High-Energy Physics, ATLAS and Trans-Pacific Collaboration Opportunities

Shawn McKee - University of Michigan

Driving eResearch Collaboration Across the Pacific Workshop

Perth, Australia, October 11th, 2007

ATLAS Detector Under construction
November 2005

Outline

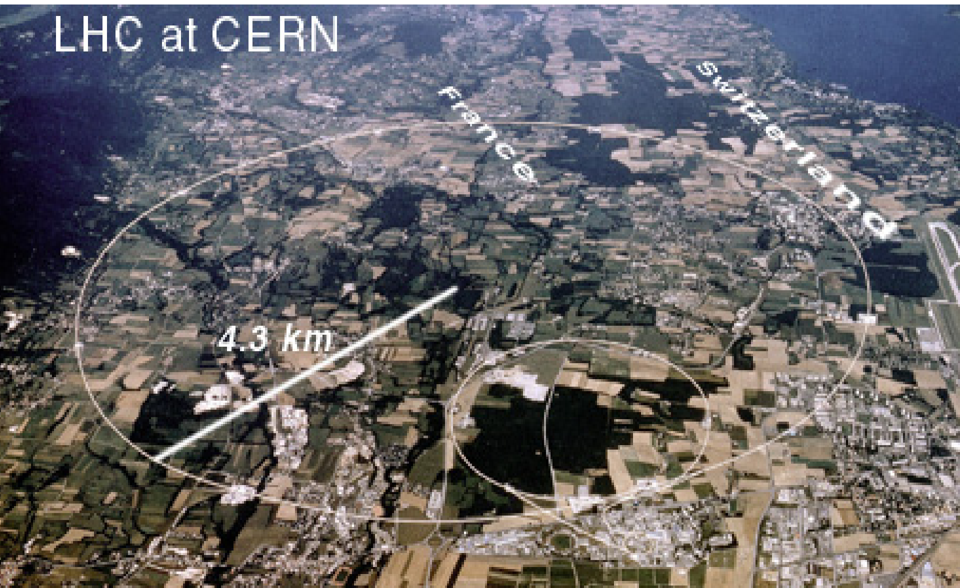
- High-Energy Physics, LHC and ATLAS
  - What are they?
  - How is ATLAS planning to do its scientific work?
- Important Issues for ATLAS (and HEP)
  - Large, distributed collaborations
  - Data intensive science
  - Significant needs for computers, storage and networks
- Opportunities for trans-Pacific collaboration…

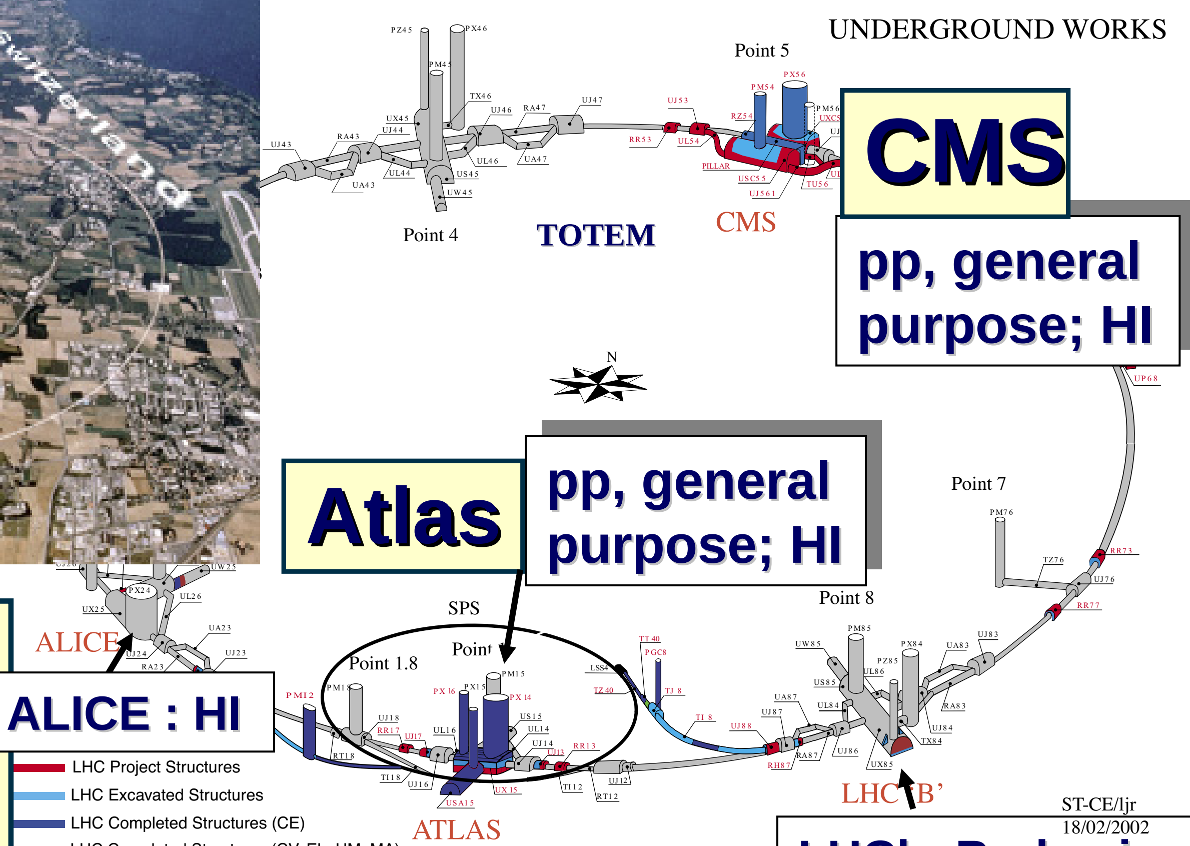
Introduction to High-Energy Physics

- High Energy physics (HEP) explores the forces and very small constituents of nature by colliding "high energy" particles and reconstructing the zoo of particles which result.
- These particle collision's can shed light on the nature and fundamental structure of matter.
- For high-energy physicists the most important scientific instrument is the LHC (Large Hadron Collider) at CERN in Geneva, Switzerland.
- One of the two primary colliding beam experiments at LHC is ATLAS which is trying to address fundamental questions in HEP.

The Large Hadron Collider (LHC)
CERN, Geneva: 2008 Start

27 km Tunnel in Switzerland & France





CMS
pp, general purpose; HI

Atlas
pp, general purpose; HI

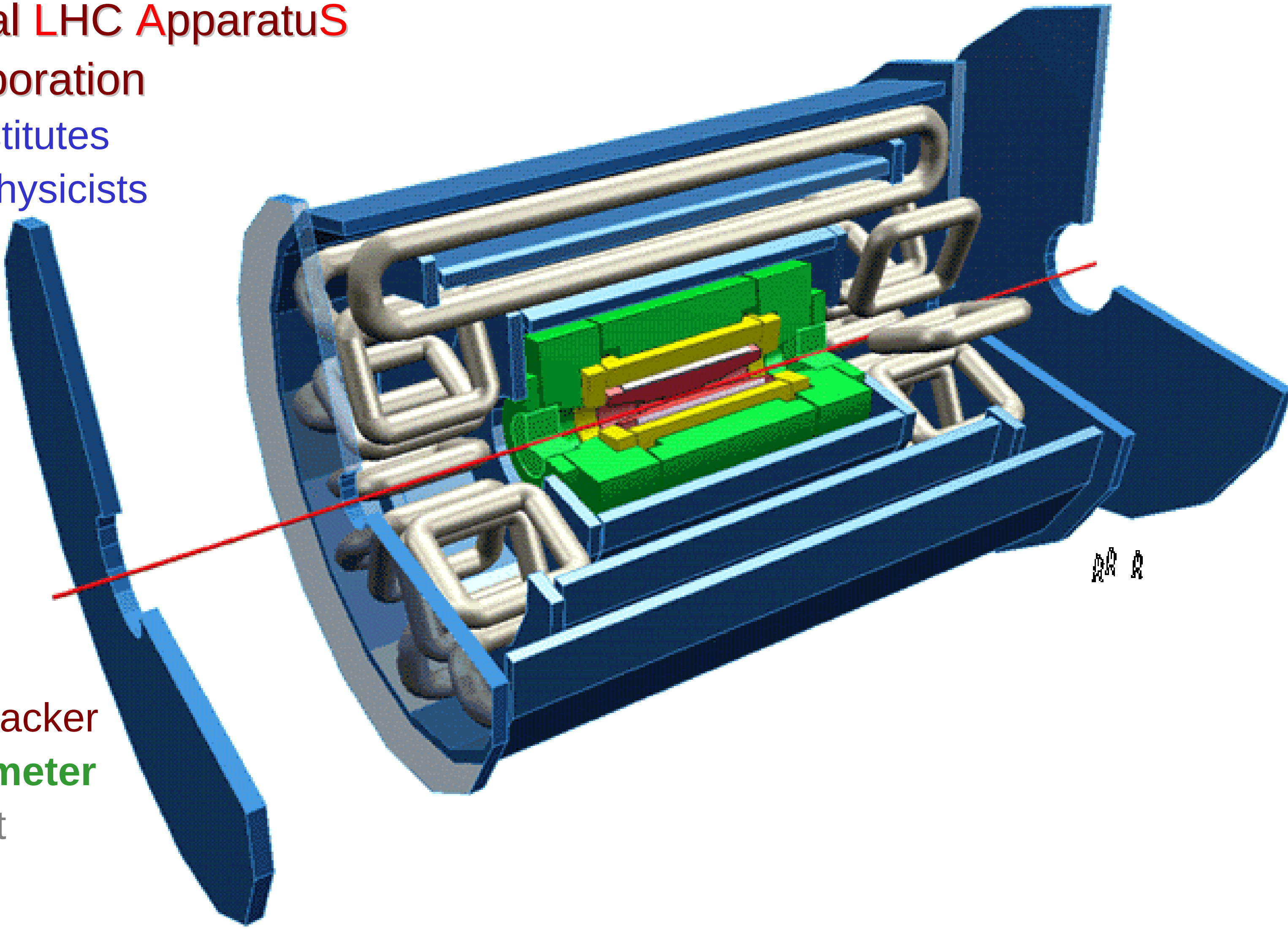
ALICE : HI

LHCb: B-physics

First Beams: Spring 2008
Physics Runs: from Summer 2008

ATLAS

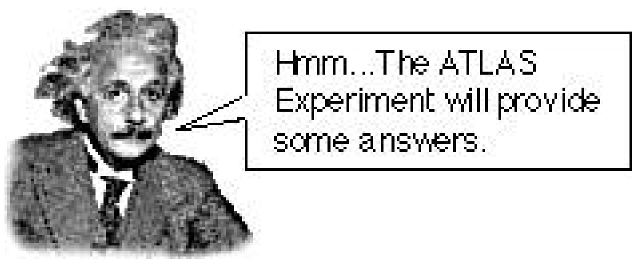
- A Torroidal LHC ApparatuS
- The collaboration
  - 150 institutes
  - 1850 physicists

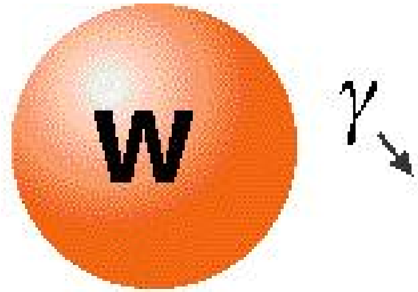


- Detector
  - Inner tracker
  - **Calorimeter**
  - Magnet
  - Muon

Physics with ATLAS: The Higgs Particle

- **The Riddle of Mass**





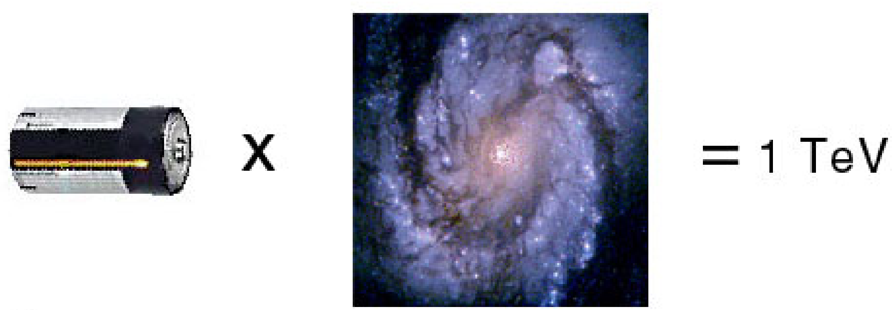
Why do some particles have large masses while the photon and neutrinos have puny masses?

- One of the main goals of the ATLAS program is to discover and study the Higgs particle. The Higgs particle is of critical importance in particle theories and is directly related to the concept of particle mass and therefore to all masses.

- ATLAS will explore many other interesting physics topics: the origins of dark matter, the nature of space-time, and extensions to the “Standard Model” (including Supersymmetry and extra dimensions)

High-Energy: From an Electron-Volt to Trillions of Electron-Volts

- Energies are often expressed in units of "*electron-volts*". An electron-volt (eV) is the energy acquired by a electron (or any particle with the same charge) when it is accelerated by a potential difference of 1 volt.
- Typical energies involved in **atomic processes** (processes such as chemical reactions or the emission of light) are of order a **few eV**. That is why batteries typically produce about 1 volt, and have to be connected in series to get much larger potentials.
- Energies in **nuclear processes** (like nuclear fission or radioactive decay) are typically of order one million electron-volts (**1 MeV**).
- The highest energy accelerator now operating (at Fermilab) accelerates protons to 1 million million electron volts (1 TeV $=10^{12}$ eV).
- The **Large Hadron Collider (LHC)** at CERN will accelerate each of two counter-rotating beams of protons to 7 TeV per proton.

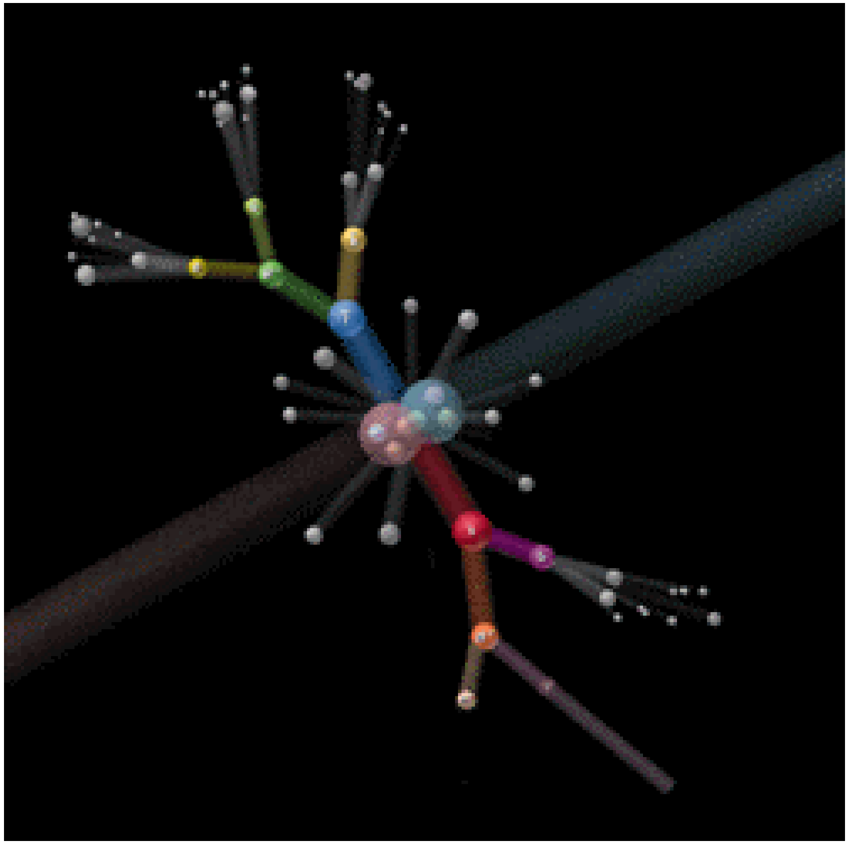


10^{12} eV is like having 1 battery for every star in our galaxy.

What is an Event in ATLAS?

ATLAS will measure the collisions of 7 TeV protons.

Each time protons collide or single particles decay is called an "event"



In the ATLAS detector there will be about a billion collision events per second, a data rate equivalent to twenty simultaneous telephone conversations by every person on the earth.

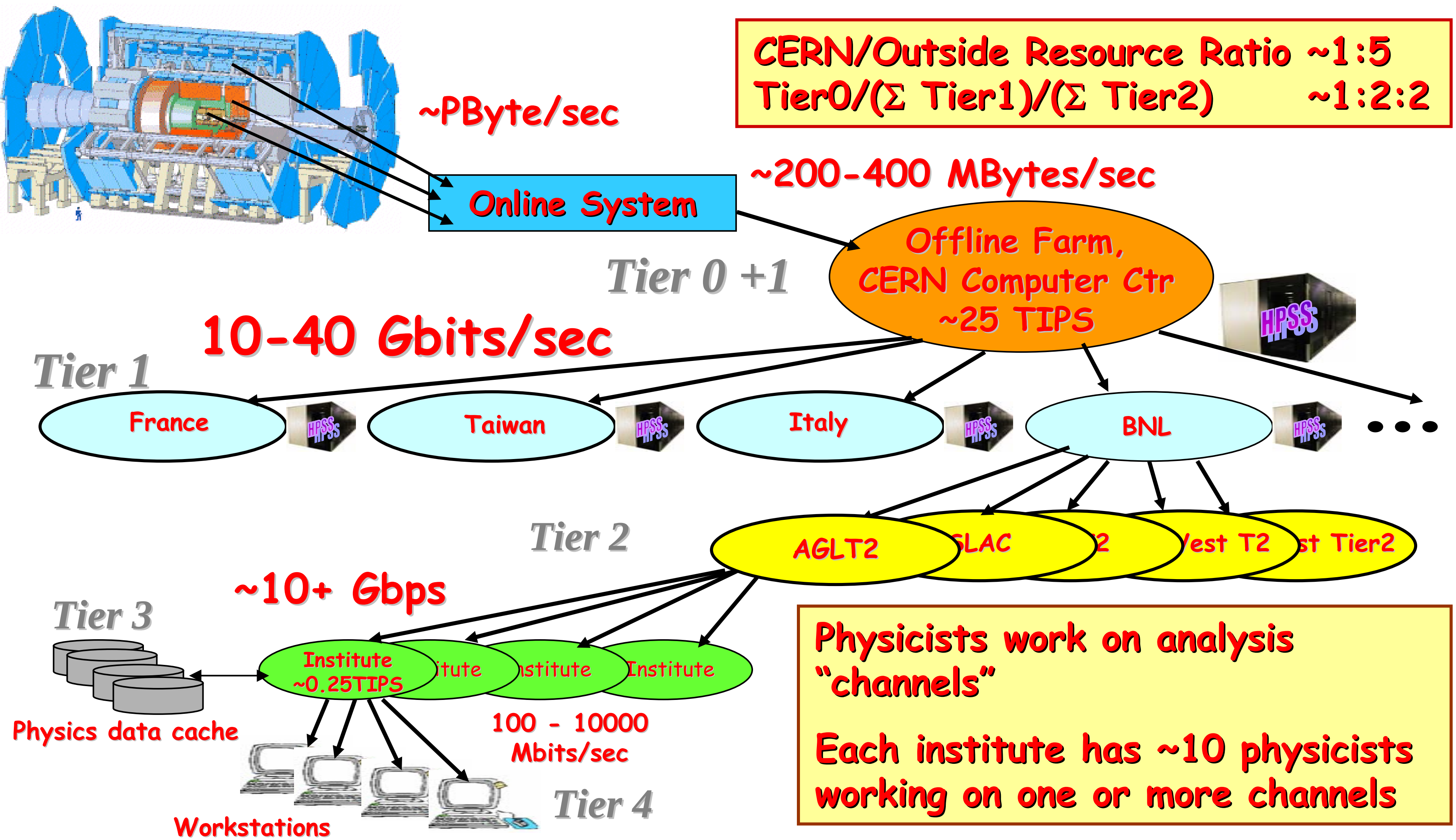
Needles in Haystacks

- When protons collide, some events are "interesting" and may tell us about exciting new particles or forces, whereas many others are "ordinary" collisions (often called "background"). The ratio of their relative rates is about **1 interesting event for 10 million background events**. One of our key needs is to separate the interesting events from the ordinary ones:
  - This will require significant computing power
  - This will require significant storage for the data and the associate network to move it all
- Furthermore the information must be sufficiently detailed and precise to allow eventual recognition of certain "events" that may only occur at the rate of one in one million-million collisions (10^{-12}), a very small fraction of the recorded events, which are a very small fraction of all events.

ATLAS Tiered Computing Model

- To address its computing needs, ATLAS has focused upon a Tiered computing model: Tier0 - Tier4
- Within this model the Tier-2s (regional centers) primarily handle Simulation & Analysis
- Implicit in this model and **central to its success** are:
  - High-performance, ubiquitous and robust **networks**
  - Grid middleware to securely **find**, **prioritize** and **manage** resources
- **Without either of these capabilities the model risks melting down or failing to deliver the required capabilities.**
- Efforts to date have (*necessarily*) focused on building the most basic capabilities and demonstrating they can work.

ATLAS's Hierarchical Computing Model



ATLAS version from Harvey Newman's original

ATLAS Computing Model Overview

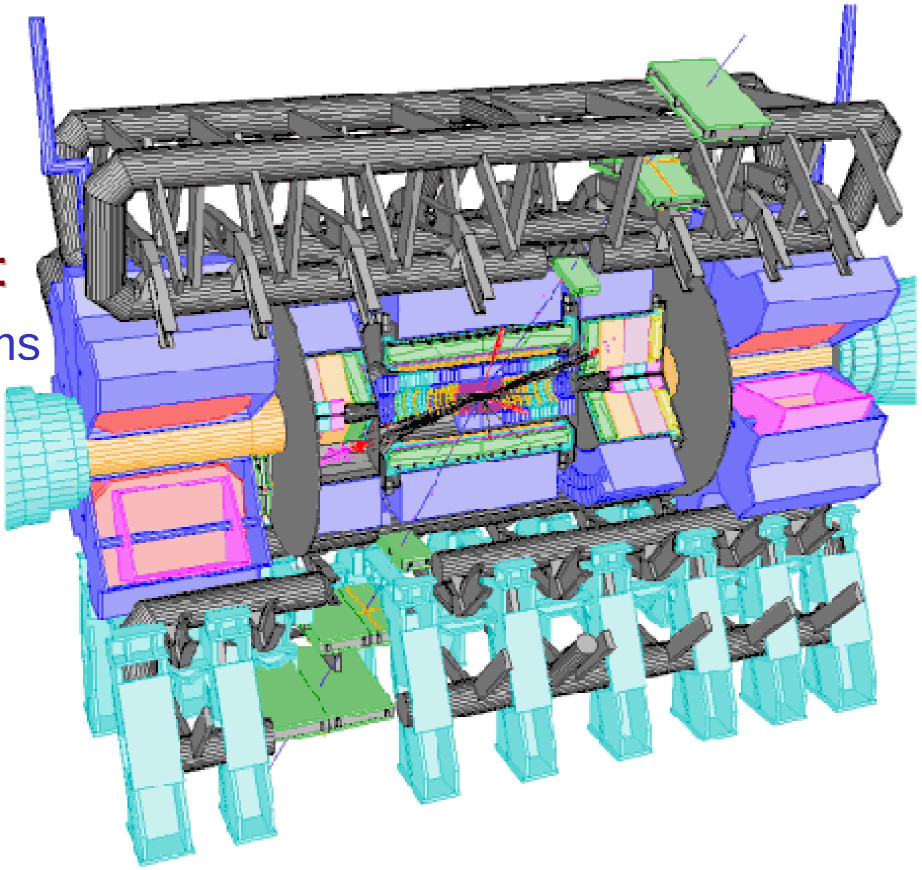
- The ATLAS hierarchical model (EF-T0-T1-T2) has specific roles and responsibilities
  - Data will be processed in stages: RAW->ESD->AOD-TAG
  - Data "production" is well-defined and scheduled
  - Roles and responsibilities are assigned within the hierarchy.
- Users will **send jobs to the data** and extract relevant data
  - typically DPDs (derived physic's data) or similar summary level data
- **Goal** is a production and analysis system with <u>seamless</u> access to <u>all</u> ATLAS grid resources
- All resources need to be managed effectively to ensure ATLAS goals are met and resource providers policy's are enforced. **Grid** infrastructure/middleware must provide this


Armenia
Australia
Austria
Azerbaijan
Belarus
Brazil
Canada
China
Germany
Danmark
Spain
Finland
France
Georgia
Greece
Italy
Israel
Japan
Kazakhstan
Morocco.
Norway
Netherlands
Portugal
Poland
Romania
Russia
Sweden
Switzerland
Slovenia
Slovakia
Turkey
UK
USA
ATLAS
Collaboration
DEB 97

ATLAS Physicist's Requirements

- The ATLAS collaboration has only ~year before it must manage large amounts of "real" data for its globally distributed collaboration.
- ATLAS physicists need the software and physical infrastructure required to:
  - Calibrate and align detector subsystems to produce well understood data
  - Realistically simulate the ATLAS detector and its underlying physics
  - **Provide access to ATLAS data globally**
  - Define, manage, search and analyze data-sets of interest

ATLAS



Wide-Area Physics and the Network

- Physicists require the network for a number of tasks:
  - Moving or accessing data (raw or analyzed) (gridftp, nfs, afs)
  - Remote collaboration and videoconferencing
  - Control of remote instruments
  - Searching for, or accessing remote information
- The developments in grid computing are placing new requirements on the network
- The goal is to have the WAN enable WAP (Wide-Area Physics) - transparent high-performance access to remote resources required by physicists

Trans-Pacific Collaboration Opportunities

- Australia has an excellent national backbone and international connectivity via AARnet/Southern Cross
- However it also has significant round-trip times to most important network destinations for high-energy physics:
  - Its Tier-1 center is in Taiwan
  - Most other ATLAS Tier-2 centers are even further away
- This presents some unique opportunities for collaboration across the Pacific

Data Movement and Service Testing

- Data movement is critical for ATLAS (and most e-Science collaborations).
  - Tuning networks and applications to enable high-performance transfer across the Pacific would not only help AU-ATLAS but all ATLAS sites; what can be done for such a “remote” connection should be easy to enable everywhere else…
- The ATLAS production/analysis systems are based upon numerous middleware components, not all of which are fully tested under applicable operating conditions
  - Australia represents one of the largest network round-trip times for ATLAS middleware to deal with.   Sometimes unusual behaviors or service failures can result from this significant change in the network connection latency, especially database systems.
  - Thorough testing of ATLAS middleware between Australia and the US could help harden the resulting systems.

Collaborative Tools and Network Monitoring

- Collaborative tools are increasingly vital for participating in large collaborations like ATLAS.
  - Using collaborative tools between Australia and the US would help to harden and improve their capabilities
  - Video-conferencing (EVO), web-lecture archives and perhaps remote control rooms are all logical things to explore
- Because of the networks fundamental role in ATLAS it would also make sense to bring up network monitoring nodes at a few locations in Australia
  - Use of PerfSONAR, MonaLISA, NDT, IEPM-BW and others would be very helpful in tracking network capability and diagnosing problems

Summary

- Within a year real LHC data will begin flowing
  - Physicists in Australia (and everywhere!) will be intently working to access and process data.
  - There will be implications for networks, collaborative systems, middleware, storage and computing resources.
- Trans-Pacific collaboration provides some unique opportunities to ensure our developing infrastructures will meet our scientific needs.
- Shortly HEP physicists will be living in "Interesting Times"…

Questions?

TCP WAN Performance

Mathis, *et. al*., Computer Communications Review v27, **3**, July 1997, demonstrated the dependence of bandwidth on network parameters:

$$BW \le \frac{0.7 \times MSS}{RTT\sqrt{PkLoss}}$$

BW - Bandwidth

MSS – Max. Segment Size

RTT – Round Trip Time

PkLoss – Packet loss rate

If you want to get 900 Mbps via TCP/IP on a WAN link from LBL to UM you need a packet loss < 1.8e-10 !! (~70 ms RTT).

To achieve ~1 gigabit across the Pacific Ocean would require a factor of 16 less packet loss.