



# Cornell University Center for Advanced Computing A Sustainable Business Model for Advanced Research Computing

David A. Lifka  
[lifka@cac.cornell.edu](mailto:lifka@cac.cornell.edu)



# My Background



- 2007 – Cornell Center for Advanced Computing – Director**
- 2010 – Weill Cornell Medical College – Director Research Computing**
- 2006 – Cornell Computing and Information Science – Adjunct Associate Professor
- 2005 – 2007 Cornell Theory Center – Director of High Performance and Innovative Computing
- 2001 – 2005 Cornell Theory Center – Chief Technical Officer
- 1999 – 2001 Cornell Theory Center – Associate Director
- 1995 – 1999 Cornell Theory Center – Senior Systems Programmer



- 1993 – 1995 Mathematics and Computer Science Division – Senior Technical Support Analyst
- 1992 – 1993 High Energy Physics Division - Petabyte Access Storage Solutions project – Associate Computer Scientist
- 1991 – 1992 Computing and Telecommunications Division – Associate Computer Scientist
- 1988 – 1991 Computing and Telecommunications Division – Assistant Computer Scientist
- 1986 – 1988 Materials Science Division – Internship



## Other Hats I Wear



Coalition for Academic Scientific Computation

2012 – Chair  
2010 – 2012 Vice Chair

**XSEDE**

Extreme Science and Engineering  
Discovery Environment

2011 – Architecture and Design –Coordinator



# Center for Advanced Computing

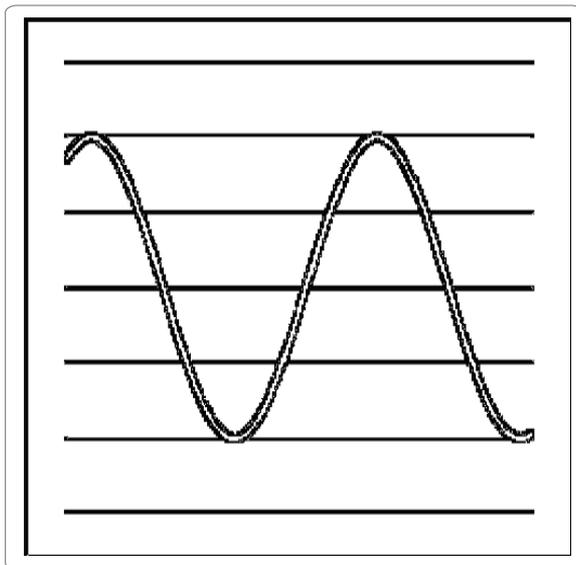
## Developing a Sustainable Business Model



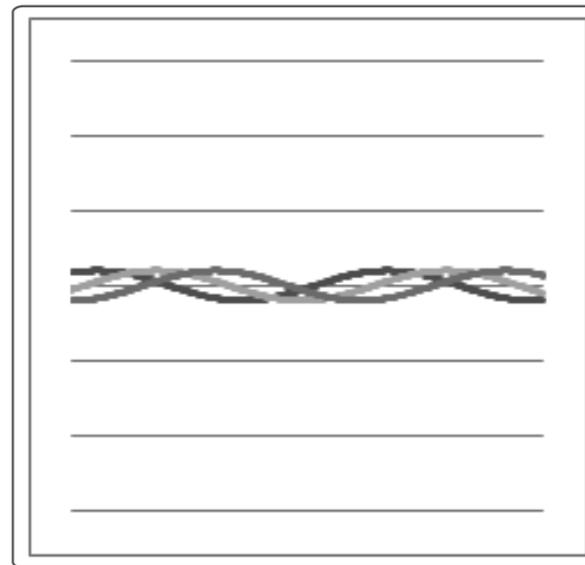


# The Funding Rollercoaster

## Changing the Cornell Research Computing Funding Model



I want to get off of this ride!



In search of sustainability...



# Research Computing Challenges

## What keeps me up at night....

### **Budget & Funding**

- Staff
- Computing
- Data storage & archival
- Networking
- Visualization
- Facilities (power, space & cooling)

### **Technology & Infrastructure**

- Constantly changing
- Expensive
- One size does NOT fit all

### **Strategic Planning**

- Understanding your researcher's needs and challenges
- Funding opportunities: leadership vs. partnership
- Campus Bridging to National CI



## 2007 – Developing a Sustainable Funding Model

- Provost, Vice-Provosts, Deans & representative faculty involved
- Understand the technical requirements and financial limitations
  - Costs for staff and services were documented and reviewed
  - Open discussions on what faculty need AND were willing to pay for
  - Develop services that meet these requirements and provide competitive rates, seeking economies of scale and scope wherever possible
    - Retain only the staff and services that faculty were willing to pay for
- Created a “Core Facility” that *serves* the Cornell research community
  - Director reports to the Vice Provost for Research
  - 80% Cost recovery required
  - Provost subsidy provided to create proper incentives for centralization
    - Must be better than graduate student labor AND cheaper
- Deans provided bridge funding to their faculty to assist transition to new model



## Goals of Sustainable Model

- A successful recovery/funding model must be institutionalized
  - Enable a broad array of researchers
  - Provide value
  - Efficient and fair
  - Provide economies of scale and scope
  - Save money
- Not based on winning a single major grant or on the efforts of a particular director and/or management team
  - Model documented so that the institution can stand behind it over time and expect it to be acted upon
  - *Model must allow room for strategic change, growth and adaptation*
  - Model should allow core staff to focus on their jobs rather than constantly worrying about funding for their jobs



## Focused on Service

*Enabling the success of Cornell researchers, collaborators and supporters  
whose work demands advanced computing solutions*

### Consulting

- Assisting with start-up packages for new faculty
- Benchmarking & performance analysis
- Proposal development & participation
- Custom programming, debugging, parallelization & optimization
- Development and support for scientific workflows
- Custom training (live & web-based)
- Desktop & remote visualization
- Strategic partnerships with HPC vendors, national computational science centers and researchers

### Computing

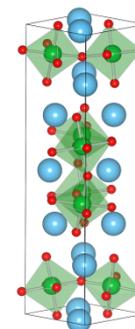
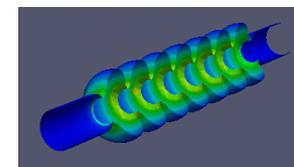
- Subscriptions for Red Cloud
- Private cluster maintenance agreements  
(*secure data hosting arrangements possible*)

### Data storage and management

- Scalable disk storage
- Archival storage Q2 2013
- Database resources
- 10Gb connectivity to the Internet

### National cyberinfrastructure partnerships

- NSF XSEDE – *Extreme Science and Engineering Discovery Environment*
- TACC Stampede – *Enabling, Enhancing, and Extending Petascale Computing for Science and Engineering*



  
**red cloud**

*For more information, contact David Lifka  
Director, Cornell University Center for Advanced Computing  
Director of Research Computing, Weill Cornell Medical College  
lifka@cac.cornell.edu*



## Foster Innovation, Adaptation and Growth

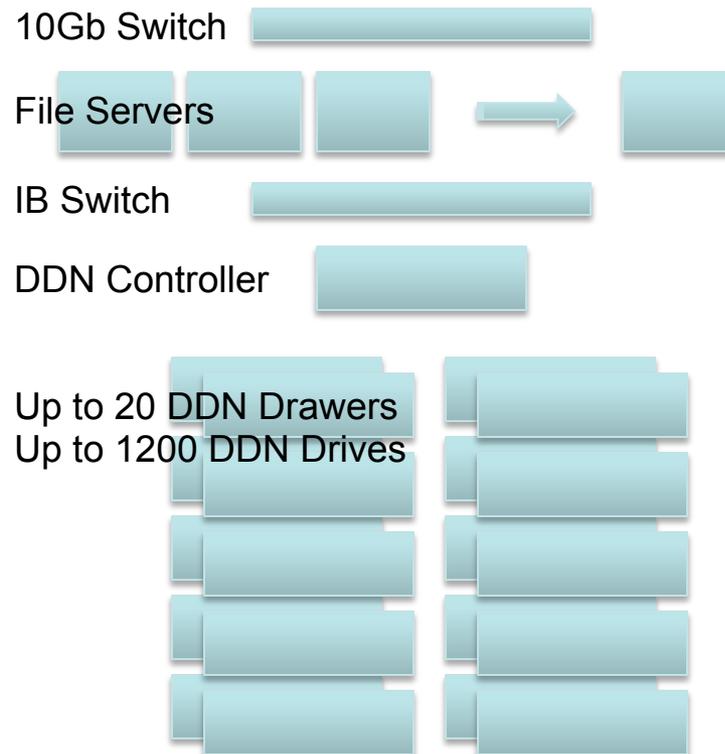
- CAC's focus on providing advanced computing services based on technologies which provide optimal economies of scale makes us more relevant to funding agencies, partners, industry and vendors
- Having a sustainable recovery model forces us to do a great job of identifying key technologies and implementing "right-sized solutions"
- Having vendor partners we can rely on to provide the best price/performance and maintenance support is essential





# DataDirect Networks Storage

- Enterprise class solution
  - Provides great economies of scale & scope
  - Expensive entry level, but given enough users becomes very affordable
  - Used for user file systems, Red Cloud EBS, Globus Online endpoints and new archival storage service
- Leveraged by other Cornell Core facilities
  - Life sciences using Gluster for performance and large single namespace
-  **redhat**
  - Received 2012 Red Hat Innovation Award
  - Medical school for remote 2<sup>nd</sup> copy of critical data
    - (Currently no PHI or PII supported)





*red cloud*

The logo for Red Cloud, consisting of three stylized, overlapping red loops above the text "red cloud" in a bold, italicized, black sans-serif font.

On-Demand Research Computing

- Infrastructure as a Service –
- Software as a Service –
- Cloud Storage Solutions –



## *red cloud* Motivation

- **Research computing means many different things...**
  - Scientific workflows have different requirements at each step
  - Cloud is only part of the solution
  - Connecting to and from other CI resources is important
- **Nobody likes a bad surprise**
  - Transparency, no hidden costs
  - Need a way to bound financial risk
  - Geographically distributed resources for disaster recovery
  - Bursting for peak periods vs. over investment in dedicate infrastructure
- **Economies of scale**
  - Sharing hardware and software where it makes sense
  - Pay for what you need, when you need it
- **Customized environments for various disciplines**
  - Collaboration tools
  - Data storage & analysis tools
  - Flexibility to support different computing models (e.g. Hadoop)



## *red cloud* Provides

### **Predictable, Reproducible, Reliable Performance**

We publish hardware specifications (CPU, RAM, network) and do not oversubscribe.

### **Convenient**

Need system up and running yesterday.

Need a big fast machine for only a few months, weeks or days.

Need a small server to run continuously.

### **No Hidden Costs**

No cost for network traffic in or out of the cloud.

### **Fast Access to Your Data**

Fast data transfers via 10Gb Ethernet in or out of the cloud at no additional charge.

Globus Online access

### **Economies of scale**

IaaS: Infrastructure

SaaS: Software

Storage: EBS, S3, NFS & CIFS

### **Expert Help**

System, application, and programming consulting are available.

### **Easy Budgeting with Subscriptions**

No billing surprises!

### **Eucalyptus is Amazon API Compatible**

Migrate when your requirements outgrow Red Cloud.



## Red Cloud IaaS Virtual Server Configurations

- **CPU** 1, 2, 4, 8, or 12 cores
- **RAM** 4GB/core
- **Network** 10Gb Ethernet *shared with all virtual servers on the same server*
- **Disk** Local scratch disk *deleted when cloud instance is terminated*
- **OS** CentOS 5.6 & CentOS 6 *user-customizable*
- **Privileges** Root access

Type	Cores	RAM	Disk
M1.small	1	4GB	20GB
C1.medium	2	8GB	40GB
M1.large	4	16GB	200GB
M1.xlarge	8	32GB	400GB
C1.xlarge	12	48GB	1000GB



# Software as a Service (SaaS)

**red cloud** with MATLAB provides:

- Seamless access to MATLAB Distributed Computing Server (MDCS) workers from your personal workstation
- Access to NVIDIA GPUs for added performance\*
- “Quick Queue” for quick turnaround and scientific gateways
- Received HPC Innovation Excellence Award in 2011 from the International Data Corporation

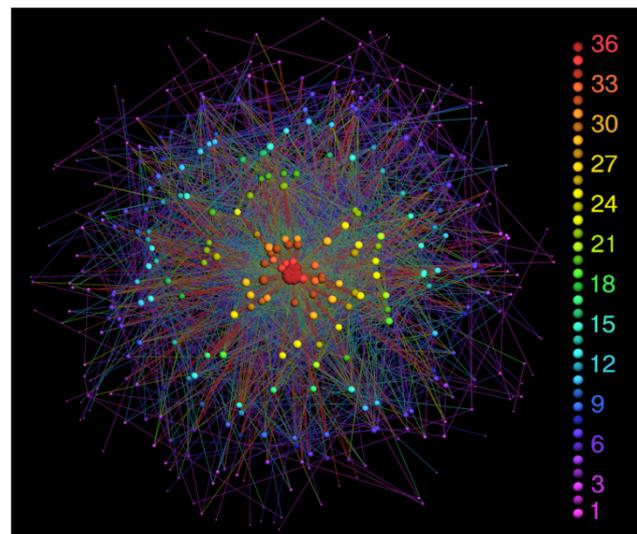


Requires MATLAB Client with Parallel Computing Toolbox

\*Use of GPUs can decrease your code run time providing optimal use of your subscription



TeraGrid  
Experimental Resource



*512-core simulation of networks of coordinated amino acid variation in Hepatitis C virus, a major cause of liver disease worldwide, enabled by the MATLAB on the TeraGrid experimental resource at CAC*



## 2012 – Hurricane Sandy Disaster brings new opportunity

### **A Flooded Mess That Was a Medical Gem**

By ANEMONA HARTOCOLLIS

Published: November 9, 2012



The federal government's emergency management chief trudged through darkened subterranean hallways covered with silt and muddy water Friday, as he toured one of New York City's top academic medical centers in the aftermath of Hurricane Sandy. The basement of the complex, NYU Langone Medical Center in Manhattan, smelled like the hold of a ship — a mixture of diesel oil and water.



4/20/13



[www.cac.cornell.edu](http://www.cac.cornell.edu)

17



# WCMC Secure Network Extension

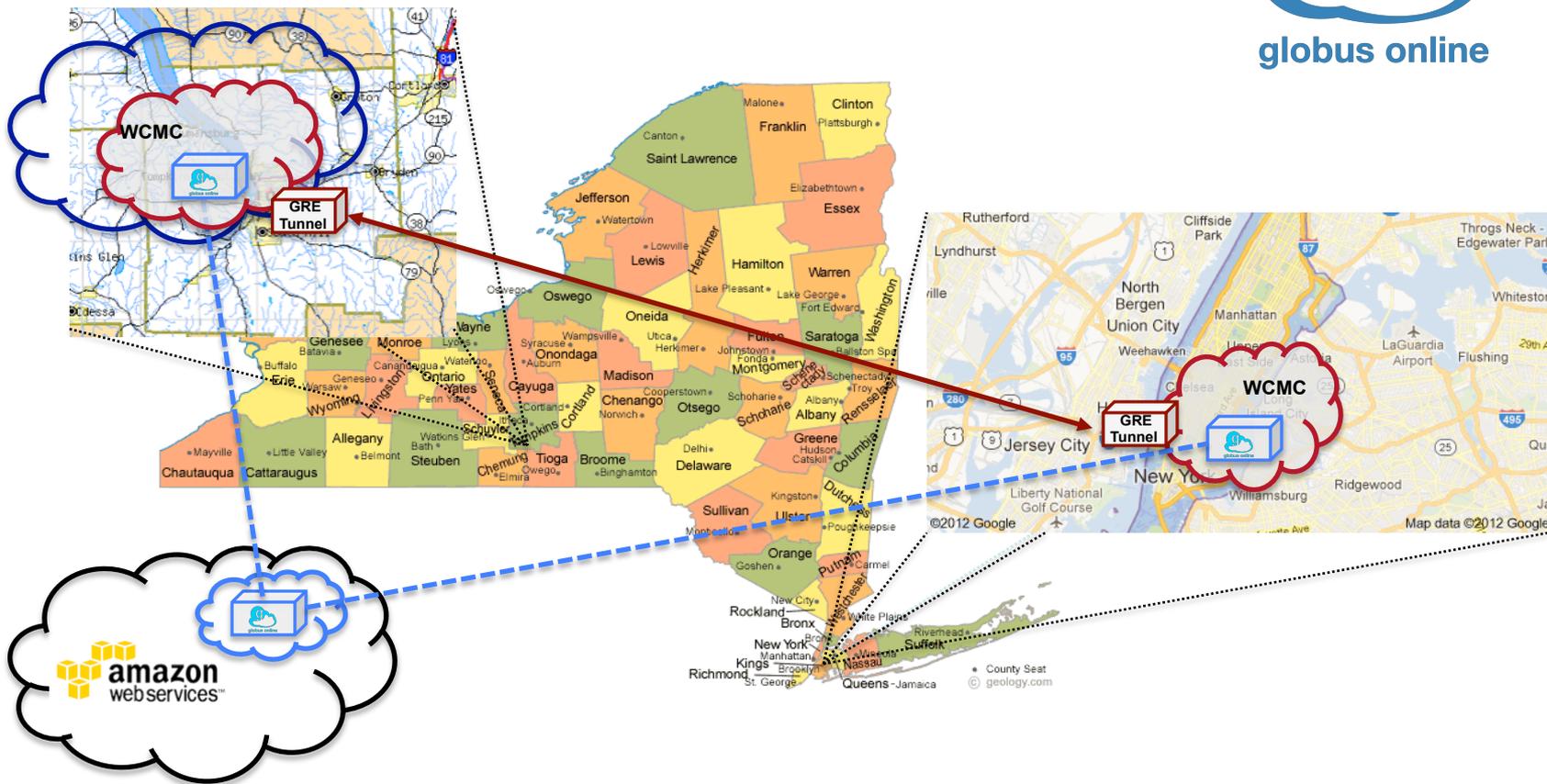
HIPAA Compliant  
Co-location Facility  
Cornell Ithaca, NY Campus





# WCMC Firewall Extension to Ithaca

## Co-location for Critical Servers



A stylized red cloud icon composed of three overlapping loops, with a small gold padlock icon attached to the right side.

# *red cloud*

## **SECURE**

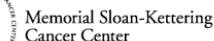
- HIPAA Compliant Secure Solutions -



# Secure Hybrid Cloud Model

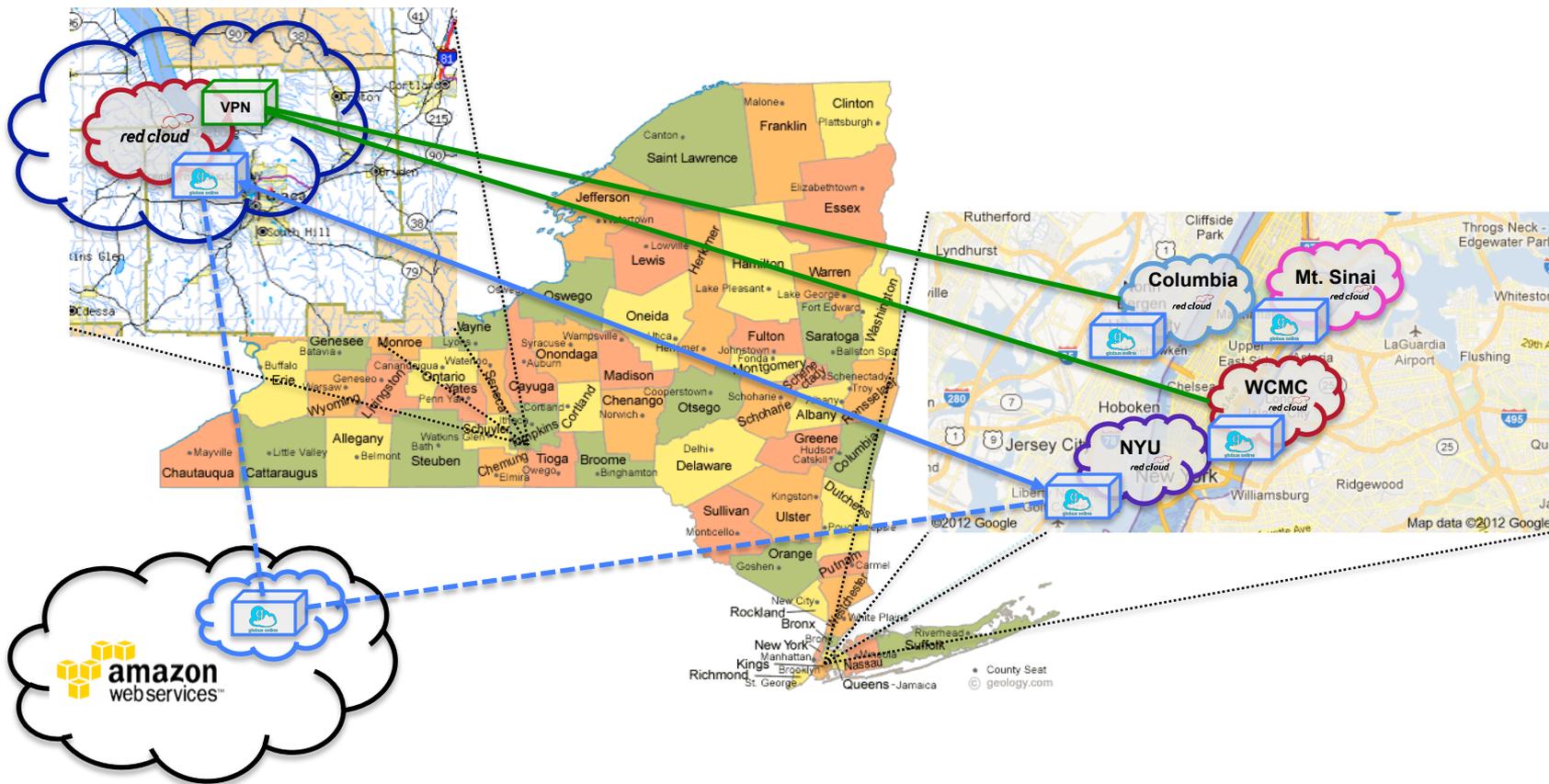
## Elasticity & Disaster Preparedness

- Provide Cloud resources locally and a remote shared resource (Ithaca):
  - IaaS Compute Resources
  - Elastic Block Storage
    - Correct price-performance for research applications
  - Archival storage
  - Provide preconfigured VMs and software tools
    - Hope to include Globus Genomics (<http://www.globus.org/genomics/>)
- Local resources important for:
  - Data locality issues (avoid transfer delays for critical analyses)
  - Supporting sustained workloads (80% utilization or better)
- Remote resources (Ithaca) important for:
  - Disaster recover/redundancy for critical systems
  - Elasticity while shared capital costs with partners (economies of scale)
- Initially a partnership between NYU Langone & WCMC
  - Understand and work through issues
    - Political, Social, Financial, Logistical & Technical
    - Start with a small group and scale as appropriate
    - Expand to HPCBMR & New York Genome Center Partners





# Red Cloud Secure New York City Institutions





# Thank You!