Cornell University Center for Advanced Computing

David A. Lifka - lifka@cac.cornell.edu
Director - Cornell University Center for Advanced Computing (CAC)
Director Research Computing - Weill Cornell Medical College (WCMC)
Adjunct Associate Professor - Cornell Computing and Information Science (CIS)

www.cac.cornell.edu/redcloud
On-Demand Research Computing
- Infrastructure as a Service -
- Software as a Service -

www.cac.cornell.edu/redcloud
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

• **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

**Expert Help**
- System, application, and programming consulting are available.

**Easy Budgeting with Subscriptions**
- No billing surprises!

**IaaS is Amazon API Compatible**
- Migrate when your requirements outgrow Red Cloud.

www.cac.cornell.edu/redcloud
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

Configuration
- **Total**
  - 64 cores, 192GB of RAM in 8 servers:
- **8 Servers**
  - 2, 2.4 GHz Xeon E5620 CPUs (2 x 4 cores)
  - 24GB of RAM (3GB/core)
  - 1 NVIDIA Tesla M2070 GPU per server
- **Gigabit Ethernet interconnect**
- **Storage**
  - 8TB mounted via CIFS
- **Software**
  - MATLAB MDCS 2011B
  - CAC parallel MATLAB client code

Requires MATLAB Client with Parallel Computing Toolbox

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

www.cac.cornell.edu/redcloud
MRI Analysis in MATLAB with GPUs

- **Goal:** accelerate existing MATLAB codes with PCT, GPUs
- **Researchers:** Steve Lantz, CAC in collaboration with Ashish Raj and Miloš Ivković, Weill-Cornell Medical College
- **Research question:** Given two different regions of the human brain, how interconnected are they?
- **Potential impact of the technology:**
  - Study of normal brain function
  - Understanding medical conditions that harm brain connections, such as multiple sclerosis, Alzheimer’s disease, brain injury
  - Surgical planning

www.cac.cornell.edu/redcloud
Connecting Two Types of MRI Data

- Analysis requires optimization of a weighted bigraph that connects two types of MRI brain data (fibers, voxels)
- Method is an iterative procedure known as “min-sum”
GPU Optimization in MATLAB

• All structs had to be converted into matrices
  – The GPU requires simple data structures for SIMD operations
  – Vectorized matrix operators are fast; nested for-loops are slow

• Resulting code is actually less complex
  – Just needed to move matrices to the GPU with gpuArray()
  – Standard operations like min, sum are polymorphic, so they use MATLAB’s built-in CUDA code on type GPUArray

• Best result (after a few tricks): 0.15 sec/iteration
  – 350x speedup over initial matrix-based version
  – 2500x speedup over initial struct-based version

www.cac.cornell.edu/redcloud
Infrastructure as a Service (IaaS)

red cloud provides on-demand:

- **Virtual Servers**
- **Storage**: Virtual disks for the virtual servers

Plus the necessary fast networking, power, space, and cooling to make everything work.

**Configuration**

<table>
<thead>
<tr>
<th>Description</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>96 cores, 384GB of RAM in 8 servers:</td>
</tr>
<tr>
<td><strong>8 Servers</strong></td>
<td>2, 2.7 GHz Xeon E5650 CPUs (2 x 6 cores)</td>
</tr>
<tr>
<td></td>
<td>48GB of RAM (4GB/core)</td>
</tr>
<tr>
<td></td>
<td>Maximum virtual server size:</td>
</tr>
<tr>
<td></td>
<td>• 12 cores, 48GB RAM</td>
</tr>
<tr>
<td><strong>10 Gigabit Ethernet interconnect</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>7.5TB</td>
</tr>
<tr>
<td></td>
<td>Accessible by virtual servers via 10Gb iSCSI</td>
</tr>
<tr>
<td><strong>Software</strong></td>
<td>Eucalyptus 2.0.3 (Open Source)</td>
</tr>
</tbody>
</table>

www.cac.cornell.edu/redcloud
Subscription-based Recovery Model

<table>
<thead>
<tr>
<th>Institution</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cornell University</td>
<td>$500/core year*</td>
</tr>
<tr>
<td>Other Academic Institutions</td>
<td>$750/core year</td>
</tr>
<tr>
<td>Cornell University</td>
<td>$750/core year</td>
</tr>
<tr>
<td>Other Academic Institutions</td>
<td>$1200/core year</td>
</tr>
</tbody>
</table>

*A core year is equal to 8585 hours

*Each subscription account includes 50GB of storage

www.cac.cornell.edu/redcloud
What if ???

<table>
<thead>
<tr>
<th></th>
<th>Consulting</th>
<th>Additional Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cornell Users</strong></td>
<td>$59.90/hour</td>
<td>$0.91/GB/year</td>
</tr>
<tr>
<td>**Other Academic</td>
<td>$85.47/hour</td>
<td>$1.45/GB/year</td>
</tr>
<tr>
<td>Institutions**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Type</th>
<th>Cores</th>
<th>RAM</th>
<th>Disk</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1.small</td>
<td>1</td>
<td>4GB</td>
<td>20GB</td>
</tr>
<tr>
<td>C1.medium</td>
<td>2</td>
<td>8GB</td>
<td>40GB</td>
</tr>
<tr>
<td>M1.large</td>
<td>4</td>
<td>16GB</td>
<td>200GB</td>
</tr>
<tr>
<td>M1.xlarge</td>
<td>8</td>
<td>32GB</td>
<td>400GB</td>
</tr>
<tr>
<td>C1.xlarge</td>
<td>12</td>
<td>48GB</td>
<td>1000GB</td>
</tr>
</tbody>
</table>

www.cac.cornell.edu/redcloud
Red Cloud IaaS Elastic Block Storage (EBS)

- Each EBS volume is a virtual disk that can be attached to or detached from a cloud virtual server on demand. Inside a virtual server, each attached EBS volume is a block device (e.g. /dev/vdb)
- User can create EBS volumes up to 1TB in increments of 1GB.
- EBS volumes connect to virtual servers in Red Cloud via iSCSI for best performance (300+ MB/sec observed in testing).
- Cloud virtual servers can serve storage to other cloud virtual servers or clients external to the Red Cloud (e.g. NFS or CIFS shares, WebDAV, databases)
- Cloud virtual servers can access network storage external to Red Cloud

www.cac.cornell.edu/redcloud
Use Cases to Consider

- **Support for Scientific Workflows**
  - Pre & post-processing of data and results
  - Data analysis
  - Globus Online for fast reliable data transfer
    - https://www.globusonline.org/

- **Collaboration**
  - Wiki hosting
  - Customized data analysis & computational environments

- **Web Portals**
  - Science Gateways
  - Domain Specific Portals
  - Hub Zero
    - http://hubzero.org/pressroom
    - http://nanohub.org

- **Education, Outreach & Training**
  - Pre-configured systems & software tools providing consistent training platform

- **Bursting**
  - Additional software and hardware on demand

www.cac.cornell.edu/redcloud
More Info

• Red Cloud Technical Documentation:  

• Red Cloud with MATLAB Technical Documentation:  

• Questions? E-mail: help@cac.cornell.edu