OMERO Image Management System

Project Plan

|  |  |
| --- | --- |
| Authors | Gregg TeHennepe |
| Creation Date | 10/19/11 11:41 AM |
| Last Updated | 10/19/11 4:38 PM |
| Version | 1.0 – DRAFT |
| Distribution | Jackson Internal |

|  |  |
| --- | --- |
| Approvals | |
|  |  |
|  |  |
|  |  |
|  |  |



Table of Contents

1. Executive Summary 3

2. Project Request 3

2.1. Background 3

2.2. Request 3

2.3. Justification 4

2.4. Assumptions 4

3. Project Scope 4

3.1. Requirements 4

3.1.1. Functional 4

3.1.2. Technical 5

3.1.3. Operational 6

3.2. Schedule 6

3.3. Budget 7

3.4. Control 7

3.4.1. Project Team 7

3.4.2. Change Management 7

4. Project Deliverables 7

4.1. Project Schedule 7

4.2. Technical Architecture 7

4.3. Implementation Plan 8

4.4. Information Security Plan 8

4.5. Business Continuance/Disaster Recovery Plan 8

4.6. Testing Plan 9

4.7. Training Plan 9

4.8. Documentation 9

5. Appendices 9

5.1. Revision History 9

5.2. Definition of Terms/Acronyms 9

5.3. References 9

5.3.1. Open Issues 9

5.3.2. Potential Users and Storage Estimates 11

# Executive Summary

*This section summarizes:*

* *the project request – who, why*
* *the recommended approach*
* *major milestones of the schedule*
* *capital and operating budgets*
* *staffing needed*
* *interdependencies with other projects*
* *project risks*

# Project Request

## Background

Multiple labs and groups including Imaging Sciences produce images from a variety of microscopy and imaging instruments from a number of different vendors. These data are created in proprietary formats, including associated metadata, and are stored in numerous fileshares and local devices across the organization.

## Request

The request for an OMERO server originated in Lindsay Shopland’s lab. Jake Bolewski joined the lab and spurred interest in putting up the server. IT checked with Imaging Sciences, who are familiar with the platform and interested in an institutional instance.

The Open Microscopy Environment (OME) is a multi-site collaborative effort among academic laboratories and a number of commercial entities that produces open tools to support data management for biological light microscopy. Designed to interact with existing commercial software, all OME formats and software are free, and all OME source code is available under the GNU General public license or through commercial license from Glencoe Software.[[1]](#footnote-1)

OME Remote Objects (OMERO) is a modern client-server software platform for visualising, managing, and annotating scientific image data. OMERO also provides components for image importing, archiving, protocol recording, and user administration. OMERO consists of a Java server, several Java client applications, as well as Python and C++ bindings and a Django-based web application.[[2]](#footnote-2)

The system is requested as soon as reasonably possible, there are no hard deadlines associated with this request.

## Justification

A centralized OMERO image server would have multiple benefits:

* would provide a single central environment for managing, curating, annotating, and sharing image data;
* provide instrument-independent formats for greater usability with analysis tools;
* centralized management and tracking of infrastructure resources (ie storage) supporting image data;
* environment for data sharing with external collaborators;
* standard APIs for integration with analysis software development

The system would likely be of interest to numerous other groups at Jackson, including the laboratories of Simon John, Patsy Nishina, Sue Ackerman, Da Ting Lin, and Kyuson Yun.

## Assumptions

*Availability of project resources, schedule in relation to other projects, etc. associated with the request. This section can also document related context that impacts this request, ie input from others beyond the requestor that will help people understand the request.*

# Project Scope

## Requirements

### Functional

In addition to the primary functional capabilities of the OMERO server (see documentation[[3]](#footnote-3)), the following needs are requested:

* Ability to provide access to external collaborators
* Ability to archive data
* Ability to exercise the API from external, locally developed software (Shopland Lab)

The OMERO software does not currently support the following requirement:

* Support for large (65,550 pixel) images from the Olympus NanoZoomer slide scanner (see associated bug fix request[[4]](#footnote-4))

A single production instance is requested.



Figure 1 – High-level functional workflow.

### Technical

From the OMERO documentation[[5]](#footnote-5):

|  |
| --- |
| The recommended OMERO.server specification we suggest for between 25-50 users is:     * Mac OS X 10.5 or later; Windows XP or later; Ubuntu 8.10 or later/Centos 5/Debian Lenny or other Linux distro * Quad core 1.33GHz Intel or AMD CPU * 8GB RAM * 250MB hard drive space for OMERO.server distribution * Java 1.5 or later * Python 2.4 or later * Hard drive space proportional to the image sizes expected (Likely between 10 and 100TB)   RAM is not going to scale linearly, particularly with the way the JVM works. You're probably going to hit a hard ceiling between 4 and 6GB for JVM size (there's really not much point in having it larger anyway). With a large database and aggressive PostgreSQL caching your RAM usage could be larger but I would surely doubt a large deployment using more than a few GBs of RAM for this purpose, it's just not cost effective.  Summary: Depending on hardware layout 16, 24 or 32GB of RAM would be ideal for your OMERO server. If you have a separate database server more than 16GB of RAM may not be of much benefit to you at all.  CPU is really not something that an OMERO system is almost ever limited by. However, when it is limited it's almost always limited by GHz and not by the CPU count. So you're not going to get a huge OMERO experience performance increase by, for example, throwing 24 cores at the problem.  Summary: Depending on hardware layout 2 x 4, 2 x 6 system core count should be more than enough. |

### Operational

The system has the following operational needs:

* Ability to report on or segregate storage resources for efficient resource management;
* Integration with existing directory services for simplified user account management;
* Standard availability – prefer 24x7, but not required;
* Non-critical recovery – in the event of a disaster this is not a priority system, and should be recovered to the last daily backup.

Responsibility for operational management of the application needs to be determined, Imaging Sciences is the most knowledgeable and plausible group.

## Schedule

There are no hard requirements or deadlines for the schedule. The system is requested as soon as reasonably possible.

## Budget

Resources needed to support the system are TBD.

## Control

### Project Team

Project team includes the following individuals:

|  |  |
| --- | --- |
| Lindsay Shopland, Assistant Professor  Jake Bolewski, Research Assistant I | Institute for Molecular Biophysics/Shopland Lab |
| Jim Denegre, Senior Manager  Mark Lessard, Microscopy Senior Technologist | Imaging Services |
| Carol Bult, Professor | Strategic Advisor to IT |
| Gregg TeHennepe, Senior Manager, Research Liaison  Jim Merritt, Manager, Systems Administration  Val Apoderado, Infrastructure Consultant | Information Technology |

### Change Management

This project is not utilizing formal change management. Change will be communicated via email and this document.

# Project Deliverables

*This section documents Information Technology’s response the to request, within the scope as described in section 3. These deliverables should be negotiated to the satisfaction of [requestor] and Information Technology.*

## Project Schedule

Installation of a pilot/test system to evaluate functionality and address open questions is under way (see section 5.3.1 for open issues).

*This will document the tasks needed to complete the project, including:*

* *when tasks start and complete*
* *who is assigned to the tasks*
* *dependencies between tasks*

## Technical Architecture

*This section documents the technical design, potentially including:*

* *Hardware architecture and configuration*
* *Software architecture and configuration*
* *Network architecture and configuration*
* *Security architecture and configuration*

*This can be an addendum to the project plan.*

## Implementation Plan

*This section documents any software required by the project:*

* *Architecture, acquisition/design, development, testing information*

*This can be an addendum to the project plan.*

## Information Security Plan

*The columns may need to be adjusted for the appropriate roles.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **User** | **Data Curator** | **App Admin** | **Sys Admin** |
| **Authorization** Who authorizes this level of access? | Imaging Services | Imaging Services | Dir, Scientific Services | CIO |
| **Authentication**  What method(s)? | Password | Password | Password | Password |
| **Access Controls**  How is access controlled, and who maintains it? | AD accounts (maintained by IT)  Data access (maintained by Imaging Services) | AD accounts (maintained by IT)  Data access (maintained by Imaging Services) | AD accounts (maintained by IT) | AD accounts (maintained by IT) |
| **Audit**  Where is the access logged? Who is responsible for reviewing it? | System & Application logs | System & Application logs | System & Application logs | System logs |

## Business Continuance/Disaster Recovery Plan

* *Document backup and recovery strategy*
* *Document BC/DR strategy to meet RTO/RPO. May require evaluating the costs of different RTOs/RPOs and selecting the most appropriate.*

## Testing Plan

* *Document testing strategy and plan*
* *Identify testing resources*

## Training Plan

* *Identify owner of training plan and implementation*
* *Document training approach*
* *Identify trainer(s) and trainee(s)*
* *Identify location and schedule*
* *Describe training documents and author(s)*

## Documentation

* *What documents will be produced in this project?*
* *Who will author/review these documents?*
* *Who will use these documents*

# Appendices

## Revision History

Document Template revision: 1.6

|  |  |  |  |
| --- | --- | --- | --- |
| **Rev** | **Revision Description** | **Date** | **Author** |
| 1.0 | Initial draft. | 10/19/11 | Gregg TeHennepe |
|  |  |  |  |
|  |  |  |  |

## Definition of Terms/Acronyms

## References

### Open Issues

Q1 - How does this system store the data?

A1 – In a [binary repository](http://www.openmicroscopy.org/site/support/omero4/server/binary-repository), with all files owned by the user that runs the server instance. I haven't found any indication of a reporting facility to report on usage.

Q2 - How much data do we estimate the system will hold? Who are other potential users?

A2 – See table in section 5.3.2.

Q3 - How are data and metadata connected?

Q4 - How long to retain the original image?

Q5 - How are data associated with groups/individuals?

A5 – [[Permissions info](http://www.openmicroscopy.org/site/support/omero4/server/permissions)] [[Security](http://www.openmicroscopy.org/site/support/omero4/server/security)]

Q6 - Does the system support multiple storage locations?

A6 – No, but [they are looking at it](http://www.openmicroscopy.org/site/support/faq/omero/can-i-store-the-repository-on-multiple-drives).

Q7 - How to migrate/import from existing fileshare system?

A7 – You can import data into the system, or utilize [OMERO.fs](http://www.openmicroscopy.org/site/support/omero4/server/fs) to monitor data in external filesystems, though it appears this is primarily in support of DropBoxes.

Q8 - Do we need to provide access to the system from outside the institution?

A8 – Yes, this needs a technical architecture. [[Security](http://www.openmicroscopy.org/site/support/omero4/server/security)]

Q9 – Is the NanoZoomer file format supported?

A9 – Yes, but only for files up to 65,550 pixels in height or width. The [underlying problem](http://trac.openmicroscopy.org.uk/ome/ticket/5092) is in Java's JPEG decoder, it's considered critical however there is no schedule for resolving this issue.

### Potential Users and Storage Estimates

The following table was developed by Mark Lessard.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Lab/Group** | **(from service machines) approximate annual data produced (Gb)** | **amount currently on Bioimage FS (Gb)** | **Current amount of Nanozoomer data (Gb)** | **Projected annual amount of Nanozoomer data (Gb)** | **Currently have their own FS** | **Size (Tb)** | **other notes** |
| Ackerman, Sue | 20 | 10 | 0 | ? | ? |  | have their own microscope systems |
| Bo Chang | 5 | 0 | 12 | 10 | ? |  |  |
| Braun, Robert | 100 | 175 | 24 | 25 | Yes |  | have their own microscope systems |
| Burgess, Robert | 100 | 0 | 0 | ? | Yes |  | have their own microscope systems |
| Cox, Greg | 150 | 0 | 0 | ? | Yes | 1 |  |
| Dai, Chengkai | 20 | 0 | 20 | 20 | ? |  |  |
| Frankel, Wayne | 100 | 5 | 0 | ? | Yes |  | have their own microscope systems |
| GRS | 20 | 22 | 405 | 500 | ? |  |  |
| Handel, Mary Ann |  | 0 | 5 | 5 | Yes | 0.371 |  |
| Harrison, David | 20 | 1 | 2 | 5 | ? |  |  |
| Histology | 5 | 30 | 65 | 50 | No |  |  |
| Lin, Da-Ting | ? | 0 | 0 | 0 | ? |  | This lab produces large 3D data stacks from a custom built 2photon microscope |
| Maser, Rick | 5 | 1.5 | 226 | 250 | ? |  |  |
| Nishina, Patsy | \*\* | 13 | 0 | 0 | ? |  | have their own microscope systems; new personel with lots of projected image analysis and capture |
| Paigen, Beverly  Korsanje, Ron | 20 | 0 | 300 | \*4000 | Yes |  | large number of nanozoomer scans and 3D reconstruction of images for strain comparison |
| Repository | 5 | 10 | 230 | 250 | ? |  |  |
| Roopenian, Derry | 50 | 50 | 0 | ? | Yes |  |  |
| Shopland, Lindsay | 1000 | 0 | 0 | 0 | Yes | 2 | have their own microscope systems |
| **Lab/Group** | **(from service machines) approximate annual data produced (Gb)** | **amount currently on Bioimage FS (Gb)** | **Current amount of Nanozoomer data (Gb)** | **Projected annual amount of Nanozoomer data (Gb)** | **Currently have their own FS** | **Size (Tb)** | **other notes** |
| Shultz, Lenny | 20 | 0 | 20 | 20 | Yes |  |  |
| John, Simon | 1500-2000 | 0 | 0 | ? | Yes |  |  |
| Sundberg, John | 5 | 100 | 135 | 150 | ? |  |  |
| Taft, Rob | 5 | 27 | 22 | 5 | No |  |  |
| Yun, Kyuson | 20 | 11 | 52 | 50 | Yes |  | have their own microscope systems |
| Zhang, Zhong Wei | 20 | 1 | 0 | 0 | ? |  | have their own microscope systems |

1. http://www.openmicroscopy.org/site/about/who-ome [↑](#footnote-ref-1)
2. http://www.openmicroscopy.org/site/support/omero4 [↑](#footnote-ref-2)
3. http://www.openmicroscopy.org/site/products/feature-list [↑](#footnote-ref-3)
4. http://trac.openmicroscopy.org.uk/ome/ticket/5092 [↑](#footnote-ref-4)
5. http://www.openmicroscopy.org/site/support/omero4/system-requirements [↑](#footnote-ref-5)