

Democratizing Supercomputers:

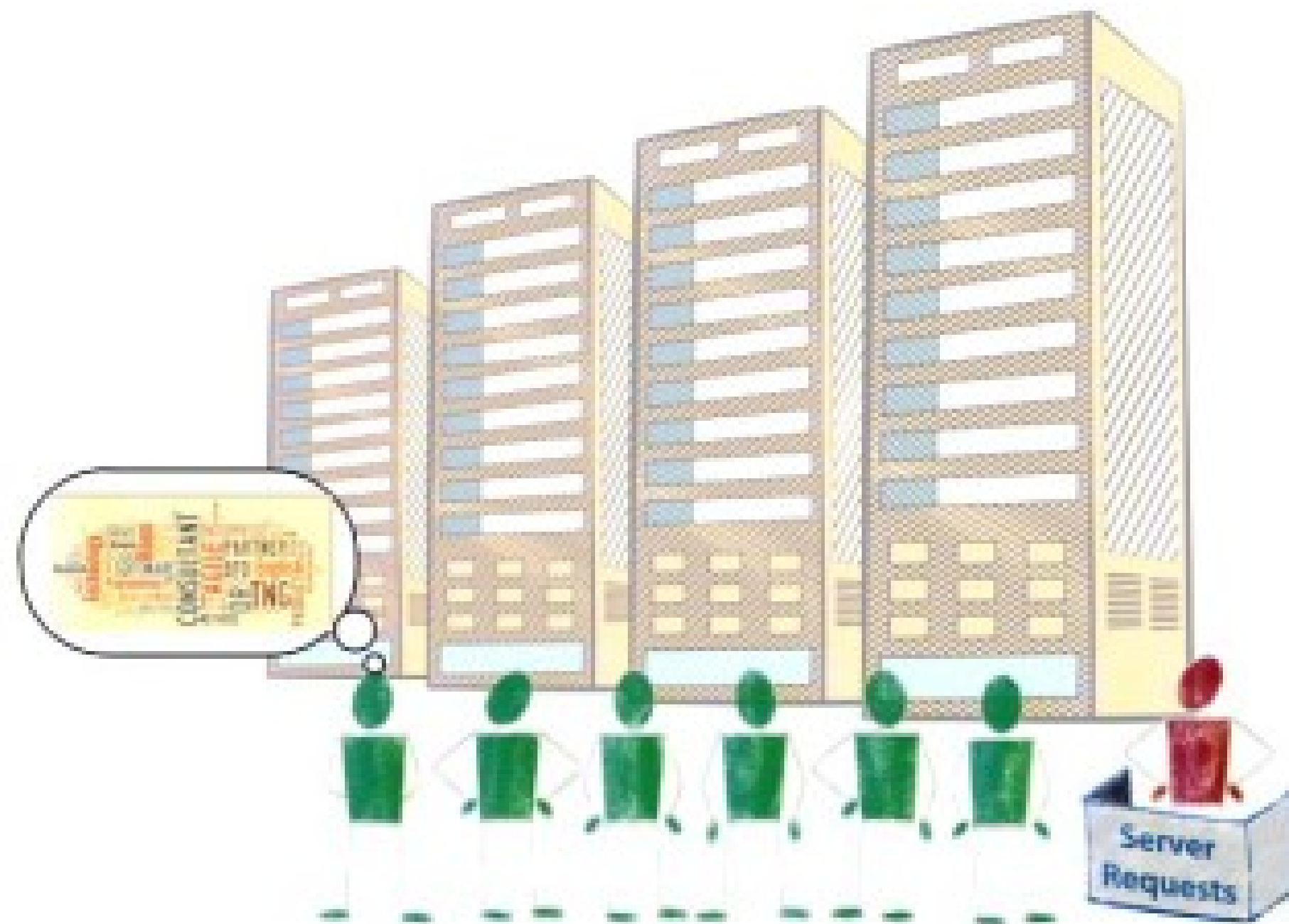
How everybody can wield infinite compute power

Constantin Gonzalez
Solutions Architect AWS
glez@amazon.de



©Amazon.com, Inc. and its affiliates. All rights reserved.

But Building/Managing IT Infrastructure is Painful



- Procurement requests
- Vendor meetings
- RFI/RFQ
- Purchase Orders
- Take Delivery
- Unpack Hardware
- Connect Hardware
- Install OS
- Install Software
- Monitor/Service/Replace

We Know, Because...



*“Amazon is a technology company
that happens to sell books online”*

– Dr. Werner Vogels, CTO Amazon.com

Amazon Web Services



- Since 2006
- Many data centers in 9 regions + China
- Hundreds of thousands of customers
- 190 countries

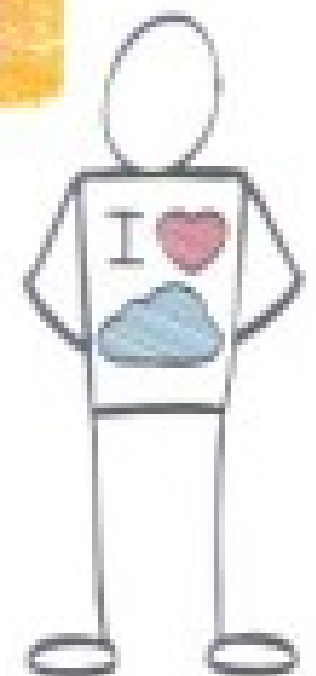
Every day, AWS adds
enough server capacity to
power a whole **\$7B**
enterprise.

amazon.com

That's about the size
of Amazon.com
in **2004**.

We Think There's a Better Way to Do IT

- Add New Dev Environment
- Add New Prod Environment
- Add New Environment in Japan
- Add 1,000 Servers
- Remove 1,000 Servers
- Deploy 1 PB Data Warehouse
- Shut down 1 PB Data Warehouse



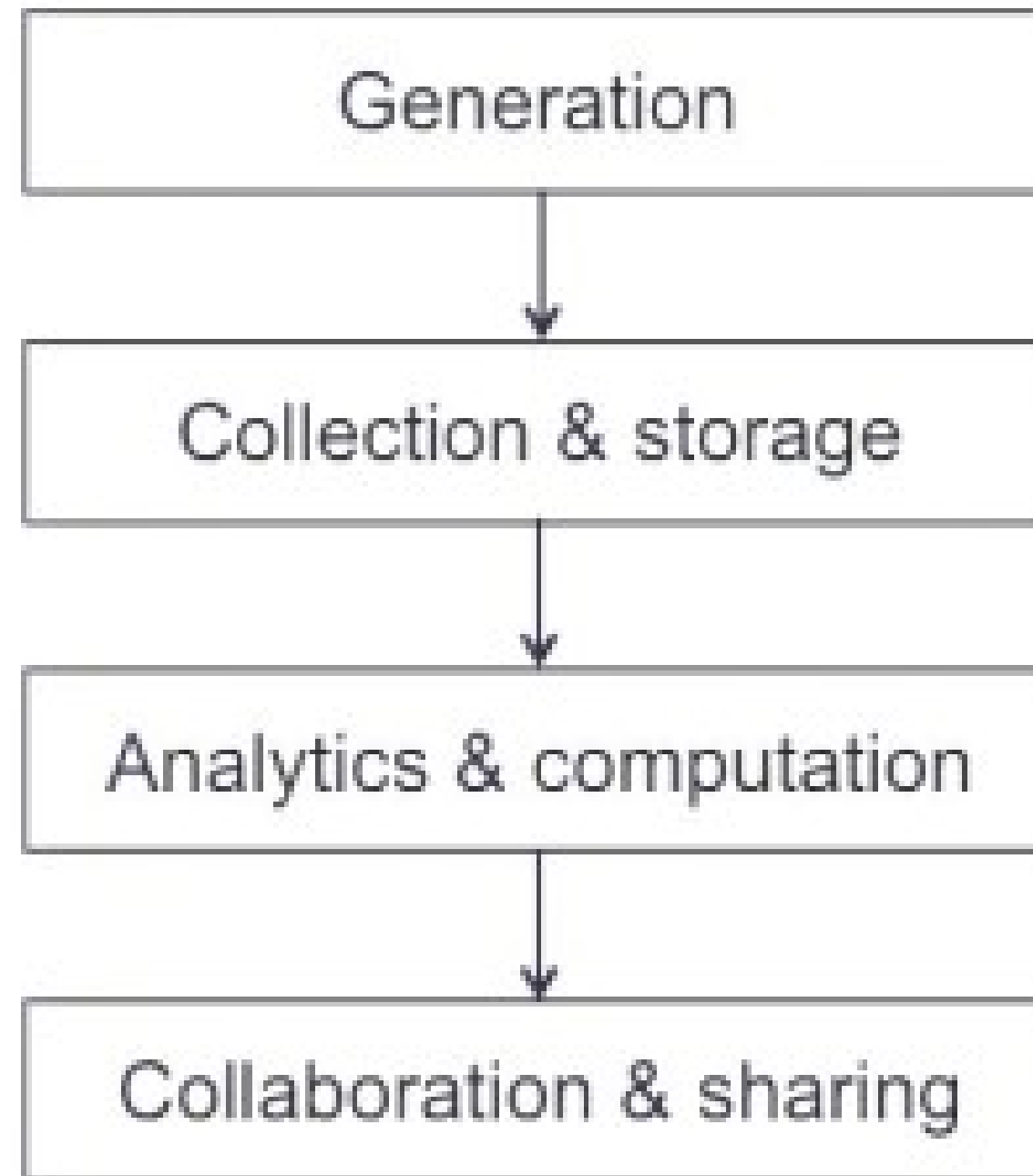
... in minutes. World wide. 100% Automated.



Big Data



Big Data/Analytics Life-Cycle



The cost of data generation
is **falling**



The volume of data
is **increasing**



Lower cost,
higher throughput



Generation



Collection & storage



Analytics & computation



Collaboration & sharing



Lower cost,
higher throughput



Generation



Collection & storage



Highly
constrained



Analytics & computation



Collaboration & sharing





Image: Don Ho Kim (Flickr), cc-by-sa
URL: <http://www.flickr.com/photos/DonHoKim/3288805122/>

Cloud Computing

Elastic & Highly Scalable

+

No capital expense

+

Pay-per-use

+

On-demand

+

Globally Available

= Remove constraints



\$0

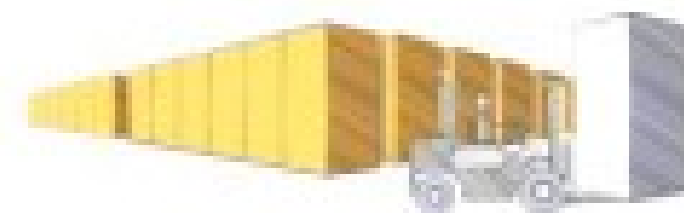
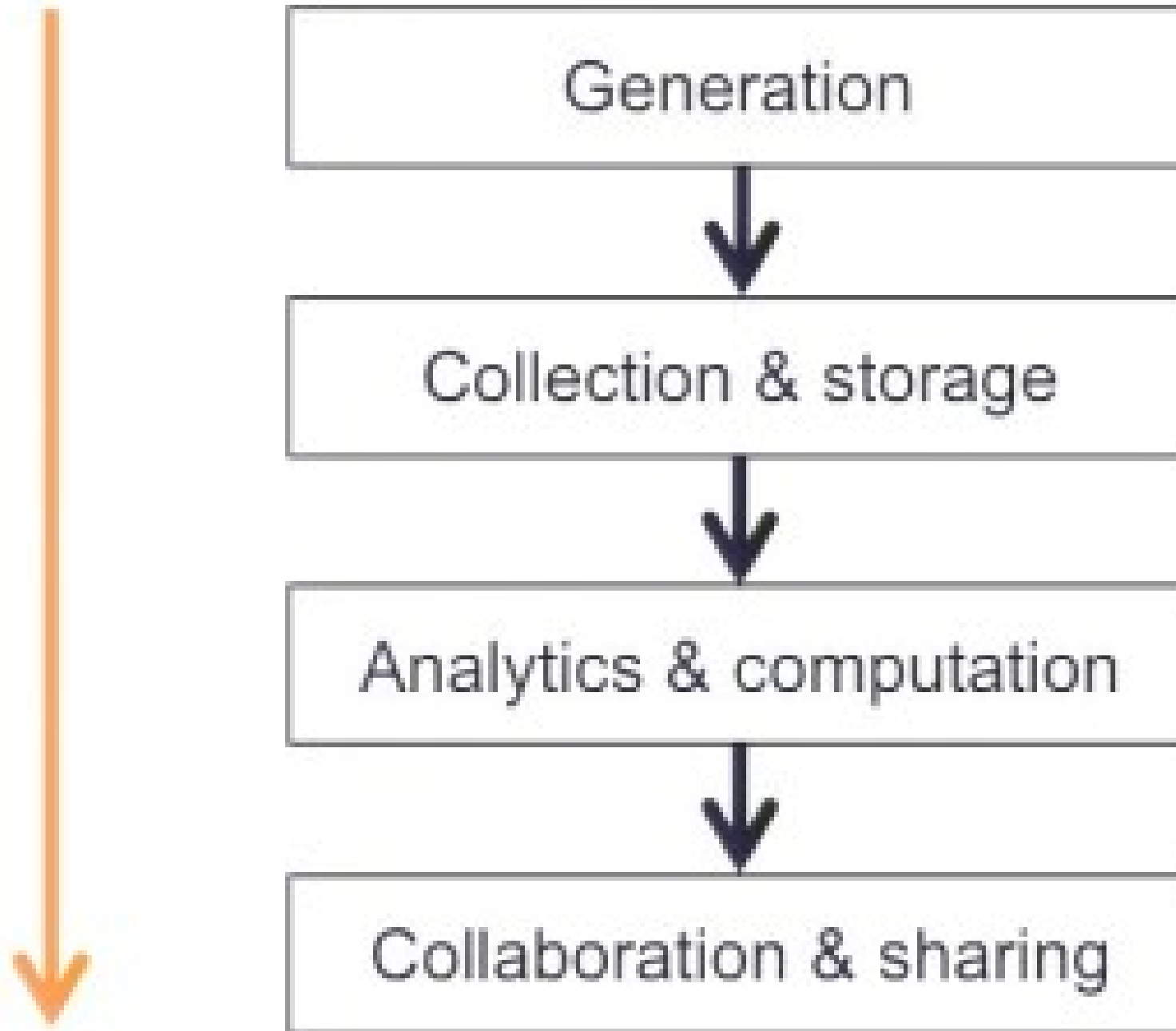




Image: Christian B. Reiger, Flickr, cc-by-2.0
Photos: iStock.com, iStock.com, iStock.com, iStock.com, iStock.com, iStock.com

Accelerated



razorfish

“who buys video games?”



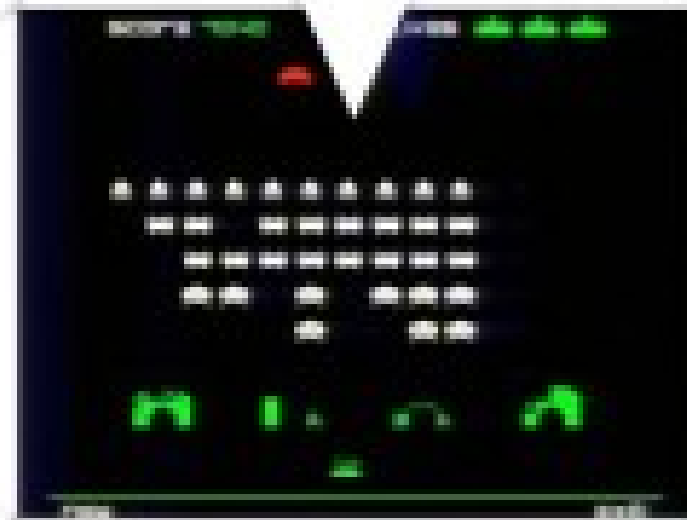
Per day:

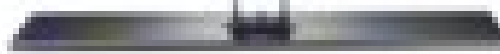
3.5 billion records

13 TB of click stream logs

71 million unique cookies







Results:

500% return on ad spend
From 2 months procurement
time to a minute



How?



Amazon Simple Storage Service (S3)

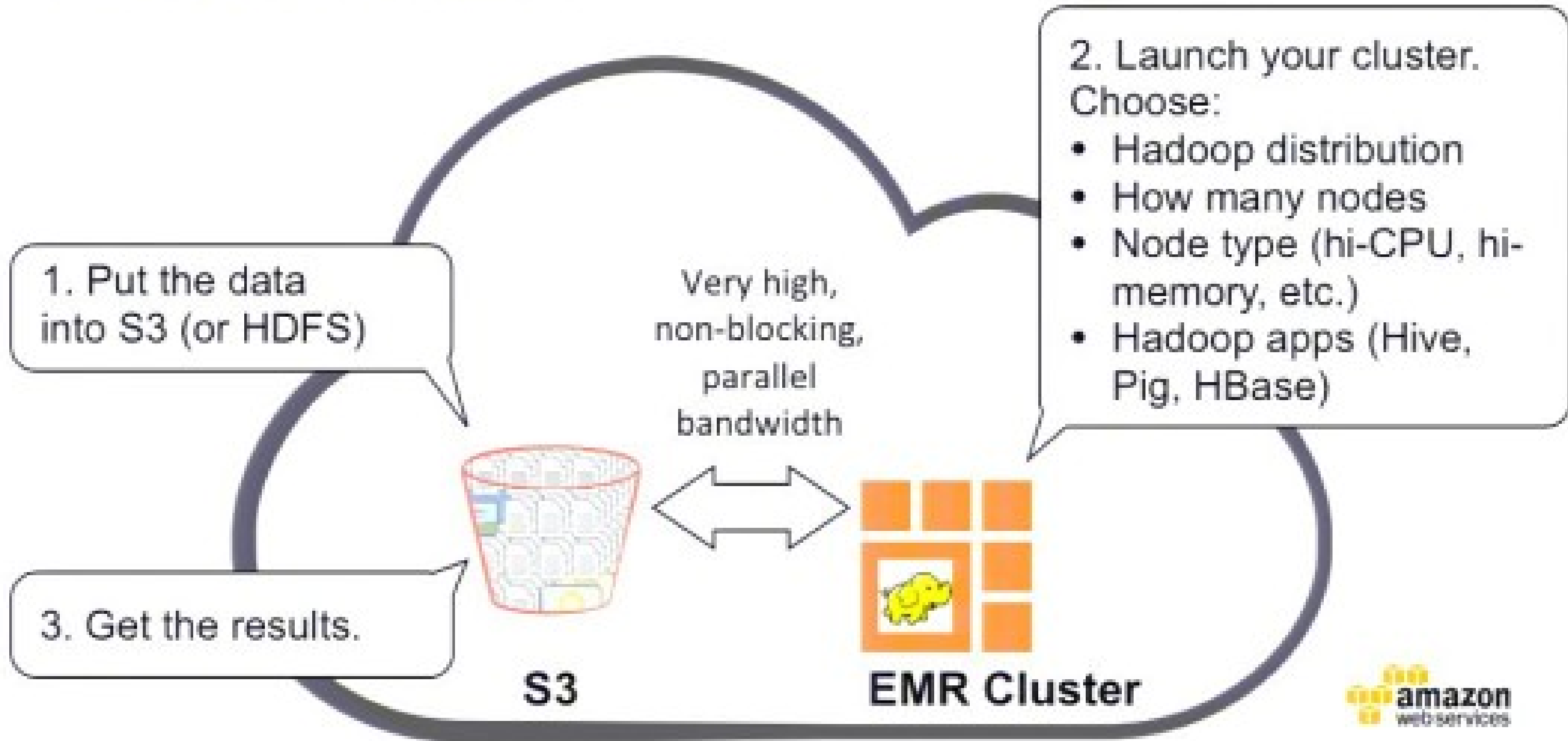


Amazon Simple Storage Service (S3)

- Durability:
99.9999999999%
- Cost:
\$0.03/GB/Month
- Capacity:
unlimited



How does ist work?



How does it work?

```
./elastic-mapreduce --create --stream \  
--ami-version 3.0.3 \  
--instance-type m1.small \  
--num-instances 100 \  
--arg "-files" \  
--arg "s3://elasticmapreduce/samples/wordcount/wordSplitter.py" \  
--input s3n://elasticmapreduce/samples/wordcount/input \  
--mapper wordSplitter.py \  
--reducer aggregate \  
--output s3n://myawsbucket/output/2014-01-16
```

How does it work?

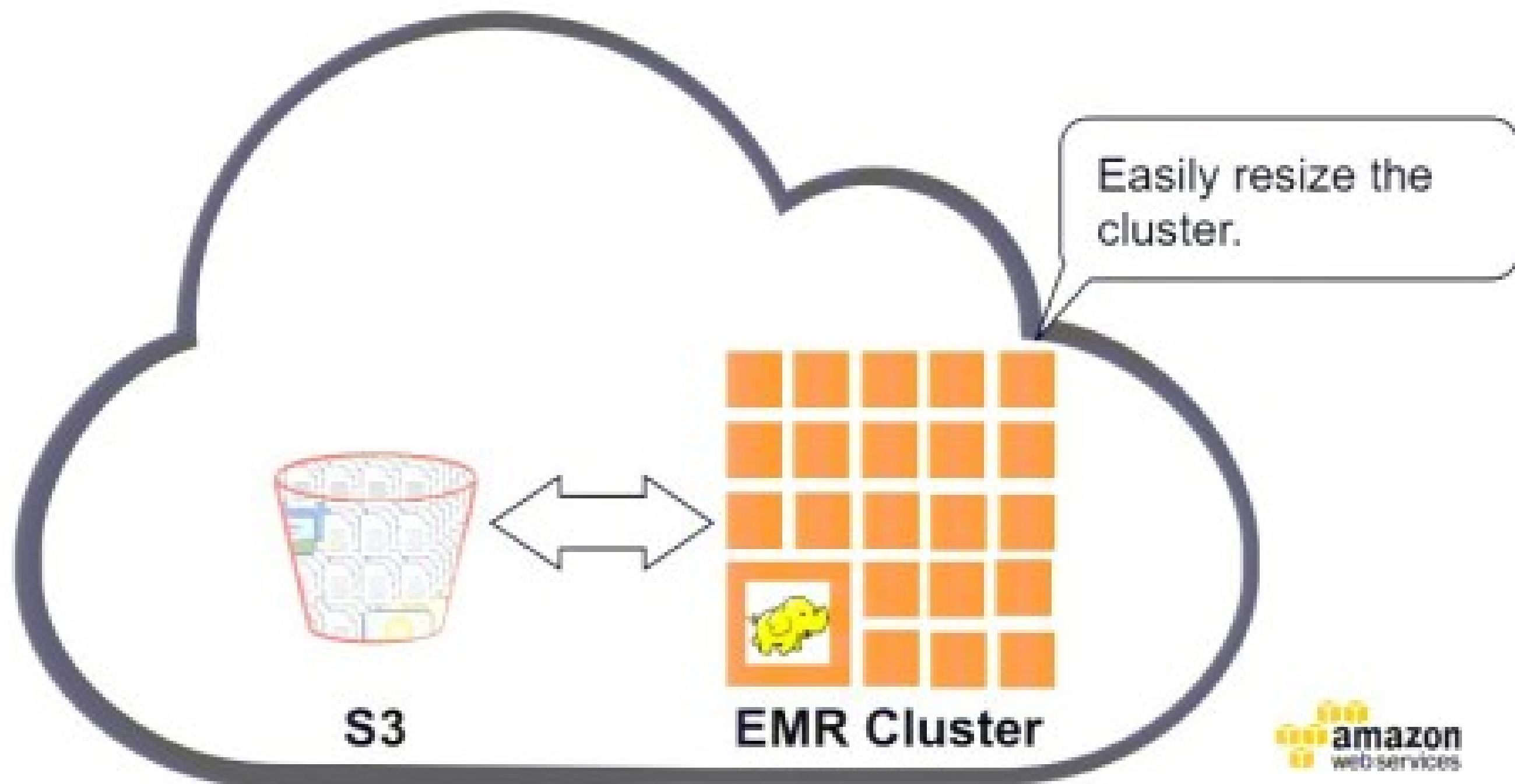
The screenshot shows the 'Create Cluster' page in the Amazon EMR console. The page is divided into several sections for configuring the cluster:

- Cluster Configuration:**
 - Cluster name:** My cluster
 - Termination protection:** Yes (checked)
 - Logging:** Enabled (checked)
 - Log folder ID location:** s3://mybucket-name-us-east-1
 - Debugging:** Enabled (checked)
- Tags:**
 - Optional: Add up to 10 tags to your EMR cluster. A tag consists of a case-sensitive key-value pair. Tags on EMR clusters are propagated to the underlying EC2 instances. [Learn more about tagging your Amazon EMR clusters.](#)
 - Table with columns: Key, Value (optional)
- Software Configuration:**
 - Hadoop distribution:** Amazon (checked)
 - AMI version:** 2.4.2.Hadoop3.3.3.1-ami
 - MapR:** (checked)
- Applications to be installed:**

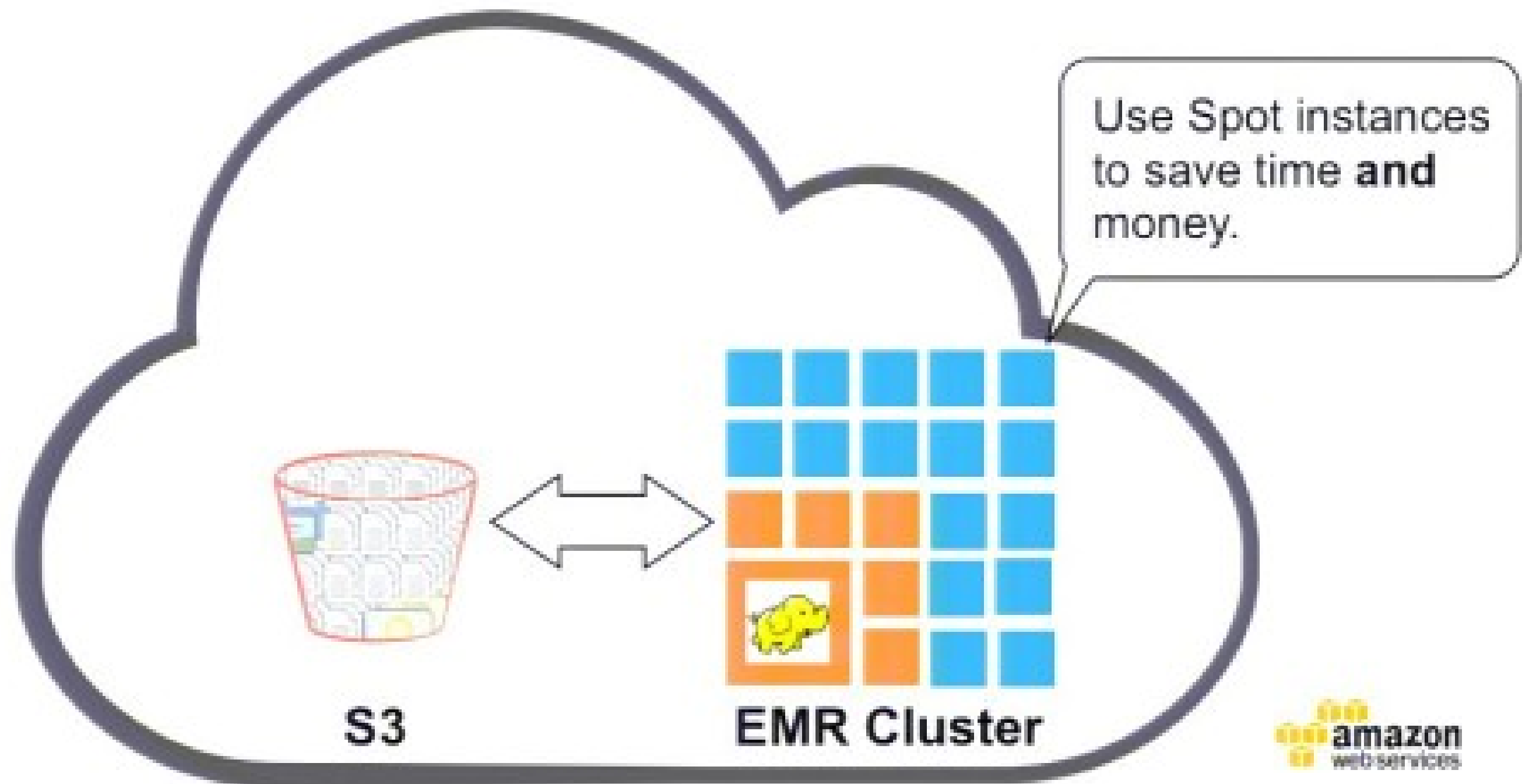
Applications to be installed	Version
Hive	2.11.0.1

On the right side, there is a 'Configure sample application' section with links for 'Prevents accidental termination of the cluster to shut down the cluster - you must turn off terminated protection. [Learn more](#)' and 'Copy the cluster's log files automatically to S3. [Learn more](#)'. Below that, there is a link for 'View logs to enable console debugging functionality (requires logging). [Learn more](#)'.

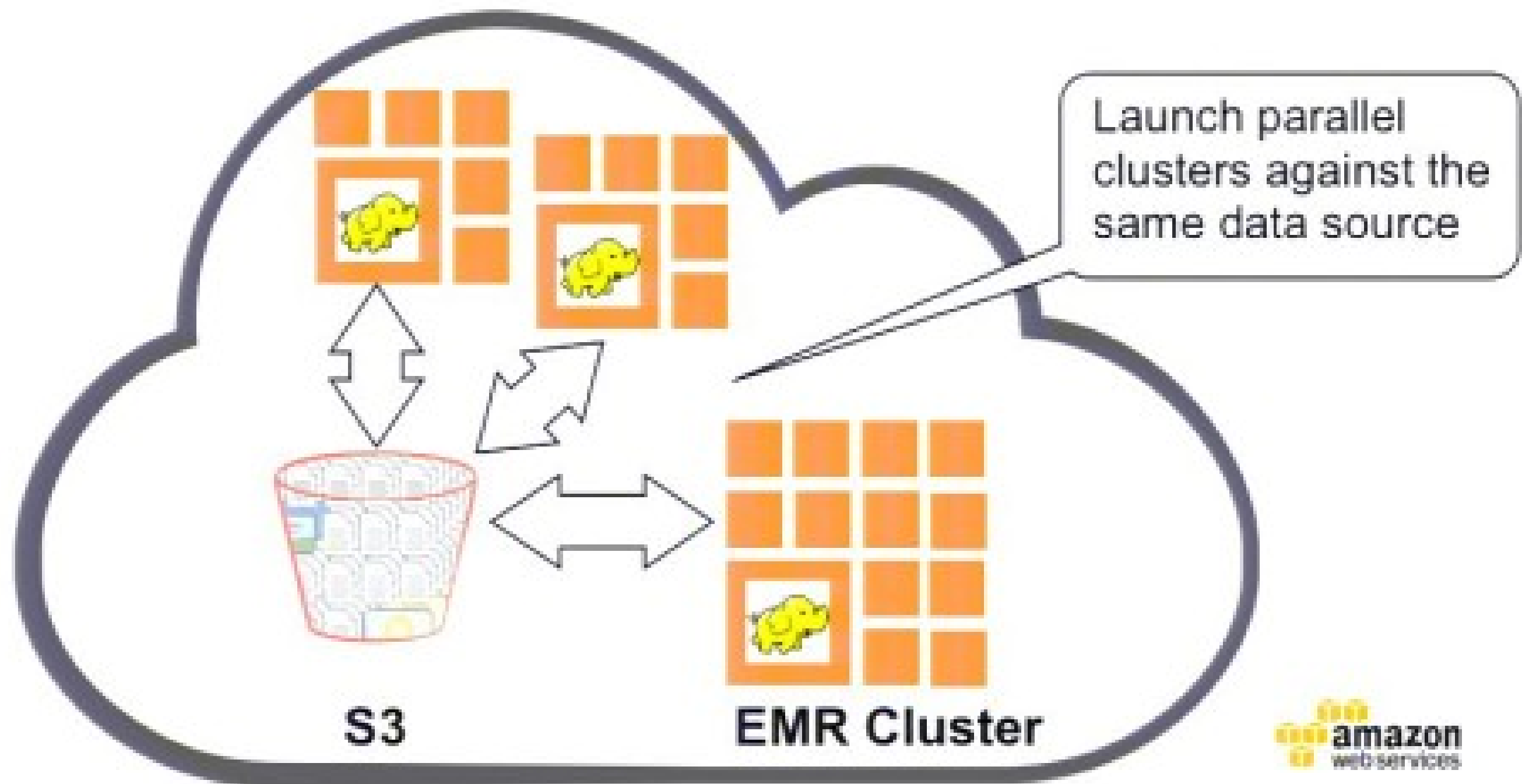
How does ist work?



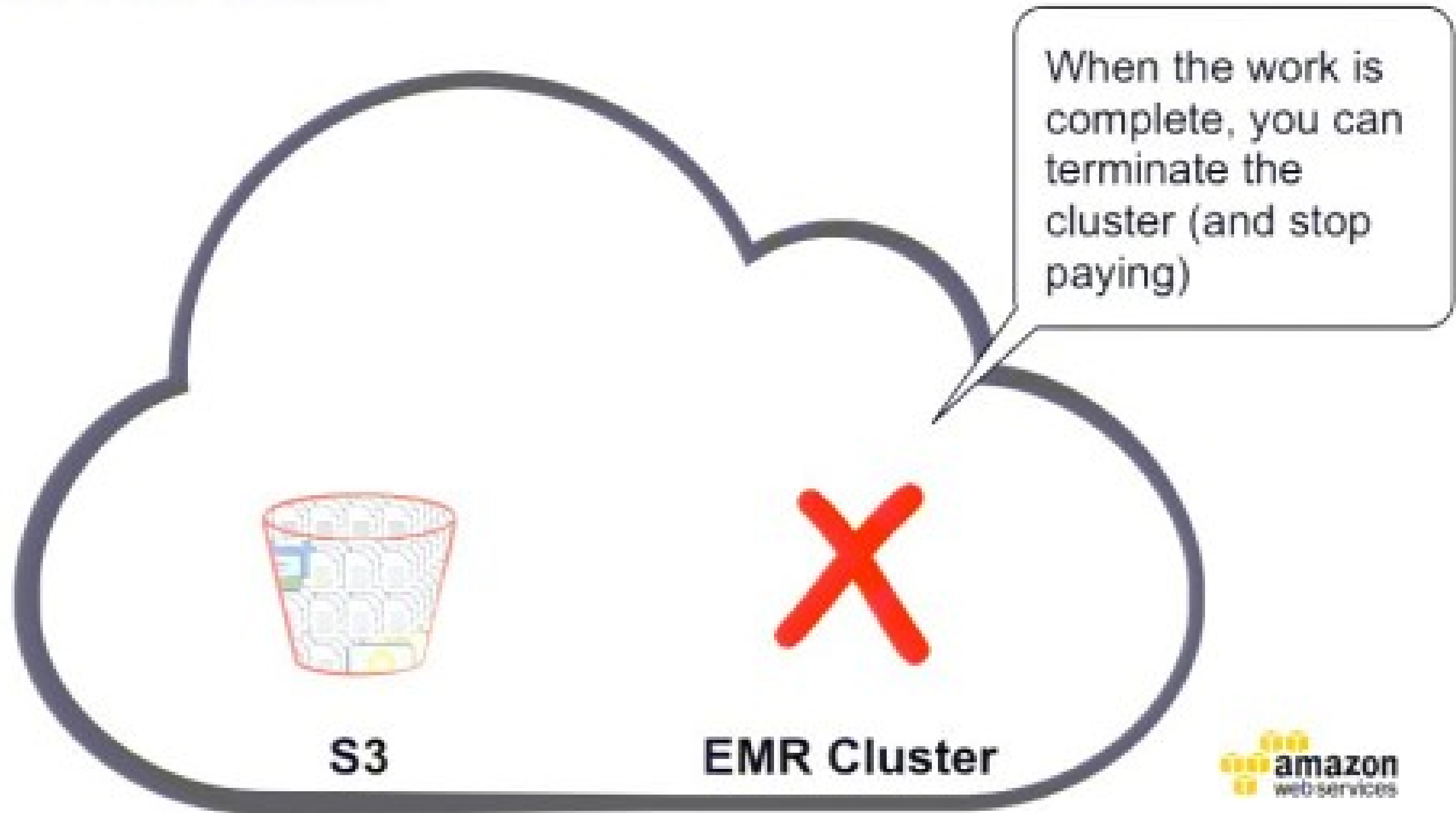
How does it work?



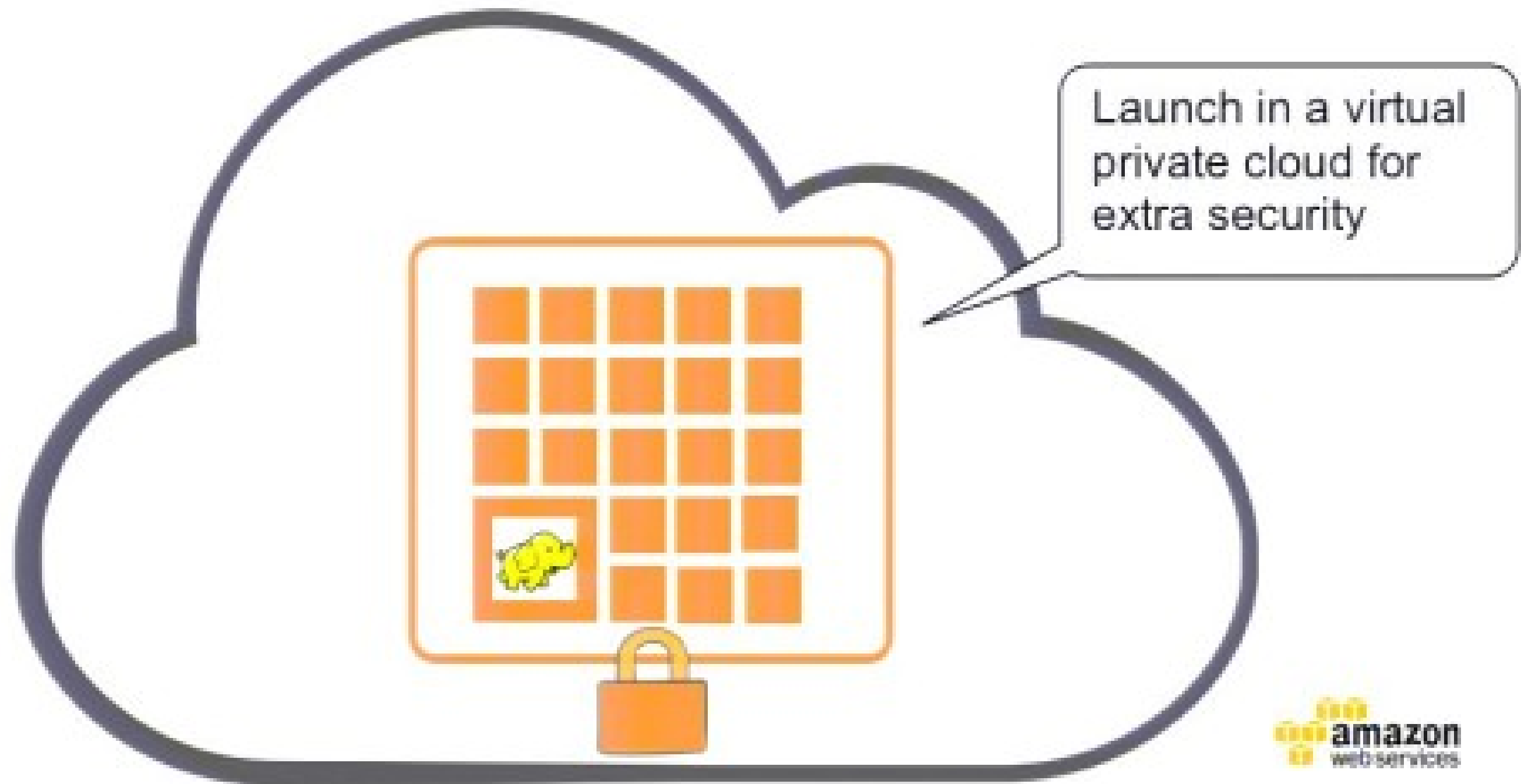
How does it work?



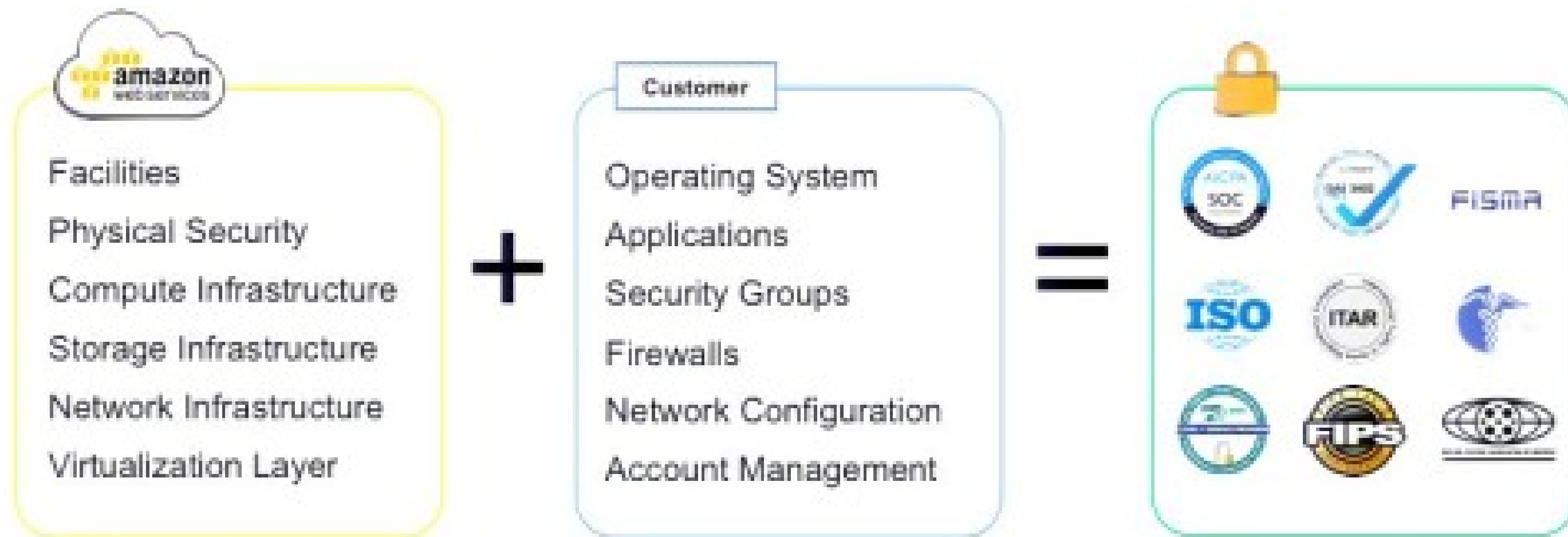
How does it work?



How does ist work?



Shared Responsibility for Security & Compliance



You keep **100% control** over your data

Thousands of Customers, 5+ Millions of Clusters



Cost to run a 100 node Elastic MapReduce Cluster

Starting at EUR 4.24/hour
(\$5.80/h)



Cost to run a 100 node Elastic MapReduce Cluster



Photos: www.flickr.com/photos/51018831@md8/5385664961/
Calgary Reviews: www.flickr.com/photos/calgaryreviews/5328300248/



What Would
YOU
Do with a
100 Node
Supercomputer?



#64 on the November Top 500 List

- Intel Xeon 2.8 GHz
- 26,496 Cores
- 1656 Instances
- **\$3166.272** per hour
- 484.2 TFlops



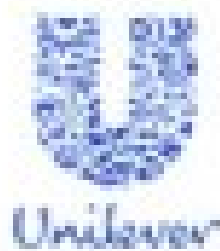
The screenshot shows the Top500 website interface. At the top, there is a navigation menu with links for Home, About, Features, Lists, Statistics, and Contact. Below the navigation, the page title is "Top500 List - November 2013". There are two explanatory paragraphs: one stating that R_{max} and R_{peak} values are in TFlops and directing users to the TOP500 description for more details; another stating that R_{peak} values are for the normal CPU clock rate and suggesting users take the Turbo CPU clock rate into account.

	Institute of Process Engineering, Chinese Academy of Sciences China	Mole-8.8 - Mole-8.8 Cluster, Intel X5820 AC 2.27 GHz, Interconnect QDR, NVIDIA 2080 IPE, Nvidia, Tyan	29440	496.6	1012.6	540
	Amazon Web Services United States	Amazon EC2 C3 Instance cluster - Amazon EC2 Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, 10G Ethernet Self-made	26496	484.2	593.5	
	United Kingdom Meteorological Office United Kingdom	Power 776, POWER7 80 3.930GHz, Custom Interconnect IBM	18432	476.0	505.6	5040

Biotech Customer quote:

“Getting onto the Top 500 list is now cheaper than running an ad on the Frankfurter Allgemeine!”





Biology and Informatics research

"The key advantage that AWS has over running this workflow on Unilever's existing cluster is the ability to scale up to a much larger number of parallel compute nodes on demand."

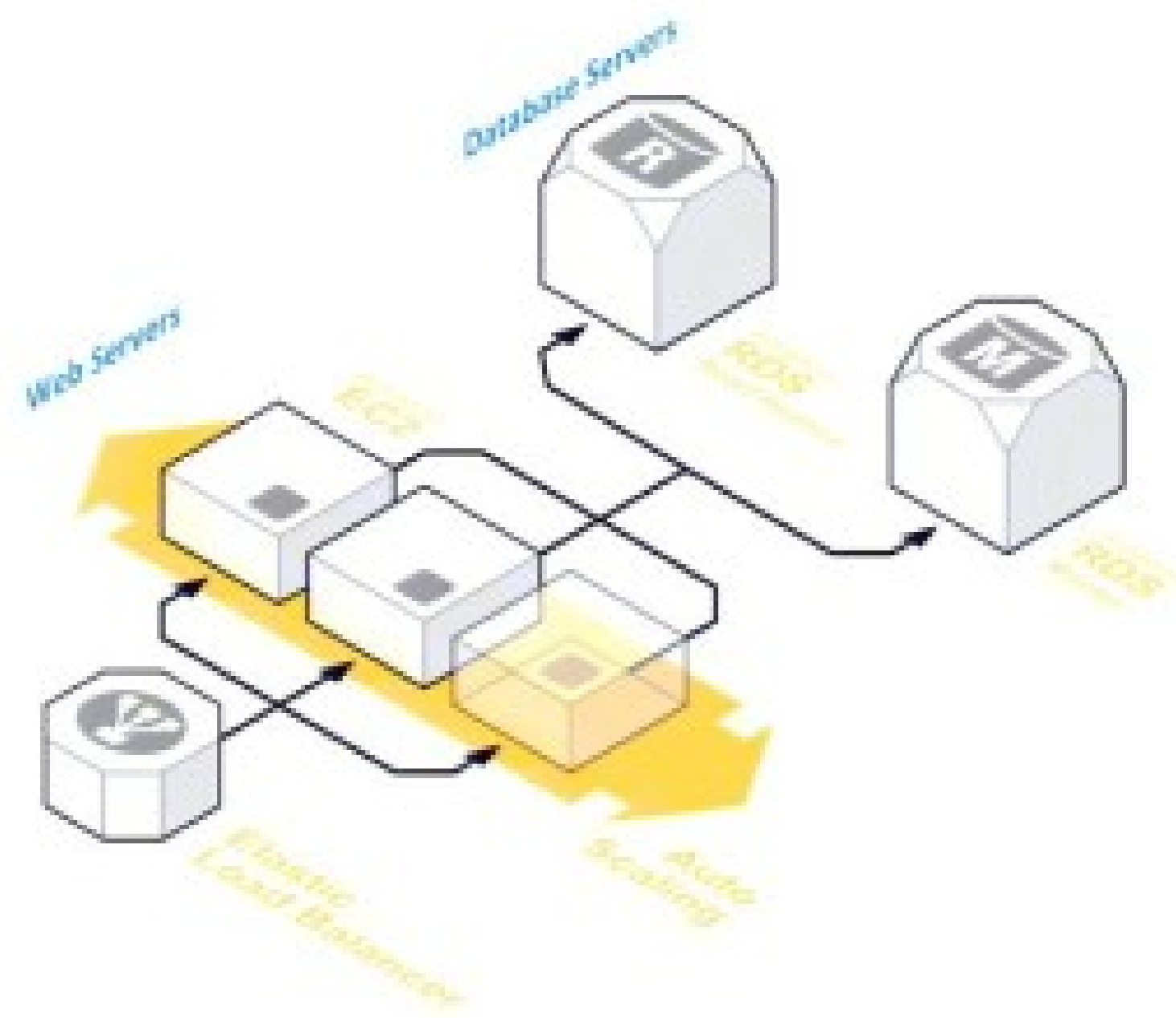
Pete Keeley, Unilever Research's eScience IT Lead for Cloud Solutions

Unilever's digital data program now
processes genetic sequences **twenty times
faster**

“But I’m not a Big Data/HPC guy...”



Web Development



Load Testing?

newsapps / beeswithmachineguns

Star 848 Fork 130

A utility for arming (creating) many bees (micro EC2 instances) to attack (load test) targets (web applications). <http://apps.chicagotribune.com/>

108 commits

1 branch

3 issues

14 contributors

Code

Issues

Pull Requests

Pages

Graphs

Network

HTTPS clone URL

<https://github.com/newsapps/beeswithmachineguns>

You can clone with HTTPS or Subversion

Clone in Desktop

Download ZIP

Switch branches/tags

Merge pull request #75 from abannafixed_header_bug

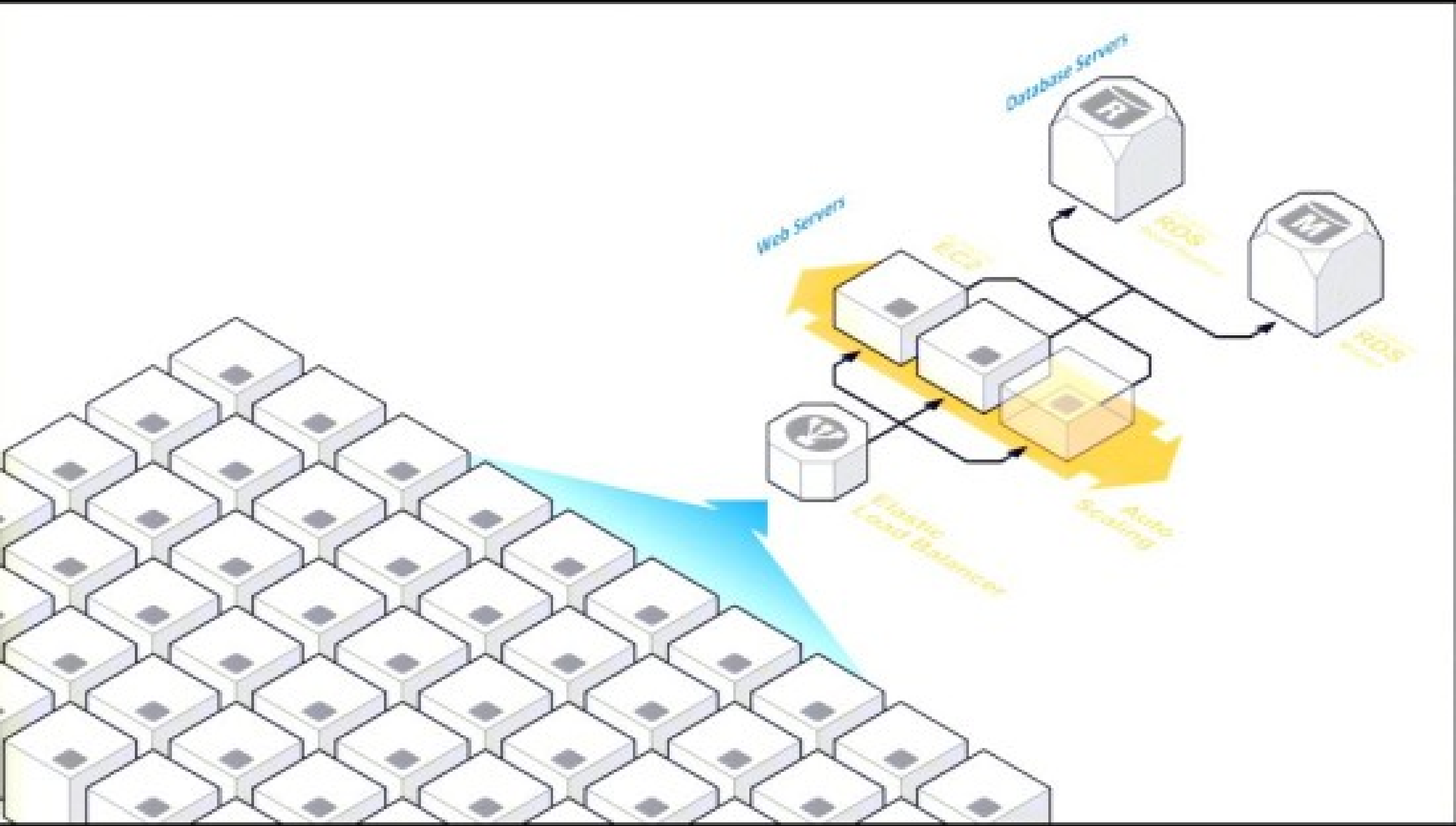
commit 13 days ago

Latest commit: f8d22ec

- beeswithmachineguns: closing up some white space. 14 days ago
- .gitignore: merging in cat_mediator branch. 5 months ago
- COPYING: Updated COPYING. 3 years ago
- README.txtile: Bees now get a simple Name tag on creation. Closes #70. 2 years ago
- bees: Wrap main main() in if __name__ == '__main__'. 2 years ago
- requirements.txt: Updated requirements to a later version. 8 months ago
- setup.py: Updated requirements to a later version. 8 months ago

README.txtile

Bees with Machine Guns!



```
$ fab -f bees.py up:2
Connecting to the hive.
Attempting to call up 2 bees.
Waiting for bees to load their machine guns...
.
.
.
.
Bee 1-df450ab4 is ready for the attack.
.
Bee 1-d4450ab6 is ready for the attack.
The swarm has assembled 2 bees.
Done.

$ fab -f bees.py attack:http://YOUR-DOMAIN.com/,1000,100
Read 2 bees from the roaster.
Connecting to the hive.
Assembling bees.
Each of 2 bees will fire 500 rounds, 10 at a time.
Stinging URL so it will be cached for the attack.
[localhost] run: curl http://YOUR-DOMAIN.com/ >> /dev/null
Organizing the swarm.
Bee 0 is joining the swarm.
Bee 1 is joining the swarm.
Bee 0 is firing his machine gun. Bang bang!
Bee 1 is firing his machine gun. Bang bang!
Bee 0 is out of ammo.
Bee 1 lost sight of the target (connection timed out).
Offensive complete.
  Target failed to fully respond to 1 bees.
  Complete requests:          500
  Requests per second:      81.670000 [#/sec] (mean)
  Time per request:         612.188000 [ms] (mean)
  50% response time:       158.000000 [ms] (mean)
  90% response time:       1841.000000 [ms] (mean)
Mission Assessment: Target successfully fended off the swarm.
The swarm is awaiting new orders.
Done.
```


EC2 Management Console | CloudWatch Management Console

amazon.com | amazon.com

Services | IAM | EC2 | Elastic MapReduce | RDS | DynamoDB | [arn:aws:iam::123456789012:role/...](#)

Region: us-east-1

Viewing: EC2 Instance Metrics | Search

EC2 Instance Metrics

Instance ID	Instance Name
i-12345678	CPURunner
i-12345678	DiskReadBytes
i-12345678	DiskReadOps
i-12345678	FreeSpaceMB

CPUUsage |
 CPUUsage |
 DiskReadBytes |
 DiskReadOps |
 FreeSpaceMB

Statistic: Maximum | Period: 1 hour

Time Range: 2016-11-16 00:00 to 2016-11-16 00:00

Absolute |
 Relative

From: 0 | 4 days ago |
 To: 0 | 1 hour ago

Green |
 Yellow |
 Red

100% CPUUsage
 90%
 80%
 70%
 60%
 50%
 40%
 30%
 20%
 10%
 0%

2016-11-16 00:00 2016-11-16 00:00 2016-11-16 00:00 2016-11-16 00:00

arn:aws:iam::123456789012:role/...

© 2016-2022, Amazon Web Services, Inc. or its affiliates. All rights reserved. | Feedback | Support | Privacy Policy | Terms of Use | An Amazon.com company

CALL OF DUTY
MW3



USING AMAZON EC2 TO SIMULATE

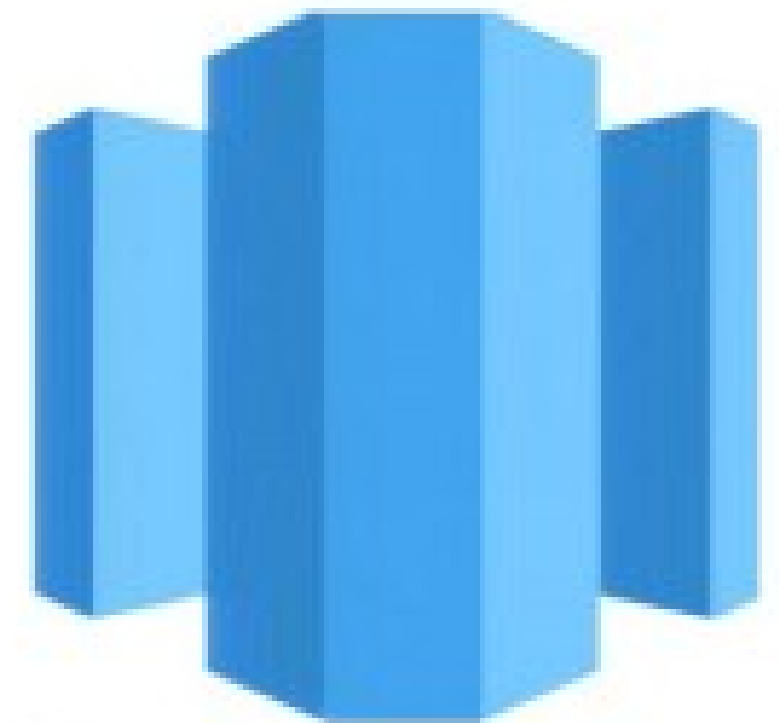
2.4 MILLION PLAYERS

Data Warehousing



Amazon Redshift

- **Easy** to provision, scale, operate
- No up-front costs, **pay-as-you-go**
- **Fast**: Columnar storage, compression, Specialized nodes
- **1-100** node clusters **2TB - 1.6PB**
- **\$999** per TB per year
- Transparent **backups, restore, failover**
- **Security** in transit, at rest, for backups, VPC



Data Warehousing the AWS way



Amazon Redshift Results

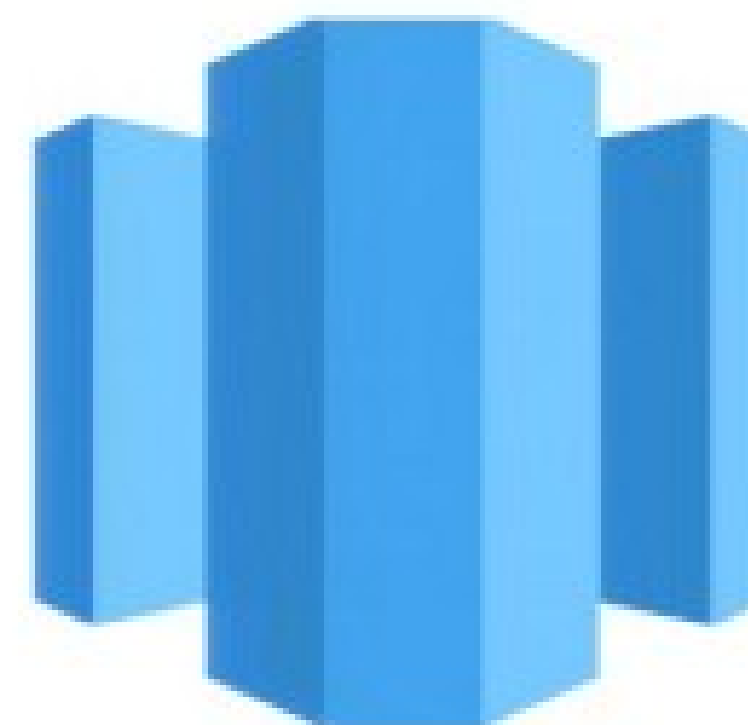
amazon.com.

Current production environment:

32 nodes, 128 CPUs, 4.2TB RAM, 1.6 PB disk

Tested 2B row data set, 6 representative

queries on a 2-node Amazon Redshift cluster



Result: 10x faster



Shop All Categories ▾

Search AWS Marketplace

GO

▸ Your Software



SAP HANA One

on SUSE Linux Enterprise Server 11 SP1

Sold by SAP

Perform real-time analysis, develop and deploy real-time applications with the SAP HANA One, an instance of SAP HANA in-memory platform on AWS. Natively built on in-memory technology and now deployed on AWS, SAP HANA One is designed to accelerate transactional processing, operational reporting, OLAP, predictive and text analysis while by-passing traditional data latency & maintenance issues created through pre-materializing views and pre-caching query results. Unlike other database management systems in the market today, the SAP HANA One on AWS streamlines both transactional (OLTP) and analytical

[Read more](#)

Customer Rating ★★★★★ (3 Customer Reviews)

Latest Version Rev (2) 1 (Other available versions)

Base Operating System Linux/Unix, SUSE Linux Enterprise Server 11 SP1

Delivery Method 64-bit Amazon Machine Image (AMI) [Learn more](#)

Support [See details below](#)

AWS Services Required Amazon EC2, Amazon EBS

- Highlights**
- SAP HANA One transforms decision processing by streamlining transactions, analytics, planning, predictive and text analysis on a single in-memory platform running on AWS so business can operate in real-time.
 - Connectivity to SAP HANA One is provided via open standards including ODBC, JDBC, ODBC, ODATA and MDX.
 - Users can provision SAP HANA One by themselves and pay as they go to deploy applications on AWS's highly available and scalable platform for applications, with no up-front

Continue

You will have an opportunity to review your order before purchasing or being charged.

Pricing Details

For region: ▾

Hourly Fees (includes SUSE Linux Enterprise Server 11 SP1)

Total hourly fees will vary by instance type and EC2 region.

EC2 Instance Type	Software	EC2	Total
Custom Compute (M5, m5.2xlarge)	\$1.90/hr	\$2.50/hr	\$4.40/hr

EBS Storage Fees ⓘ

\$0.10 / GB / Month for Standard EBS Storage

Reserves On-Demand EC2 pricing unless for Reserved and Spot instances, which will be lower. [See pricing details.](#)

The software is built on a version of EC2 that includes SUSE Linux Enterprise Server 11 SP1 ⓘ

Data transfer fees not included. ⓘ





Scalability test results for SAP Business Warehouse using In-Memory Data Fabric

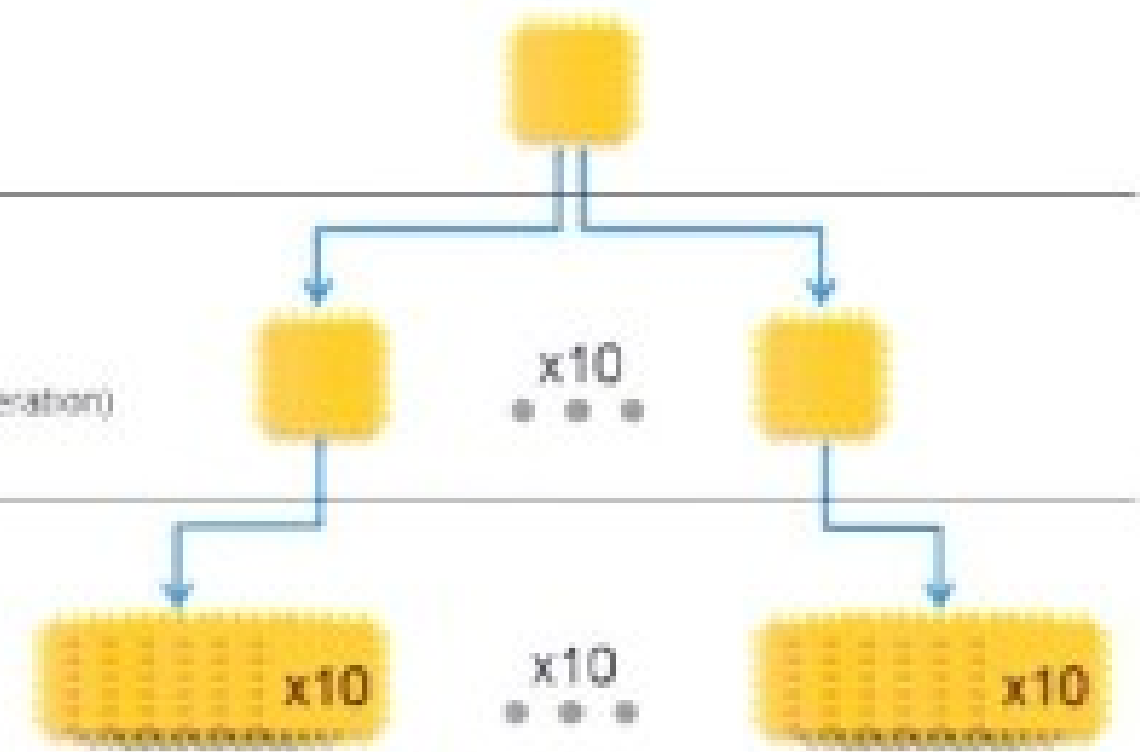
Posted by [Suresh Kumar Reddy Venkatasaiam](#) [Blog](#) on Mar 26, 2014 6:38:13 AM

- 8 Million rows per second loaded on 100 SAP HANA Instances
- 60 Billion rows analyzed in 330 milliseconds
- Near linear scalability

Level 2 Federation
(1 Main HANA Instance)

Level 1 Federation
(10 HANA Instances for federation)

Data Instances
(100 HANA Instances)



But:

It's Not About Size...





"On the Internet, nobody knows you're a dog."



It's about:

Democratization of IT



And...



“Want to increase innovation?
Lower the cost of failure.”

– Joi Ito



Experiment
often



Fail quickly,
at low cost



More
Innovation



S&P Capital IQ

THE SCALABLE NATURE OF THE CLOUD ALLOWS US TO FIRE UP CLUSTERS OF COMPUTERS TO PROCESS DATA USING HADOOP, AND THOSE CLUSTERS CAN BE CONFIGURABLE IN TERMS OF THE NUMBER OF NODES, THE AMOUNT OF STORAGE AND THE AMOUNT OF PROCESSING POWER

- Jeff Stenberg
Chief Data Scientist

S&P
CAPITAL IQ



Provides data to
4200+ top global
investment firms

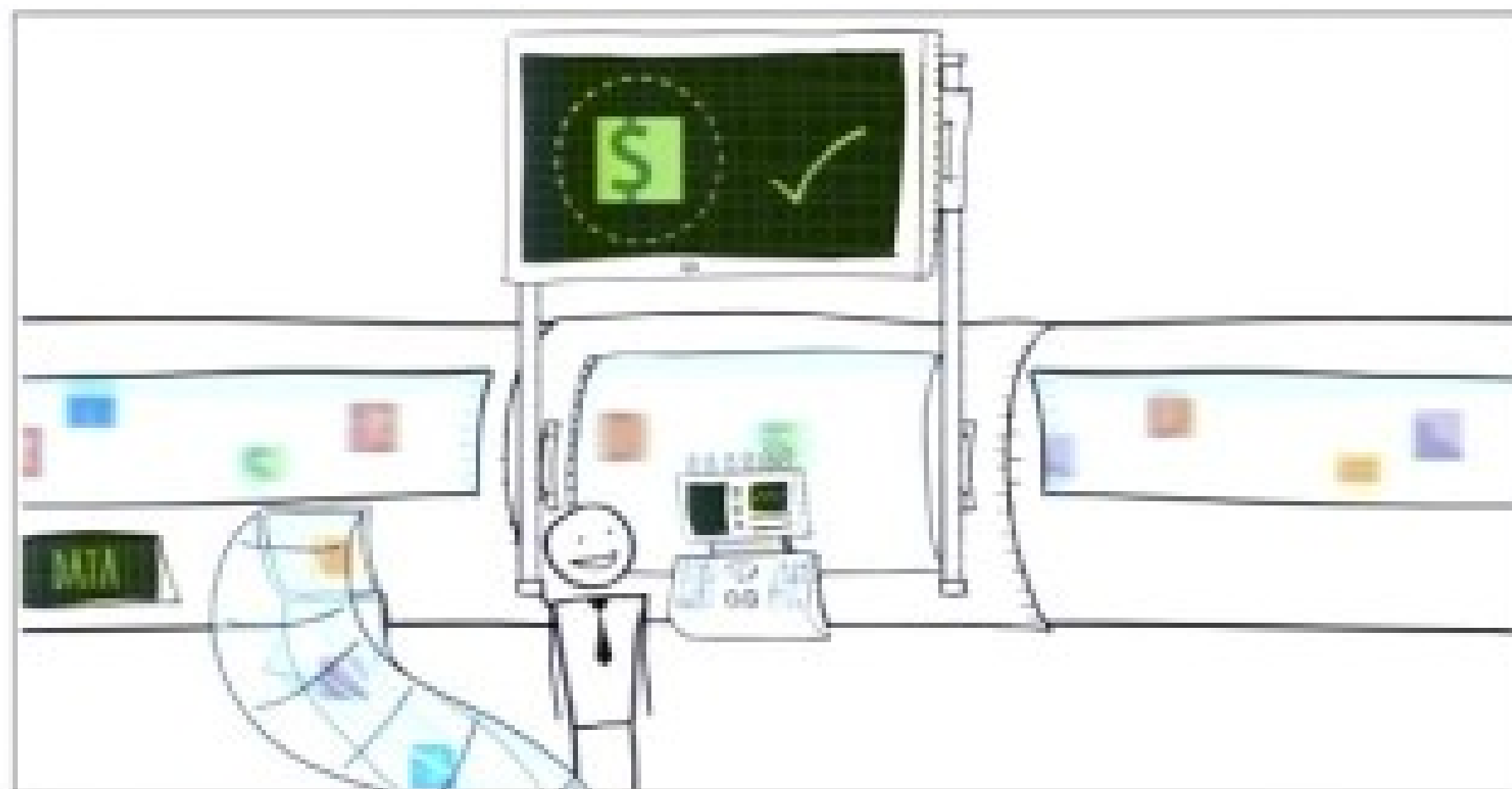
Launched Hadoop
faster, Learned
Hadoop faster



Real Time

New: Amazon Kinesis

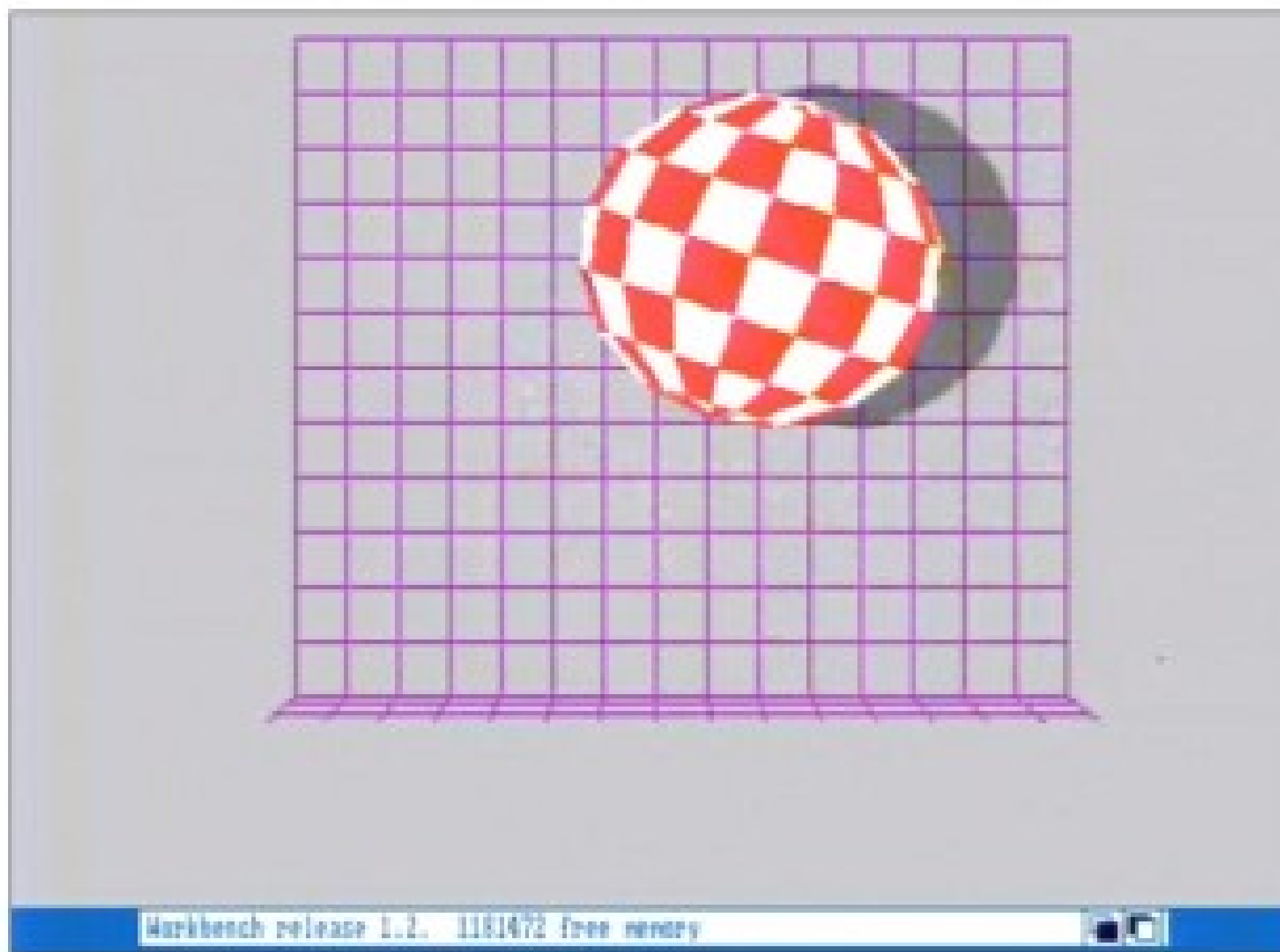
- Real-time processing
- Massive scale
- Integrated
- Use cases:
 - Real-time log analysis
 - Sensor data analytics
 - Social media monitoring
 - Financial transactions
 - Online machine learning



Amazon Kinesis Data Flow



Demo



Amazon Kinesis Stream List > Stream Details

Kinesis Help

Delete Stream

Summary for KinesisDemo-HttpReferrerPairs

Open Shards: 3

Open shards both receive data from producers and make data available to consumers. This is the default state for a shard.

Closed Shards: 1

Closed shards no longer accept data from producers. Data in these shards is still available to consumers, but only for 24 hours. Resharding operations result in closed shards.

Monitoring

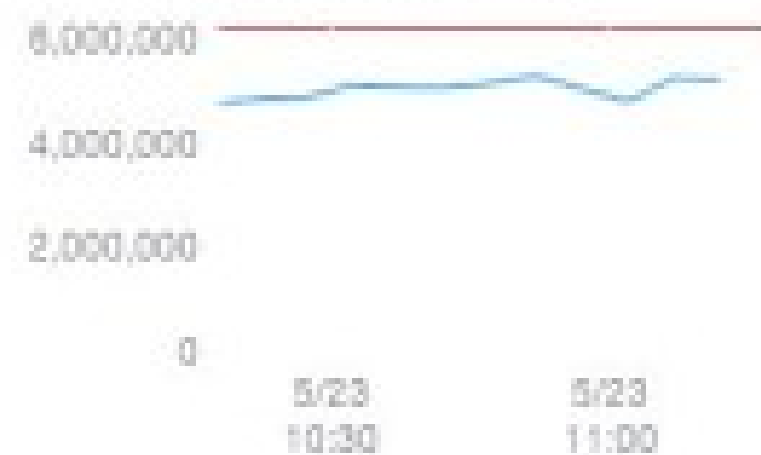
Time range: Last hour All graphs are displayed in the UTC time zone.



Write Capacity (Bytes/second)



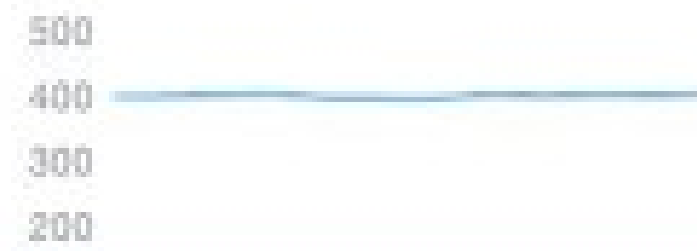
Read Capacity (Bytes/second)



Put Latency (Milliseconds)



Get Latency (Milliseconds)



Amazon Kinesis

Streams List

Kinesis Help

Create Stream

Delete Stream

Filter:

Stream Name

Number of Shards

Status

Amazon Kinesis

Stream List

Kinesis Help

Create Stream

Delete Stream

Filter:



Stream Name	Number of Shards	Status
KinesisDemo-HttpReferrerPairs	3	ACTIVE

Amazon Kinesis Stream List > Stream Details

Kinesis Help

Delete Stream

Summary for KinesisDemo-HttpReferrerPairs

Open Shards: 3

Open shards both receive data from producers and make data available to consumers. This is the default state for a shard.

Closed Shards: 0

Closed shards no longer accept data from producers. Data in these shards is still available to consumers, but only for 24 hours. Resharding operations result in closed shards.

Monitoring

Time range: Last hour All graphs are displayed in the UTC time zone.



Write Capacity (Bytes/second)



Read Capacity (Bytes/second)



Put Latency (Milliseconds)



Get Latency (Milliseconds)



- EC2 Dashboard
- Events
- Tags
- Reports
- INSTANCES
 - Instances
 - Spot Requests
 - Reserved Instances
- IMAGES
 - AMIs
 - Bundle Tasks
- ELASTIC BLOCK STORE
 - Volumes
 - Snapshots
- NETWORK & SECURITY
 - Security Groups
 - Elastic IPs
 - Placement Groups
 - Load Balancers
 - Key Pairs
 - Network Interfaces

Launch Instance Cancel Actions

Filter: All instances All instance types Search Instances

1 to 5 of 5 instances

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm
KinesisDemo	i-a9922afa	m1.large	us-east-1b	running	2/2 checks...	None

Select an instance above

- EC2 Dashboard
- Events
- Tags
- Reports
- INSTANCES
 - Instances
 - Spot Requests
 - Reserved Instances
- IMAGES
 - AMIs
 - Bundle Tasks
- ELASTIC BLOCK STORE
 - Volumes
 - Snapshots
- NETWORK & SECURITY
 - Security Groups
 - Elastic IPs
 - Placement Groups
 - Load Balancers
 - Key Pairs
 - Network Interfaces

Resources

You are using the following Amazon EC2 resources in the US East (N. Virginia) region:

5 Running Instances	3 Elastic IPs
5 Volumes	2 Snapshots
6 Key Pairs	0 Load Balancers
0 Placement Groups	21 Security Groups

Focus on application development and offload database management to AWS - [Try Amazon RDS Now!](#) Hide

Create Instance

To start using Amazon EC2 you will want to launch a virtual server, known as an Amazon EC2 instance.

[Launch Instance](#)

Note: Your instances will launch in the US East (N. Virginia) region

Service Health

Service Status:

US East (N. Virginia):
This service is operating normally

Availability Zone Status:

Scheduled Events

US East (N. Virginia):
No events

Account Attributes

Supported Platforms

- EC2
- VPC

Additional Information

- [Getting Started Guide](#)
- [Documentation](#)
- [All EC2 Resources](#)
- [Forums](#)
- [Pricing](#)
- [Contact Us](#)

AWS Marketplace

Find free software trial products in the AWS Marketplace from the [EC2 Launch Wizard](#).

Or try these popular AMIs:

- [Vyatta Virtual Router/Firewall/VPN](#)
- Provided by Vyatta, Inc.
- Rating
- Pay by the hour for software and AWS usage

- EC2 Dashboard
- Events
- Tags
- Reports
- INSTANCES
 - Instances
 - Spot Requests
 - Reserved Instances
- IMAGES
 - AMIs
 - Bundle Tasks
- ELASTIC BLOCK STORE
 - Volumes
 - Snapshots
- NETWORK & SECURITY
 - Security Groups
 - Elastic IPs
 - Placement Groups
 - Load Balancers
 - Key Pairs
 - Network Interfaces

Launch Instance Cancel Actions

Filter: **All instances** **All instance types** Search Instances

<< 1 to 5 of 5 instances >>

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm
KinesisCons...	i-3574ad65	m1.small	us-east-1a	running	2/2 checks...	None
KinesisCons...	i-9574cdc6	m1.small	us-east-1b	running	2/2 checks...	None
KinesisCons...	i-253bd977	m1.small	us-east-1d	running	2/2 checks...	None
KinesisDemo	i-20ff4173	m1.medium	us-east-1b	running	2/2 checks...	None
KinesisDemo	i-a9922afa	m1.large	us-east-1b	running	2/2 checks...	None

Select an instance above ☰ ☱ ☲

- EC2 Dashboard
- Events
- Tags
- Reports
- INSTANCES
 - Instances
 - Spot Requests
 - Reserved Instances
- IMAGES
 - AMIs
 - Bundle Tasks
- ELASTIC BLOCK STORE
 - Volumes
 - Snapshots
- NETWORK & SECURITY
 - Security Groups
 - Elastic IPs
 - Placement Groups
 - Load Balancers
 - Key Pairs
 - Network Interfaces

[Launch Instance](#)
Actions ▾
↻ ⚙ ⚠

Filter: All instances ▾ All instance types ▾

X

<< < 1 to 5 of 5 instances > >>

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm
KinesisCons...	i-3574ed65	m1.small	us-east-1a	running	2/2 checks...	None
KinesisCons...	i-9574cdc6	m1.small	us-east-1b	running	2/2 checks...	None
KinesisCons...	i-253bd977	m1.small	us-east-1d	running	2/2 checks...	None
KinesisDemo	i-20ff4173	m1.medium	us-east-1b	running	2/2 checks...	None
KinesisDemo	i-a9922afa	m1.large	us-east-1b	running	2/2 checks...	None

Instances: i-3574ed65 (KinesisConsumerApp), i-9574cdc6 (KinesisConsumerApp), i-253bd977 (KinesisConsumerApp)

Description	Status Checks	Monitoring	Tags
i-3574ed65: ec2-54-237-188-167.compute-1.amazonaws.com			
i-9574cdc6: ec2-23-22-5-5.compute-1.amazonaws.com			
i-253bd977: ec2-54-83-145-168.compute-1.amazonaws.com			

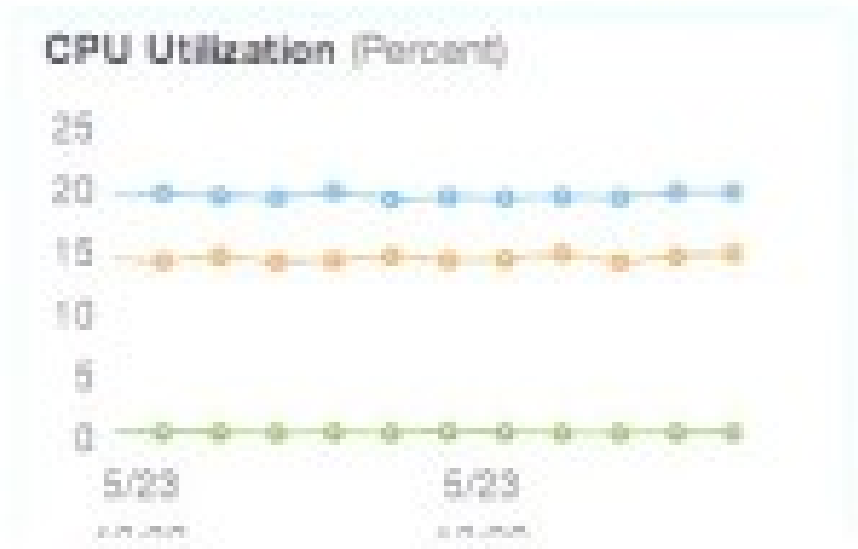
- EC2 Dashboard
- Events
- Tags
- Reports
- INSTANCES
 - Instances
 - Spot Requests
 - Reserved Instances
- IMAGES
 - AMIs
 - Bundle Tasks
- ELASTIC BLOCK STORE
 - Volumes
 - Snapshots
- NETWORK & SECURITY
 - Security Groups
 - Elastic IPs
 - Placement Groups
 - Load Balancers
 - Key Pairs
 - Network Interfaces

[Launch Instance](#)
Cancel
Actions

Filter: All instances All instance types

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm
KinesisCons...	i-3574ed65	m1.small	us-east-1a	running	2/2 checks...	None
KinesisCons...	i-9574cdc6	m1.small	us-east-1b	running	2/2 checks...	None
KinesisCons...	i-253bd977	m1.small	us-east-1d	running	2/2 checks...	None
KinesisDemo	i-20ff4173	m1.medium	us-east-1b	running	2/2 checks...	None
KinesisDemo	i-a9922afa	m1.large	us-east-1b	running	2/2 checks...	None

Below are your CloudWatch metrics for the selected resources (a maximum of 10). Click on a graph to see an expanded view. All times shown are in UTC. [View all CloudWatch metrics](#)



- Amazon Redshift
- Clusters
- Snapshots
- Security**
- Parameter Groups
- Reserved Nodes
- Events

- Security Groups**
- Subnet Groups
- HSM Connections
- HSM Certificates

Filter: Search... X Viewing 1 of 1 Cluster Security Groups

Name	Status	Description
default	authorized	default

Cluster Security Group: default

Connection Type	Details	Status	Action
CIDR/IP	CIDR/IP: 80.187.106.10/32	authorized	Revoke
CIDR/IP	CIDR/IP: 80.187.107.22/32	authorized	Revoke
EC2 Security Group	AWS Account ID: 9794006868-12 EC2 Security Group: kinesiredshiftapp-ec2securitygroup-10cfjcg4f5dd	authorized	Revoke

Connection Type: CIDR/IP : CIDR/IP of your current machine: 80.187.106.10/32 **Authorize**

CIDR/IP to Authorize*: 80.187.106.10/32

- Amazon Redshift
- Clusters
- Snapshots
- Security
- Parameter Groups
- Reserved Nodes
- Events

Cluster: Configuration Status Performance Queries Loads

Cluster: kinesisredshiftapp-kinesiscluster-n2ea2k3gg7bl

Cluster Database Backup

Cluster Properties

Cluster Name: [kinesisredshiftapp-kinesiscluster-n2ea2k3gg7bl](#)

Cluster Type: Multi Node

Node Type: dw1.xlarge

Nodes: 2

Zone: us-east-1d

Created Time: 2014 May 22 14:50:13 UTC+2

Cluster Version: 1.0.789

Cluster Security Groups: [default](#) (active)

Cluster Parameter Group: [default:redshift-1.0](#) (in-sync)

Cluster Status

Cluster Status: available

Database Health: healthy

In Maintenance Mode: no

Parameter Group Apply Status: in-sync

Pending Modified Values: None

Cluster Database Properties

Endpoint: [kinesisredshiftapp-kinesiscluster-n2ea2k3gg7bl.c7faa3buzm4i.us-east-1.redshift.amazonaws.com](#)

Port: 5439

Backup, Audit Logging, and Maintenance

Automated Snapshot Retention Period: 1

Cross-Region Snapshots Enabled: No

Audit Logging Enabled: No

Maintenance Window: mon:07:00-

- Amazon Redshift
- Clusters**
- Snapshots
- Security
- Parameter Groups
- Reserved Nodes
- Events

Cluster: Configuration Status Performance Queries Loads

Cluster Properties

Cluster Name: [kinesisredshiftapp-kinesiscluster-n2ea2k3gg7bl](#)

Cluster Type: **Multi Node**

Node Type: **dw1.xlarge**

Nodes: **2**

Zone: **us-east-1d**

Created Time: **2014 May 22 14:50:13 UTC+2**

Cluster Version: **1.0.789**

Cluster Security Groups: [default \(active\)](#)

Cluster Parameter Group: [default.redshift-1.0 \(in-sync\)](#)

Cluster Database Properties

Endpoint: [kinesisredshiftapp-kinesiscluster-n2ea2k3gg7bl.c7fae3buzm6i.us-east-1.redshift.amazonaws.com](#)

Port: **8192**

Database Name: **kinesisdatabase**

Master Username: **dbuser**

Encrypted: **No**

Cluster Status

Cluster Status: **available**

Database Health: **healthy**

In Maintenance Mode: **no**

Parameter Group Apply Status: **in-sync**

Pending Modified Values: **None**

Backup, Audit Logging, and Maintenance

Automated Snapshot Retention Period: **1**

Cross-Region Snapshots Enabled: **No**

Audit Logging Enabled: **No**

Maintenance Window: **mon:07:00-mon:07:30**

Allow Version Upgrade: **Yes**

User=dbuser, URL=jdbc:postgresql://kinesisredshiftapt

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referer like 'google';
3 SELECT * FROM kinesisbasictable LIMIT 200;
4
```

I

Messages

Empty message area

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referrer like 'google';
3 SELECT * FROM kinesisbasictable LIMIT 200;
4
```

I

Result Messages

id	resource	referrer
1	/getting_started.html	http://www.amazon.com
6	/getting_started.html	http://www.reddit.com
11	/tips.html	http://www.amazon.com
14	/users.html	http://www.yahoo.com
20	/users.html	http://www.yahoo.com
1	/services.html	http://www.bing.com
6	/careers.html	http://www.reddit.com
11	/help.html	http://www.amazon.com
14	/careers.html	http://www.bing.com
20	/services.html	http://www.yahoo.com
1	/index.html	http://www.google.com
6	/index.html	http://www.yahoo.com
11	/blog.html	http://www.reddit.com
14	/blog.html	/index.html
20	/blog.html	/index.html



User=dbuser, URL=jdbc:postgresql://kinesisredshiftapi

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referrer like 'google?';
3 SELECT * FROM kinesisbasictable LIMIT 200;
4
```

I

Result Messages

id	resource	referrer
1	/getting_started.html	http://www.amazon.com
6	/getting_started.html	http://www.reddit.com
11	/tips.html	http://www.amazon.com
14	/users.html	http://www.yahoo.com
20	/users.html	http://www.yahoo.com
1	/services.html	http://www.bing.com
6	/careers.html	http://www.reddit.com
11	/help.html	http://www.amazon.com
14	/careers.html	http://www.bing.com
20	/services.html	http://www.yahoo.com
1	/index.html	http://www.google.com
6	/index.html	http://www.yahoo.com
11	/blog.html	http://www.reddit.com
14	/blog.html	/index.html
20	/blog.html	/index.html

User=dbuser, URL=jdbc:postgresql://kinesisredshiftapt

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;  
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referrer like 'google';  
3 SELECT * FROM kinesisbasictable LIMIT 200;  
4
```

Result Messages

count
4765904

User=dbuser, URL=jdbc:postgresql://kinesisredshiftapt

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;  
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referer like 'google*';  
3 SELECT * FROM kinesisbasictable LIMIT 200;  
4
```

Result Messages

count
4765904

Navigation icons: Run, Stop, Refresh, Undo, Redo, Home, Back, Forward, Find, Filter, Sort, Copy, Paste, Print, Reload, Save, Close. User=dbuser, URL=jdbc:postgresql://kinesisredshiftapt

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;  
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referer like 'google*';  
3 SELECT * FROM kinesisbasictable LIMIT 200;  
4
```

Result Messages

count
4765904

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referred like 'google';
3 SELECT * FROM kinesisbasictable LIMIT 200;
4
```

Messages

Empty message area for execution output.



Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referer like 'google';
3 SELECT * FROM kinesisbasictable LIMIT 200;
4
```

Result Messages

count
699466

Statement 1

```
1 SELECT COUNT(*) FROM kinesisbasictable;
2 SELECT COUNT(*) FROM kinesisbasictable WHERE referer like 'google*';
3 SELECT * FROM kinesisbasictable LIMIT 200;
4
```

Result Messages

count
699466

2014-05-22_Democratizing_Supercomputers.key

New Play View Guides Themes Masters Text Box Shapes Table Charts Comment Mask Alpha Group Ungroup Front Back

Auto-shrink Stroke

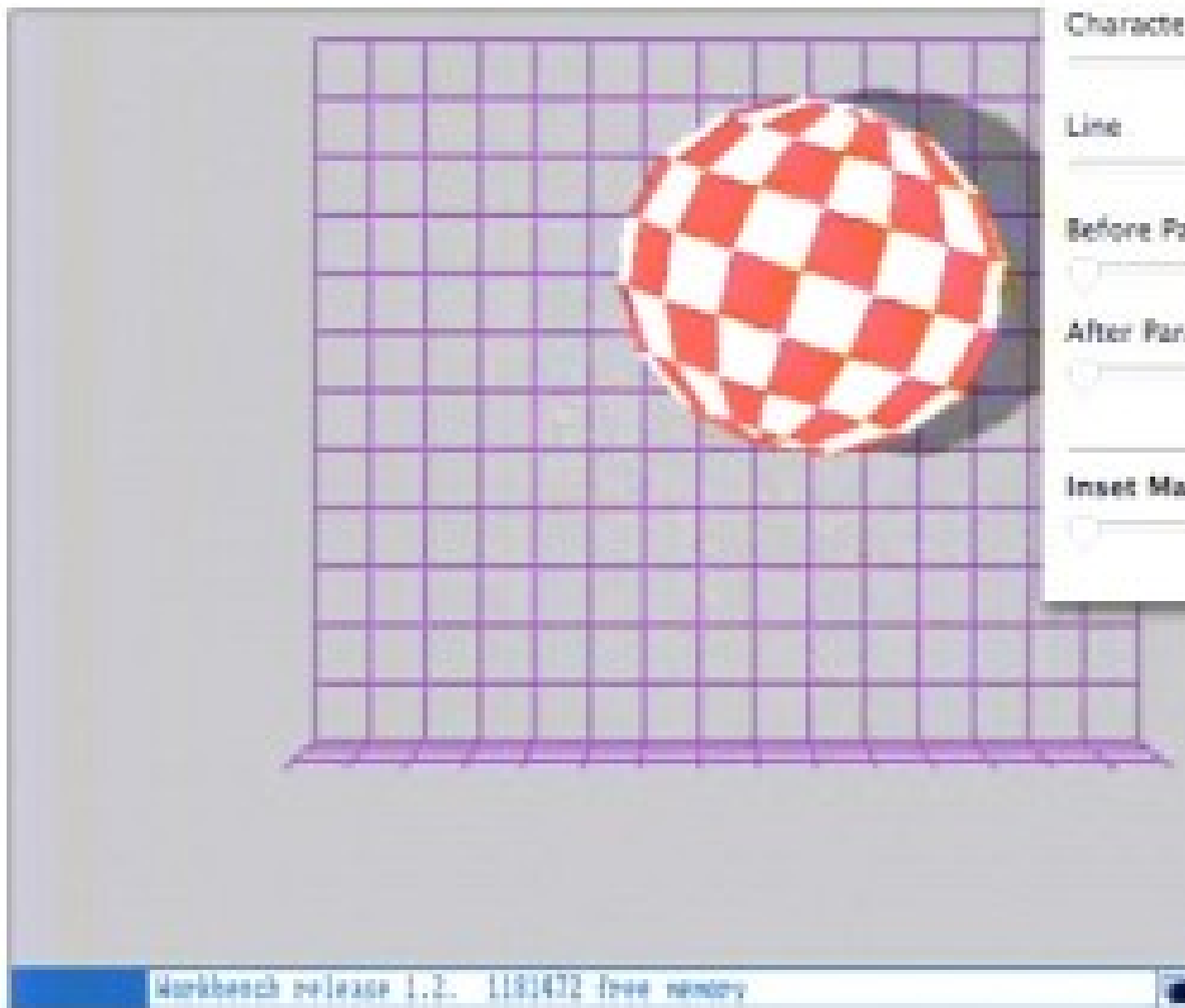
Master Slides

- Default - Title Slide
- ✓ Default - Title Only
- Default - Blank

Slides

- 62
- 63
- 64
- 65
- 66
- 67

Demo



Text

Text Columns Bullets

Color & Alignment

Spacing

Character

Line

Before Paragraph

After Paragraph

Inset Margin

2014-05-22_Democrati

New Play View Guides Themes Masters Text Box Shapes Table Charts

Play Slideshow ⌘P

Record Slideshow

Clear Recording

Rehearse Slideshow

Customize Presenter Display...

Text

Text Columns Bullets

Color & Alignment

Spacing

Automatically Shrink Text

Character

Line

Before Paragraph

After Paragraph

Inset Margin

Master Slides

Default - Title Slide

Default - Title Only

Default - Blank

Slides

62

63

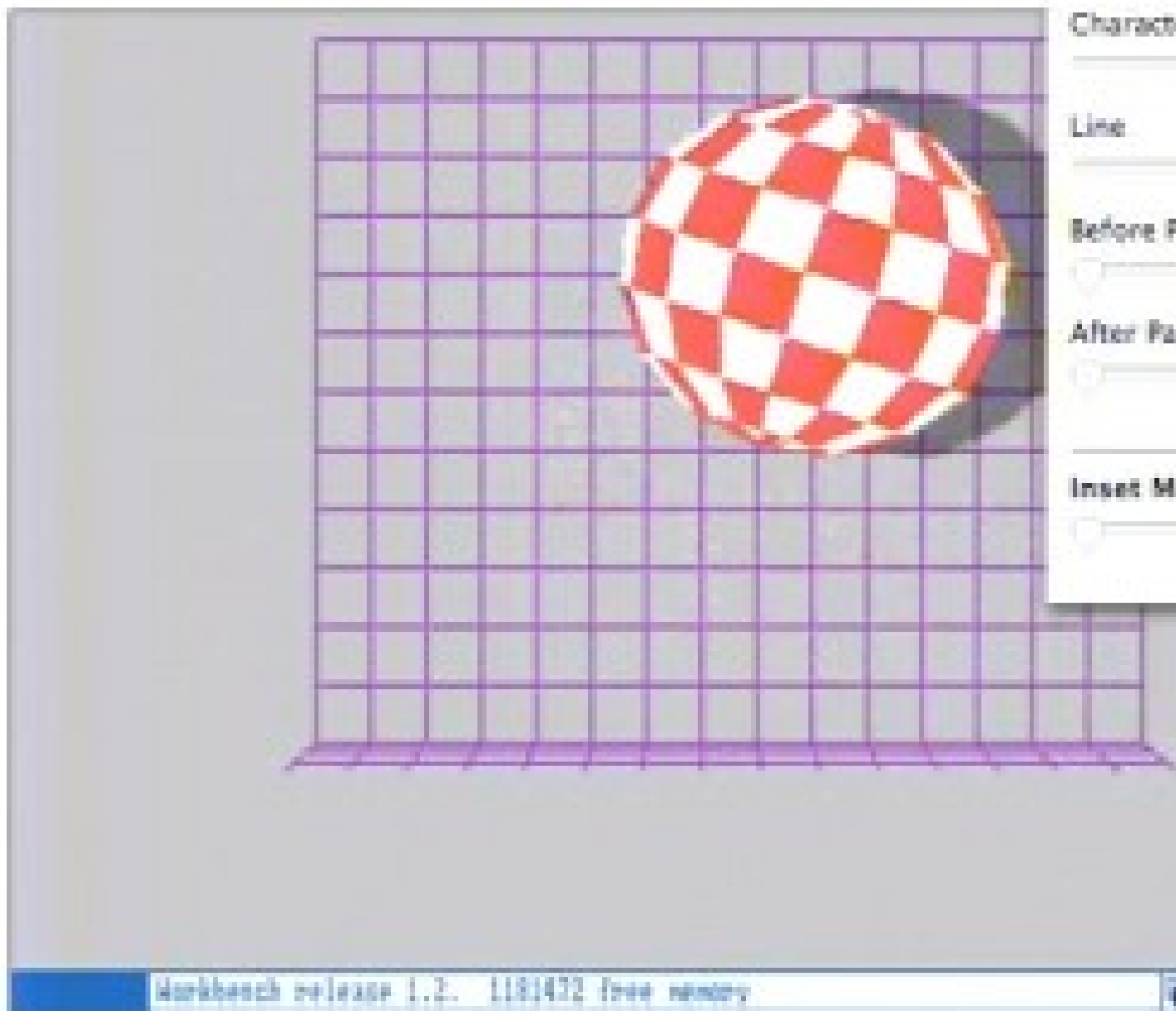
64

65

66

67

Demo



What Would
YOU
Do with a
100 Node
Supercomputer?



Try It Out Now



aws.amazon.com





Thank you!

Constantin Gonzalez
Solutions Architect
glez@amazon.de