

ATLAS at NAF, user perspective

Wolfgang Ehrenfeld[†] Ivor Fleck[‡] Oliver Rosenthal[‡]
• Hendrik Czirr[‡]

DESY[†]
University of Siegen[‡]

4th Annual Helmholtz Alliance meeting 2010, Dresden



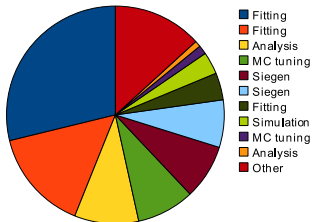
Bundesministerium
für Bildung
und Forschung



ATLAS usage of NAF CPU Time

CPU time / CPU days

NAF ATLAS



- Accounting interval: Aug.1st - Nov.29th 2010,
- Total CPU time: 21362 d ($\sim 12\%$ of NAF CPUs),
- Total wall clk time: 58228 d ($\sim 33\%$ of NAF CPUs)
- Nominal ATLAS share: $\sim 25\%$.

- ATLAS members actively working on the NAF batch system (within nominal share),
- Need to watch ratio between wall clk and CPU time,
- Power user activities: Analysis, Fitting, MC tuning and Simulation,
- 'Plain' data analysis (pre-made data/MC Ntuples) does usually not qualify as power user,
- User example: Top Physics Group Siegen.

Manpower:

1 Professor, 1 PostDoc, 3 PhD. Students and 3 Master Students.

Projects:

- Measurement of the top quark charge using $pp \rightarrow t\bar{t}\gamma$.
- Determination of the electric/magnetic form factors of the top photon coupling.
- Studies on color sextet scalar particles. Possible decay channel: $\phi_6 \rightarrow tt$.

Problem:

No official MC samples for e.g.: exotic top quark charge or color sextet particles.

- ⇒ Use NAF for private MC production and preparation of ROOT Ntuples for further analysis.
- ⇒ First look at data with regard to $t\bar{t}\gamma$, using the GRID from LXPLUS.

Standard User:

E.g. Myself

- Production of private Monte Carlo Samples,
 - Samples with ~ 60000 Full Simulation events, e.g. $pp \rightarrow t\bar{t}\bar{t}$ (Whizard2 + Pythia, .lhe files produced in Siegen),
 - SM background for $pp \rightarrow \phi_6\bar{\phi}_6 \rightarrow t\bar{t}\bar{t}$, no official samples available.
- Production of ROOT Ntuples from AODs,
 - By using private Dumpers or official tools, e.g. D2PD maker from Athena TopPhys Cache (ATLAS Top Working Group software package).
 - ⇒ No extensive disk space usage (~ 150 GB),
 - ⇒ Few different software releases in parallel.

Power User:

E.g. O. Rosenthal

- Large scale production of private Monte Carlo Samples,
 - Samples with ~ 180000 Full Simulation events, e.g. $pp \rightarrow t\bar{t}\gamma$ (Whizard+Pythia), for direct top charge measurement.
 - No official samples for $pp \rightarrow t\bar{t}\gamma$, yet,
 - Extensive MC studies, e.g. HERWIG+PHOTOS vs. Pythia,
 - Plans to migrate MC/Ntuple production to the GRID.
- Production of D3PDs in different Athena production caches.
 - ⇒ Extensive disk space usage,
 - ⇒ Many different software releases,
 - ⇒ Extensive use of the batch system (> 100 jobs in parallel) for long periods.

Production of private Monte Carlo Samples:

- Use NAF batch system, **good availability of computing power**,
- Sometimes difficult to define the best 'size' of a job,
⇒ More events per job:
computing time \uparrow , job priority \downarrow , number of jobs to be resubmitted in case of errors \downarrow .
- ⇒ Less events per job: vice versa.

ROOT Ntuple production:

- **Frequent changes to new software releases (due to bug fixes, etc.)**,
- ⇒ Setup of new releases is very easy.
- **Latest software releases available on NAF**,
- ⇒ Not only base releases: Working group related patch releases, e.g. Athena TopPhys cache.

Personal experience:

- + Easy to use,
- + Good availability of computing power,
- + Comfortable setup of new software releases,
- + Lots of disc space available:
 - Std. quota 500 GB of user scratch space, plus temporary disk space on Lustre and work group server scratch space on local machines,
- + Easy access and management of files (e.g. compared to CASTOR at LXPLUS),
- + Good user support,
- Home directory relatively small. Older software releases have to be deleted frequently.
- Very slow download speed from the NAF.