Infrastructure for Research Computing in West Virginia

Jack A. Smith, PhD

Cyberinfrastructure Coordinator

Division of Science and Research

WV Higher Education Policy Commission

Personal Background

- **WV Higher Education Policy Commission**
  - Division of Science and Research (WV EPSCoR)
 - Cyberinfrastructure Coordinator

- **Marshall University** (part-time)
  - Center for Environmental, Geotechnical, and Applied Sciences (CEGAS)
  - Virtual environments for underground mine safety training
  - Visualization Lab
  - High-Performance Computing cluster (*Big Green*)

- **XSEDE** (volunteer)
  - Marshall Campus Champion
  - Regional Champion
  - ECSS Fellow (H-T Nucleosome Modeling)
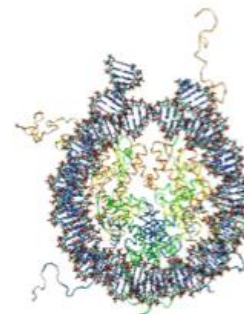
Personal Background

- **Mid-Atlantic Technology, Research, and Innovation Center**
  - Advanced Engineering Systems
  - Contract R&D for NASA IV&V, NASA Langley, DHS, Aurora Algae,…

- **Spatial Integrated Systems**
  - Autonomous Maritime Navigation (Navy contract)
  - Real-time multi-sensor data fusion with 3D vision

- **Union Carbide and Dow Chemical** (23 years)
  - Member of Catalyst Skills Center – Central R&D
  - Application of computational chemistry to catalyst discovery and optimization
  - Cheminformatics – high-throughput experiments and computation

- **University of Florida**
  - PhD –Quantum Chemistry & Solid State Physics
  - Member of the Quantum Theory Project (QTP)

Outline

Present

- NSF EPSCoR RII Projects
- HPC Clusters
- Visualization
- Broadband Deployment in WV
- WVNET
- National R&E Networks
- National HPC Resources
- XSEDE Campus Champions

Future

- GENI Rack
- Science DMZs
- Globus Online
- SDN
- InCommon & CILogon
- Eduroam
- NSF CC*IEE
- Condo of Condos
- HPC Portal / Gateway

NSF EPSCoR RII Projects

- **NanoSAFE (Track 1)**
  - **Shared Facilities, including HPC clusters**

- **CI-TRAIN (Track 2)**
  - **Collaborative project between WV and AR**
  - **Cyberinfrastructure**
    - HPC clusters
    - Visualization centers
    - Virtual collaboration (videoconferencing)
  - **Training and outreach**
    - CI Days at Marshall and WVSU

- **RII-C2**
  - **Cyberinfrastructure (networking) upgrades at WVU and Marshall**
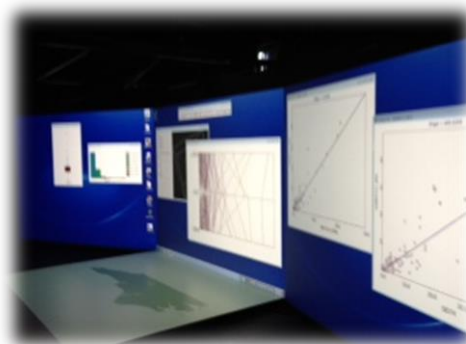 - **Internet2 SEGP Sponsorship**
    - Marshall → WVNET

HPC Clusters

- **"Mountaineer" and "Spruce Knob" @ WVU**
- **"Big Green" @ Marshall**
- **"Stinger" @ WVSU**

Visualization

- **Visualization Lab @ Marshall**
- **Geo-visualizaton CAVE @ WVU**
- **Viz Wall @ WVSU**

Broadband Deployment in WV

- **BTOP (NTIA)**
  - $126MM
  - 2400 miles of fiber
  - Routers at 1500+ Community Anchor Institutions (CAI)

- **WVU-GB (NRAO) link**
  - Supplement to BTOP award

- **Surplus routers to higher education institutions**
  - Left out of the state BTOP proposal
    - Separate proposal by HEPC not awarded

- **WV Broadband Deployment Council (BDC)**
  - State-funded grants
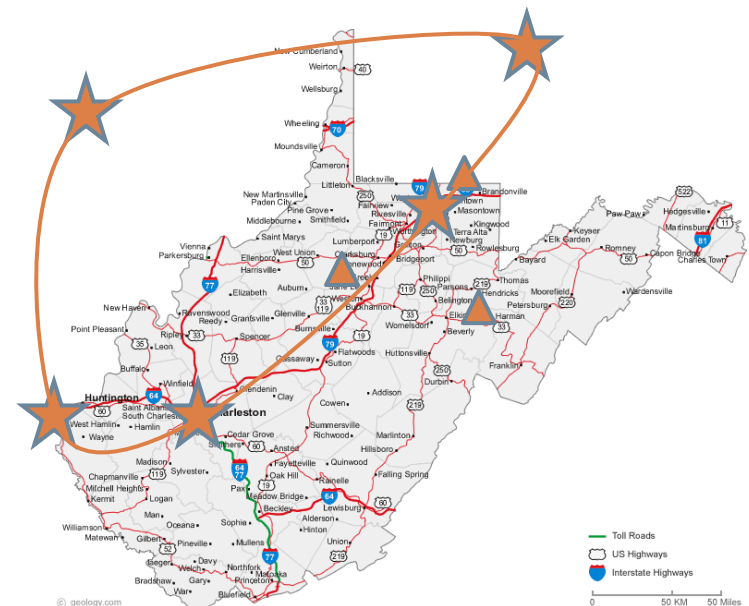   - Underserved regions
    - Last mile
    - Demand promotion

WVNET
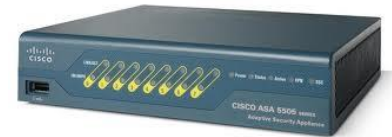
- **10Gig Ring**

  - Pittsburgh (3ROX) – Internet2, NRL
  - Morgantown (WVU)
  - Charleston (Capital Complex)
  - Huntington (Marshall)
  - Columbus (OARnet) – Internet2

- **Connections to:**

  - NOAA (Fairmont)
  - NRAO (GB)
  - NETL (Morgantown)

National R&E Networks

□ **Internet2**

    □ SEGP via EPSCoR RII C2 grant

        ■ Administered by WVNET (was Marshall)

        ■ Allows all state schools to be sponsored members

    □ OARnet, 3ROX

    □ NET+ Services

    □ InCommon

        ■ Single-Sign-On (SSO)

National HPC Resources

- **XSEDE**
  - Formerly the *TeraGrid*
  - 12 Supercomputing Centers
  - New attention being paid to
    - Big Data
    - High-Throughput Computing
    - Campus Bridging

- **Others HPC Grids**
  - OSG (Open Science Grid)
  - FutureGrid

- **Cloud HPC (and Data) Provisioning**
  - Amazon's *AWS/EC2*
  - Google's *App Engine*
  - Microsoft's *Azure*

# XSEDE Campus Champions

- Don McLaughlin (WVU)
  - Workgroup on EOT
- Jack Smith (Marshall, HEPC)
  - ECSS Fellowship
- Sridhar Malkaram (WVSU)

GENI Racks

- **G**lobal **E**nvironment for **N**etwork **I**nnovations
- **Platform for Campus or Regional Cloud Services**
- **Federation of Shared Resources**
  - Computing
  - Data Storage
  - Network Infrastructure
- **Software-Defined Networking (SD**
- **WiMAX**
- **Help Provision Science DMZs**
- **WVNET-Cisco Partnership**
  - Looking for Apps and Campus Champions
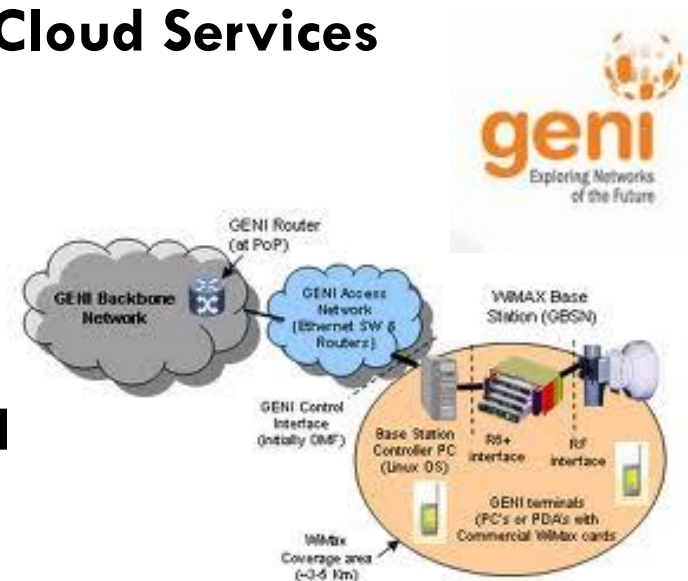
Science DMZs

- **"Friction Free" Network Path**
  - Virtual circuit
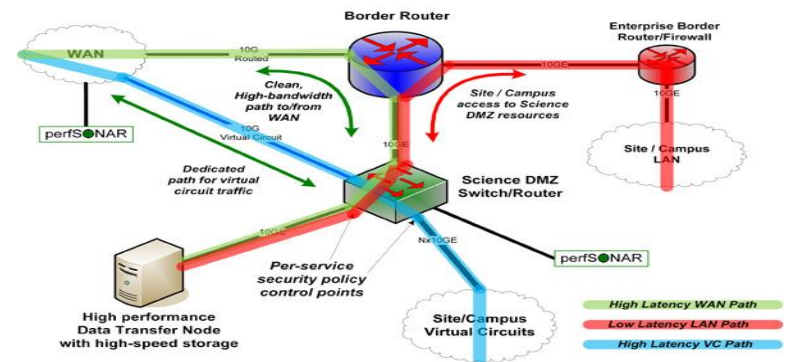 - Limited firewall overhead

- **Data Transfer Node (DTN)**
  - Generally sits on border
  - *Globus* endpoint

- **Performance Measurement**
  - *perfSONAR*

- **Switches that support SDN**
  - *OpenFlow*

Globus Online

Globus *was developed to enable grid computing: by connecting computing resources, the data can be freed from its initial source, and made portable—even if it's huge.*

- a project of the Computation Institute, which is a partnership between The **University of Chicago** and **Argonne National Laboratory**

- supported by funding from the **Department of Energy,** the **National Science Foundation**, and the **National Institutes of Health**

SDN

- **S**oftware-**D**efined **N**etworking
  - Decoupling of data and control planes
  - Helping cloud computing blur the distinction between computers and networks
  - *OpenFlow* – from the Open Networking Foundation (ONF)
  - Becoming widely adopted (Google, Facebook,…)
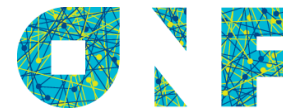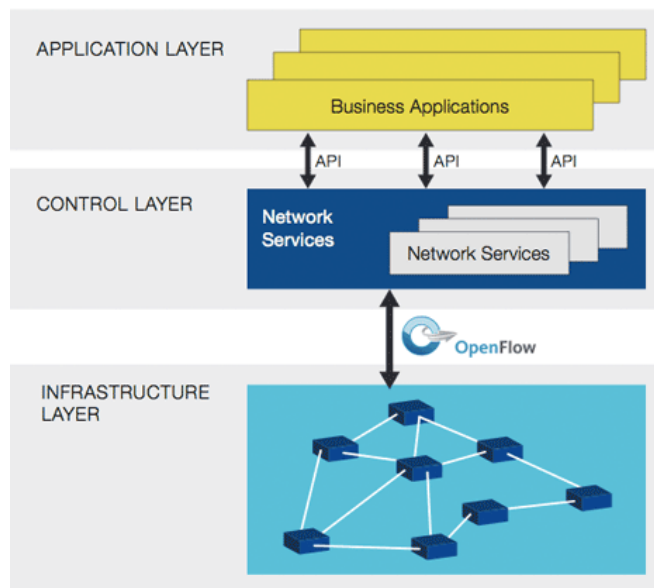
OPEN NETWORKING FOUNDATION

InCommon & CILogon

- **InCommon**
  - Provides a secure and privacy-preserving **trust fabric** for research and higher education, and their partners, in the US.
  - Operates an **identity management federation**, a related assurance program, and offers **certificate and multifactor authentication services**
  - Operated by Internet2
- **CILogon**
  - A Federated **X.509 Certification Authority** for CyberInfrastructure Logon
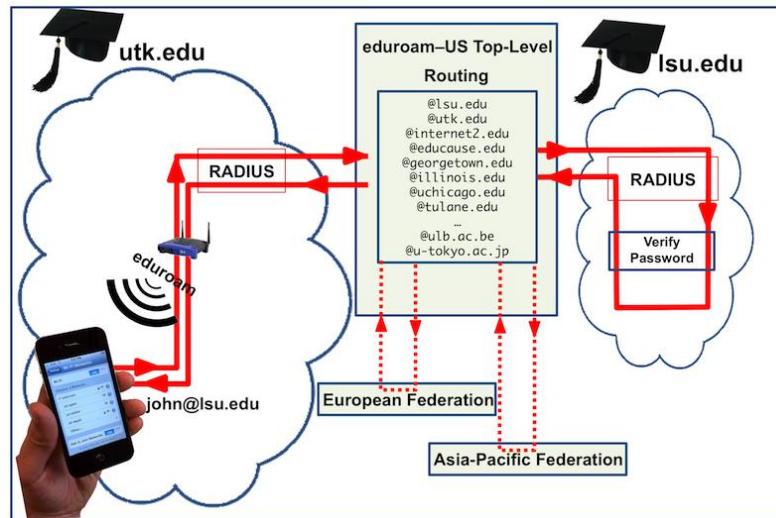 - Developed at NCSA

Eduroam

- *Eduroam (**edu**cation **roam**ing) is the secure worldwide federated network access service developed for the international research and education community.*
  - *Combines the power of 802.1X, SSL, and RADIUS to create a standards based global trust fabric*

NSF CC*IEE

- NSF Grant
  - Directorate for Computer & Information Science & Engineering   Division of Advanced Cyberinfrastructure   Division of Computer and Network Systems
- Campus Cyberinfrastructure - Infrastructure, Innovation and Engineering Program (CC*IIE)
- Follow-up to CC*NIE
  - Focused on Science DMZs and Networking
- Broadened to 5 Areas
- Next solicitation expected Spring 2015
- Internet2-sponsored workshops

Condo of Condos

*The ACI-REF **consortium** includes six institutions that embrace the **condominium computing model**. We are dedicated to forging a nationwide **alliance of educators** to empower local campus researchers to be more effective users of advanced cyberinfrastructure (ACI). In particular, we seek to work with the "academic missing middle" of ACI users—those scholars and faculty members who traditionally have not benefited from the power of massively scaled cluster computing but who recognize that their research requires access to more compute power than can be provided by their desktop machines.*

$5.3M NSF ACI Grant to Clemson

ACI-REF
Advanced CyberInfrastructure - Research and Education Facilitators
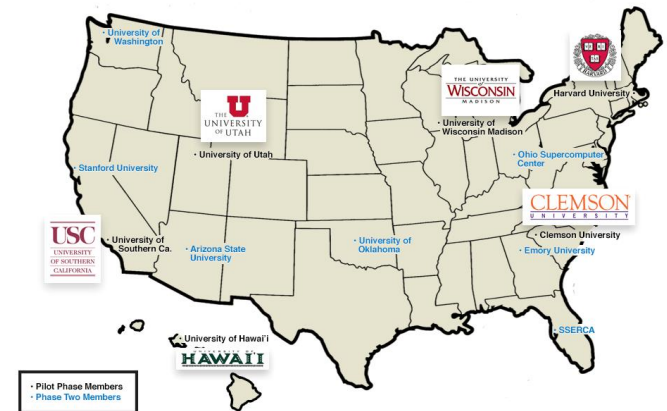
HPC Portal / Gateway

- **HUBZero (*Purdue*)**
  - Web-based Collaboration Framework
- **Airavata (*Apache*)**
  - Gateway and Workflow Development for Distributed Environments
- **BigBlueButton**
  - Online Collaborative Learning Environment
- **Moodle**
  - Online Learning Management System (LMS)

- **Common Statewide Portal**
- **Open Source Tools**
- **Community Developed**