Thoughts on running ATLAS Data Analysis in Containers on GCP

R. Mashinistov (UTA) and S. Panitkin (BNL)

UNIVERSITY OF TEXAS ARLINGTON

BROOKHAVEN NATIONAL LABORATORY

Outline

- Introduction
- ATLAS user analysis
- ATLAS Containers
- ATLAS containerized analysis on GCP

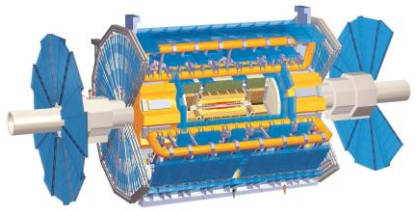
Introduction

- This is just a set of slides for ongoing discussion with Google about possible common projects

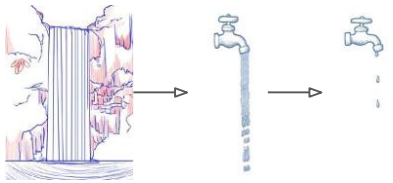
ATLAS User Analysis

- ATLAS user analysis is typically the final step in the data processing chain that leads to physics discovery, publication, etc.
- But it can be any user activity that is need for ATLAS – i.e. private simulations, calibrations, ML related activities (model training, etc), etc
- Often referred to as “chaotic” activity. Nevertheless all analyses are operated under physics working group auspices, to promote coordination, information exchange, quality control, software and knowledge sharing as well as proper resource usage
- Many people are involved in data analysis on the Grid: ~ 700 users per month
- Analysis process is typically iterative and I/O intensive (mostly input intensive)
- Fast job turn around is very important

The data processing chain



2 level, online system (HW+SW)



Reduce event rate from 40 MHz (60TB/s) to 1kHz (1.6GB/s) based on signatures
Event size ~1.6MB

Trigger

Raw data (RAW)

Reconstruction

Analysis Object Data (AOD)

Derivation

Organized production

$$S = \mathrm{i} \int \mathrm{d}^4x \, \mathcal{L}(x)$$

Generation

Event generator output (EVNT)

Simulation

Simulated detector output (RDO)

Reconstruction

Analysis Object Data (AOD)

Derivation

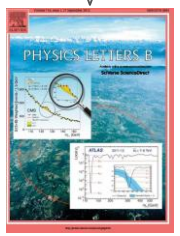
Derived AOD (DAOD)

Chaotic analysis

Detector data

Analysis

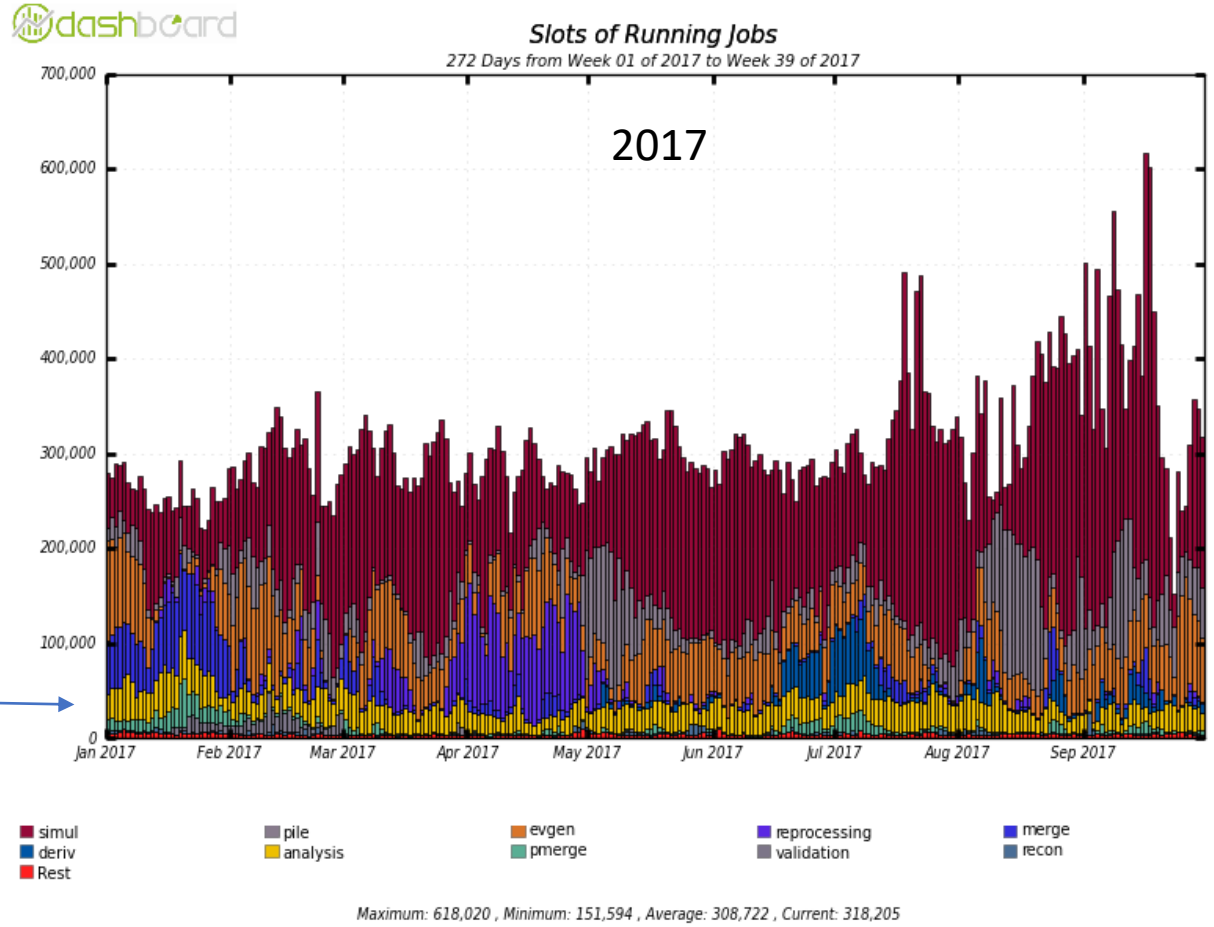
Simulated data



Online
Tier-0
Grid

From F. Barreiro Megino

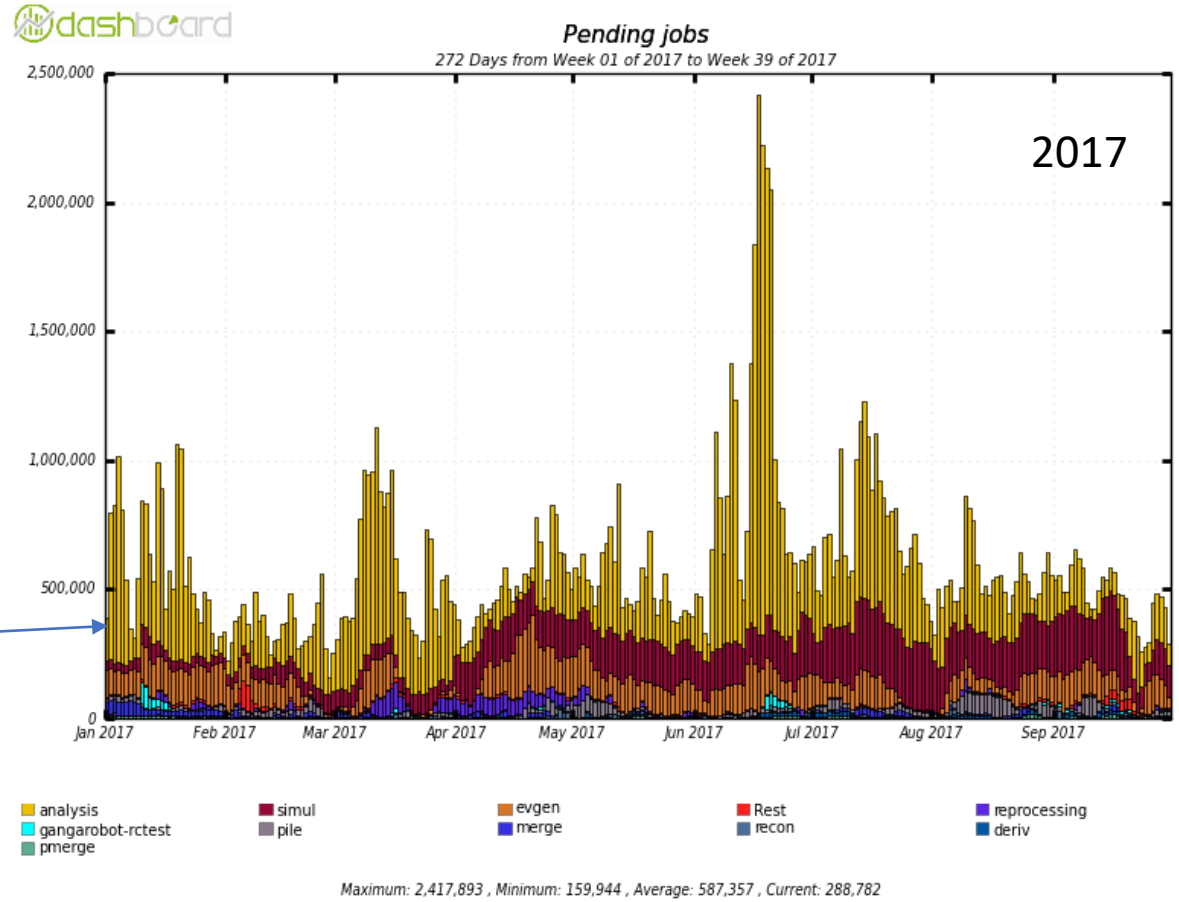
ATLAS User Analysis



analysis

~20% of ATLAS computing resources are dedicated to user analysis -> ~30k slots
~700 users per month on average , 1k+ per year
Analysis jobs on the Grid are managed by PanDA

ATLAS User Analysis



analysis

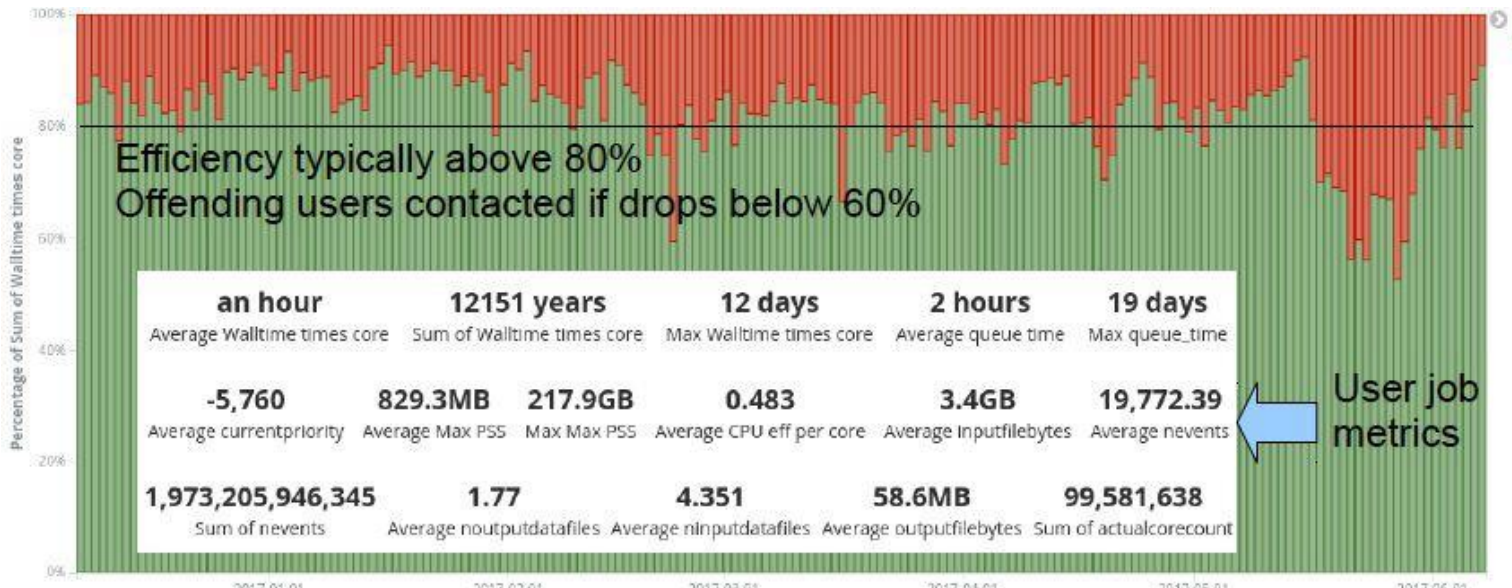
Constant demand for resources

Occasional large spikes in number of submitted jobs

Distributed Analysis: last 6 months

F. Legger

Percentage of Sum of Walltime times core

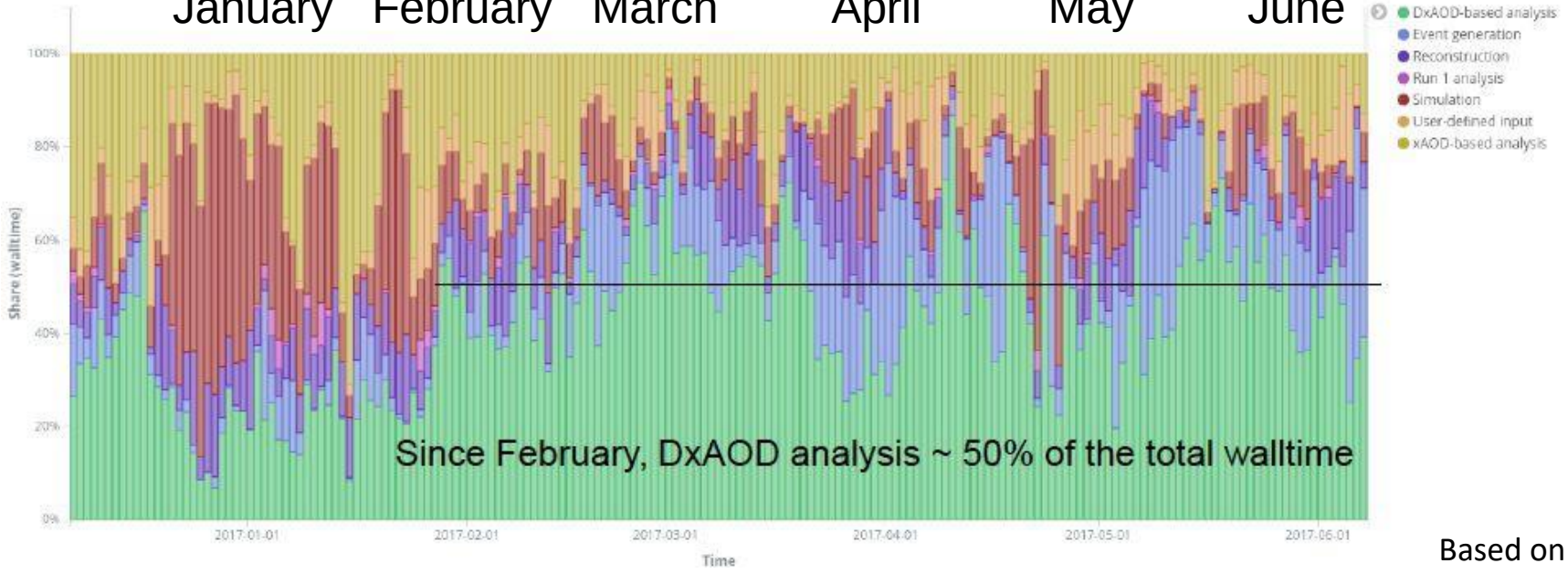


finished
failed

Many small, short, I/O intensive jobs
Data level parallelism can be exploited in many cases
Fast turn around is crucial

January February March April May June

Share (Walltime)



Based on N. Ozturk slide in Valencia 06.2017

Containers in ATLAS. I

- The most efficient “virtualization” technique
  - Native performance as compared to true virtualization
  - Effectively using a custom set of OS libs and software apart from sharing the kernel with host OS (similar to chroot)
- Providing isolation
  - PID namespace
  - Network configuration
  - UID/GID
  - Filesystem, device access
- Better suited for batch execution than VMs
- Two container systems systems are under tests in ATLAS
  - Docker and Singularity

From A. Filipcic at ATLAS TIM 09.2017

Containers in ATLAS. II

- Docker is better suited for full encapsulation
  - Full software and environment stack
  - Arbitrary workflow or service execution
  - Instances can be long lived
- Singularity:
  - OS encapsulation but use as much as possible from host OS
  - lower initialization latency

Docker and Singularity can use identical images, the usage of either is purely context dependent.

From A. Filipcic at ATLAS TIM 09.2017

Containers in ATLAS. III

- Deploy the ATLAS-controlled singularity, use more sites
  - Wrapper? Time scale?
- Find a common solution for grid and HPCs (and the rest)
  - Common procedure/recipe to create images (and unpacked images)
- Make the image creation part of software release process
  - Store the images in cvmfs, fat images in datasets…
- Documentation and installation instructions for deployment

From A. Filipcic at ATLAS TIM 09.2017

ATLAS Containers on GCP

- GCP supports containers, container orchestration and image management
  - Docker, rkt, Kubernetes, Container registry, etc
- In conjunction with the idea of ATLAS Data Lake on GCP running data analyses on GCE seems to be a natural
- GCE elasticity and scale seem to be well suited for ATLAS data analysis requirements:
  - Fast turn around, often spiky demand for resources, I/O intensity,…
- Possible connection with other GCP services like Cloud Machine Learning Engine
- ADC and ATLAS Physics Groups should be involved in selection of physics analyses suitable for running in containers on GCP and creation of containers for GCP
  - It should be representative cross-section of analysis activities
  - As well as several software frameworks used for analyses – Athena, Root, etc

Summary

- Container technology is well suited for running user workloads on heterogeneous infrastructure
- ATLAS started testing of containers usage on the Grid in
- GCP supports containers, container orchestration and image management
- Google Cloud elasticity and scale seem to be well suited for ATLAS data analysis requirements
- It would be interesting to try to run selected data analyses on GCP in containers