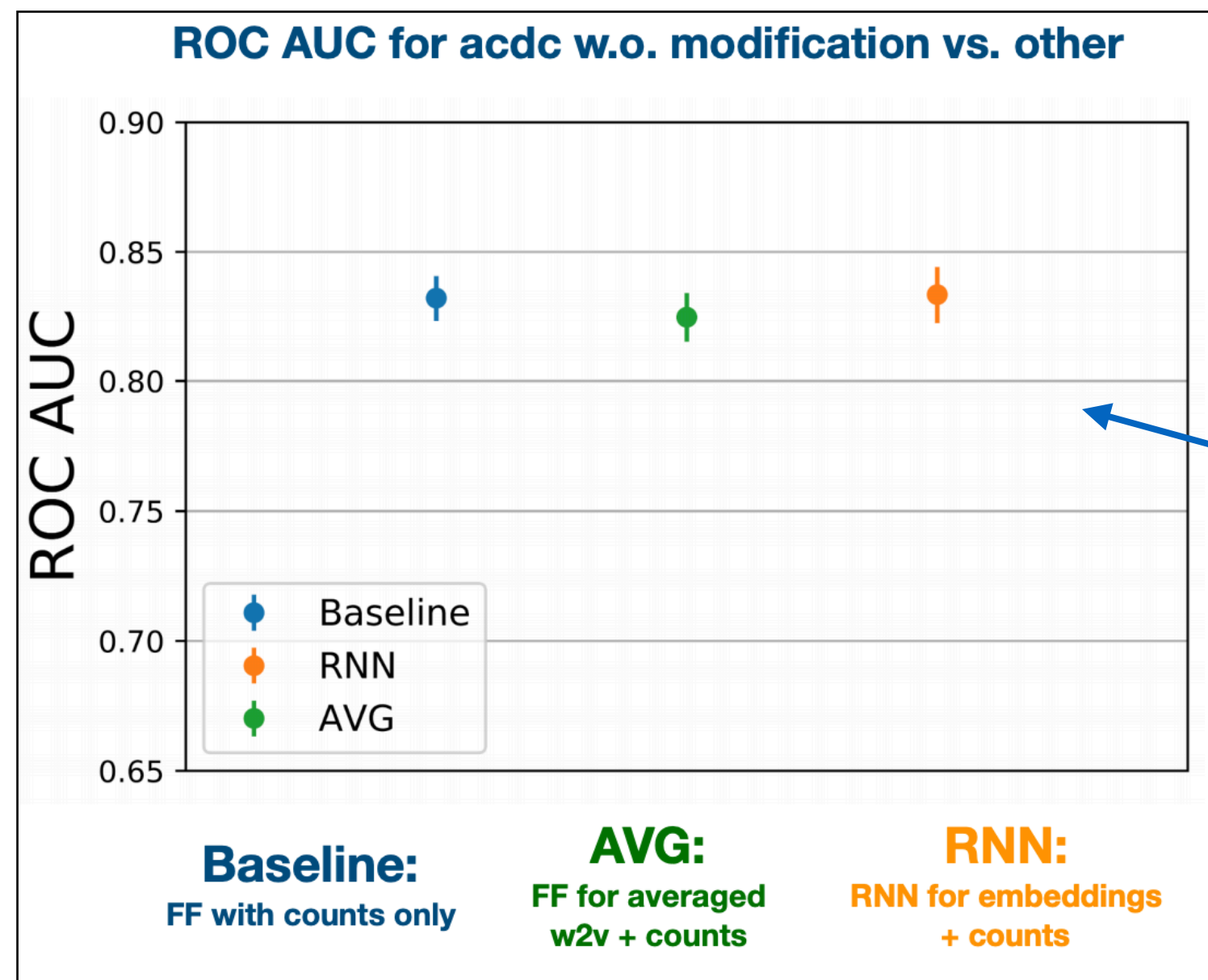
- ▶ WLCG trend towards the federation of resources



- ▶ 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 Learning model, which predict operator's action from error information**

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

Fri, Oct 11, 2019

System 4:43 AM
@Yoji Hasegawa and 2 others joined the team.

Rucio OpInt

webhook BOT 11:08 AM
I am the Rucio bot and detected some potential issues affecting transfers.

- During the last 3 hours the transfers efficiency on site IL-TAU-HEP is below 50% (39.325843%)

Main error (20.220588%): SOURCE [70] srm-ifce err: Communication error on send, err: [SE][Ls][]
http://dpm.lhep.unibe.ch:8446/srm/managerv2: CGI-gSOAP running on fts707.cern.ch reports could not open connection to dpm.lhep.unibe.ch:8446\\
No GGUS/TEAM ticket found so far
Please have a look

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”

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%

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

The screenshot shows a Sonic Pi terminal window with the following code:

```

19 use_synth :gnoise
20 n = sync "/osc/ping"
21 puts n
22 # use_synth :blade
23 # play :re3, release: 0.02, amp: 1
24 use_synth :pulse
25 play 80, attack: 0, sustain: 0, decay: 0.03, release: 0.05, res: 0.95, cutoff: 80
26 end
27
28 live_loop :badness do
29 use_real_time
30 n = sync "/osc/badness"
31 puts n
32 use_synth :prophet
33 #play 30, attack: 0, sustain: 1, decay: 0.1, release: 0.1, res: 0.91, width: 2, sustain_level: 0.1
34 play 30, attack: 0, decay: 0.05, release: 1.0, res: 0.9, amp: 0.2
35 end
36
37 live_loop :goodness do
38 use_real_time
39 n = sync "/osc/goodness"
40 puts n
41 #sample :misc_cineboom
42 duration = Math.log10([n[0], 1].max)
43 #duration = [duration, 1].min
44 print duration
45 use_synth :hollow
46 play 50, attack: 0, decay: 0.05, release: duration, res: 0.9, amp: 0.5
47 #play 30, attack: 0, sustain: 1, decay: 0.1, release: 0.1, res: 0.91, width: 2, sustain_level: 0.1
48 end

```

Below the code, there is a graph titled "Mono" showing a fluctuating signal. To the right of the graph, there is a log window showing the following output:

```

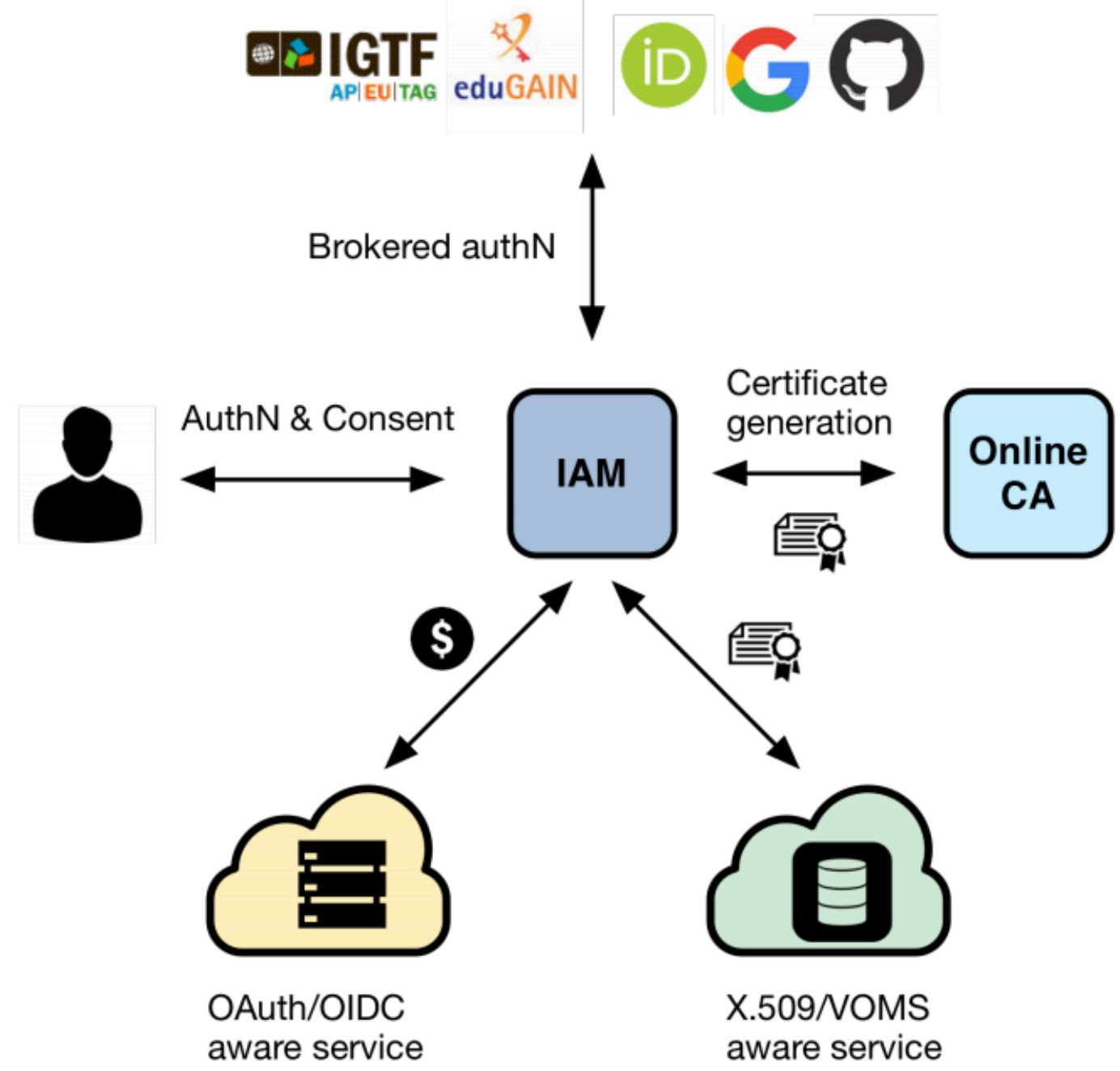
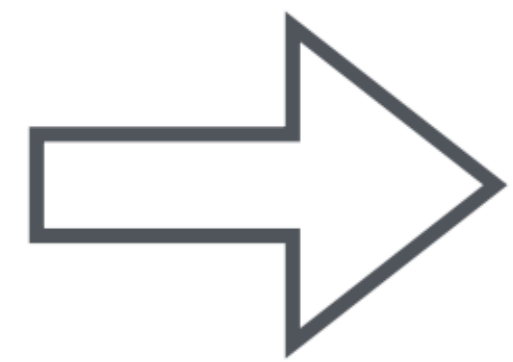
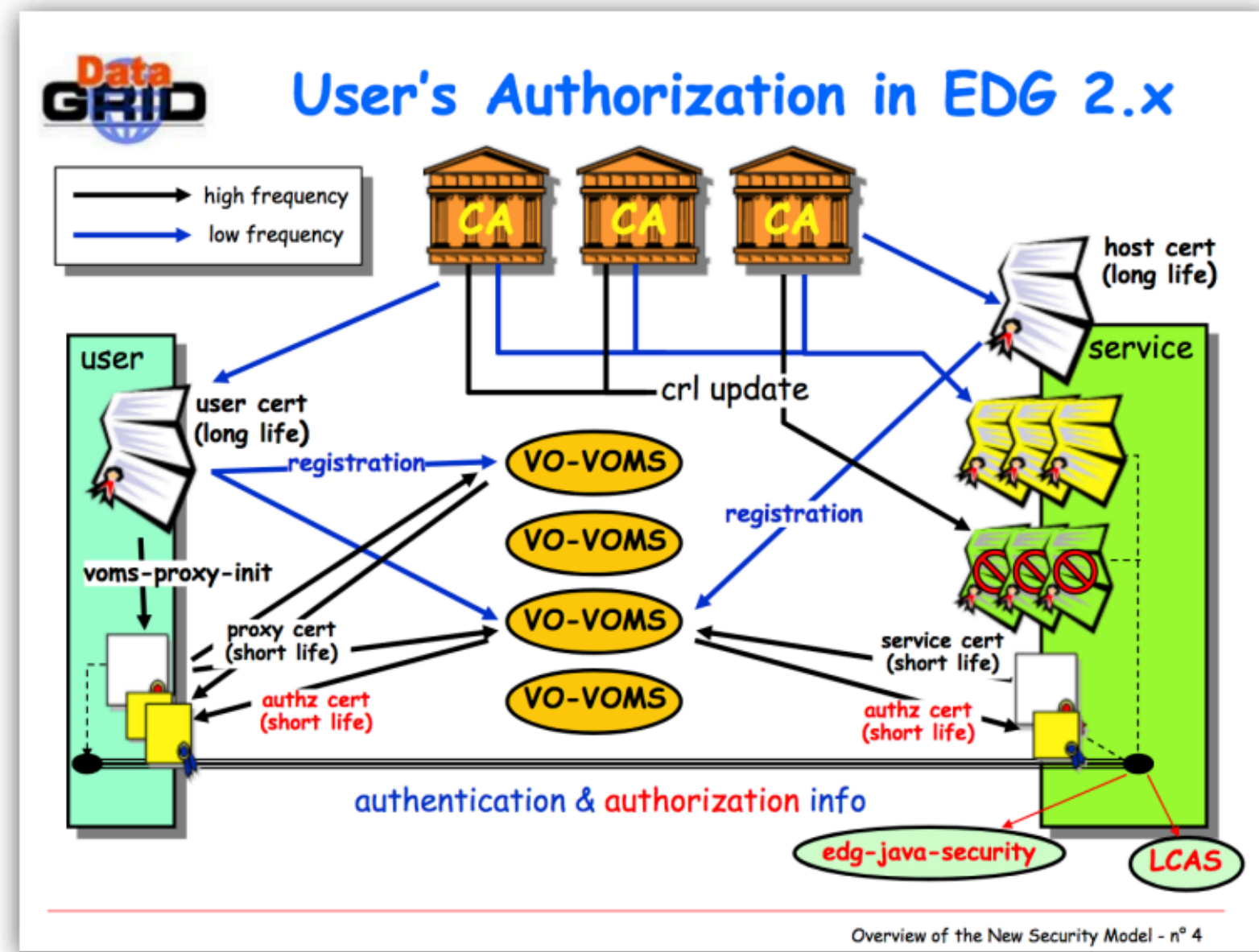
[33, 66222, -0.38476438999176025, 90.42853546142570]
1.5185139398778875
{run: 1, time: 1216.3497, thread: :live_loop_goodness}
[57, 2153, -0.48587466955184937, 88.91180849316406]
1.7558748556724915
{run: 1, time: 1216.4388, thread: :live_loop_goodness}
[52, 3772, 0.1477447748184204, 97.21617126464844]
1.7160033436347992
{run: 1, time: 1216.5303, thread: :live_loop_goodness}
[57, 2159, -0.48587466955184937, 88.91180849316406]
1.7558748556724915

```

At the bottom of the terminal, there is a list of running processes for the subtlenoise project, including various analysis and monitoring tasks.

Identity

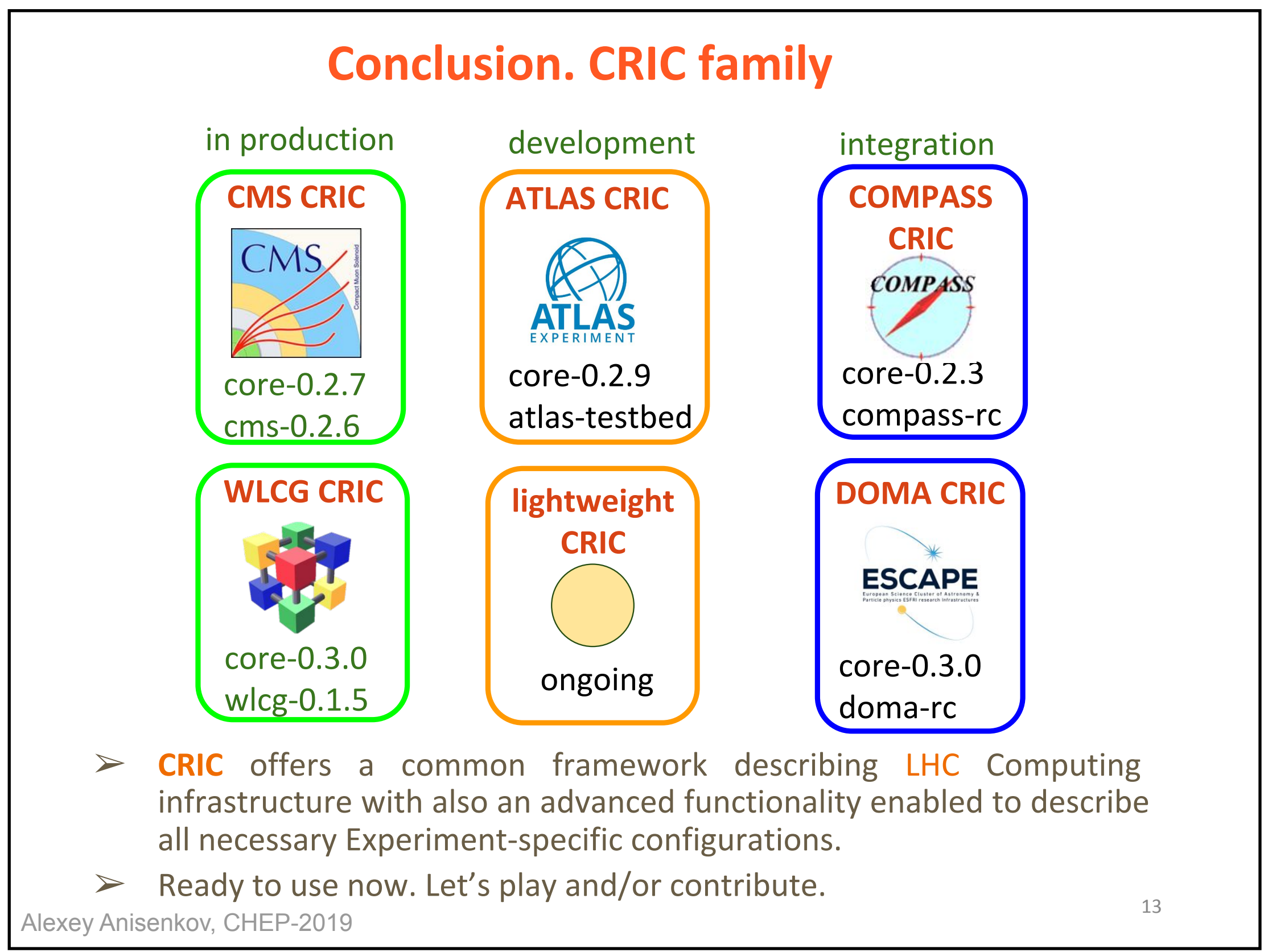
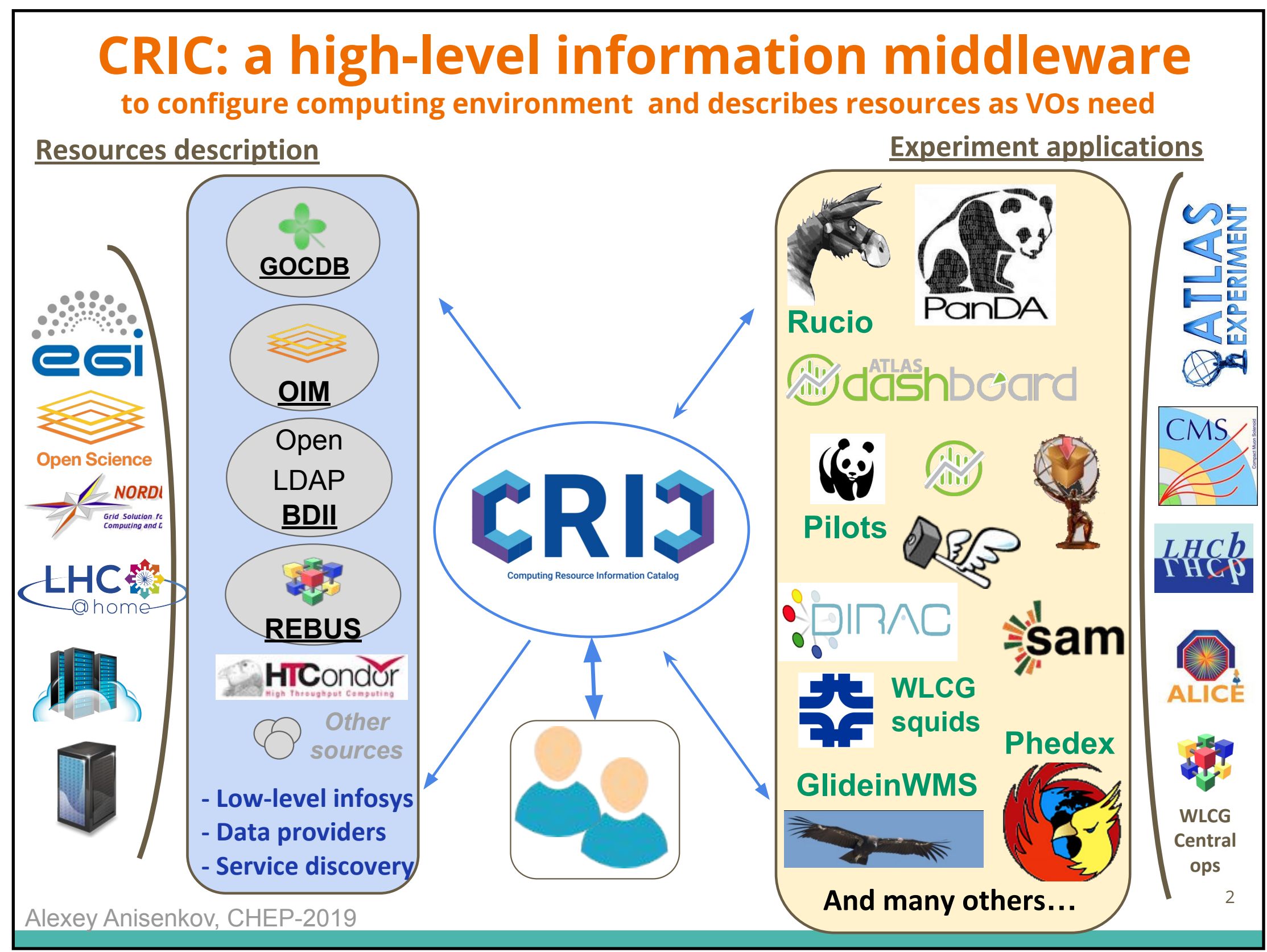
- ▶ WLCG authorization from x509 to Tokens is a hot topic



- 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

Site cost Tier-1 survey

- Launched in September 2018 a survey among Tier-1 sites (and open to Tier-2s) to understand their costs for CPU, disk and tape
 - Questionnaire available [here](#)
 - Eight Tier-1s and one Tier-2 answered
- Average costs
 - CPU: €10.3/HS06, -12%/y
 - Disk: €126/TB, -15%/y
 - Tape: €22/TB, -14%/y
- 20-50% spread among sites

Source: R. Vernet



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!