
Alone in the dark
making big data
luke hammond / QBI

Mouse Hippocampus

Nika Mohannak / Meunier Lab

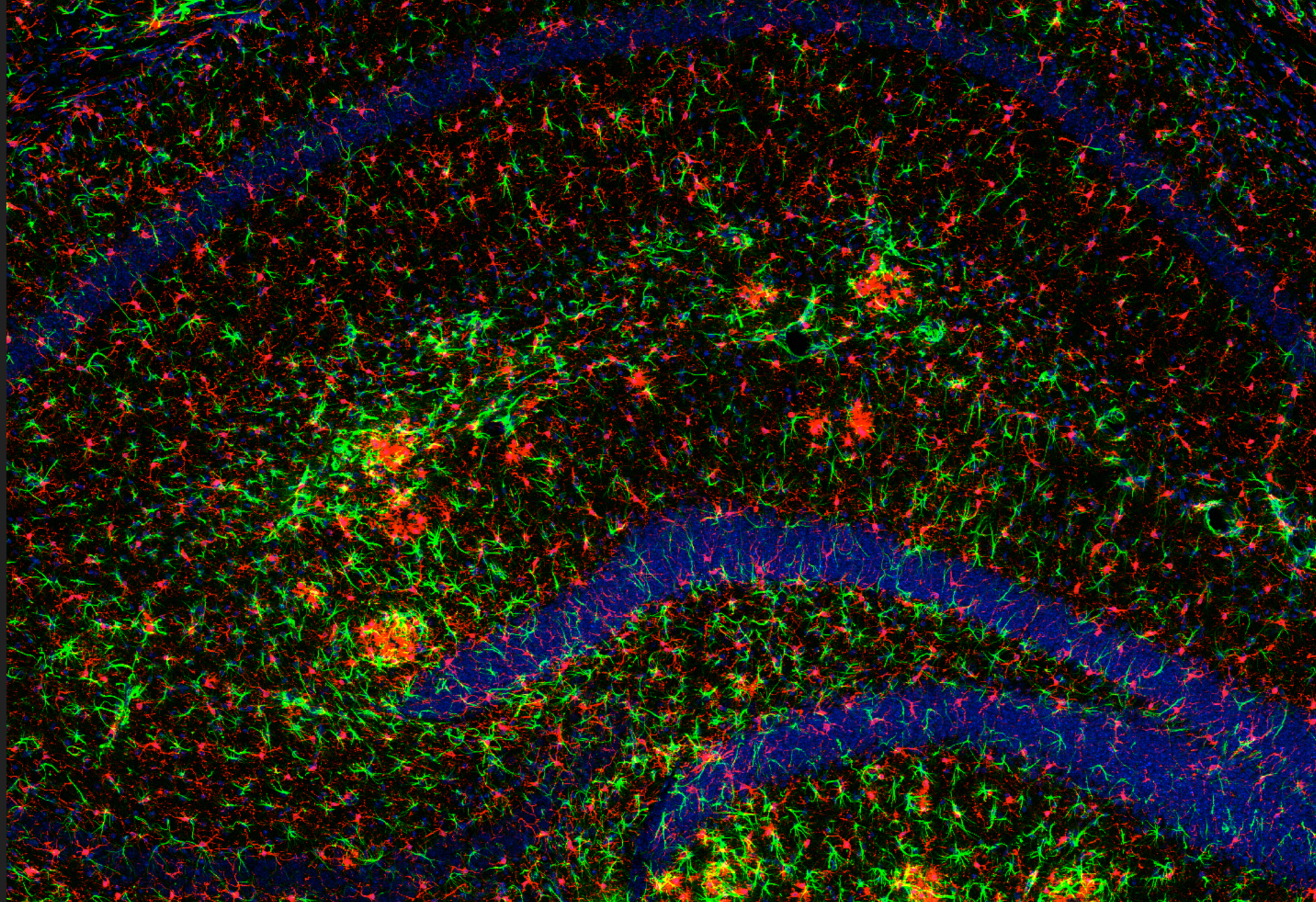
AxioImager + ApoTome

20x 0.8 NA Water Objective

0.3 x 0.3 µm XY

1.2 Z step size

12 slices x 6 tiles



Zebrafish / GCaMP6

Dr. Jeremy Ullmann

Yokogawa W1 SDC

40x 1.2 NA Water Objective

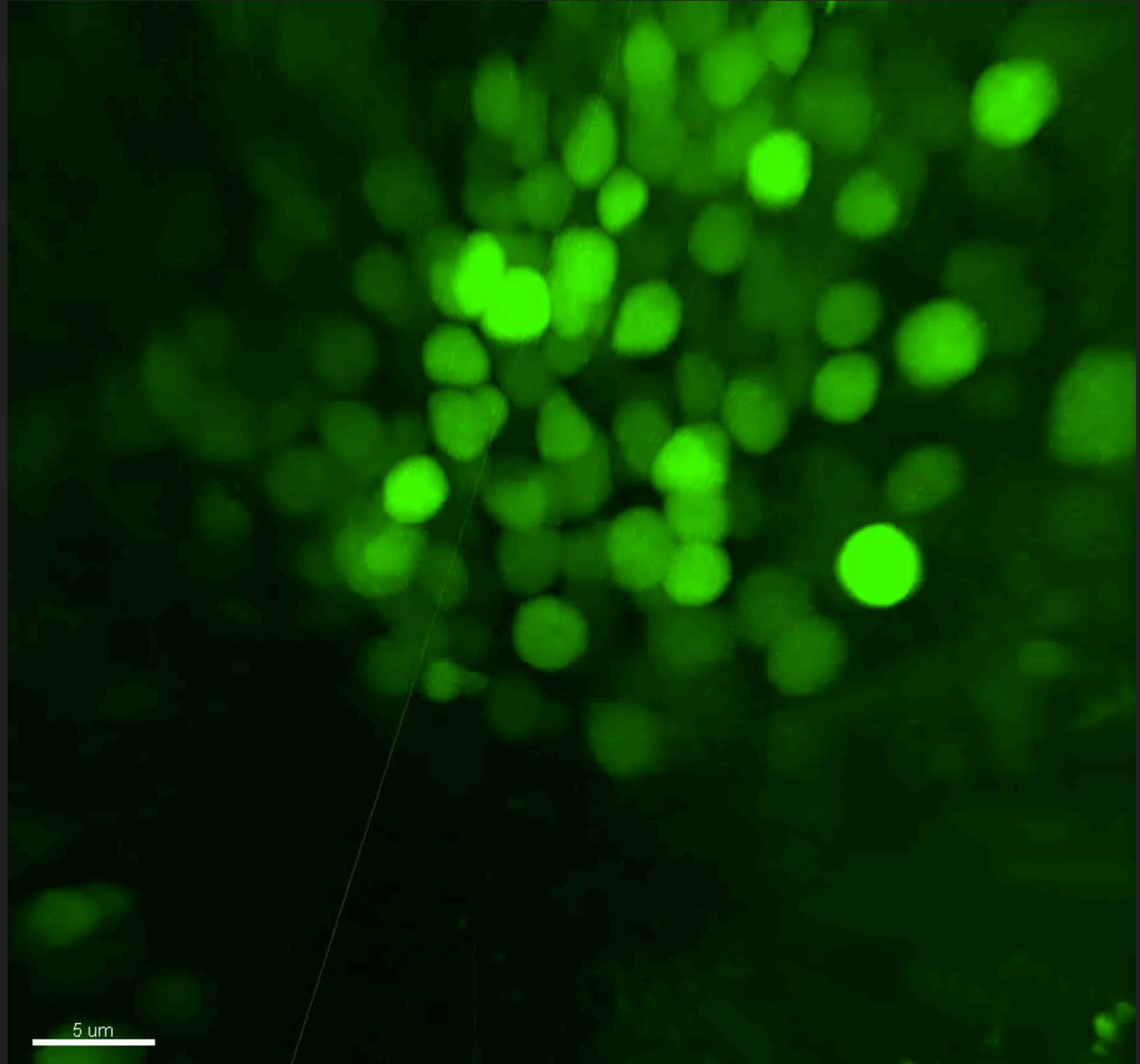
0.15 x 0.15 µm XY

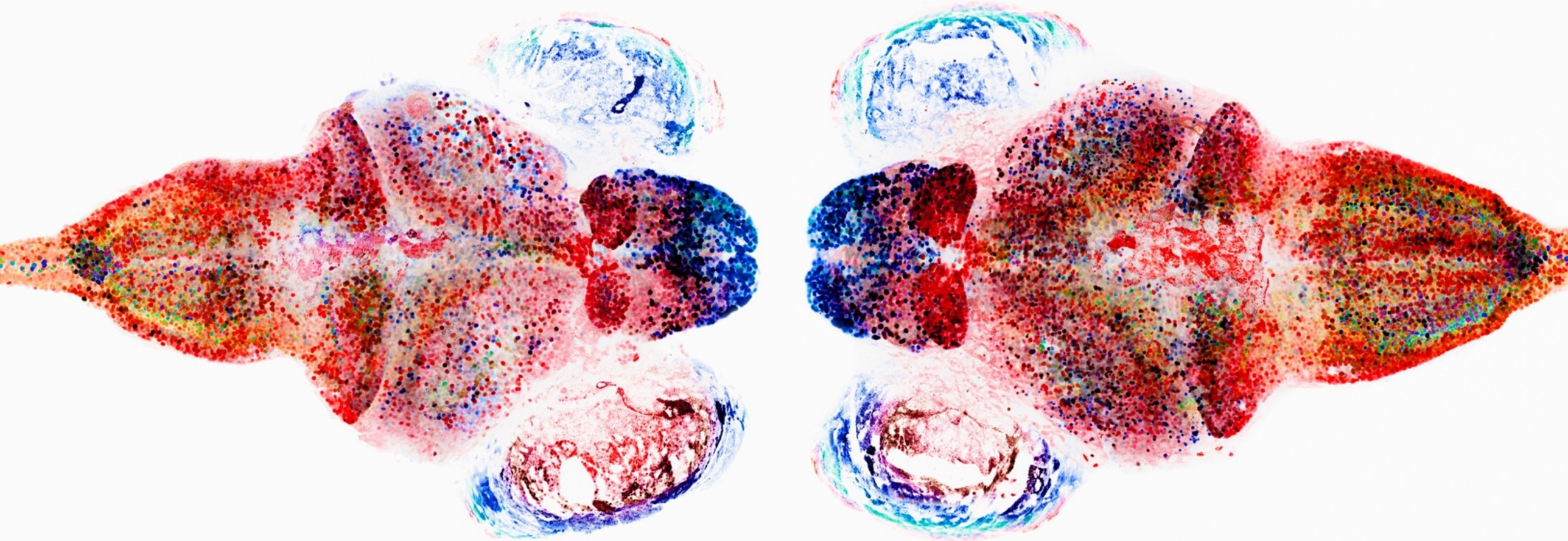
0.33 Z step size

800 slices x 6 tiles

20 minutes acquisition

40 GB data set

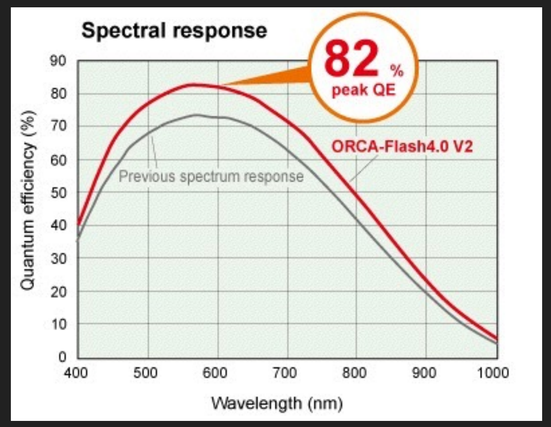




Jeremy Ullmann

sCMOS Cameras

Hamamatsu Flash 4v2 high QE / 2015



100 frames / second

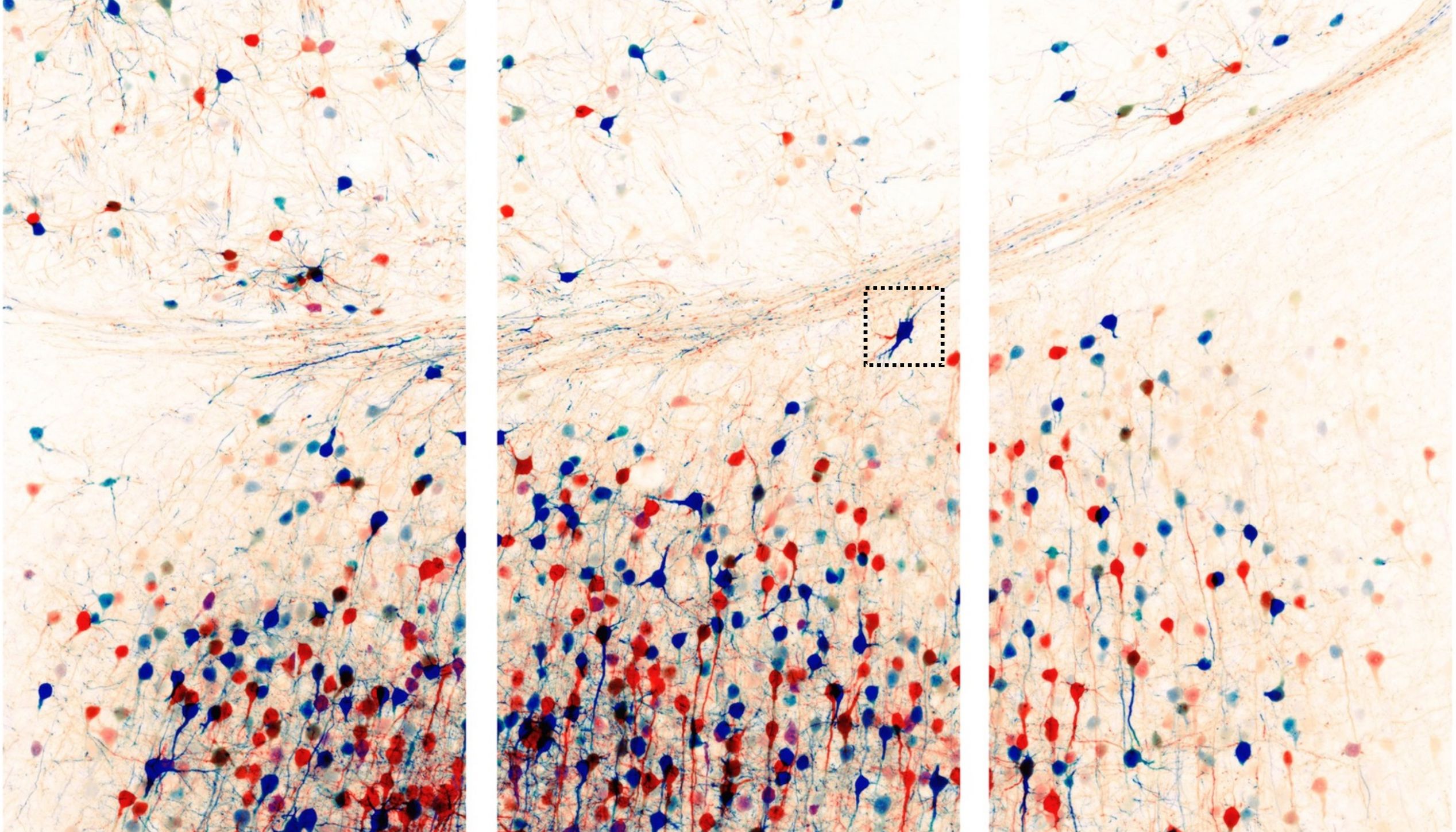
10 MB / frame

=

1GB / second

PMT = ~20%

GaAsP = ~45%



LSM 710

63x 1.4 NA Oil Objective

0.64 x 0.64 µm XY

0.13 µm Z step size

180 slices x~20µm

1024 x 1024 pixels

scan speed 7 (1.58µs)

1 frame = 4 seconds

z stack = 12 minutes

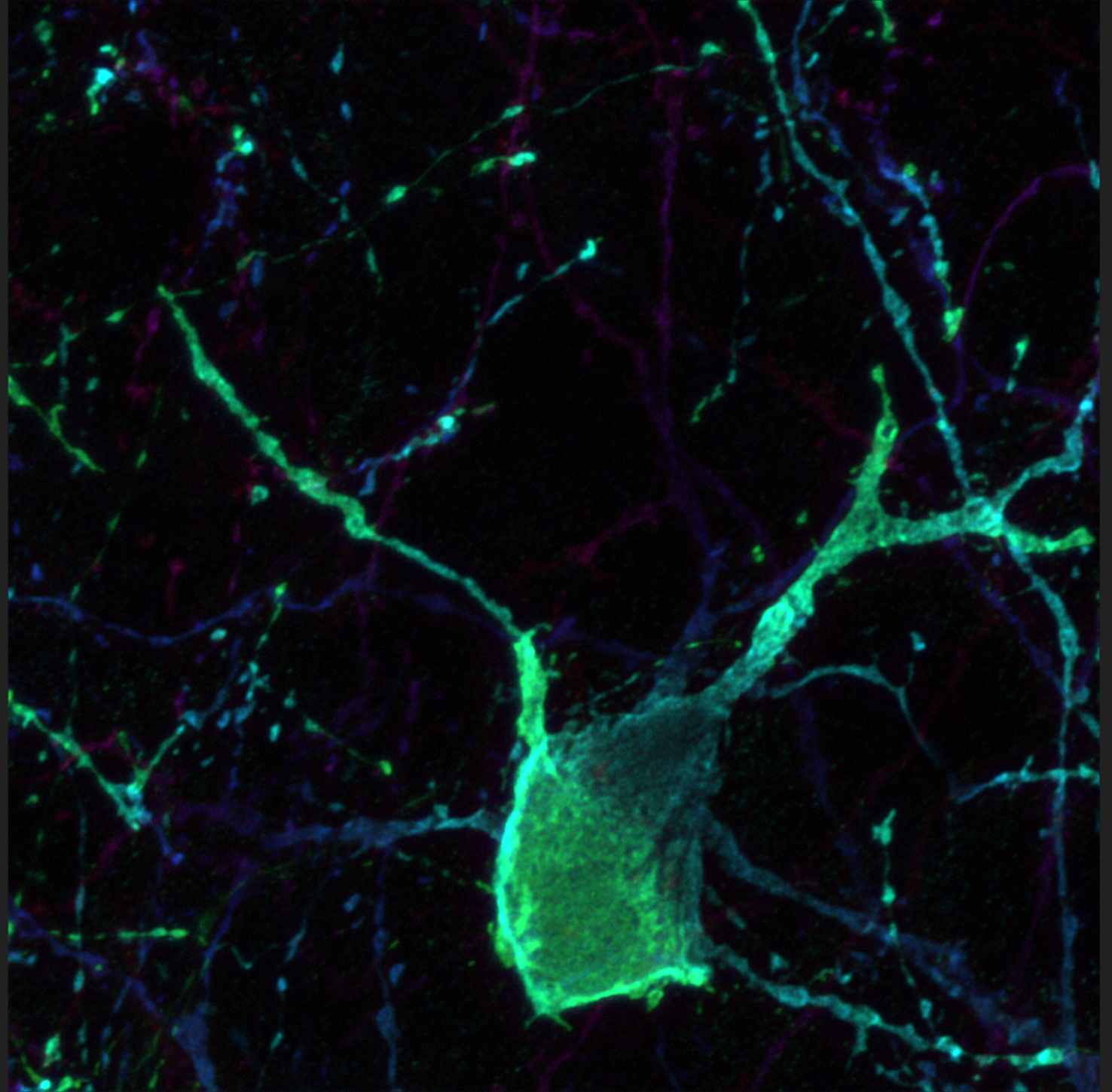
total size = 360MB

Or

0.4 z step

z stack = 4 minutes

total size = 120MB



Yokogawa W1

100x 1.4 NA Oil Objective

0.64 x 0.64 µm XY

0.13 µm Z step size

180 slices x~20µm

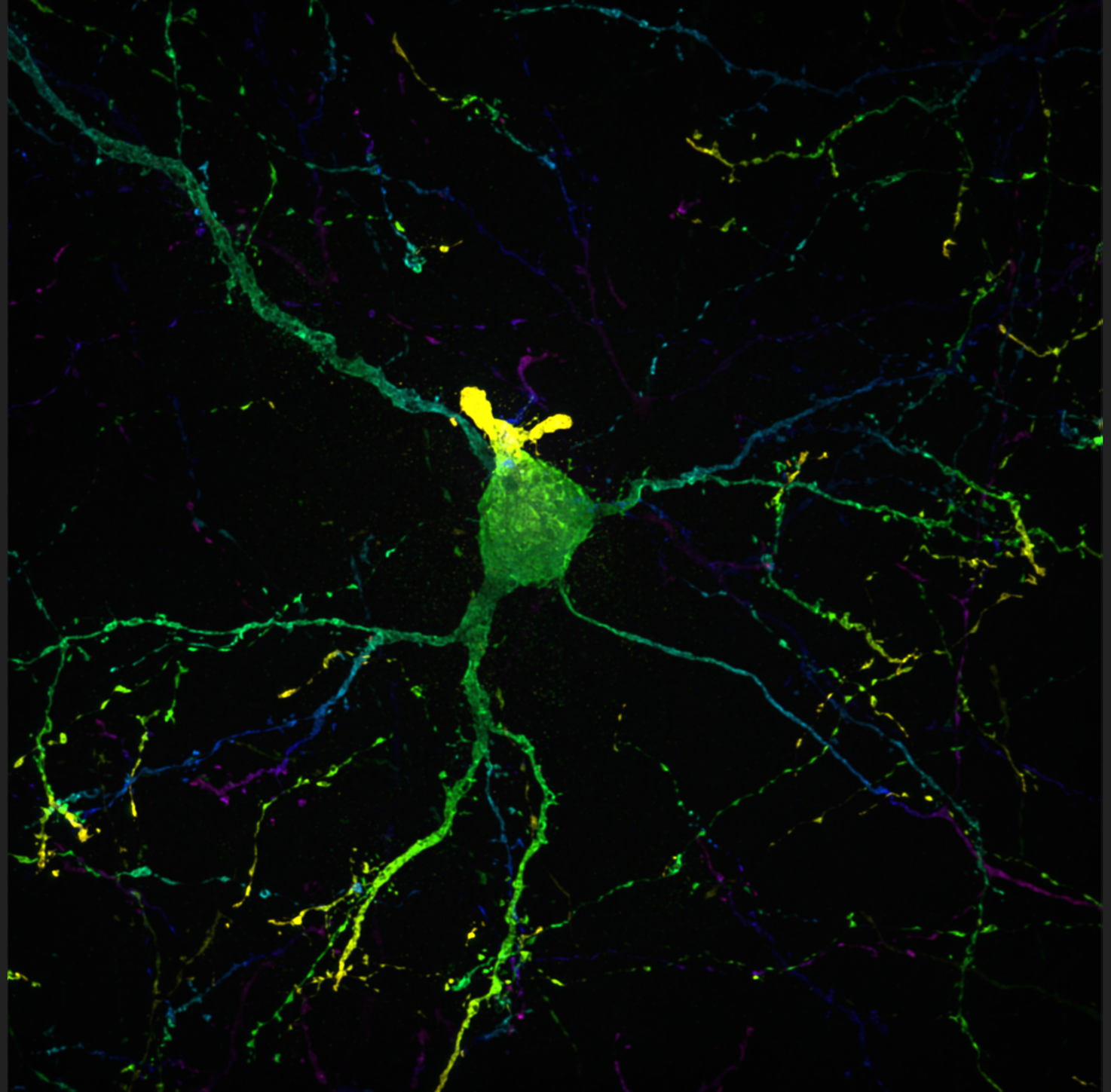
2048 x 2048 pixels

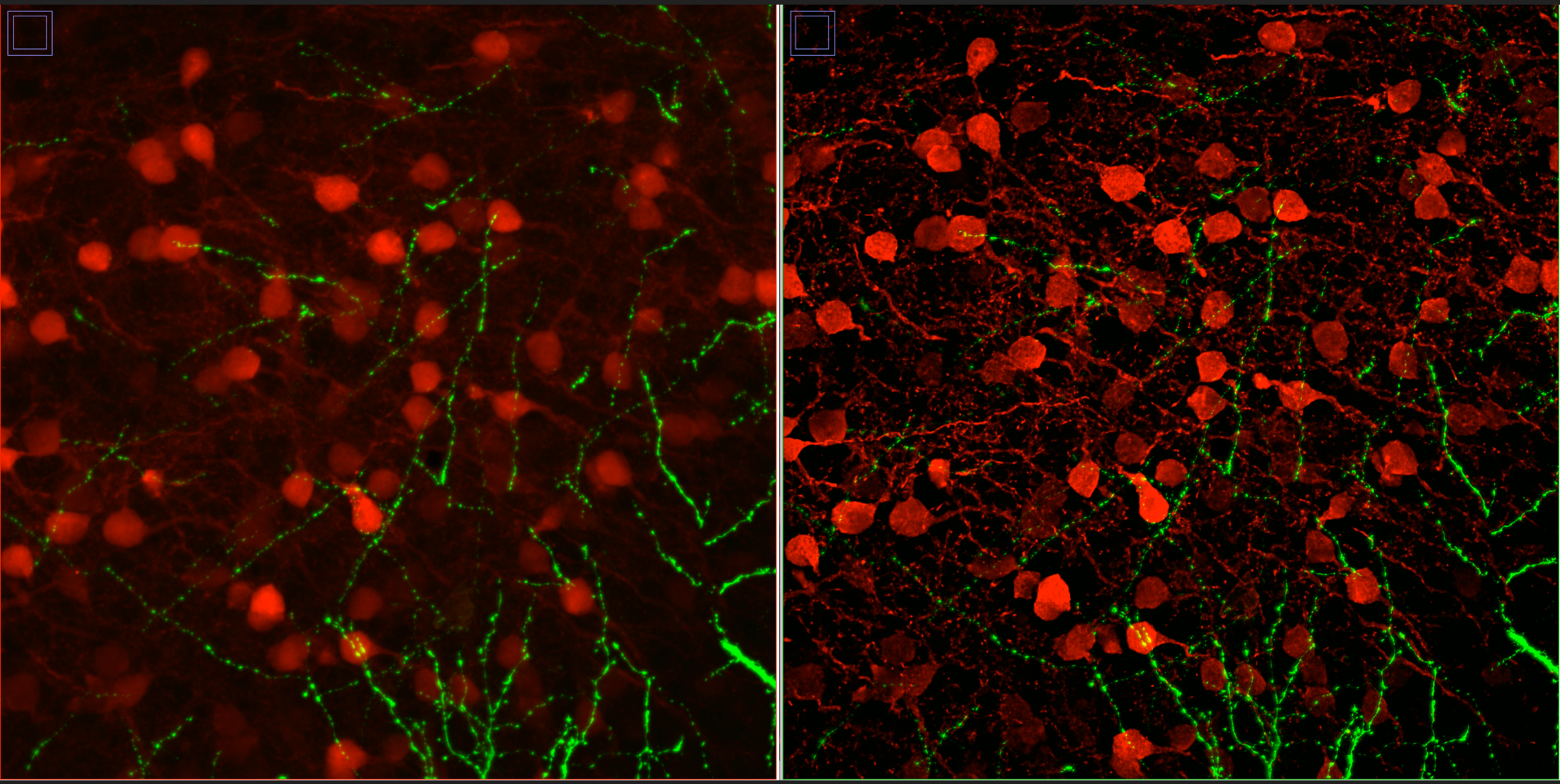
100msec exposure

1 frame = 0.1 seconds

z stack = 50 seconds

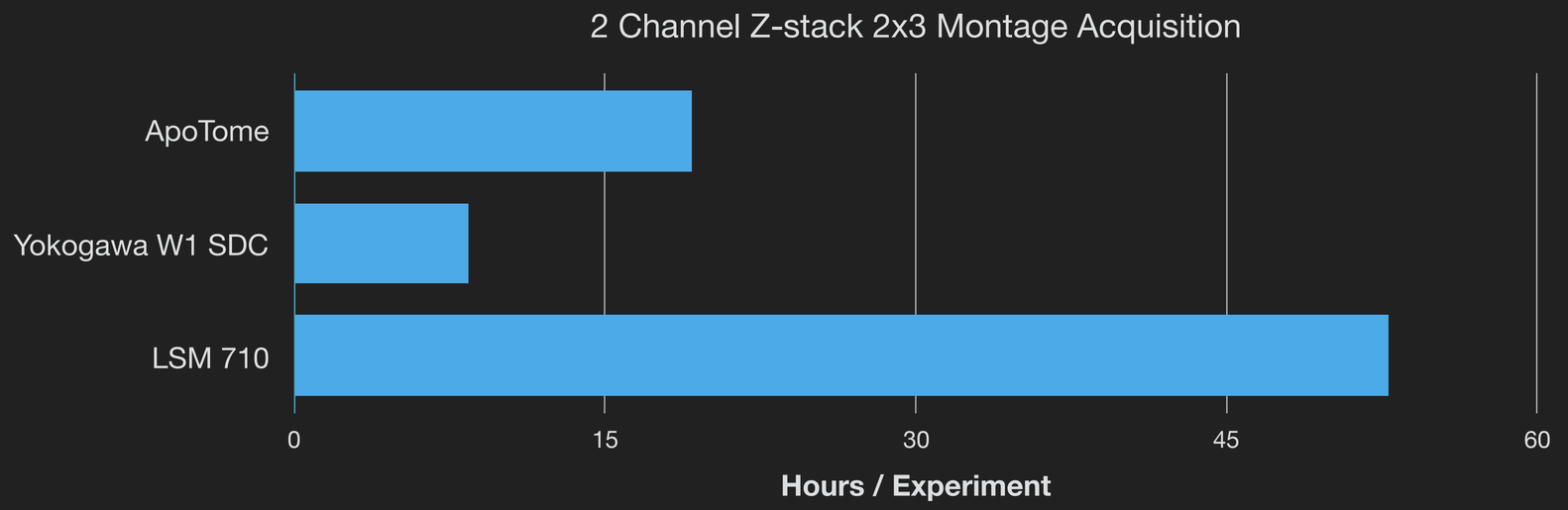
total size = 1.5GB





Diskovery 20x

2048 x 2048 0.3µm

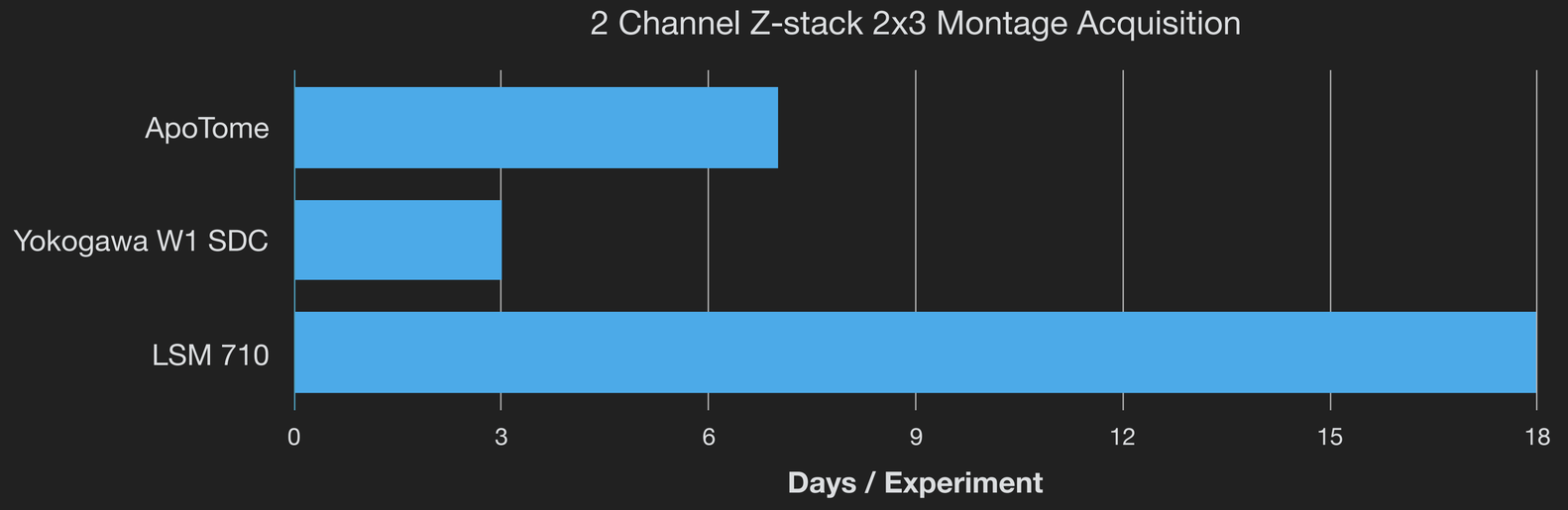


Assuming 2 Conditions

With 6 Animals / Condition

Imaging 12 tissue sections with a 2x3 montage and 15 slices

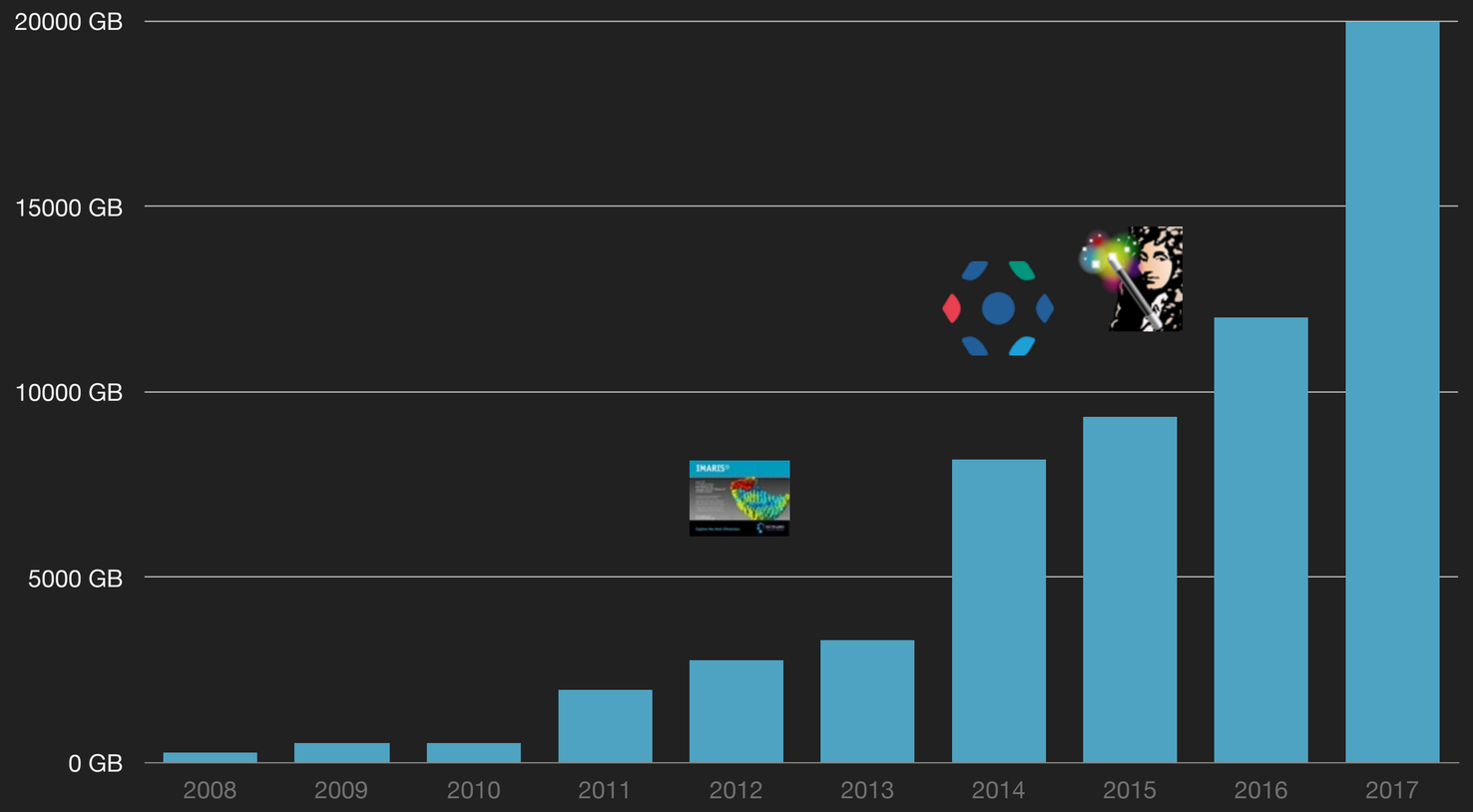
2 channels / 300 msec exposure



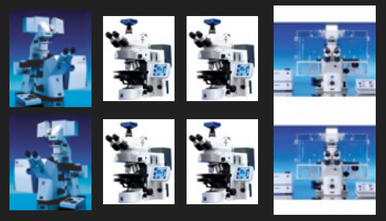
Assuming 3 hours of imaging each day

QBI Microscopy Facility Growth

Data generation
GB /month



2008



2009

2010

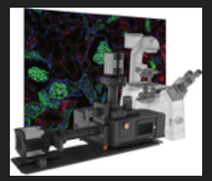
2011

2012

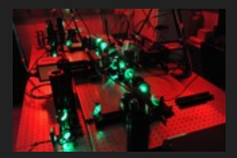
2013

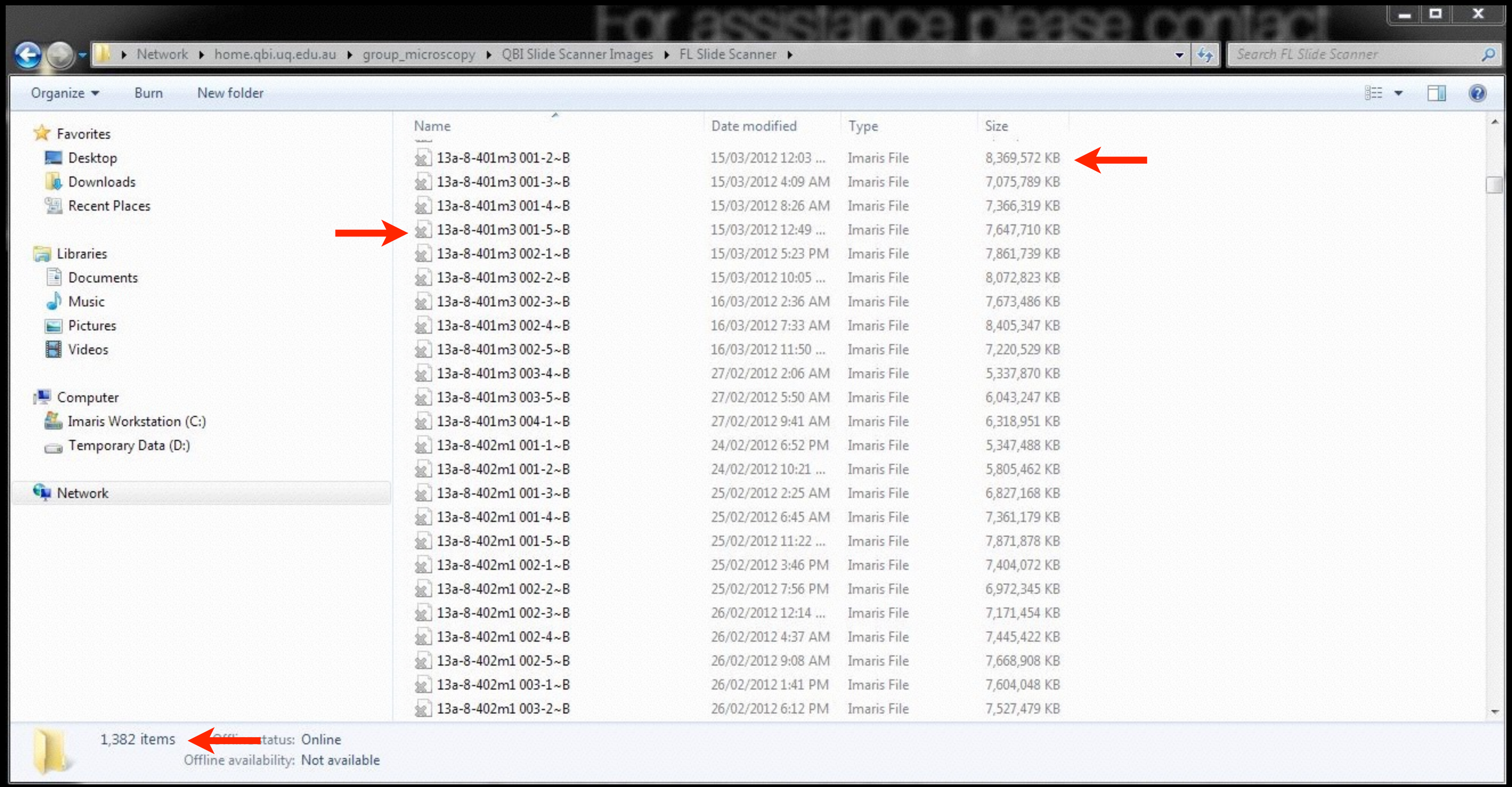
2014

2015

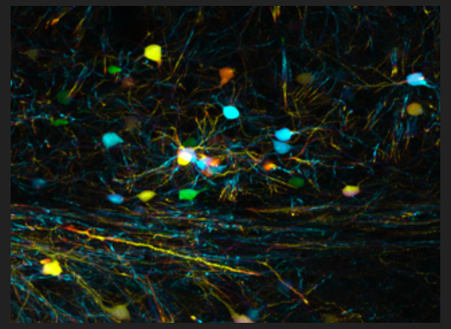


2016



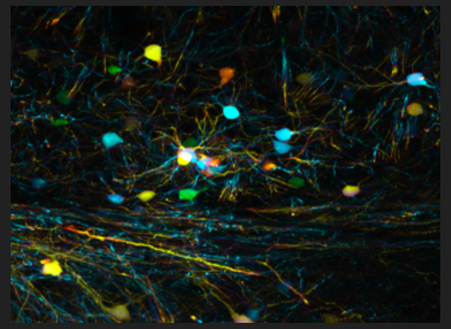

for assistance please contact
Network ▸ home.qbi.uq.edu.au ▸ group_microscopy ▸ QBI Slide Scanner Images ▸ FL Slide Scanner ▸
Search FL Slide Scanner
Organize
Burn
New folder
Favorites
Desktop
Downloads
Recent Places
Libraries
Documents
Music
Pictures
Videos
Computer
Imaris Workstation (C:)
Temporary Data (D:)
Network
Name
Date modified
Type
Size
13a-8-401m3 001-2~B 15/03/2012 12:03 ... Imaris File 8,369,572 KB
13a-8-401m3 001-3~B 15/03/2012 4:09 AM Imaris File 7,075,789 KB
13a-8-401m3 001-4~B 15/03/2012 8:26 AM Imaris File 7,366,319 KB
13a-8-401m3 001-5~B 15/03/2012 12:49 ... Imaris File 7,647,710 KB
13a-8-401m3 002-1~B 15/03/2012 5:23 PM Imaris File 7,861,739 KB
13a-8-401m3 002-2~B 15/03/2012 10:05 ... Imaris File 8,072,823 KB
13a-8-401m3 002-3~B 16/03/2012 2:36 AM Imaris File 7,673,486 KB
13a-8-401m3 002-4~B 16/03/2012 7:33 AM Imaris File 8,405,347 KB
13a-8-401m3 002-5~B 16/03/2012 11:50 ... Imaris File 7,220,529 KB
13a-8-401m3 003-4~B 27/02/2012 2:06 AM Imaris File 5,337,870 KB
13a-8-401m3 003-5~B 27/02/2012 5:50 AM Imaris File 6,043,247 KB
13a-8-401m3 004-1~B 27/02/2012 9:41 AM Imaris File 6,318,951 KB
13a-8-402m1 001-1~B 24/02/2012 6:52 PM Imaris File 5,347,488 KB
13a-8-402m1 001-2~B 24/02/2012 10:21 ... Imaris File 5,805,462 KB
13a-8-402m1 001-3~B 25/02/2012 2:25 AM Imaris File 6,827,168 KB
13a-8-402m1 001-4~B 25/02/2012 6:45 AM Imaris File 7,361,179 KB
13a-8-402m1 001-5~B 25/02/2012 11:22 ... Imaris File 7,871,878 KB
13a-8-402m1 002-1~B 25/02/2012 3:46 PM Imaris File 7,404,072 KB
13a-8-402m1 002-2~B 25/02/2012 7:56 PM Imaris File 6,972,345 KB
13a-8-402m1 002-3~B 26/02/2012 12:14 ... Imaris File 7,171,454 KB
13a-8-402m1 002-4~B 26/02/2012 4:37 AM Imaris File 7,445,422 KB
13a-8-402m1 002-5~B 26/02/2012 9:08 AM Imaris File 7,668,908 KB
13a-8-402m1 003-1~B 26/02/2012 1:41 PM Imaris File 7,604,048 KB
13a-8-402m1 003-2~B 26/02/2012 6:12 PM Imaris File 7,527,479 KB
1,382 items
Offline status: Online
Offline availability: Not available

Acquisition



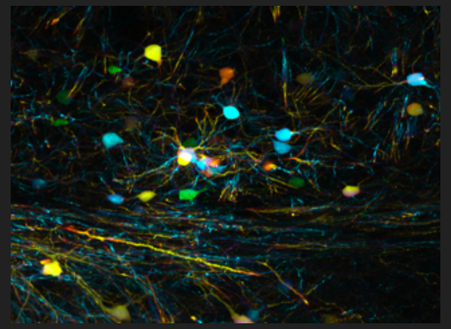
3.5 GB

Server



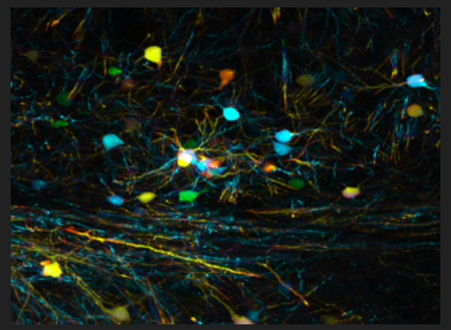
7 GB

Duplicate to
shared folder



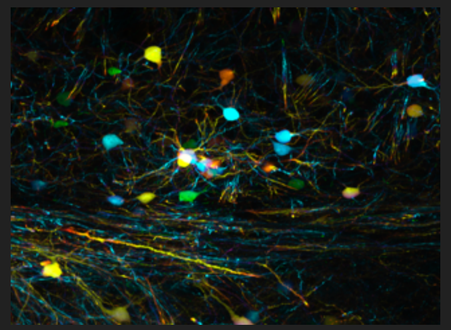
10.5 GB

Analysed to
server



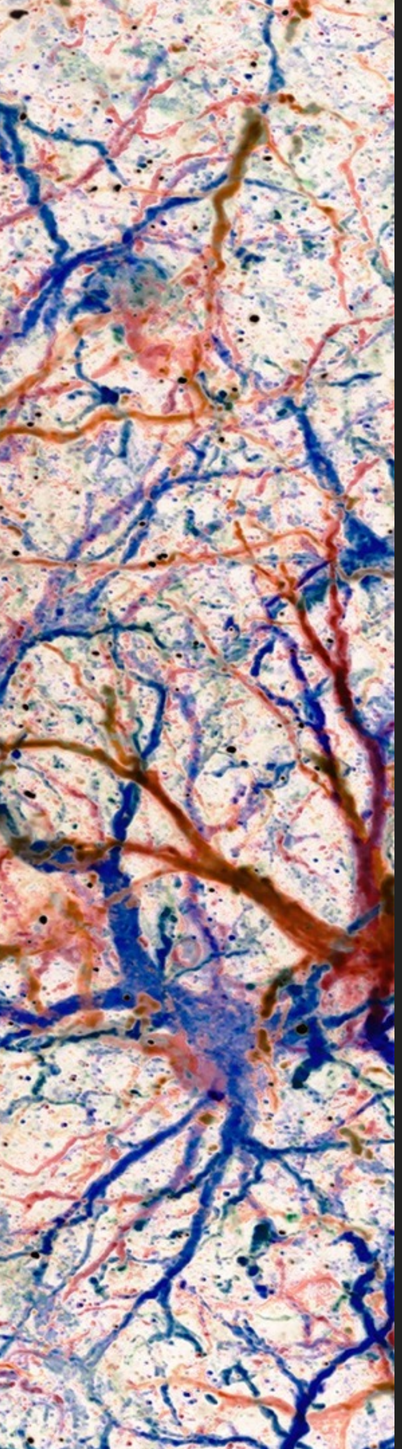
14.5 GB

Copied to
private folder

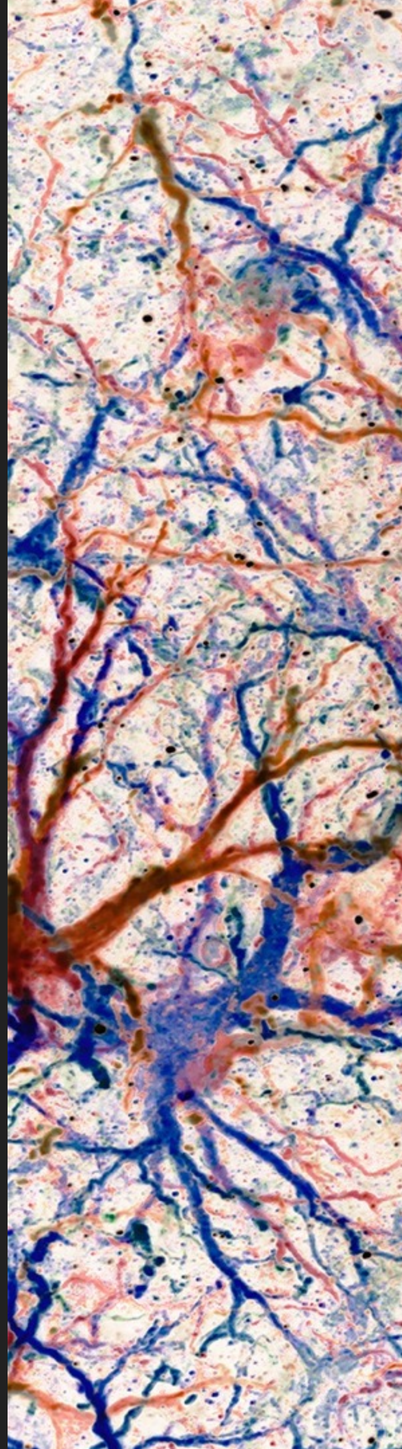


18 GB

1. Need to be able to preview image data before accessing it
2. Data preservation when researchers leave lab
3. Prevent unnecessary duplication of image data
4. Facilitate server based analysis pipelines and deter ad hoc manual approaches
5. Reanalysis and sharing of analysis tools



OMERO @ QBI



Polaris

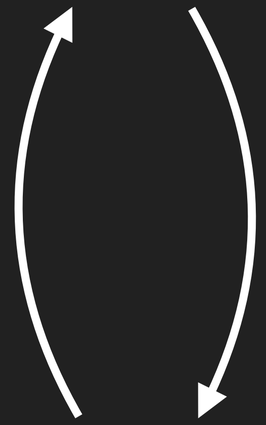
nectarcloud

openstack

OMERO

16 VCPU

64GB RAM

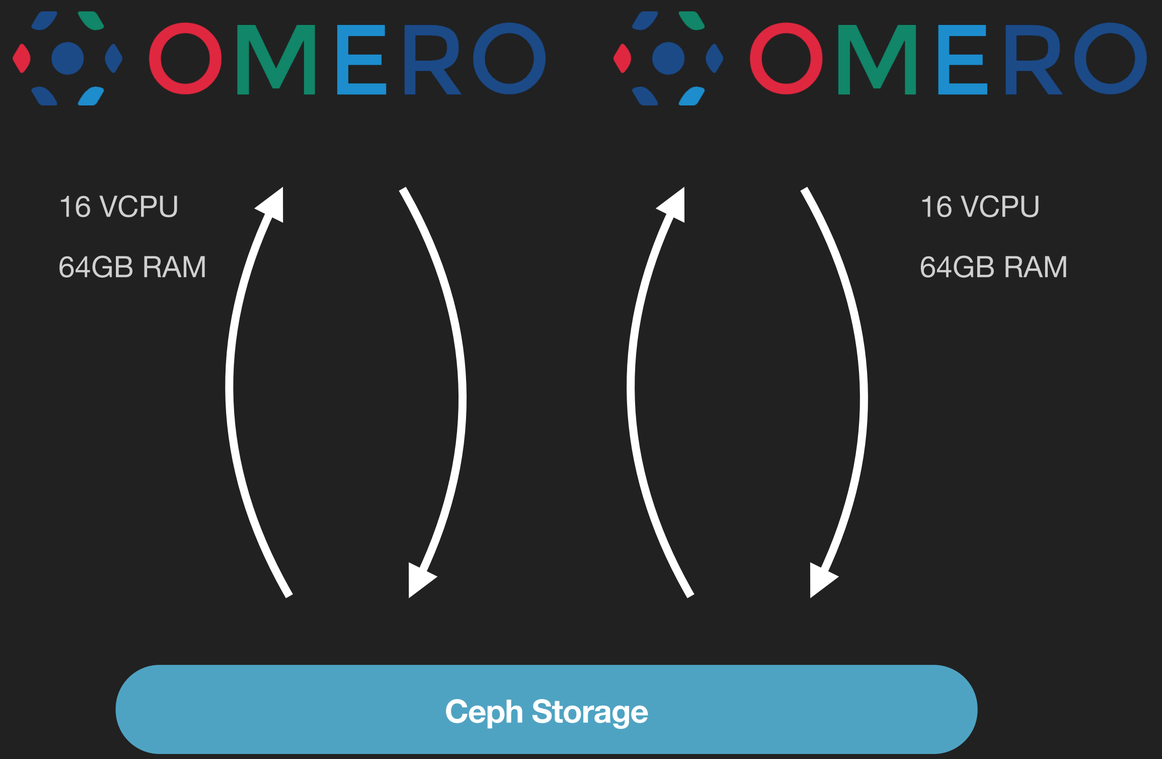


Ceph Storage

Polaris

nectarcloud

openstack



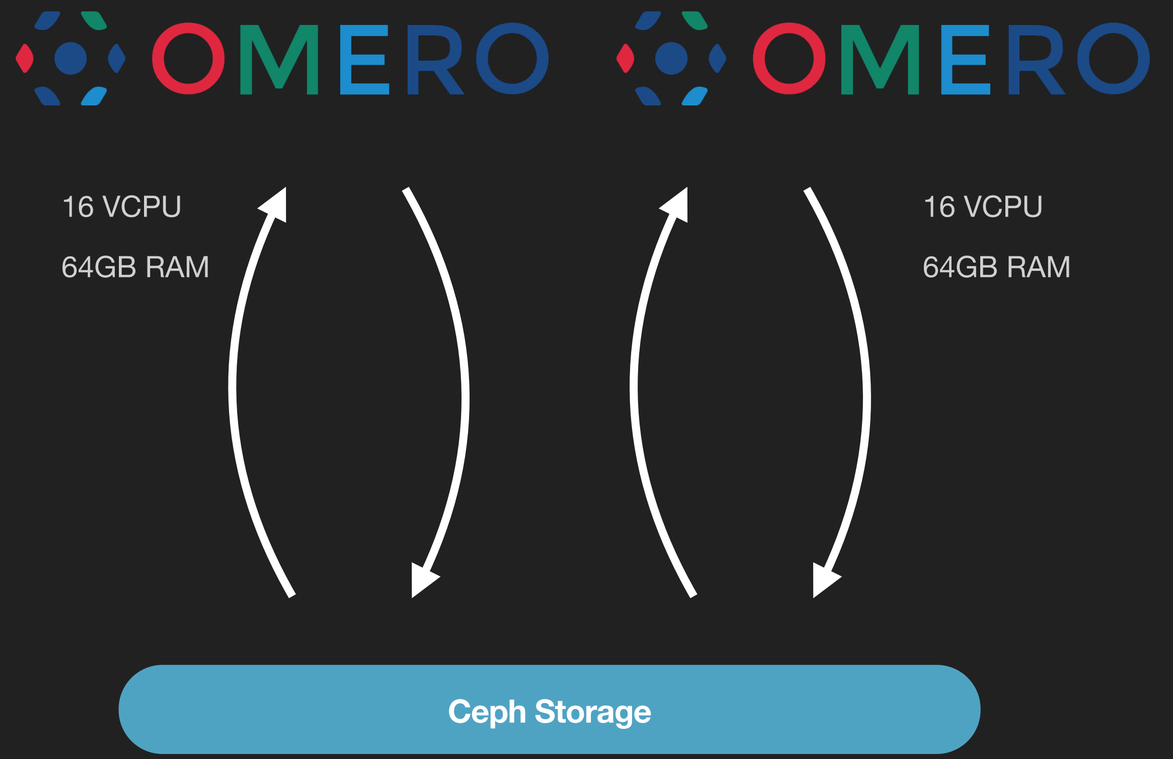
If OMERO gets too busy (e.g. >90% load) a new instance is spawned

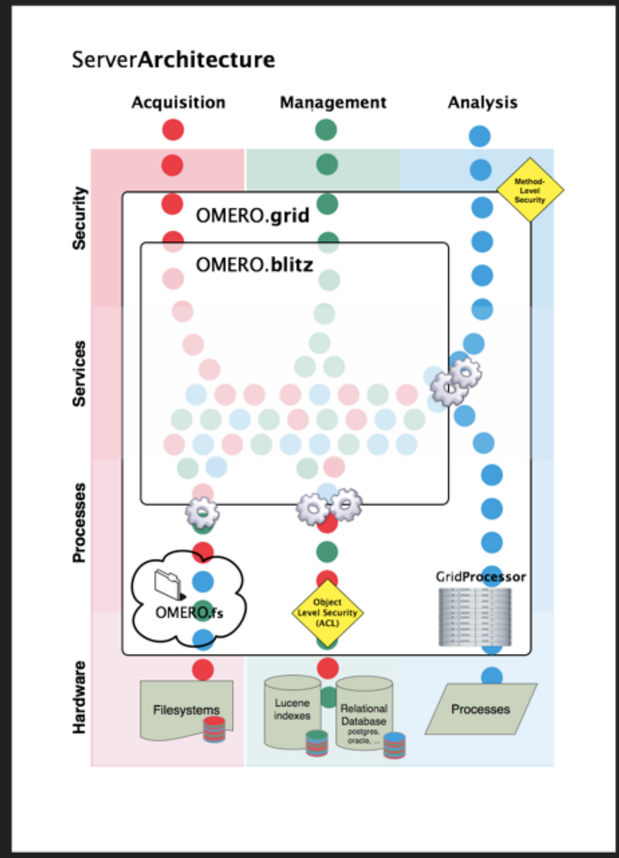
1. Need to be able to preview image data before accessing it
2. Data preservation when researchers leave lab
3. Prevent unnecessary duplication of image data
4. Facilitate server based analysis pipelines and deter ad hoc manual approaches
5. Reanalysis and sharing of analysis tools

Polaris

nectarcloud

openstack





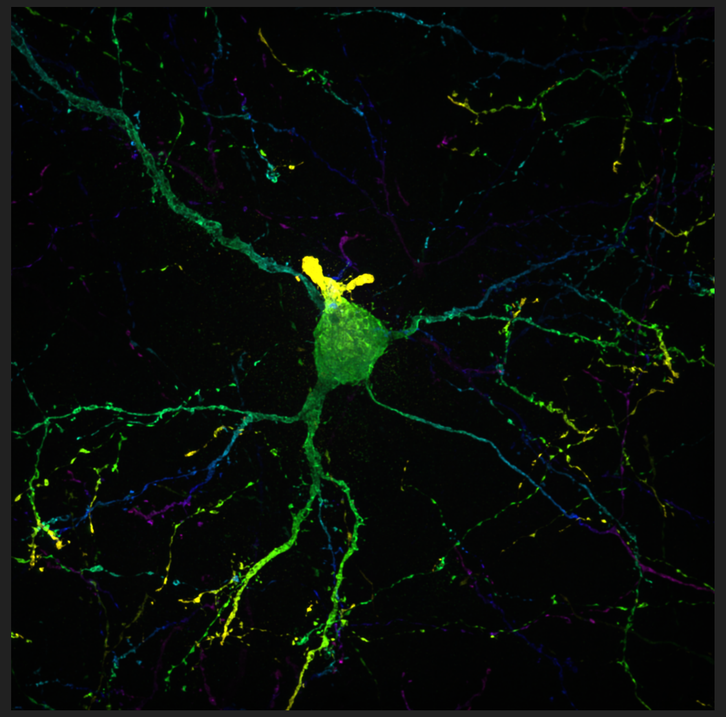
Openstack : Orchestration

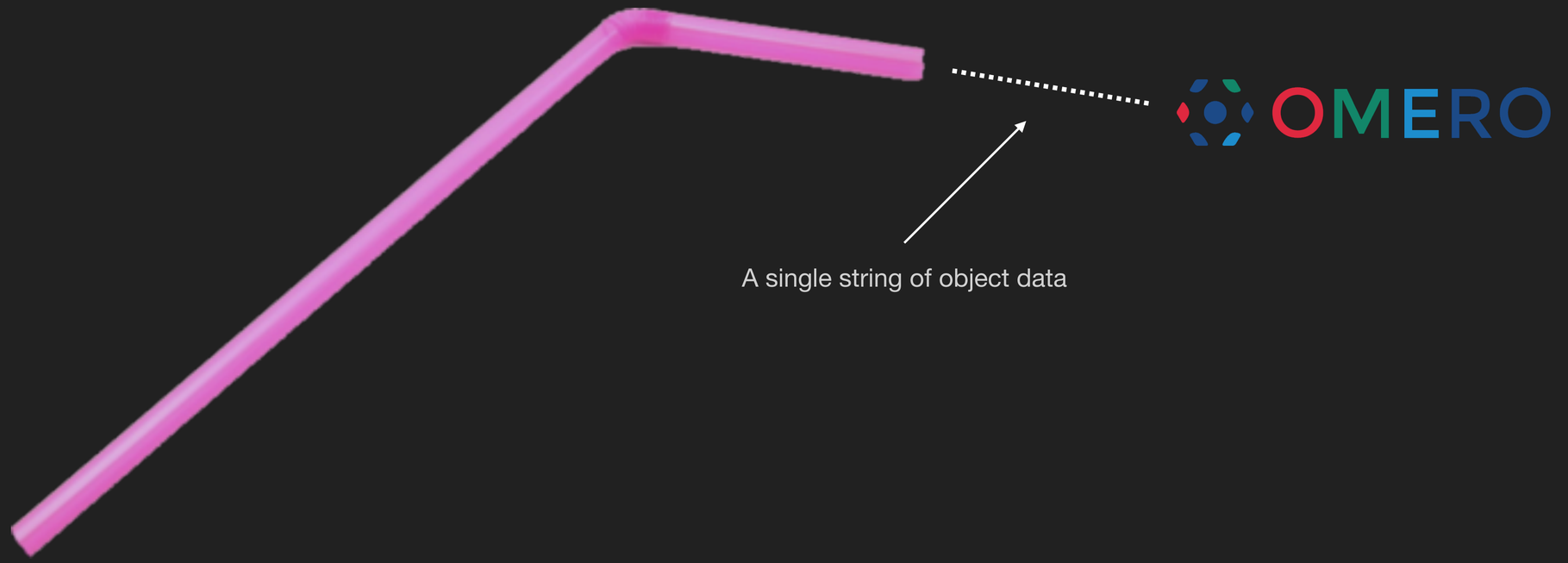
If OMERO gets too busy (e.g. >90% load) a new instance is spawned

BUT new instances are read only to Ceph Storage

The Blitz Gateway

The Blitz Gateway





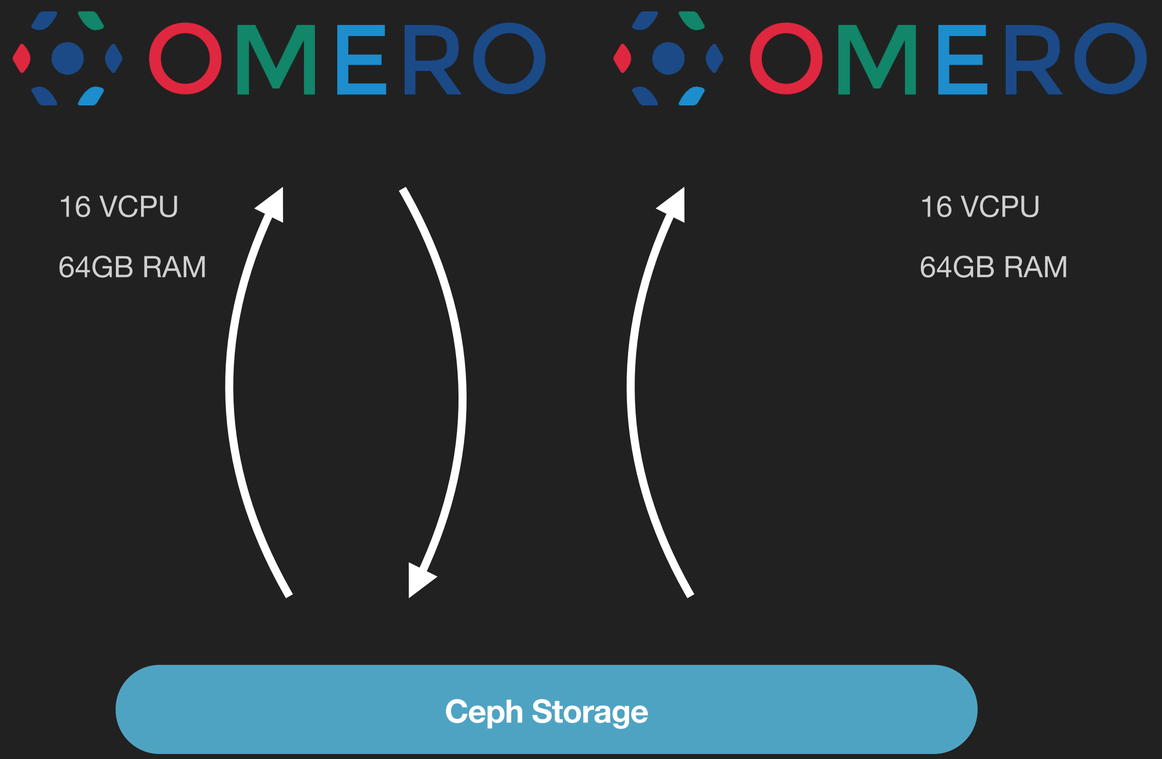
A single string of object data

* Blitz is also non scalable

Polaris

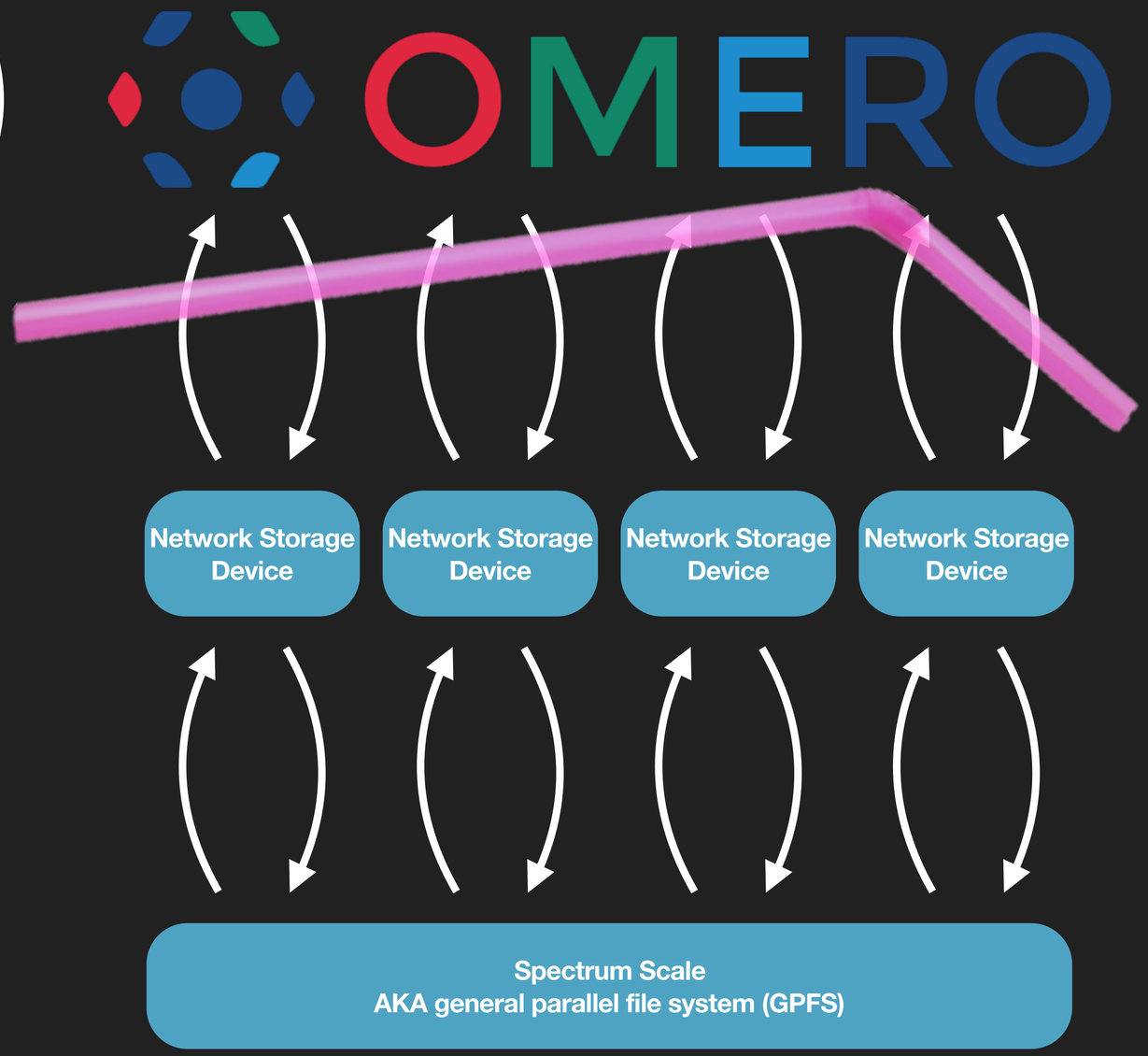
nectarcloud

openstack



Polaris

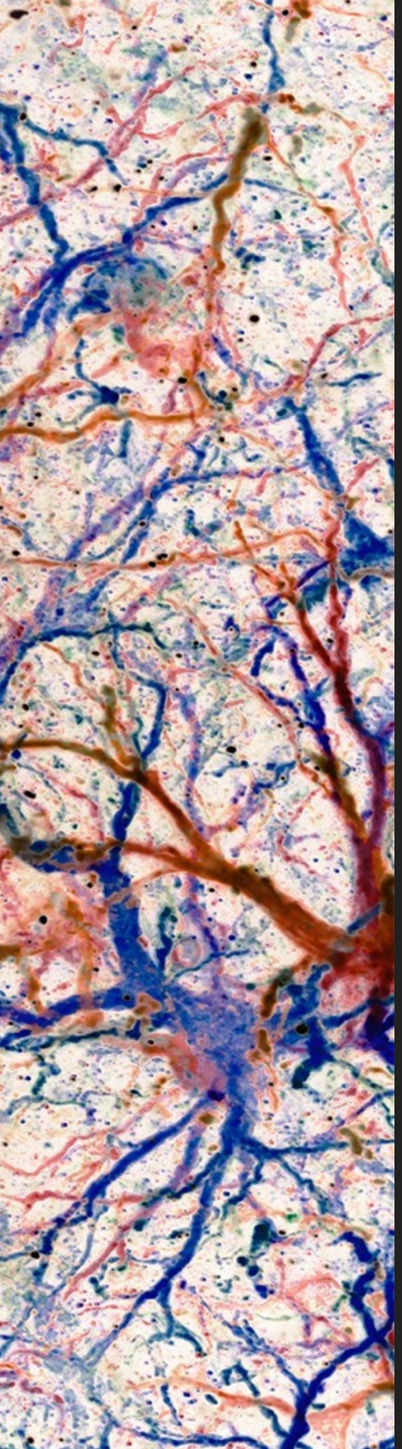
nectarcloud



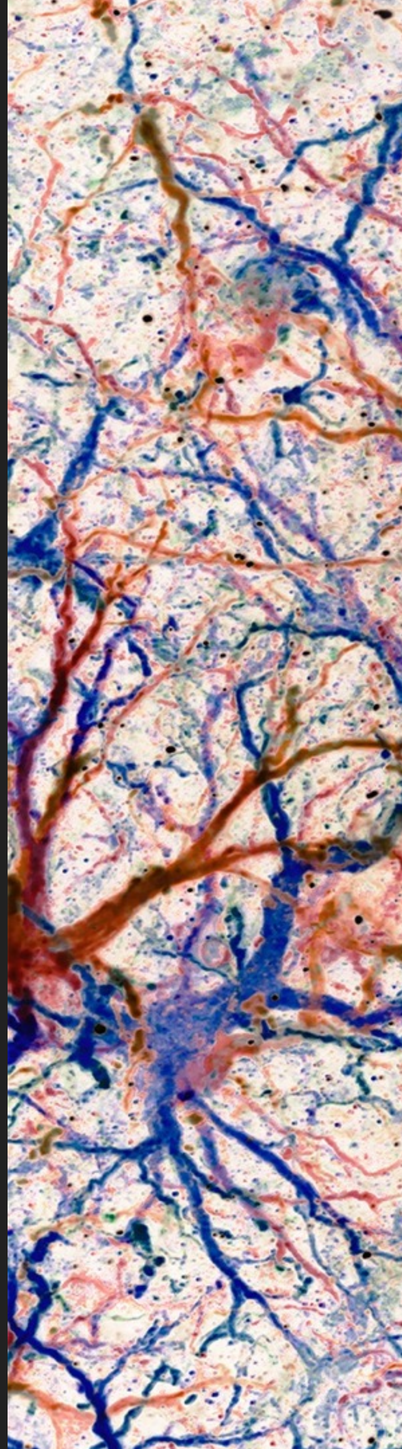
16 VCPU

64GB RAM

* no longer requires openstack as a single instance is very fast



OMERO + big data




dotroom

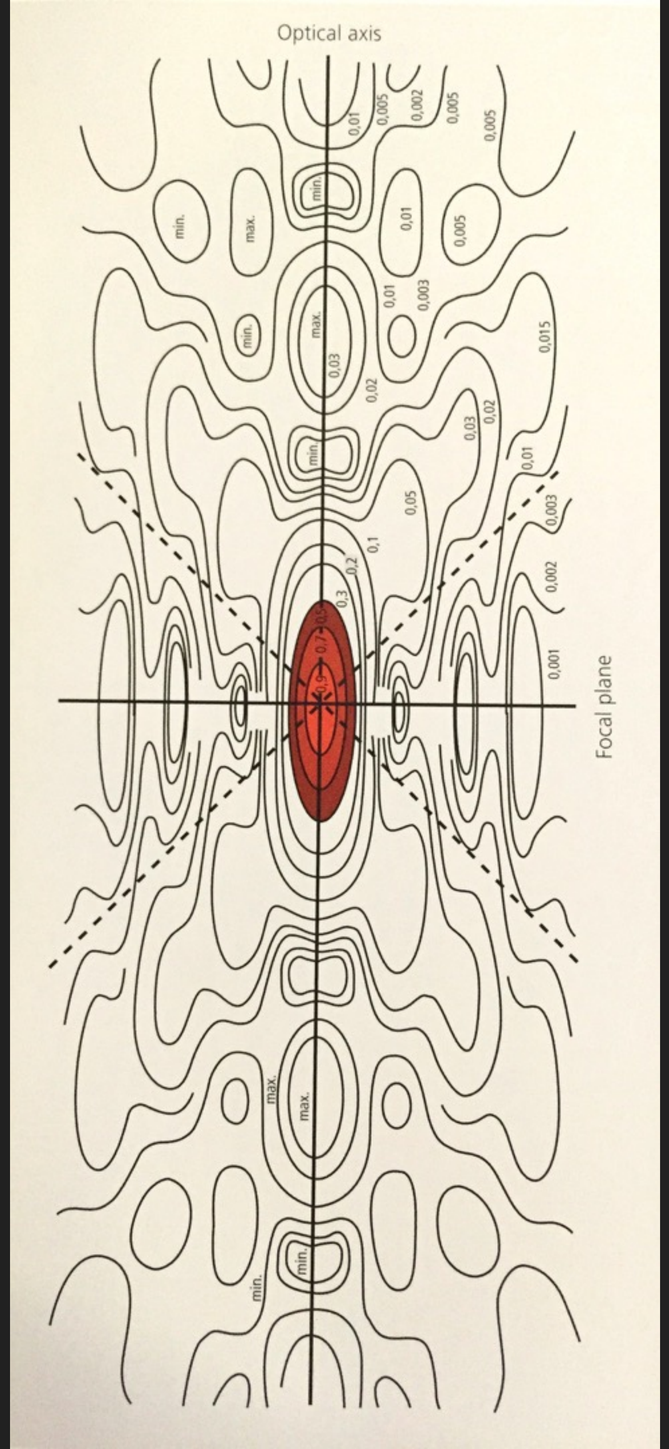

IMARIS 8
Setting main frame...
Imaris x64
8.1.2 [Jun 3 2015]
Build 36825 for x64
Copyright © 1993-2015 Bitplane AG
www.bitplane.com
welcome@bitplane.com
BITPLANE
an Oxford Instruments company
bitplane.com


WHEN
SIZE
MATTERS

Deconvolution

Increases image resolution
Increases contrast / SNR
Makes analysis easier / possible

SLOW - 10 - 20 minutes/stack

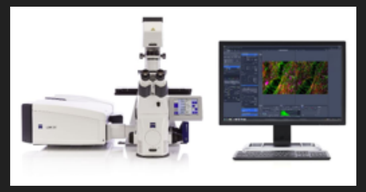


Huygens GPU Deconvolution server

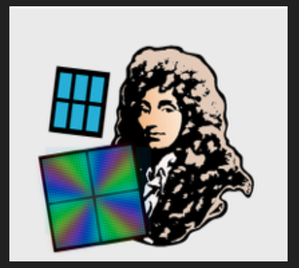
Accessible from anywhere in QBI
Remote and batchable processing

FAST - 20 - 30 seconds / stack

..but teaching researchers to connect
to a remote computer is not always easy



OMERO



Large nVidia K80 GPU array

OMERO.Fs

work directly with OMERO.Fs and avoid front end / blitz + ICE framework

direct FTP access API or similar to allow queued uploading with correct permissions

back end workflows can make full use of hardware

Essentials required to meet the needs of advanced imaging facilities:

1. Support for large scale images

- Support for large scale images should not be limited. Images 10GB+ are becoming the norm and light-sheet techniques expand this further.

- Data Ingestion:
  - 56432 x 152960px (15.3GB)
  - ~20 mins (6mins to upload, remainder to process into OMERO)
- Extracting ROIs:
  - 77119 x 149376px (7.45GB)
    - 2xpolygon ROIs each (20113 x 31321 px) (1.75 GB)
    - 47 mins
  - 56432 x 152960px (15.3GB)
    - extracting 9xROIs (each 20113 x 31321 px)
    - Session timeout after 6 hrs - not completed

Essentials required to meet the needs of advanced imaging facilities:

1. Support for large scale images
   - An exporter client similar to the importer could help to manage this problem and allow the appropriate download procedures to occur in the background.
   - Can we avoid the Blitz gateway?
     - Appears to constrain the handling of binary data and it is assumed that this is where most of the problems in image processing lie
   - We need a parallel mechanism
   - Hadoop FS - works in objects and blocks + add nodes and get more power

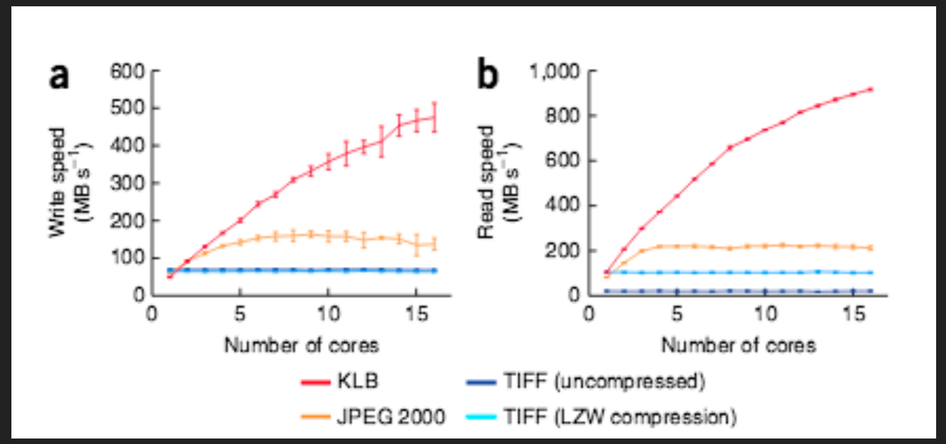
- Requires a solution that is built to scale and can drive up hardware utilisation

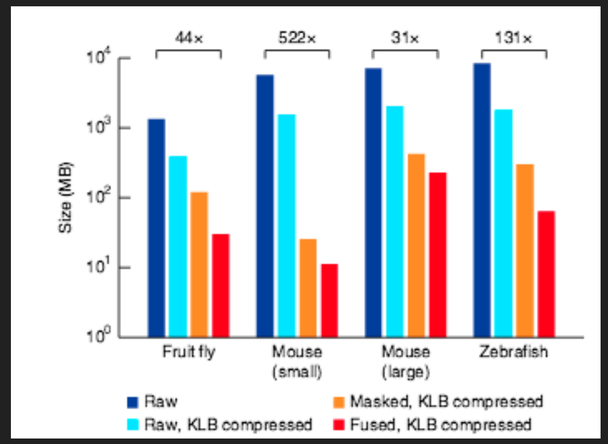
Essentials required to meet the needs of advanced imaging facilities:

- Can we avoid working in raw pixel data?
  - Support for compressed image data rather than having to deal with raw uncompressed pixel data could help speed things up. The time required to manage raw pixel data compared to compressed pixel data can be 10 to 100 fold difference – with datasets of 100's of images, this has a major impact on the user.
  - Attempts to handle ROI extraction in large images by tiling the image has resulted in scripts taking over an hour to complete, beyond the standard timeout of the user session. If the pixel data was compressed, this processing time could be drastically reduced.
  - Attempts to increase the session timeouts has resulted in broken sessions – better integration between the Python and Java API might solve this.

Essentials required to meet the needs of advanced imaging facilities:

- Alternatives to OME TIFF
  - TIF and LZW-compressed TIFF = no benefit from multicore CPU architecture
  - HDF5 format - data blocks for fast retrieval but slow write (no parallel writing)
  - Keller Lab Block (KLB) - data blocks for fast retrieval, parallel write so scales with # of CPU





Efficient processing and analysis of large-scale light-sheet microscopy data

Fernando Amat, Burkhard Höckendorf, Yinan Wan, William C Lemon, Katie McDole & Philipp J Keller

NATURE PROTOCOLS | VOL.10 NO.11 | 2015 | 1679

Essentials required to meet the needs of advanced imaging facilities:

2. Support for external processing

- A simple and unhindered means of transferring data in and out of OMERO for external processing (e.g. GPU-cluster analysis)
- One possibility: A direct FTP access API similar to the OMERO Importer to enable queued uploading to OMERO with appropriate permissions and relationships (as these would be large image files, preferably avoiding the restrictions of the ICE framework).

Has the trust of researchers and a great community behind it

but needs scalability and improved hardware utilisation

Microscopy has transitioned from a time of single snapshots to an era of comprehensive imaging

OMERO needs to make this same transition

Parallel file systems help, but only go so far

by liberating OMERO from the blitz gateway imaging research everywhere will become easier and faster

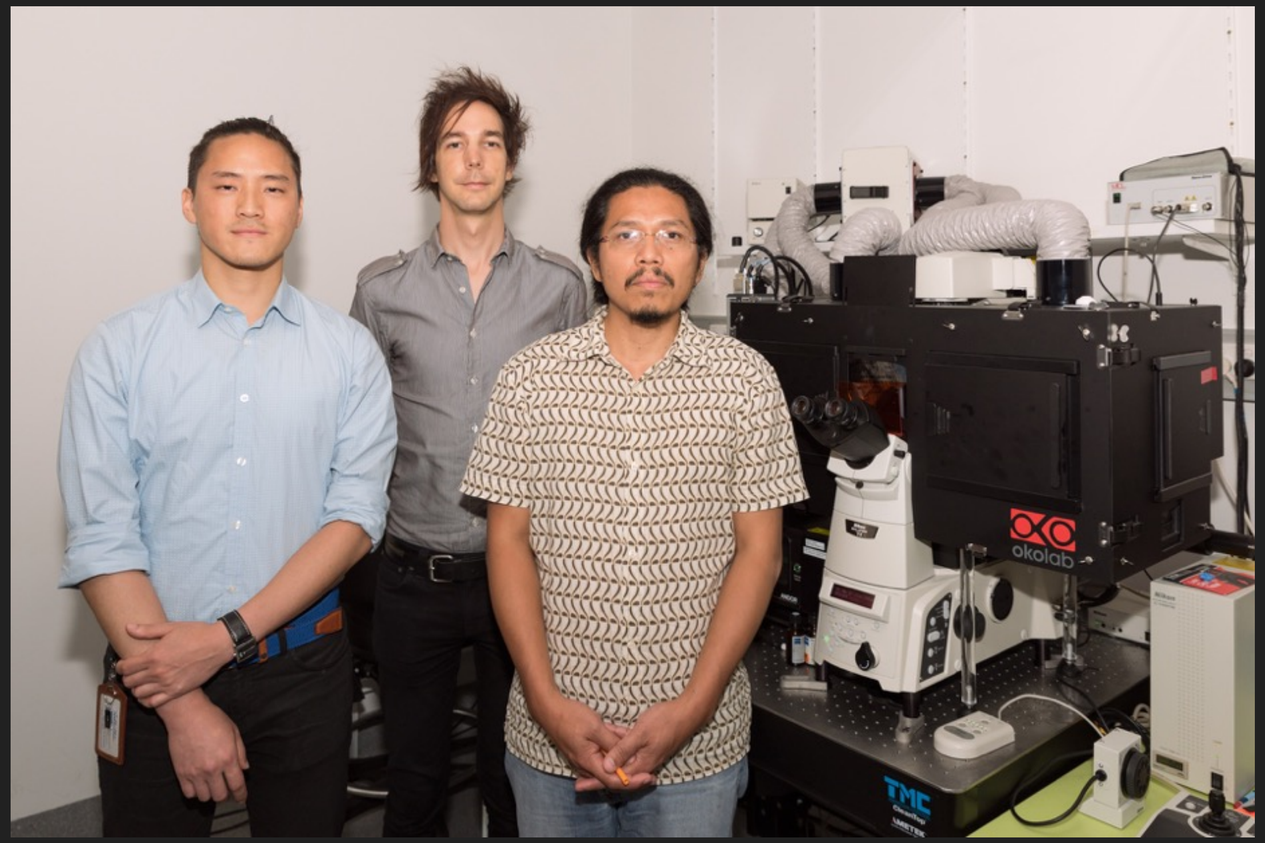
Thank you

Jason Swedlow and the OME team!

Jeremy Ullmann / Boston Children's Hospital

Rumelo Amor and Arthur Chien / QBI Microscopy

Jake Carroll / QBI IT



Arthur Chien Rumelo Amor

Jake Carroll

Liz Cooper-Williams
Dan Matthews

The University Of Queensland AUSTRALIA | QBI
Queensland Brain Institute