

## MISO-lims

The MISO-lims laboratory information management system (LIMS) is an open-source LIMS software used and supported by the Genome Sequencing Informatics group (GSI) at OICR and is the only known, purpose-built LIMS supporting functionality specifically for genomics core facilities. Ontario Joint Genomics Program (OJGP) member genome centres are required to use MISO-lims as the standard LIMS for accredited assays offered through OJGP. This document covers the primary considerations for centres needing to deploy the software. If you have any questions, please contact [inquiries@ojgp.ca](mailto:inquiries@ojgp.ca).

### Cloud vs. standalone

In choosing between cloud versus a standalone setup you should consider whether your institution's IT policies will allow OJGP to access your environment to provide support to a standalone instance of MISO. The instance will also have to be able to connect to the OJGP network through institutional firewalls. To avoid these potential barriers as well as technical issues such as differences across operating systems and configurations, OJGP recommends running MISO in the cloud. However, OJGP will provide support for standalone operations where possible.

### Choosing a provider

With offerings from Microsoft (Azure), Amazon (AWS) and Google (GCP), the best choice of cloud provider is left up to member centres and depending on its size, budget and current computing ecosystem.

Pricing models for commercial clouds are complex. Below are three possible setups to help guide your decision about the number and cost of the virtual machines (VMs) needed.

### Ideal setup (recommended for a large genome centre)

The following VMs are recommended:

- MISO web service staging: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- MISO db staging: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- MISO web service production: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- MISO db production: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- RunScanner production: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- Pinery production: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB

## Typical setup (recommended for a small genome centre)

It is recommended to run at least three VMs:

- MISO-staging web service and db: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- MISO-production web service and db: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- RunScanner and Pinery production: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB

## Minimal setup (recommended only for trial or supporting a small research lab)

GSI has been able to successfully compact MISO + Pinery into the following for a small research lab:

- MISO-staging web service, db, Pinery: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB
- MISO-production web service, db, Pinery: Ubuntu OS, RAM 16GB, 2 CPU, Disk 25GB

## Cost considerations -Case study

A research lab using the minimal setup described above is currently running MISO in the cloud using MS Azure virtual machines. They switched from hosting MISO on their own servers due to the expense of maintaining their own hardware. According to their calculations, the total cost of ownership over time is in favour of the commercial cloud solution. The cost of this minimal footprint is about \$3,000 to \$4,000 per year but a core lab operation should budget for more. The current rule of thumb is about \$150 per VM a month. Costs can be saved if the staging instance is turned off except when it is needed to do updates.

Other factors in commercial cloud pricing include network IO, region, backups, etc. Note that for some configurations disk storage is billed separately from the VM.

## Private cloud

OICR Genomics has been evaluating private cloud and has found that MISO + Pinery + RunScanner is running well using OpenStack with similar to the above "Ideal" machine specs:

- MISO web service staging: RAM 8GB, 4 CPU, Disk 30GB
- MISO db staging: RAM 8GB, 4 CPU, Disk 30GB
- MISO web service production: RAM 16GB, 2 CPU, Disk 25GB
- MISO db production: RAM 8GB, 4 CPU, Disk 30GB
- RunScanner staging: RAM 16GB, 2 CPU, Disk 25GB
- RunScanner production: RAM 16GB, 2 CPU, Disk 25GB
- Pinery production: RAM 32GB, 2 CPU, Disk 40GB
- Pinery staging: RAM 16GB, 2 CPU, Disk 40GB

## Support provided by OJGP

OJGP (via OICR's GSI group) currently does not administer MISO instances in commercial cloud interfaces. OJGP only provides machine minimum specifications, and logs into these instances to do MISO setup and maintenance work (using SSH). OJGP will provide initial training for member centre administrator(s) to get them up and running. If a centre already has a standalone MISO instance, OJGP will assist in upgrading its MISO and/or migrate to the cloud. GSI requires a backup of the MISO instance and can work with member centres to restore its database in a new instance (wherever that be on physical hardware, private cloud, commercial cloud).

## Glossary

OJGP – Ontario Joint Genomics Program

OICR – Ontario Institute for Cancer Research

GSI – Genome Sequencing Informatics

Pinery – software which acts as a data extraction layer that works with MISO to feed reporting.

LIMS – laboratory information management system

MISO – an open-source LIMS supported by GSI

VM – virtual machine

DB – database

Production – day to day operating environment

Staging – an environment used for software development, testing or migration

Cloud – a highly available, scalable, network accessible internal (private) or external (commercial) VM environment providing infrastructure as a service (IaaS), platform as a service (PaaS) and/or software as a service (SaaS) solution.

SSH – Secure shell

