

# Sim@P1: Using Cloudscheduler for offline processing on the ATLAS HLT farm

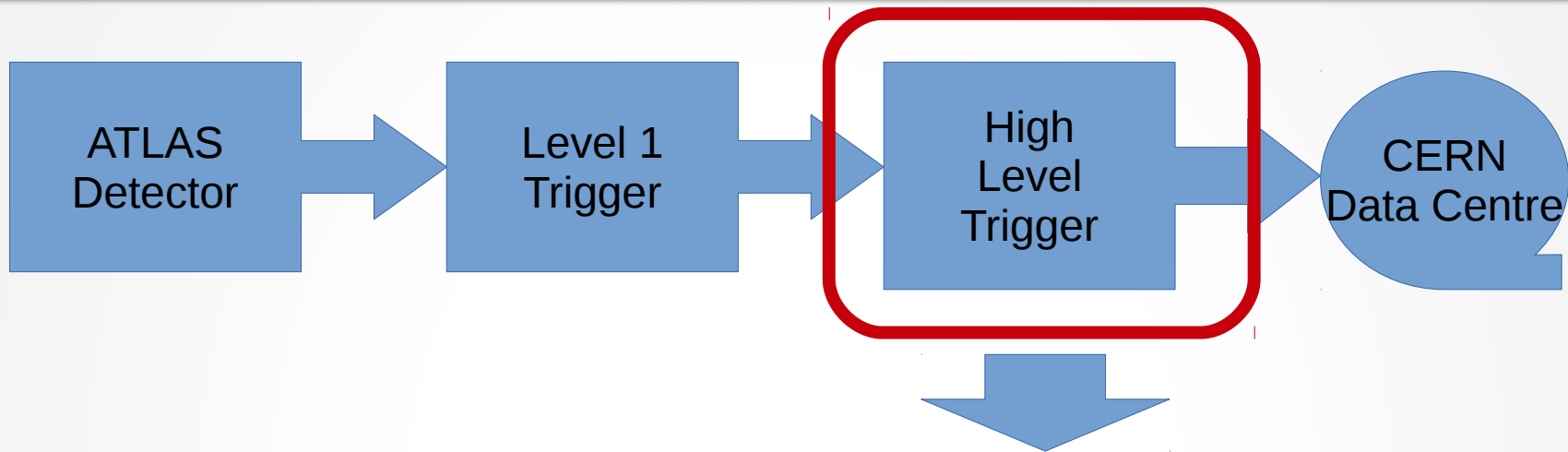
**F Berghaus** for the Sim@P1 team

on behalf of the ATLAS Collaboration

# Outline

- Definition: What is Sim@P1
- Status: Current operation
- Plan: Integration of Cloudscheduler

# What is Sim@P1?

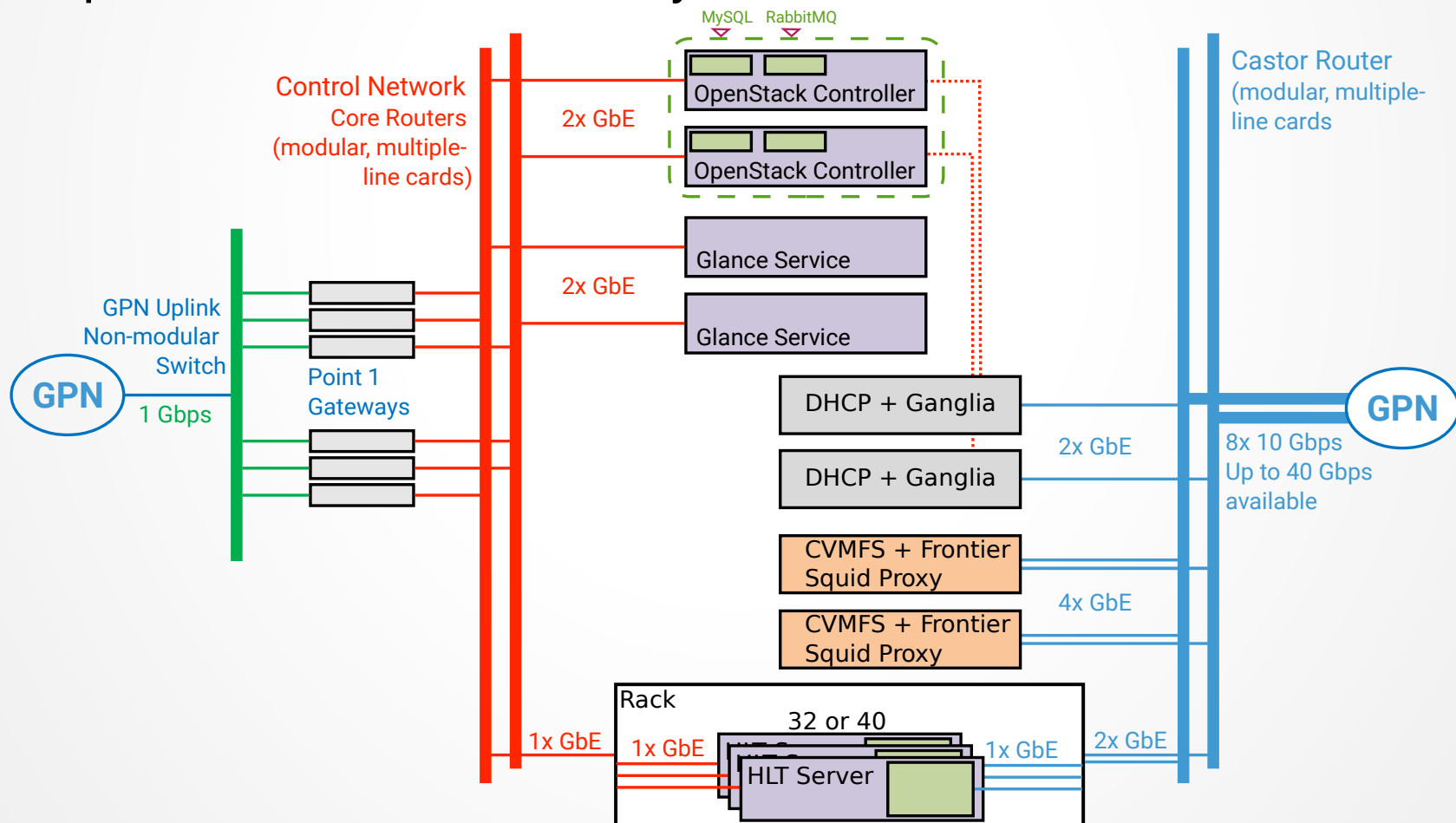


Racks	Servers per rack	Cores per node	RAM per node	RAM per core	Total cores
1-4, 6-13, 94, 95	32	16	~24 Gbyte	~1.5 Gbyte	10K
64-69	40	16			
16-26, 75-77	32	48	~64 Gbyte	~1.3 Gbyte	64K
70-74, 79-90	40	48			
44-54	40	56	64 Gbyte	~1.1 GByte	
Total: 58					74K

Sim@P1 = Simulation at point one

# Sim@P1: Current Operation

- Dedicated VLAN for *offline* access to list of hosts in CERN General Purpose Network
- Compute resources isolated by virtualization

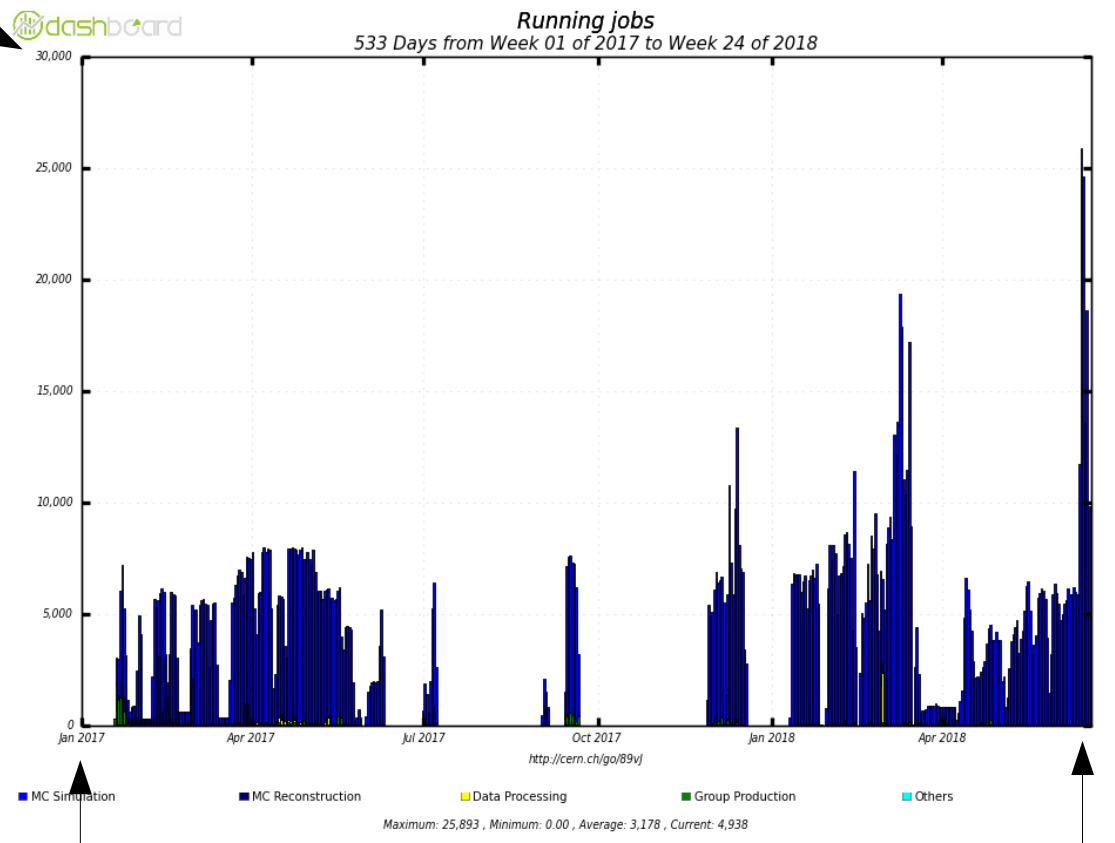


# Sim@P1: Current Operation

- Boot:
  - Puppet launches nova on worker nodes
  - Puppet executes scripts to launch instances
  - Instances connect to condor
    - CM + 2 Sched in GPN
- APF submits to each Sched
- Shutdown:
  - Puppet kills nova on worker nodes
  - Puppet calls cleanup scripts

30k

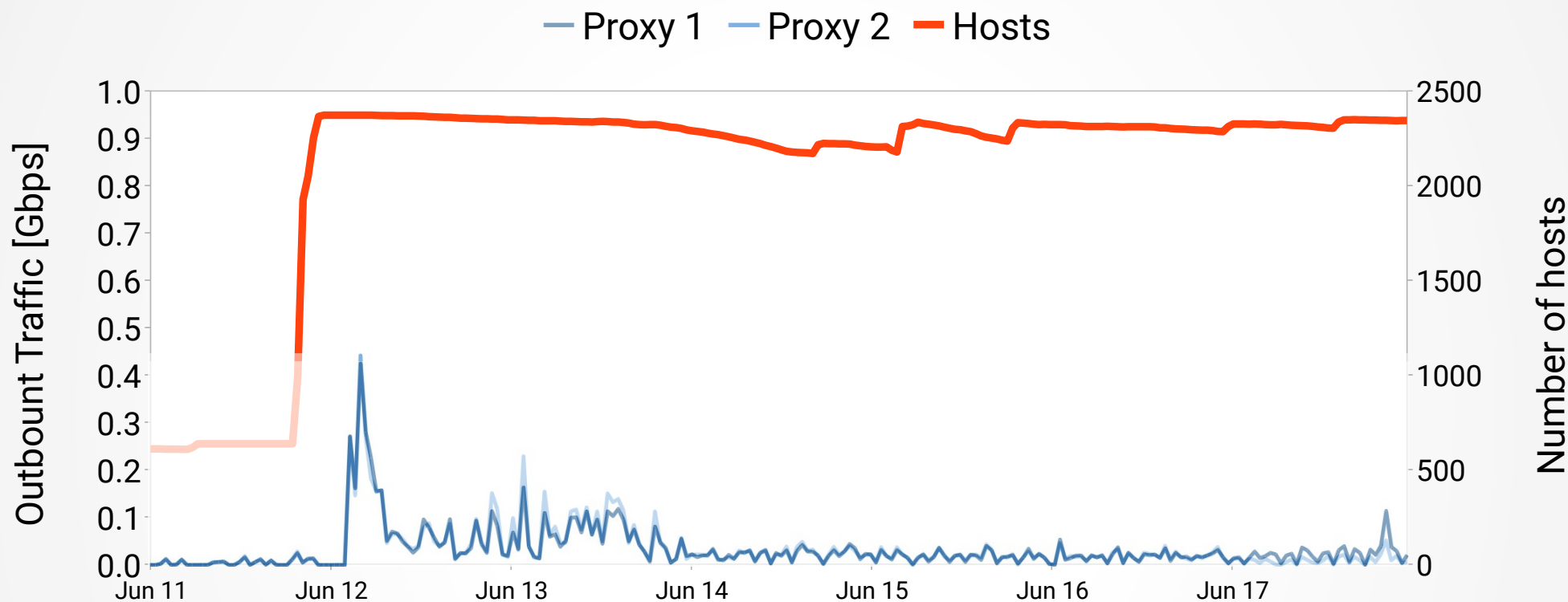
dashboard



Jan 2017

Jun 2018

# CernVM at Point 1



- 20MB CernVM3 micro-kernel distributed from glance
- CernVM3 caches in ATLAS software and operating system
- Two SQUID servers in P1 are sufficient to provide software

# Issues with current operation

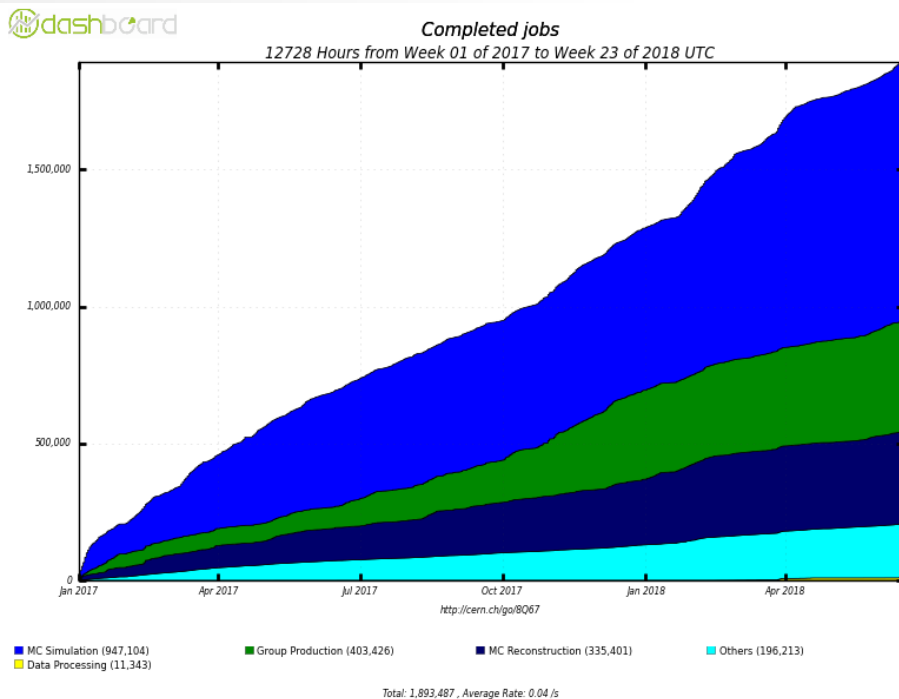
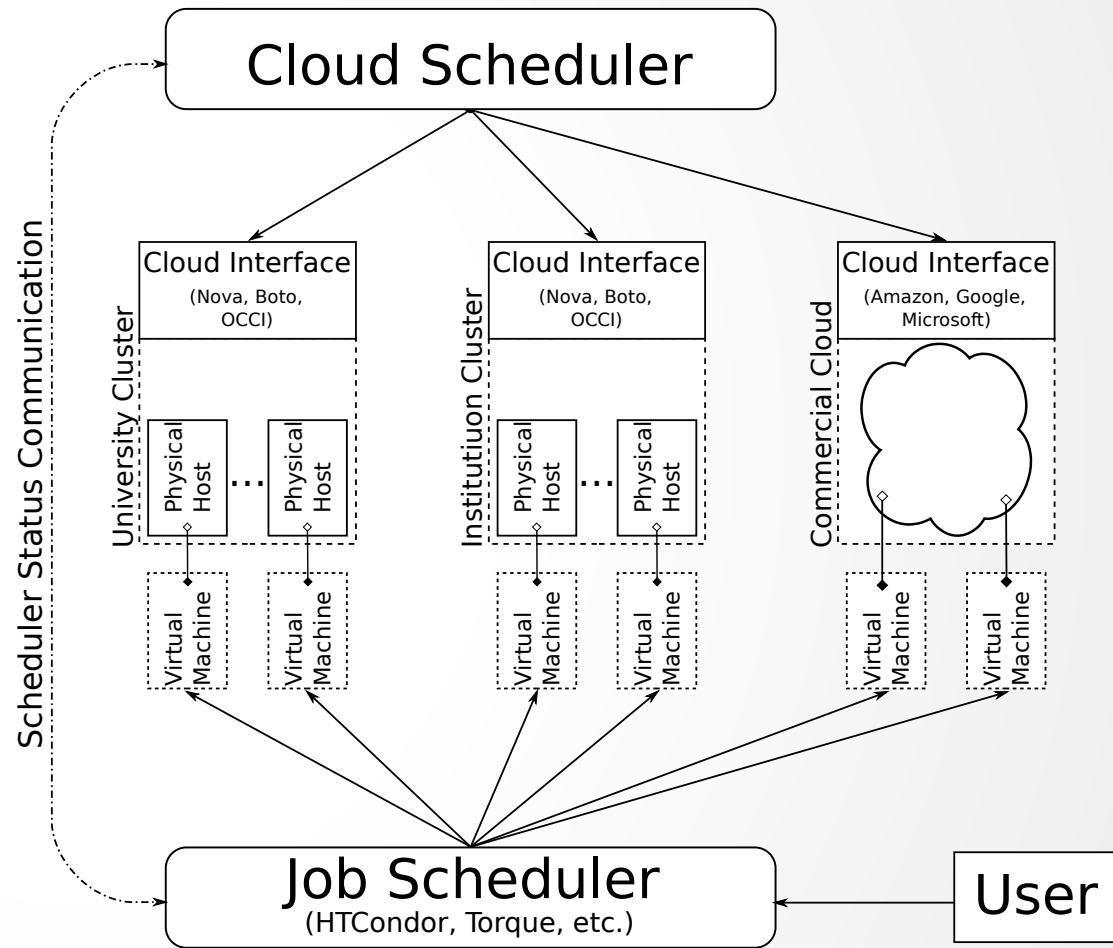
- Hard to maintain:
  - Many undocumented scripts
  - Scripts spread over many servers in P1 and in GPN
- No error handling for running instances
- Hard to update or modify

# Proposal for Sim@P1



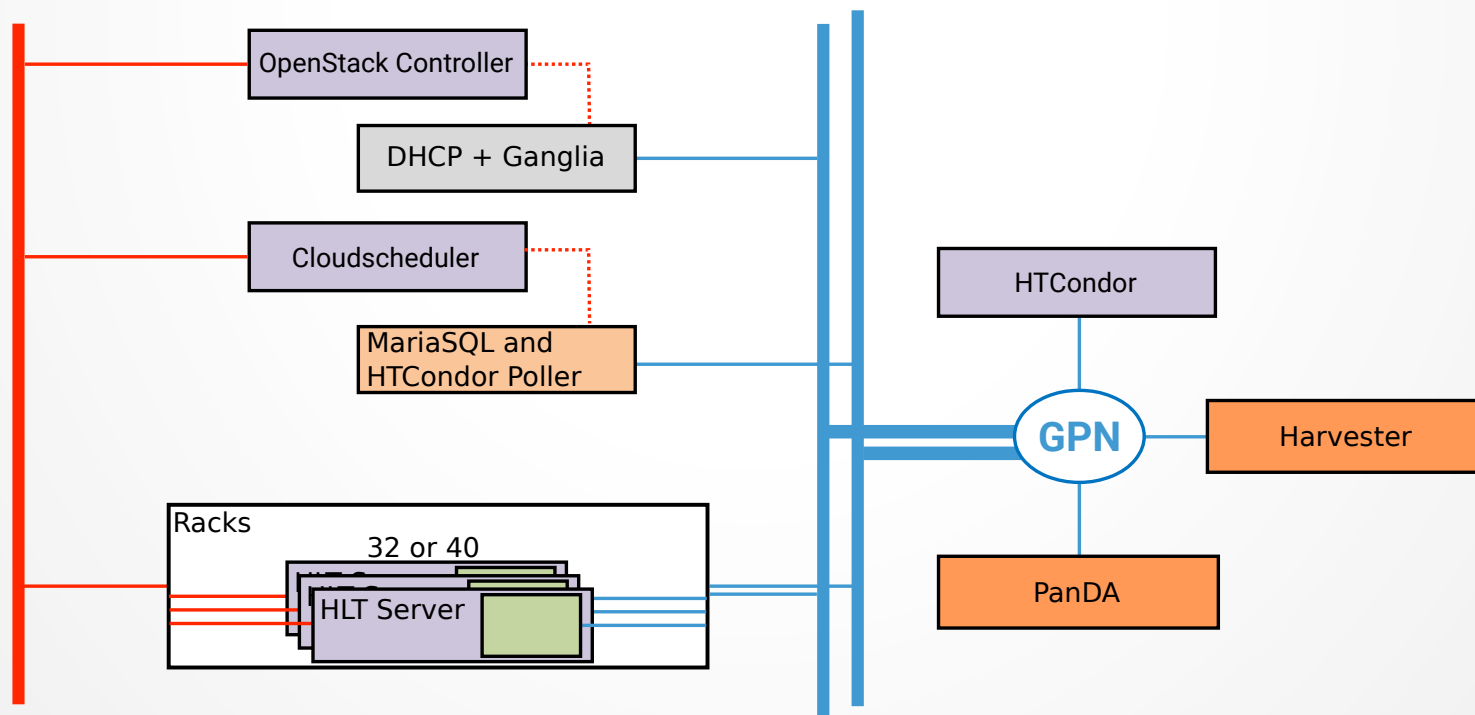
# Cloudscheduler

- Batch system on distributed cloud infrastructure
- In production for offline processing for
  - ATLAS (2012 - present)



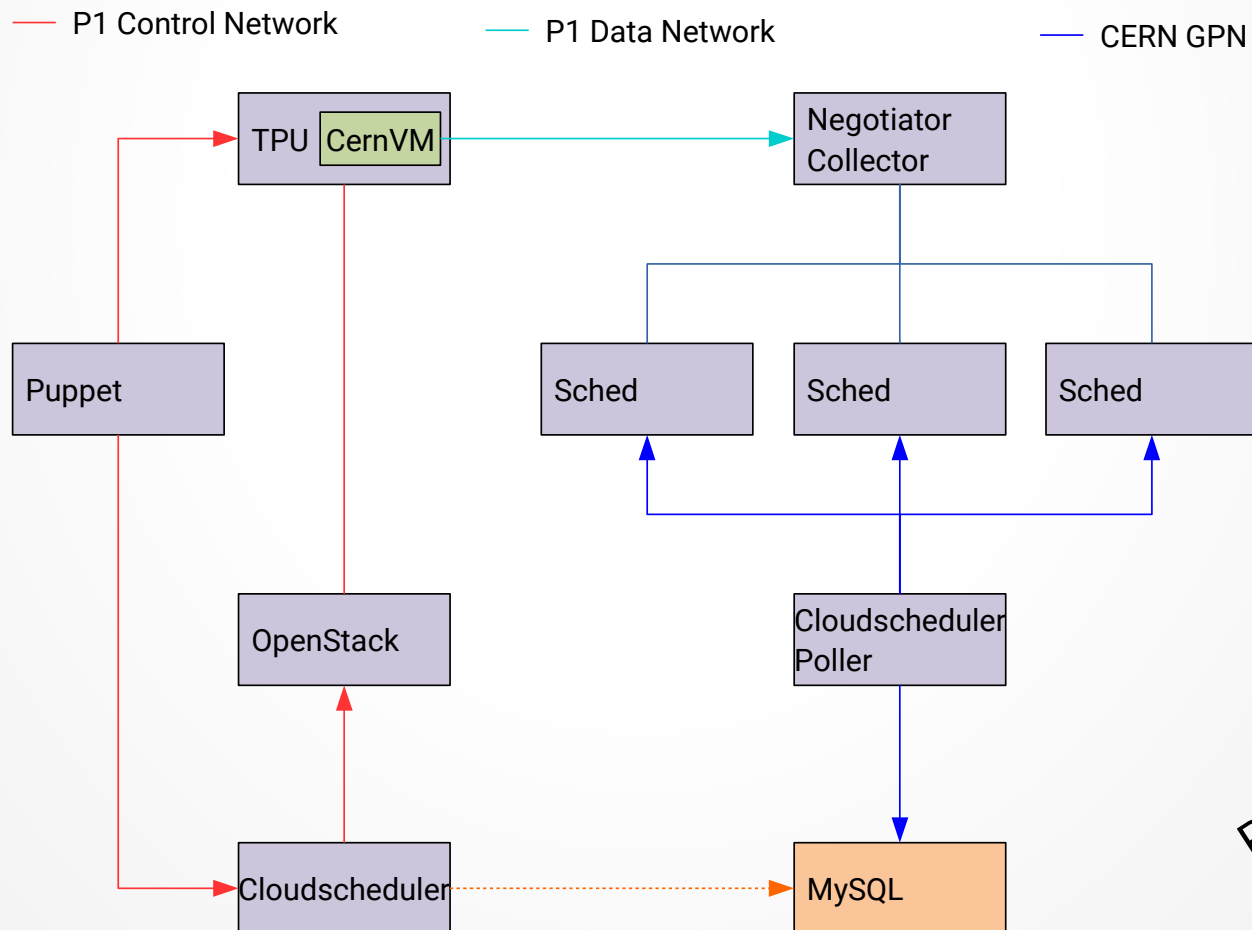
# Cloudscheduler at Point 1

- Proposal for long shutdown two [LS2]:
  - Cloudscheduler & OpenStack run in P1 Network
  - Polling thread and HTCondor run in CERN GPN
  - Cloudscheduler and polling thread interact with database



# Cloudscheduler at Point 1

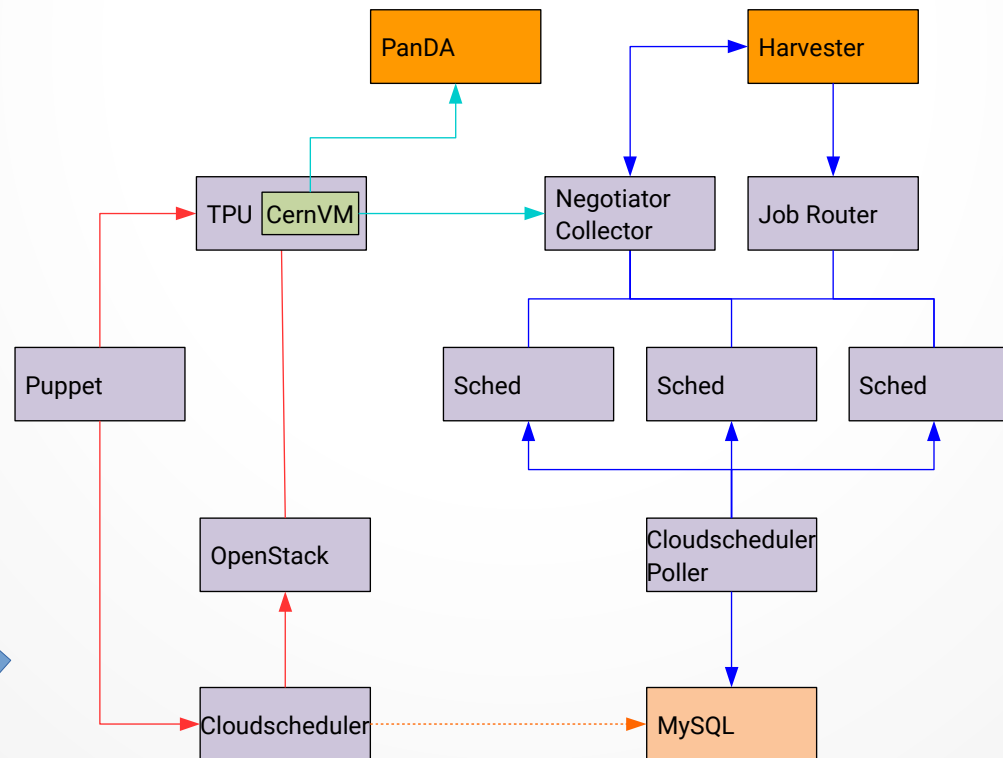
- Communication flow for Cloudscheduler
- Requires channel to database between P1 & GPN



Proposal  
Under Evaluation

# Harvester Job Submission

- Harvester pull mechanism allows job-specific resource request
- Condor reports resources availability to Harvester to improve PanDA job brokering
- Condor job router balance jobs across multiple schedulers



Harvester : an edge service harvesting heterogeneous resources for ATLAS  
T Maenow, Track 3 @ 11:15 on Thursday

Proposal Under Evaluation

# Summary

- Sim@P1 is successfully operating
- Cloudscheduler setup to ease operation under evaluation
- PanDA Harvester setup for job more flexible job submission

# Thanks to many contributors

## Cloudscheduler Team

K Casteels, C Driemel,  
M Ebert, C Leavett-Brown,  
M Paterson, R Seuster,  
R Sobie, R P Taylor, T Weiss-  
Gibbons

## Sim@P1 Team

A Di Girolamo, C Lee, P Love,  
J Schovancova, R Walker

## TDAQ Team

F Brasolin, D A Scannicchio,  
M E Pozo Astigarraga

# P1 Network Upgrades

- Tentative future

