Research Computing Services: Current and Future

Erik Deumens

Nov 7, 2013

Ready, Working, Planning

- Services offered today
- Work in progress
- Planning for new services

Ready…

Research Computing Services

- Compute
- Storage
- Training
- Consulting

UF Data Center

HiPerGator supercomputer

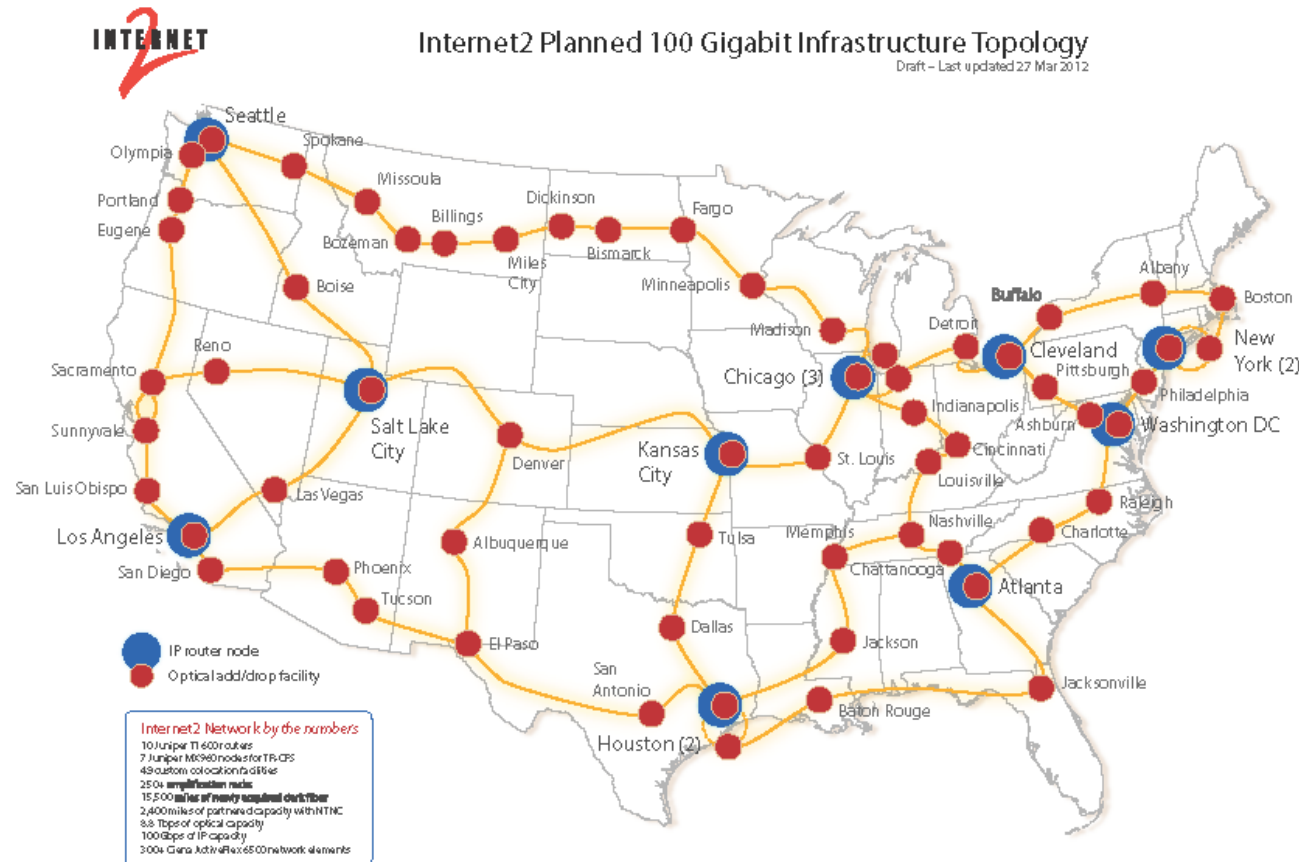Nov 7, 2013   6

Compute and storage

▸ Over 22,000 cores total

▸ HiPerGator has
  ◦ 16,000 cores,
  ◦ 65 TB RAM,
  ◦ 2 PB scratch disk

▸ Buy some for your group:
  ◦ $200 per core for 5 years
  ◦ Matching Program makes a deal you cannot refuse!
  ◦ $250 per TB per year for replicated long-term storage

Internet2 Innovation Platform

Internet2 Planned 100 Gigabit Infrastructure Topology
Draft – Last updated 27 Mar 2012

Working…

Work in progress

▸ We are working on several projects
  ◦ to be completed in the first three months of 2014.

▸ Globus Online endpoints
▸ Collaborative research data storage
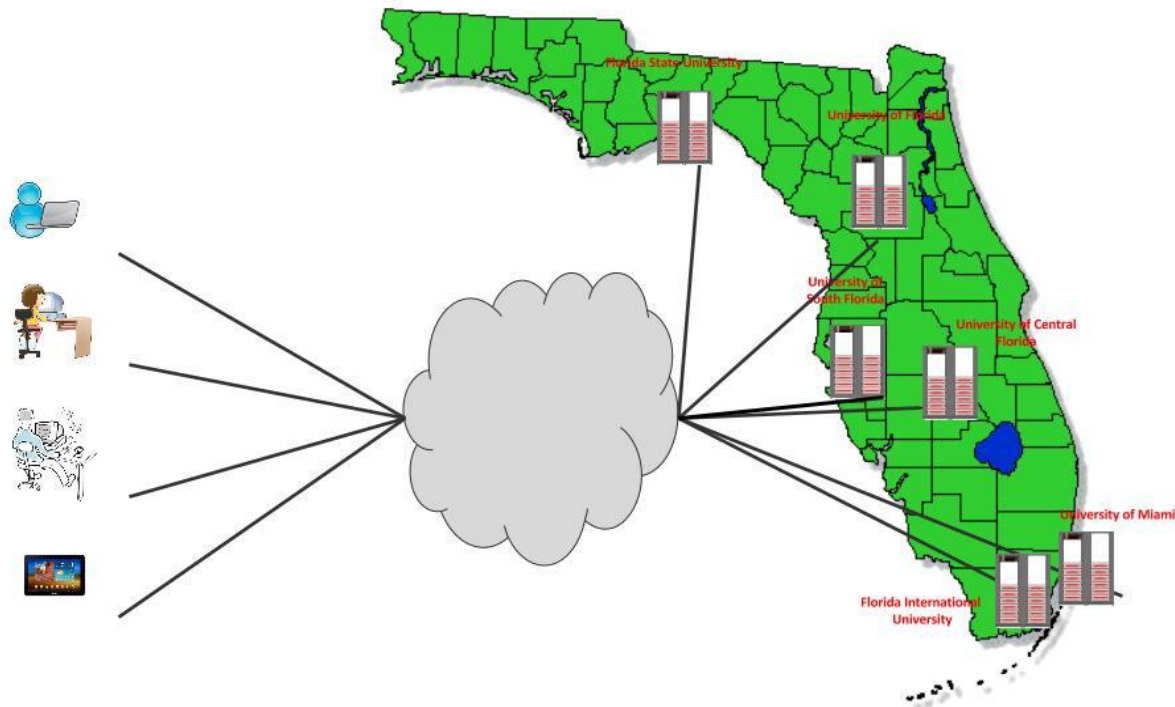▸ MRI 2012 Gatorcloud cluster
▸ HIPAA compliant facility

Globus Online endpoints

▸ 12 data transfer nodes are in place

▸ Part of 200 G Campus Research Network
  ◦ Now called Science DMZ

▸ Close to 100 G connection

▸ Easy tool to control data transfers
  ◦ Between published endpoints
  ◦ From laptop and tablet

Collaborative research data

- The long term storage will be replaced
  - Object based
  - CIFS access to researcher machines
  - Dropbox-like upload, download, sync
- Policy-based replication
  - RAID-style, replication once or twice, and off-site
- Simple sharing process for collaborators
  - No AD or GatorLink needed
- Price to be determined
  - Initial cost same as now $250 per TB per year

Sunshine State Education Research Computing Alliance

Gatorcloud cluster

- MRI 2012 grant with PI Andy Li
- Run VMs in Openstack environment (KVM)
- Support for Big Data analytics: Hadoop, Pig
- Integrated with HiPerGator and its storage systems

HIPAA compliant facility

- Research data storage
- Compute cluster with scratch storage for processing
- Secure gateway to control access
- Use VMs (VDI)
  - Allow researchers to work from any device
  - Support complex computation work flows

HIPAA compliant facility (2)
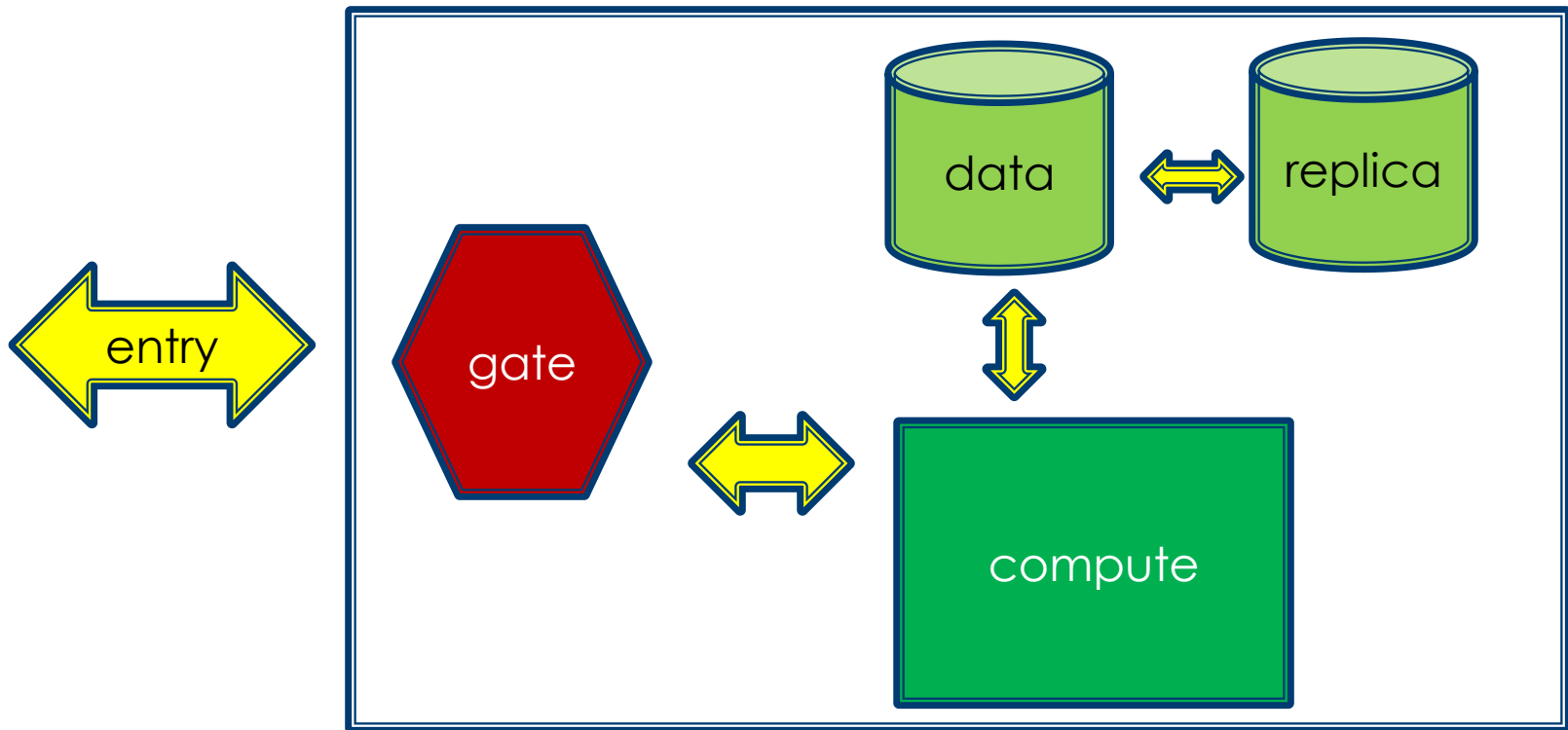
Planning...

Planning for new services

▸ Existing services are basic

▸ Many researchers could use

▸ But they need tools…
  ◦ Researchers
  ◦ Connecting layer:
    • Web portal, virtual apps, database servers
  ◦ Basic services:
    • Compute, Storage, Network

Some requests already received…

Active web access

▸ Researchers have data on RC resources
  ◦ They want to share for collaboration
    • With authentication
    • Some data may be public
  ◦ The web page is not static but needs to support active content
    • Search for data
    • Allow data upload for some special processing
  ◦ Researchers do not have or want to have admin responsibility for the server, only for their data
    • A cloud VM would not work
  ◦ Functionaily should remain after the project
    • Indefinitely…

Applications

- Researchers want to have
  - Easy access, like UF Apps, to
  - Windows OS applications, like ArcGIS, working on
  - Research data that can be
    - Processed by HiPerGator
    - Shared across 100 G Internet2 across the world
- Currently this is only supported for Linux applications

Visualization support

▸ Researchers need
  ◦ Tools,
  ◦ Training,
  ◦ Consulting
▸ On advanced visualization in many domains
▸ To have a visualization interest group on campus would be great
  ◦ Help researchers get started
  ◦ Advise RC where best to invest resources

Big Data analytics

▸ Researchers need an environment
  ◦ With the modern analytics tools like Hadoop, Pig, ready to go
  ◦ At scale for large data sets with many cores and GPUs

Computer science experiments

▸ Computer science researchers need
  ◦ An environment where they can carry out
  ◦ Complex experiments
  ◦ At scale
  ◦ That involve kernel and network changes
  ◦ And can break things
▸ A virtual cluster of VMs in a virtual network will enable this
  ◦ Simple and quick provisioning is needed

Requests?