Putting Eggs in Many Baskets: Data Considerations in the Cloud

Rob Futrick, CTO

CYCLECOMPUTING

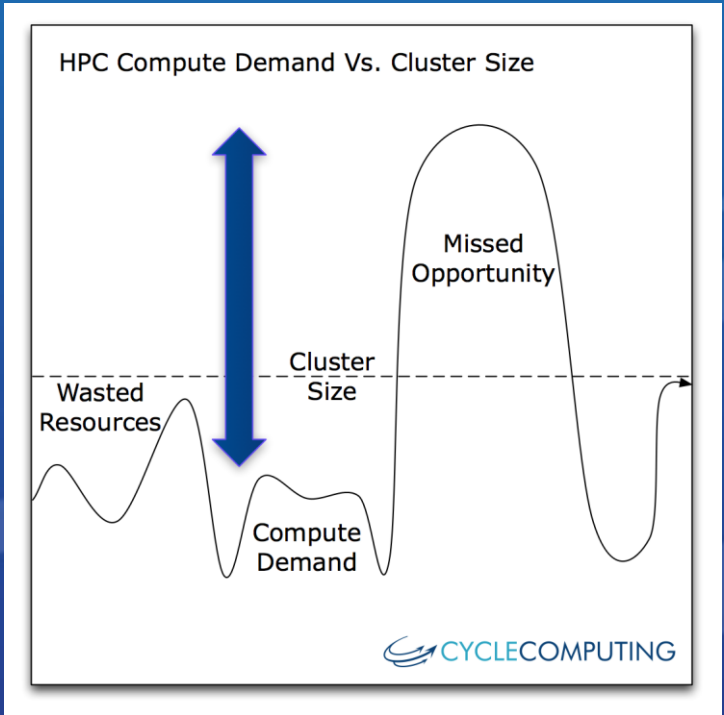
We believe utility access to technical computing power
accelerates discovery & invention

CYCLECOMPUTING

The Innovation Bottleneck:

Scientists/Engineers
forced to size their work to the
infrastructure their organization bought

Limitations of fixed infrastructure



Too small when needed most,
Too large every other time…

- Upfront CapEx anchors you to aging servers
- Costly administration
- Miss opportunities to do better risk management, product design, science, engineering

CYCLECOMPUTING

Our mission:
Write software to make
utility technical computing
easy for anyone,
on any resources,
at any scale

CYCLECOMPUTING

As an example…

Many users use 40 - 4000 cores,
but let's talk about an example:

World's first PetaFLOPS
((Rmax+Rpeak)/2)
Throughput Cloud Cluster

CYCLECOMPUTING

Study Rules Out Global Warming Being a Natural Fluctuation With 99% Certainty

Posted by **Soulskill** on Saturday April 12, 2014 @11:39AM
from the let's-blame-the-dinosaurs dept.

An anonymous reader writes

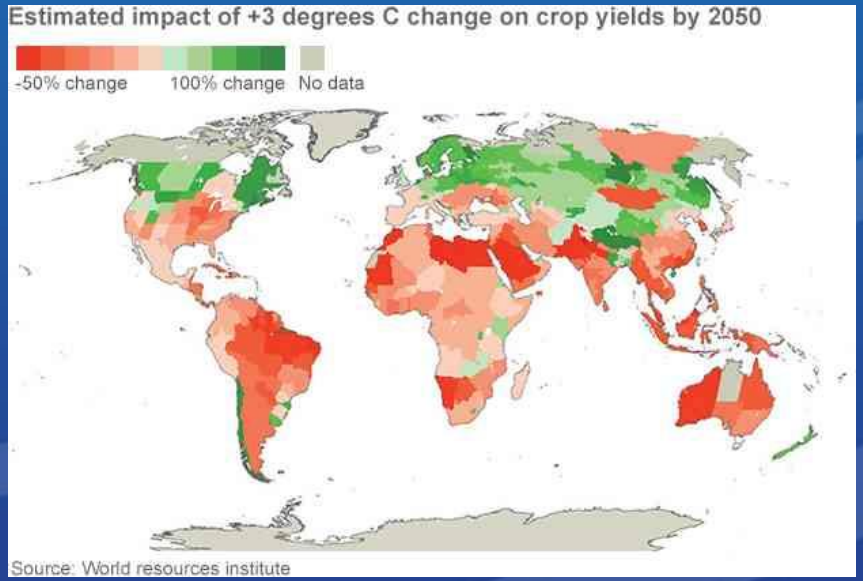
> "A study out of McGill University sought to examine historical temperature data going back 500 years in order to determine the likelihood that global warming was caused by natural fluctuations in the earth's climate. The study concluded there was less than a 1% chance the warming could be attributed to simple fluctuations. 'The climate reconstructions take into account a variety of gauges found in nature, such as tree rings, ice cores, and lake sediments. And the fluctuation-analysis techniques make it possible to understand the temperature variations over wide ranges of time scales. For the industrial era, Lovejoy's analysis uses carbon-dioxide from the burning of fossil fuels as a proxy for all man-made climate influences – a simplification justified by the tight relationship between global economic activity and the emission of greenhouse gases and particulate pollution, he says. ... His study [also] predicts, with 95% confidence, that a doubling of carbon-dioxide levels in the atmosphere would cause the climate to warm by between 2.5 and 4.2 degrees Celsius. That range is more precise than – but in line with — the IPCC's prediction that temperatures would rise by 1.5 to 4.5 degrees Celsius if CO2 concentrations double.'"

Read the **849** comments

x news x statistics x collapse x cuethedeniers x climatechange story

CYCLECOMPUTING

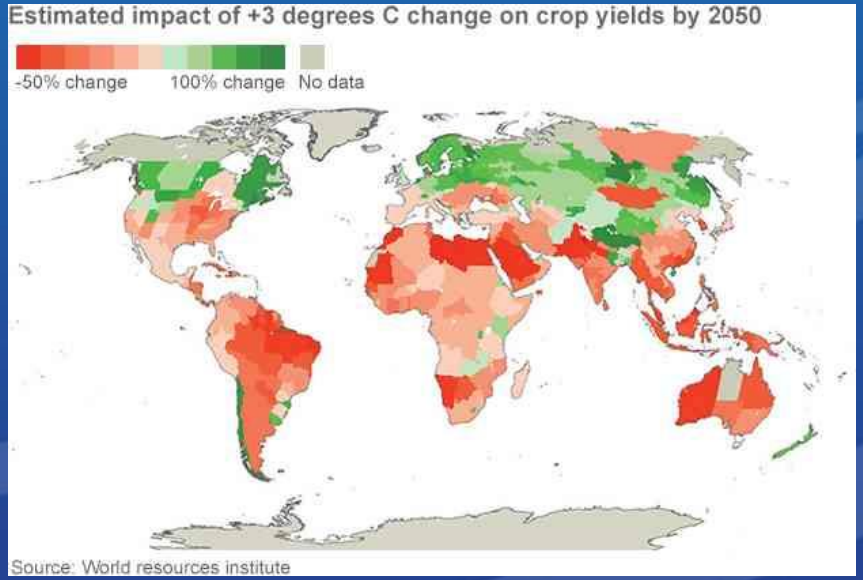
What do you think?



- Much of the worlds “bread basket” land will be hotter and drier
- Ocean warming is decreasing fish populations / catches

CYCLECOMPUTING

First, buy land in Canada?



CYCLECOMPUTING

Solar Energy

Wind power

Nuclear Fission energy

GeoThermal

Sure! But there have to be engineer-able solutions too.

Climate Engineering

Nuclear Fusion

BioFuels

CYCLECOMPUTING

Designing Solar Materials

The challenge is efficiency - turning photons to electricity

The number of possible materials is limitless:

- Need to separate the right compounds from the useless ones
- Researcher Mark Thompson, PhD:

> “If the 20th century was the century of silicon, the 21st will be all organic. Problem is, how do we find the right material without spending the entire 21st century looking for it?”

Needle in a Haystack Challenge:

205,000 compounds
totaling 2,312,959 core-hours
or 264 core-years

CYCLECOMPUTING

205,000 molecules
264 years of computing

16,788 Spot Instances, 156,314 cores

CYCLECOMPUTING

205,000 molecules
264 years of computing

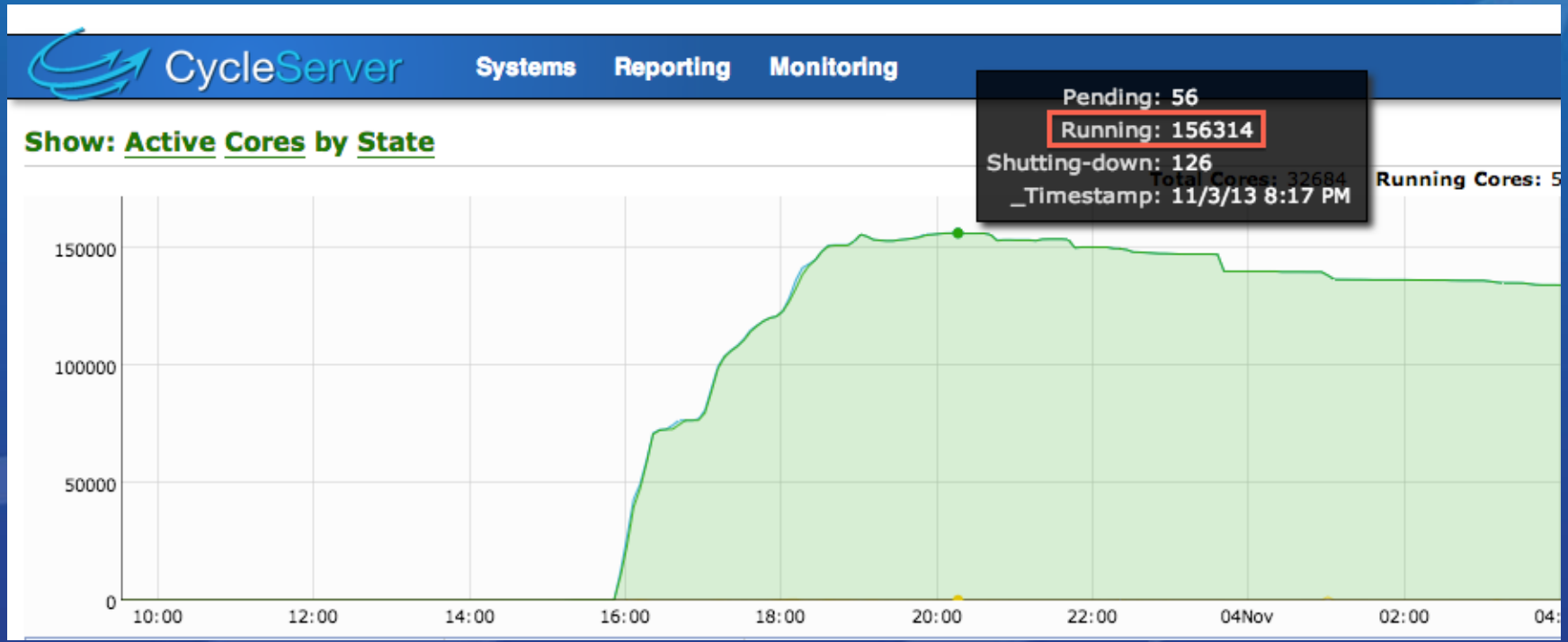
156,314 cores =
1.21 PetaFLOPS (~Rpeak)

CYCLECOMPUTING

8-Region Deployment

CYCLECOMPUTING

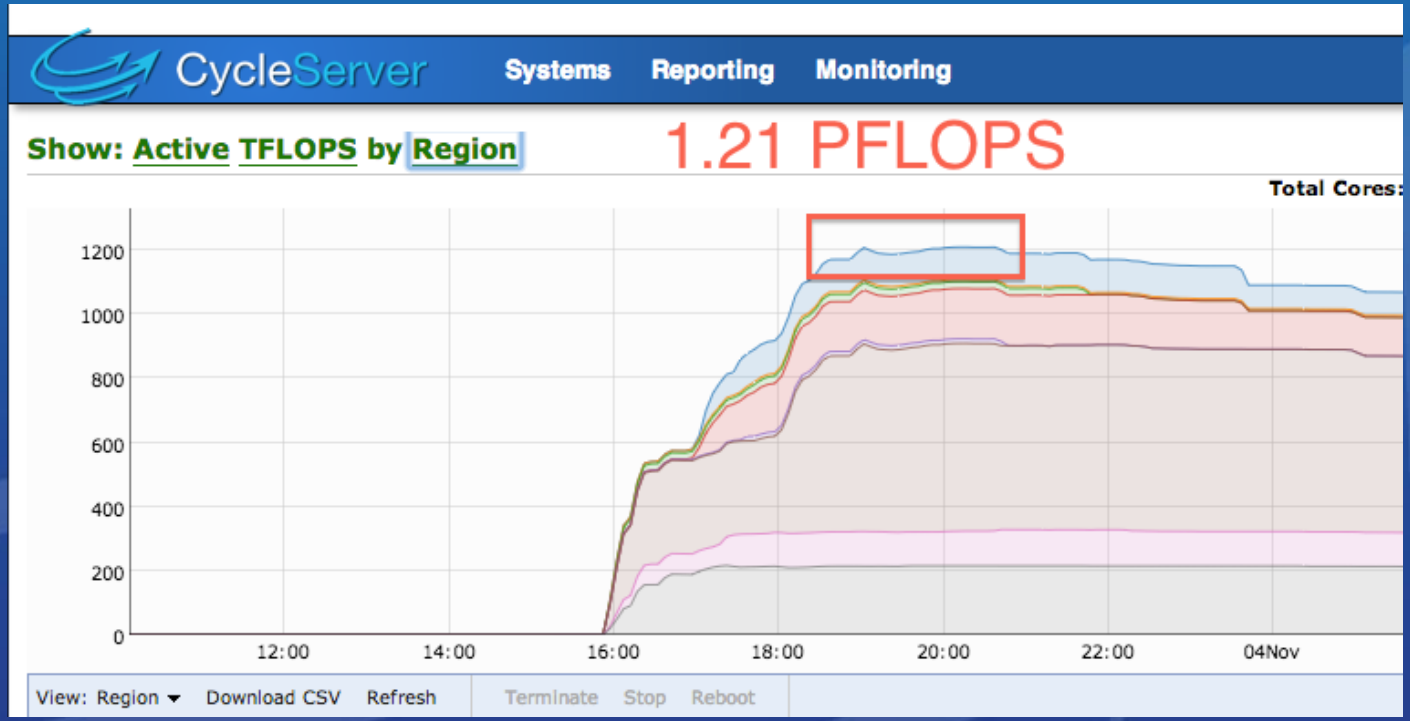
1.21 PetaFLOPS (Rmax+Rpeak)/2, 156,314 cores



CYCLECOMPUTING

Each individual task was MPI, using a single, entire machine

Benchmark individual machines



CYCLECOMPUTING

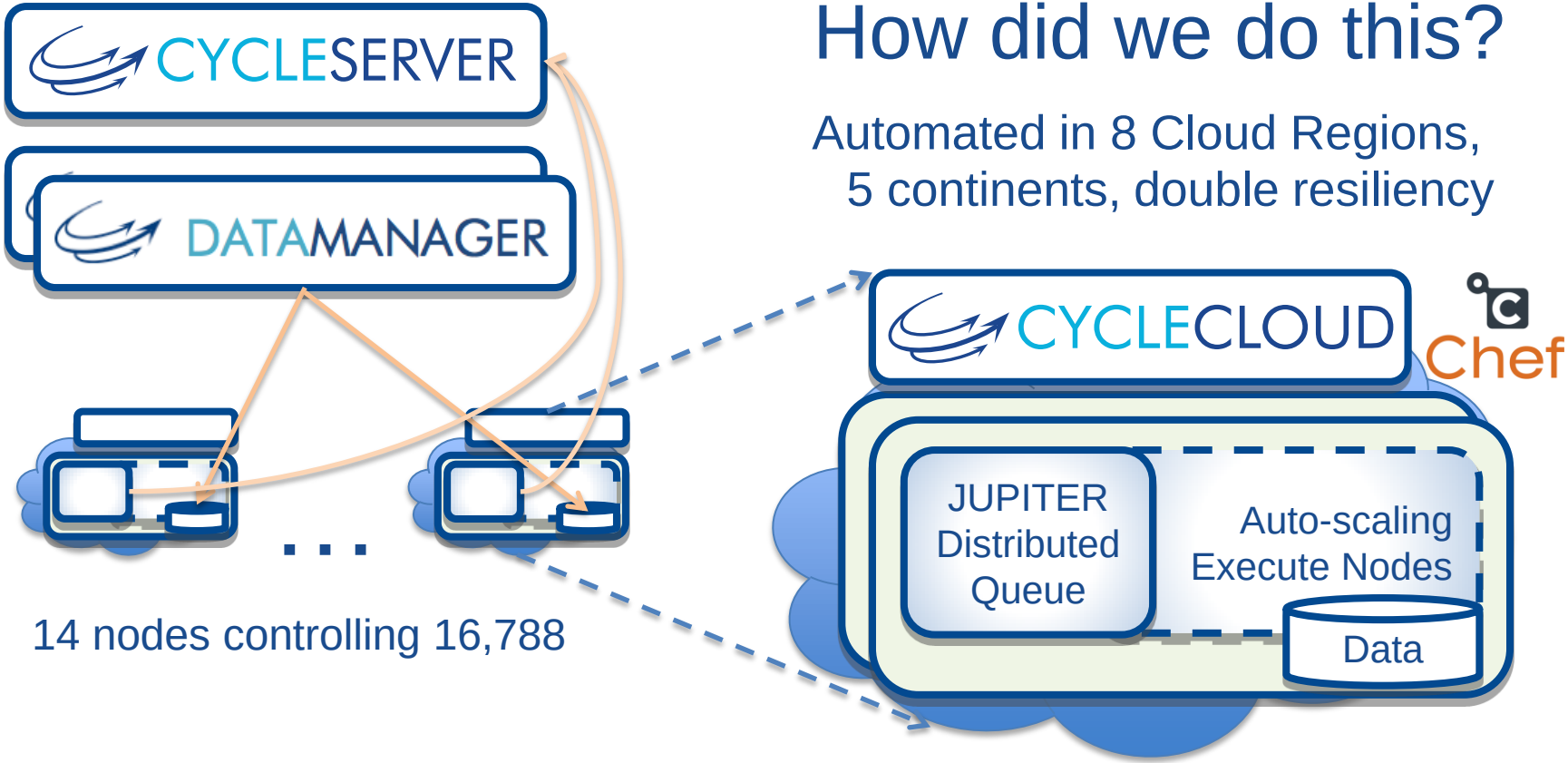
205,000 molecules
264 years of computing

Done in 18 hours
Access to $68M system for $33k

CYCLECOMPUTING

How did we do this?

Automated in 8 Cloud Regions,
5 continents, double resiliency



Now Dr. Mark Thompson

- Is 264 compute years closer to making efficient solar a reality using organic semiconductors

CYCLECOMPUTING

Important?

Not to me anymore ;)

CYCLECOMPUTING

Large Memory

Interconnect Sensitive

Whole sample set analysis

Large Grid MPI Runs

We see across all workloads

Needle in a Haystack runs

Hi I/O
Big Data runs

Interactive, SOA

CYCLECOMPUTING

Users want to decrease
Lean manufacturing's 'Cycle time'
=
(prep time + queue time + run time)

Everyone we work with faces this problem

SLED/PS	Insurance	Financial Services	Life Sciences	Manufacturing & Electronics	Energy, Media & Other
Morgridge Institute for Research	The Hartford	JPMorganChase	Novartis	Johnson & Johnson	ConocoPhillips
Purdue University	Pacific Life	Bank of America	Pfizer	BAE Systems	EA
Lockheed Martin	SunAmerica the retirement specialist	NYSE Euronext	Merck	Altera	Schrödinger
	White Mountains Insurance Group	SIG Susquehanna International Group, LLP	Pacific Biosciences	life technologies	IDA
		Fidelity Investments	Lilly	Aerospace Assuring Space Mission Success	
			Genentech	cadence	
			genomic health		

CYCLECOMPUTING

External resources
(Open Science Grid, cloud)
offer a solution!

How do we get there?

CYCLECOMPUTING

In particular how do we deal with data in our workflows?

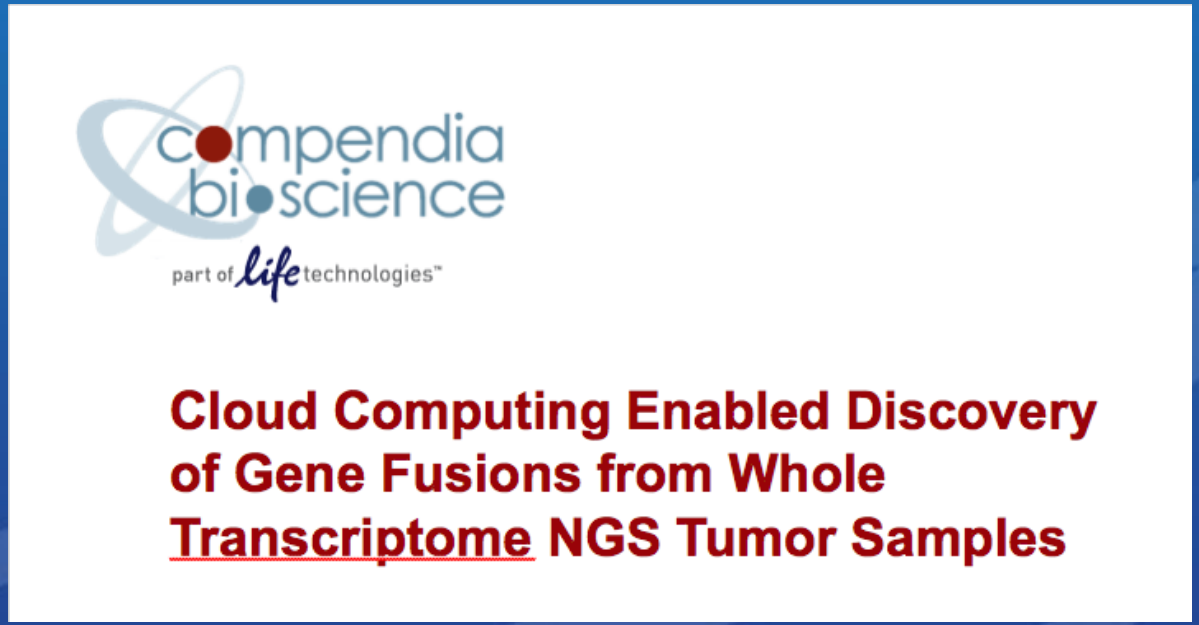
Several of the options…

- Local disk
- Object Store / Simple Storage Service (S3)
- Shared FS
  - NFS
  - Parallel file system (Lustre, Gluster)
- NoSQL DB

Let’s do this through examples

Compendia BioSciences (Local Disk)

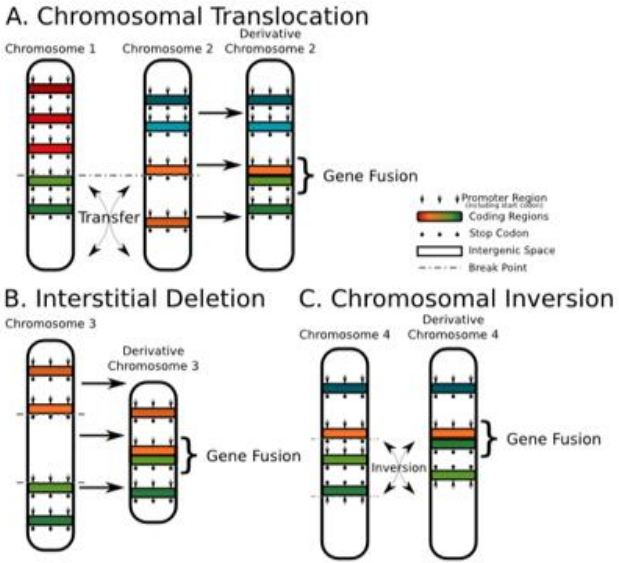
CYCLECOMPUTING



The Cancer Genome Atlas

CYCLECOMPUTING

Gene Fusion Overview



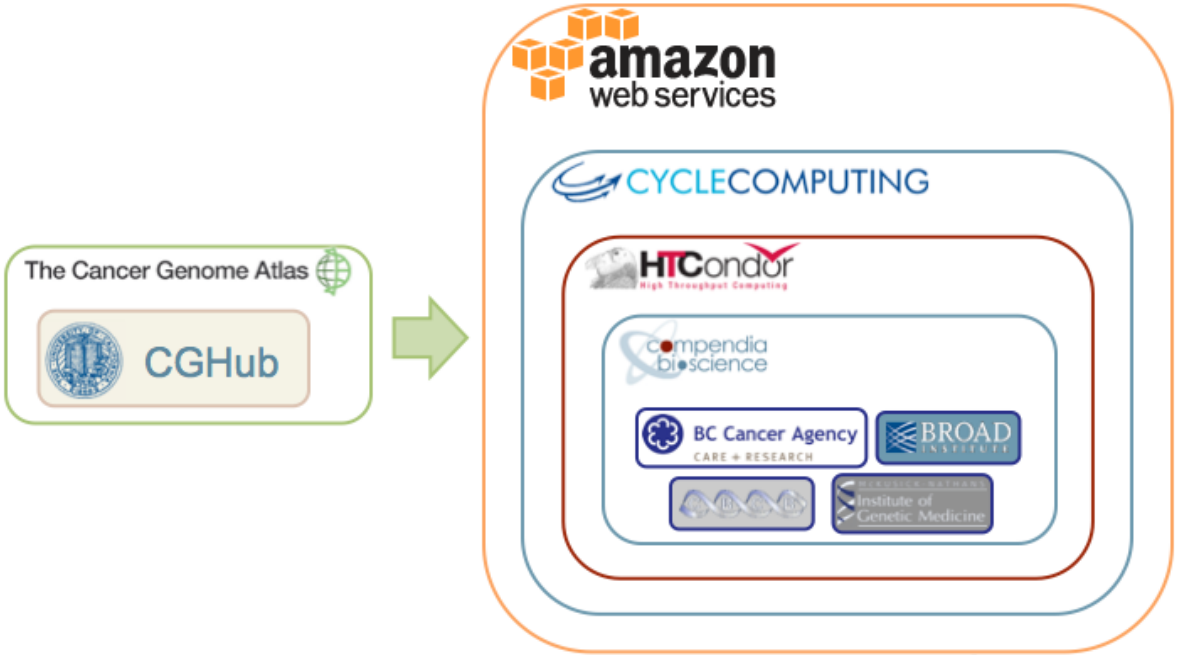
http://en.wikipedia.org/wiki/Gene_fusion

compendia bioscience
part of life technologies

Click to add text

CYCLECOMPUTING

Fusion Calling Platform

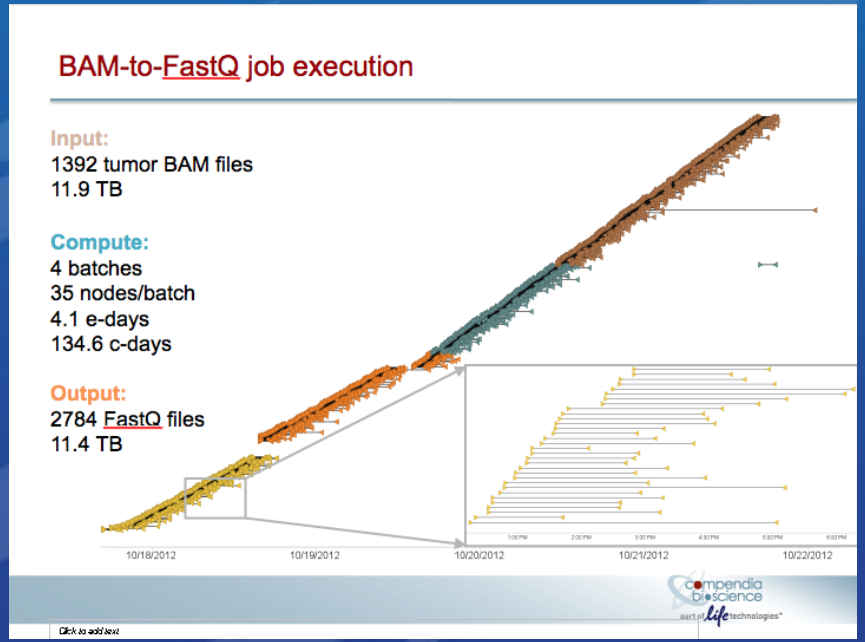


Click to add text

CYCLECOMPUTING

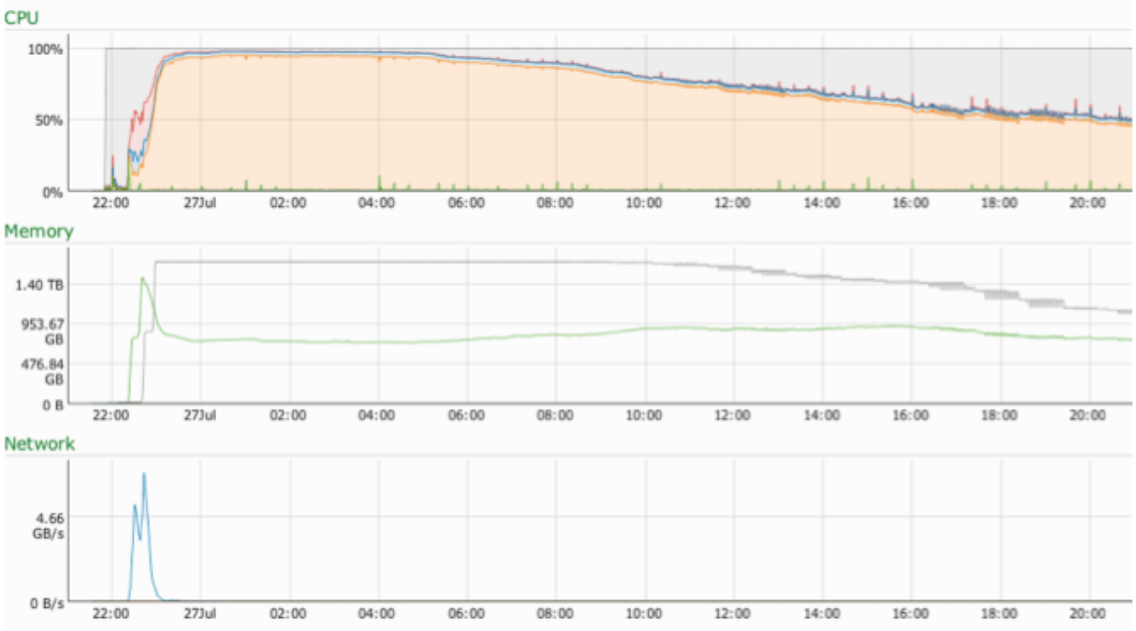
As described at TriMed Conference

- Stream data out of TCGA into S3 & Local machines (concurrently on 8k cores)
- Run analysis and then place results in S3
- No Shared FS, but all nodes are up for a long time downloading (8000 cores * xfer)



CYCLECOMPUTING

Defuse node utilization

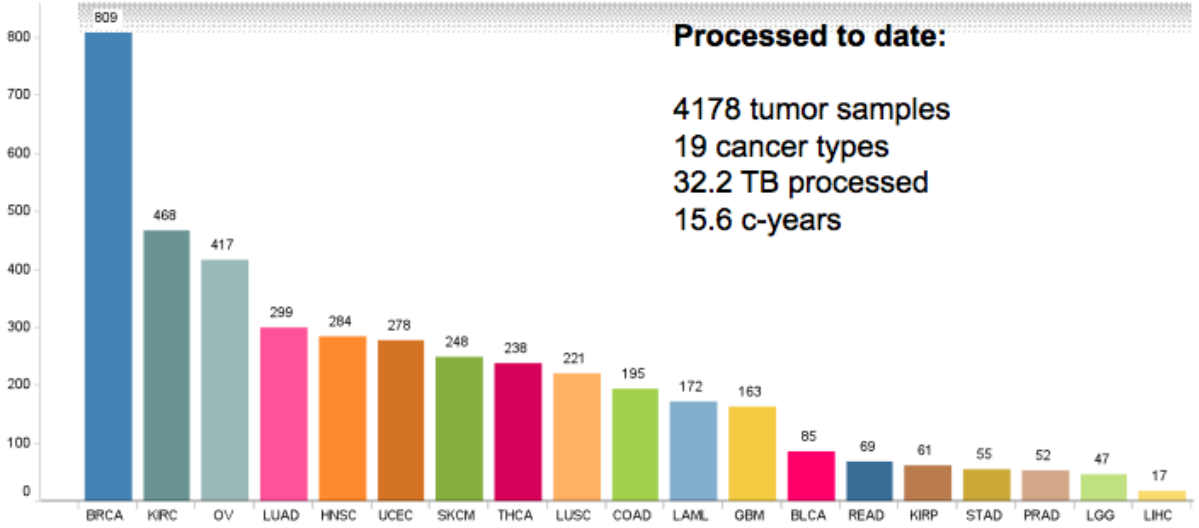


compendia bioscience
part of life technologies

Click to add text

CYCLECOMPUTING

Fusion Caller Processing Summary



Processed to date:

4178 tumor samples
19 cancer types
32.2 TB processed
15.6 c-years

As of 12/31/2012

Click to add text

compendia bioscience
part of life technologies

CYCLECOMPUTING

Pros & Cons of Local Disk

Pros:

- Typically no application changes
- Data encryption is easy; No shared key management
- Highest speed access to data once local

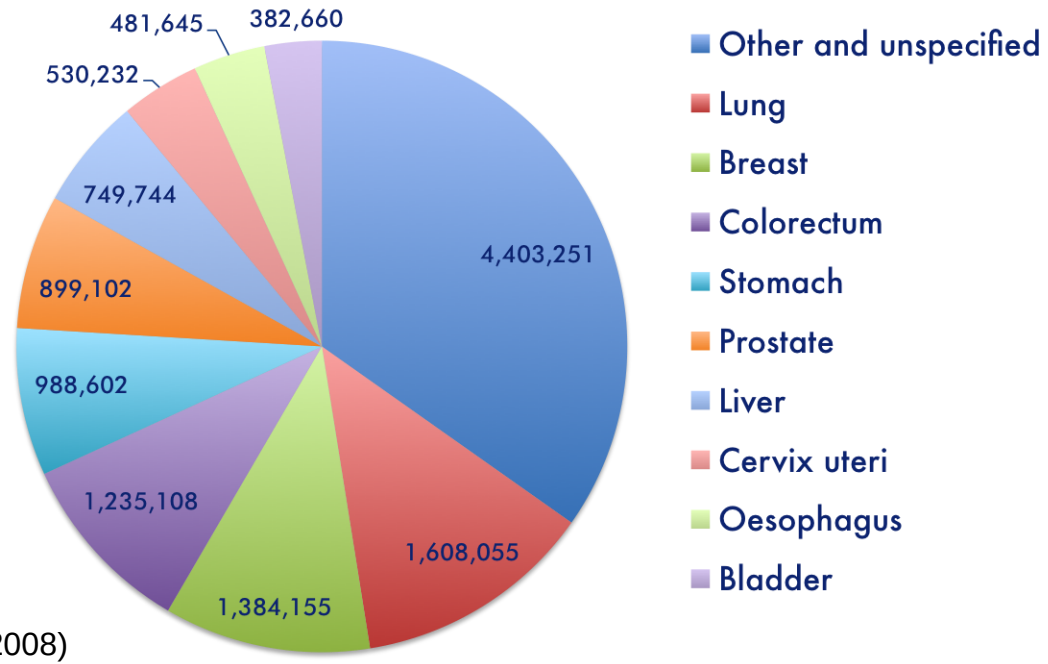
Cons:

- Performance cost/impact of all nodes downloading data
- Only works for completely mutually exclusive workflows
- Possibly more expensive; Transferring to a shared filer and then running may be cheaper

Novartis
(S3 sandboxes)

CYCLECOMPUTING

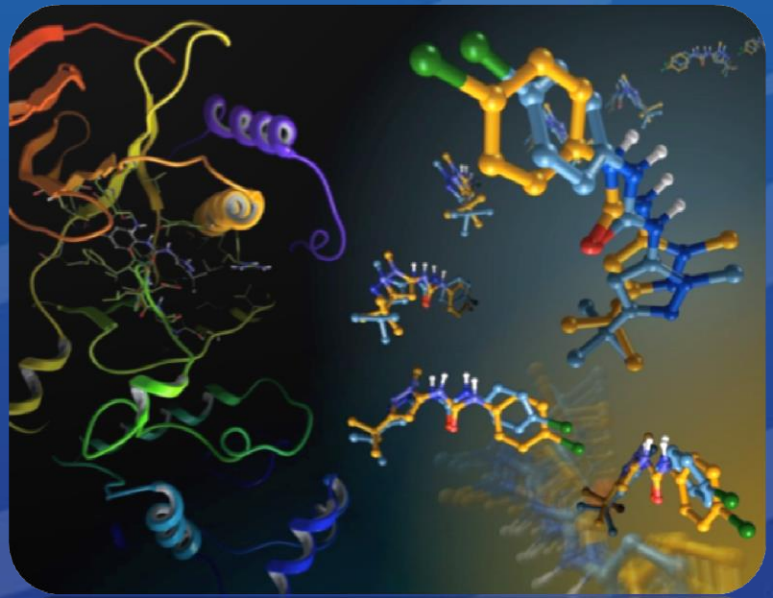
New Cases of Cancer per Year: 12.66 Million



(W.H.O./Globocan 2008)

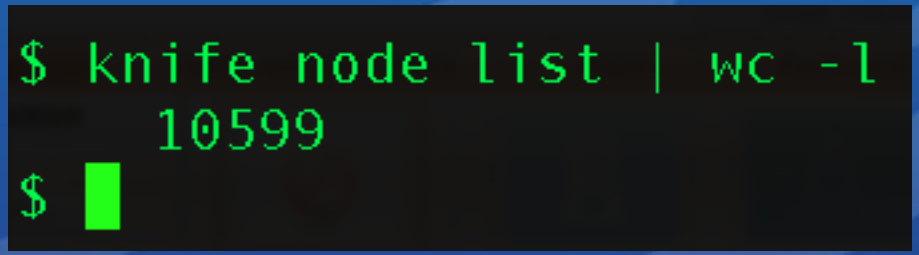
Every day is
crucial and **costly**

Challenge: Novartis run 341,700 hours of docking against a cancer target; impossible to do in-house.



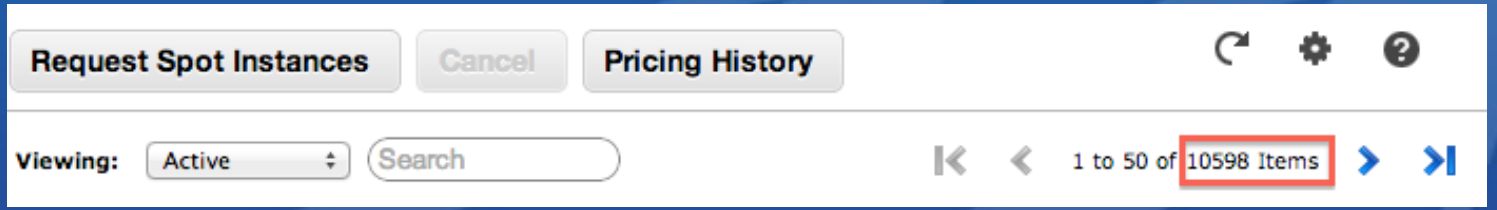
CYCLECOMPUTING

Most Recent Utility Supercomputer server count:



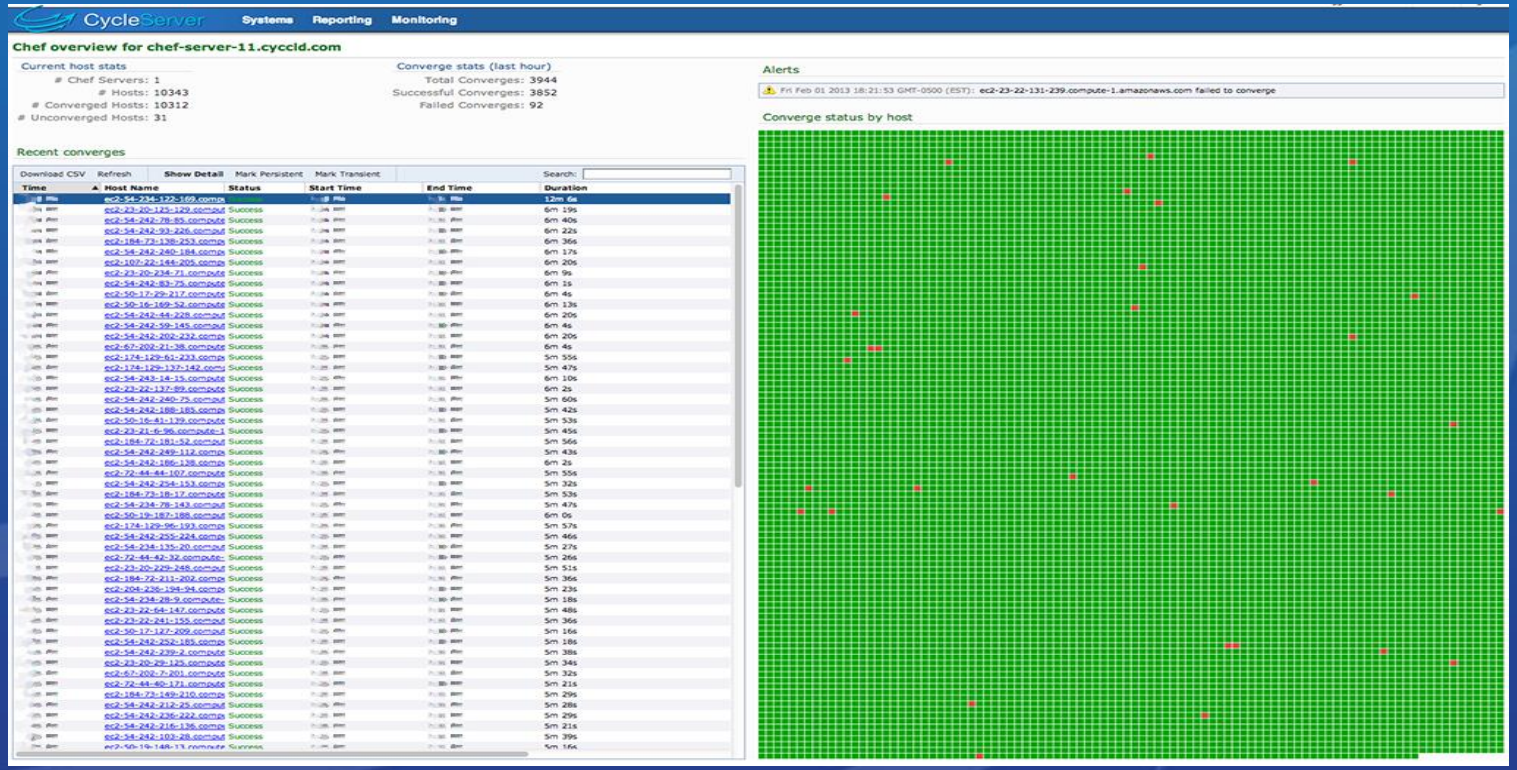
CYCLECOMPUTING

AWS Console view (the only part that rendered):



CYCLECOMPUTING

Cycle Computing's cluster view:



CYCLECOMPUTING

Metric	Count
Compute Hours of Science	341,700 hours
Compute Days of Science	14,238 days
Compute Years of Science	39 years
AWS Instance Count	10,600 instances

CycleCloud, HTCondor, & Cloud finished impossible workload, $44 Million in servers, …

CYCLECOMPUTING

39 years of drug design in 11 hours
on 10,600 servers for < $4,372

CYCLECOMPUTING

Does this lead to new drugs?

Novartis announced
3 new compounds going
into screening based upon
this 1 run

Pros & Cons of S3

Pros:

- Parallel scalability; high throughput access to data
- Only bottlenecks occur on S3 access
- Can easily and greatly increase capacity

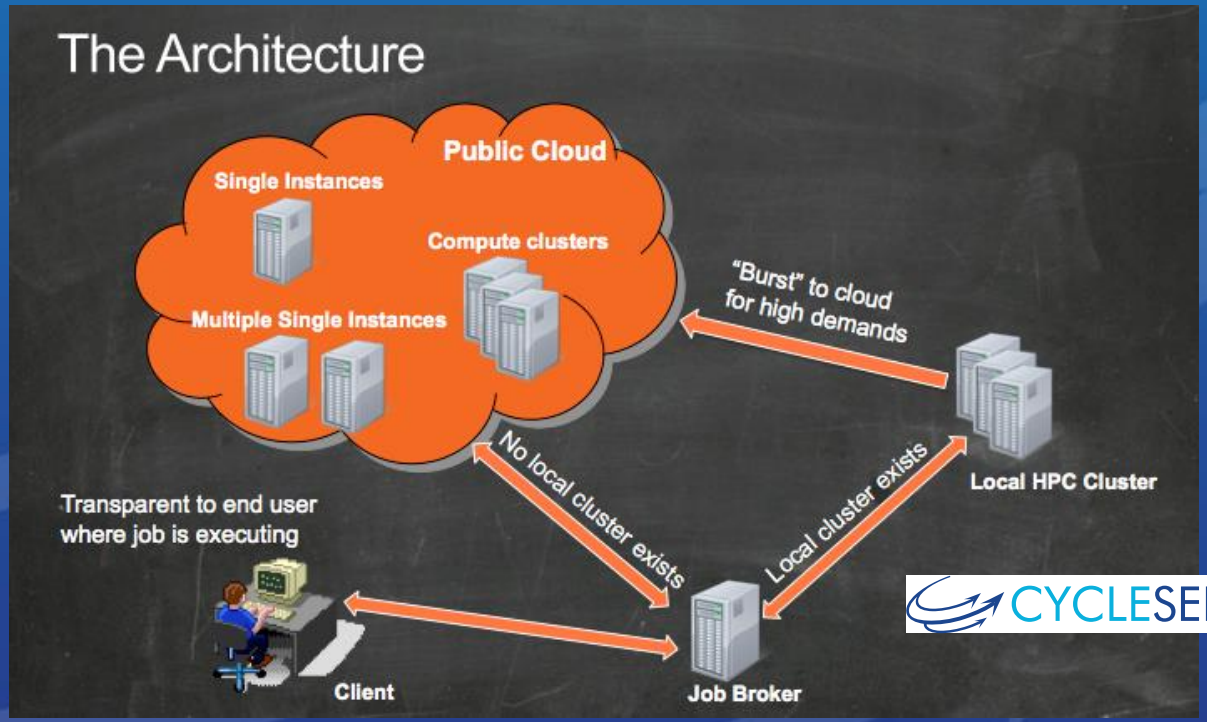
Cons:

- Only works for completely mutually exclusive workflows
- Native S3 access requires application changes
- Non-native S3 access can be unstable
- Latency can affect performance

Johnson & Johnson (Single node NFS)

CYCLECOMPUTING

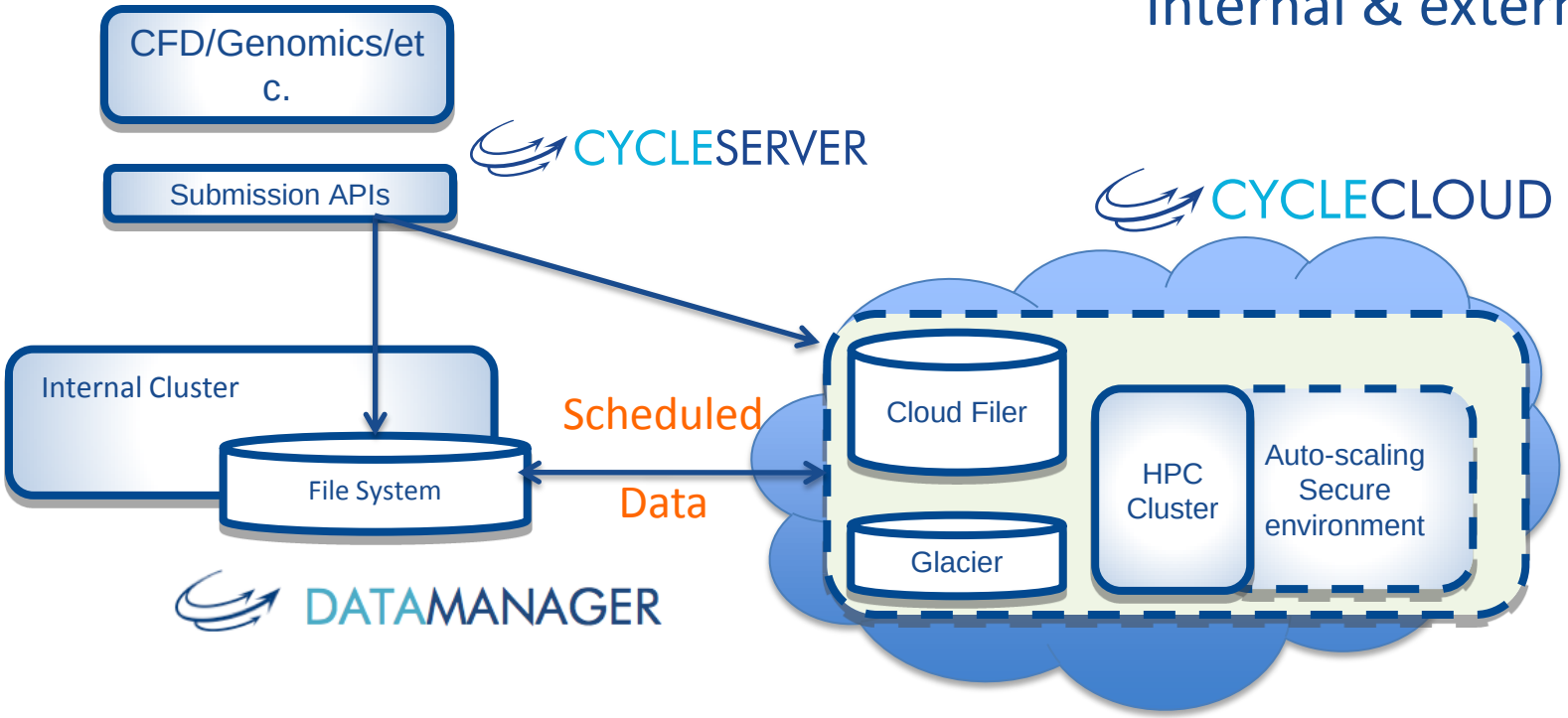
JnJ @ AWS re:Invent 2012



CYCLECOMPUTING

JnJ Burst use case

Internal & external



(Patents Pending)

CYCLE COMPUTING

Pros & Cons of NFS

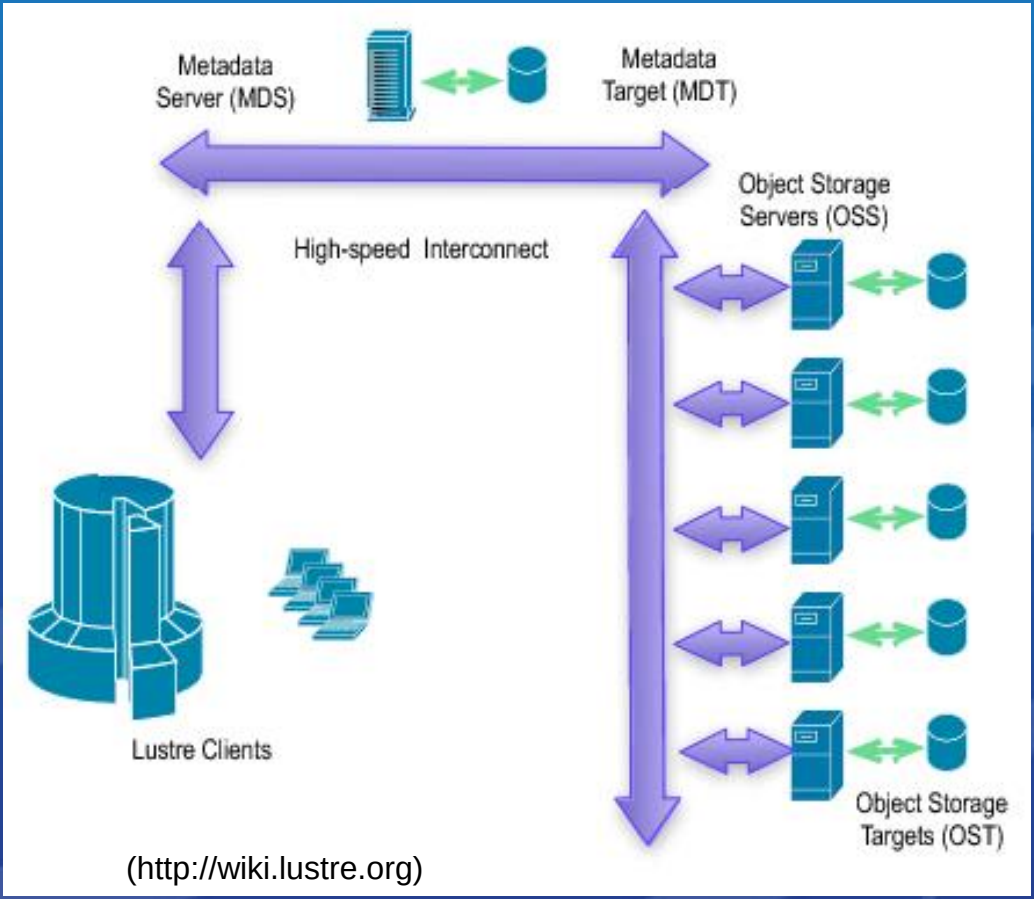
Pros:

- Typically no application changes
- Cheapest at small scale
- Easy to encrypt data
- Performance great at (small) scale and/or under some access patterns
- Great platform support

Cons:

- Filer can easily become performance bottleneck
- Not easily expandable
- Not fault tolerant; Single point of failure
- Backup and recovery

Parallel Filesystem (Gluster, Lustre)



(http://wiki.lustre.org)

Pros & Cons of Parallel FS

Pros:

- Easily expand capacity
- Read performance scalability
- Data integrity

Cons:

- Greatly increased administration complexity
- Performance for “small” files can be atrocious
- Poor platform support
- Data integrity and backup
- Still has single points of failure

NoSQL DB (Cassandra, MongoDB)

CYCLECOMPUTING

(http://strata.oreilly.com/2012/02/nosql-non-relational-database.html)

Pros & Cons of NoSQL DB

Pros:

- Best performance for appropriate data sets
- Data backup and integrity
- Good platform support

Cons:

- Only appropriate for certain data sets and access patterns
- Requires application changes
- Application developer and Administration complexity
- Potential single point of failure

Large Memory

Interconnect Sensitive

Whole sample set analysis

Large Grid MPI Runs

That's a survey of different workloads

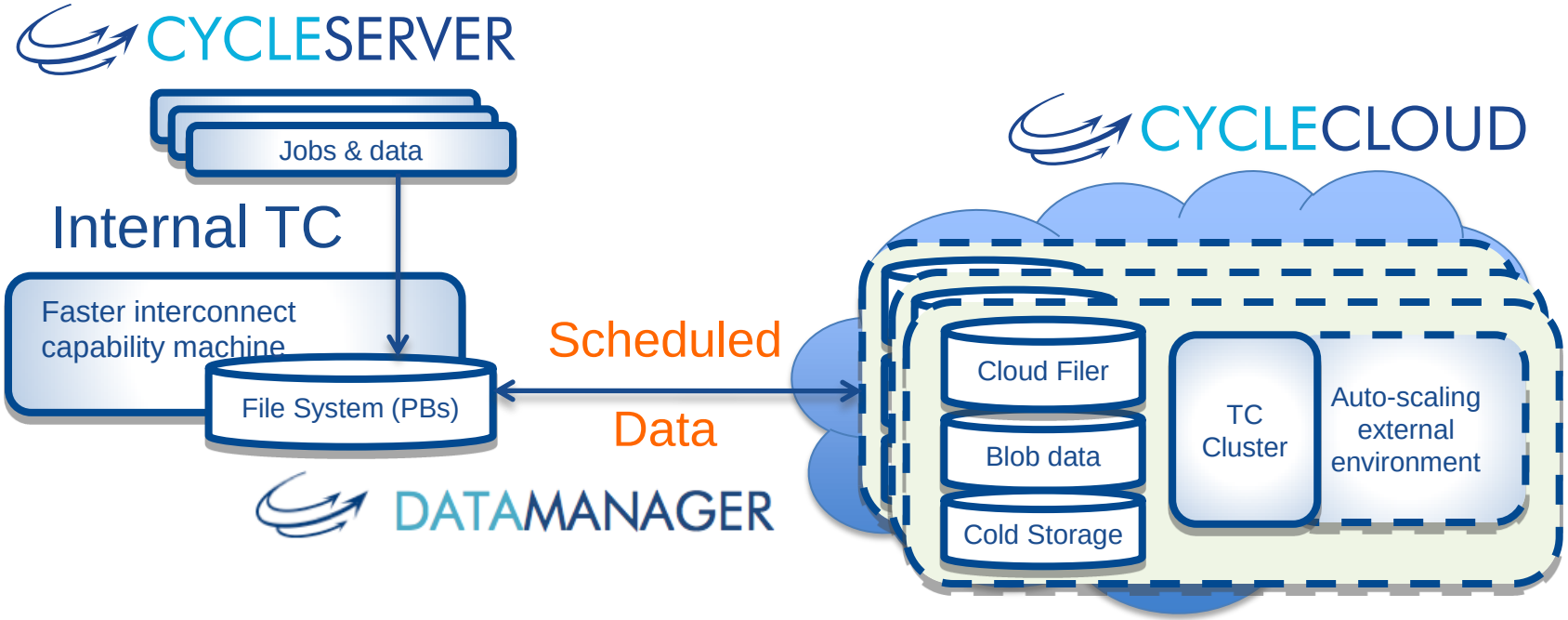
Needle in a Haystack runs

Hi I/O
Big Data runs

Interactive, SOA

CYCLECOMPUTING

Depending upon your use case



There are a lot more examples…

NOVARTIS

Life Sciences

- **39.5 years of drug compound computations in 9 hours**, at a total cost of **$4,372**
- **10,000 server cluster** seamlessly spanned US/EU regions
- Advanced 3 new otherwise unknown compounds in wet lab

SCHRÖDINGER.

Life Sciences

- **156,000-core utility supercomputer** in the Cloud
- Used $68M in servers for 18 hours for $33,000
- Simulated 205,000 materials **(264 years) in 18 hours**

Nuclear engineering — Utilities / Energy

- Approximately 600-core MPI workloads run in Cloud
- Ran workloads in months rather than years
- Introduction to production in 3 weeks

AEROSPACE
Assuring Space Mission Success

Rocket Design & Simulation

- Moving HPC workloads to cloud for burst capacity
- Amazon GovCloud, Security, and key management
- Up to 1000s of cores for burst / agility

PACIFIC LIFE

Asset & Liability Modeling (Insurance / Risk Mgmt)

- Completes monthly/quarterly runs 5-10x faster
- Use 1600 cores in the cloud to shorten elapsed run time from ~10 days to ~ 1-2 days

Johnson&Johnson

Manufacturing & Design

- Enable end user on-demand access to 1000s of cores
- Avoid cost of buying new servers
- Accelerated science and CAD/CAM process

CYCLECOMPUTING

Hopefully you see various data placement options

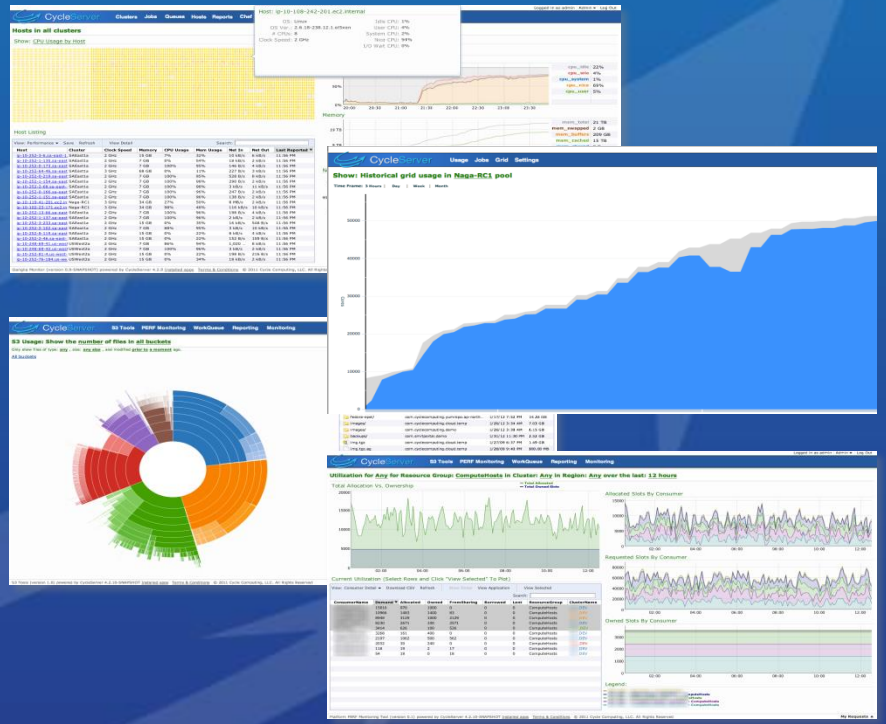
Each have pros and cons…

- Local disk
- Simple Storage Service (S3)
- Shared FS
  - NFS
  - Parallel file system (Lustre, Gluster)
- NoSQL DB

We write software to do this…

Cycle easily orchestrates workloads and data access to local and Cloud TC

- Scales from 100 - 100,000's of cores
- Handles errors, reliability
- Schedules data movement
- Secures, encrypts and audits
- Provides reporting and chargeback
- Automates spot bidding
- Supports Enterprise operations



CYCLECOMPUTING

Does this resonate with you?

We're hiring like crazy:
Software developers,
HPC engineers,
devops, sales, etc.

jobs@
cyclecomputing.com

CYCLECOMPUTING

Now hopefully…

You'll keep these tools in mind

as you use HTCondor

to accelerate your science!

CYCLECOMPUTING