

Atlas Software Infrastructure

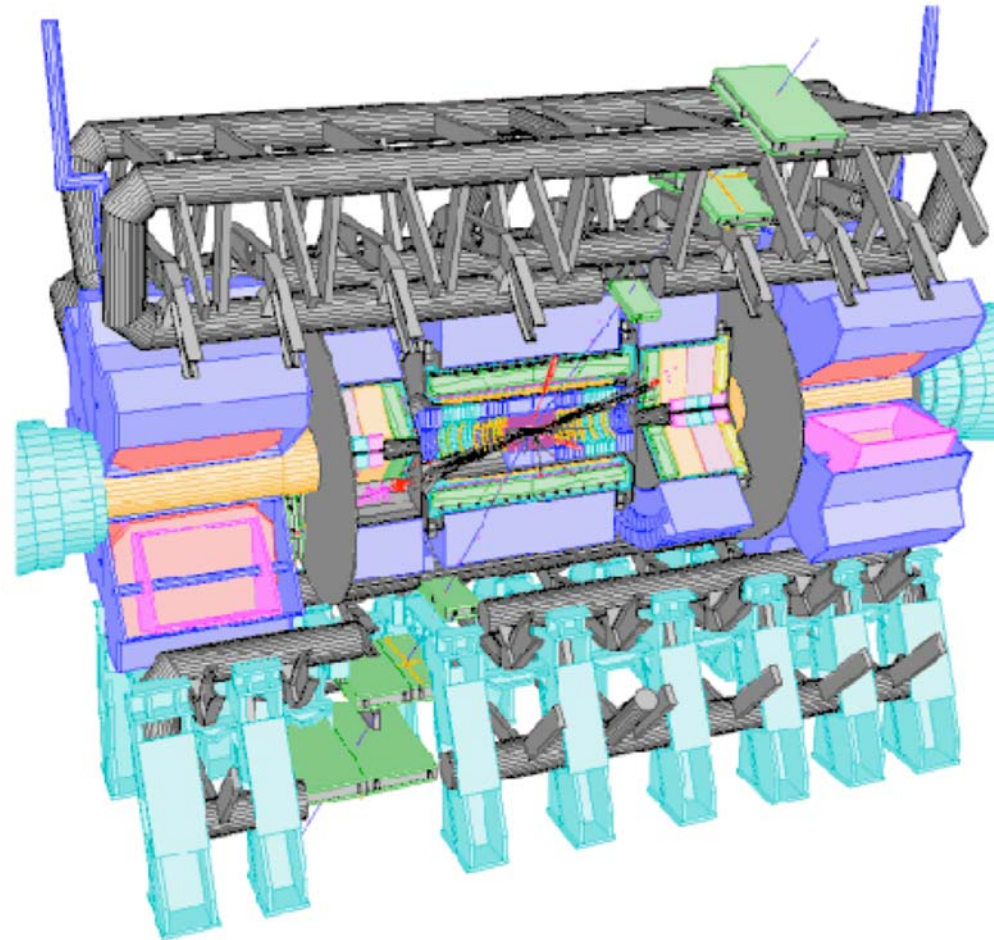
Frederick Luehring

Indiana University

MWT2 Kick-Off Meeting at IUPUI

Friday May 13, 2005

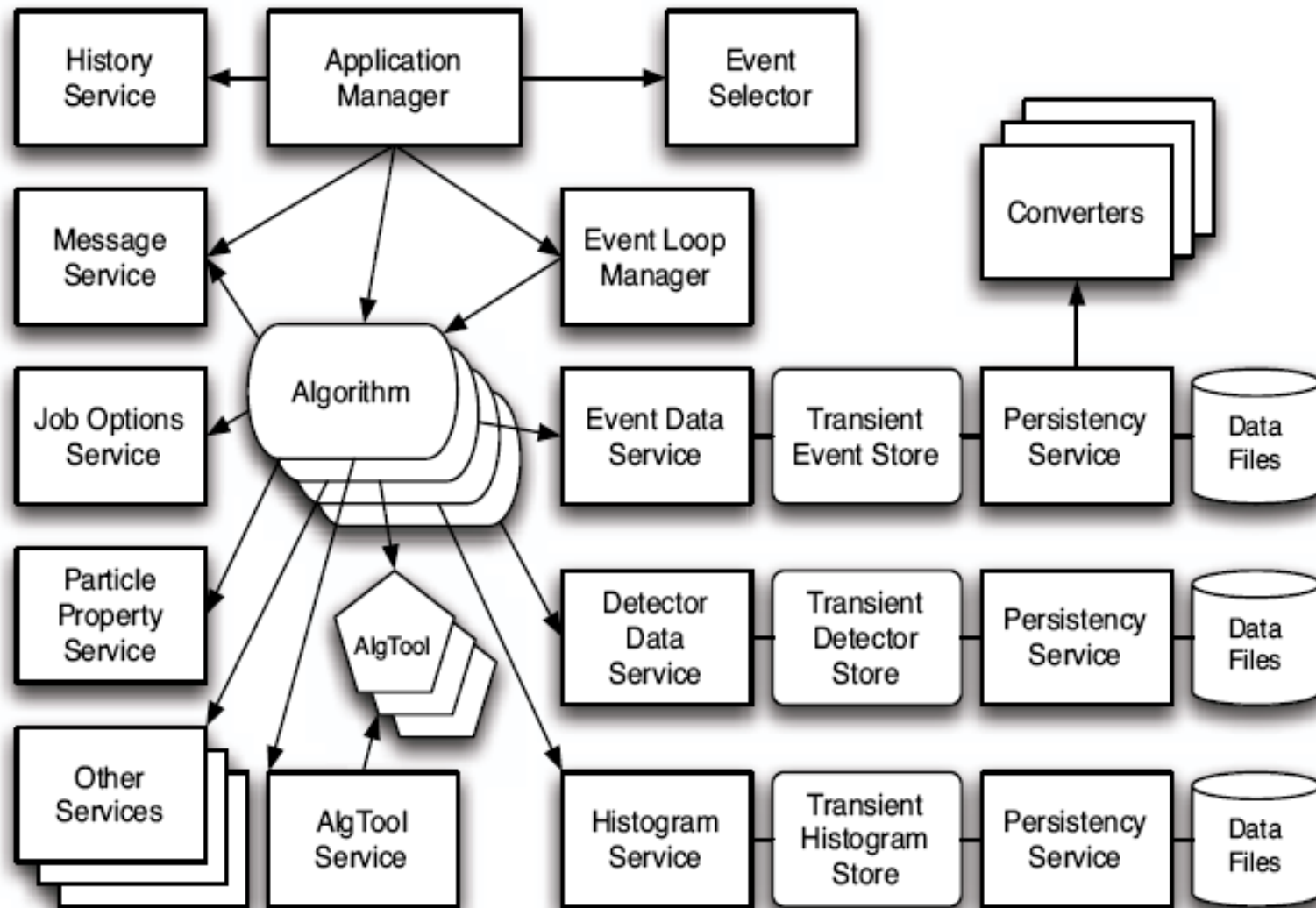
The Atlas Detector



Introduction

- The Atlas offline software known as Athena is:
 - Big and complex
 - Over 1000 packages (mainly C++)
 - Over a million lines of code
 - Install is ~5 GB (with external packages)
 - Takes about 24 hours to build on a single fast machine
 - Object-Oriented
 - Maintained by many people
 - ~500 code developers
 - Infrastructure team has ~20 people (mostly volunteers).
 - Only two librarians
 - Comprehensive:
 - Generation
 - Simulation
 - Reconstruction
 - Analysis

Athena SW Architecture



Data and Database

- Atlas will generate ~1 Petabyte of event data / year.
 - This does not include other data about the detector conditions (e.g. voltages on detectors, positions of detection elements).
 - This also does not include simulated event (Monte Carlo) data.
- Managing all of the data is a massive job and multiple databases will be used:
 - ROOT (home grown physics product)
 - MySQL
 - Oracle
- Dealing with all this data is the single most difficult task faced by the Atlas Computing project.

Atlas Offline Software Installation

- The Atlas software is installed from a “kit” that is constructed from Pacman and tar balls.
 - Instructions:
<https://uimon.cern.ch/twiki/bin/view/Atlas/InstallingAtlasSoftware>
 - Pacman command:

```
% pacman -get http://atlas.web.cern.ch/Atlas/GROUPS/  
SOFTWARE/OO/pacman/slc3/cache:10.0.0/AtlasRelease-  
slc3_ia32-opt
```
- Currently the supported platforms are RH 7.3 and SLC3 (essentially RHEL 3.0) running on IA32.
- Work is underway to support Opteron first in the 32 bit mode and later (end of the year?) in 64 bit mode.

Software Installation Validation

- Once the software is installed the correctness of the installation can be validated using a tool called KitValidation.
- KitValidation uses simulated data of typical Atlas Physics events to test that the software is installed correctly on a site.
- There are about 20 tests in the KitValidation suite and they test most parts of the software from the basics like whether data can be read and written to persistent storage up to whether a simulated Higgs boson is identified correctly.
 - It is essential to have such a tool when the software will be installed on hundreds of sites all over the world.

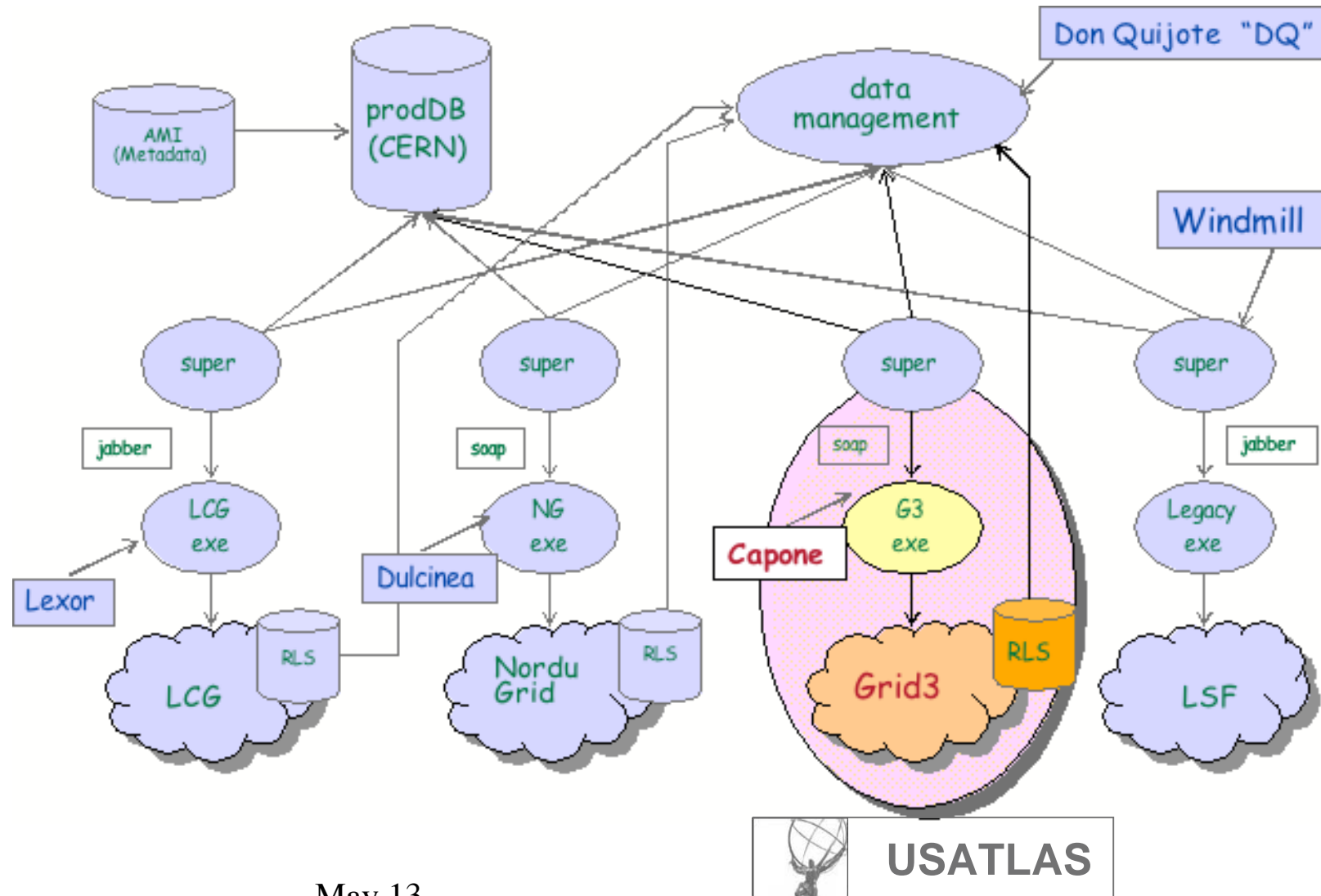
Source Code and Release Areas

- Atlas has four types of releases:
 - Nightly releases (built every night) but normally built incrementally so that they can be ready before morning.
 - Development releases (built every 3-4 weeks) and numbered X.1.0, X.2.0,...
 - Production releases (built ever 6 months) and numbered X.0.0
 - Bug fix production releases (built as needed) and numbered X.0.1, X.0.2,...
- All releases are available on AFS. The developer and production releases are found at:
`/afs/cern.ch/atlas/software/dist/X.Y.Z`
- The ATLAS CVS repository is also on AFS but is only accessible by kserver and ssh access.

Software Version Control

- A web driven tool called Tag Collector 2 is used to select which CVS version of each software package goes into a release.
 - Each developer can select for his or her package what version is used.
 - Related software packages are grouped into containers and the container manager can override the chosen version.
 - A software coordinator sits on top of the whole system and is responsible for coordinating between the developers when things don't build properly.

ATLAS Production System

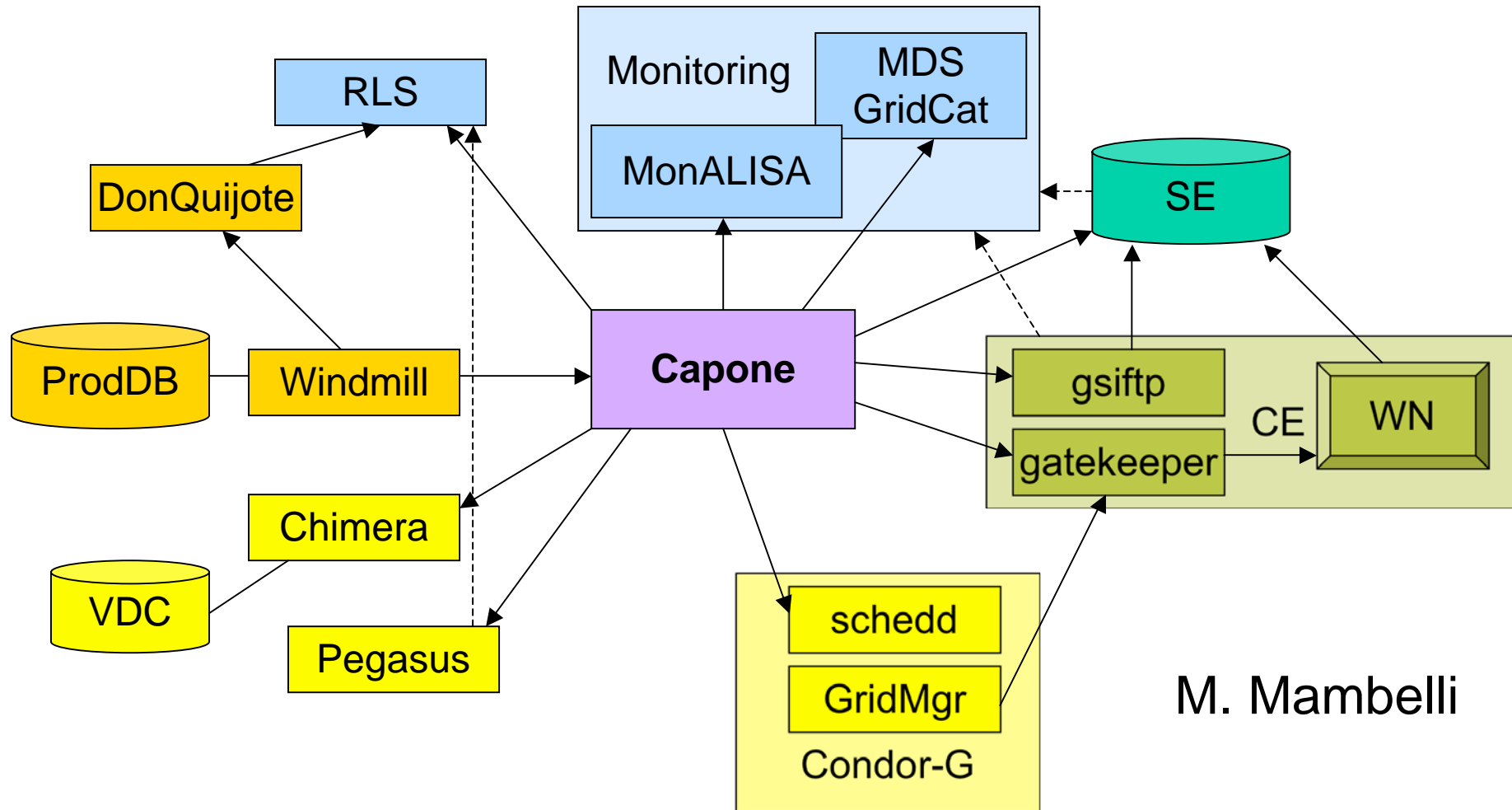


May 13,
2005

Atlas SW Infrastructure

10

Capone Executor for Grid3



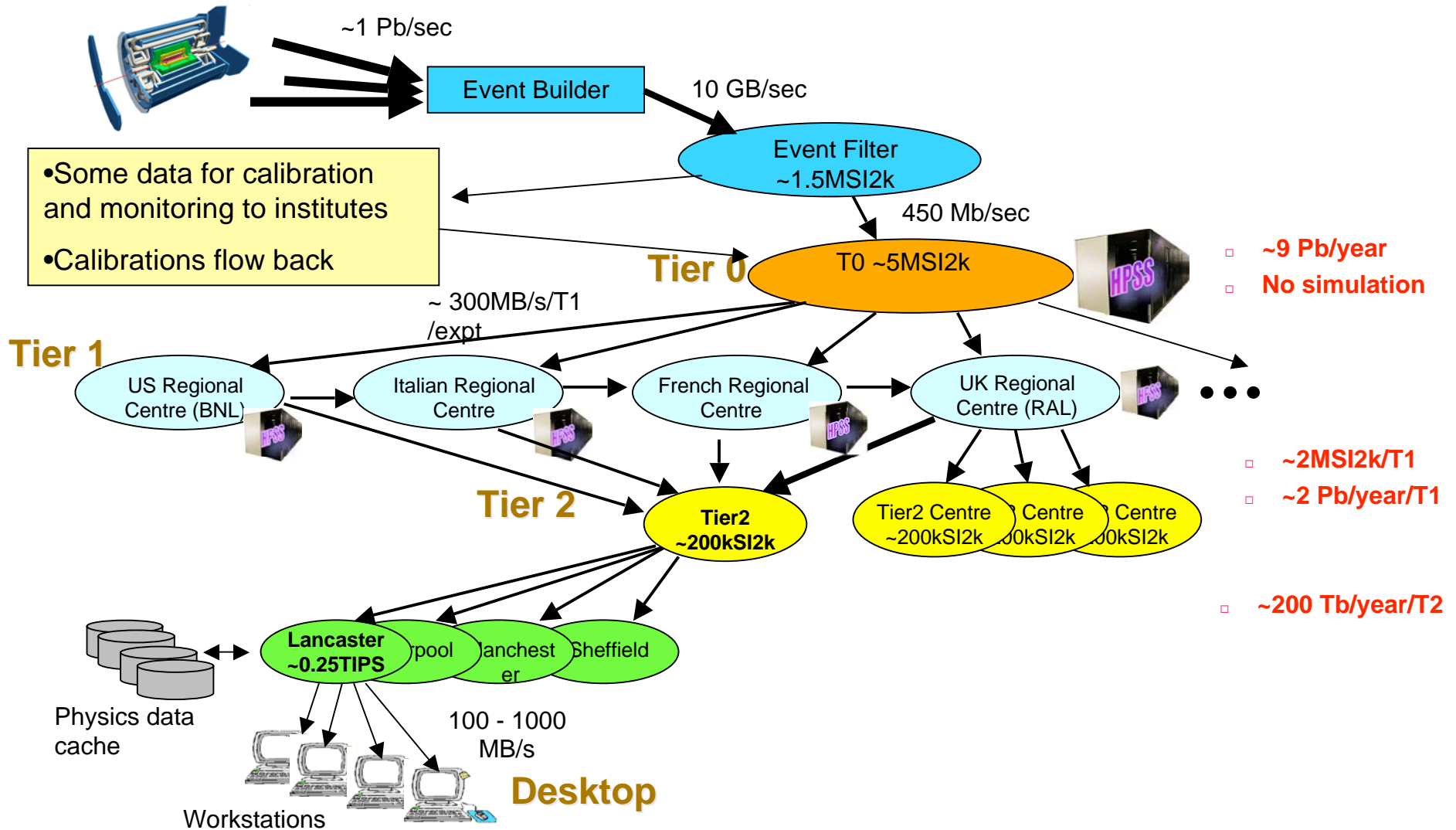
M. Mambelli

May 13,
2005

Atlas SW Infrastructure

11

Background: LHC Computing Model



May 13,
2005

Final Thoughts

- The Atlas detector is one of the most complex scientific instruments ever built.
- The Atlas detector will produce a huge amount of data that is recorded automatically.
- The offline software and hardware must be able to:
 - Keep up with rate of data production
 - Keep track of all the data without losing any of it
 - To find the interesting events in the mountain of data
 - To respond quickly when interesting events are found
- The MWT2 must be able to work in this environment.