

Día Virtual sobre microscopía

Procesamiento de imágenes de microscopía: cálculo y almacenamiento con apoyo de redes de alta velocidad

Mauricio Cerda, Victor Castañeda



*Laboratory of Scientific Image Analysis (SCIAN-Lab)
Biomedical Neuroscience Institute (BNI)
Programa de Anatomía y Biología Del Desarrollo,
ICBM, Facultad de Medicina,
Universidad de Chile*

Santiago, 16.06.2015



Basic Science

R&D

Human Capital Formation
 Magister Informática Médica

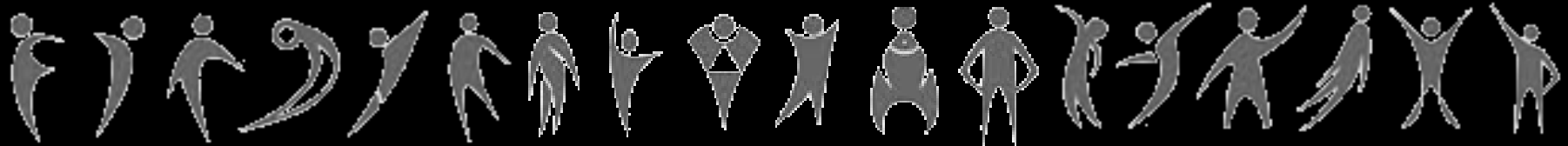
Basic Science

FONDECYT
 CONICYT

FONDEF
 CONICYT

DAAD / DFG

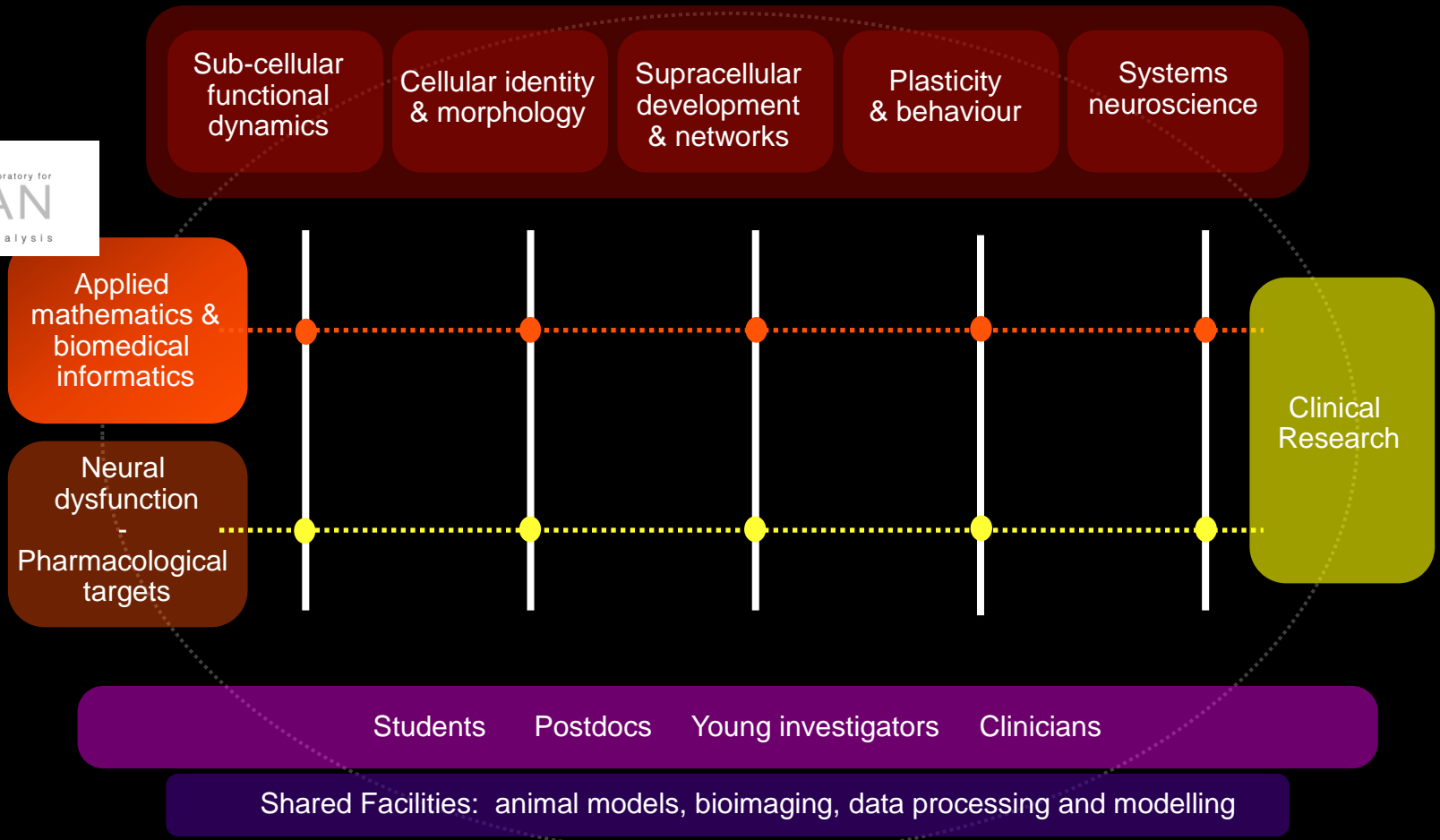
ICM

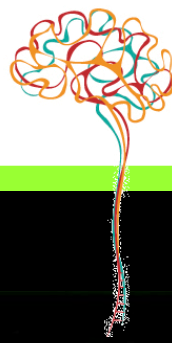


SCIAN-Lab Members

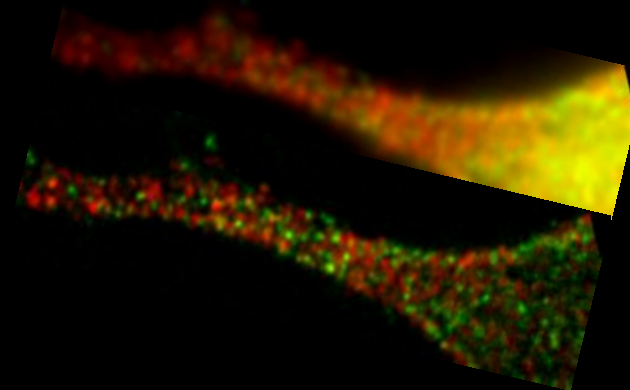
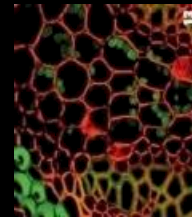
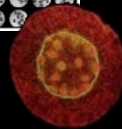
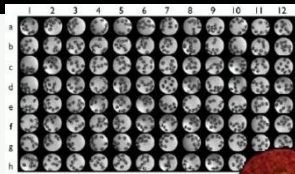
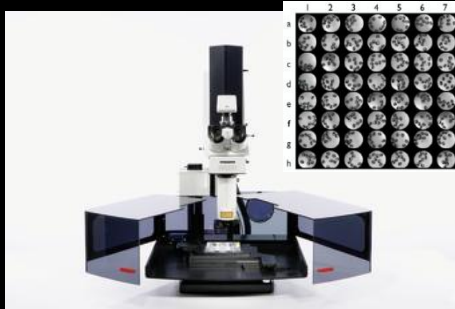
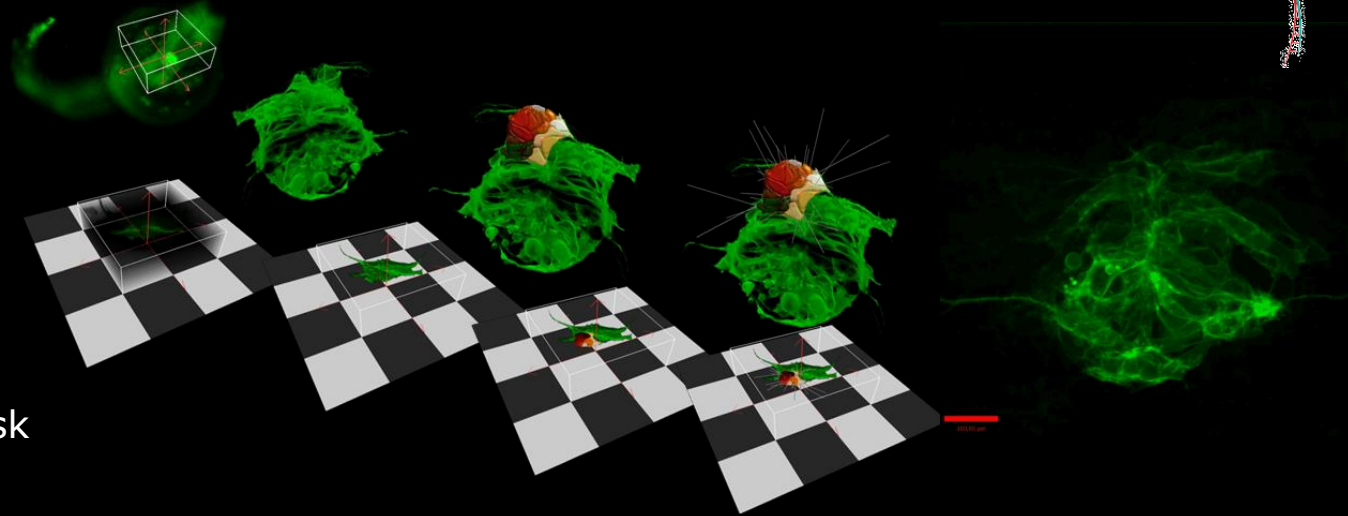
PI	•	Biophysics
PostDocs	•••	Biology / Computer Science / Electric Engineer
PhD - students	••••	Computer Sc / Electric Engineer / Biology
Master - students	•••••• (20)	Medical Technology / Electric Engineer / Medical Informatics
Undergraduate	•	Computer Sc
Research – Assistants	•••••	Medicine / Computer Sc / Electrical Engineer / Biology
Technicians	•••	Biotechnology / Labtechnician / Administration

Organización BNI

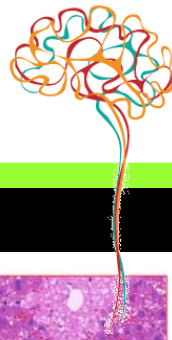




Perkin Elmer Spinning Disk (2008)



Leica TCS LSI: Super Zoom Confocal + Superresolution Optical Fluctuation Imaging SOFI (2013)

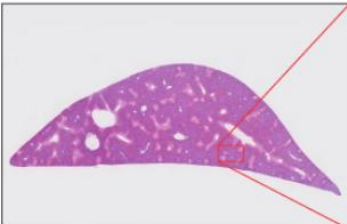


- One slide ~30Gb uncompressed
- High speed (30s per slide)
- 350 Slides per batch
- Z-Stack capacity
- 3-channel fluorescence
- Scans 2.5x7.5cm @ 230nm pixel

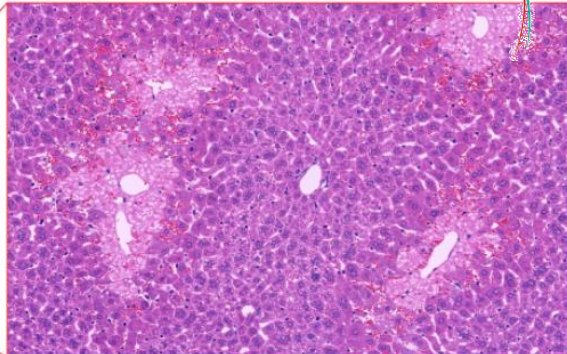
Hamamatsu Tissue Scanner, whole slide.
(2014)

● Toxicity test using H&E staining

NanoZoomer is available for a wide range of applications including toxicity evaluation.



▲ Liver in mouse with a dose of acetaminophen in 4 hours (Whole Image)

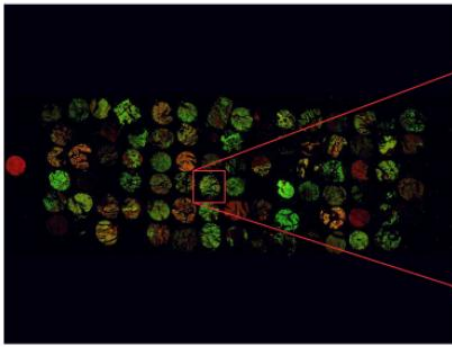
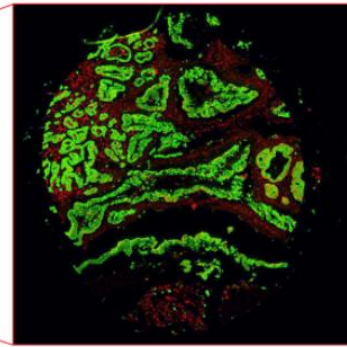


▲ Liver in mouse with a dose of acetaminophen in 4 hours (Magnified Image)

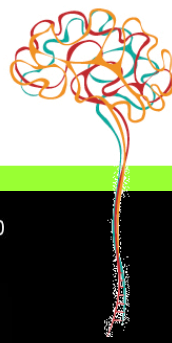
Provision of Glass Slide
 Courtesy of Dr. June Kanino Division of Toxicology, Biological Safety Research Center, National Institute of Health Sciences, Japan

HAMAMATSU
 PHOTON IS OUR BUSINESS

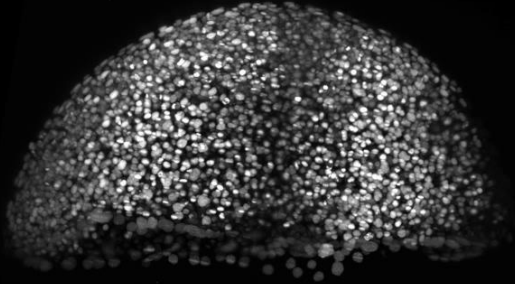
● Observing multi-wavelength fluorescence image of TMA

TMA images were acquired in the course of joint research with the National Cancer Center Research Institute, Department of Chemotherapy and Department of Pathology.



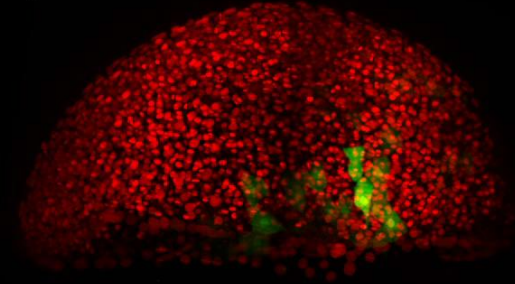
H2B-mCherry



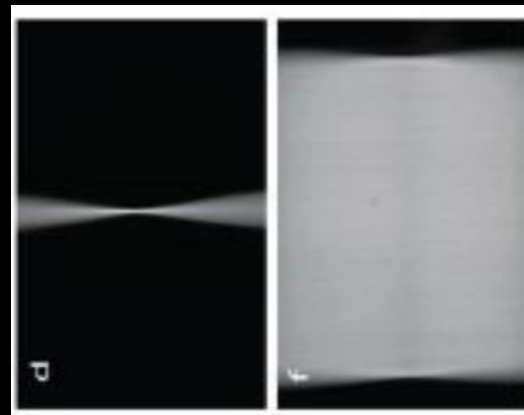
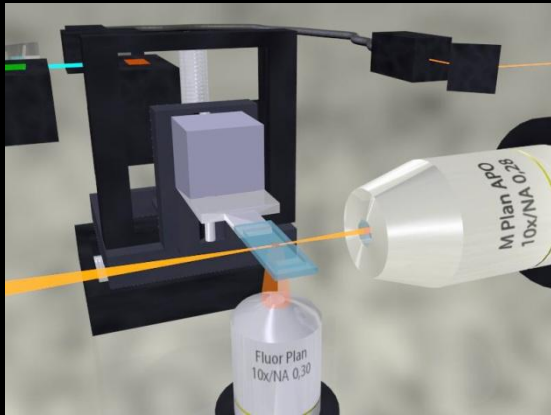
crestin::GFP



merge



00:00:00

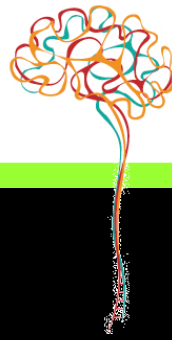


Pulgar, Keller, Concha, unpublished

Keller et al, Science 322, 2008

Lightsheet microscope, custom build in Germany-Chile
(2015-2016)

Microscopio	Fecha Instalación	Experimento típico (GB)
Spinning disk confocal	2008	2
Super Zoom confocal	2013	20
Tissue Scanner	2014	30
Ligth-sheet	2015	5000



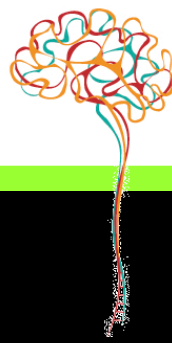
¿Cómo lo manejamos hasta 2014? ... ¡Lo mejor que podemos!



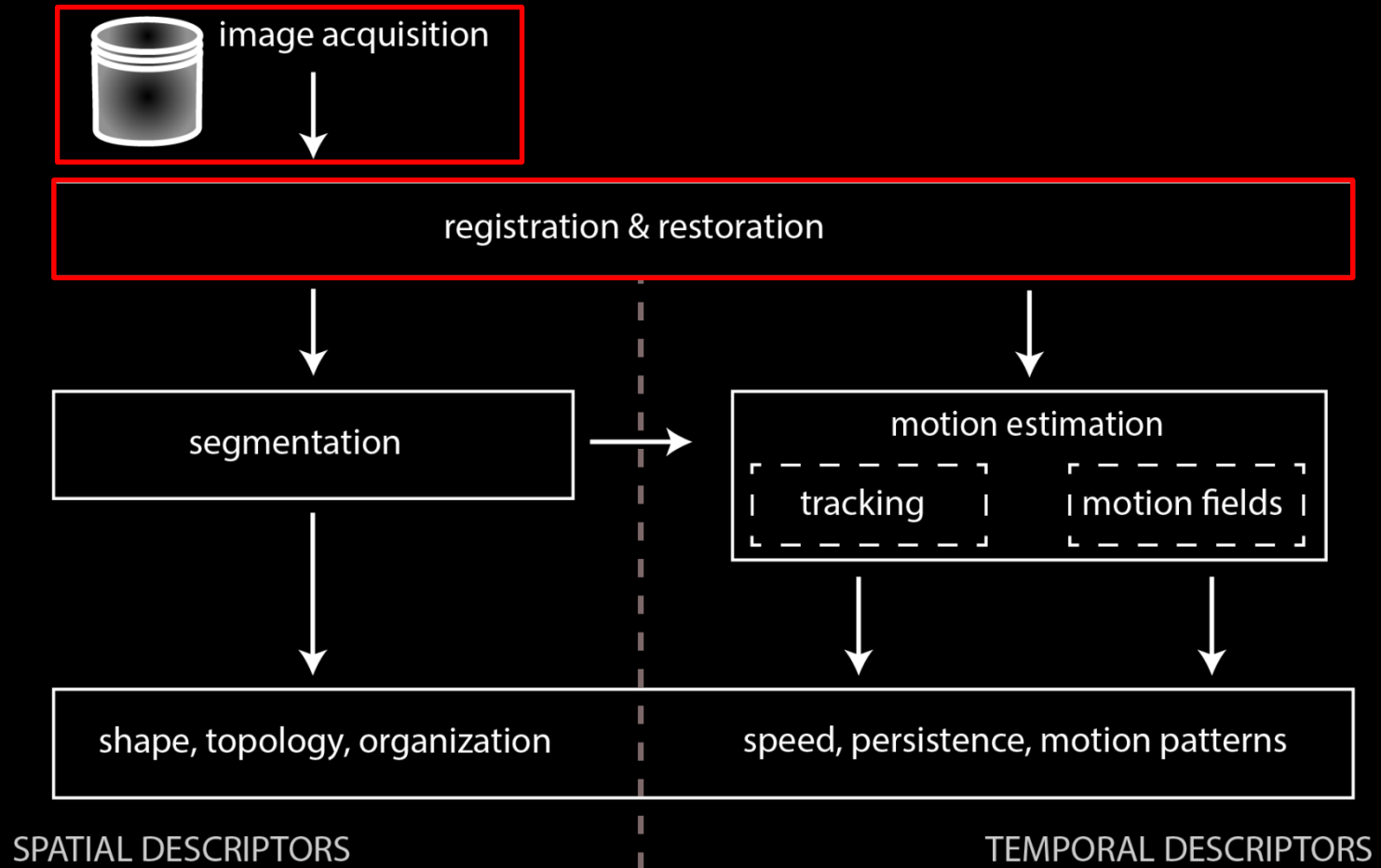
- 6 maquinas Debian RAID 6 (en RAID "60")
- 30 TB a investigadores (60 TB raw)
- Red 1 Gbps interna
- 5 servidores de cálculo (BNI)
- 1 UPS

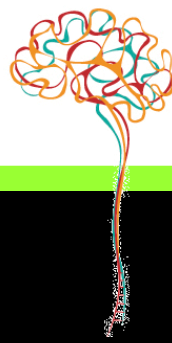
- equipos baratos
- facil uso desde windows
- uptime pobre (~90%)
- alto costo mantención (1 ing. + 1 postdoc)

¡ no escala a la próxima generación de microscopios!



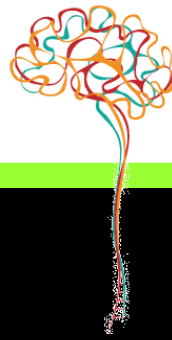
¿Qué hacemos típicamente con un set de datos?





¿Qué se requiere para escalar el almacenamiento y pre-procesamiento para microscopías como lightsheet?

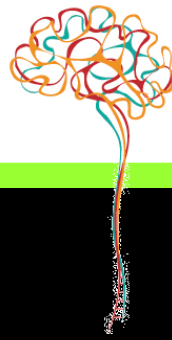
- Red rápida (> 1 Gbps)
- Potencia de cálculo (clusters, GPU, cloud)
- Crear/Adaptar algoritmos y Software
- *Almacenamiento*



Santiago-2015
10 Gbps Facultad Medicina – Reuna/NLHPC
Distancia 3 KM



BioMed-HPC
Red de Biología y Medicina
Computacional de Alto Rendimiento



En Chile existe potencia de cálculo disponible, pero distribuida...



NLHPC
National Laboratory
for High Performance
Computing
Chile

Leftraru
CLUSTER NLHPC

2640
cores

70
TFlops

56Gbps
Infiniband
FDR

Capacidad
de cómputo *

Conexión

Santiago-2015
2640 cores + Xeon Phi
Distancia 6 KM
10 Gbps



Viña del Mar-2014
500 cores + 6 GPU Tesla
Distancia 100 KM
1 Gbps



Temuco-2013
96 cores (ALTIX)
Distancia 600 KM
1 Gbps

A common pre-processing step in microscopy imaging is deconvolution

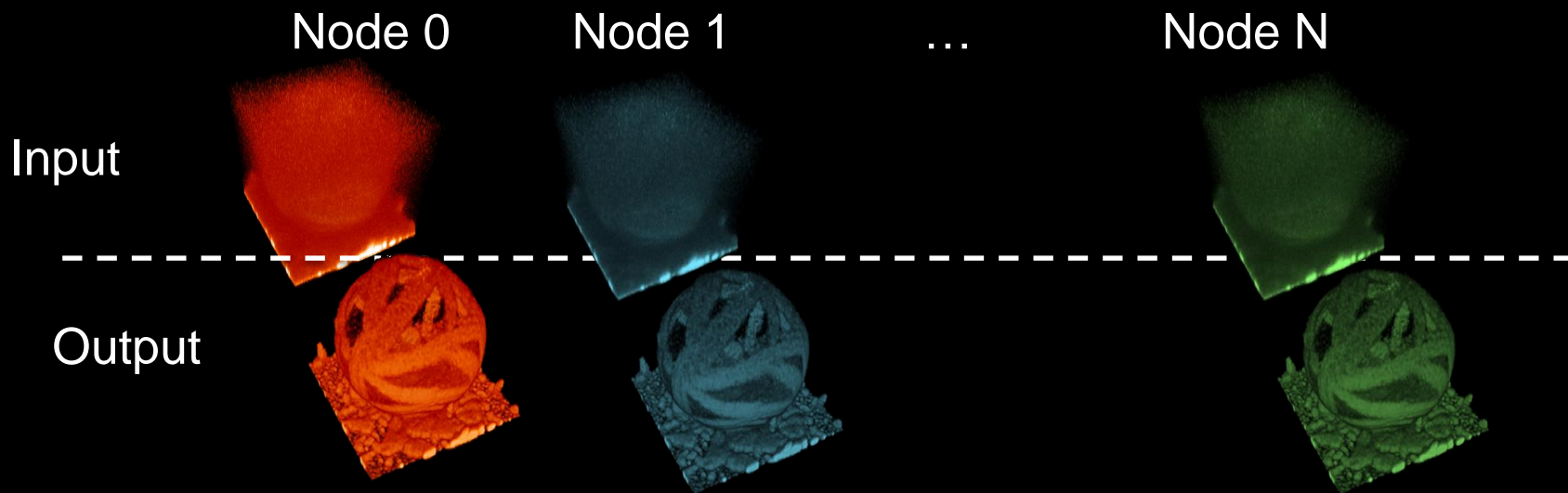


-Deconvolution algorithms estimate **B** from observed data **A**

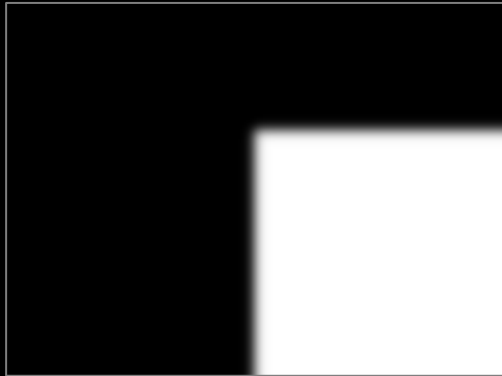
- + At the lab we use Maximum Likelihood Expectation estimation.
- + It takes 10-50 minutes per stack $I(x,y,z,t)$
- + Fac. Medicine has 5 dedicated deconvolution servers (BNI).

Restoration algorithm: HPC deconvolution

Deconvolution is highly parallelizable! 😊



Restoration algorithm: HPC deconvolution



Input stack

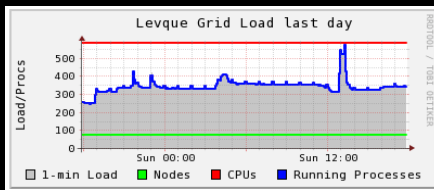


Estimated deconv. (Levque cluster)

In 2014 cluster (Levque):

- 1 512x512x64 stack (16 MB) -> 2 [m].
- 150 512x512x64 stacks (2.4 GB) -> 2* [m]

*



Exact deconv.

```
Laboratorio Nacional de Computacion de Alto Rendimiento (NLHPC)  
Centro de Modelamiento Matematico (CMM)  
Universidad de Chile
```

```
IMPORTANTE: NO EJECUTAR PROCESOS EN ESTE NODO POR T>10Min  
PARA ESTO DEBEN DE USARSE LAS COLAS DE EJECUCION
```

```
.....  
-bash: warning: setlocale: LC_CTYPE: cannot change locale (UTF-8): No such fi  
le or directory  
[mcerda@leftraru4 ~]$ █
```

Problema

- Interfaz SSH para biólogos (!)

Trabajo Actual

- Colaboración con empresa de software de deconvolución
- Adaptación de Software web para encolamiento de trabajos de deconvolución.

- Importantes avances técnicos en los últimos años en F-Med con el objetivo de manejar microscopía TB.
- La red es un elemento central que permite utilizar recursos aún dispersos.
- Utilizar recursos dispersos es especialmente crítico para una Facultad de Medicina.
- Trabajo actual incluye consolidar pipelines existentes (deconvolución, segmentación).

Incorporar la solución de almacenamiento de datos a la red (SASIBA 2016):

- 300 TB (hasta 1 PB) para todo el campus
- Servidores dedicados a estudios clínicos y housing
- Conexión 10 Gbps a: Campus, REUNA, Clusters, Universidad
- Proyecto financiado por fondos concursables del Gobierno - FONDEQUIP (~250.000 US\$)