UNIVERSITY OF ALBERTA FACULTY OF SCIENCE

The Cirrus Research Computing Cloud

Faculty of Science ITOC
April 2018

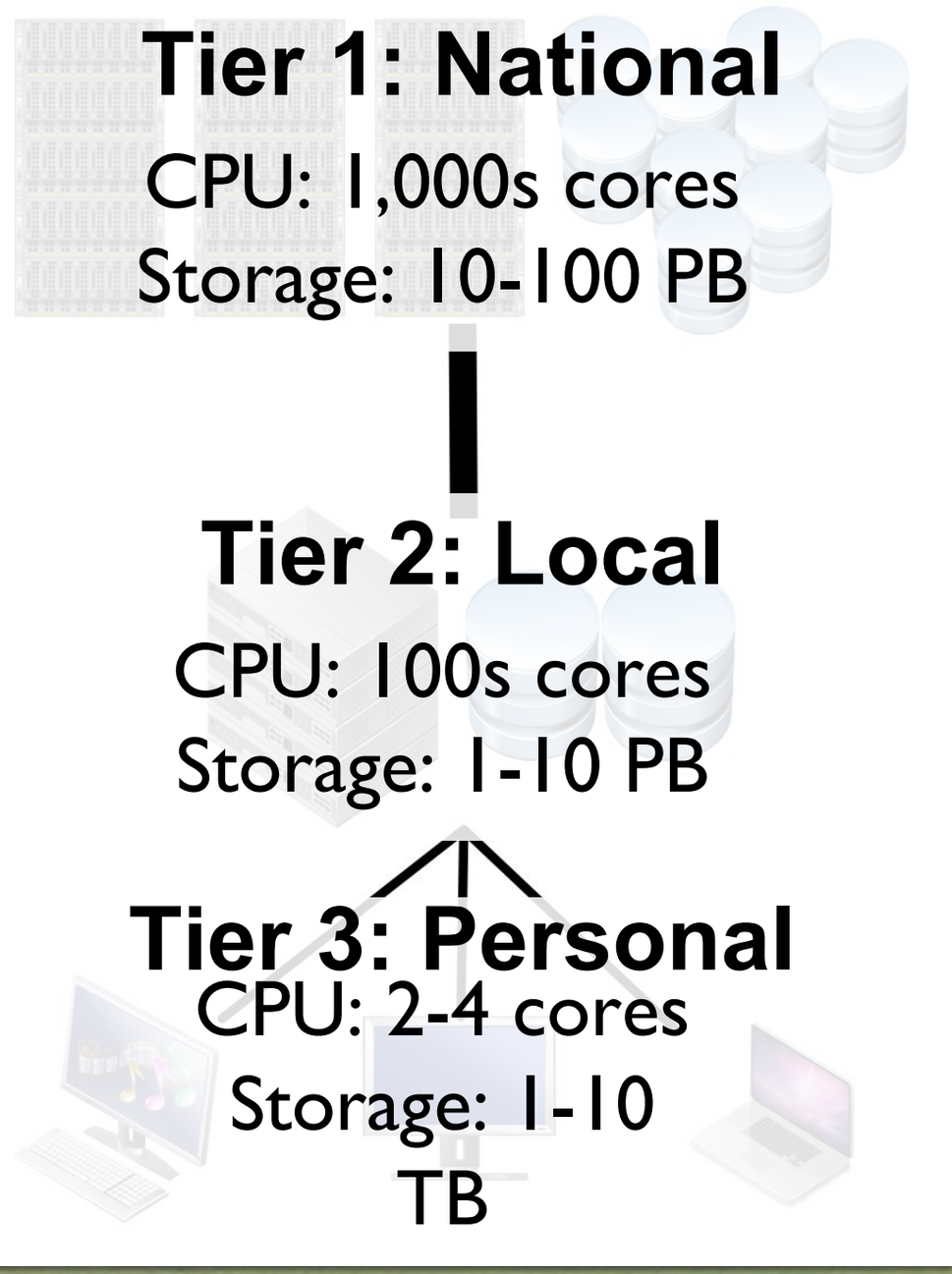
Philosophy behind Cirrus

Goal: save researchers time and effort by simplifying IT support and insulating them from external hassles while still giving them full control over their machines

- Provide something as close to "buy-your-own" server as possible with many of the advantages of the cloud
- Cirrus is governed by the Faculty IT Oversight Committee which contains a researcher from each department
- Allows us to negotiate as a Faculty when dealing with future changes to University IT policies

“Tier-2” Computing

- Not designed to replace large scale Compute Canada facilities
  - More than desktop less than CC clusters
- Management reflects this
  - Hardware looked after by IST in ECHA data centre
  - SRIT can administer VMs and operating systems on VMs
  - OpenStack managed by Compute Canada team in Sherbrooke
- Will eventually allow jobs to burst out onto the CC clouds and return data back to Cirrus



Cirrus

- Cirrus is a local, faculty controlled cloud built for research use (a Tier 2 type facility)
  - Cluster infrastructure purchased by faculty
    - Controller nodes, network switches etc.
  - Resources for research paid for by researcher
    - Only buy the cores, memory and storage which you need
    - Only pay once
    - You own what you paid for for the hardware lifetime and are the exclusive user of that share unless you choose to share your resources
- Purchasing simple: speed code transfer
  - No need for RFPs, quotes, specifications etc.
  - Cost designed to be very competitive with a standalone server purchase…but with faster network, RAID storage etc.

Management

- All the hardware lives in the ECHA data centre and is managed by IST
  - Warranty replacements, installs of new machines etc.
- Cirrus runs an OpenStack installation which is managed by the Compute Canada cloud team in USherbrooke
  - Same configuration (as far as possible) as the new Compute Canada clouds
  - Uses your Compute Canada CCDB account name for login
- Future plan is to integrate with the CC clouds
  - Allow researchers to setup and test jobs on Cirrus and then burst out onto the far larger CC clouds
  - Will also allow output to be copied back to Cirrus for local analysis of results
- Note: not just limited to CC clouds, will investigate Amazon, Cybera and possibly even IST VMware cloud integration based on demand

Cloud Advantages

- Hardware configurable through a webpage
  - It is easy to reconfigure a cluster of 10 single core machines with 8GB of memory each into a single machine with 10 cores and 80GB of memory
  - Current limit is 10 cores and 128GB memory
    - Hardware supports a maximum of 24 cores and 360GB but allocation would be hard but if 10 cores are not enough talk to your local ITOC rep and we can see what we can do
- Software installation hugely simplified
  - Once you setup one machine with all the software you need all it takes is a few mouse clicks to clone that machine as many times as you have resources to run it
- If resources are not being used by you then it is easy to share them with others
  - However sharing will be limited at first since the top priority is to ensure that the resources you own are always available for your use
  - We need to have data on average resource use before we can increase sharing
    - Once we have this data the number of cores in your allocation will likely increase
    - Cannot do the same (over allocate) with memory or disk storage

Cloud Advantages

- Cirrus is built with high speed 10Gb network links
  - File and data transfer speeds ~10 times faster than 1Gb links typically seen in many research clusters and with your desktop
- Bulk storage can be configured with backups
  - Snapshots of the file system taken automatically at regular intervals
  - Like time machine on a Mac: all snapshots accessed as directories and files can be copied from them into the current file system by users (…but through a CLI not a fancy GUI!)
- Hardware maintenance entirely taken care of
  - All warranty replacements are done centrally: no action is needed by researchers
  - …and because a virtual machine can run on any node in the cloud if one node goes down due to a fault you can immediately restart your machine on another node while repairs are taking place

Cirrus vs. Own Servers

	Cirrus	Own Servers
Purchase	Select the CPU and storage you want; pay by speed code transfer; resources usually instantly there.	Come up with specifications; get quotes; purchase hardware; wait 2-3+ weeks for delivery; arrange installation.
Repair	**Do nothing! No delay, no effort!** Your VM can run on a different node while IST fixes the issue.	Diagnose problem; call vendor and arrange a replacement part; wait for new part; install; unable to use machine until done
Software Installation	Install on one machine and clone; some images with existing OS+software provided	Install OS and software on each machine by hand; configure accounts, network each time
Sharing	Same physical resource can run any virtual machine; resources can be shared and hardware reconfigured with a few clicks	Hard/impossible to share unless other researchers use the same OS and software and the same hardware specifications

Cirrus "Share" Prices

- Researchers purchase shares for lifetime of the equipment
  - 3 years guaranteed under warranty
  - 3-5 years "best effort": cannibalize machines for parts to keep others up and running
    - some possible resource degradation
    - Faculty **may** help with some repairs
  - 5+ years: run for as long as is reasonable
    - i.e. until failure or until cost (financial or time) becomes unsustainable

Resource Share Type	Price
CPU unit (2 vCPUs + 15GB mem)	$650
High Perform. Storage (100GB)	$80
Bulk Storage (1TB)	$125
GPU slot	$450 (see next slide)

Cirrus GPU "Share" Prices

- GPU pricing is constrained by the number of slots available that can hold GPU cards
- The T630 host machines can accommodate upto 4 GPUs
- The cost of the host machine is recovered in the cost of the CPU units it has available, RAM available and then usage of GPU slots.
- It has been determined that a GPU slot will cost $450
- In addition, you need, at least, two CPU units (4 vCPUs) to qualify for a GPU slot
- And you have to purchase the GPU card as well
- However, if you purchase 6 CPU units (12 vCPUS) you get one GPU slot for free, and you can purchase 2 more GPU slots
- See next slide for pricing of Cirrus configurations with GPUs

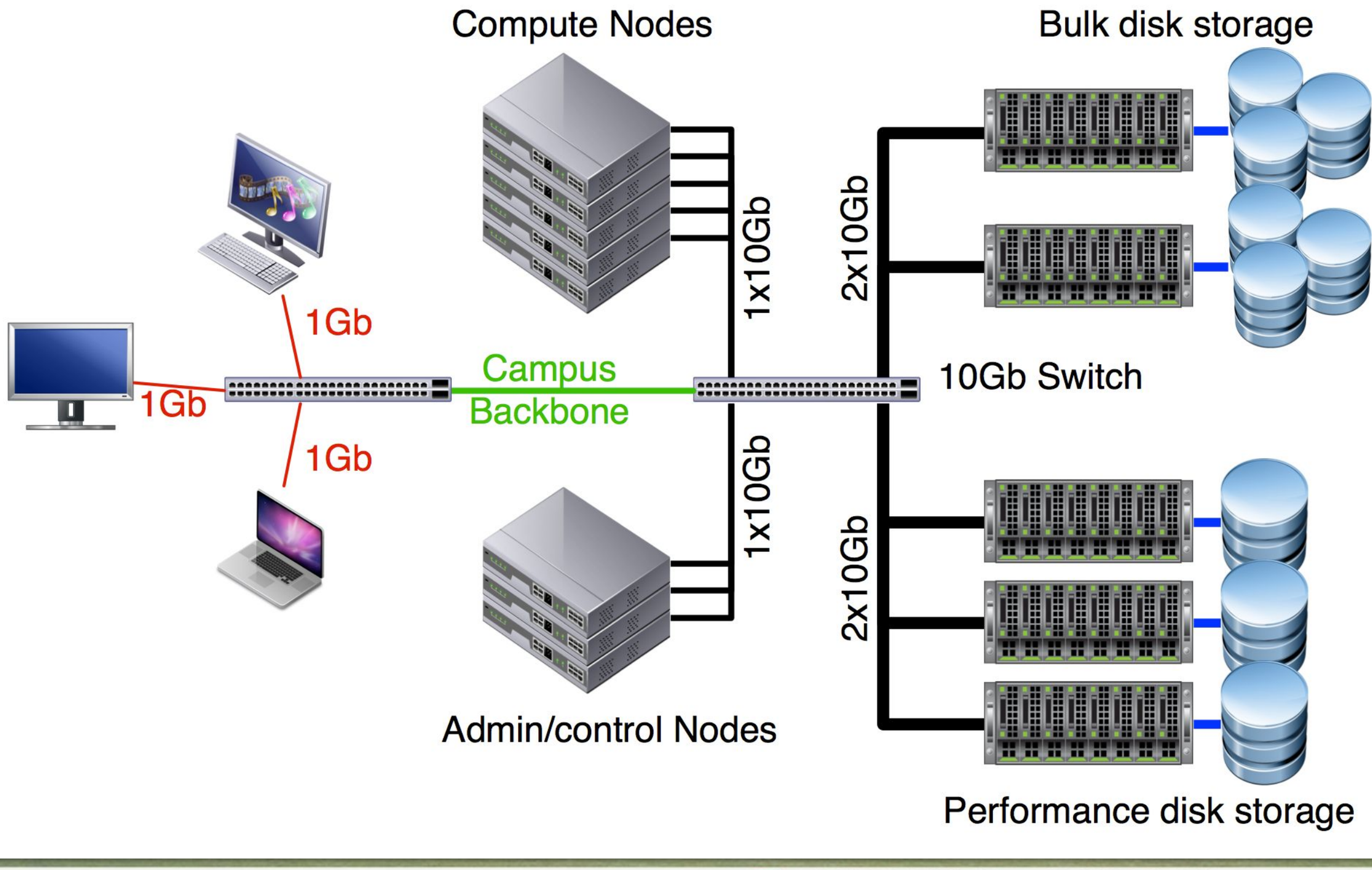
Cirrus Example Costs

- Minimal configuration - 1 CPU unit + 1 high performance storage unit - $550 + $80 = $630
  - usually want at least 1TB of bulk storage so $630 + $125 is $755
- Minimal configuration with a single GPU - 2 CPU units + 1 high performance storage unit + 1 GPU slot + price of GPU card = $1100 + $80 + $450 + price of GPU card = $1630 + price of GPU card
  - Add 1 TB of scratch storage - $1755 + price of GPU card
- 6 CPU units + 3 high performance storage units + 3 GPU slots + price of GPU cards = $3300 + $240 + 1 free GPU slot + $900 for two extra GPU slots = $4440 + price of 3 GPU cards
- SRIT purchased 6 CPU units, 3 high performance storage units and 8 TB of bulk storage (4 TB of scratch and 2TB of "work" storage) prior to price change on CPU units
  - 6 * $650 + 3 * $80 + 8 * $125 = $5140

Cirrus Gotchas

- Windows Licensing
  - If volume licensing scheme changes will not be able to spin up older images that used previous volume licensing scheme
- Previous lack of IT support
  - Independent researchers want to concentrate on their research not on getting virtual machines to work for them.
  - SRIT is now available to help these researchers use Cirrus
- File sharing outside of Cirrus
  - Limited to webdavfs and sshfs now, no SMB shares or NFS
  - Many researchers want their research data to be accessible outside of Cirrus, so this limitation puts the brakes on many other usages of Cirrus
- How do you use spare cycles of other projects who want to share?
  - Currently there is no "easy" way to share resources between projects
  - SRIT will be looking into this in 2018 to see what can be done

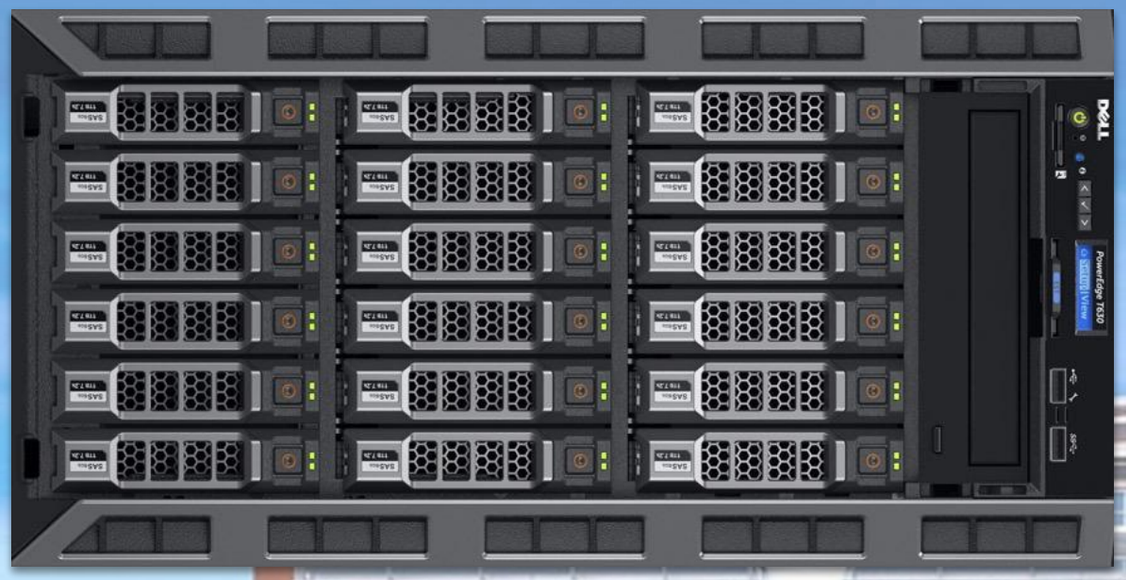
Cirrus Design



Technical Specs

- Compute: 408 cores (816 vCPUs) with 5 TB RAM
  - 5 Dell PowerEdge R730 with dual, 12 core Intel Xeon E5-2600 v3 CPUs and 320GB RAM - 1PE R730 with 2 GPUs
  - 36 DELL T630 with dual, 4 core Intel Xeon E5-2650 v4 2.2GHz CPUs and 128GB RAM - 4 PCIe slots per T630 for GPUs - 48 GPUs
  - 2 PowerEdge R330 - 2x16GB Memory-2 x 240 GB SSD - 2x480 GB SSD (LDAP Servers)
- Performance Storage:
  - 3 Dell PowerEdge R730xd with 2x400GB SSD and 10x6TB HDDs, 2x10Gb net
  - Runs Ceph with data stored in triplicate
- Bulk Storage: 1.07 PB of usable storage
  - 2 Storinator 60 storage units with 30x6TB WD Red HDDs, dual 10Gb network links
  - 1 DS7500 with 90x10TB HDDs, multiple 10Gb network links
  - Runs FreeNAS with disks arranged in 6-disk RAID Z2 and 10-disk RAID-Z3 arrangements under ZFS
    - 3 or 4 disks can fail before data is lost in these two types of RAID configurations

GPUs



- GPUs are supported
  - NVidia K2A, 1080Ti, Titan, Tesla,
  - AMD GPUs have not been tested yet
- If you are interested in using GPUs talk to your ITOC rep
- You can purchase GPUs and put them into Cirrus but may have to purchase GPU slot (see pricing slide)
- Currently 14 uninhabited GPU "slots" in host servers

Existing Projects

- Astrophysics - G. Sivakoff
  - 30 CPU units, 22 Ceph storage units, 80 bulk storage units
- Mauricio Sacchi
  - 25 CPU units, 10 Ceph storage units
- Erin Bayne
  - 18 CPU units, 40 Ceph storage units, 325 bulk storage units
- IceCube - Claudio Kopper
  - 298 CPU units, 100 Ceph storage units, 105 bulk storage units, 48 GPUs
- 14 other projects of various sizes

Interested?

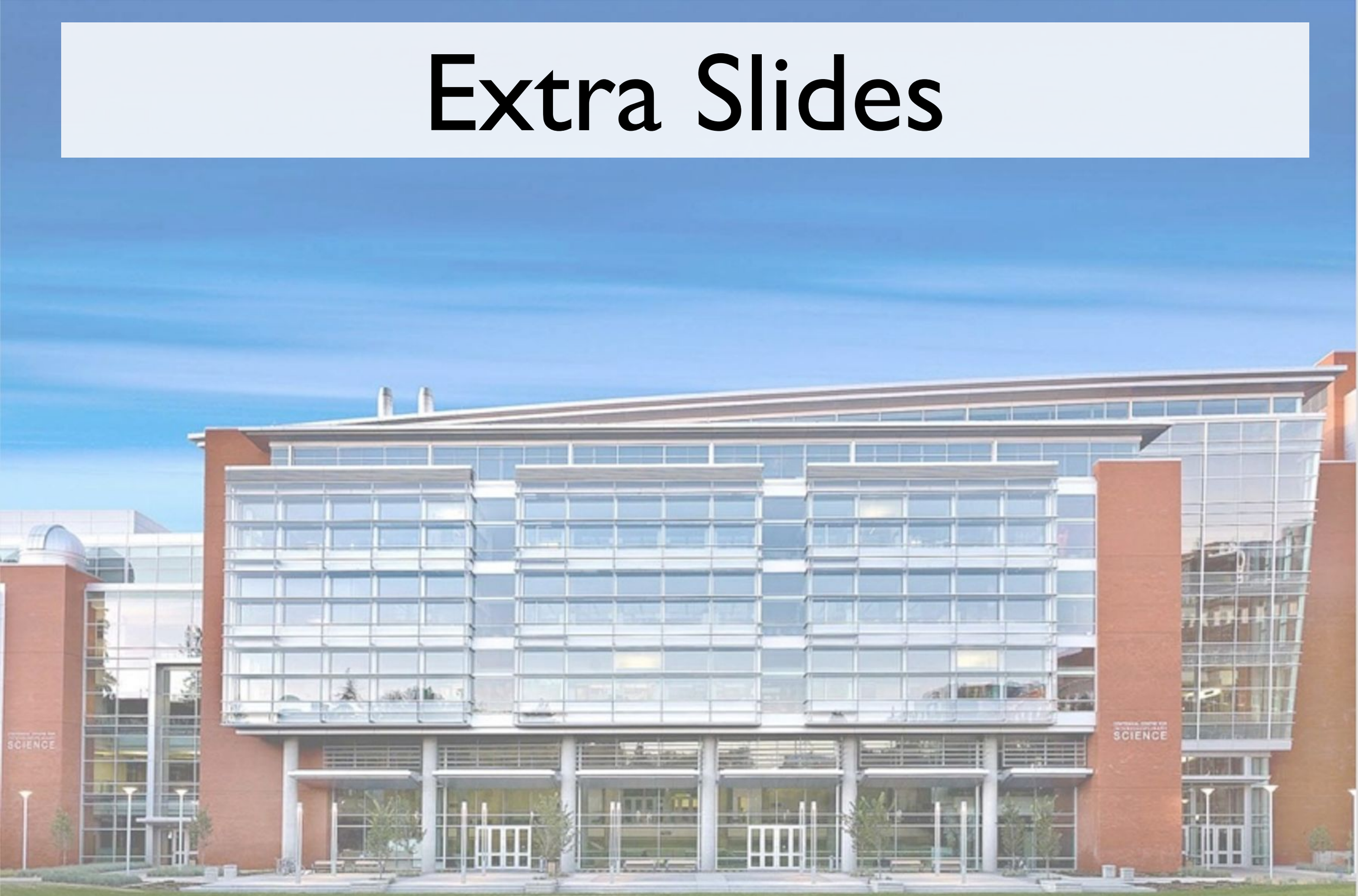
- If Cirrus sounds like something you would be interested in using talk to your department ITOC rep
  - Temporary use of a VM or two can be provided if you want to try the system out first
- If you have special computing needs that you don't think we currently meet talk to your ITOC rep
  - If others have the same requirements we can look at adding them to Cirrus
  - …but designed primarily for research not fixed-scheduled teaching classes
    - Design uses "cheap", commodity hardware/software not "enterprise-class": if a critical component fails it may take a few hours to fix (<1 day but >1 lecture)

ITOC Representatives

- Bio-science: Erin Bayne
- Chemistry: Mariusz Klobukowski (Chair)
- CompSci: Nilanjan Ray
- EAS: Paul Meyers, Mary-Jane Turnell
- Maths: George Peschke
- Physics: Roger Moore
- Psychology: Tom Johnson
- Faculty: Kerrie Johnston, Carol Smith, Rod Johnson

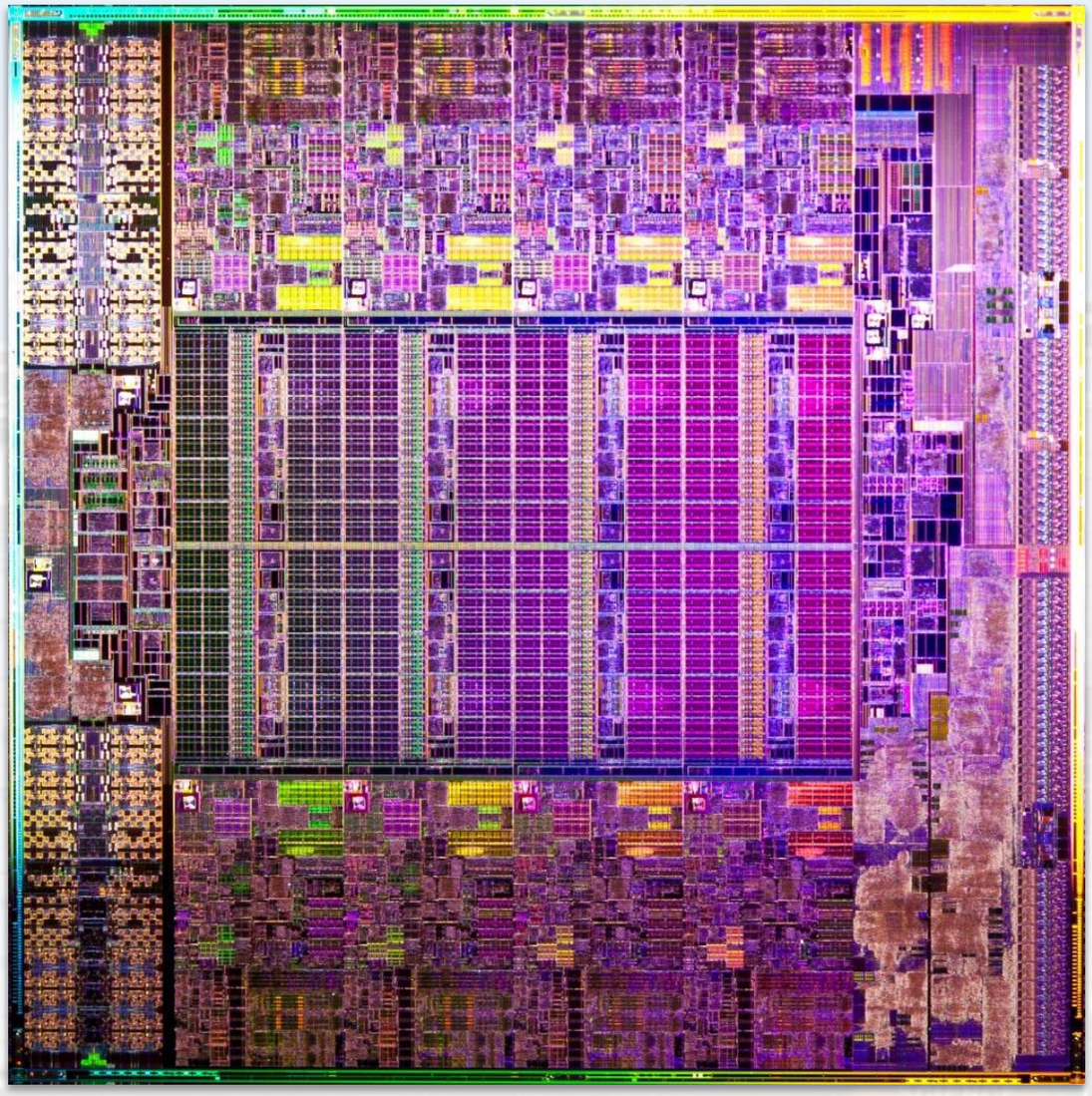
The End

Extra Slides



Modern CPUs

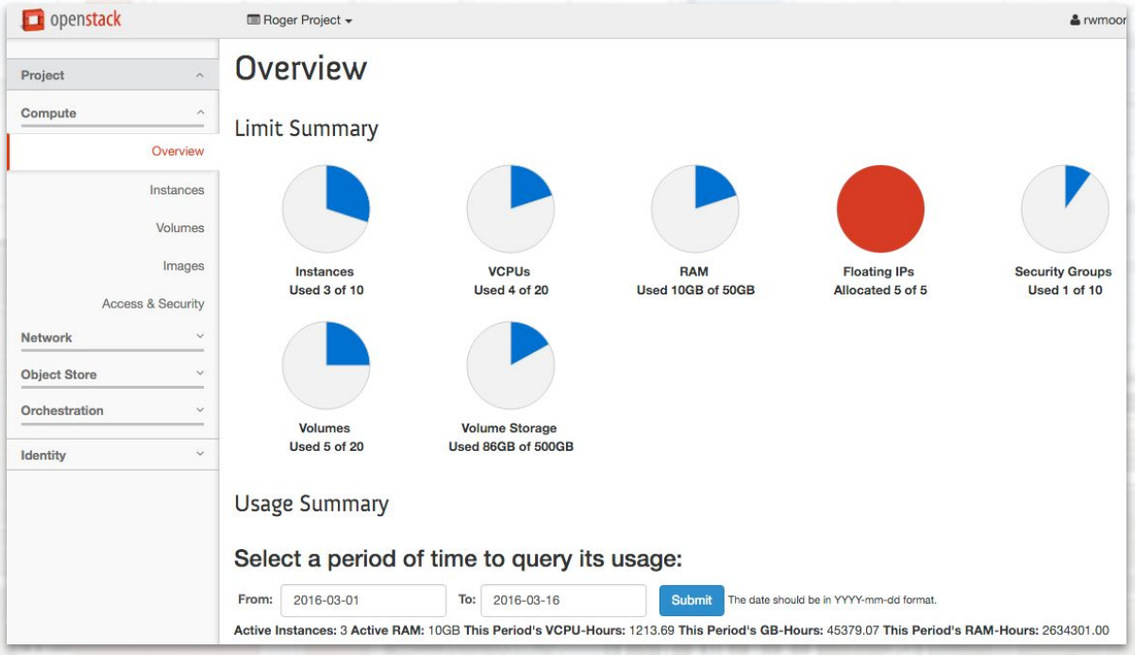
- Modern CPUs contain multiple CPU cores
  - 1 core = 1 independent processing unit
  - Hyper-threading provides an additional performance boost
- CPU growth now in cores and capability not clock frequency
  - 10+ cores per CPU not uncommon
- Support for "virtualization"
  - Enables a single CPU to be split into multiple, independent virtual machines



8-core Intel Xeon

Cirrus: 2x10 core Xeons

What is Cloud Computing?



- Cloud computing is a physical cluster which runs virtual machines
- Unlike a typical cluster there is no one operating system
  - Each virtual machine has its own operating system
    - Mix Windows, Linux; not OS X (license requires Apple hardware)
- Virtual machines and be created and reconfigured by a web interface
  - CPUs, memory or disk can be added or removed by the click of a mouse