GMOD in the Cloud
http://gmod.org/wiki/Cloud
Scott Cain¹ and the GMOD Consortium
¹Ontario Institute for Cancer Research, Toronto, Ontario, Canada, M5G 0A3

With decreasing costs of sequencing technology there is an increasing need for computing resources and tools to work with the data. An alternative to building and maintaining a large computing infrastructure in-house is to use existing networked computing systems ("cloud computing") to make use of preconfigured and extensible servers. Here we present several efforts implementing GMOD software tools using Amazon Web Services (AWS). First, a GBrowse2 server has been prepared with options for importing available data from Amazon storage (EBS) for several common organisms and as well as tools for creating additional rendering servers as needed to improve the responsiveness of the website. Next, a community annotation server is available with Chado, GBrowse2, JBrowse and Tripal to facilitate taking computational annotations (for instance from MAKER) and making them available to interested users in the community to review and improve upon. Additionally there are two data analysis cloud images: one for sequence annotation based on GMOD tools called CloVR, and a cloud implementation of Galaxy called Cloudman.

---

### GMOD in the Cloud

#### Installed software:

- Tripal: Drupal-based web frontend
- Chado: Generic organism DB schema
- GBrowse: Venerable genome browser
- Sample data: Saccharomyces cerevisiae

### GBrowse2

GBrowse2 is a highly configurable and flexible web-based genome browser used by hundreds of organizations.

#### Features:

- Rendering servers
- Fixed cluster size
  - 5 nodes: computation cost: $5
  - 20 nodes: computation cost: $15
- Dynamic cluster size
  - 1 to 16 nodes: computation cost: $20

#### Benefits:

- To speed up GBrowse, tracks can be rendered by any number of rendering slaves, spreading out the work.
- The admin interface lets you add and remove slave rendering nodes as needed, as well as mount standard data sets.

### Galaxy CloudMan

Use Galaxy's CloudMan to get the power of Galaxy without the usage or data restrictions.

#### Instructions:

1. Create AWS account and find the current CloudMan AMI.
2. Create CloudMan instance.
3. Provide some CloudMan specific parameters (e.g., how many machines to allow in the cluster).
4. Go to the URL that corresponds to the head node.

#### Benefits:

- The local virtual machine monitors the analysis cluster in the cloud, and the master node in the cloud directs computational work in the cluster.

### CloVR

Cloud Virtual Resource is an automated sequence analysis pipeline powered by GMOD tools Workflow and Ergatis.

#### Features:

- Installed software:
  - Sample data: Fly
  - Sample data: Human
  - Sample data: Yeast

#### Instructions:

1. Go to the URL that corresponds to the head node.
2. Provide some CloudMan specific parameters (e.g., how many machines to allow in the cluster).
4. Provide some CloudMan specific parameters (e.g., how many machines to allow in the cluster).
5. Go to the URL that corresponds to the head node.

#### Benefits:

- Powerful computation on the cheap!

---

GBrowse2

#### Example Image 1

![GBrowse2](https://example.com/gbrowse2.png)

#### Example Image 2

![GBrowse2](https://example.com/gbrowse2.png)

---

CloVR

#### Example Image 3

![CloVR](https://example.com/clovr.png)

#### Example Image 4

![CloVR](https://example.com/clovr.png)