Status of WLCG FCPPL project

- Status of Beijing site
- Activities over last year
- Ongoing work and prospects for next year

CHEN Gang & POGGIOLI Luc
FCPPL 2018, Marseille

Last year activities on one page

- Grid Operation and development

  - Participation of Chinese & French colleagues to monthly ATLAS and technical computing French meetings + vidyo/Skype/…

  - Sharing of expertise and tools : grid middleware and experiment specific **(Xiaofei YAN (闫晓飞)**

  - Network monitoring (**Fazhi QI/齐法制**)

- New: Hong-Kong joining! New Tier-2 for ATLAS.
  CUHK, HKUST, HKU **Ming-chung Chu (朱明中)**, **Luis Flores Castillo, Roger Wong**

- HPCs for ATLAS simulation

  - Collaboration between IHEP, CNIC and European partners

  - IHEP **Wenjing WU (伍 文静), Xiaofei YAN (闫晓飞)**

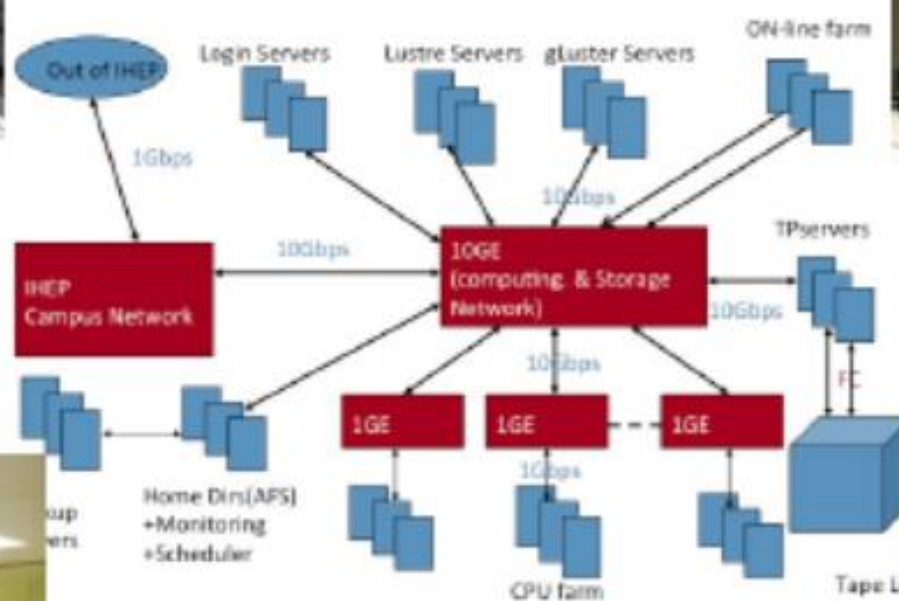- ATLAS@home: Volunteer computing project **Wenjing WU (伍 文静)**

Beijing Computing infrastructure (Local, Cloud, WLCG)

~18000 CPU cores
Water cooling

5PB tape library

10PB disk space

Power supply system

ATLAS Tier-2:
- 450 slots, 400TB
- Significant increase planned

Beijing Tier-2 - status
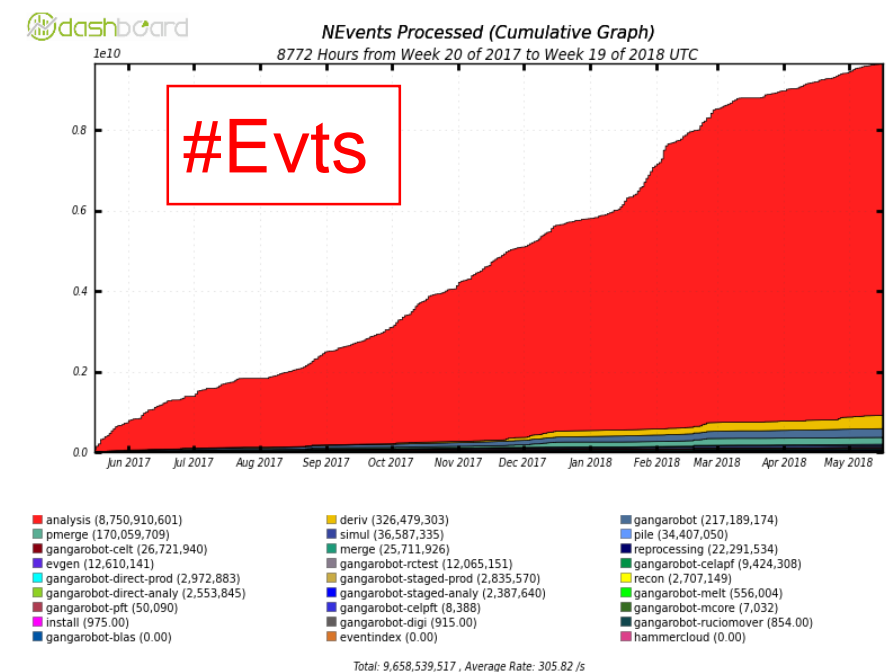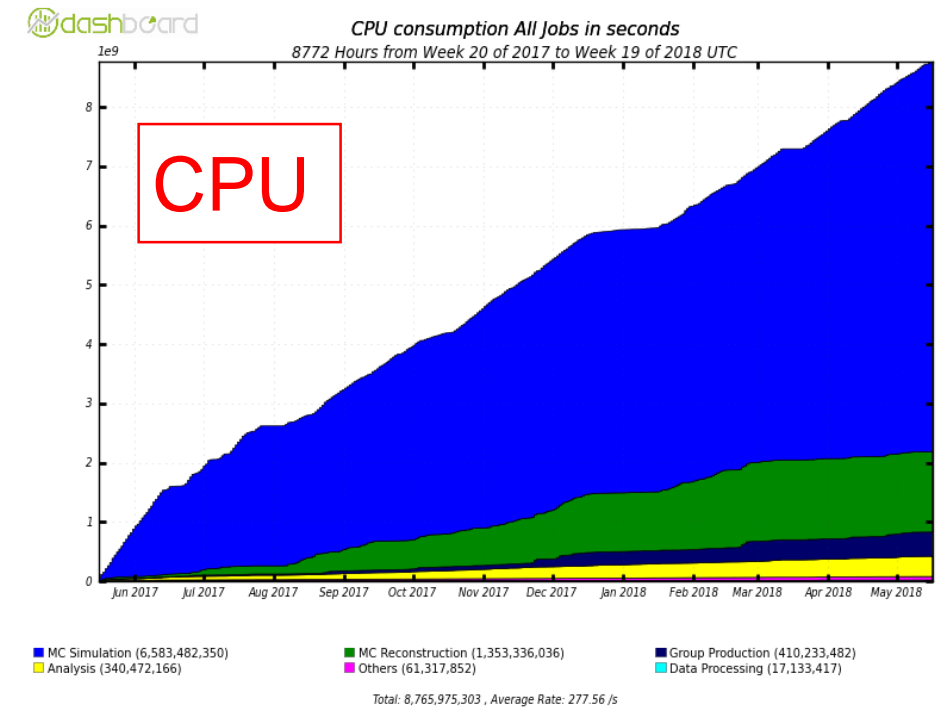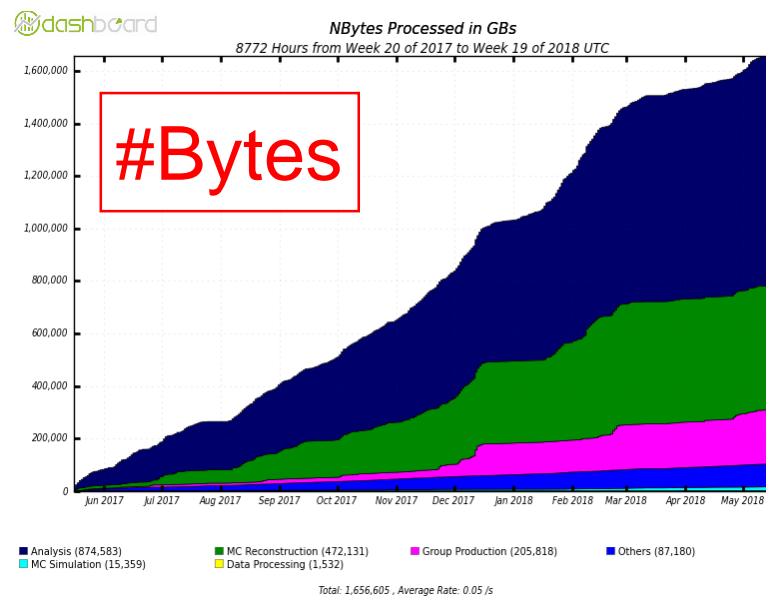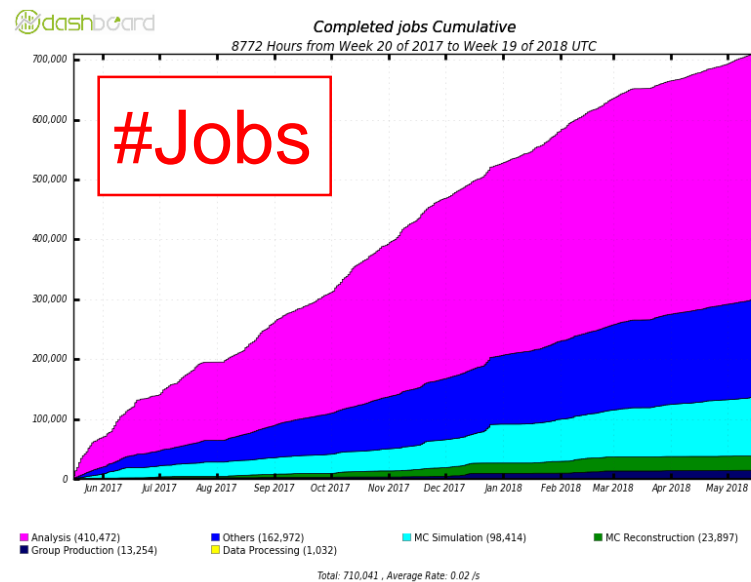

From 2017/05-2018/05

Site availability: Excellent over last year (**96.00%**)

Site reliability: **98.7%**, ranked at 45th of 130 ATLAS sites

Data Processed at Beijing ATLAS Tier-2

‣ **2.5Million** CPU hours

‣ **0.70 Million** jobs completed

‣ **1.6PB** data, **10Billion** events processed

Completed jobs Cumulative
8772 Hours from Week 20 of 2017 to Week 19 of 2018 UTC

#Jobs

■ Analysis (410,472)   ■ Others (162,972)   ■ MC Simulation (98,414)   ■ MC Reconstruction (23,897)
■ Group Production (13,254)   ■ Data Processing (1,032)

Total: 710,041 , Average Rate: 0.02 /s

CPU consumption All Jobs in seconds
8772 Hours from Week 20 of 2017 to Week 19 of 2018 UTC

CPU

■ MC Simulation (6,583,482,350)   ■ MC Reconstruction (1,353,336,036)   ■ Group Production (410,233,482)
■ Analysis (340,472,166)   ■ Others (61,317,852)   ■ Data Processing (17,133,417)

Total: 8,765,975,303 , Average Rate: 277.56 /s

NBytes Processed in GBs
8772 Hours from Week 20 of 2017 to Week 19 of 2018 UTC

#Bytes

■ Analysis (874,583)   ■ MC Reconstruction (472,131)   ■ Group Production (205,818)   ■ Others (87,180)
■ MC Simulation (15,359)   ■ Data Processing (1,532)

Total: 1,656,605 , Average Rate: 0.05 /s

NEvents Processed (Cumulative Graph)
8772 Hours from Week 20 of 2017 to Week 19 of 2018 UTC

#Evts

■ analysis (8,750,910,601)   ■ deriv (326,479,303)   ■ gangarobot (217,189,174)
■ pmerge (170,059,709)   ■ simul (36,587,335)   ■ pile (34,407,050)
■ gangarobot-celt (26,721,940)   ■ merge (25,711,926)   ■ reprocessing (22,291,534)
■ evgen (12,610,141)   ■ gangarobot-rctest (12,065,151)   ■ gangarobot-celapf (9,424,308)
■ gangarobot-direct-prod (2,972,883)   ■ gangarobot-staged-prod (2,835,570)   ■ recon (2,707,149)
■ gangarobot-direct-analy (2,553,845)   ■ gangarobot-staged-analy (2,387,640)   ■ gangarobot-melt (556,004)
■ gangarobot-pft (50,090)   ■ gangarobot-celpft (8,388)   ■ gangarobot-mcore (7,032)
■ install (975.00)   ■ gangarobot-digi (915.00)   ■ gangarobot-ruciomover (854.00)
■ gangarobot-blas (0.00)   ■ eventindex (0.00)   ■ hammercloud (0.00)

Total: 9,658,539,517 , Average Rate: 305.82 /s

Interface of HPCs from CAS

▸High Performance Computers

- **Large** number of spare CPU cycles at HPCs & not used by 'standard' HPC applications

  - Promising opportunistic resource
  - Perfectly adapted to ATLAs Monte-Carlo simulation jobs (low I/O, CPU intensive)

- **China** host some of the largest HPC facilities worldwide

  • Collaboration between European collaborators, IHEP & CNIC (Computer Network Information Center)

  • HPC centres from CAS interconnected through ScGrid

  • Active since Fall 2015

  • Event simulation on HPCs & Storage at IHEP

中国科学院
计算机网络信息中心
Computer Network Information Center
Chinese Academy of Sciences

Supercomputing Center of Chinese Academy of Sciences

Scientific Computing Grid (ScGrid)

☐ Supercomputing Environment in CAS

☐ computing resources
  - 300T flops CPU
  - 3000T flops GPU

☐ 30+ Institutes from CAS
  - SCCAS: Head Center
  - 8 branch centers
  - 18 regional centers
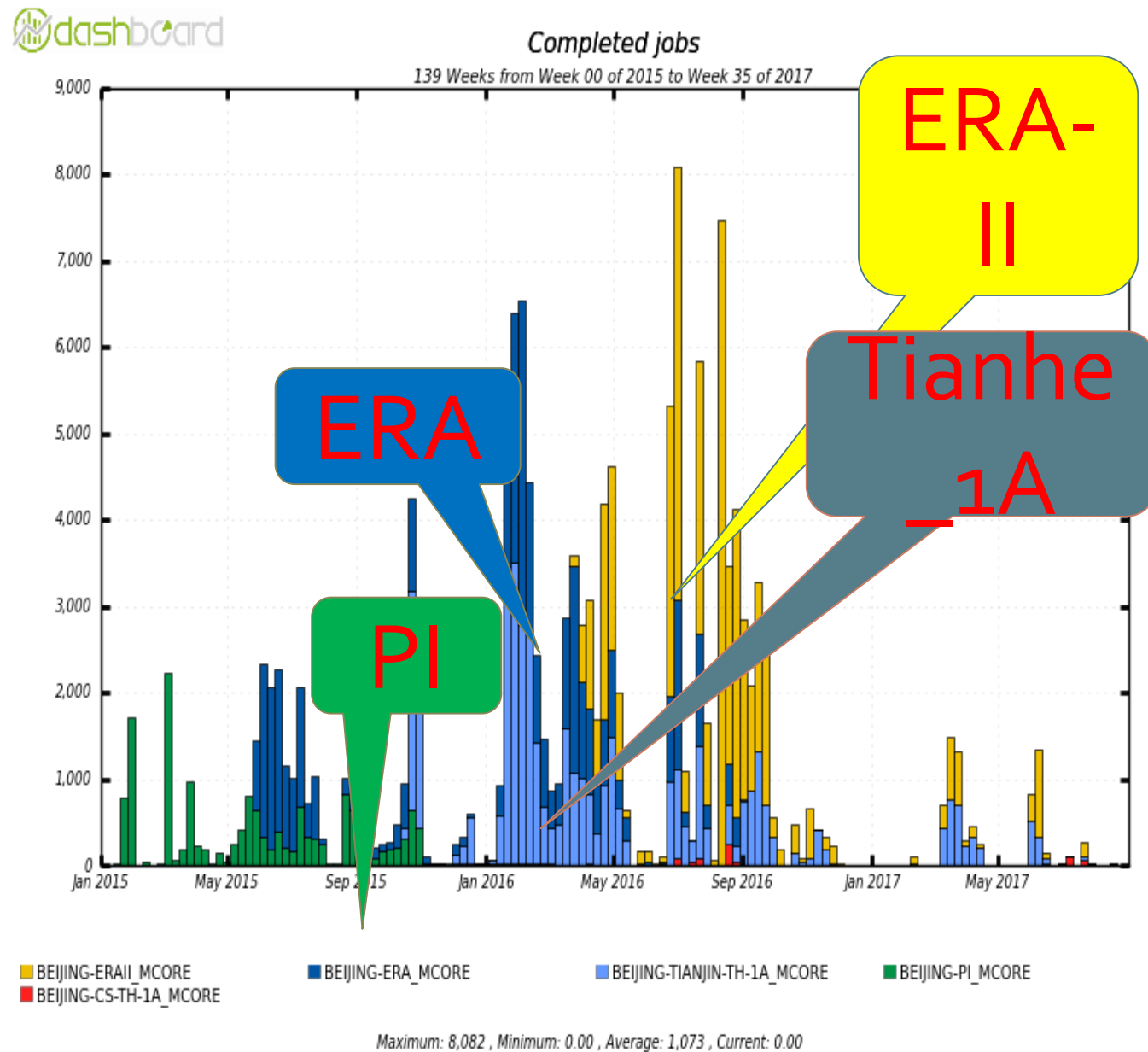 - 11 GPU clusters

21

Huairou campus of CNIC

Chinese: 元（Yuan）

English: ERA
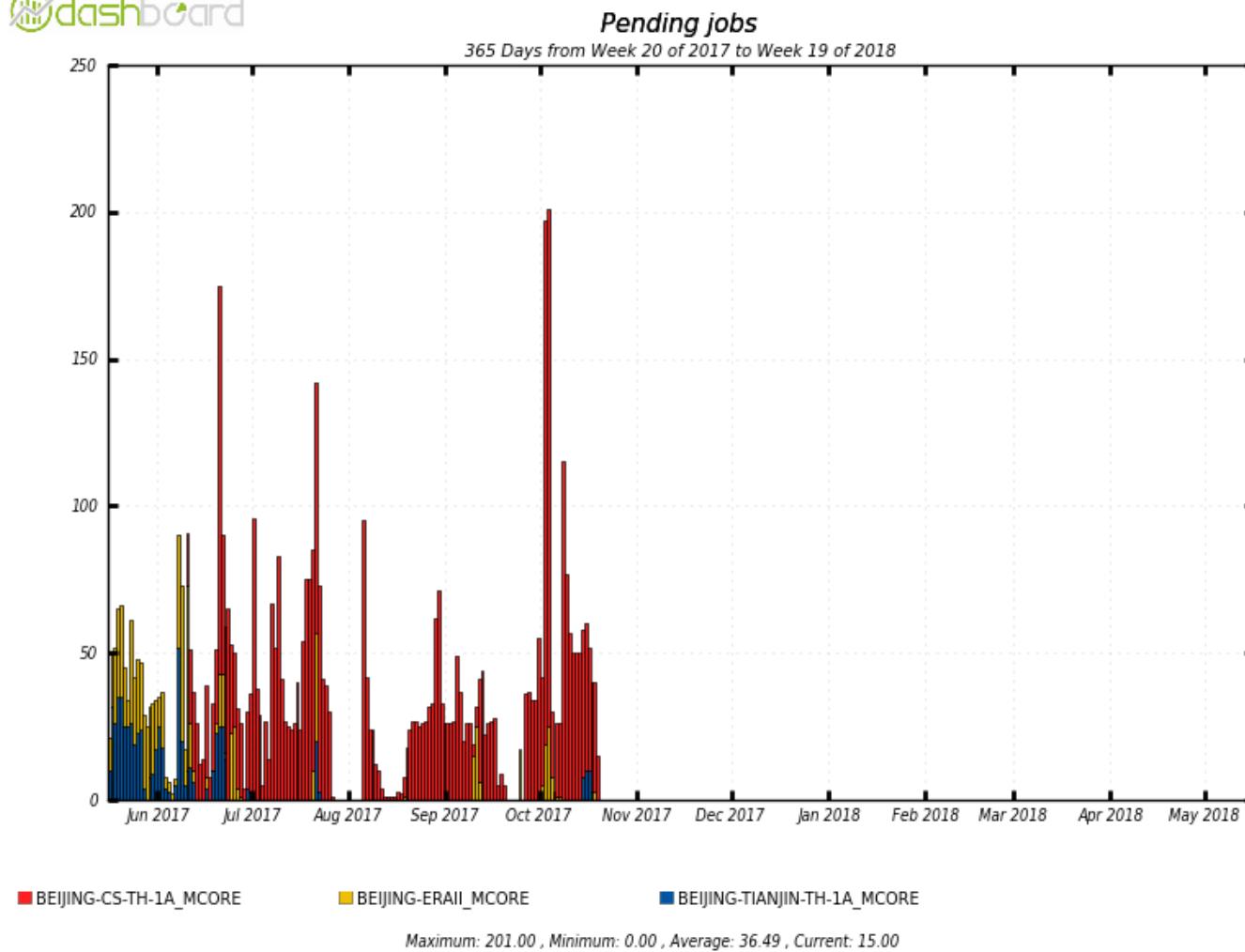
Tianhe-1A :TOP500 2nd, June 2011

Chinese HPC usage
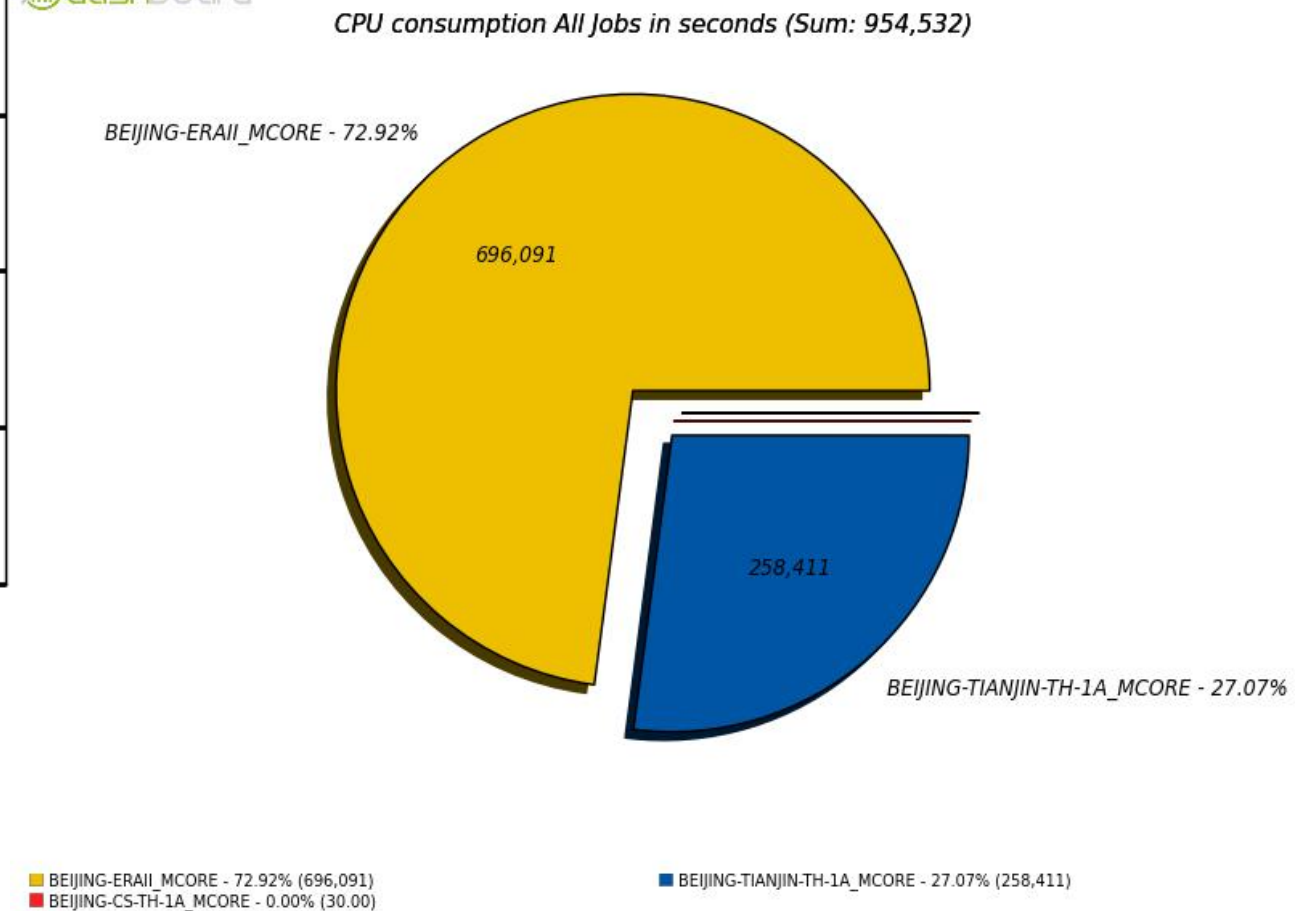

‣ Running Mcore jobs
  24 core (Era-II)
  12 core (Tianhe_1A)

‣ 4 HPCs were used, 2 remain active

‣ Max. used of Era-II: **1400** cores/day
  Max. used of Tianhe_1A: **600** cores/day

‣ Currently active:
  Era-II and Tianhe_1A
  Both PI and Era stopped providing

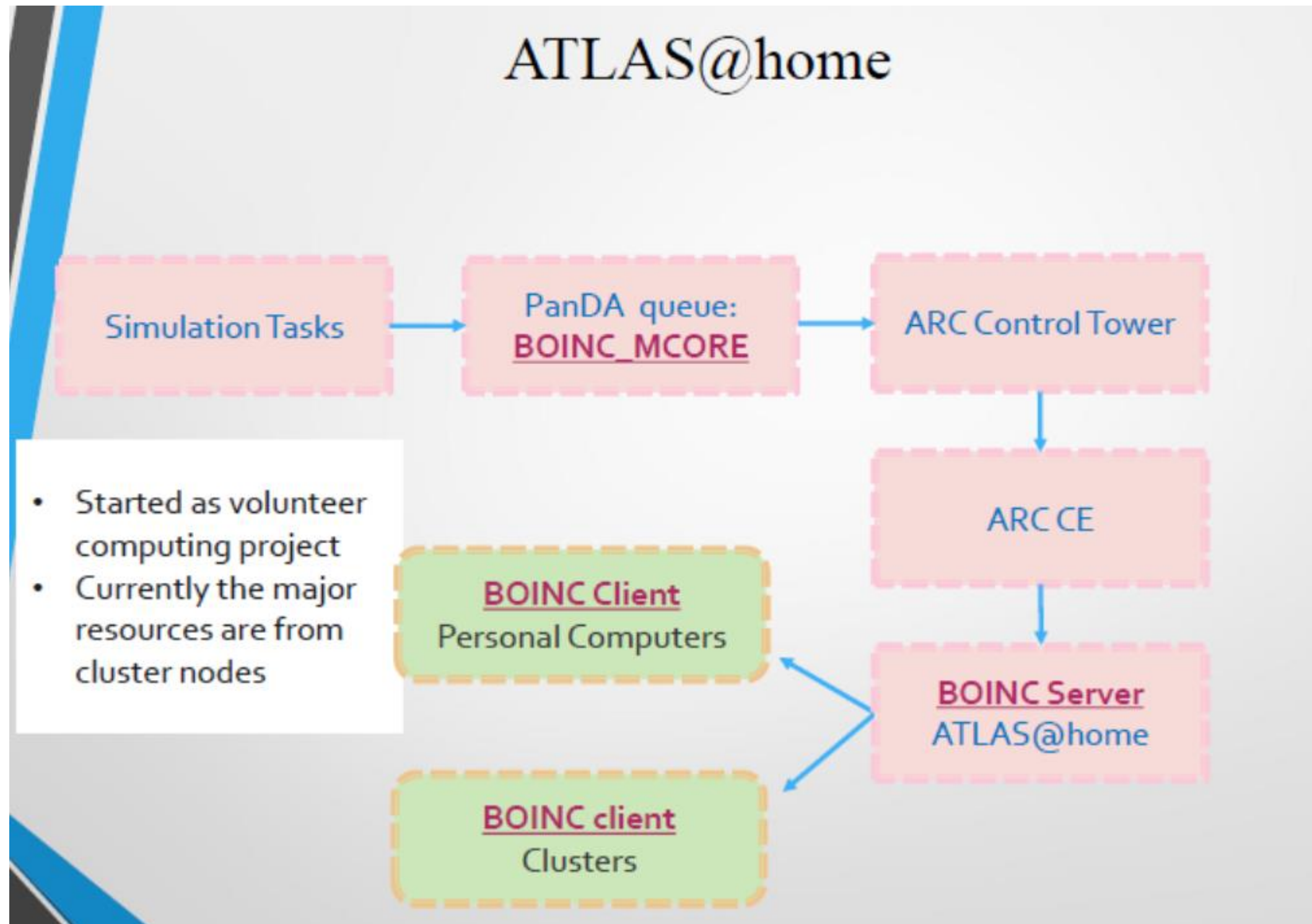# IHEP (ATLAS) HPC production activity since May 2017

Pending jobs

CPU consumption

9

Workplan on HPCs

- Chinese HPCs can provide a considerable amount of CPUs for ATLAS simulation
- Was not fully used last year due to lack of manpower investment
  - Technical issues
  - Communication with HPC centers
  - Moving to "Production" becomes the issue (cost, availability)
    - True for **ALL** HPCs throughout the world
- Hopefully can be improved

  -Reactivate the usage of current available 2 HPCs (solve software distribution issues)
  - Add 1 or more HPCs to the Free resource pool
  - French-Chinese joint effort via FCPPL can definitely help
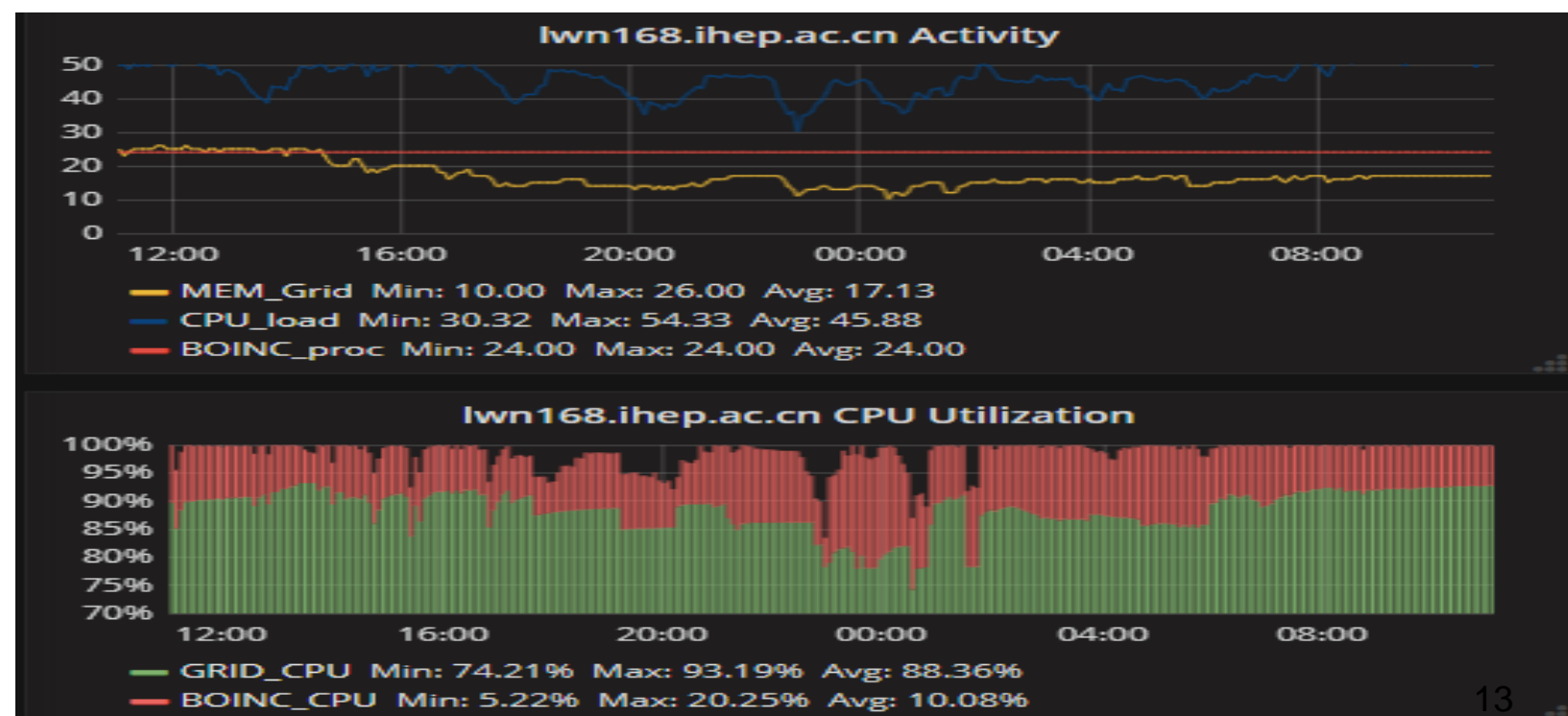
ATLAS@HOME

ATLAS@Home: Current scale in ATLAS

l jobs of _All ATLAS sites_ in the past 10 days

CPU consumption Good Jobs in seconds (Sum: 219,860,593,718)

BOINC:
3.52%

Avg. BOINC core power: 11HS06

CPU days:
8950 per day

✓ Ranked as the 6th site in terms of good CPU time
✓ Equivalent to a site with **140KHS06**, given the avg. CPU util. of ATLAS site is less than 0.7

AGLT2 - 2.81% (6,169,791,511)
CERN-P1 - 2.10% (4,623,043,072)
PRAGUELCG2 - 2.01% (4,419,683,011)
TOKYO-LCG2 - 1.73% (3,802,277,094)
NDGF-T1 - 1.58% (3,484,470,235)
DESY-HH - 1.53% (2,347,041,000)

BU_ATLAS_TIER2 - 2.17% (4,768,698,655)
FZK-LCG2 - 2.06% (4,525,046,454)
SWT2_CPB - 1.95% (4,277,432,763)
SIGNET - 1.71% (3,761,534,275)
LRZ-LMU - 1.54% (3,387,474,454)

ATLAS@home as a simulation site

NEvents Processed in MEvents (Million Events) (Sum: 463.00)

MWT2 - 12.17%

BOINC: 8.83%
Avg. **4.1M** events per day

CERN-PROD - 13.20%

✓ Runs only simulation jobs, CPU intensive
✓ Ranked as the 3rd site for simulation jobs
✓ Over 50% of the ATLAS computing resources are used by simulation

CERN-PROD - 13.20% (61.00)
BOINC - 8.83% (41.00)
CERN-P1 - 3.54% (16.00)
LBNL_DSD_ITB - 2.96% (14.00)
TOKYO-LCG2 - 2.59% (12.00)

MWT2 - 12.17% (56.00)
BNL-ATLAS - 7.35% (34.00)
IN2P3-CC - 3.09% (14.00)
OLCF - 2.51% (12.00)
MPPMU - 2.09% (10.00)

Perfectly adapted to Monte-Carlo simulation jobs
- Low I/O, CPU intensive
- Thanks to 'Event Service', possible to optimize theses opportunistic resources by processing 1 event at a time

12

New: Backfilling of ATLAS sites

‣ Sites are not always fully used (downtime, task brokering glitches, switching from single to multi core resource)

‣ Idea: Run 2 jobs on each core

   - 1 grid job with normal priority, 1 BOINC job with the lowest priority

   - Linux uses "non preemptive" scheduling for CPU cycles, which means high priority jobs occupies CPU until it releases the CPU voluntarily

‣ Example: Beijing Tier-2 on 100 days

Grid jobs *Walltime Util.* is **87.8%**,
Grid CPU Util. is **65.6%**,

BOINC exploit an **extra 23%** of CPU time,

Node CPU Util. reaches **89%**

lwn168.ihep.ac.cn Activity

MEM_Grid Min: 10.00 Max: 26.00 Avg: 17.13
CPU_load Min: 30.32 Max: 54.33 Avg: 45.88
BOINC_proc Min: 24.00 Max: 24.00 Avg: 24.00

lwn168.ihep.ac.cn CPU Utilization

GRID_CPU Min: 74.21% Max: 93.19% Avg: 88.36%
BOINC_CPU Min: 5.22% Max: 20.25% Avg: 10.08%

13

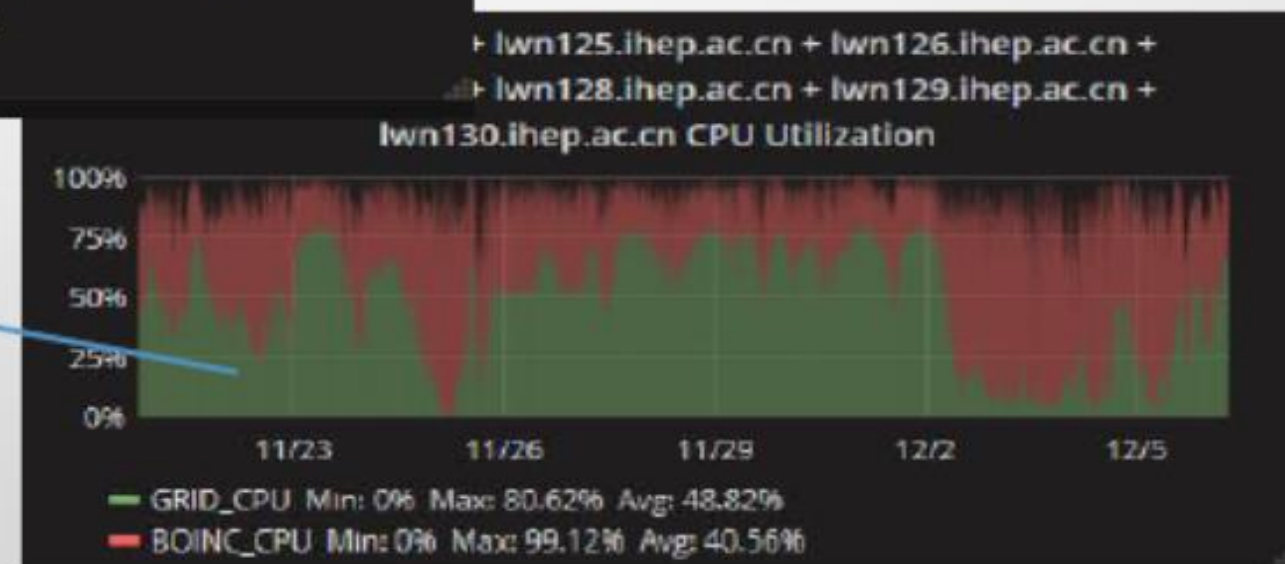# One step further:
Dynamical BOINC ReSource Configuration

‣ Automatize and dynamically allocate BOINC jobs on top of standard Grid jobs

On the same set of nodes (ATLAS single core analysis nodes), CPU Util. is increased by **15%** (from **74.4 % to 89.4%**) while the Grid workload is the same (48%). **15% increase of BOINC CPU**
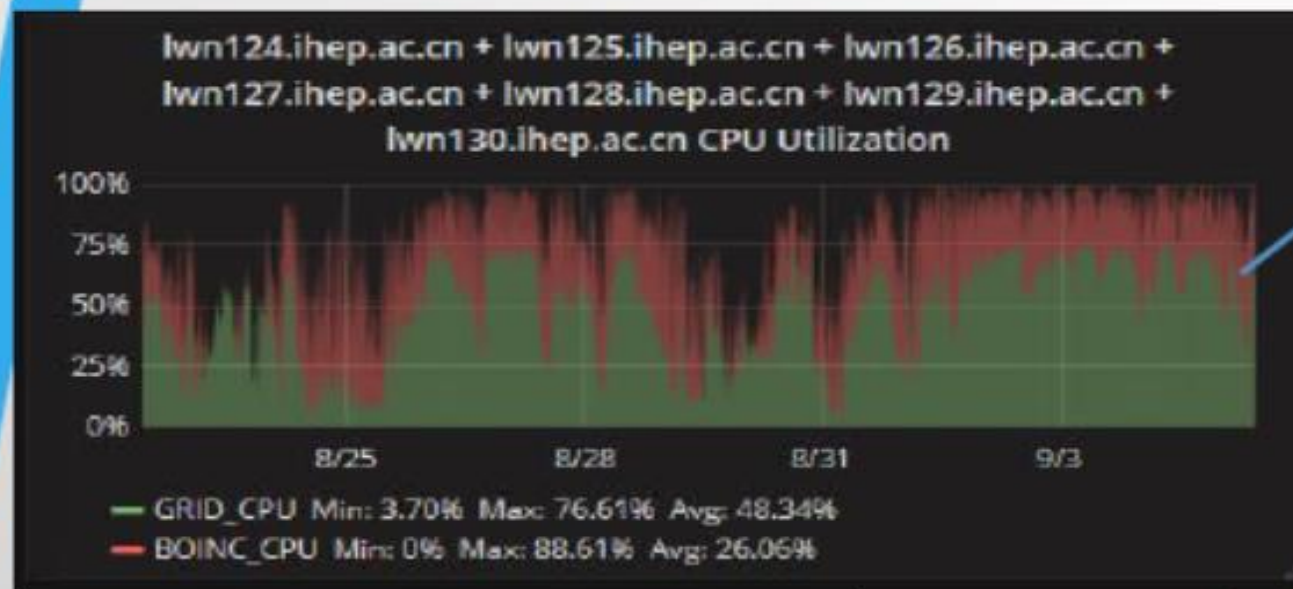
ATLAS@home prospects

- ATLAS@home is becoming a big resource contributor to ATLAS, and the resource is stable and reliable

- Backfilling on the BEIJING ATLAS grid site exploit an extra of 28% CPU(6 months), on regular cluster is 46%.

- Sites are encouraged to use ATLAS@home to harness their non official ATLAS computing resources and Backfilling running it on the clusters.

- Dynamical BOINC configuration makes sure:

  - Efficiently exploit the available resource

  - Not to affect the Non BOINC jobs

  - Hide the BOINC details to site admins.

Hong-Kong Tier-2

‣ Experimental Physics group in HK: Cluster of 3 universities

- Joined ATLAS in 2014

‣ HK-LCG2: Collaborative project of the 3 universities

- Located at CUHK

- Got the green light to become a ATLAS Tier-2 in December 2017

‣ Resources

- 1,008 cores and 436TB disk storage (1pB at Summer)

- Will add 700 cores & 500TB in 3 years

- Manpower: 6 IT staff (manages also Central HPC platform and some smaller computing clusters)
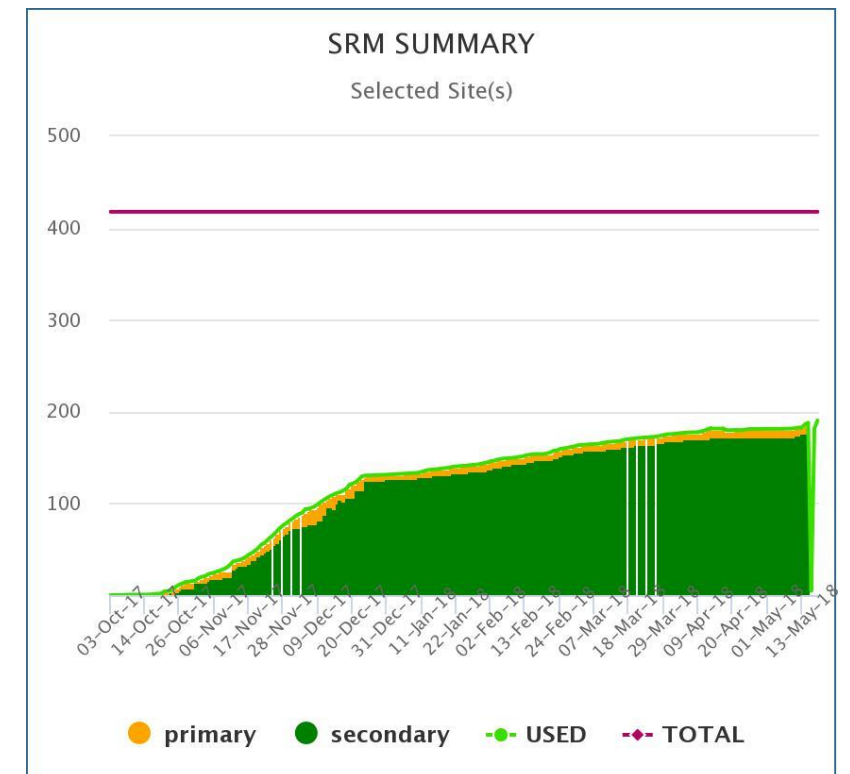
Hong-Kong Tier-2

‣ Setting-up started end 2015

- Continuous exchange between HK, Beijing IHEP, French experts (At Tier-1 and various Tier-2s)

- Remotely, but also face to face during dedicated meetings thanks to FCPPL
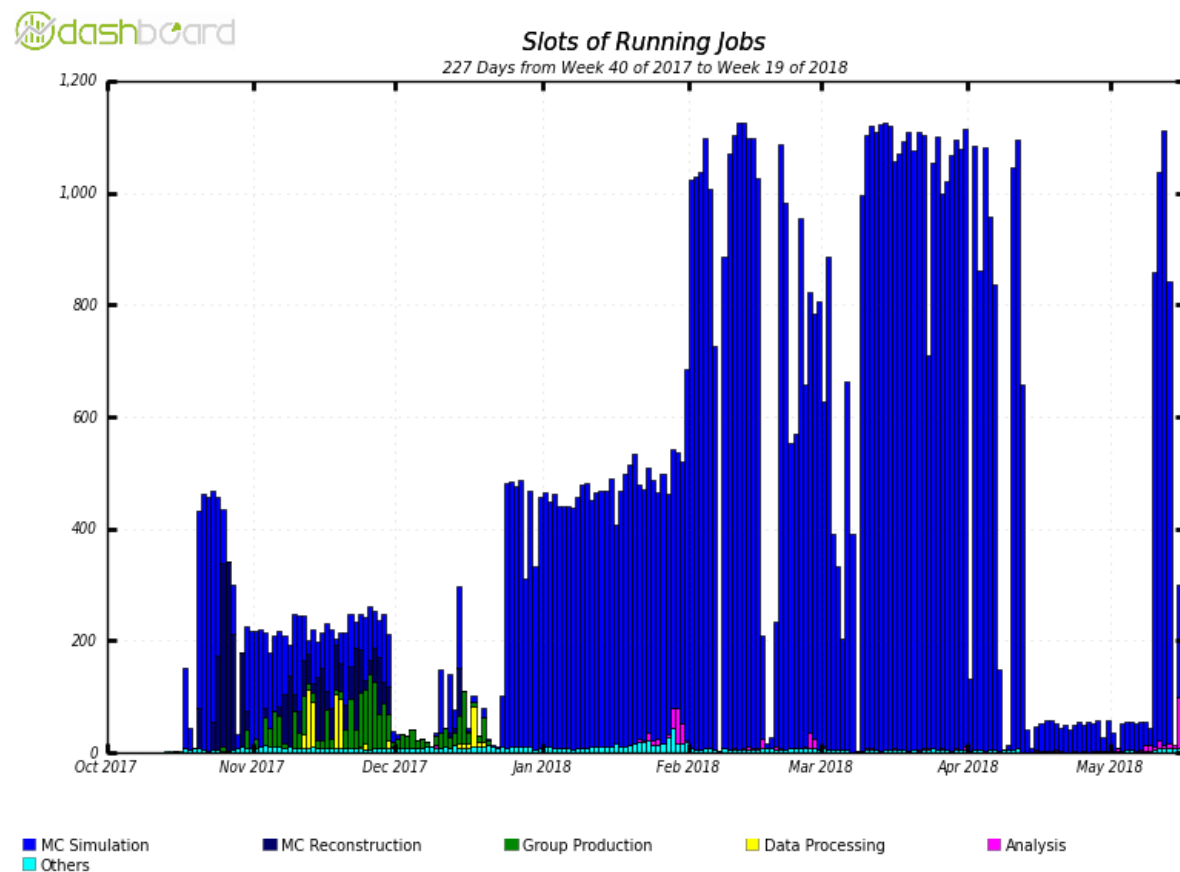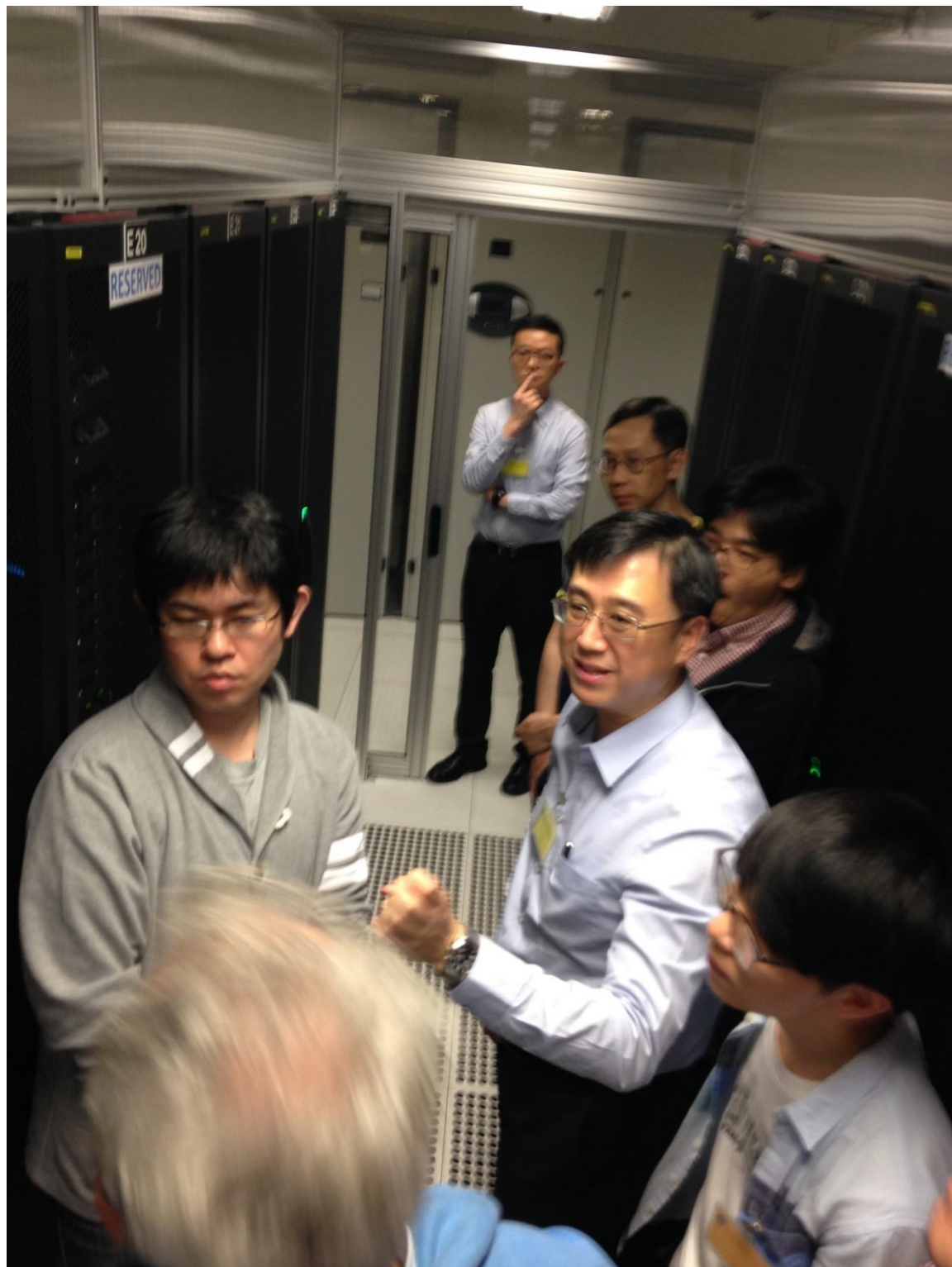

Running slots

Storage

Hong-Kong Tier-2

Visit of the Computing room, April 2018

Signing the EGI agreement with
NGI_China responsible,
April 2018

Meetings and workshops

Annual workshop of the French-cloud in Hong-Kong, April 2018
Participants from : Beijing, Hong-Kong, Tokyo, France

19