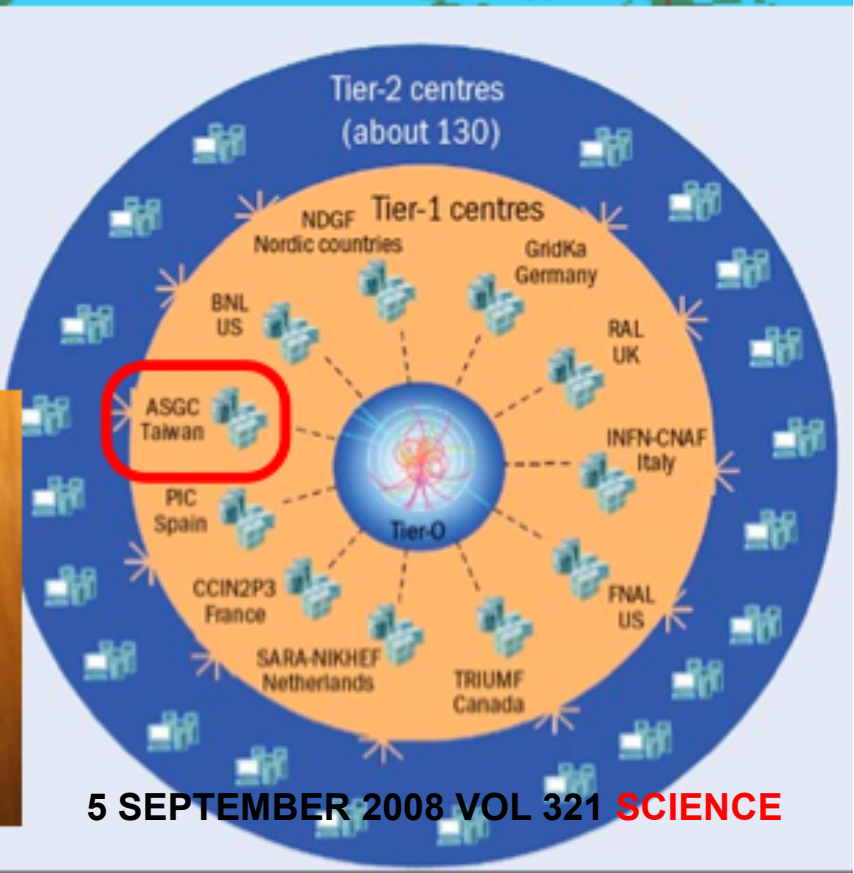
Platform support & benchmark for CryoEM in ASGC

Eric Yen, Felix Lee
ASGC
CryoEM workshop, ISGC2018

After WWW, CERN decided in 2000 to develop and deploy distributed computing to support a new generation of big data driven research. Middleware had been developed for geographically distributed cloud centers to share resources with applications to particle physics research over the past 10 years.

ASGC was established to join the international collaboration centered at CERN to build a platform for distributed computing to support e-Science.

WWW was Invented at CERN

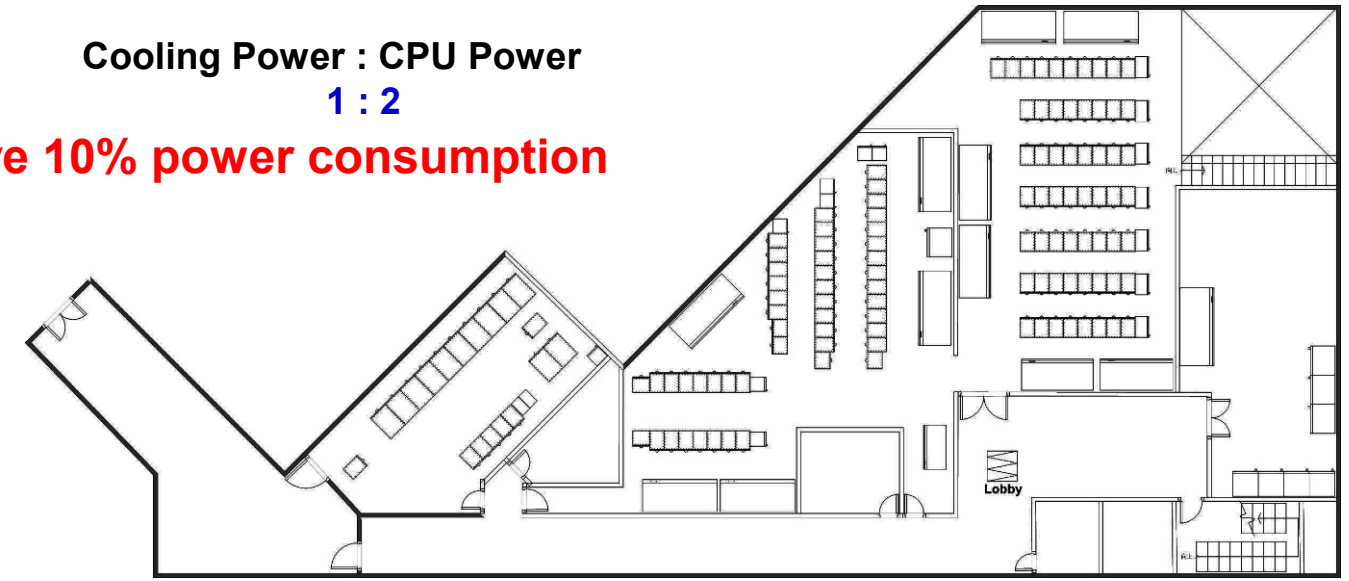


5 SEPTEMBER 2008 VOL 321 SCIENCE

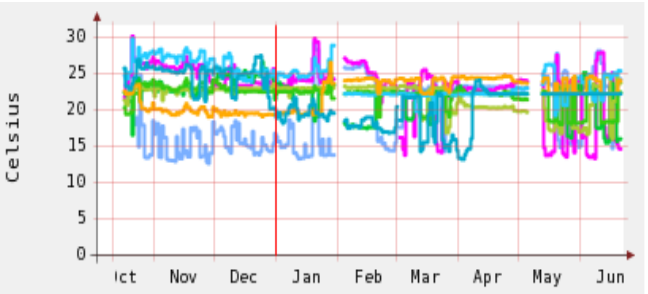
ASGC Computing Center

Cooling Power : CPU Power
1 : 2

No UPS to save 10% power consumption



- **Total Capacity**
- 2MW, 400 tons AHUs
- 99 racks
- ~ 800 m²
- **Resources**
- 21,000 CPU Cores
- 68,000 CUDA Cores
- 14 PB Disks
- **Rack Space Usage (Racks)**
- AS e-Science: 53.4 (53.9%)
- RCEC: 8.6 (8.6%)
- IPAS: 7.8 (7.9%)
- ASCC: 1.9 (2.0%)
- IES: 1.3 (1.3%)
- Free: 26 (26.3%)



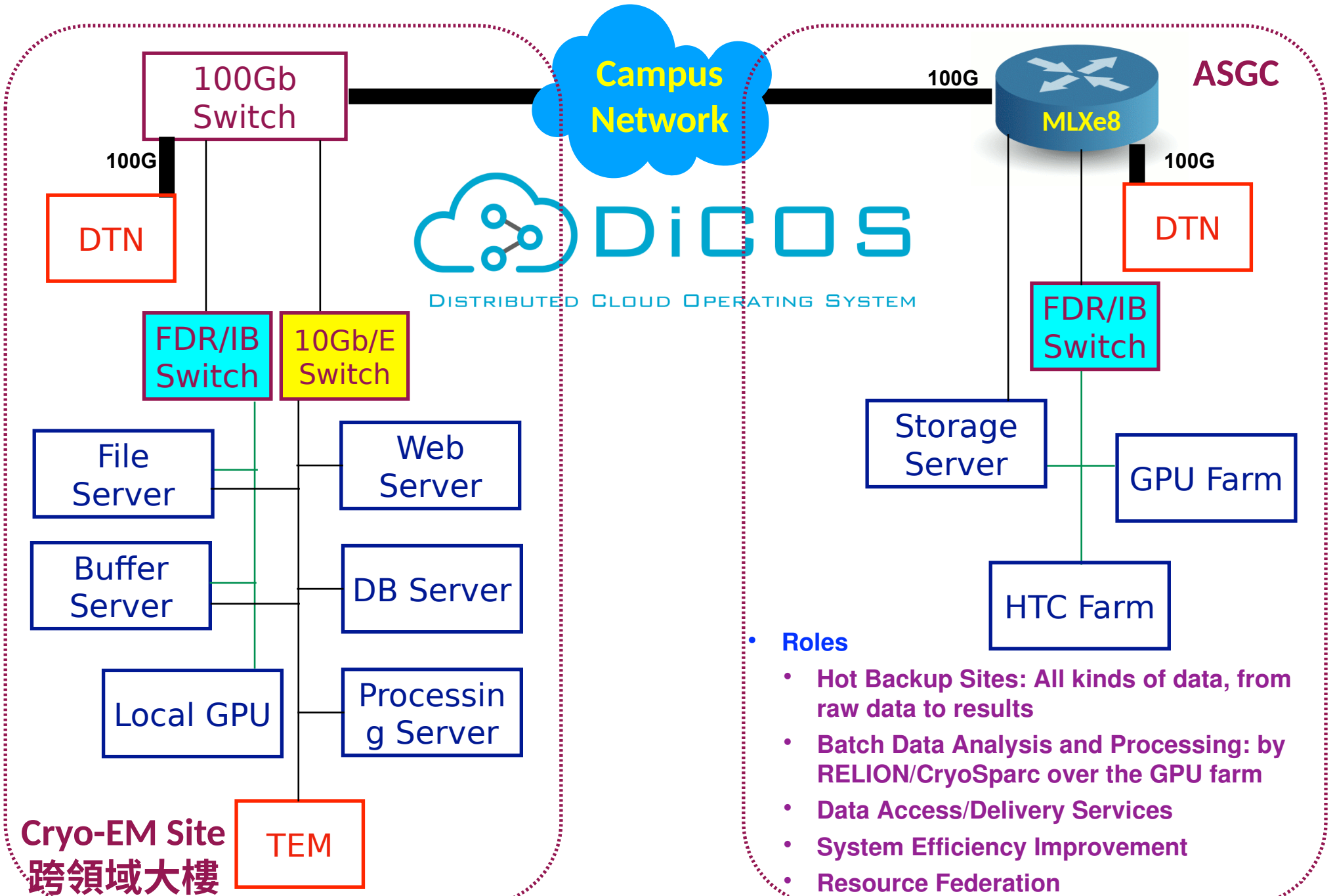
Monitoring the power consumption and temperature of every piece of equipment every 10 seconds.

All software used are opensource codes developed by ASGC and an international collaboration led by CERN

ASGC GPU farm

- Two models available
- Consumer GPU
  - GTX 1080Ti: 64 in total.
  - 8 * GPUs per compute box.
  - 128GB system memory.
  - 1TB SSD local hard-drive.
- Server GPU
  - Tesla p100: 16 in total
  - 4 * GPUs per compute box.
  - 128GB system memory
  - 1TB SSD local hard-drive
  - NV-link
- 8 * v100 model is on-going.

Supporting Cryo-EM Applications



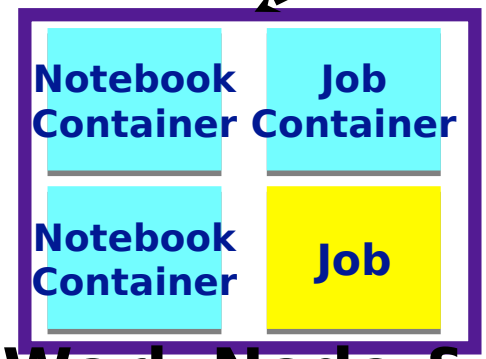
- **Roles**
  - **Hot Backup Sites: All kinds of data, from raw data to results**
  - **Batch Data Analysis and Processing: by RELION/CryoSparc over the GPU farm**
  - **Data Access/Delivery Services**
  - **System Efficiency Improvement**
  - **Resource Federation**

DTN=Data Transfer Node; IB = InfiniBand

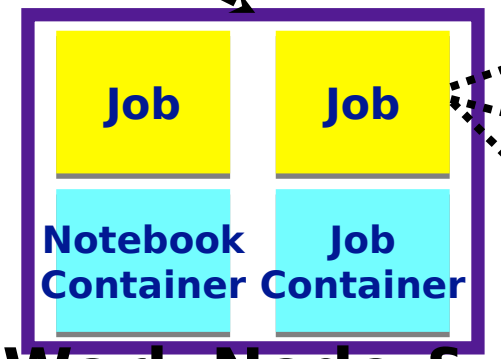
Supporting Cryo-EM Applications by DiCOS

DiCOS WebUI/Web Terminal

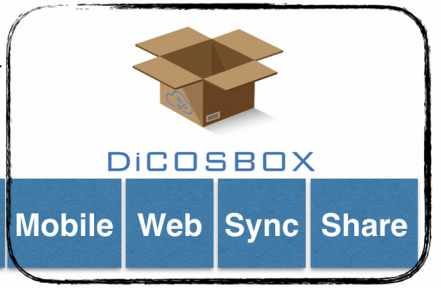
Scheduler Distributed Job Management



Work Node & Local Disk



Work Node & Local Disk



CVMFS (Software)

DiCOS
Distributed Cloud Operating System

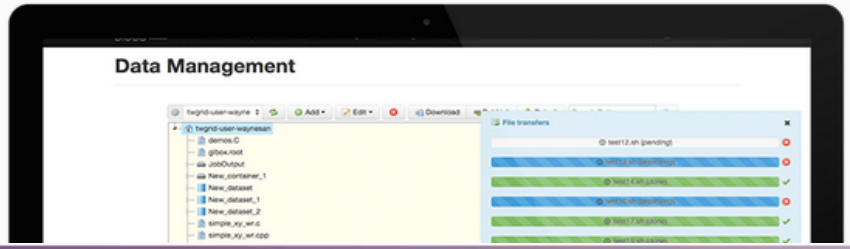
DiCOS web UI

https://dicos.grid.sinica.edu.tw/

DiCOS Beta | About | Job Submission | Job Monitoring | Data Management | Wiki | API | Contact | Terminal | Felix Lee

DiCOS

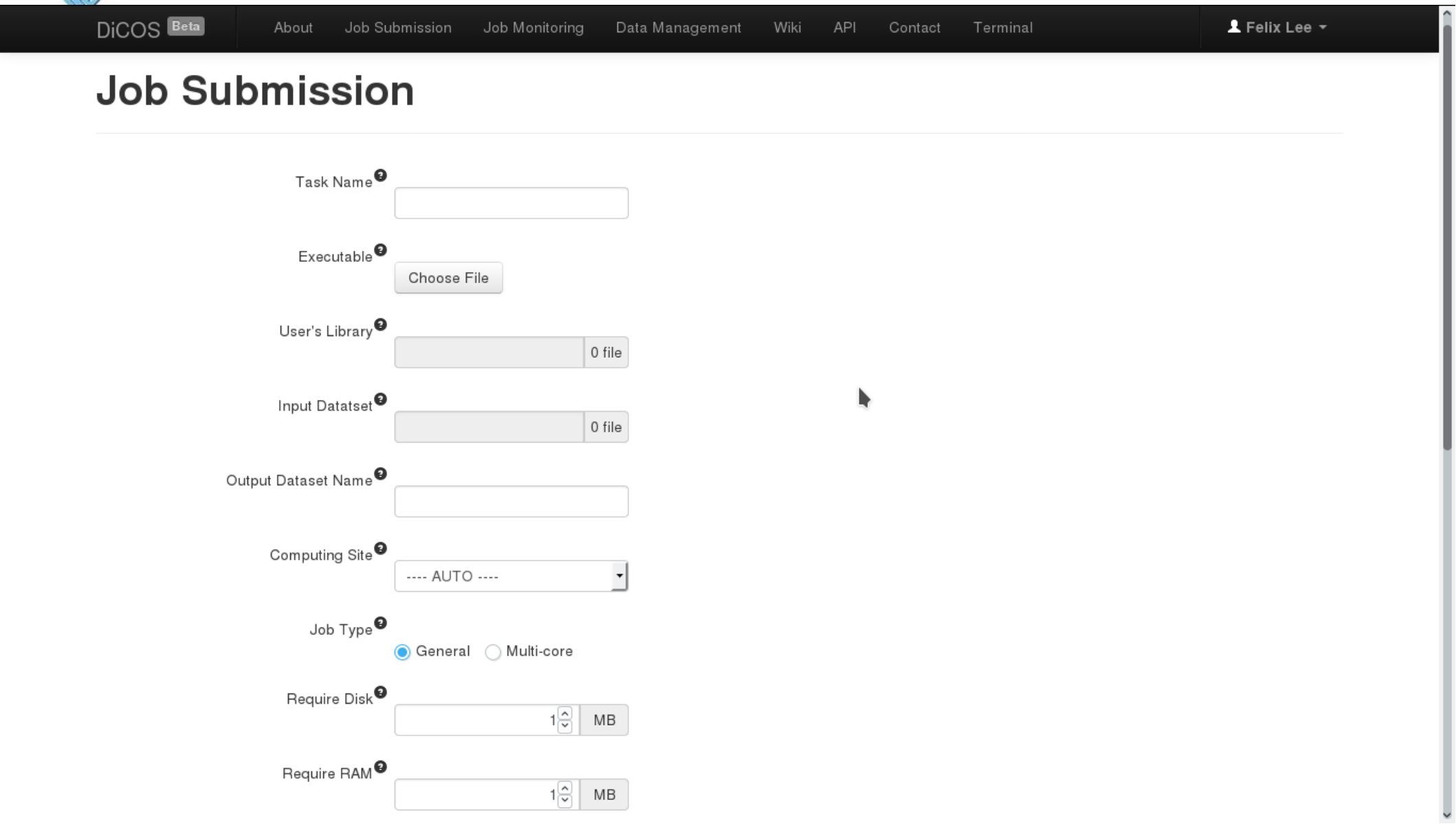
Distributed Cloud Operating System to support big data analysis and e-Science.



About DiCOS

DiCOS is the operating system for the distributed Cloud

DiCOS Job submission



DiCOS Job Management

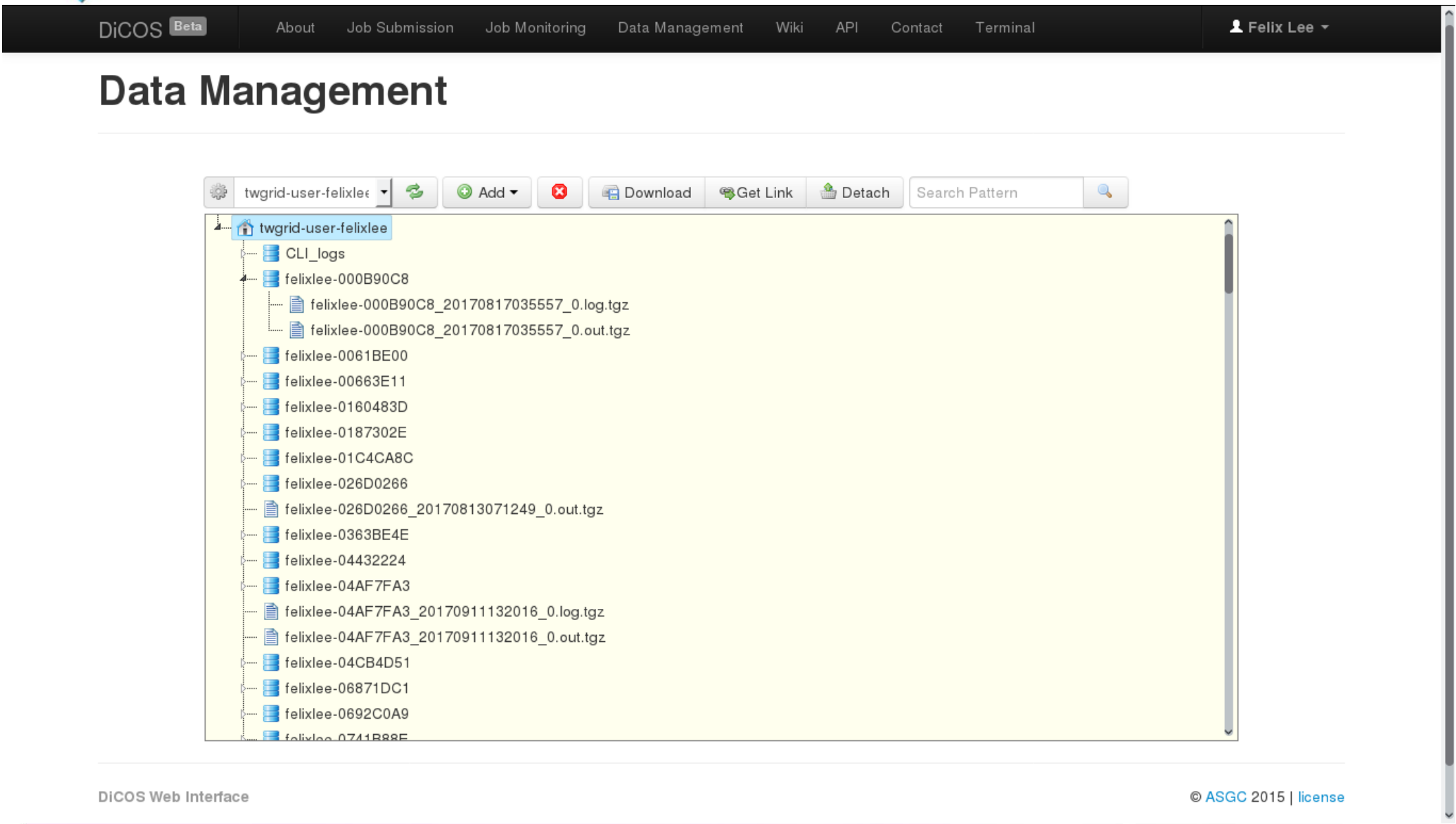
DiCOS Beta | About | Job Submission | Job Monitoring | Data Management | Wiki | API | Contact | Terminal | Felix Lee

Jobs Finished (Count: 22)

Display maximum 2000 rows

Select	PandaID	Job Name	Queue	Create Time	Status
☐	2290106	taiwan.873058066.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:44.000Z	finished
☐	2290114	taiwan.873058102.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:46.000Z	finished
☐	2290118	taiwan.873058105.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:46.000Z	finished
☐	2290120	taiwan.873058106.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:46.000Z	finished
☐	2290122	taiwan.873058113.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290124	taiwan.873058114.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290126	taiwan.873058115.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290128	taiwan.873058116.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290130	taiwan.873058117.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290132	taiwan.873058118.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290134	taiwan.873058119.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290136	taiwan.873058120.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290138	taiwan.873058121.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290140	taiwan.873058122.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290142	taiwan.873058123.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:47.000Z	finished
☐	2290146	taiwan.873058125.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:48.000Z	finished
☐	2290148	taiwan.873058126.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:48.000Z	finished
☐	2290150	taiwan.873058127.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:48.000Z	finished
☐	2290152	taiwan.873058128.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:39:48.000Z	finished
☐	2290762	taiwan.873058508.pr.pl1.eca...	ANALY_TAIWAN_CLOUD_PROD	2017-04-04T14:40:14.000Z	finished

DiCOS Data Management



DiCOS web terminal

DiCOS Beta | About | Job Submission | Job Monitoring | Data Management | Wiki | API | Contact | Terminal | Felix Lee

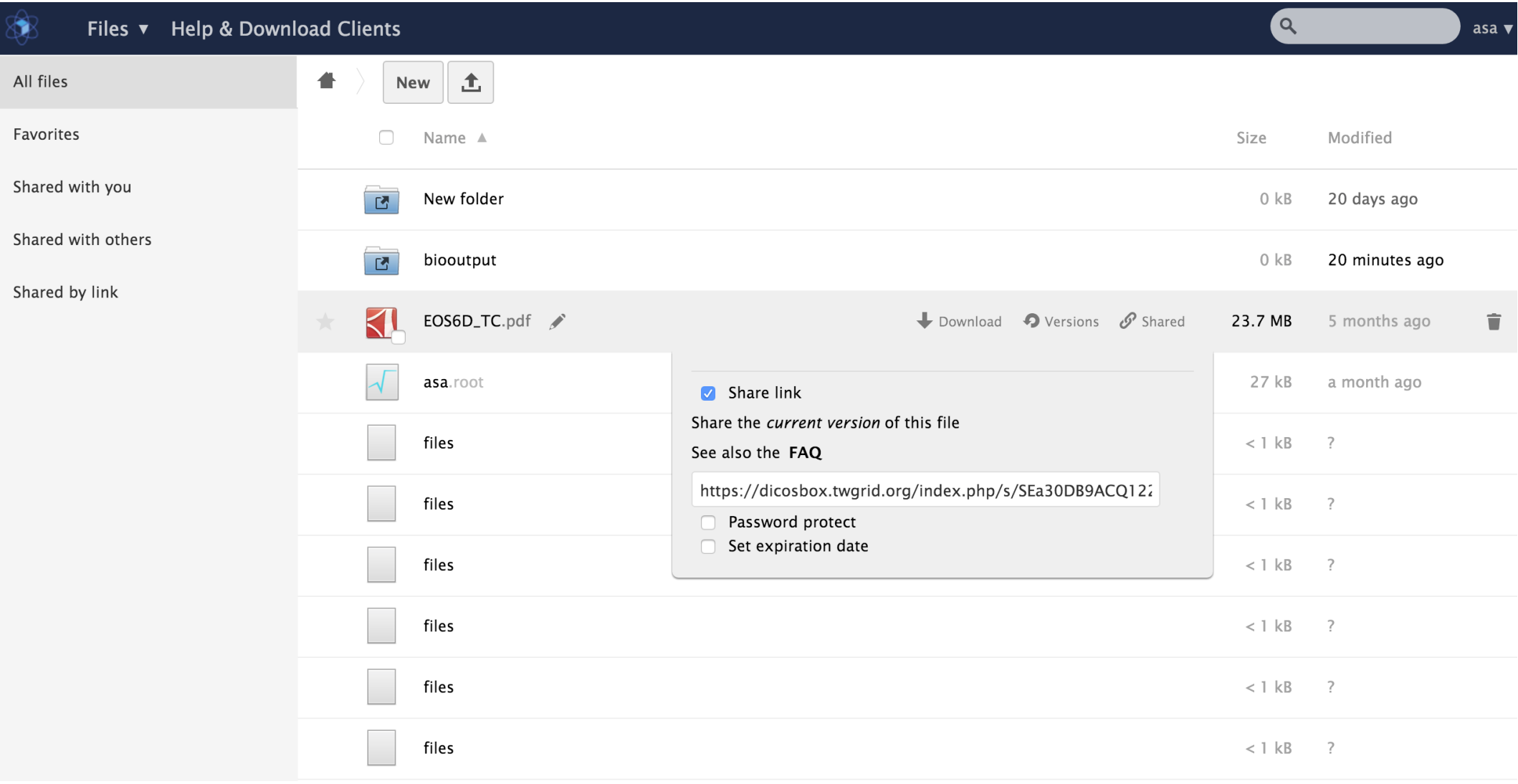
```
[felixlee@dicos-ui02 ~]$
[felixlee@dicos-ui02 ~]$ export X509_USER_PROXY=$HOME/.globus/felix.proxy
[felixlee@dicos-ui02 ~]$ voms-proxy-info -all
subject   : /C=TW/O=AS/OU=GRID/CN=Felix Dteam Lee 153280/CN=proxy
issuer    : /C=TW/O=AS/OU=GRID/CN=Felix Dteam Lee 153280
identity  : /C=TW/O=AS/OU=GRID/CN=Felix Dteam Lee 153280
type      : full legacy globus proxy
strength  : 2048
path      : /asgc_ui_home/felix/.globus/felix.proxy
timeleft  : 120:18:56
key usage : Digital Signature, Key Encipherment, Data Encipherment, Key Agreement
=== VO ams02.cern.ch extension information ===
VO        : ams02.cern.ch
subject   : /C=TW/O=AS/OU=GRID/CN=Felix Dteam Lee 153280
issuer    : /C=TW/O=AS/OU=GRID/CN=voms.grid.sinica.edu.tw
attribute : /ams02.cern.ch/Role=production/Capability=NULL
attribute : /ams02.cern.ch/Role=NULL/Capability=NULL
timeleft  : 120:18:56
uri       : voms.grid.sinica.edu.tw:15016

[felixlee@dicos-ui02 ~]$
[felixlee@dicos-ui02 ~]$ dicos job status
   PandaID  Name                        Status
-----------------------------------------------
   2301292  taiwan.873059581.he4.pl1  activated
   2301290  taiwan.873059580.he4.pl1  activated
   2301288  taiwan.873059579.he4.pl1  activated
   2301286  taiwan.873059578.he4.pl1  activated
   2301284  taiwan.873059577.he4.pl1  activated
   2301282  taiwan.873059576.he4.pl1  activated
   2301280  taiwan.873059575.he4.pl1  activated
   2301278  taiwan.873059574.he4.pl1  activated
   2301276  taiwan.873059573.he4.pl1  activated
   2301274  taiwan.873059572.he4.pl1  activated
   2301272  taiwan.873059571.he4.pl1  activated
   2301270  taiwan.873059570.he4.pl1  activated
   2301268  taiwan.873059569.he4.pl1  activated
   2301266  taiwan.873059568.he4.pl1  activated
   2301264  taiwan.873059567.he4.pl1  activated
[felixlee@dicos-ui02 ~]$
```

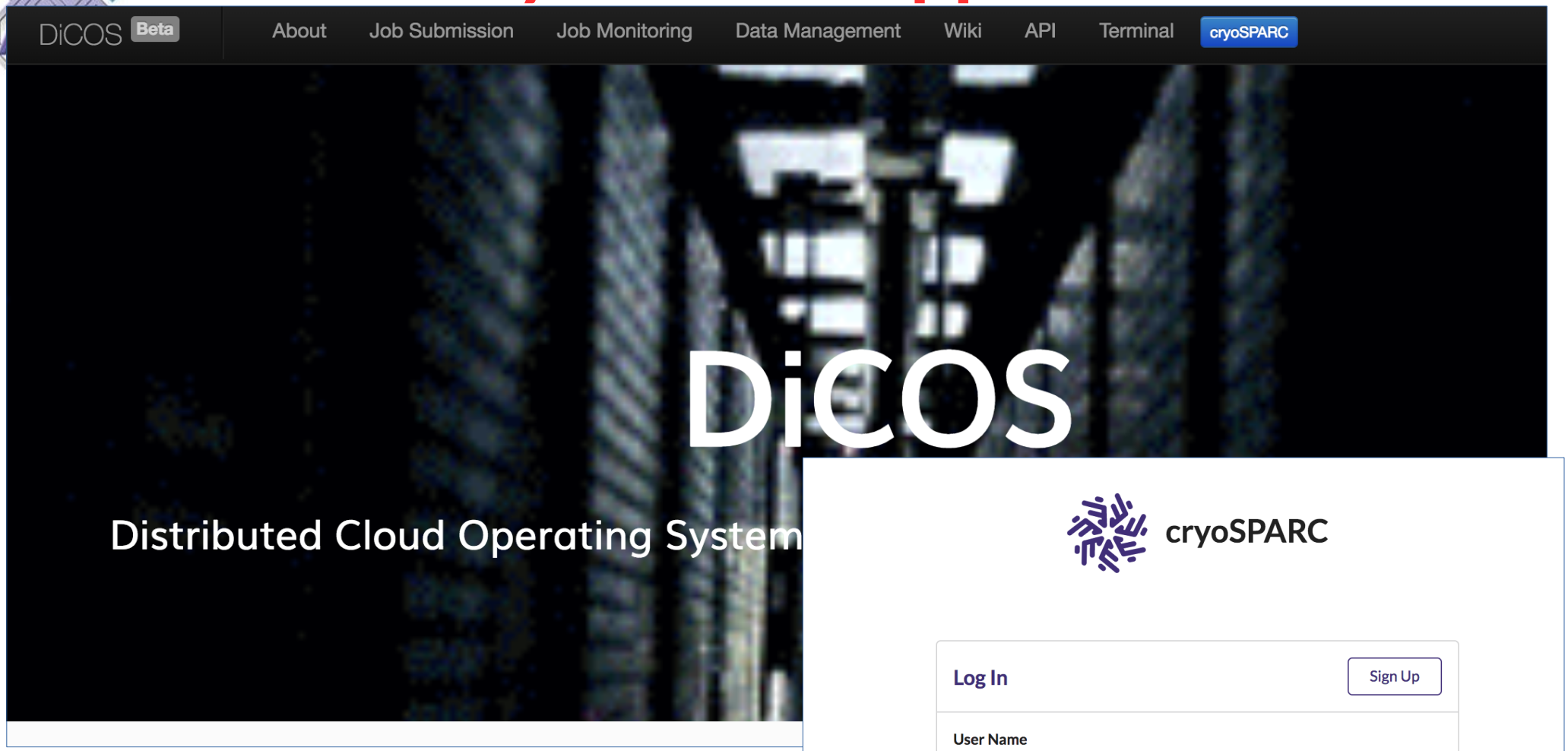
3:31:37 PM

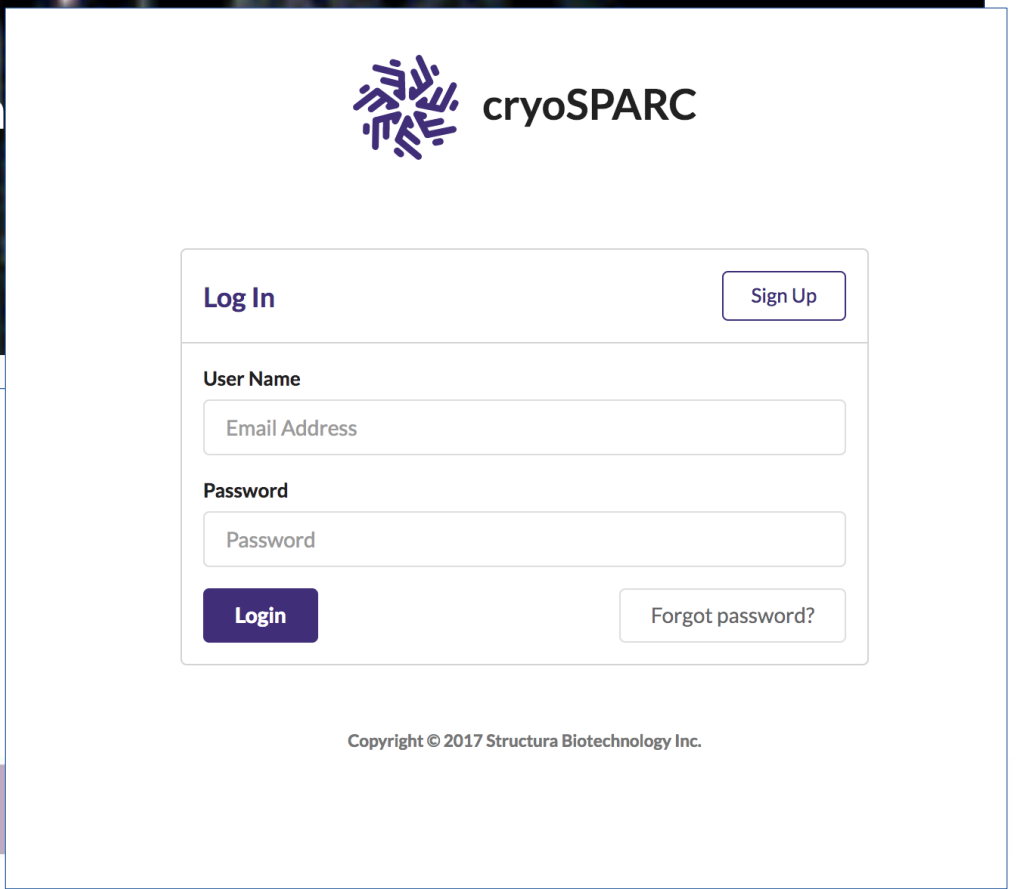
DiCOSBOX

Cloud storage, easy to use.
Default quota: 1TB per user.



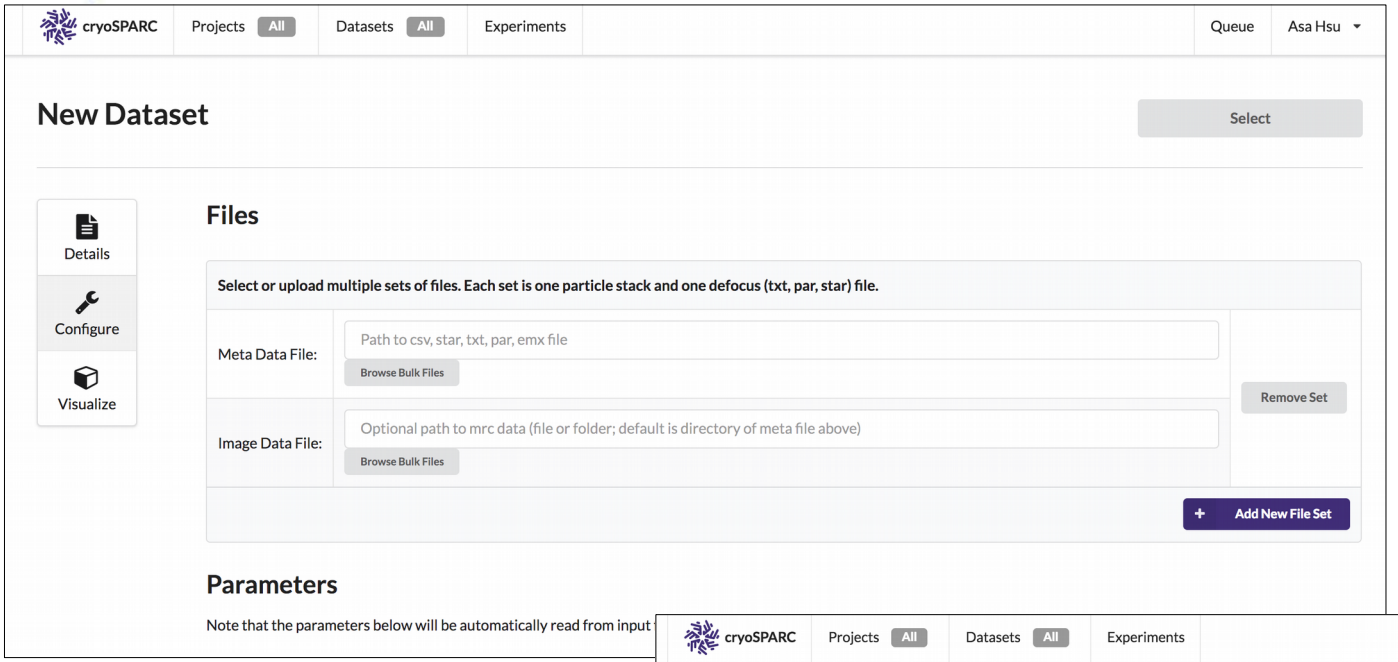
CryoSPARC Support

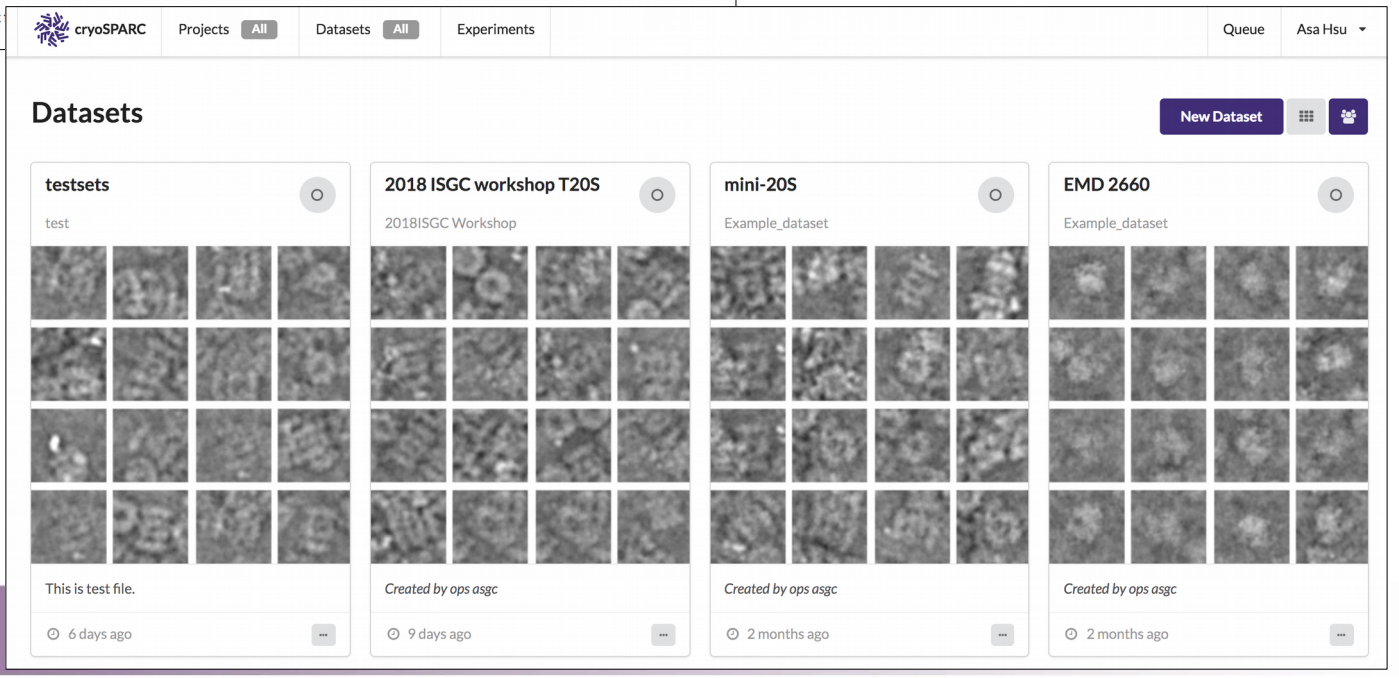




SSO login integration is on going

CryoSPARC Support





Benchmark for RELION

H/W SPEC	Benchmark	Time[hr]	nr MPIs	OS
CPU: E5-2650 v4 2160 cores	Class3D	23+	2160	CentOS7
GPU 4*p100	Class3D	2:08	9	CentOS7
GPU 8*GTX1080Ti	Class3D	2:02	9	CentOS7
GPU 4*GTX1080Ti (from relion website)	Class3D	2:26	5	SL7

Benchmark for CryoSparc

- Many thanks to Dr. Kuen-Phon Wu for doing benchmark with our facilities.

80S ribosome, 105247 particles	IBC 1080Ti	ASGC-1080Ti	ASGC-P100
2D classification(150 classes)	59.4	91.9	92.5
Ab initio Reconstruction	27.0	34.1	34.2
Homoegeneous Refinement	46.6	53.9	45.8
Resolution of homogeneous model	3.09 A	3.1 A	3.1 A
Heterogeneous Refinement	46.7	67.2	100.3
Heterogeneous population	51.2/28.7/14/5.9 (%)	54.3/24.4/14.8/6.3 (%)	52.7/29.3/11.4/6.4 (%)
Local Resolution	41.2	43.1	47.2

20S proteasome miniset, 9109 particles	IBC 1080Ti	ASGC-1080Ti	ASGC-P100
2D classification	9.4	12.2	12.3
Ab initio Reconstruction	11.5	17.6	15.6
Homoegeneous Refinement	124.3	66.0	56.9
Resolution of homogeneous model	3.38 A	3.43 A	3.5 A
Local Resolution	89.4	94.6	103.9

Thank you!

DiCOS how-to wiki: https://dicos.grid.sinica.edu.tw/wiki/

Contacts: felix@twgrid.org, Eric.Yen@twgrid.org