Cloud Computing: Is It Powerful, And Can It Lead To Bankruptcy?

G. Bruce Berriman
gbb@ipac.caltech.edu
http://astrocompute.wordpress.com
Infrared Processing and Analysis Center, Caltech

Space Telescope Science Institute, February 2012.

Collaborators

Ewa Deelman

Gideon Juve

Mats Rynge

Jens-S Völcker

Information Sciences Institute, USC

Cloud Computing Is Not A New Idea!

- Rewind to the early1960's ….
- John McCarthy … "computation delivered as a public utility in…. the same way as water and power."

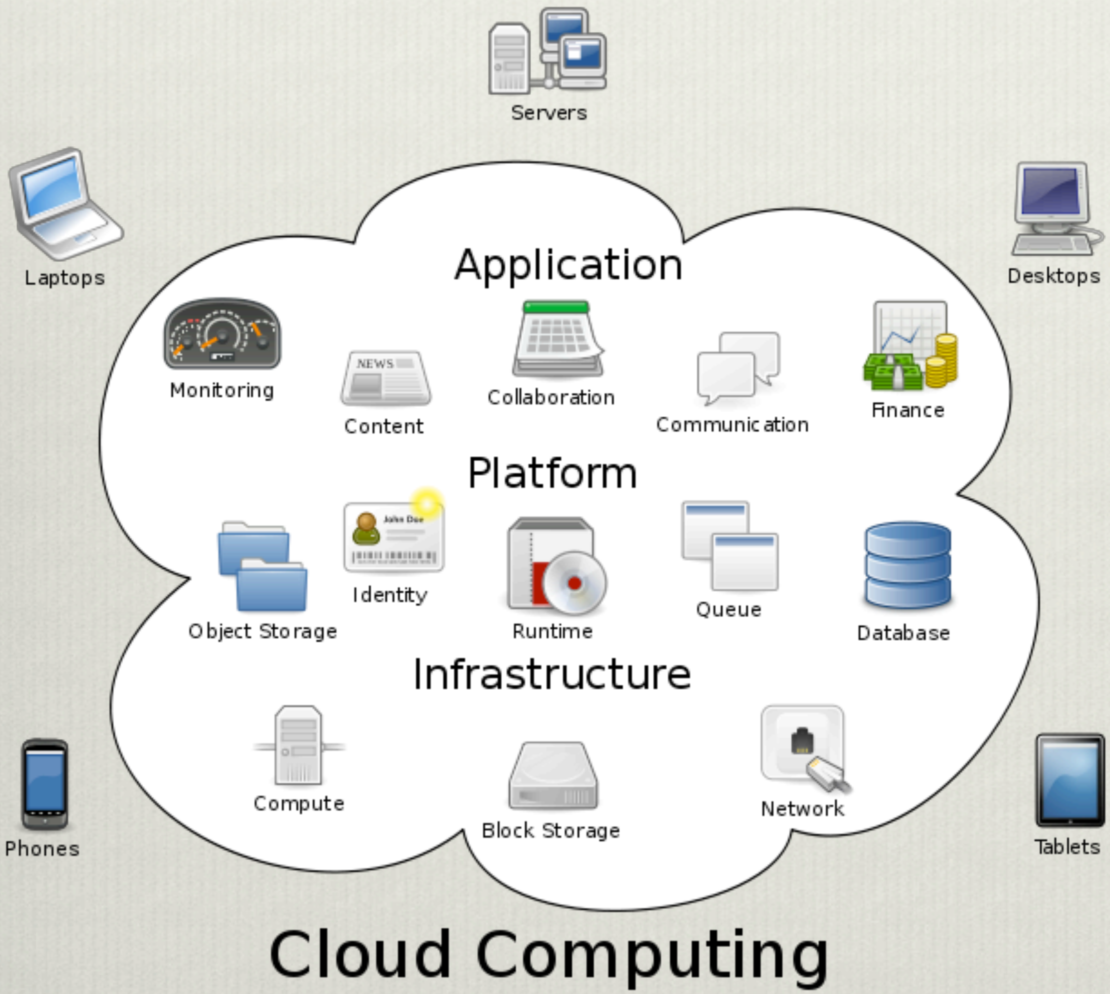
- J. C. R. Licklider … "the intergalactic computer network"

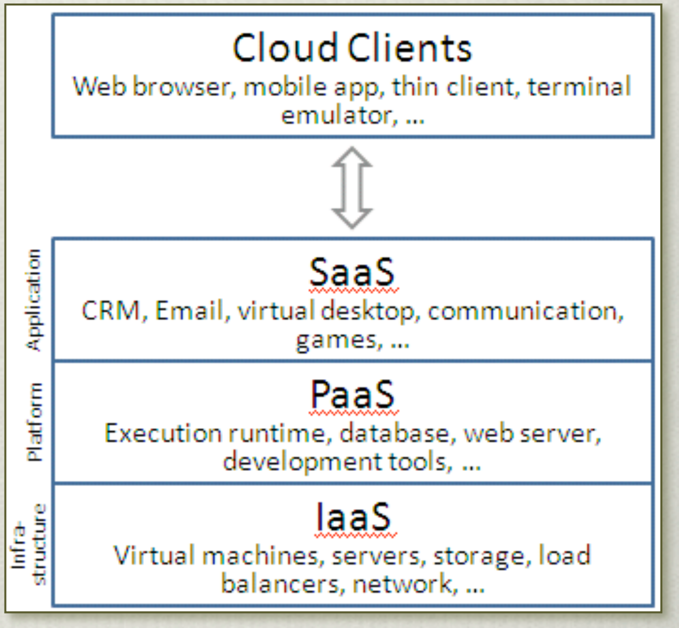
  "It seems to me to be interesting and important to develop a capability for **integrated network operation** … we would have …perhaps six or eight **small** computers, and a great assortment of disc files and magnetic tape units … all churning away"

The Idea Was Dormant For 35 Years

- … until the internet started to offer significant bandwidth.
- Salesforce.com, Amazon Web Services started to offer applications over the internet (1999-2002).
- Amazon Elastic Compute cloud (2006) offered first widely accessible on-demand computing.
- By 2009, browser based apps such as Google Apps had hit their stride.

Cloud Service Layers

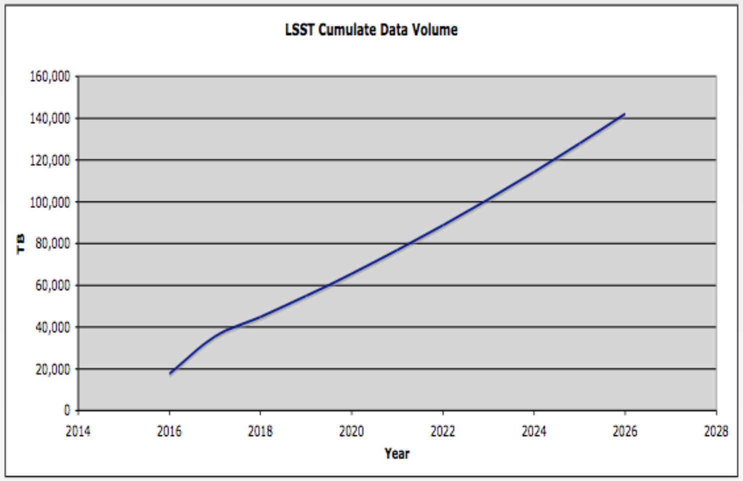


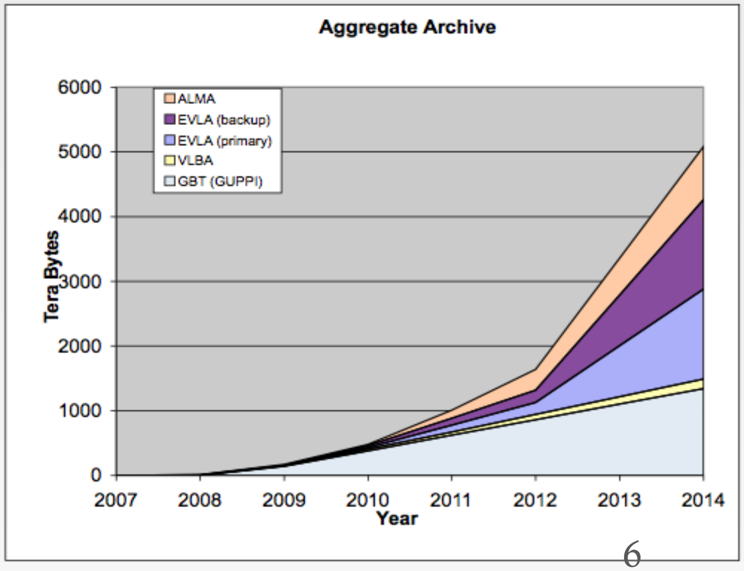


http://en.wikipedia.org/wiki/Cloud_computing

Developing A New Business Model For Astronomical Computing

- Astronomy is already a data intensive science
  - Over 1 PB served electronically through data centers and archives.
  - By 2020, as much as 60-120 PB on-line .
- Astro2010 recognized the need for demand high performance computing on massive, distributed data sets.





How Useful Is Cloud Computing For Scientific Workflow Applications?

- Loosely-coupled parallel applications
  - Many domains: astronomy, biology, earth science.
  - Potentially very large: 10 K tasks common, >1M not uncommon
  - Potentially data-intensive: 10 GB common, >1TB not uncommon
- Data communicated via files
  - Shared storage system, or network transfers

1. What is performance/cost of different cloud resource configurations for workflow applications?
2. What is performance of cloud and workflow applications?
3. What is virtualization overhead on cloud?

Getting Started With Cloud Computing

All you need is a credit card and connect to http://aws.amazon.com/ec2/

This looks cheap!

Amazon EC2 front page:

Region: US East (Virginia)

	Linux/UNIX Usage	Windows Usage
Standard On-Demand Instances		
Small (Default)	$0.085 per hour	$0.12 per hour
Large	$0.34 per hour	$0.48 per hour
Extra Large	$0.68 per hour	$0.96 per hour
Micro On-Demand Instances		
Micro	$0.02 per hour	$0.03 per hour
Hi-Memory On-Demand Instances		
Extra Large	$0.50 per hour	$0.62 per hour
Double Extra Large	$1.00 per hour	$1.24 per hour
Quadruple Extra Large	$2.00 per hour	$2.48 per hour
Hi-CPU On-Demand Instances		
Medium	$0.17 per hour	$0.29 per hour
Extra Large	$0.68 per hour	$1.16 per hour
Cluster Compute Instances		
Quadruple Extra Large	$1.60 per hour	N/A*
Cluster GPU Instances		
Quadruple Extra Large	$2.10 per hour	N/A*

* Windows® is not currently available for Cluster Compute or Cluster GPU Instances

Commercial Providers
Amazon.com EC2
AT&T Synaptic Hosting
GNi Dedicated Hosting
IBM Computing on Demand
Rackspace Cloud Servers
Savvis Open Cloud
ServePath GoGrid
Skytap Virtual Lab
3Tera
Unisys Secure
Verizon Computing
Zimory Gateway
Science Clouds
FutureGrid
NERSC Magellan
NASA Nebula

"Little sins add up …"

OS	EC2 Instace	Demand Type	Cost / Hr	Hours	Length	Total
Windows	HCPU Extra Large	OnDemand	$1.16	8,736	Year	$10,133.76
Windows	Extra Large	OnDemand	$0.96	8,736	Year	$8,386.56
Linux/UNIX	Extra Large	OnDemand	$0.68	8,736	Year	$5,940.48
Linux/UNIX	HCPU Extra Large	OnDemand	$0.68	8,736	Year	$5,940.48
Linux/UNIX	Large	OnDemand	$0.68	8,736	Year	$5,940.48
Windows	HCPU Extra Large	Reserved	$0.50	8,736	Year	$4,368.00
Windows	Large	OnDemand	$0.48	8,736	Year	$4,193.28
Windows	HCPU Medium	OnDemand	$0.29	8,736	Year	$2,533.44
Linux/UNIX	Extra Large	Reserved	$0.24	8,736	Year	$2,096.64
Linux/UNIX	HCPU Extra Large	Reserved	$0.24	8,736	Year	$2,096.64
Linux/UNIX	HCPU Medium	OnDemand	$0.17	8,736	Year	$1,485.12
Linux/UNIX	Large	Reserved	$0.12	8,736	Year	$1,048.32
Windows	Small	OnDemand	$0.12	8,736	Year	$1,048.32

Instance	Memory (MB)	Virtual Core	ECU	ECU per Core	Storage (GB)	I/O	Platform
Micro Instance	633	1	2	-			32/64 bit
Small Instance - default	1740.8	1	1	1	160	Moderate	32 bit
Large Instance	7680	2	4	2	850	High	64 bit
Extra Large Instance	15360	4	8	2	1690		64 bit
High-Memory Extra Large Instance	17510.4	2	6.5	3.25	420	Moderate	64 bit
High-Memory Double Extra Large Instance	35020.8	4	13	3.25	850	High	
High-Memory Quadruple Extra Large Instance	70041.6	8	26	3.25	1690		64 bit
High-CPU Extra Large Instance	7168	8	20	2.5	1690	High	64 bit
Cluster Compute Quadruple Extra Large Instance*	18432	8	33.5	4.1875	1690	Very Large	64 bit
Cluster GPU Quadruple Extra Large Instance **	18432	8	33.5	4.1875	1690	Very Large	64 bit

Annual Costs!

… and that's not all. You pay for:

-Transferring data into the cloud
-Transferring them back out again
-Storage while you are processing (or sitting idle)
-Storage of the VM and your own software
-Special services: virtual private cloud…

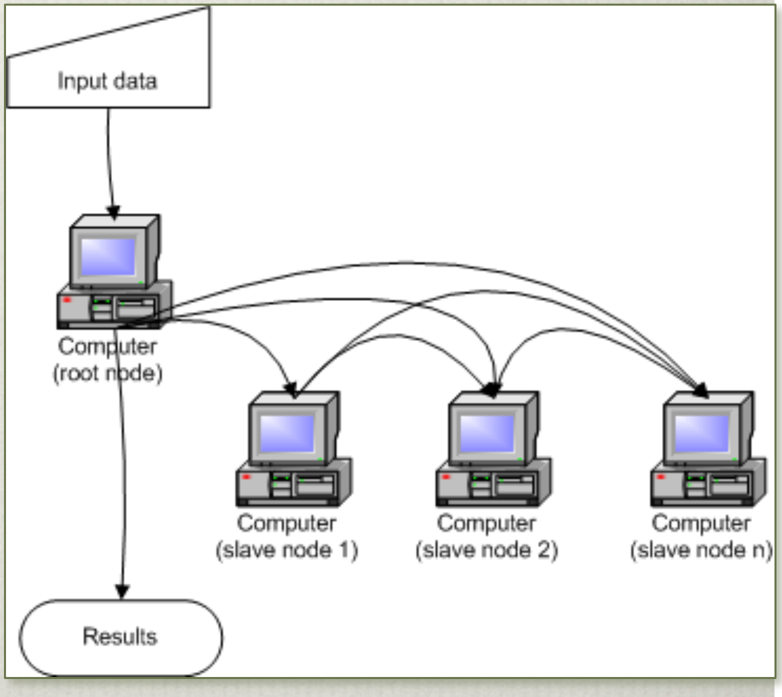
See Manav Gupta's blog post *http://manavg.wordpress.com/2010/12/01/amazon-ec2-costs-a-reality-check/*

What Amazon EC2 Does

- Creates as many independent virtual machines as you wish.
- Reserves the storage space you need.
- Gives you a refund if their equipment fails.
- Bills you.

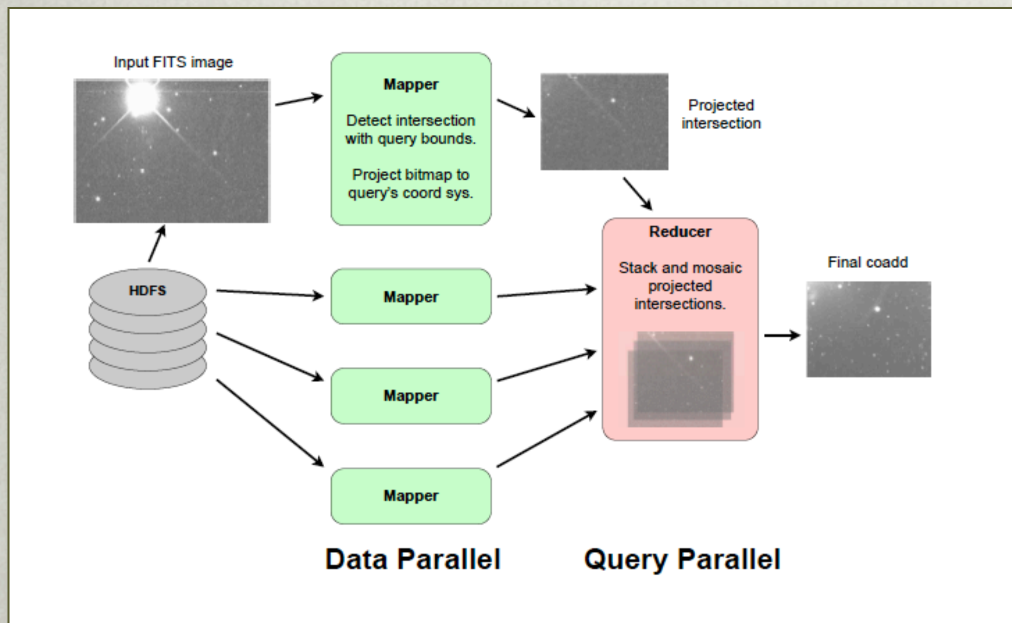
What You Have To Do

- Configure the virtual machines and create your environment
- Load all your software and input data
- Manage and maintain
- Working at scale:
  - Adapt applications to new computing models in cloud (e.g. MapReduce) or
  - Adapt the cloud to recreate the environment in which the app has run - **a virtual cluster.**



Computational Models: MapReduce

- Wiley et al (2011) "Astronomy in the Cloud: Using MapReduce for Image Co-Addition" PASP, 123, 366.
- SDSS Image co-addition of 20 TB data
  - 100,000 files processed in 3 minutes on 400 cores
  - Considerable effort to adapt co-addition to Hadoop



Map: partitions input into smaller sub-problems, and distributes them to worker nodes.

Reduce: collects answers to sub-problems and combines them to form the output

Pegasus Workflow Management System

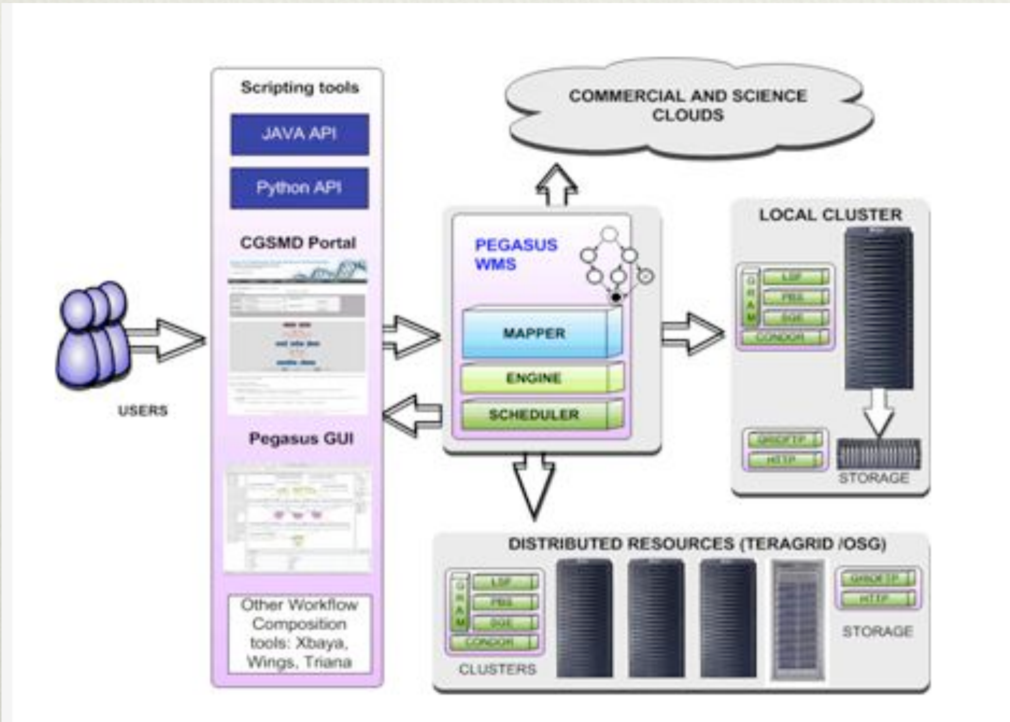
http://pegasus.isi.edu

- Converts abstract workflow to a concrete workflow
- No special requirements on infrastructure

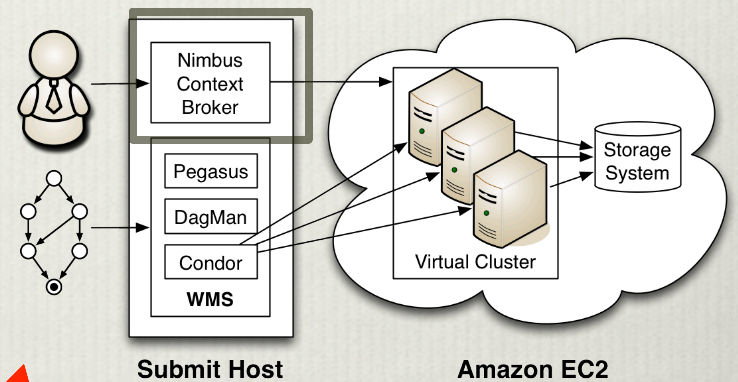
Pegasus	Maps tasks and data to executable resources; Performance optimizer
DAGMan	Workflow engine - Tracks dependencies, releases tasks, retries tasks
Condor Schedd	Task manager; schedules and dispatches tasks (and data) to resources

Cyberinfrastructure: Local machine, cluster, Condor pool, Grid, Cloud,…

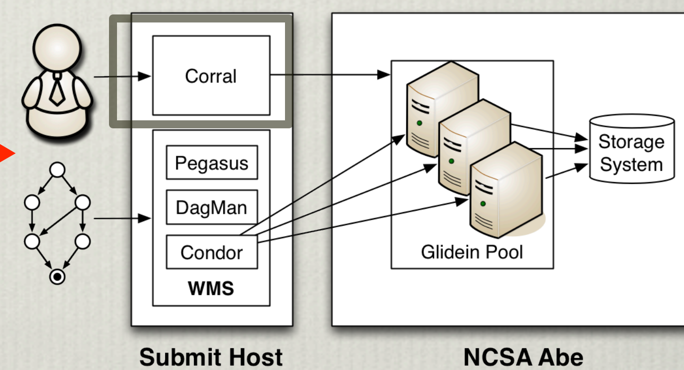
Comparing Clusters With Clouds



Amazon EC2



NCSA Abe - high-performance cluster.



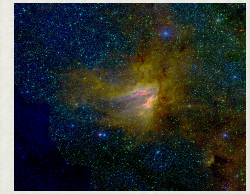
Set up computationally equivalent configurations on Amazon and Abe

Nimbus Context Broker – toolkit for configuring virtual clusters.

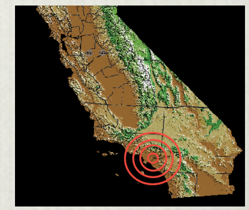
Corral – resource provisioning tool for grids.

The Applications

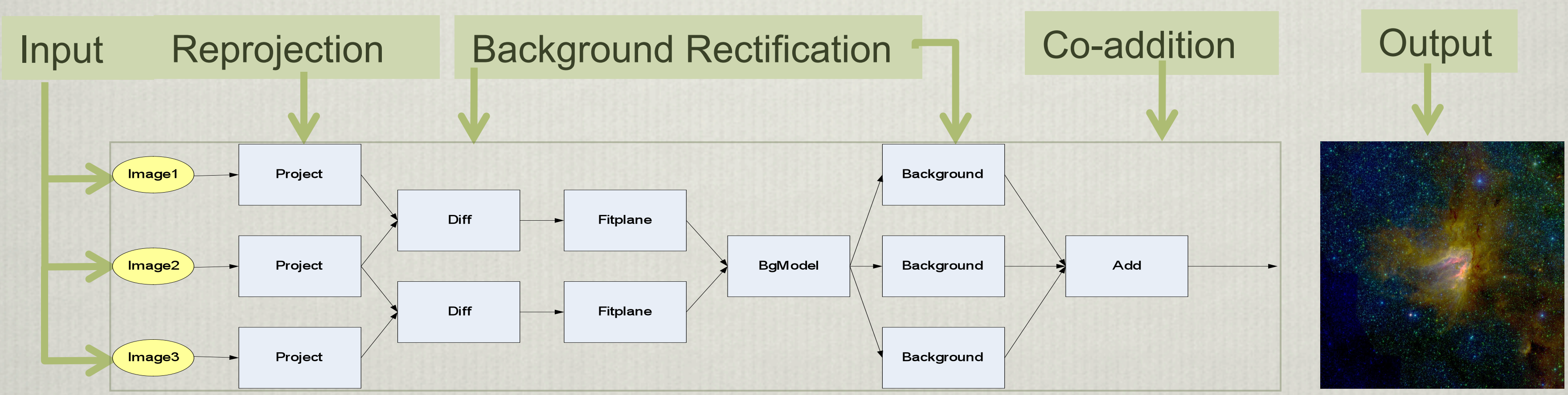
Montage **(http://montage.ipac.caltech.edu)** creates science-grade image mosaics from multiple input images.



Broadband calculates seismograms from simulated earthquakes.



Epigenome maps short DNA segments collected with gene sequencing machines to a reference genome.



Montage Workflow

Characteristics of Workflows

Workflow Specifications for this Study

Application	Workflow	# Tasks	Input	Output
Montage	8 deg. sq. mosaic of M16, 2MASS K-band	10,429	4.2 GB	7.9 GB
Broadband	4 earthquake sources, 5 sites	320	6 GB	160 MB
Epigenome	Maps DNA sequences to ref. chromosome 21	81	1.8 GB	300 MB

Resource Usage of the Three Workflow Applications

Application	I/O	Memory	CPU
Montage	High	Low	Low
Broadband	Medium	High	Medium
Epigenome	Low	Medium	High

Computing Resources

Type	Arch	CPU	Cores	Memory	Network	Storage	Price
Amazon EC2							
m1.small	32-bit	2.0-2.6 GHz Opteron	1-2	1.7 GB	1-Gbps Ethernet	Local	$0.10/hr
m1.large	64-bit	2.0-2.6 GHz Opteron	2	7.5 GB	1-Gbps Ethernet	Local	$0.40/hr
m1.xlarge	64-bit	2.0-2.6 GHz Opteron	4	15 GB	1-Gbps Ethernet	Local	$0.80/hr
c1.medium	32-bit	2.33-2.66 GHz Xeon	2	1.7 GB	1-Gbps Ethernet	Local	$0.20/hr
c1.xlarge	64-bit	2.0-2.66 GHz Xeon	8	7.5 GB	1-Gbps Ethernet	Local	$0.80/hr
Abe							
abe.local	64-bit	2.33 GHz Xeon	8	8 GB	10-Gbps InfiniBand	Local	…
abe.lustre	64-bit	2.33 GHz Xeon	8	8 GB	10-Gbps InfiniBand	Lustre	…

Processors and OS

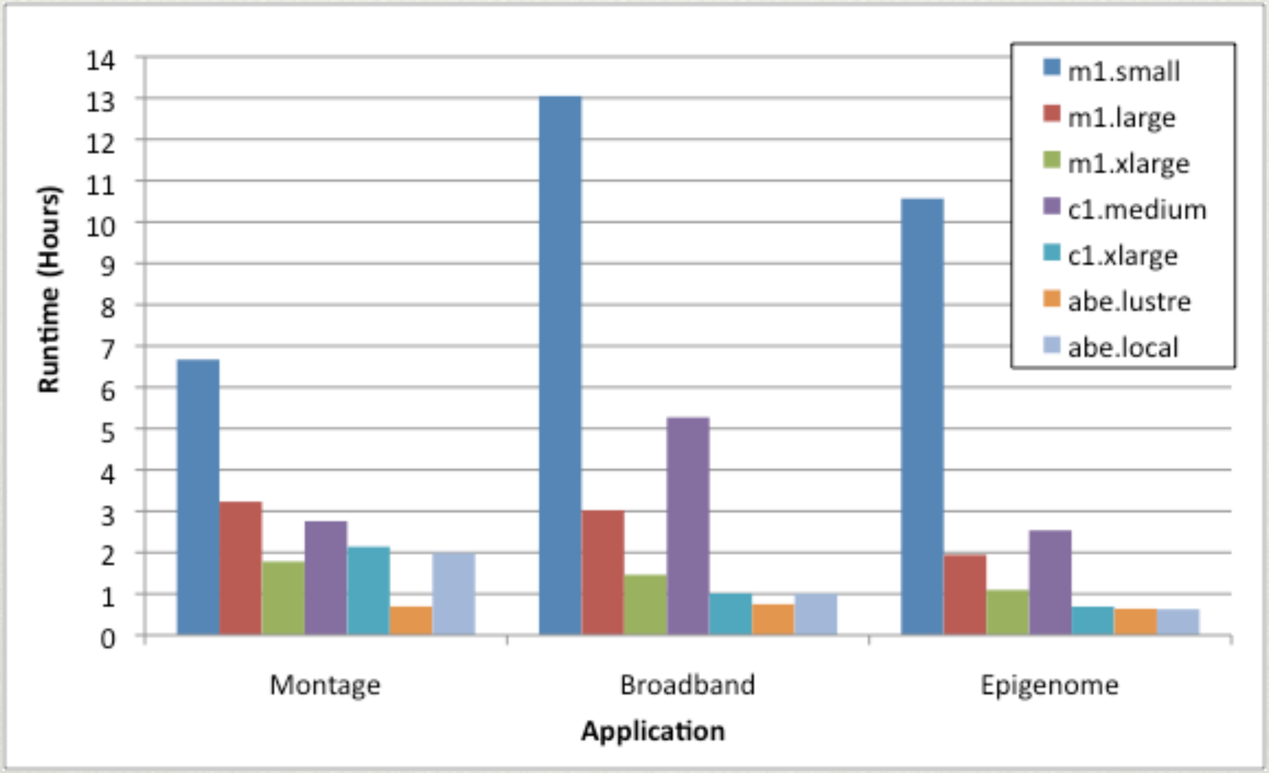
- Amazon offers wide selection of processors.
- Ran Linux Red Hat Enterprise with VMWare
- *c1.xlarge* and *abe.local* are equivalent – estimate overhead due to virtualization
- *abe.lustre* and *abe.local* differ only in file system

Networks and File Systems

- HPC systems use high-performance network and parallel file systems
- Amazon EC2 uses commodity hardware

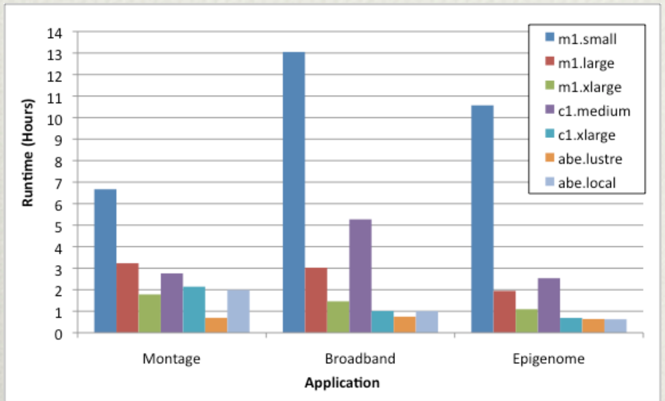
➔Ran all processes on single, multi-core nodes. Used local and parallel file system on Abe.

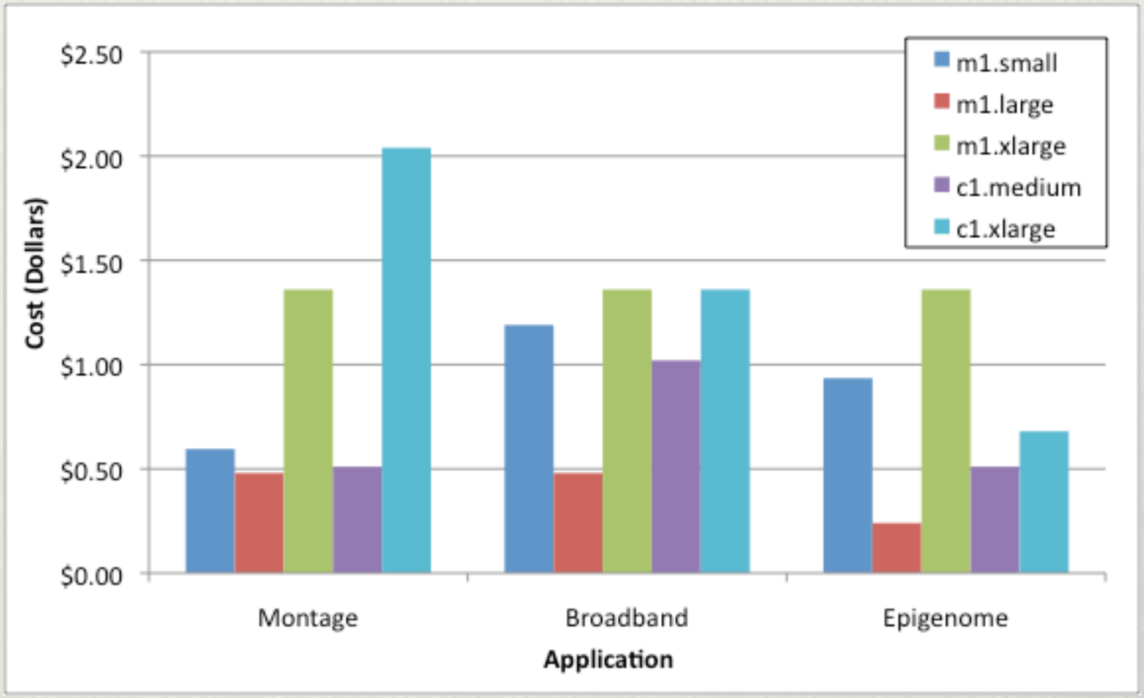
Performance Results



- Virtualization Overhead <10%
- Large differences in performance between the resources and between the applications
- The parallel file system on *abe.lustre* offers a big performance advantage of x3 for Montage

How Much Did It Cost?

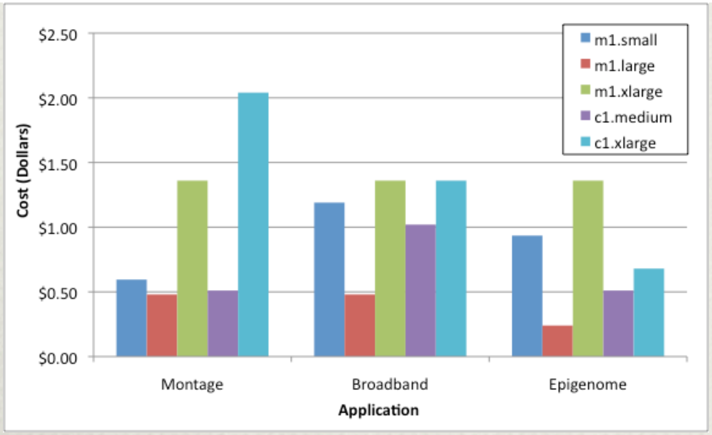




Instance	Cost $/hr
m1.small	0.10
m1.large	0.40
m1.xlarge	0.80
c1.medium	0.20
c1.xlarge	0.80

- Clear trade-off between performance and cost.
- **Most powerful processor *c1.xlarge* offers 3x the performance of *m1.small* – but at 4x the cost.**
- Most cost-effective processor for Montage is *c1.medium* – 20% performance loss over *m1.small*, but 4x lower cost.

Data Transfer Costs



Operation	Cost $/GB
Transfer In	0.10
Transfer Out	0.17

Transfer Rates

- ❖ Amazon charges different rates for transferring data into the cloud and back out again.
- ❖ Transfer-out costs are the higher of the two.

Application	Input (GB)	Output (GB)	Logs (MB)
Montage	4.2	7.9	40
Broadband	4.1	0.16	5.5
Epigenome	1.8	0.3	3.3

Application	Input	Output	Logs	Total
Montage	$0.42	$1.32	<$0.01	$1.75
Broadband	$0.40	$0.03	<$0.01	$0.43
Epigenome	$0.18	$0.05	<0.01	$0.23

Transfer Costs

- ❖ For Montage, the **cost to transfer data out of the cloud is higher** than monthly storage and processing costs.
- ❖ For Broadband and Epigenome, **processing incurs the biggest costs**.

Storage Costs

Data Storage Charges

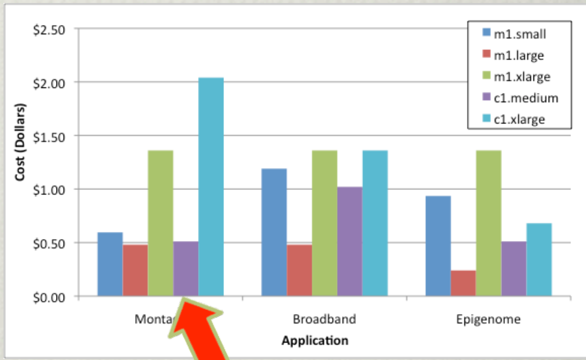
- Amazon charges for storing Virtual Machines (VM) and user's applications in local disk
- It also charges for storing data in persistent network-attached Elastic Block Storage (EBS).

Storage Rates

Item	Charges $
Storage of VM's in local Disk (S3)	0.15/GB-Month
Storage of data in EBS disk	0.10/GB-Month

Storage Volumes

Application	Input (GB)	Output (GB)	Logs (MB)
Montage	4.2	7.9	40
Broadband	4.1	0.16	5.5
Epigenome	1.8	0.3	3.3



Storage Costs

Application	Data ($)	VM ($)	Monthly Cost ($)
Montage	$0.95	$0.12	$1.07
Broadband	$0.02	$0.10	$0.12
Epigenome	$0.22	$0.10	$0.32

Montage Storage Costs Exceed Most Cost-Effective Processor Costs

The bottom line for Montage

Item	Best Value	Best Performance
	c1.medium	*c1.xlarge*
Transfer Data In	$ 0.42	$ 0.42
Processing	$ 0.55	$ 2.45
Storage/month	$ 1.07	$ 1.07
Transfer Out	$ 1.32	$ 1.32
Totals	$ 3.36	$ 5.26

4.5x the processor cost for 20% better performance

Cost-Effective Mosaic Service

Local Option

Item	Cost ($)
12 TB RAID 5 disk farm and enclosure (3 yr support)	12,000
Dell 2650 Xeon quad-core processor, 1 TB staging area	5,000
Power, cooling and administration	6,000
Total 3-year Cost	**23,000**
Cost per mosaic	**0.64**

-2MASS image data set
- 1,000 x 4 square degree mosaics/month

Amazon EBS Option

Item	Price ($)
Input Transfer (10 TB)	1,024.00
Output Transfer (24 TB)	3,691.41
Storage (10 TB)	36,864.00
Input I/O (10 TB)	24.58
Output I/O (27 TB)	67.50
Compute (c1.medium)	4,467.60
Total	**46,139.08**
Cost per mosaic	**1.28**

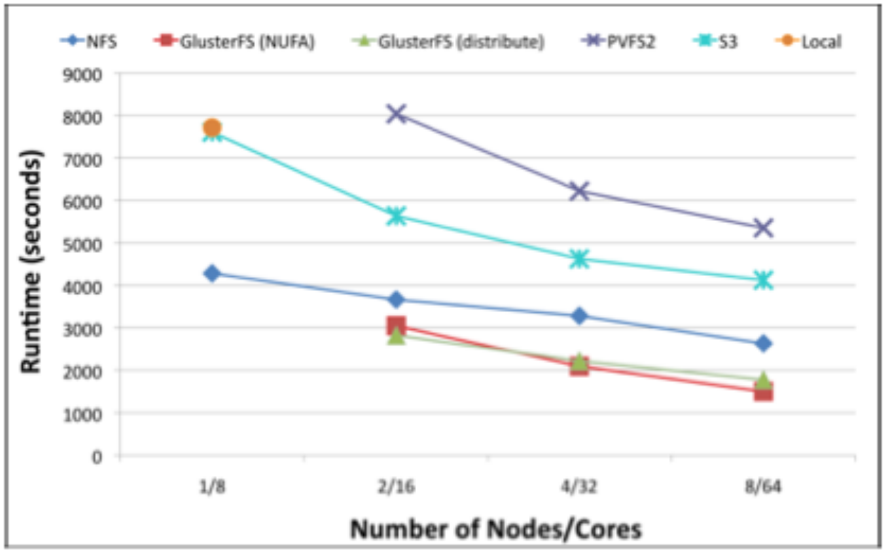
Amazon S3 Options

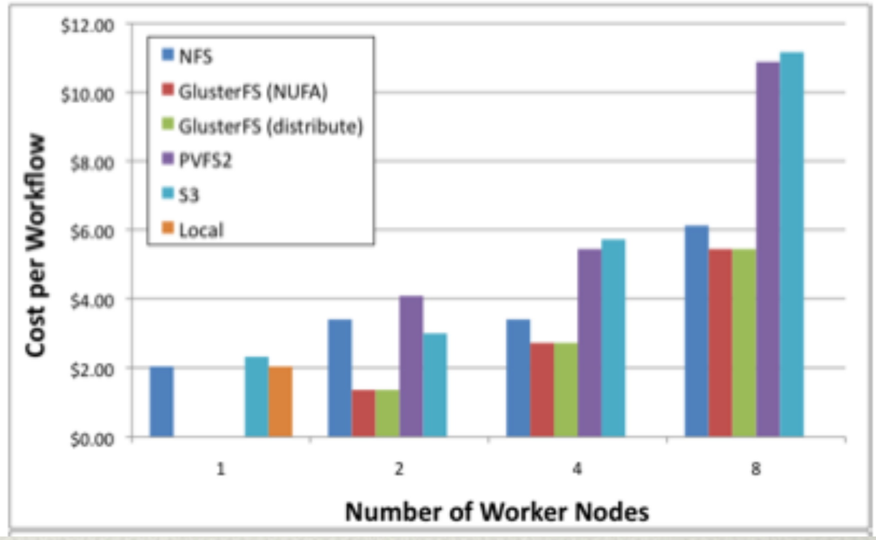
Item	Price ($)
Input Transfer (10 TB)	1,024.00
Output Transfer (24 TB)	3,691.41
PUT Ops (5.24 M)	52.43
GET Ops (14.4 M)	14.40
Compute (c1.medium)	4,467.60
Normal Storage (10 TB)	55,296.00
Total w/ Normal Storage	**64,545.84**
Cost per mosaic (Normal)	**1.79**
Reduced Redundancy Storage (10TB)	36,864.00
Total w/ RR Storage	**46,113.84**
Cost per mosaic (RR)	**1.28**

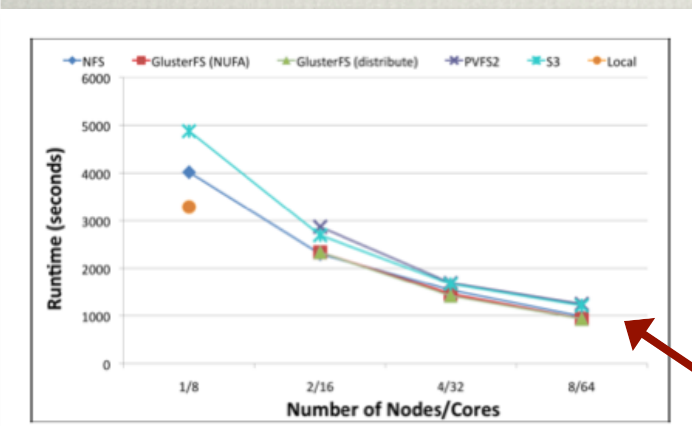
Amazon cost is 2X local!

Just To Keep It Interesting …

Running the Montage Workflow With Different File Storage Systems



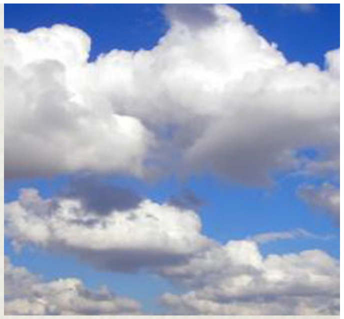




Cf. Epigenome

Cost and performance vary widely with different types of file storage depending on how storage architecture handles lots of small files

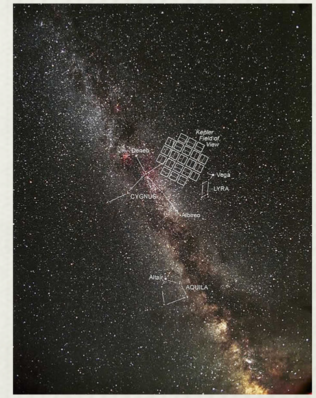
When Should I Use The Cloud?

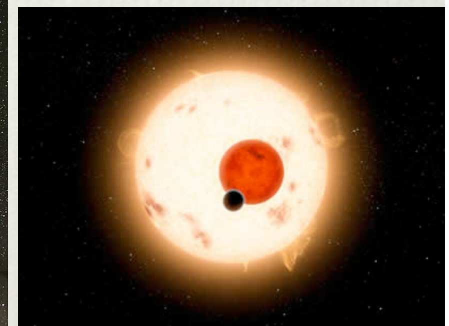


- **The answer is….it depends on your application and use case.**
- **Recommended best practice: Perform a cost-benefit analysis to identify the most cost-effective processing and data storage strategy. Tools to support this would be beneficial.**
- Amazon offers the best value
  - For compute- and memory-bound applications.
  - For one-time bulk-processing tasks, providing excess capacity under load, and running test-beds.
- Parallel file systems and high-speed networks offer the best performance for I/O-bound applications.
- Mass storage is **very** expensive on Amazon EC2

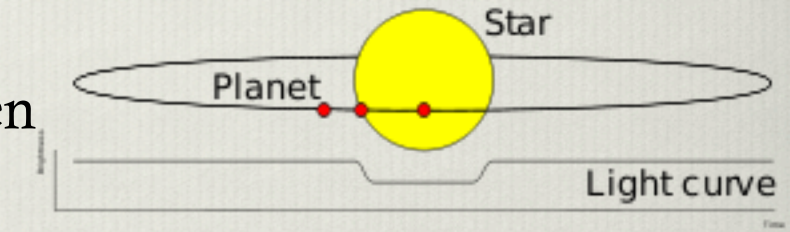
Hunting Exoplanets with Kepler

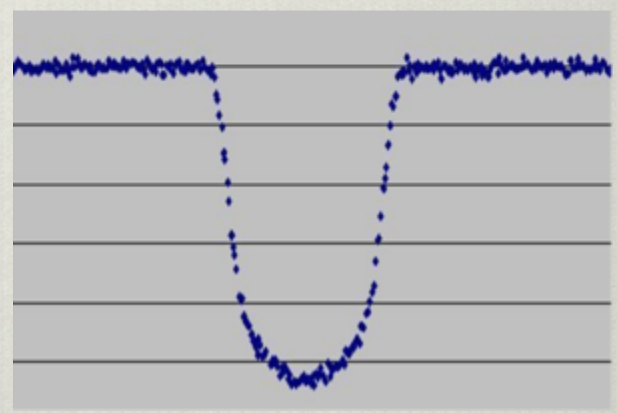
http://kepler.nasa.gov





- ❖ Kepler continuously monitors the brightness of over 175,000 stars.
  - ❖ Search for periodic dips in signals as Earth-like planets transit in front of host star.
- ❖ Currently 380,000 light curves have been released.
- ❖ 20+ new planets have been discovered, nearly 2,000 candidate planets.
- ❖ Can we perform a bulk analysis of all the data to search for these periodic signals?

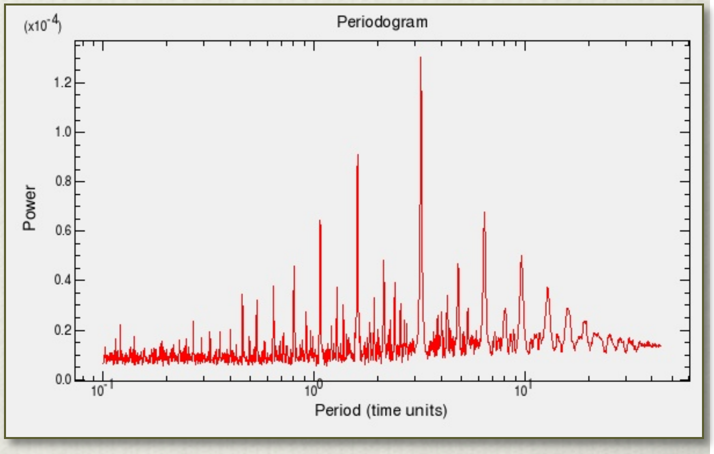


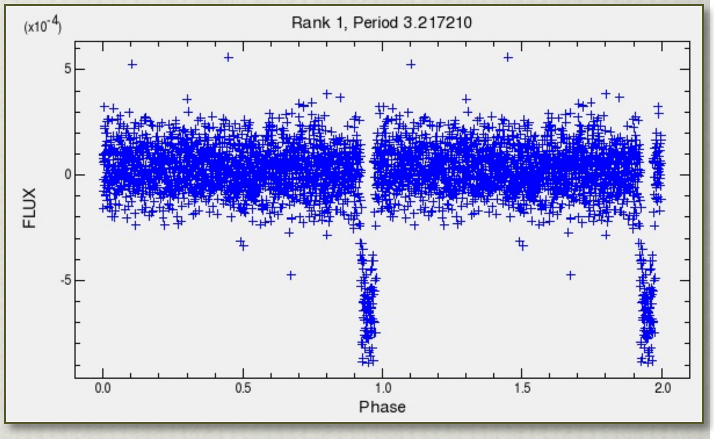


Kepler 6-b transit

Digging Out Exoplanets with Periodograms

- A **periodogram** calculates the significance of different frequencies in time-series data to identify periodic signals.
- *NASA Exoplanet Database* Periodogram tool
  - Fast, portable implementation in C
  - Easily scalable: each frequency sampled independently of all other frequencies
- Calculations are slow: 1 hour for 100,000-200,000 points typical of Kepler light curves.
- How can we process the entire data set? Candidate for the cloud: "high-burst," processor-bound, easily parallelizable.





Kepler Periodogram Atlas

- Compute periodogram atlas for public Kepler dataset
  - Use 128 processor cores in parallel on Amazon EC2 and TeraGrid
  - ~210K light curves X 3 algorithms

Run	Algorithm	Optimization
1 (EC1)	Lomb-Scargle	Sinusoids
2 (EC1)	Box-Least Squares	Box
3 (TG)	Plavchan	Unrestricted

		Run 1 (EC2)	Run 2 (EC2)	Run 3 (TeraGrid)
Runtimes	**Tasks**	631992	631992	631992
	Mean Task Runtime	7.44 sec	6.34 sec	285 sec
	Jobs	25401	25401	25401
	Mean Job Runtime	3.08 min	2.62 min	118 min
	Total CPU Time	1304 hr	1113 hr	50019 hr
	Total Wall Time	16.5 hr	26.8 hr	448 hr
Inputs	**Input Files**	210664	210664	210664
	Mean Input Size	0.084 MB	0.084 MB	0.084 MB
	Total Input Size	17.3 GB	17.3 GB	17.3 GB
Outputs	**Output Files**	1263984	1263984	1263984
	Mean Output Size	0.171 MB	0.124 MB	5.019 MB
	Total Output Size	105.3 GB	76.52 GB	3097.87 GB
Cost	**Compute Cost**	$179.52	$94.61	$4,874.24
	Output Cost	$15.80	$11.48	$464.68
	Total Cost	**$195.32**	**$106.08**	**$5,338.92**

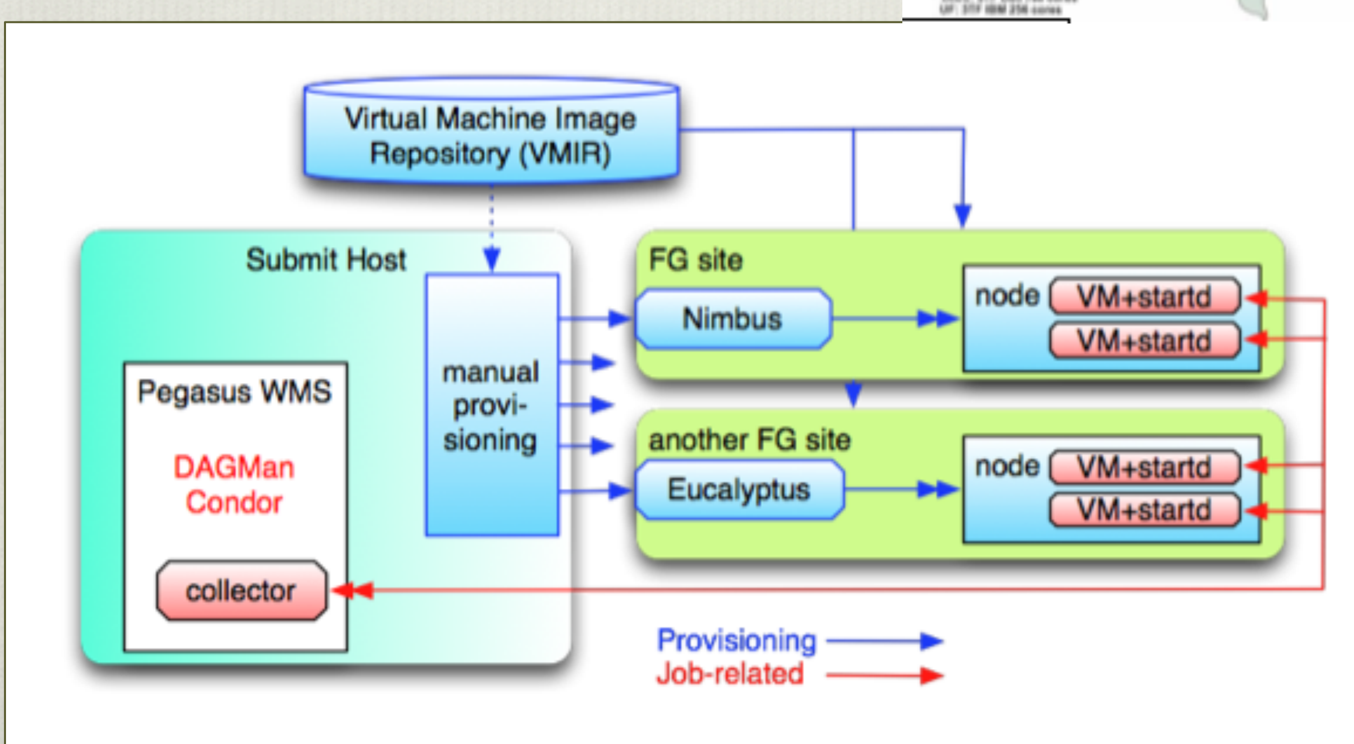
Compute is ~10X Transfer

Estimated cost

Amazon: 16 x c1.xlarge instances = 128 cores
Ranger: 8-16 x 16 core nodes = 128-256 cores

Digging Out Exoplanets On Academic Clouds

- FutureGrid test bed for Cloud Computing
  - 6 centers across the U.S.
  - Nimbus, Eucalyptus
  - http://www.futuregrid.org/



Computing Periodograms on Academic Clouds

Site	CPU	RAM (SW)	Walltime	Cum. Dur.	Speed-Up
Magellan	8 x 2.6 GHz	19 (0) GB	5.2 h	226.6 h	43.6
Amazon	8 x 2.3 GHz	7 (0) GB	7.2 h	295.8 h	41.1
FutureGrid	8 x 2.5 GHz	29 (½) GB	5.7 h	248.0 h	43.5

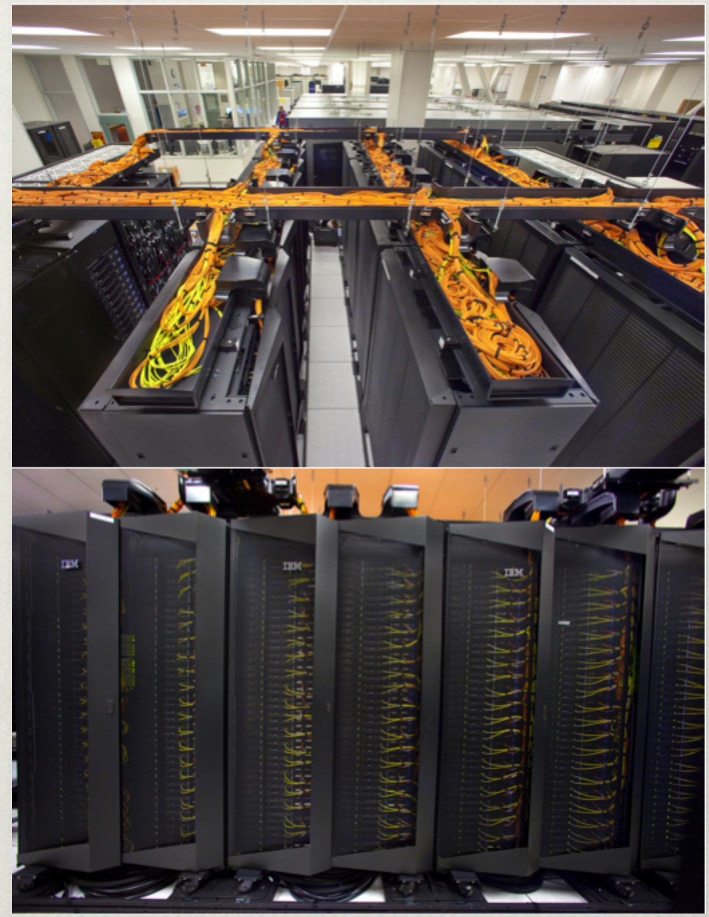
- 33 K periodograms with Plavchan algorithm
- Given 48 physical cores
  - Speed-up ≈ 43 considered *pretty good*
  - AWS cost ≈ $31:
    - 7.2 h x 6 x c1.large ≈ $29
    - 1.8 GB in + 9.9 GB out ≈ $2
- Results encouraging.

Failure To Provision VM's The Main Problem

	Sierra	Magellan
Requests	20	28
Failed to start	2	0
No route to host	0	5
Invalid IP	1	0
No public IP	1	0
Request timed out	1	1
Insufficient resources	0	6
Total failures	5	12

Summation

- Cloud computing is powerful when applied in the right places: "high-burst" applications, especially when they are processor or memory bound.
  - Be careful of how costs can spiral esp. when storing or moving data.
  - Always perform a cost-benefit analysis
  - When mature, academic clouds may offer a realistic alternative to commercial clouds



BUT ----

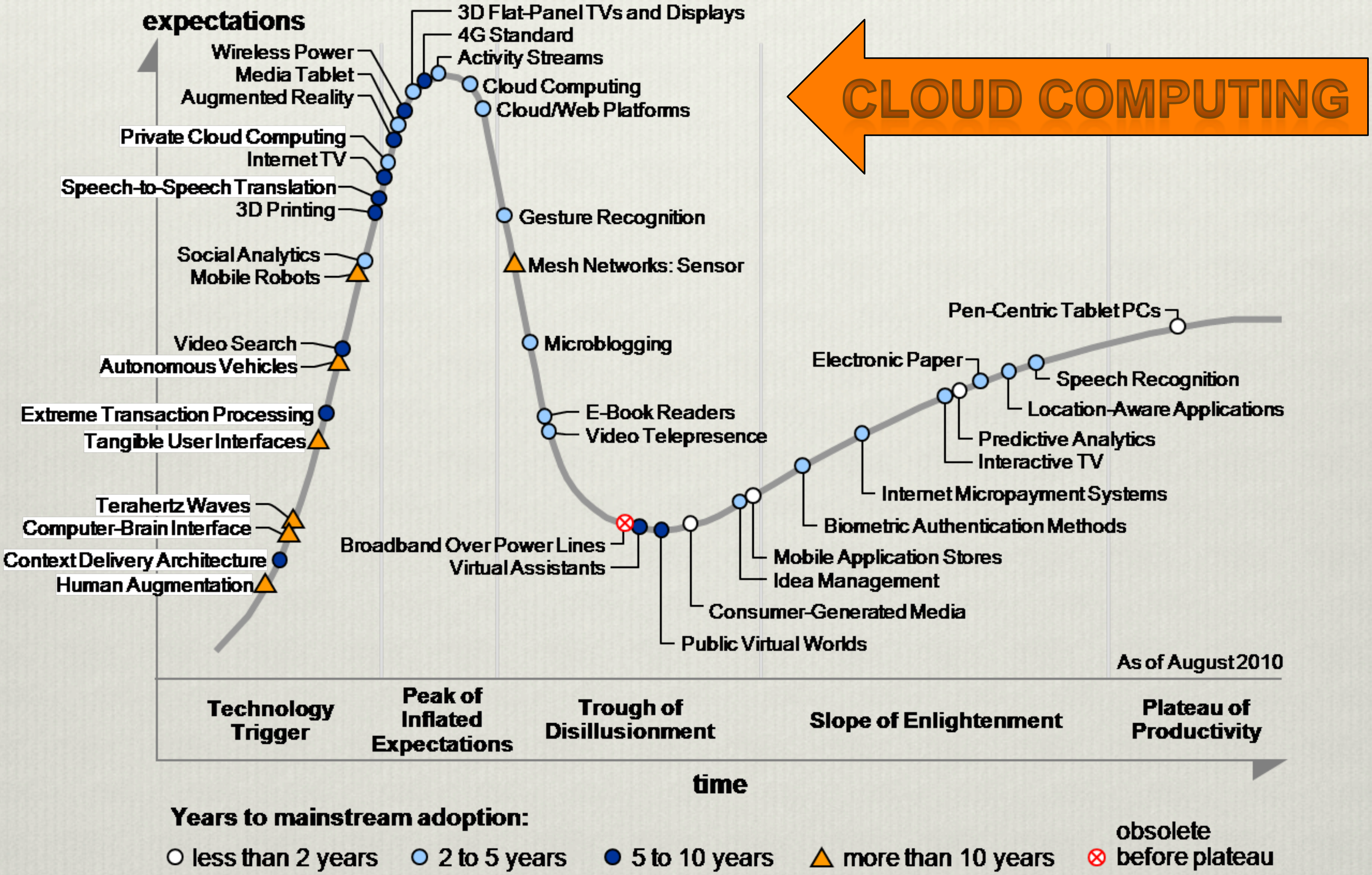
Caveat Emptor!

- **"Cloud Computing as it exists today is not ready for High Performance Computing because**
  - Large overheads to convert to Cloud environments
  - Virtual instances under perform bare-metal systems and
  - The cloud is less cost-effective than most large centers"

  Shane Canon et al. (2011). "Debunking some Common Misconceptions of Science in the Cloud." *Science Cloud Workshop, San Jose, CA. http://datasys.cs.iit.edu/events/ScienceCloud2011/*

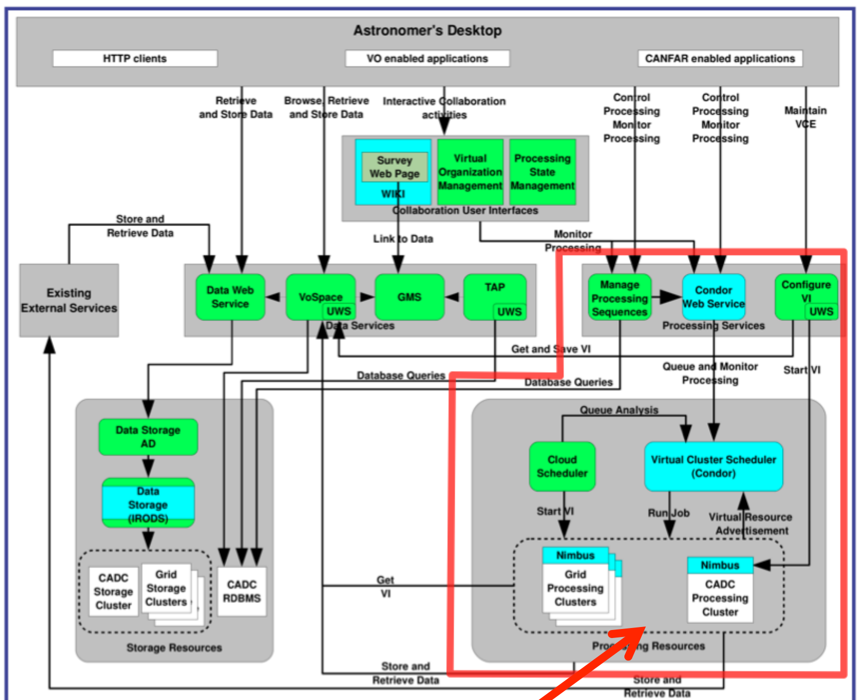
- Similar Conclusions in **Magellan Final Report (December 2011)** *http://science.energy.gov/ascr/*

Gartner's Emerging Technologies Hype Cycle



CANFAR: Cloud Technology in Action in Astronomy

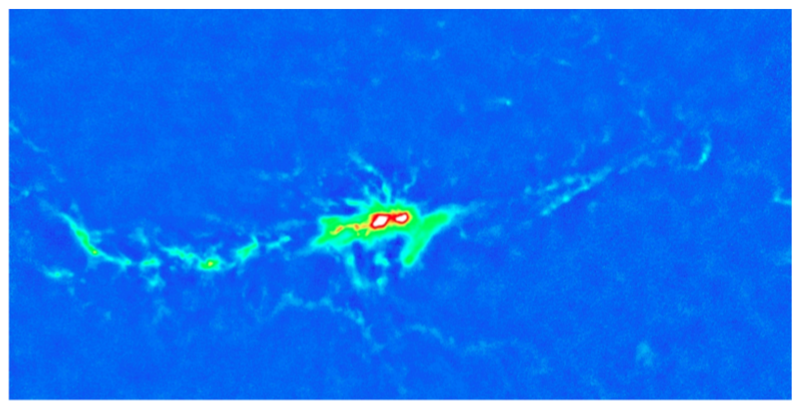
"The Canadian Advanced Network For Astronomical Research (CANFAR) is an operational system for the delivery, processing, storage, analysis, and distribution of very large astronomical datasets. The goal of CANFAR is to support large Canadian astronomy projects."



Nimbus open-source toolkit. IaaS tools to support scientific computing.

Project Name	Project Lead	Institution	Telescope
SCUBA-2 Cosmology Legacy Survey	Mark Halpern	UBC	JCMT
SCUBA-2 All-Sky Survey	Douglas Scott	UBC	JCMT
Next Generation Virgo Survey	Laura Ferrarese	NRC-HIA	CFHT
Pan-Andromeda Archaeological Survey	Harvey Richer	UBC	CFHT
Time Variable Sky	Chris Pritchet	UVic	CFHT
Canada-France Ecliptic Plane Survey Simulator	JJ Kavelaars	NRC-HIA	CFHT
Shapes and Photometric Redshifts for Large Surveys	Ludo Van Waerbeke	UBC	CFHT

Projects currently using the CANFAR processing system



Dust emission from Orion Molecular Cloud at 850um processed using CANFAR (JCMT SCUBA-2 map thanks to Ed Chapin and Mark Halpern, UBC)

Where Can I Learn More?

- **The Application of Cloud Computing to Scientific Workflows: A Study of Cost and Performance**. G. Berriman et al. 2012**.** Invited Review Paper Submitted to Special e-Science Edition of Philosophical Transactions of the Royal Society A.

- **Scientific Workflow Applications on Amazon EC2.** G. Juve et al. Cloud Computing Workshop in Conjunction with e-Science 2009 (Oxford, UK). **http://arxiv.org/abs/1005.2718**

- **Data Sharing Options for Scientific Workflows on Amazon EC2**, G. Juve et al. Proceedings of Supercomputing 10 (SC10), 2010. **http://arxiv.org/abs/1010.4822**

- **The Application of Cloud Computing to Astronomy: A Study of Cost and Performance.** G. B. Berriman et al. 2010. Proceedings of "e-Science in Astronomy" Workshop. Brisbane. **http://arxiv.org/abs/1006.4860**

- **Astronomy in the Cloud: Using MapReduce for Image Co-Addition.** K. Wiley et al. 2011. PASP, 123, 366.

- **Magellan Final Report, December 2011.** http://science.energy.gov/ascr/. Summary: http://www.isgtw.org/feature/assessing-science-cloud

- Bruce Berriman's blog, "Astronomy Computing Today," at **http://astrocompute.wordpress.com**