

# Thanks!

- Galaxy core team, contributors and sponsors
- GCC 2013 Organizing Committee
- University of Oslo
- MIT StarCluster team
  - <http://star.mit.edu/cluster/docs/latest/index.html>
- Iowa Institute of Human Genetics
  - <http://www.medicine.uiowa.edu/humangenetics>

# Galaxy on an elastic HPC cluster at AWS

## A cost-effective and agile solution

srinivas-maddhi@uiowa.edu, Iowa Institute Of Human Genetics

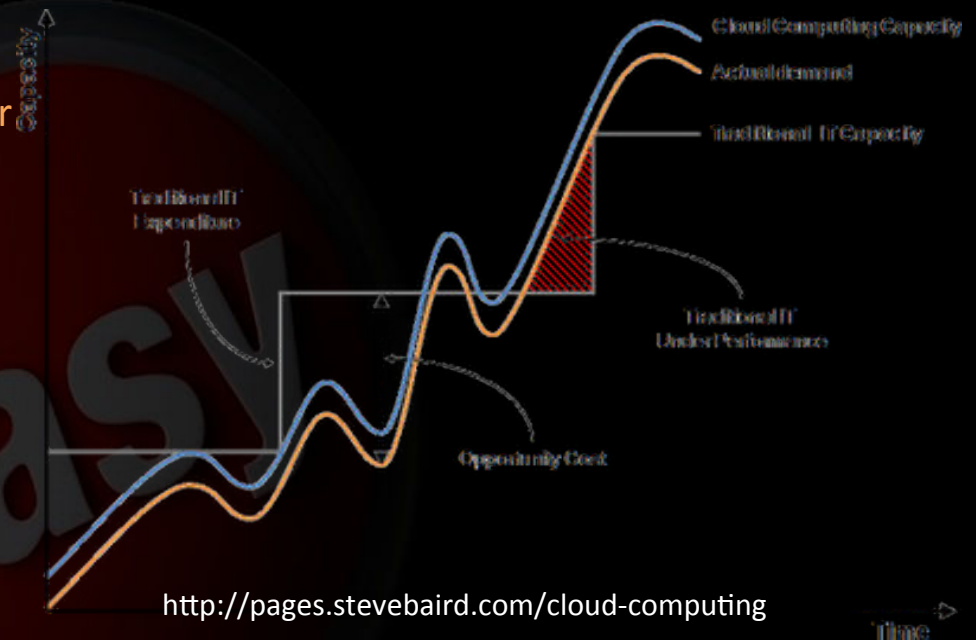
- **Bi-directionally auto-scaling HPC cluster**

- 10 gbps Ethernet
- Shared storage
- Sun Grid Engine 6

- **Cluster nodes (number based on load)**

- 32 cores (ht)
- 60.5 gB RAM
- 4 x 840 gB local storage
- Ubuntu 12.04

- **Galaxy with PostgreSQL**



### Steps:

- Create AWS account and EC2 instance
- Build elastic HPC Cluster using MIT StarCluster
- Install, configure and run Galaxy

### All that you need are:

1. A web browser
2. A credit card
3. ssh or Putty program

# Create An AWS Account [http://aws.amazon.com/]

EC2 Management Console - Pentadactyl

Services Edit IIG N. Virginia Help

### EC2 Dashboard

- Events
- Tags
- 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:

1 Running Instance	0 Elastic IPs
8 Volumes	0 Snapshots
2 Key Pairs	0 Load Balancers
1 Placement Group	8 Security Groups

[Optimize your resources' cost, performance and security with AWS Trusted Advisor](#) [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:**

- us-east-1a: Availability zone is operating normally
- us-east-1b: Availability zone is operating normally
- us-east-1c: Availability zone is operating normally

[Service Health Dashboard](#)

### Scheduled Events

**US East (N. Virginia):**

No events

### Account Attributes

#### Supported Platforms

- EC2-Classical
- EC2-VPC

#### Additional Information

- [Getting Started Guide](#)
- [Documentation](#)
- [All EC2 Resources](#)
- [Forums](#)
- [Pricing](#)
- [Report an Issue](#)

#### Popular AMIs on AWS Marketplace

- [Debian GNU/Linux](#)  
Provided by Debian  
Rating ★★★★★  
Free Software, pay only for AWS usage  
[View all Operating Systems](#)
- [Couchbase Server - Community Edition](#)  
Provided by Couchbase  
Rating ★★★★★  
Free Software, pay only for AWS usage  
[View all Databases](#)
- [LAMP Stack powered by BitNami](#)  
Provided by BitNami  
Rating ★★★★★  
Free Software, pay only for AWS usage  
[View all Application Stacks](#)

[Find more software on AWS Marketplace](#)

© 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

<https://console.aws.amazon.com/ec2/v2/home?region=us-east-1#> [1/3] Pop

# Create a management instance (a virtual machine)

EC2 Management Console - Pentadactyl

Services Edit

IHG N. Virginia Help

EC2 Dashboard  
Events  
Tags

Launch Instance Actions

Viewing: All Instances All Instance Types (Search) 1 to 8 of 8 Instances

### Create a New Instance

Select an option below:

- Classic Wizard  
Launch an On-Demand or Spot instance using the classic wizard with fine-grained control over how it is launched.
- Quick Launch Wizard  
Launch an On-Demand instance using an editable, default configuration so that you can get started in the cloud as quickly as possible.
- AWS Marketplace  
AWS Marketplace is an online store where you can find and buy software that runs on AWS. Launch with 1-Click and pay by the hour.

**Name Your Instance:** gcc13 Pick a meaningful name, e.g. Web Server

**Choose a Key Pair:**  
Public/private key pairs allow you to securely connect to your instance after it launches.

Select Existing  Create New  None

Name: kp01 Download

Please note that you need to download the key pair before you can continue.

**Choose a Launch Configuration:**

- More Amazon Machine Images** NEW  
Search through public and AWS Marketplace AMIs or choose from your own custom AMIs.
- Amazon Linux AMI 2013.03.1**  
The Amazon Linux AMI is an EBS-backed, PV-GRUB image. It includes Linux 3.4, AWS tools, and repository access to multiple versions of MySQL, PostgreSQL, Python, Ruby, and Tomcat. 64 bit  32 bit Free tier eligible
- Red Hat Enterprise Linux 6.4**  
Red Hat Enterprise Linux version 6.4, EBS-boot. 64 bit  32 bit Free tier eligible
- SUSE Linux Enterprise Server 11**  
SUSE Linux Enterprise Server 11 Service Pack 2 basic install, EBS boot 64 bit  32 bit Free tier eligible
- Ubuntu Server 12.04.2 LTS**  
Ubuntu Server 12.04.2 LTS with support available from Canonical (<http://www.ubuntu.com/cloud/services>). 64 bit  32 bit Free tier eligible

**Note:** You can customize your settings in the next step. Continue

[Submit Feedback](#) [Getting Started Guide](#)

Instances

After an Amazon Machine Image (AMI) is launched, the resulting running system is called an instance. By default, you can run up to 20 instances in a region. If you need more than 20 instances, you must request a limit through the Request Form.

Instances are not running unless they fail or you manually stop them. When this happens, the instance (the "instance store" or "ephemeral storage") is no longer available. Also, any Amazon Elastic Block Store (EBS) volumes that were attached at the time of the instance launch continue to exist until you manually delete them.

For more information about working with instances, go to the Amazon EC2 User Guide and Using Instances in the Amazon EC2 User Guide for Amazon Elastic Compute Cloud User Guide.

You can launch an instance from the console, from the AMIs page, or from the EC2 Dashboard. For help with the console, see the Amazon EC2 Instances page.

**Instance Details**

View the instance's details. To view an instance's details, select an instance. The instance's properties appear on the Instance Details tab in the lower pane.

**Instance**

Additional data you can add to an instance. This includes a key and an optional value. For more information about using tags, go to Using Tags in the Amazon Elastic Compute Cloud User Guide.

**Note:** Tags whose keys begin with aws: have been created by AWS and cannot be edited or deleted.

**To tag an instance**

Feedback

© 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

INSERT -- <https://console.aws.amazon.com/ec2/home?region=us-east-1#Instances> [1/3] Top



# Install, config and start Galaxy; initiate auto-scaling

### allow port 8080 access from world

```
[ec2-user@ip-10-164-2-121 ~]$ aws ec2 authorize-security-group-ingress --group-name @sc-gcc13 --ip-protocol tcp --from-port 8080 --to-port 8080 --cidr-ip 0.0.0.0/0
```

### log in to master node

```
[ec2-user@ip-10-164-2-121 ~]$ starcluster sshmaster gcc13
```

### install PostgreSQL, create DB user and DB

```
root@master:~# apt-get update; apt-get install postgresql; sudo -i -u postgres createuser -DPRS gx; sudo -i -u postgres createdb -O gx gx
```

### create directory for Galaxy; su to Galaxy user

```
[root@master ~]# mkdir /gxds/gx; chown gx:gx /gxds/gx; su - gx
```

### install Galaxy

```
[gx@master ~]$ cd /gxds/gx; hg clone https://bitbucket.org/galaxy/galaxy-dist/; cd galaxy-dist; hg update stable
```

### create universe\_wsgi.ini

```
[gx@master galaxy-dist]$ sed -e 's/^\[server:main\]\[server:main\]\n\nhost=0.0.0.0/' \
-e 's/^\[app:main\]\[app:main\]\n\nndatabase_connection=postgres://\//gx?host=\var/run/postgresql/' \
universe_wsgi.ini.sample > universe_wsgi.ini
```

### create job\_conf.xml

```
[gx@master galaxy-dist]$ cat > job_conf.xml <<EOF
```

```
<?xml version="1.0"?>
```

```
<job_conf>
```

```
<plugins>
```

```
<plugin id="drmaa" type="runner" load="galaxy.jobs.runners.drmaa:DRMAAJobRunner"/>
```

```
</plugins>
```

```
<handlers default="main">
```

```
<handler id="main"/>
```

```
</handlers>
```

```
<destinations default="remote_cluster">
```

```
<destination id="remote_cluster" runner="drmaa" tags="longjobs"/>
```

```
</destinations>
```

```
</job_conf>
```

```
EOF
```

### start Galaxy

```
[gx@master galaxy-dist]$ ./run.sh
```

## process some jobs in Galaxy

...

### start StarCluster auto-scaling process to allow minimum of 1 node and maximum of 16 nodes

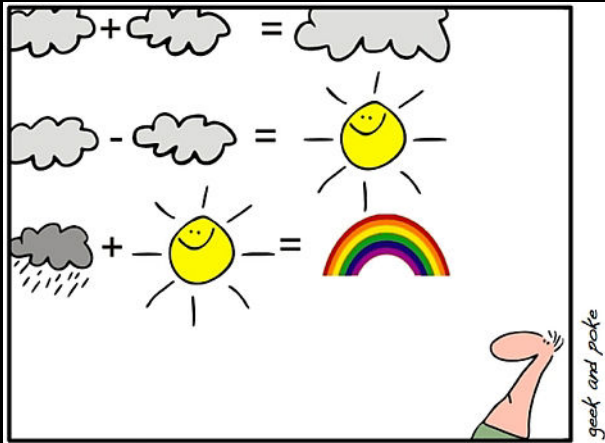
```
[ec2-user@ip-10-164-2-121 ~]$ starcluster loadbalance -m 16 -n 1 gcc13
```

### stop cluster; restart cluster; terminate cluster

```
[ec2-user@ip-10-164-2-121 ~]$ starcluster stop gcc13
```

```
[ec2-user@ip-10-164-2-121 ~]$ starcluster start -x gcc13
```

```
[ec2-user@ip-10-164-2-121 ~]$ starcluster terminate gcc13
```



**SIMPLY EXPLAINED - PART 17:  
CLOUD COMPUTING**