ATLAS Tiered Computing Model

- To address its computing needs, ATLAS has focused upon a Tiered computing model: **Tier0 - Tier4**.
- ATLAS has different access policies from CMS in that (at least initially) this hierarchy is enforced by restricting access via Tier-1 “clouds”. Effectively, the models are very similar.
- Implicit in this model and **central to its success** are:
  - High-performance, ubiquitous and robust **networks**
  - Grid middleware to help **find**, **prioritize** and **manage** resources
- **Without both of these capabilities the model risks melting down or failing to deliver the required capabilities.**

ATLAS Computing Model Overview

- The ATLAS hierarchical model (EF(Event Farm)-T0-T1-T2) has specific roles and responsibilities
  - Data will be processed in stages: RAW->ESD (**E**vent **S**ummary **D**ata)->AOD (**A**nalysis **O**bject **D**ata)->DPD(**D**erived **P**hysics **D**ata)
  - Data "production" is well-defined and scheduled
  - Roles and responsibilities are assigned within the hierarchy.
- Users **send jobs to the data** and extract relevant data
  - Output is typically DPDs or similar summary level data
  - DPDs are stored either in "user" space or as part of regular data management for officially produced DPDs
- **Goal** is a production and analysis system with seamless access to all 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

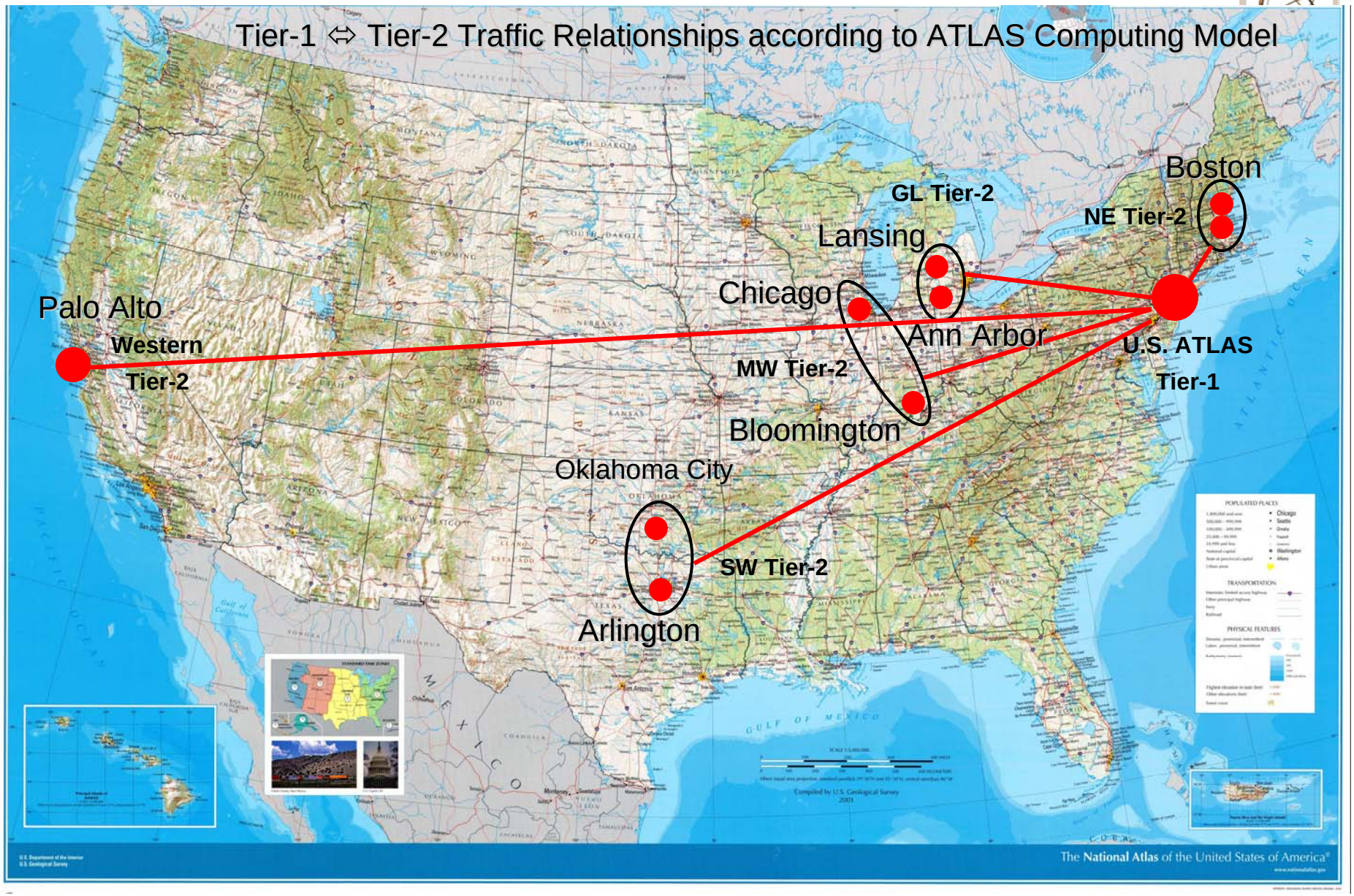
ATLAS Facilities and Roles

- Event Filter Farm at CERN
  - **Assembles data (at CERN) into a stream** to the Tier 0 Center
- Tier 0 Center at CERN
  - **Data archiving:** Raw data to mass storage at CERN and to Tier 1 centers
  - **Production:** Fast production of Event Summary Data (ESD) and Analysis Object Data (AOD)
  - **Distribution:** ESD, AOD to Tier 1 centers and mass storage at CERN
- Tier 1 Centers distributed worldwide (10 centers)
  - **Data steward: Re-reconstruction** of raw data they archive, producing new ESD, AOD
  - **Coordinated access** to full ESD and AOD (all AOD, 20-100% of ESD depending upon site)
- Tier 2 Centers distributed worldwide (approximately 30 centers)
  - **Monte Carlo Simulation**, producing ESD, AOD, ESD, AOD sent to Tier 1 centers
  - On demand user **physics analysis** of shared datasets
- Tier 3 Centers distributed worldwide
  - **Physics analysis**
- A CERN Analysis Facility
  - Analysis
  - Enhanced access to ESD and RAW/calibration data on demand

US ATLAS Infrastructure

- There is a Tier-1 center in the US at BNL and 5 Tier-2 (org.) centers
  - 9 different sites
- The Tier-1 center currently has two 10GE connections to the WAN
- All US Tier-2 centers either have (or plan to have) 10GE+
  - Great Lakes Tier-2 – 2 x 10GE
    - UM and MSU
  - Mid West Tier-2 – 10GE at UC
    - Univ. of Chicago and Indiana University
  - North East Tier-2 – 10GE at BU
    - Boston University and Harvard University
  - South West Tier-2 – 10GE at OU
    - Oklahoma University, Univ. of Texas - Arlington
  - Western Tier-2 – SLAC has implemented redundant 2 x 10GE
- Today US ATLAS fully relies on “best-effort” IP Services provided by ESnet and Internet2, and eventually Metropolitan Network Service Providers

Tier-1 ⇔ Tier-2 Traffic Relationships according to ATLAS Computing Model



Examining "Required" Network Capacity

- Production data flows are well understood and planned for within ATLAS.
- However, perhaps one of the most significant components of network usage has not been well tested yet: User Analysis (both Planned and "Chaotic")
  - We expect **significant** traffic related to 100's of physicists in the US undertaking individual and group analysis projects. What are the **requirements** to support this?
- Traffic is likely to be large at all scales in the hierarchy as large physics groups, down to individuals, pursue their scientific goals**.**
  - **An important point is that if physicists have access to end-systems and networks capable of high-bandwidth they will use them to the full extent possible**
- One important aspect for physicists is how quickly new ideas can be tested and "turned around" on our architecture. The burst capacity of the network and end-systems is central in determining user latency. **Our goal is to support such work on human timescales to expedite "time-to-discovery".**

Data Movement and Service Testing

- Data access is critical for ATLAS
  - Finding any I/O bottlenecks and removing them where feasible is an effective way to improve our infrastructure.
  - Tuning networks/applications to enable high-performance transfers within the US and across the Atlantic
- Efforts to date have (*necessarily*) focused on building the most fundamental capabilities and demonstrating they can work.
- Because of our successes in getting our software and hardware facilities operational we are now increasing our efforts on improving the overall throughput of our systems.

Estimating Useful Bandwidth Needs

- Physicists can obviously benefit from being able to quickly access datasets, including refreshing stale “caches” and getting the newest reconstructed data as quickly as possible
- We have single systems at BNL and some Tier-2’s which can source or sink data in excess of 400MB/sec (~$25K)
- We have lower performance systems in large numbers that also produce large data rates in aggregate.
- Given that Tier-2’s can afford such systems it makes sense to provision one or more 10GE connections. Three 400MB/sec systems could saturate a 10GE link.
  - Looking at predictable vs best-effort Tier-1 ⇔ Tier-2 connectivity
  - Must remain affordable

Some Ideas …

Boston
GL Tier-2
NE Tier-2
Lansing
ManLan
Ann Arbor
Chicago
Palo Alto
Western
Tier-2
BNL-SLAC via ESnet
MW Tier-2
U.S. ATLAS
Tier-1
Bloomington
Oklahoma City
SW Tier-2
Arlington
Major Exchange Points
The National Atlas of the United States of America

US ATLAS Summary

- This year LHC data begins flowing and ATLAS Physicists globally will be intently working to access and process data…there will be implications for networks, storage systems and computing resources.
- US ATLAS Planning provides for a sufficient network infrastructure at LHC startup:
  - **Tier-1: Multiple 10 Gbps capacity (10GE LHCOPN, 10GE US)**
  - **Typical Tier-2: 10 Gbps**
  - **Typical Tier-3: 1 (to 10) Gbps (depends on # of physicists and scale of resources)**