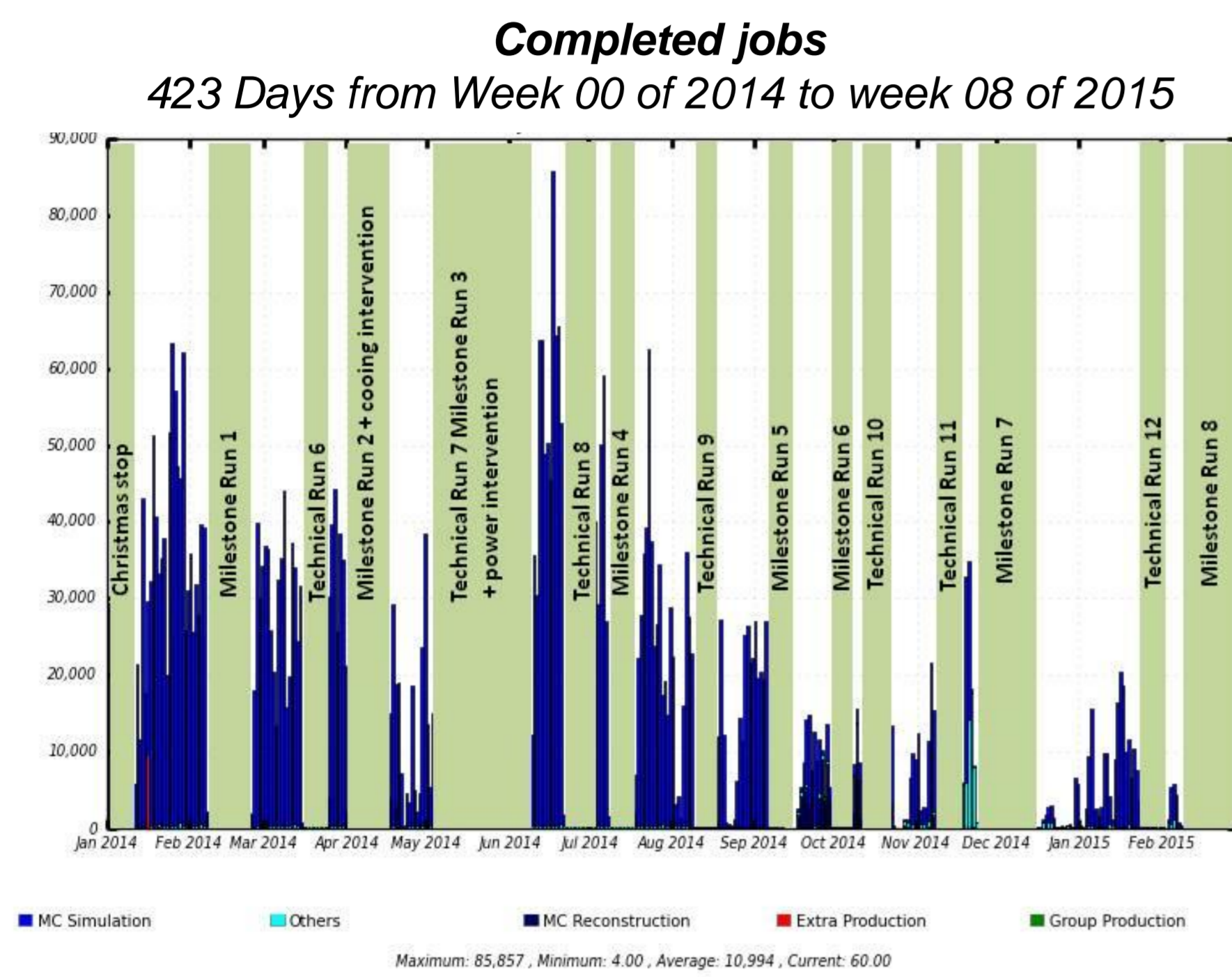


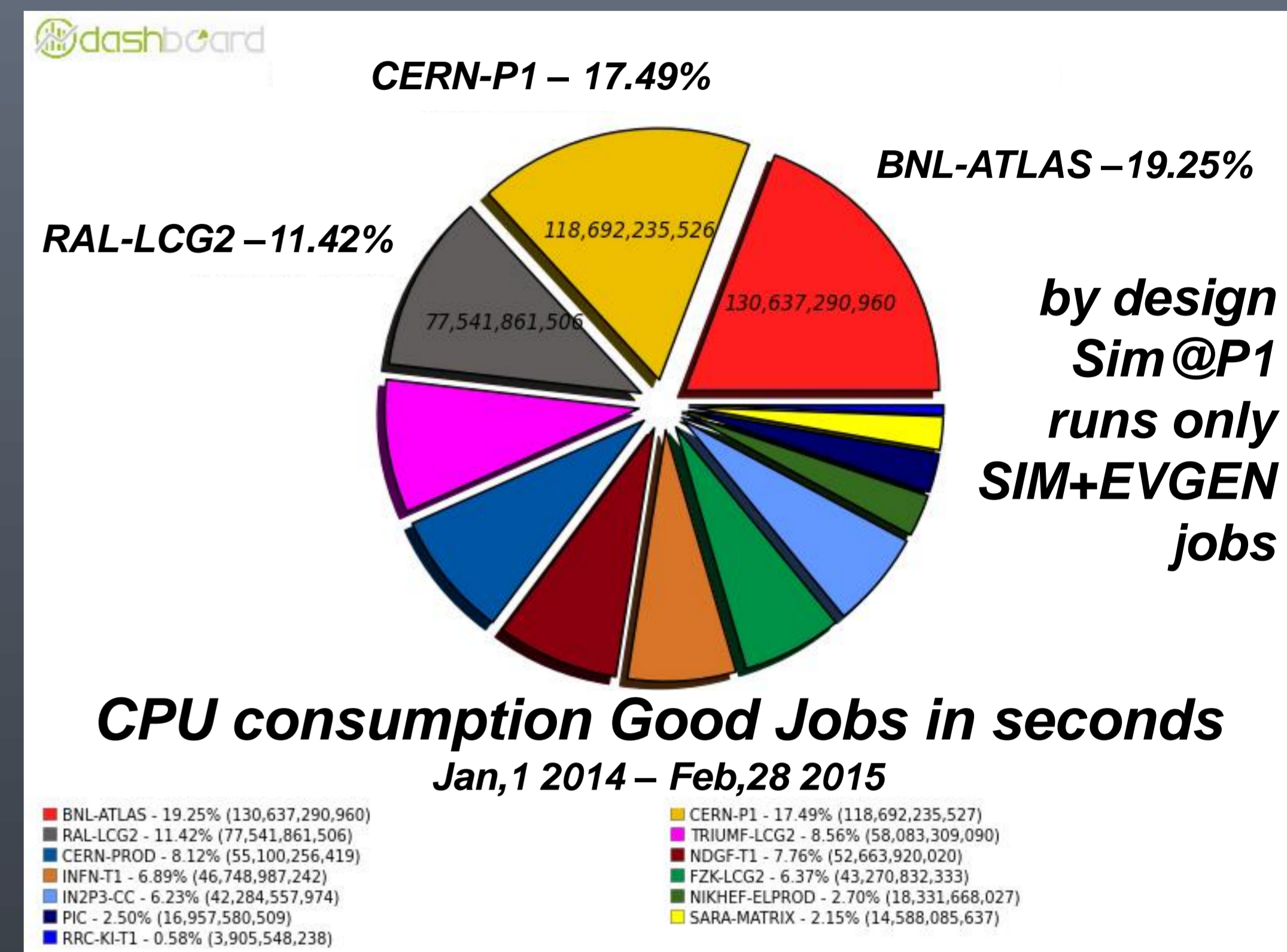
Design, Results, Evolution and Status of the ATLAS simulation in Point1 project

S. Ballestrero [1], S.M. Batraneanu [2], F. Brasolin [3], C. Contescu [2,4], D. Fazio [2], A. Di Girolamo [2], C.J. Lee [1,2], M.E. Pozo Astigarraga [2], D.A. Scannicchio [5], A. Sedov [6], M.S. Twomey [7], F. Wang [8], A. Zaytsev [9]

The simulation in Point1 project, based on OpenStack, uses in an opportunistic way the resources of the TDAQ High Level Trigger (HLT) farm of the ATLAS experiment. More than 1300 compute nodes (CNs) running up to 2700 VMs are exploited for running event generation and Monte Carlo production jobs, mostly CPU and not I/O bound, for a total of up to 22K parallel running jobs



After the setup phase, during 2014 Sim@P1 was one of the most productive ATLAS sites, although the farm was used also for other activities: milestone run, technical run, infrastructure maintenance. It delivered more than 33 million CPU-hours and it generated more than 1.1 billion Monte Carlo events since Jan 1, 2014



LHC run 2: a new challenge

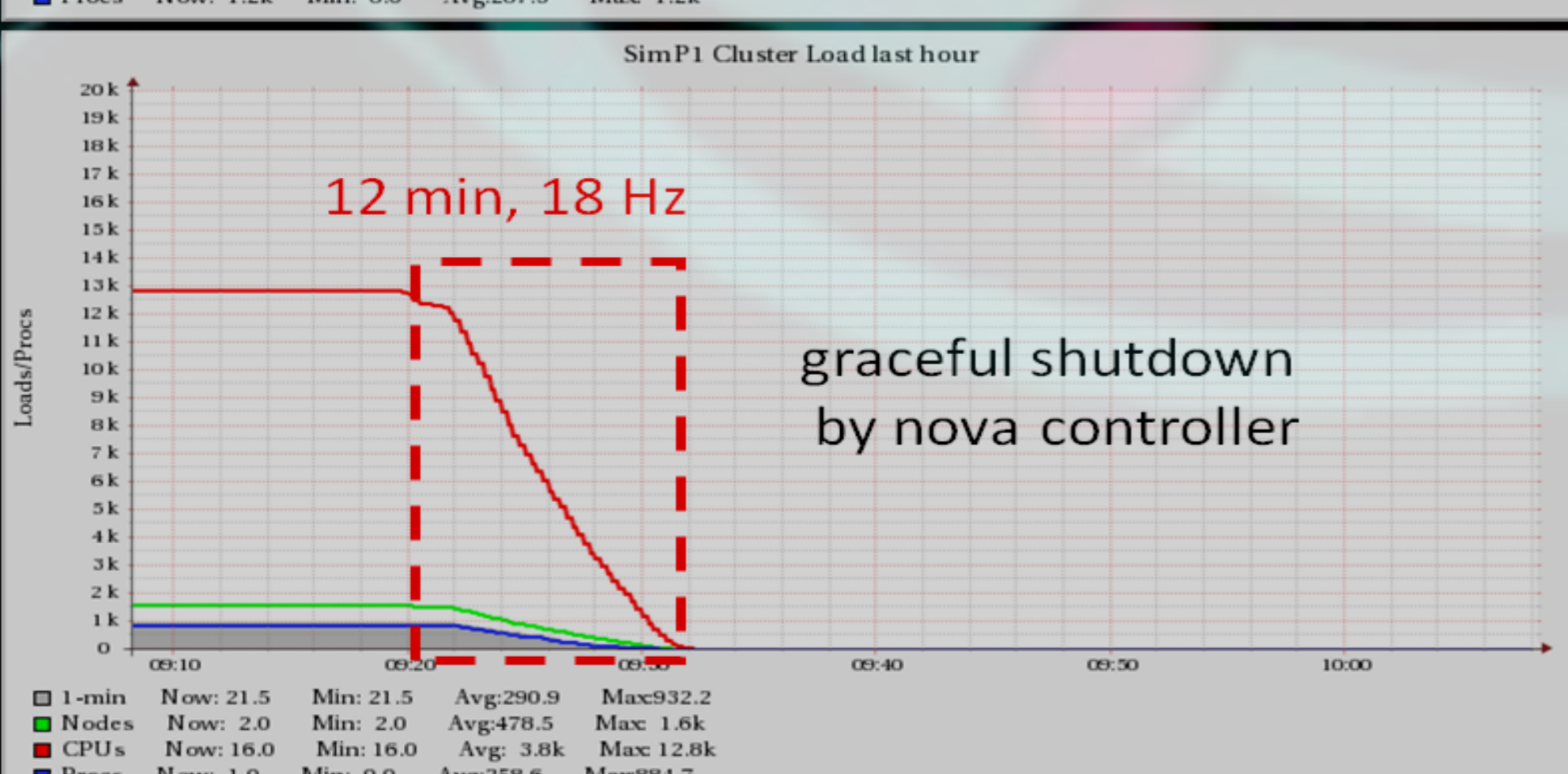
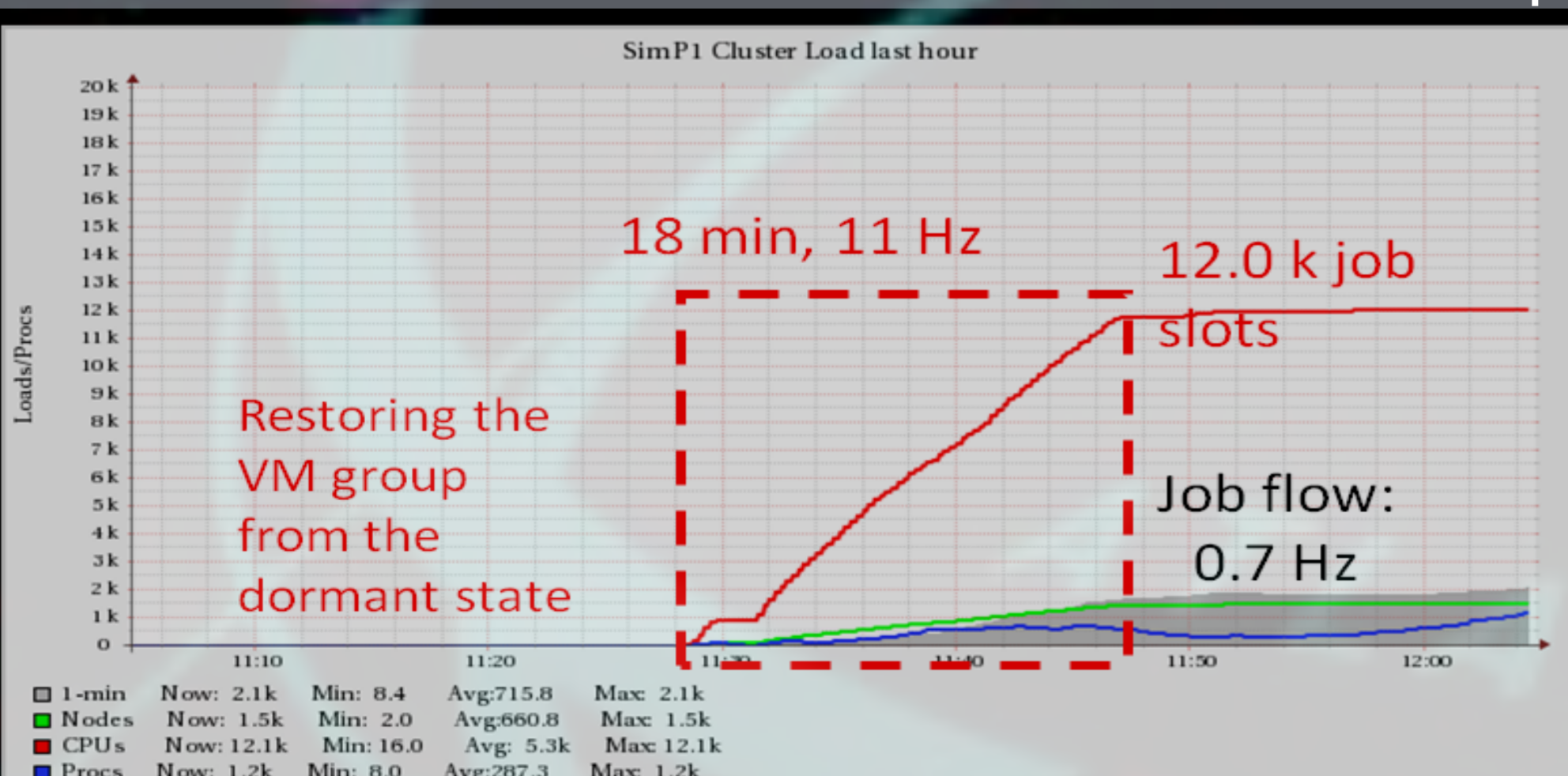
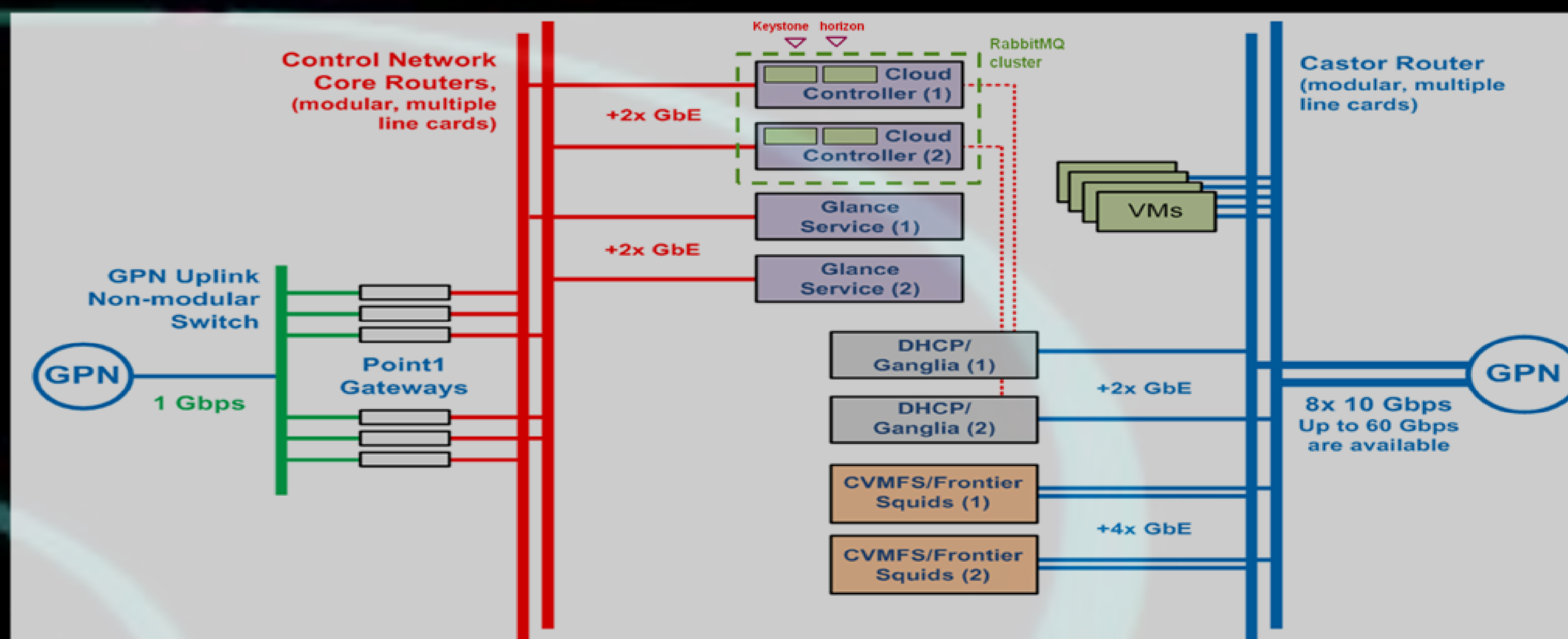
Sim@P1 will use the HLT resources when not needed for data taking, even on short notice and for short periods. New automated procedures are required for fast switching from TDAQ to Sim@P1 state and back

operations: fast switching procedures

- part of the racks are always in TDAQ mode
- the other resources are switched when a LHC stop longer than 24h is foreseen and no other tests or interventions are planned
- the switching will be performed through a WEB gui by the ATLAS Control Room shifters
- ATLAS online have the full control of the HLT farm
- the switching is completely automated and fast:**
 - from Sim@P1 to TDAQ: ~ 12 mins. Time needed by the Nova servers for a graceful VM shutdown and to update the database (DB)
 - from TDAQ to Sim@P1: ~ 1 h. Time needed for the VMs creation when the VM image is already available on the CNs, check the Nova DB, restart bad VMs, deal with bad CNs.
- This is faster than the incoming jobs flow
- emergency switch to TDAQ mode even faster: ~ 100 s.
- no VMs deletion and no Nova DB clean up

new OpenStack Icehouse vers. 2014.1.3

- components used: *Nova, Keystone, Glance, Nova-Network*
- new CLI wrappers for handling bulk operations for large VM groups instantiation/deletion
- improved security: additional RPC layer on Nova DB provided by Nova Conductor, for the price of higher CPU consumption
- ref: <https://www.openstack.org/>



dynamic partition

- implemented through HTCondor
- permit a more effective use of the available computing resources
- the same WN (VM) can run both Single-core and Multi-core jobs, via:
 - *Group fair share mechanism*: allows to set up the expected fractions that can be used by each type of job
 - *Defrag tool*: checks current requirements and allows draining and migration of the WNs from one group to another

The success of the Sim@P1 project is due to the continuous, combined efforts of the ATLAS TDAQ SysAdmins and NetAdmins teams, CERN IT SDC, the RHIC & ATLAS Computing Facility (RACF) at BNL and the University of Wisconsin-Madison. OpenStack allows the separation of the admin duties between the groups in charge of the HLT farm infrastructure, the VM group management and the Sim@P1 WLCG site operations

openstack™
CLOUD SOFTWARE