ATLAS Connect Usability from a Tier3 Center Perspective

Tier3 Users • UChicago ATLAS

Joint US ATLAS Software and Computing / Physics Support Meeting, LBNL
August 20, 2014

THE UNIVERSITY OF CHICAGO

Introduction

- UChicago Tier3 setup
- Analysis topics and use cases
- Wish list
- Future plans
- Summary and Conclusions

Current UChicago Tier3 Setup

- Our group has about 15 active users
- Two interactive analysis nodes (uct3-s1, uct3-s3)
- 19 worker nodes of various vintages

  - 4 newer (Dell R410), 15 legacy (very old, out of warranty)

- HTCondor is the scheduler, configured to optionally flock to the ATLAS Connect factories
- 3 storage servers (110 TB usable storage in xrootd); some Posix (70 TB) -- mostly legacy usage

Analysis topics

- SUSY analysis
  - All hadronic R-parity violating SUSY: Direct stop production and decay (4 jet final states) and gluino mediated neutralino production (high multiplicity final states)
- Exotic Higgs production
  - High mass resonance decay to HH (4 b final states)
- Phase I trigger upgrade R&D
  - **g**lobal **F**eature **EX**traction (gFEX) studies
- **F**ast**T**rac**K**er studies
  - Track reconstruction for L2 triggers

How we use ATLAS Connect

- Had SLC6 first (before UCT3) which helped for testing/compiling compatible software
- Submit directly from a UCT3 analysis node
- Or, login to the ATLAS Connect login node, submit from there
- Handles a number of processing tasks between the grid and interactive steps

How we use FAXbox

- Why?
  - Better interface - more like a normal filesystem. Having path, easy to access.
  - Used for histograms mostly? Some datasets
  - Easy to share files with collaborators from different institutions
- Medium-term storage (few months) and sharing of privately produced samples (mostly ntuples/D3PDs)
- Medium-term storage (few months) and sharing of small sub-samples of centrally produced datasets (ntuples/D3PDs) for local code development
- Long-term storage and sharing of specific samples (ntuples/D3PDs) for use by undergraduate research assistants
- Long-term storage and sharing of very specific, small, datasets for detector group analysis
  - Monitoring files from ATHENA jobs (~1-10 MB)
  - Legacy ESD files from cosmics runs (~1-10 GB)

Use Case 1: Exotics

- Store MC ntuples and skimmed data ntuples on LOCALGROUPDISK (~15TB)
  - Skimming done on GRID
  - Some datasets also on T3 XrootD as there was a point where LOCALGROUPDISK was having issues
- Develop locally on T3 interactive node
- Submit analysis jobs to Condor from T3
  - Takes a few of hours to run over (w/o systematics)

Use Case 2: gFEX

- Python analysis framework using privately produced D3PD (derived NTUP_COMMON) upgrade samples as input
    - Datasets are replaced relatively often (every 2-3 months at the moment)
- Login to Connect (login.usatlas.org)
- Run makefiles to produce config files based on lists of datasets
- Submit 10's - 100's of jobs to Connect that pull private D3PDs from FAXbox via root://faxbox.usatlas.org//user/…
    - Files are transferred to or accessed directly from Condor nodes
    - Jobs can have multiple, serial executions with multiple outputs
- Pickled files are then copied back to Connect (for now; no link from Tier3 to FAXbox)
    - 100's of output plots are generated locally/interactively in under 5 minutes and can be stored on faxbox
    - Direct access and further processing possible using output data saved on Connect/FAXbox
- **Private** repository of example code here: https://github.com/ATLASConnect/tutorial-PythonCondor

Use Case 3: Quark-Gluon Tagging

- Use ProofAna based code to run over NTUPles. Tested and submitted from UCT3.
- Submit skim code to the GRID to reduce event count
  - Also update calibrations + add extra variables
  - Skim is very loose (GRL + triggers only)
  - Events converted to ProofAna's native format
  - Output transferred to UC_LOCALGROUDISK using DaTRI (~10 Tb)
- Use Condor to make histograms
  - Use FAX to access skimmed files, allows access before files are transferred to LOCALGROUPDISK
  - Output histograms are transferred to FAXBOX for archiving (~200Mb for all samples)
  - DAGMan handles job failures and histogram merging
- Use desktop computer to further process the histograms
  - Code written using PyROOT
  - Histograms transferred to local disk using Globus Connect Personal from `connect#faxbox` endpoint
  - Faster response time during interactive work
  - Works offline
  - Access to tools (up-to-date python, diff GUI, pdf viewer…) not installed on UCT3

Use Case 4: H→ZZ→4l Analysis

- TSelector based analysis tested on FAX accessible samples using uct3-s1 and uct3-s3 (c++ only)
  - D3PD samples are located on LOCALGROUPDISK, flat ntuple output are put to my faxbox area
  - input MC + Data D3PDs ~ 7TB
    - These files were, on average, replaced about 3 times over 2 years
  - output ntuples are ~ 0.4 TB, including repetitive systematics (~10GB without sys)
- Analysis submitted to dagman owned condor jobs
  - submitted from uct3, not atlas connect
  - Dagman monitors jobs, retries failures, and combines output - it is essential!
    - A hadding script hadds a maximum of 10 files at a time from faxbox to /share/t3data3/, where these intermediate files are further hadded, transferred to faxbox, and deleted
    - this is a public folder so it can be accessed by nodes
  - ~15,000 jobs, each about 6 hours (for full systematics on all data, signal, and irreducible background)
    - Most analysis runs were done without systematics, roughly 50x fewer aggregate job-hours
- Plotting code done locally on s3 with png/eps/txt output to /share/t3data2/ (c++ heavy, very little python)
  - sometimes plotting done with condor on tier3 accessible nodes
- Next Dijet analysis will likely use a similar strategy

Use Case 5: All Hadronic SUSY Analysis

- **Storage**
  - Full privately produced datasets and subset of centrally produced datasets (mostly ntuple/D3PDs) stored on FAXbox (~2 TB)
  - Large centrally produced datasets stored on UC_LOCALGROUPDISK (~10 TB)
- **Analysis**
  - Code development on private laptop (faster for interactive work)
  - ProofAna (C++/ROOT) based analysis
  - Submit batch jobs to Condor from ATLAS Connect for small to medium sized datasets (up to a few TB)
    - Datasets stored on FAXbox accessed via: root://faxbox.usatlas.org://...
    - Using FAX to access datasets on UC_LOCALGROUPDISK
  - Submit to GRID for huge jobs (full datasets of several TB)
  - Output histograms automatically transferred and stored on FAXbox
  - Further analysis done on private laptop (scp files over from FAXbox)

Wish list

- Condor job monitoring - a graphical interface: where jobs are running, status
- Make it easier to switch between, and move data between Connect, uct3 and lxplus
- Better management of scientific python packages
- Data management of LOCALGROUPDISK
  - Probably have hundreds of TBs that are no longer used, but no easy way to know what is on there and how to delete

Future Plans

- Our goal is to offload as much of the traditional analysis "batch" processing to the grid
- Locally we would like to upgrade our interactive nodes, and perhaps add to local storage

Conclusions

- We've found ATLAS Connect as a useful service to extend Tier3-like processing tasks
- FAXbox is very convenient when sharing files with collaborators at other institutions

Contributors

- M. Wu (Exotics)
- D. Miller, G. Stark (gFEX Analysis)
- K. Krizka (QG Tagging)
- J. Dandoy (H→ZZ→4l Analysis)
- D. Miller, J. Olsson (SUSY Analysis)