National
Astronomical
Research
Institute of
Thailand

Introduction to Chalawan HPDA system
&
Getting Started on Workshop Hands-On computing

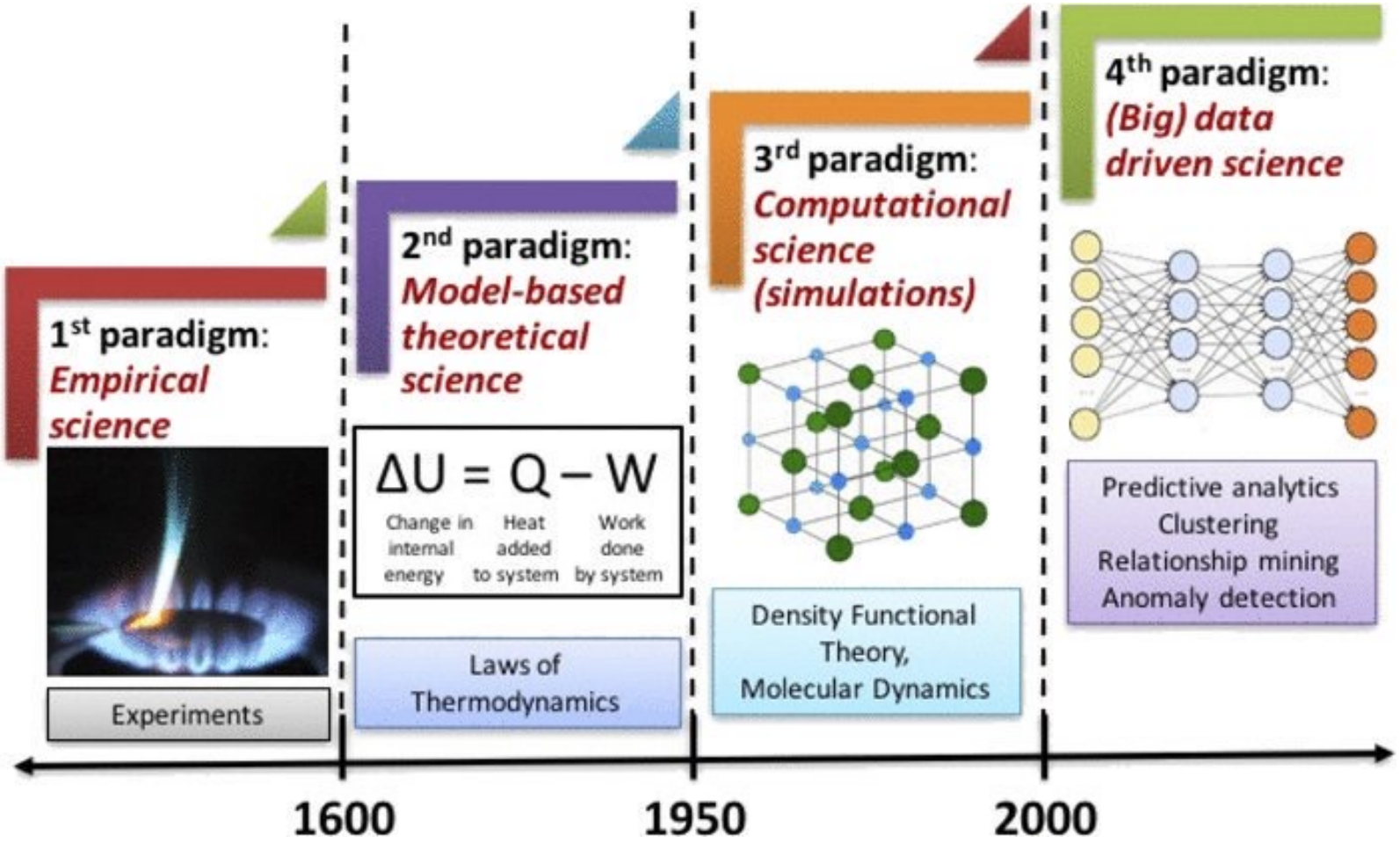
Utane Sawangwit

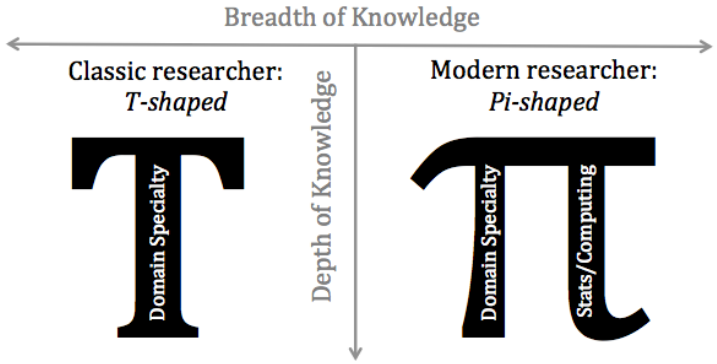
Researcher & Acting CIT Director

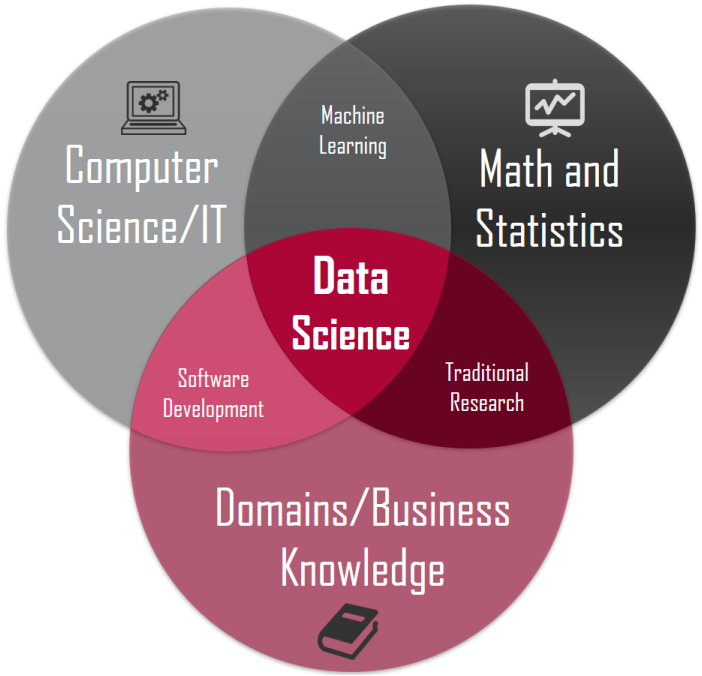
Thai National Observatory

CHALAWAN
High Performance Computing

Paradigm shift in scientific research

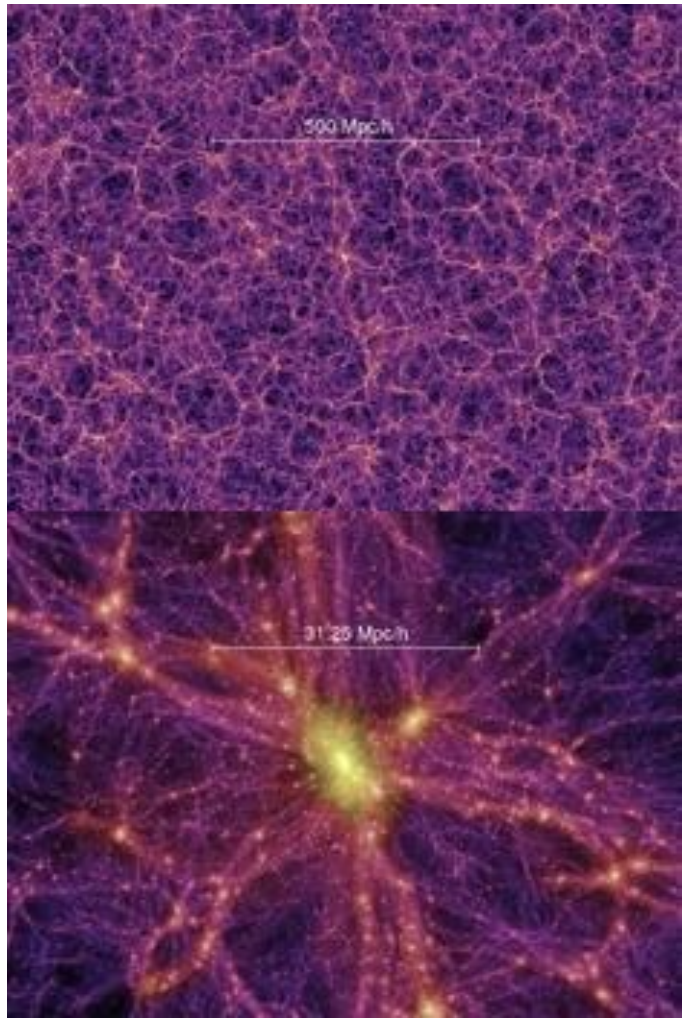




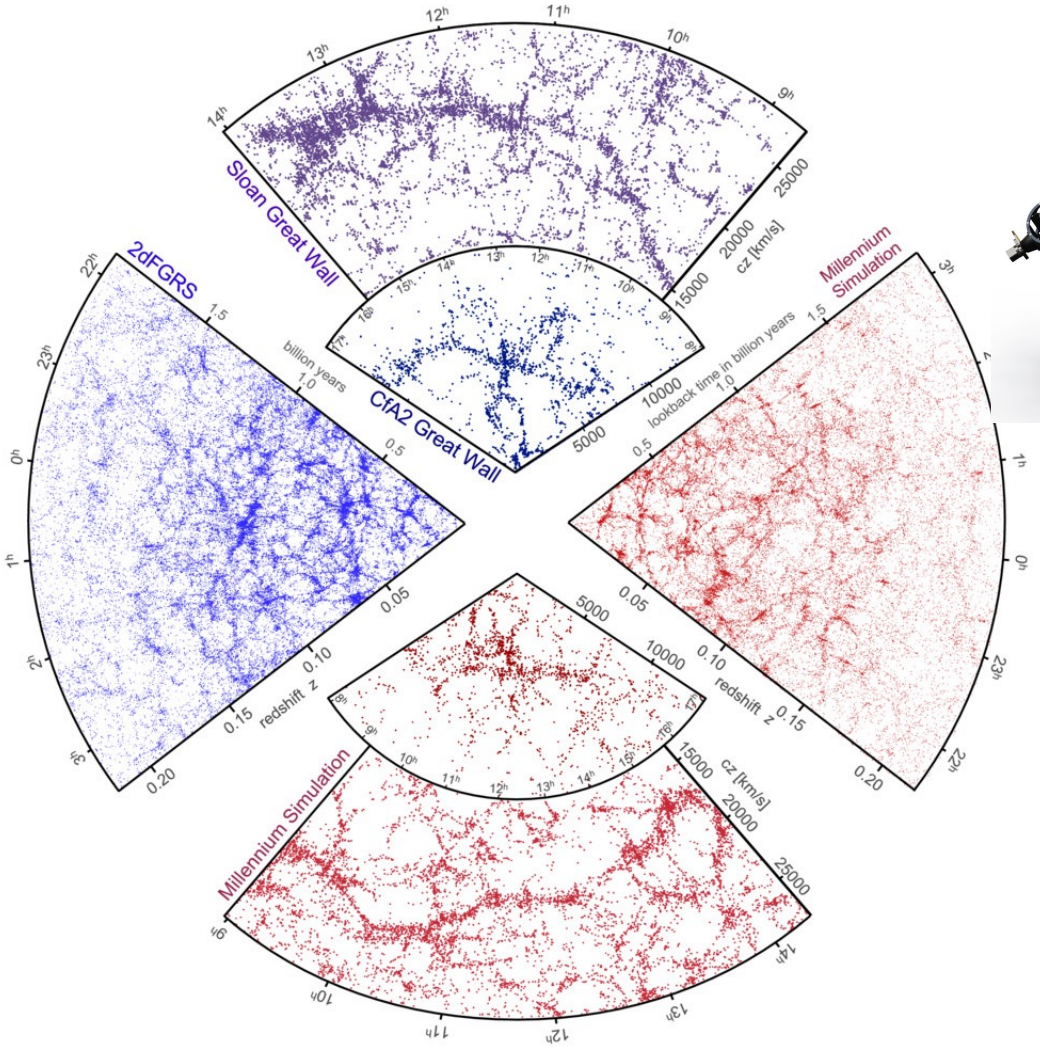


From Big Bang to Big Data

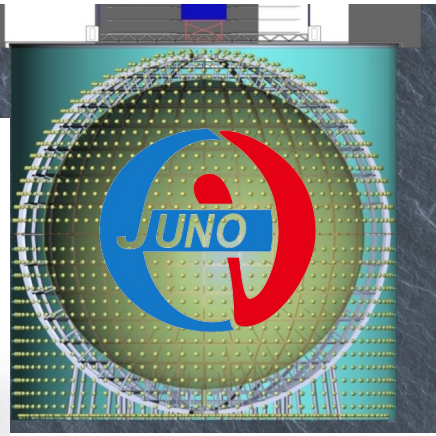
Millennium Simulation (2000)

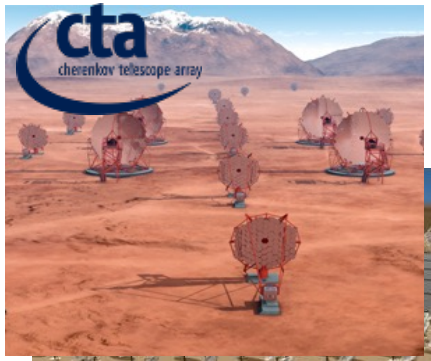


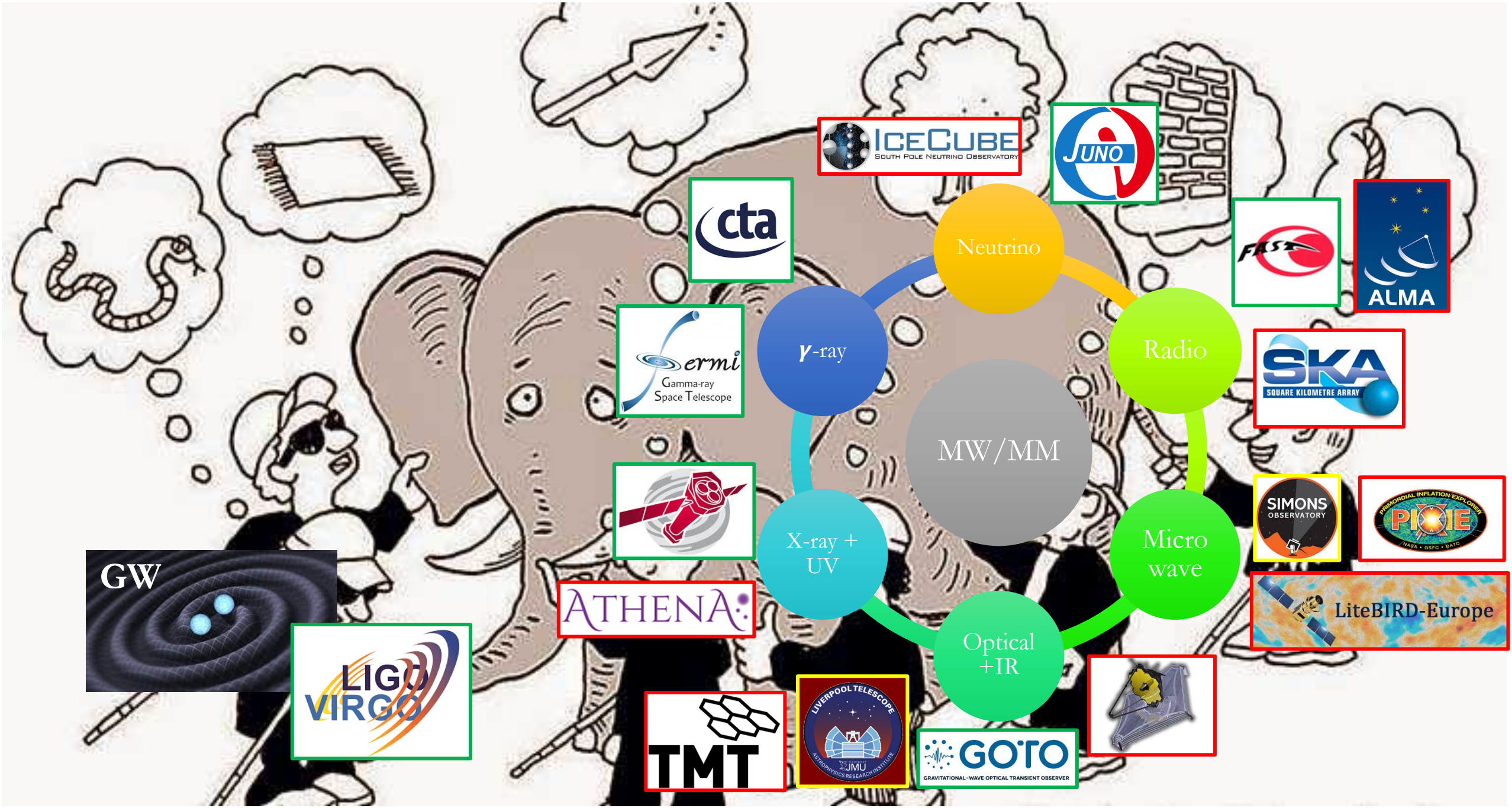
Real Universe versus Simulations (2000s onwards)



Age of Data Deluge (Now)

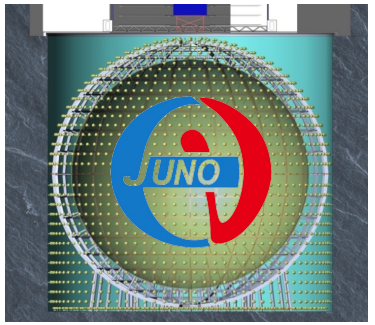




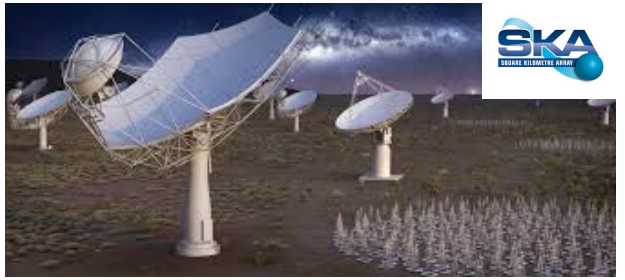

Neutrino
γ-ray
Radio
MW/MM
X-ray + UV
Micro wave
Optical +IR
IceCube
South Pole Neutrino Observatory
JUNO
cta
FAST
ALMA
Fermi
Gamma-ray Space Telescope
SKA
SQUARE KILOMETRE ARRAY
SIMONS OBSERVATORY
PIXIE
PRIMORDIAL INFLATION EXPLORER
NASA • GSFC • BATC
LiteBIRD-Europe
GW
ATHENA
LIGO
VIRGO
TMT
LIVERPOOL TELESCOPE
ASTROPHYSICS RESEARCH INSTITUTE
JMU
GOTO
GRAVITATIONAL-WAVE OPTICAL TRANSIENT OBSERVER

Expected data rate in the coming 1-3 years

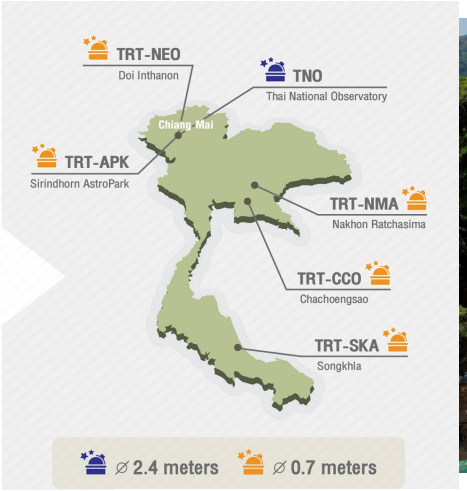
- Large Data volume, constantly require processing and analytics
  - **GOTO: 150-300TB/yr**
  - **POLARBEAR: 10TB/yr**
  - **TNRT: 200-700TB/yr**
  - **JUNO: 200TB/yr (reconstruction and analysis data only)**
  - **CTA: 1GB/s/event**
  - **Other indep. Experiments & surveys**



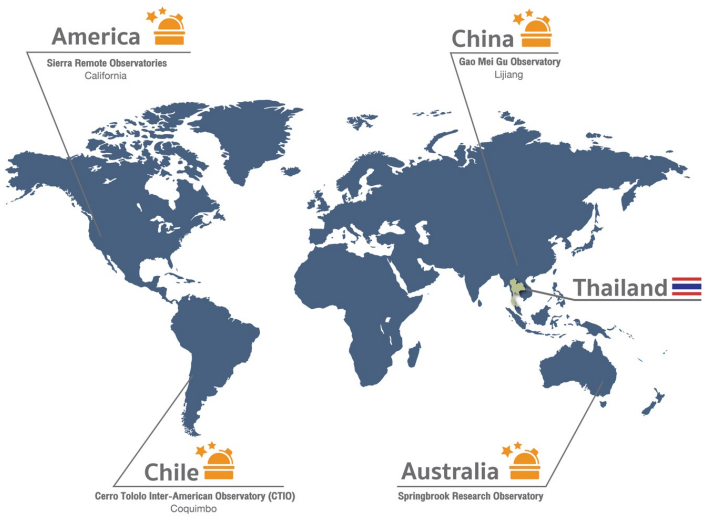
- **Future**: Opportunity for joining world-class experiments **e.g. as a regional data & processing centre node for SKA (160TB/s) (à la CERN-WLCG)**



Scientific & Engineering computing need @ NARIT



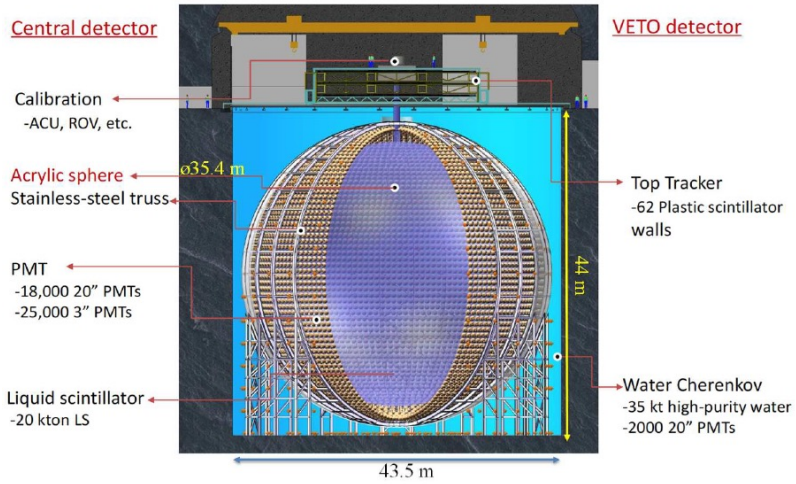
Thai Robotic Telescope Network

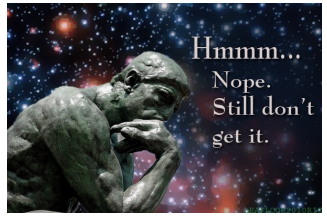


0.6–0.7 meter reflecting telescopes

GOTO
GRAVITATIONAL-WAVE OPTICAL TRANSIENT OBSERVER

JUNO





- Mission-critical analytics (experiments/surveys)
- Data Intensive analytics + ML & DL (observations/simulation)
- High-resolution data and model visualization
- Optimizations

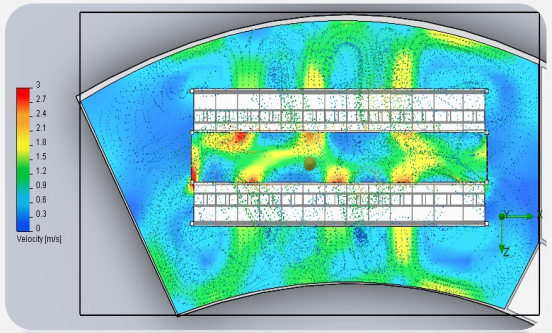
Why?: Our Objectives

HPC & Data Handling Advance Lab.

Complex Analysis & Simulations for Sci. Research

เพิ่มขีดความสามารถให้นักวิจัยไทย → นวัตกรรมใหม่ & งานวิจัยแนวหน้าของโลก

Green Data Centre



HPC Cluster

Time is Money

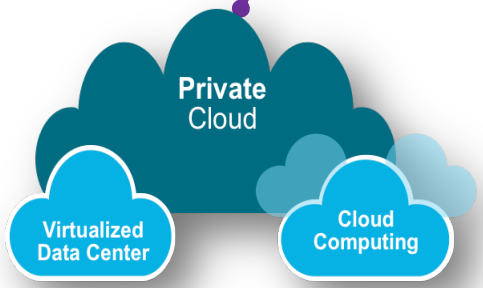
ให้บริการทั้งด้านการวิจัย วิชาการ และภาคธุรกิจ SME สำหรับระบบฮาร์ดแวร์และการให้คำปรึกษาจากบุคลากรที่มีความเชี่ยวชาญ

→ ลดเวลาในการนำความคิดไปปฏิบัติจริง ลดต้นทุน

Big Data Analytic & Machine Learning

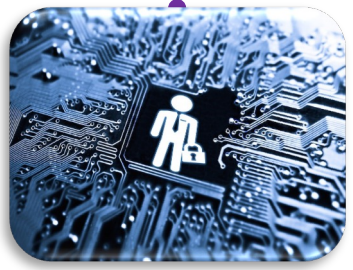
การตัดสินใจด้านนโยบายและธุรกิจที่ใช้ข้อมูลขนาดใหญ่และสถิติขั้นสูง

→ ผลลัพธ์คุ้มค่ากับการลงทุน แก้ปัญหาถูกจุด



การให้บริการ HPC on Private cloud

→ เข้าถึงบริการได้ง่าย รวดเร็ว ปลอดภัย ลดเวลาในการเตรียมการเพื่อ deployment



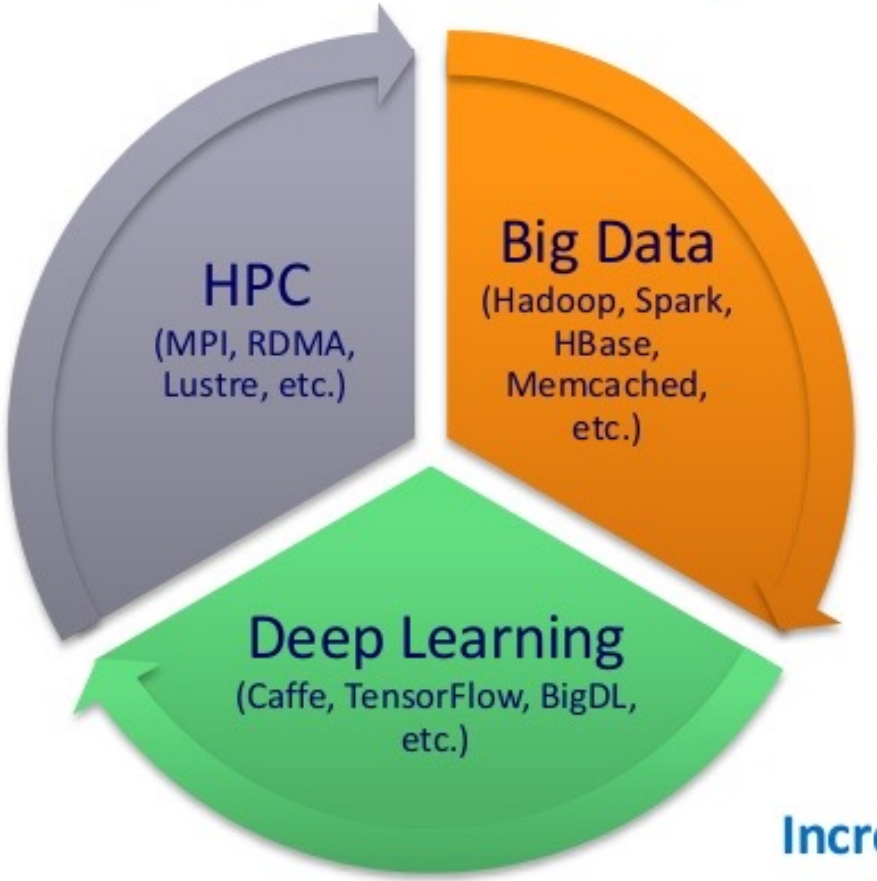
IT Oriented workforce

พัฒนาทรัพยากรบุคคลที่มีทักษะด้านเทคโนโลยีสารสนเทศและการวิเคราะห์ขั้นสูง → แรงงานคุณภาพสำหรับธุรกิจที่ขับเคลื่อนด้วย innovation & Data driven

High-Performance Data Analytics (HPDA)

HPC + Big Data Analytics + AI → **HPDA**

Increasing Usage of HPC, Big Data and Deep Learning



Convergence of HPC, Big Data, and Deep Learning!

Increasing Need to Run these applications on the Cloud!!

Chalawan HPC Cluster

CASTOR (2016)

- CPU Intel Xeon processor, ~ 560 cores
- RAM (TOTAL) ~ 2 TB

POLLUX (2519)

- CPU Intel Xeon processor, ~ 100 cores
- GPU Nvidia Tesla V100, 12 cards
- RAM (TOTAL) ~ 1 TB

SPICA (2522)

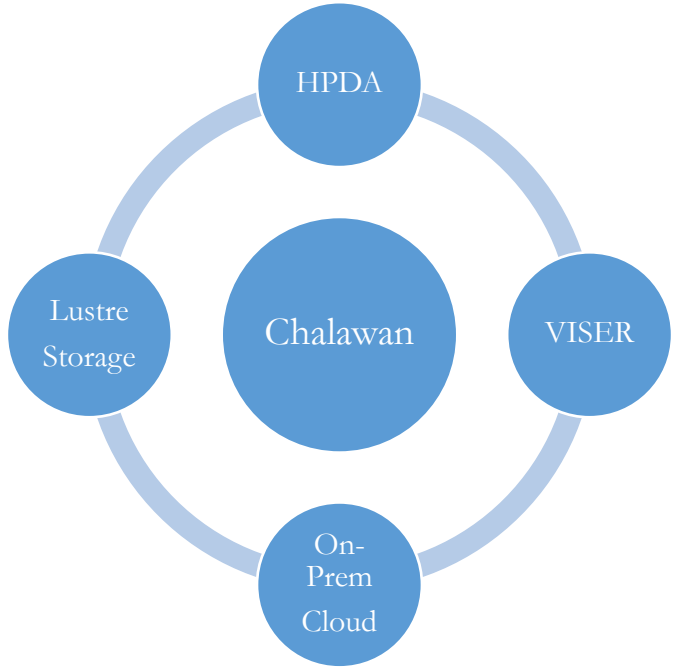
- CPU AMD, ~ 1,024 cores (+A100 GPU)
- RAM (TOTAL) ~ 8 TB
- STORAGE ~ 2PB + (2PB upgrade in 2023)

NARIT · Thai National Observatory · กระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม · United Nations Educational, Scientific and Cultural Organization · ITCA · National e-Science Infrastructure Consortium

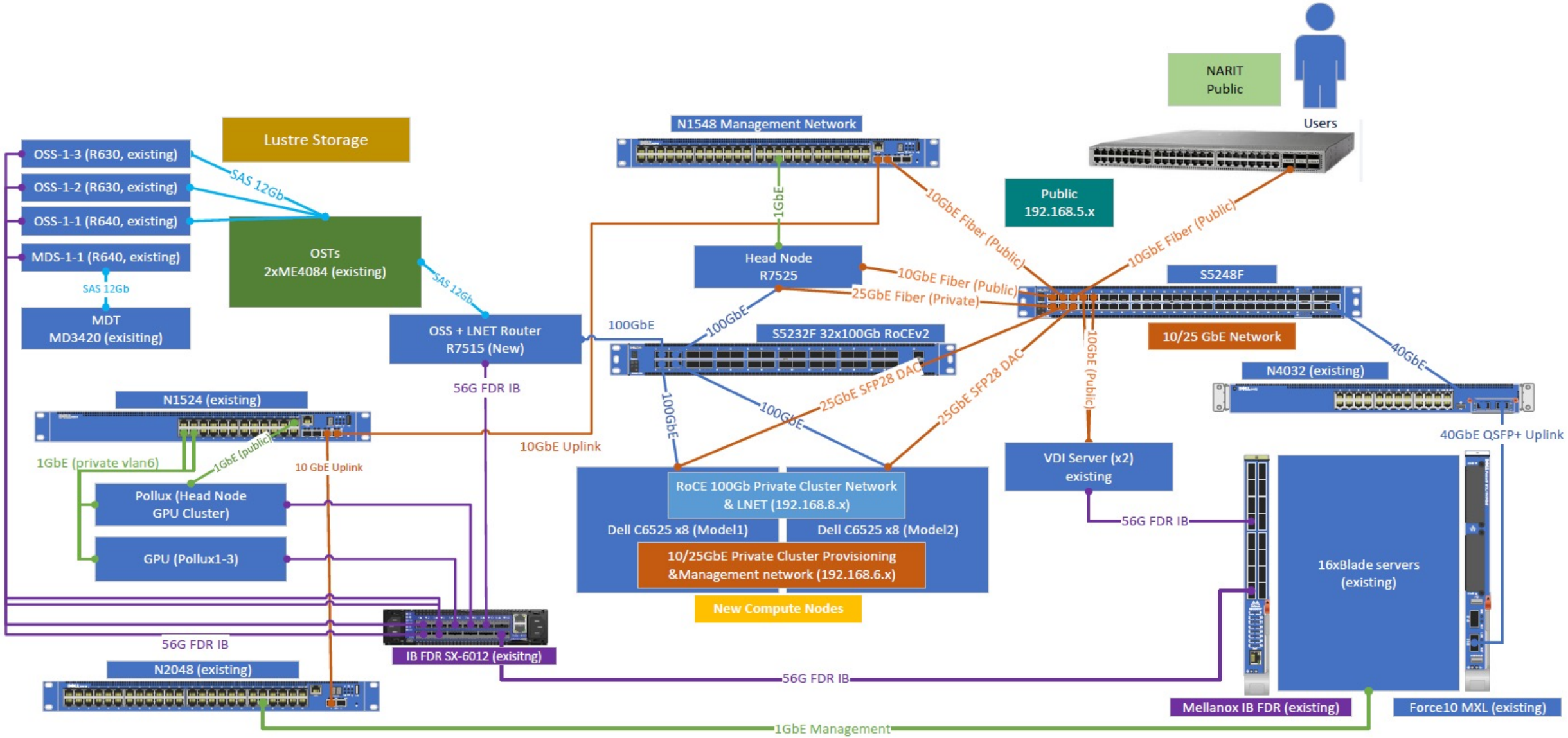
http://chalawan.narit.or.th

Chalawan Cluster

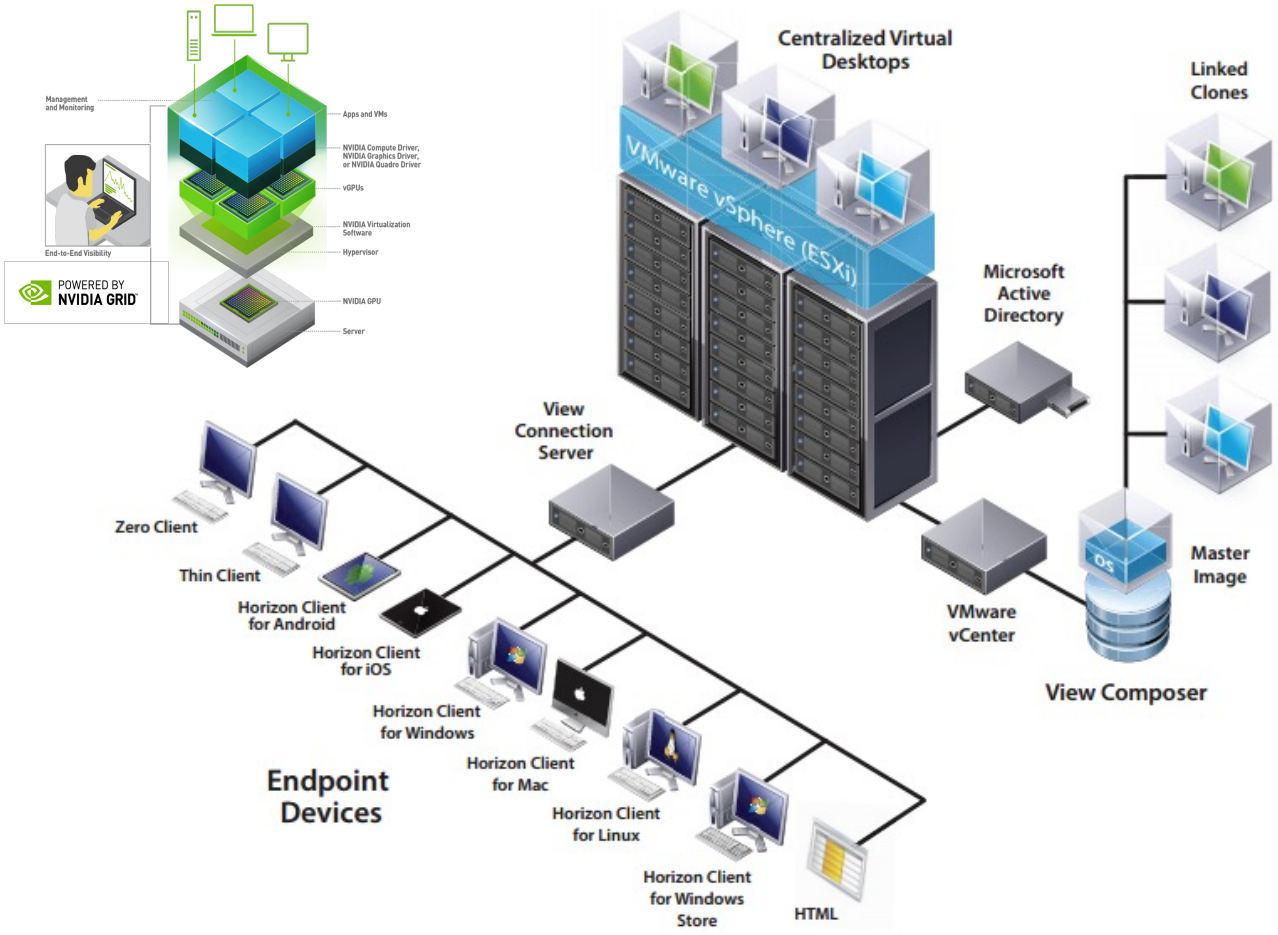
- **HPDA Computes** (Gen1: Castor, Gen2: Pollux, Gen3: Spica)
  - 1600 CPU cores (2.5-2.6GHz, 28c,64c per node)
  - 12TB memory (128GB,512GB per node)
  - GPU: Nvidia 12xV100 (32GB), 1xA100 (80GB)
  - Local Storage: NVMe Flash
  - Interconnect: 56Gbps IB, 100Gbps RoCE
- **Lustre Storage**: 4.7 PB
- **Virtual Infrastructure for Scientific and Engineering Research computing (VISER)**
  - VDI: Horizon View 20CCU (virtual workstation: vWS)
  - vGPU: 2 x A40, 1 x A10
  - Storage: Local SSD and Lustre Storage
- **On-prem cloud** (in development)
  - OpenStack on top of Proxmox cluster
  - Ceph Storage: Block & CephFS

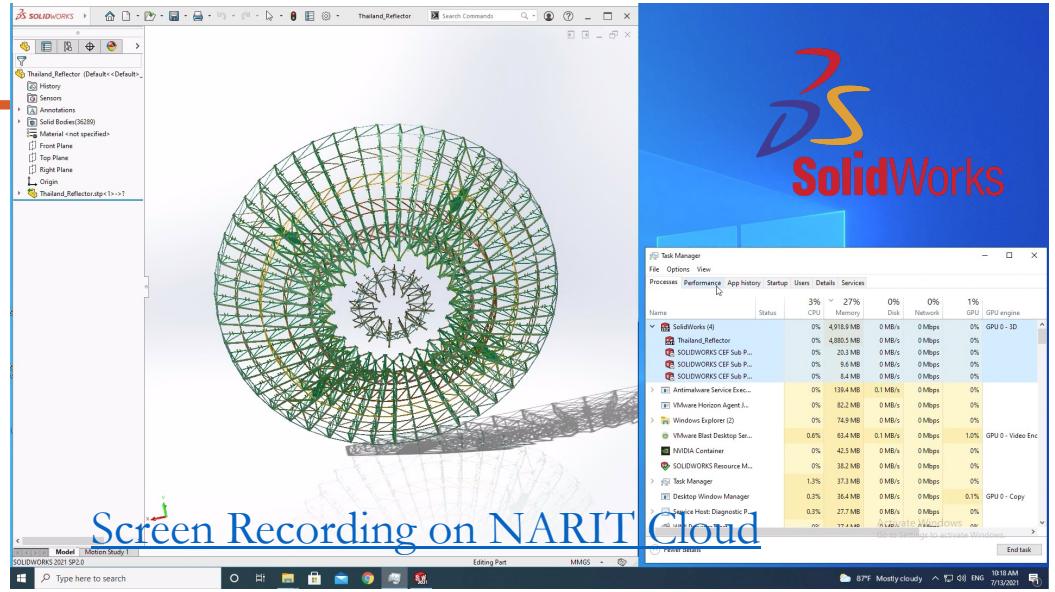


Chalawan Cluster (contd.)



VISER

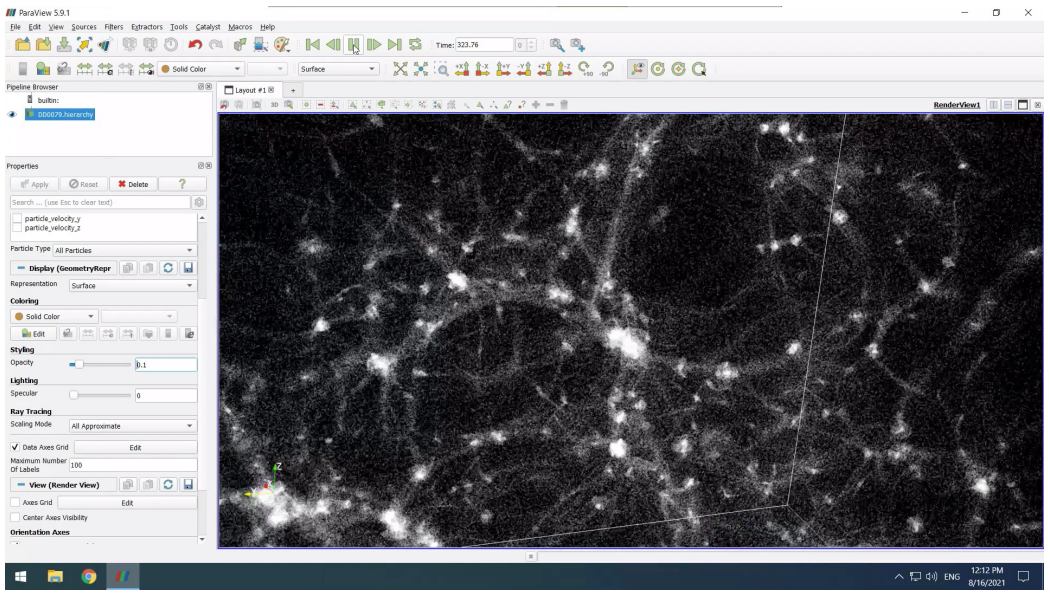




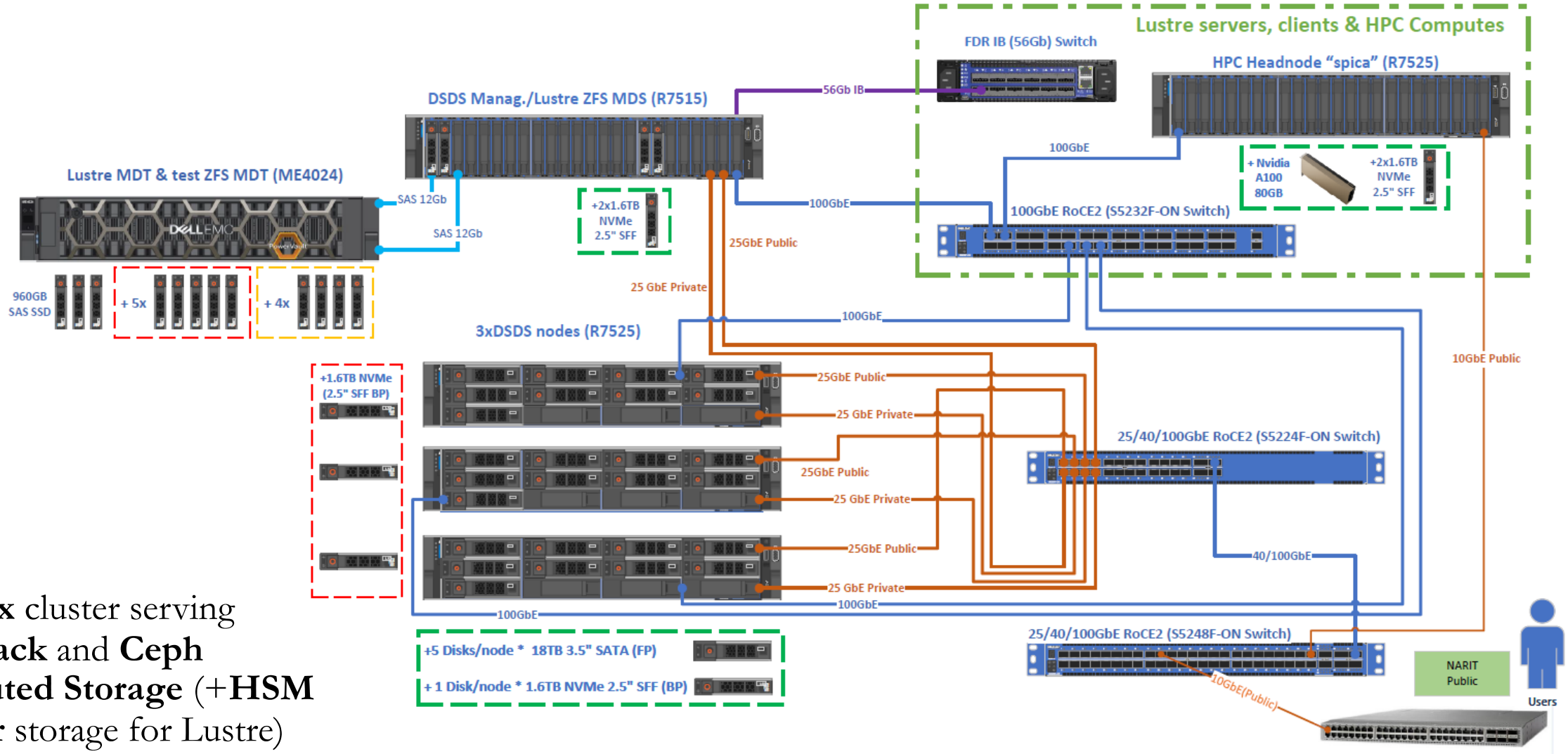
Screen Recording on NARIT Cloud

ParaView
Parallel Visualization Application

Screen Recording on NARIT Cloud



On-Prem Cloud (in development)



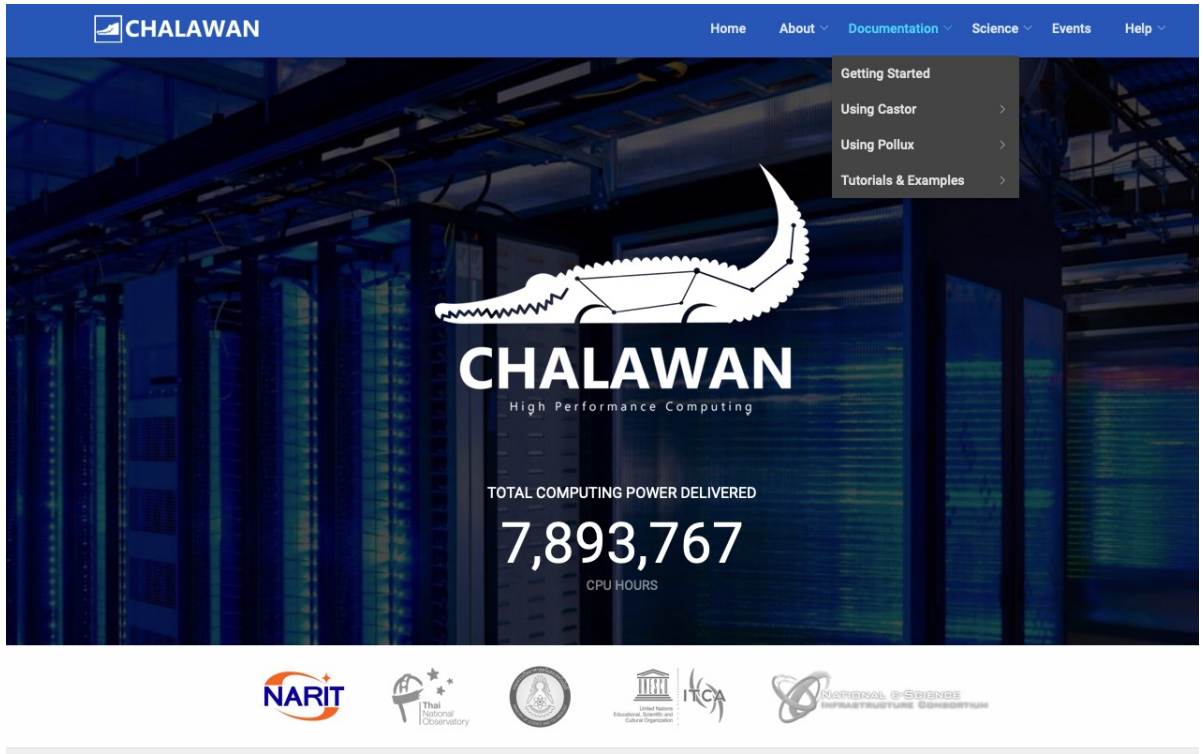
- **Proxmox** cluster serving **OpenStack** and **Ceph Distributed Storage** (+**HSM cold-tier** storage for Lustre)

Monitoring system (Grafana + Influxdb + Telegraf)

Emphasize on easy access for users

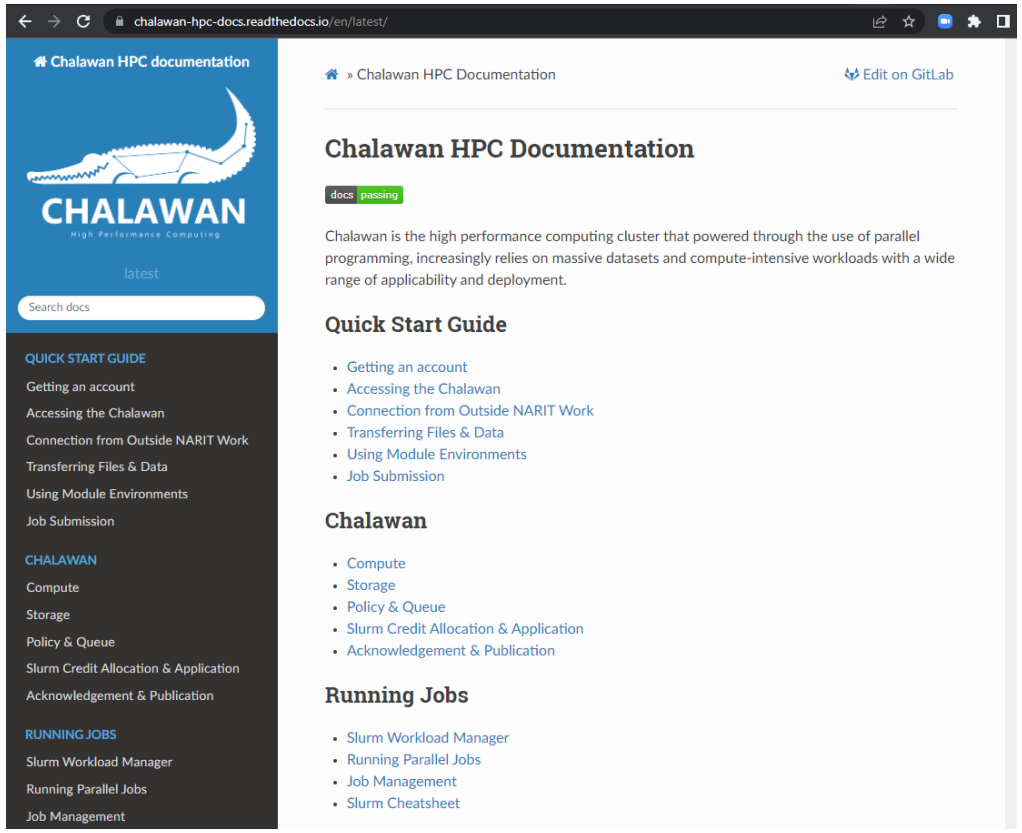


- Online Tutorials and applying for an account @ http://chalawan.narit.or.th
- Available to Thai users outside NARIT via **National e-Science Consortium**



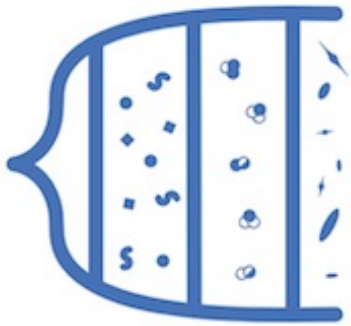
Chalawan Homepage and online proposal

https://chalawan-hpc-docs.readthedocs.io/en/latest/



Online documentation and user guide

Research and Development Highlights



Cosmology

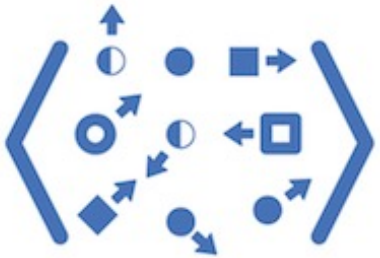
High Energy Astrophysics

Active Galactic Nuclei & Black Hole

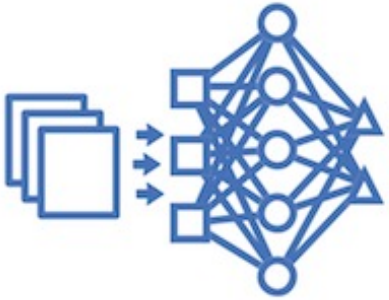
Exoplanet

Atmospherics Science

Condensed Matter Physics

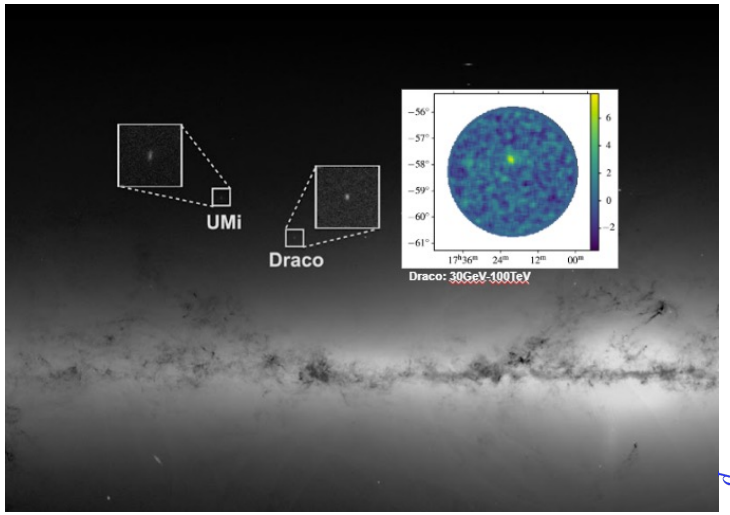


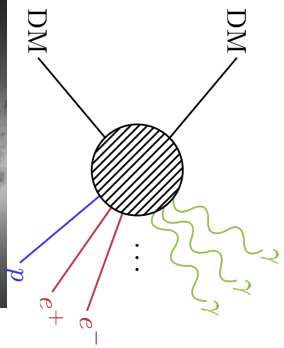
Statistical Mechanics



Big Data & Machine Learning

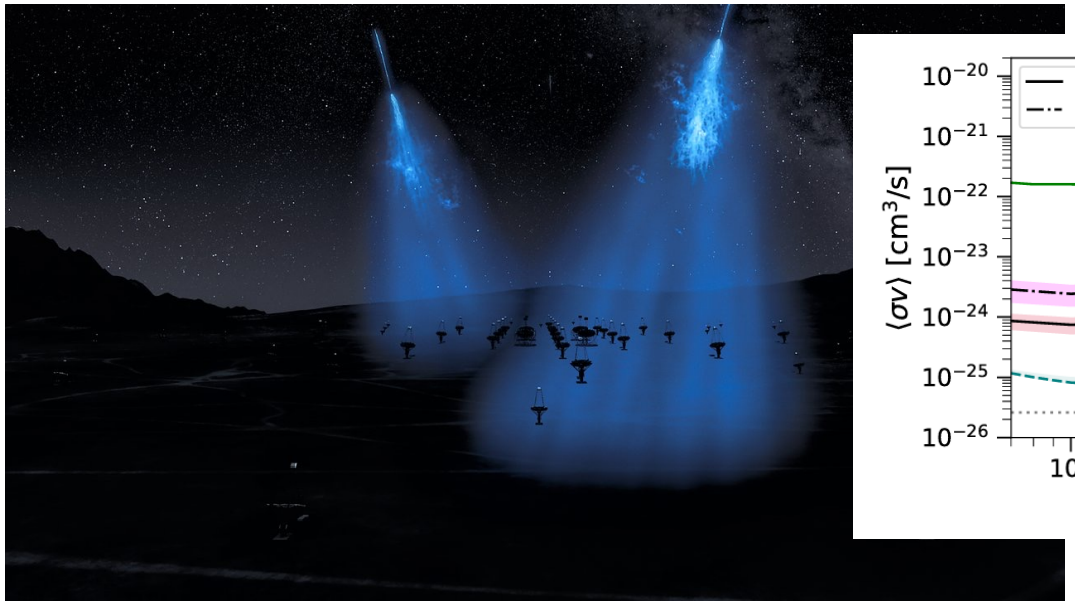
Dark Matter with CTA

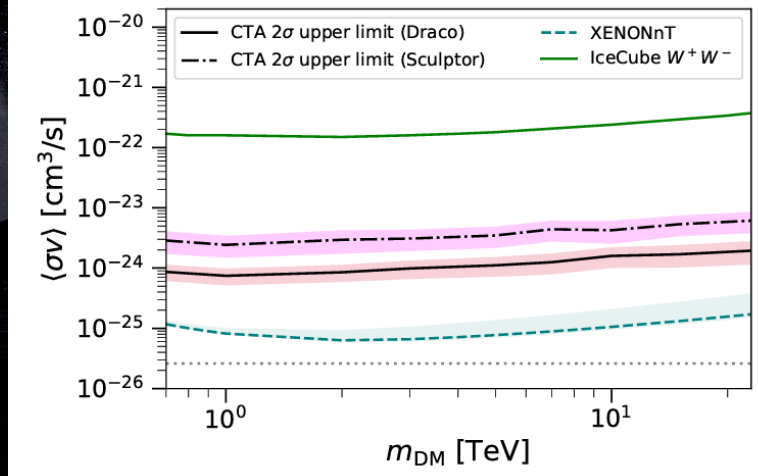




1st Thai-led CTA Sci paper

"CTA sensitivity on TeV scale-dark matter models with complementary limits from direct detection", Journal of Cosmology and Astroparticle Physics (JCAP),05,038, 2022





IOPscience Journals Books Publishing Support Login

Journal of Cosmology and Astroparticle Physics

PAPER

CTA sensitivity on TeV scale dark matter models with complementary limits from direct detection

C. Duangchan[1], C. Pongkitivanichkul[2], P. Uttayarat[3], A. Jardin-Blicq[4,5], M. Wechakama[1], T. Klangburam[2,5], W. Treesukrat[3], D. Samart[2], U. Sawangwit[5], A. Aguirre-Santaella[6,7] + Show full author list

Published 23 May 2022 • © 2022 IOP Publishing Ltd and Sissa Medialab

Journal of Cosmology and Astroparticle Physics, Volume 2022, May 2022

Citation C. Duangchan *et al* JCAP05(2022)038

Article PDF

+ Article information

Abstract

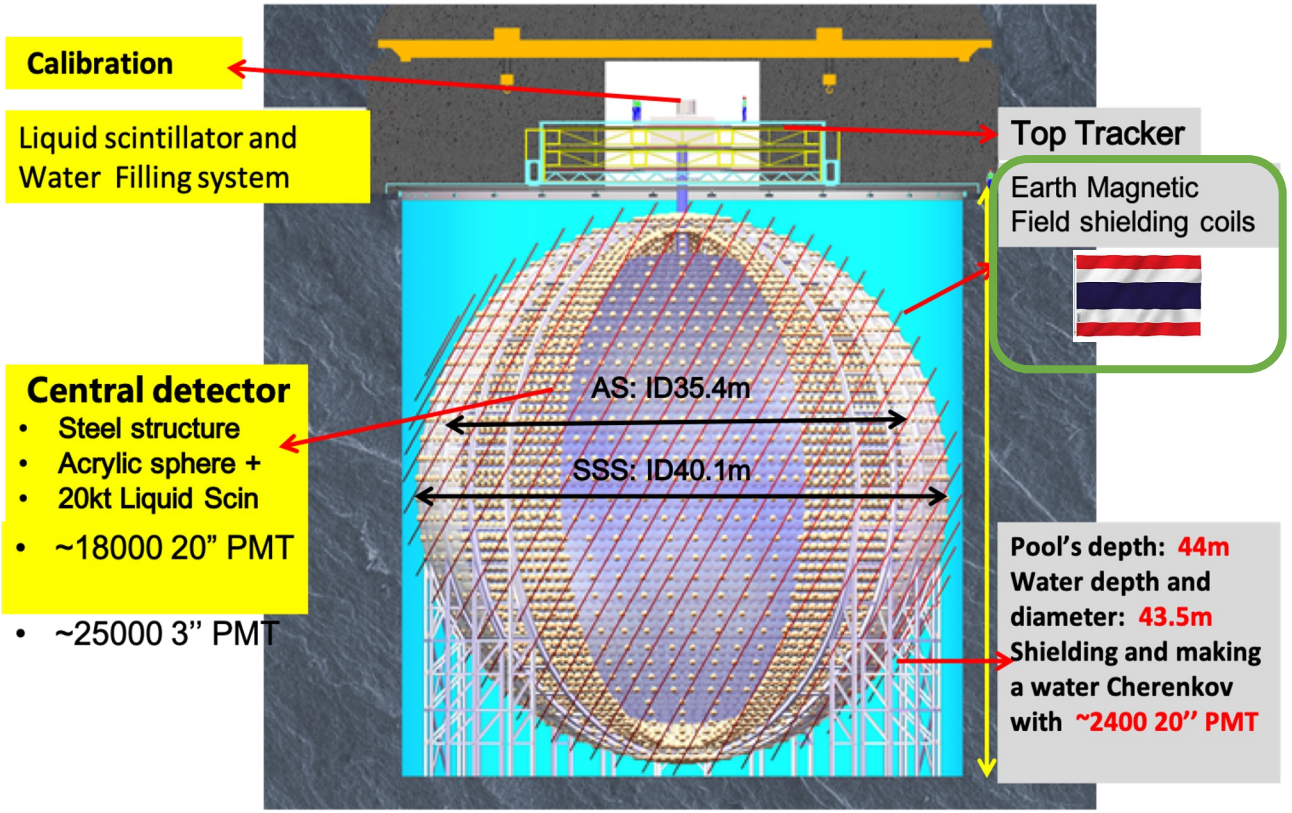
With ever increasing pressure from collider physics and direct detection experiments, particle physics models of TeV scale dark matter are gaining more attention. In this work, we consider two realizations of the class of scalar portal dark matter scenarios — the inverse seesaw model and the inert doublet model. Observations by the Cherenkov Telescope Array (CTA) of very-high-energy γ rays from dark matter annihilation in the context of these models are simulated for the Draco and Sculptor dwarf spheroidal galaxies, and later analyzed using `ctools`. We study the potential of CTA for the 5σ detection of a dark matter annihilation signal. In the absence of a signal, we also derive the 2σ upper limits on the annihilation cross-section. We compare our projected CTA sensitivity against the projected sensitivity of the next generation of direct detection experiment, i.e. XENONnT. Although the limits from CTA are significantly improved compared with the previous generations of γ-ray experiments, they are still ~2 orders of magnitude above the thermal relic cross-section for the considered targets.

JUNO Experiment

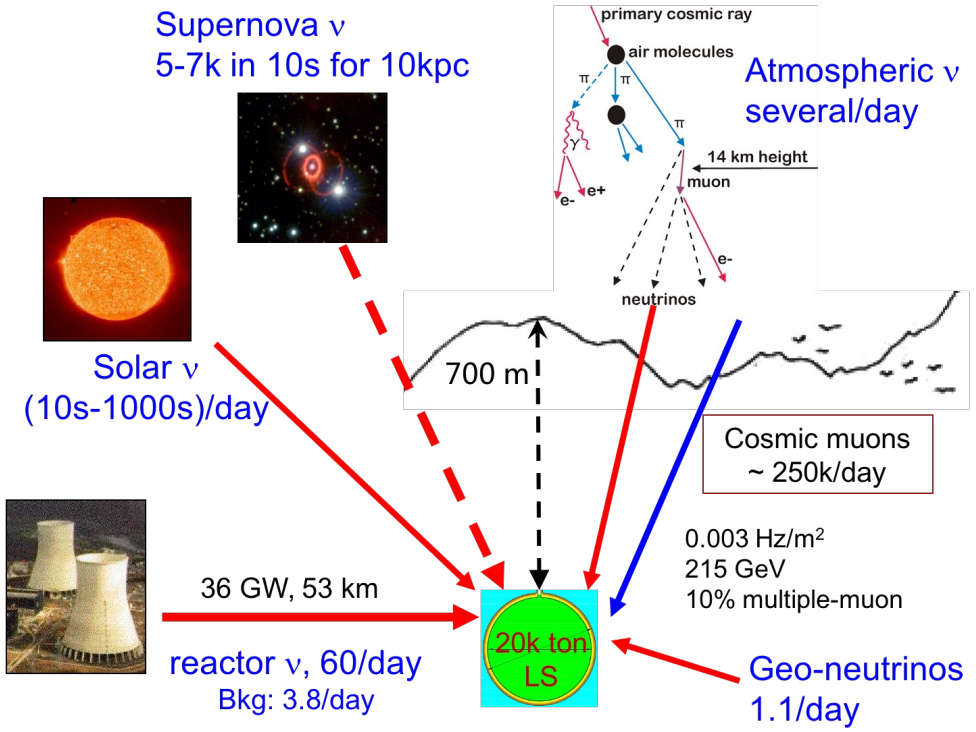
Jiangmen Underground Neutrino Observatory

- A multiple-purpose neutrino experiment, 300M$ budget (led by IHEP, China)
- 77 member institutes from 17 countries
- Thai-JUNO Consortium (NARIT, SUT & Chula)
- Thai contribution: Design and cost for the EMF shielding system (12M THB)

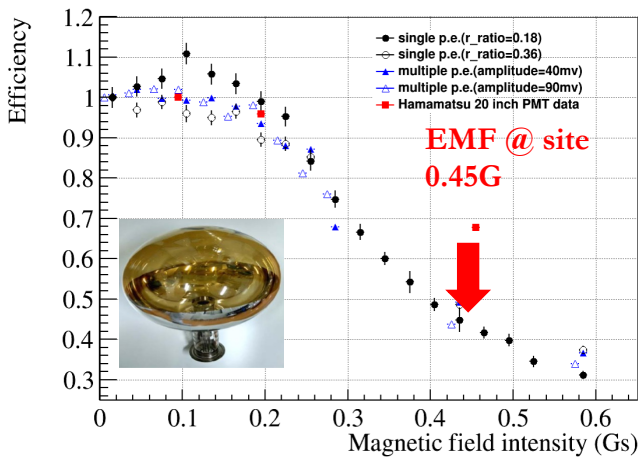
- 20 kton LS detector
- 3% energy resolution
- 700m underground
- Rich physics & High-Energy Astrophysics possibilities

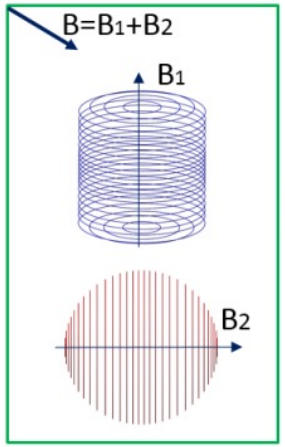


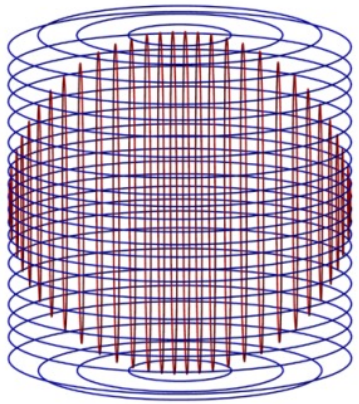
AS: Acrylic Sphere; SSS: Stainless Steel Structure

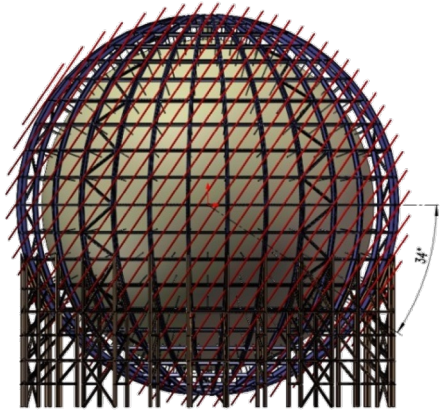


EMF Shielding Coil Optimization (2017 – 2019)

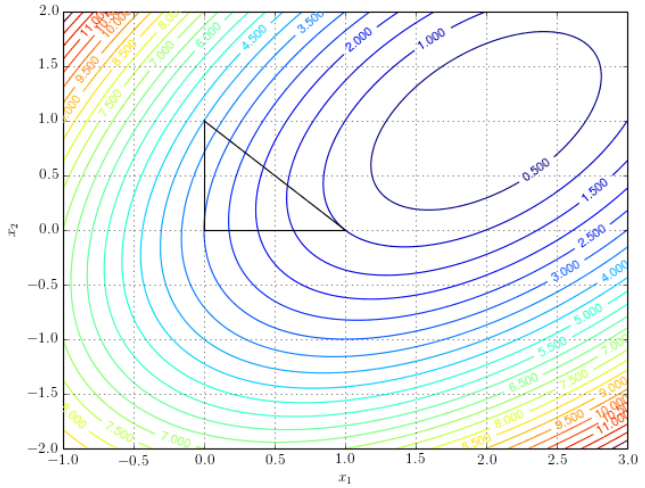






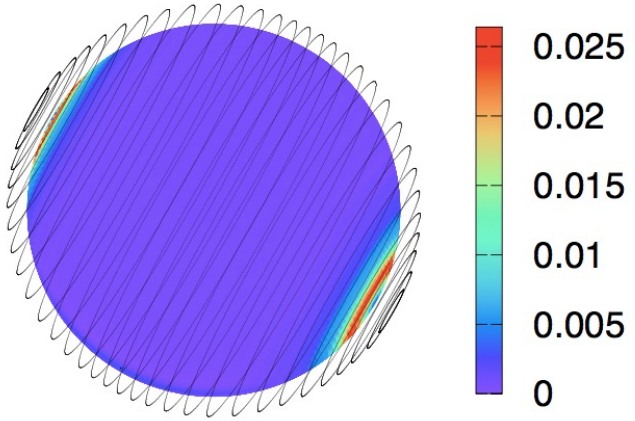


- **A lot of study was done, use different geometries and distance of the coils.**
- **One set/two set of coils.**
- **We chose one sets of coils：**
  - Installation：there are no big difference between one sets or two sets coils(we can have support structure on SS frame)
  - Coils and installation work quantity reduce ~50%;save cost；
  - One sets of coils performance are better than two sets as left figure shown.



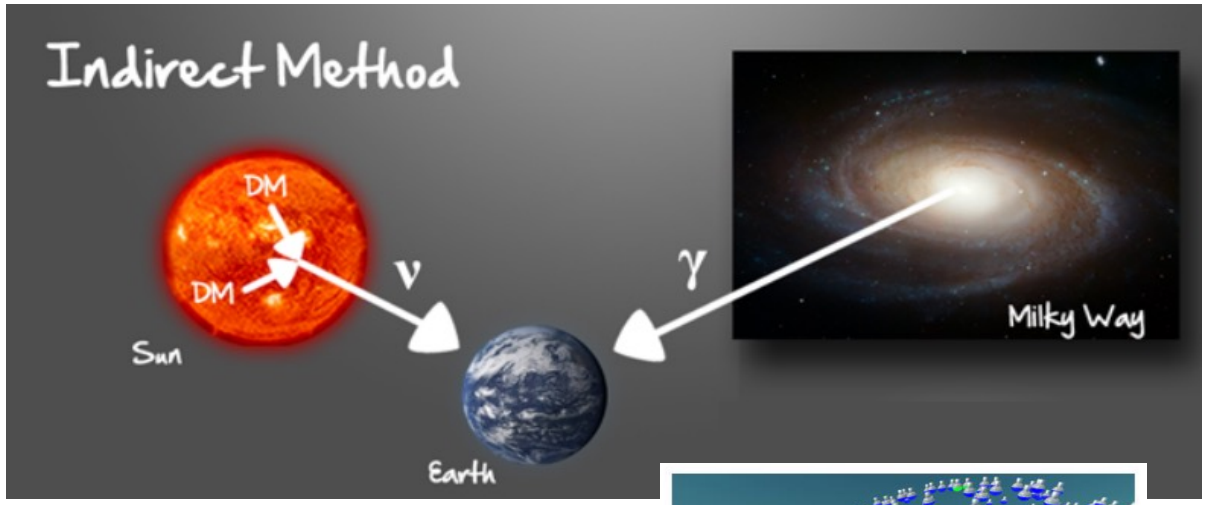
Simplex Method (Nelder and Mead)

B_{res} [G] of circular coils design



Requirement Reduce EMF @ experiment to < 10% (0.045G) ☑

Indirect Dark Matter detection with JUNO



Spectrum of neutrino from source → Neutrino interaction → Parameters in individual detector

- DM: WIMPSIM package [6]
- BG: JUNO software

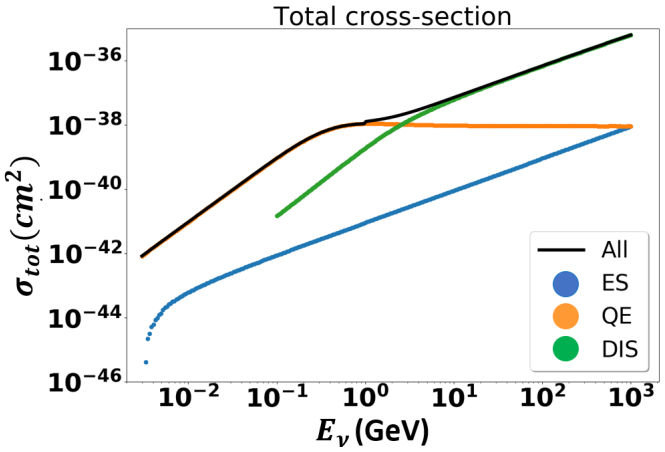
- 1-100 MeV: $\nu - e^-$ elastic scattering int.
- > 0.1– 2 GeV: QE $\nu - N$ scattering int.
- > 2 GeV: $\nu - N$ DIS int.

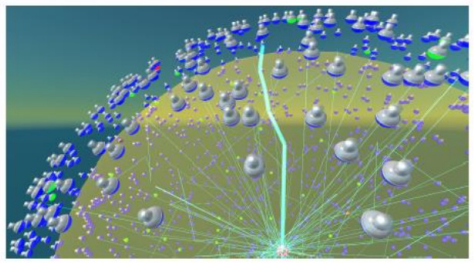
- Num. of particle
- Detector efficiency
  - energy

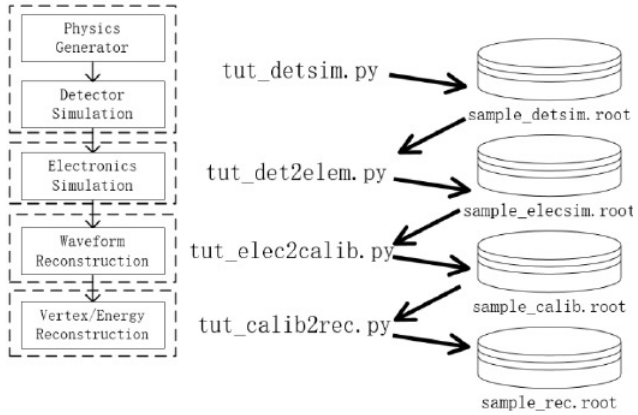
Signal from electronic device. → Analysis data 1 → Sensitivity of detectors → Comparison

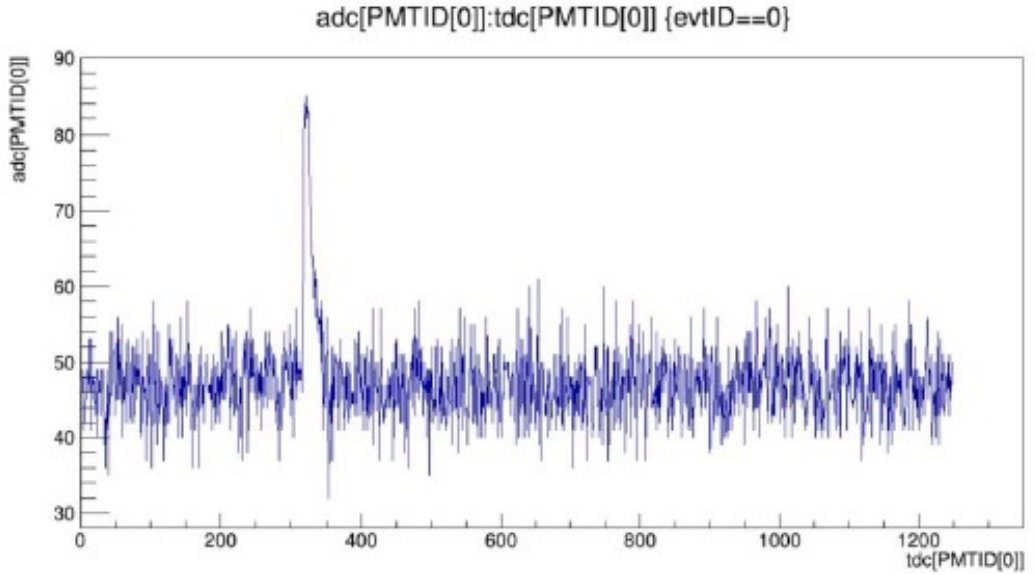
- The pulse of photon
- PSD, ML

- Compare with other experiments

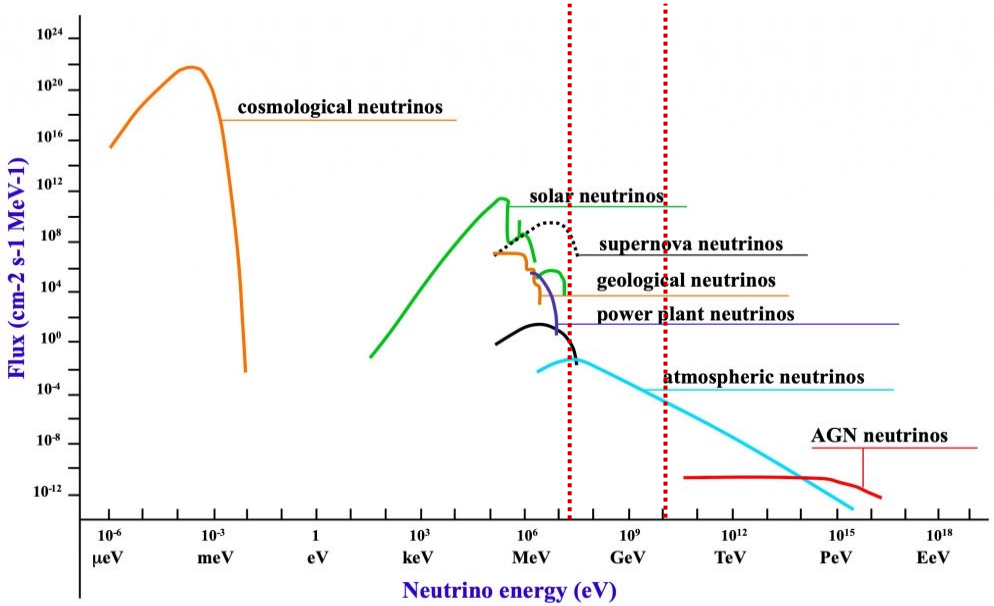


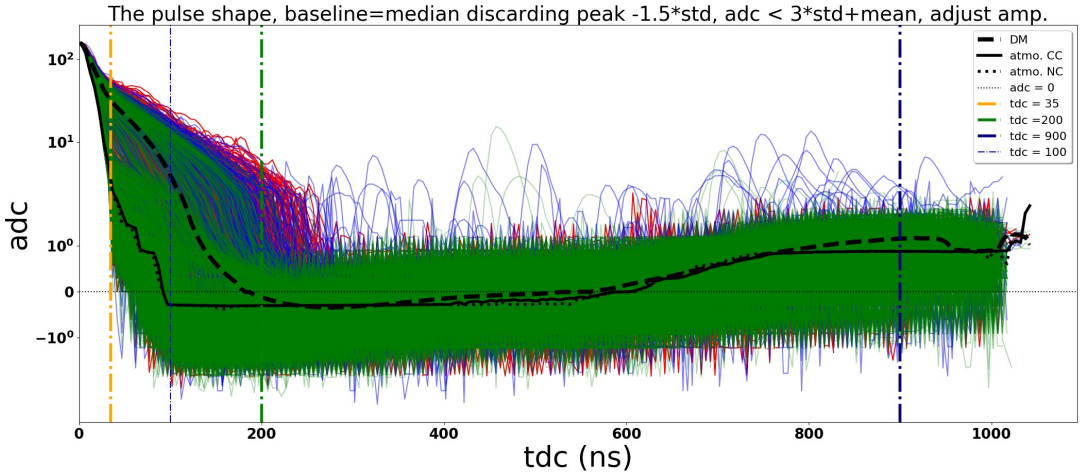


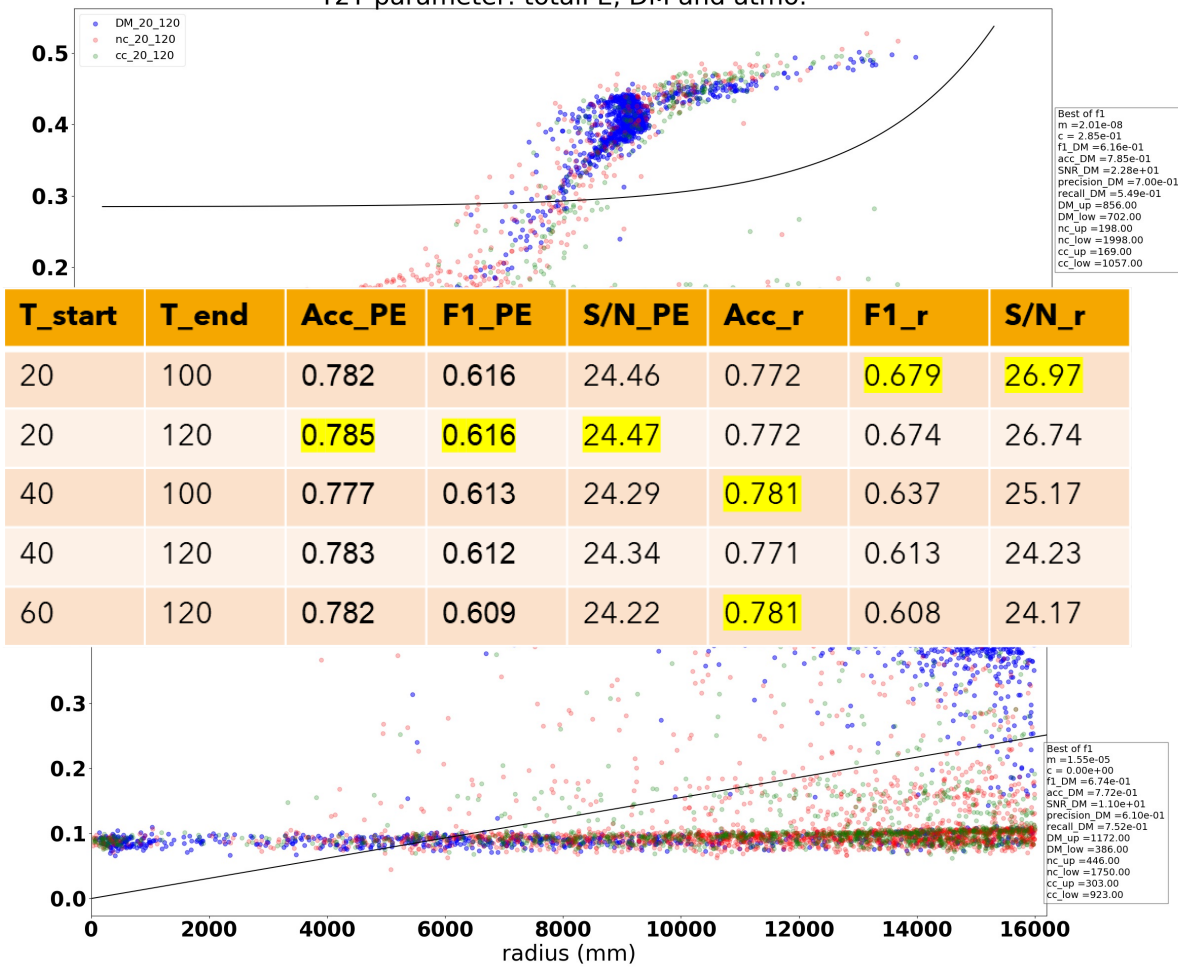




Signal/Background rejection/classification





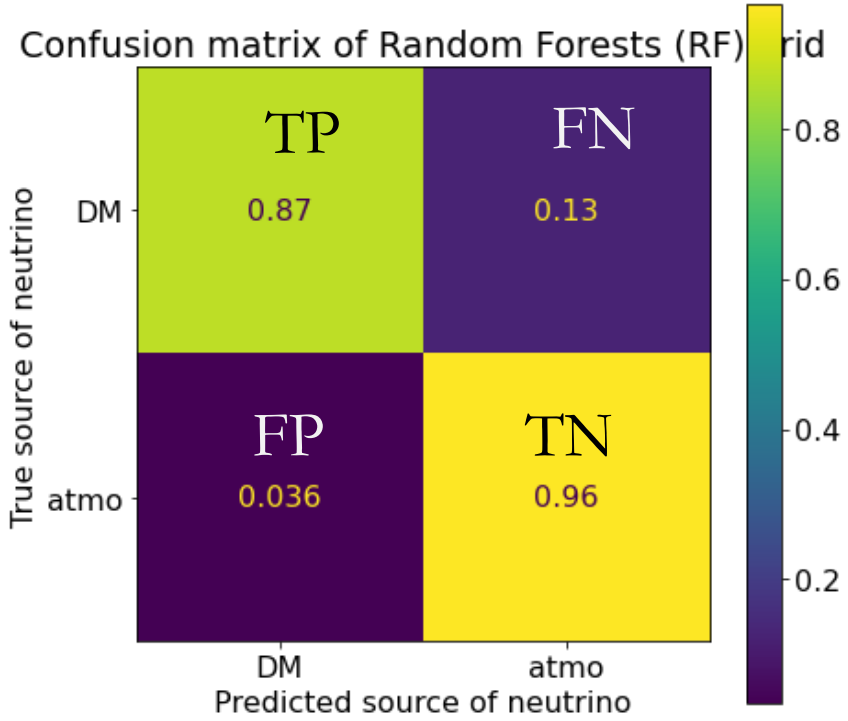


T_start	T_end	Acc_PE	F1_PE	S/N_PE	Acc_r	F1_r	S/N_r
20	100	0.782	0.616	24.46	0.772	0.679	26.97
20	120	0.785	0.616	24.47	0.772	0.674	26.74
40	100	0.777	0.613	24.29	0.781	0.637	25.17
40	120	0.783	0.612	24.34	0.771	0.613	24.23
60	120	0.782	0.609	24.22	0.781	0.608	24.17

$$\text{Tail to total} = T2T = \frac{\int_{t_{start}}^{t_{end}} adc}{\int adc}$$

Application of Machine Learning

- Machine Learning (ML) is used to classify the source of neutrino by many parameters.
- The algorithm of classification:
  - Support Vector Machine (SVM)
  - Decision Tree (DT)
  - Random Forest (RF)
  - K-nearest nearby (KNN)
- The training : test = 80 : 20.
- The dataset of atmo. ~ 3,000 and DM. ~ 1,500 but both source generate from ~2,000 neutrino events.



t2t	avgfwhm	avgfwtm	avgsnr	edep	totalPE	rec_x	rec_y	rec_z	rec_r	rec_t	label
0.088016	12.43918	26.72496	8.621387	3270.438	4546137	3780.269	10148.83	-1589.33	10946.01	75.97283	atmo
0.251026	19.63414	51.80563	7.443627	237.1081	337709	-2759.03	-11581.7	9036.714	14946.91	216.639	atmo
0.124321	13.75763	30.68443	8.418034	37.42637	23317	5681.028	-8821.68	-125.831	10493.43	158.0702	atmo
0.315048	22.26553	58.2368	7.41682	353.5436	440749	2690.307	-13075.2	1381.208	13420.41	187.137	atmo
0.095424	12.38672	26.493	8.574283	105.4499	110068	-8981.42	-3343.99	1125.98	9649.66	57.47731	atmo
0.103059	12.52885	26.55711	8.573019	550.2937	659049	-6886.27	3497.977	8289.013	11329.8	167.69	atmo
0.104565	12.87998	27.68803	8.646714	2242.681	3192779	-6322.04	-6629.24	12463.32	15467.69	192.1136	atmo

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FN+TN}$$

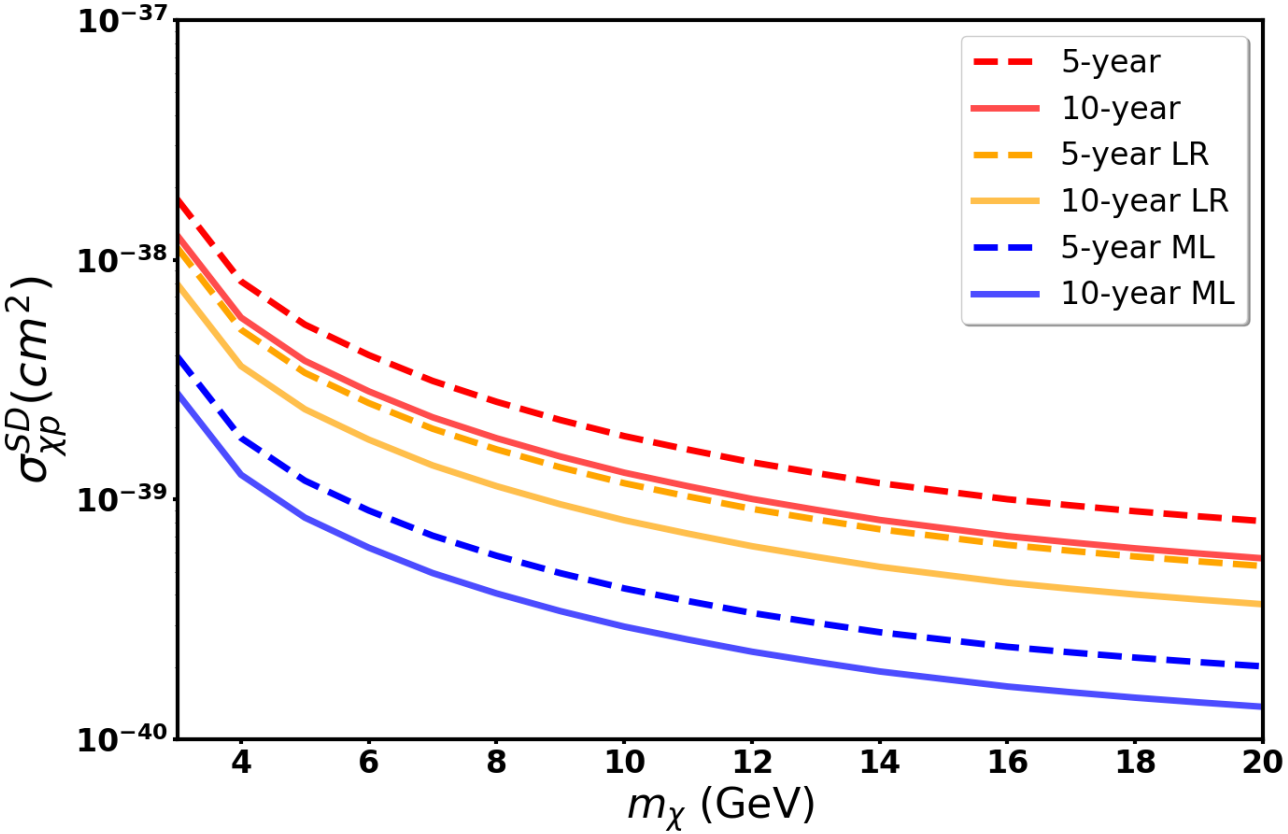
$$\text{F1 score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

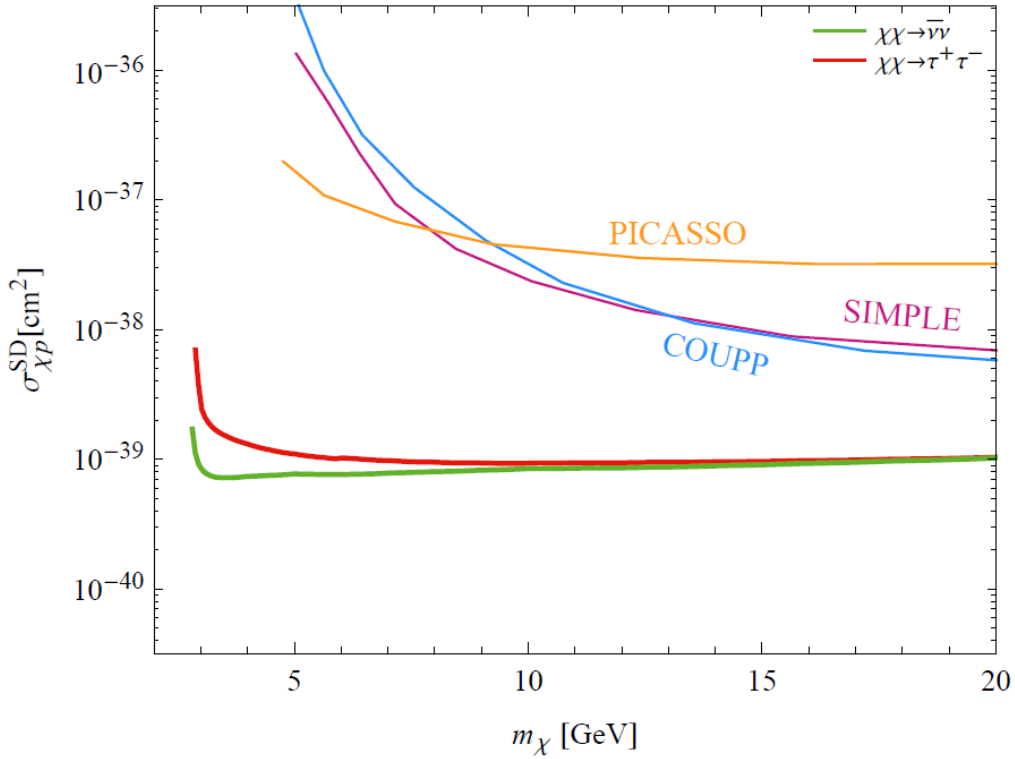
$$\text{Recall} = \frac{TP}{TP+FN}$$

WIMP Sensitivity: 5 & 10-year experiments

This Work



Ann et al (2016) "JUNO Yellow Book"



The Team

Partner Institutes

NATIONAL E-SCIENCE INFRASTRUCTURE CONSORTIUM

THE UNIVERSITY OF WARWICK

Durham University

The University Of Sheffield.

UNIVERSITY OF PORTSMOUTH

LIVERPOOL JOHN MOORES UNIVERSITY

ICRAR

+2 new members
(PostDoc & RA)

- Alumni (Japan)

- Alumni (CU)

- Alumni (UK)

- Alumni (S. Korea)

Chalawan on the Cloud

slurm workload manager

jupyter

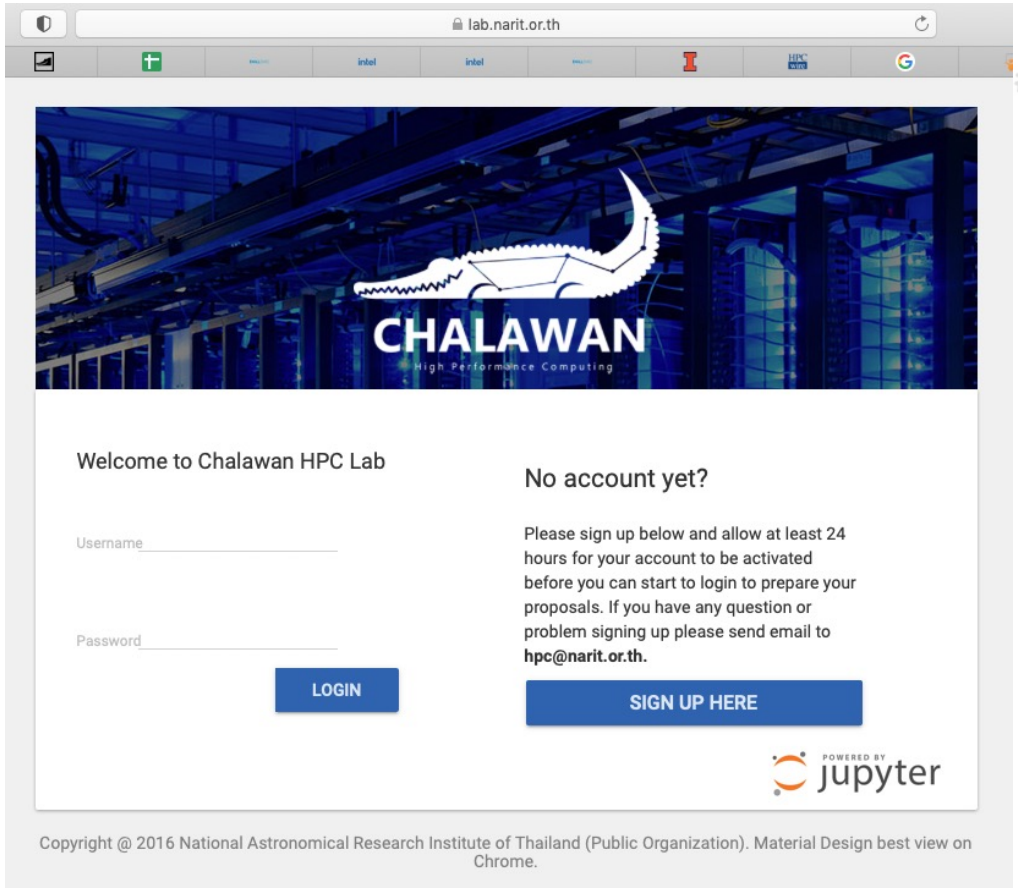
julia

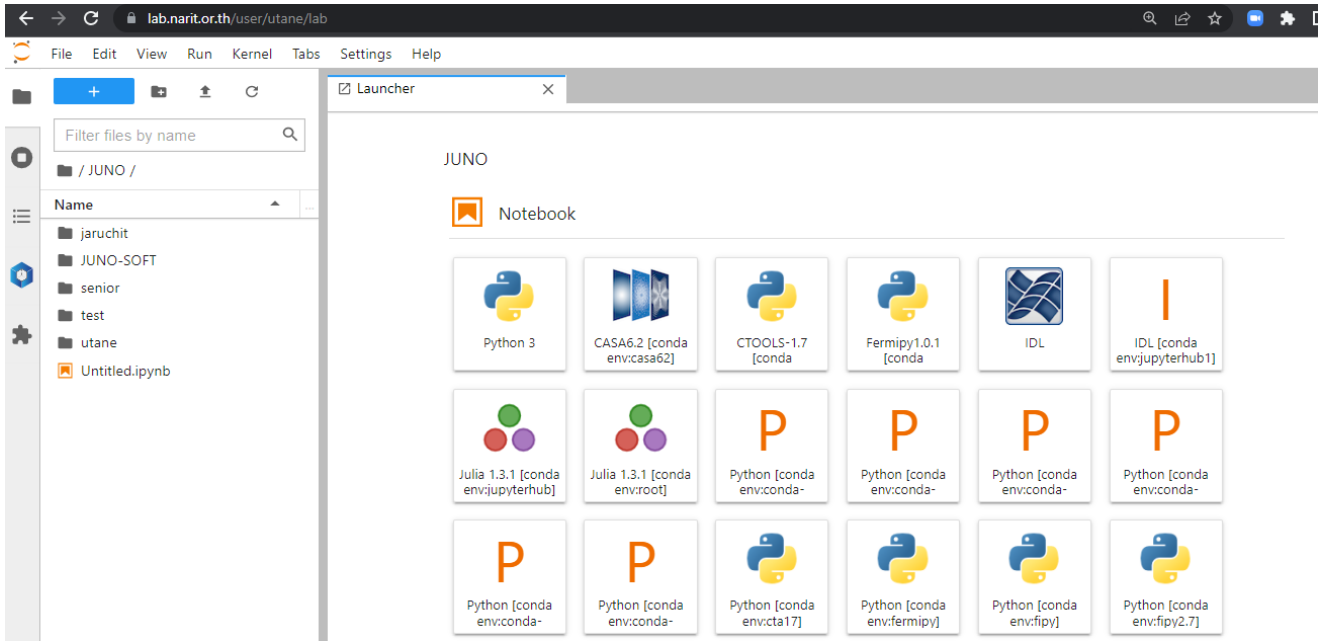
python

CASA Common Astronomy Software Applications

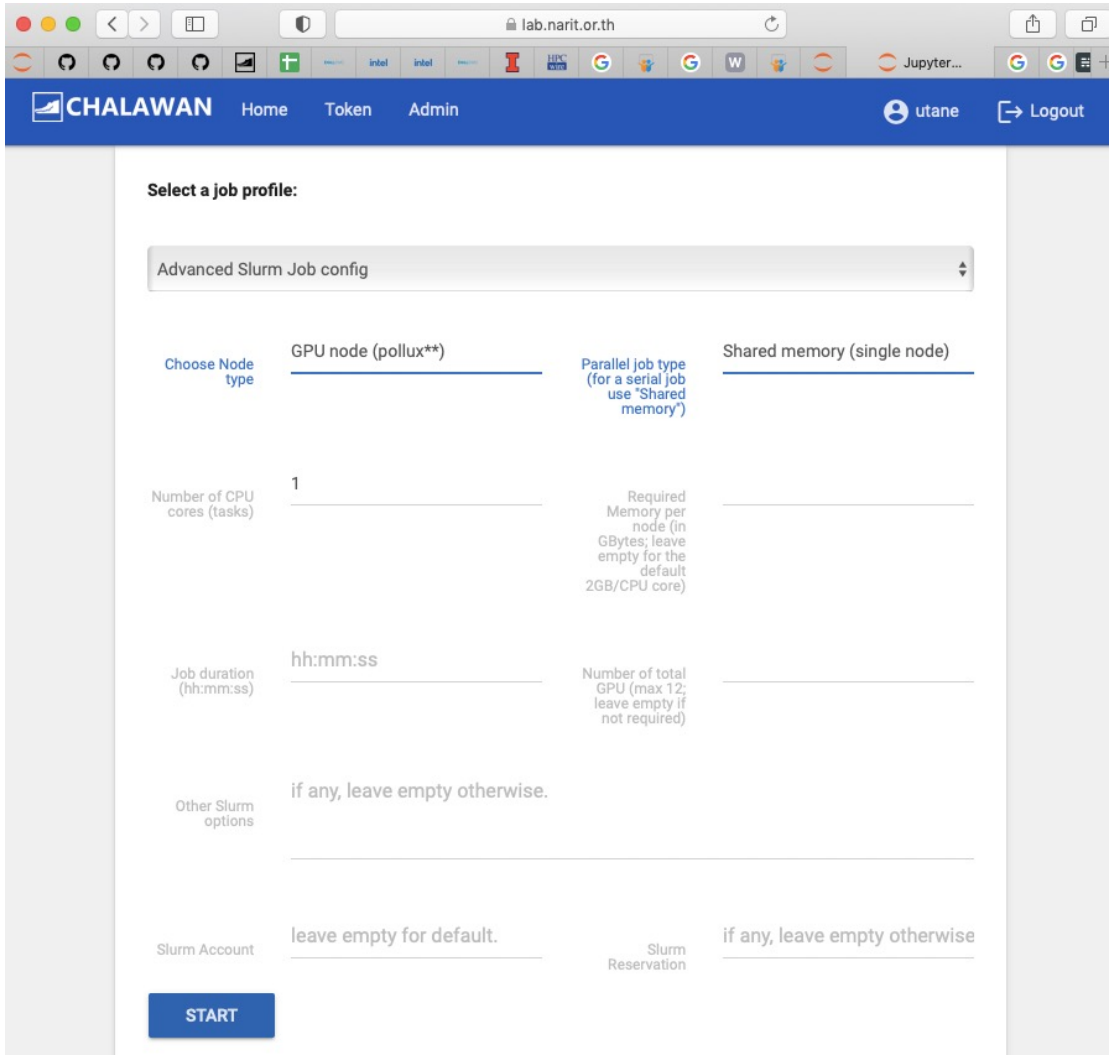
R

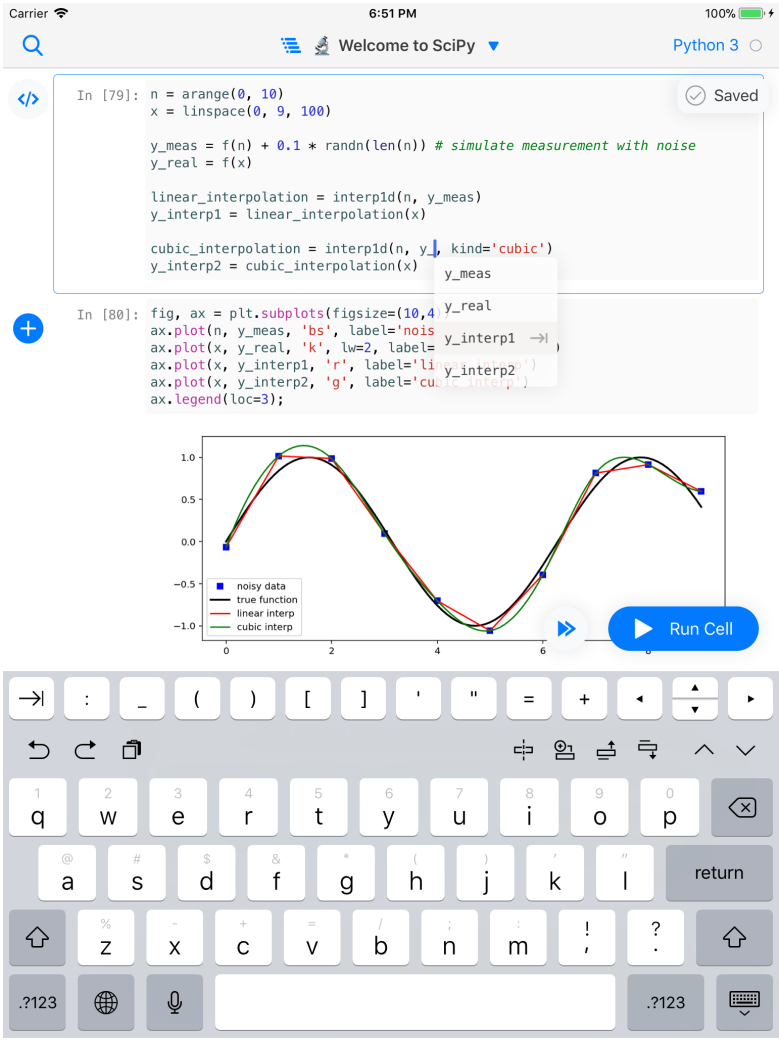
IDL

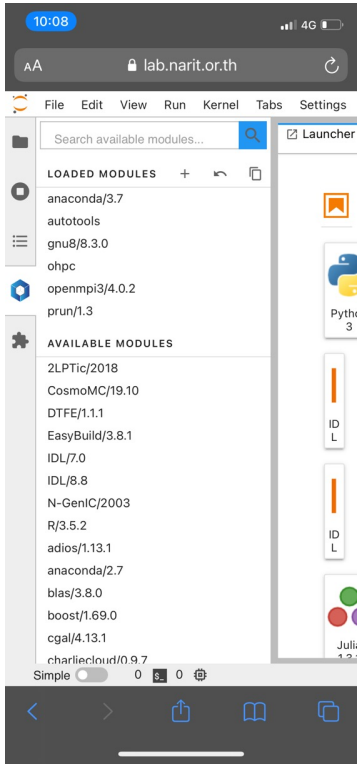




Chalawan on the Cloud (contd.)



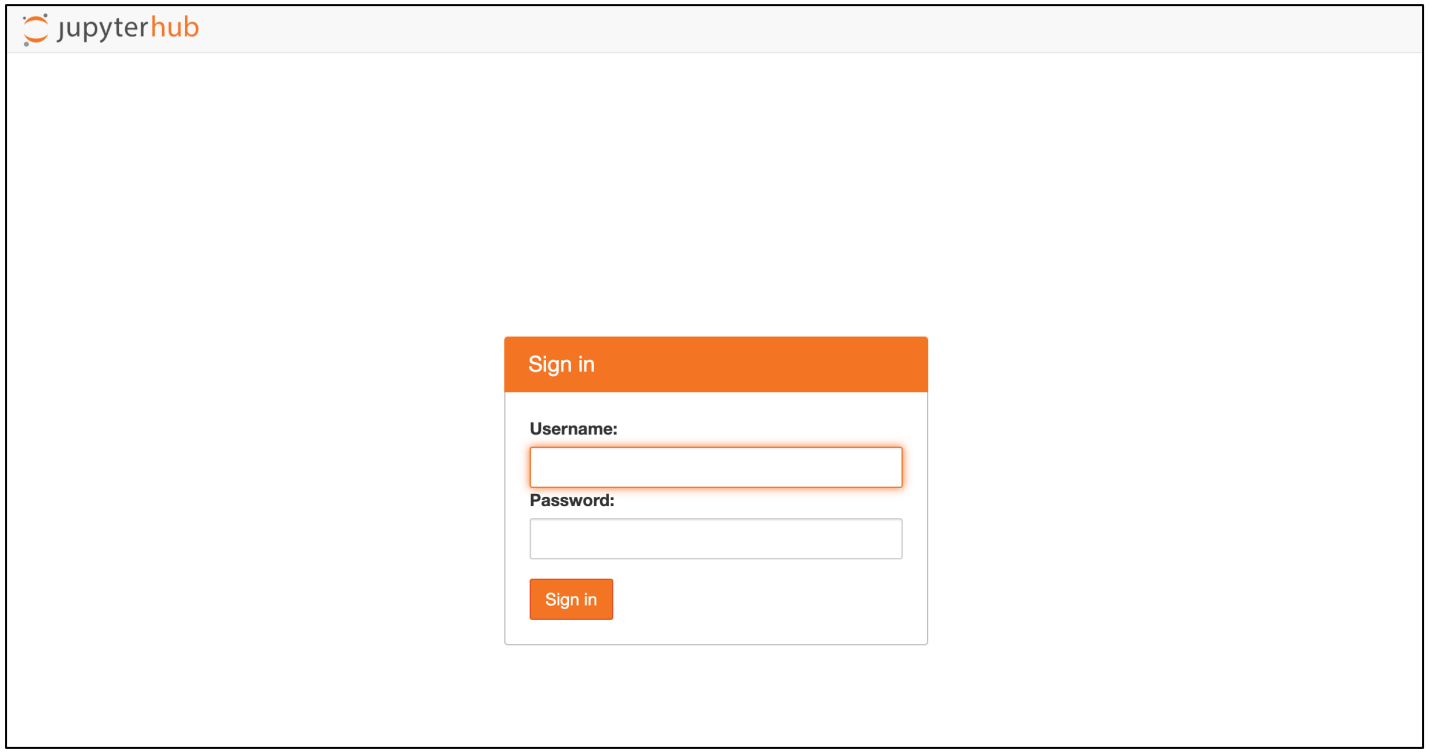




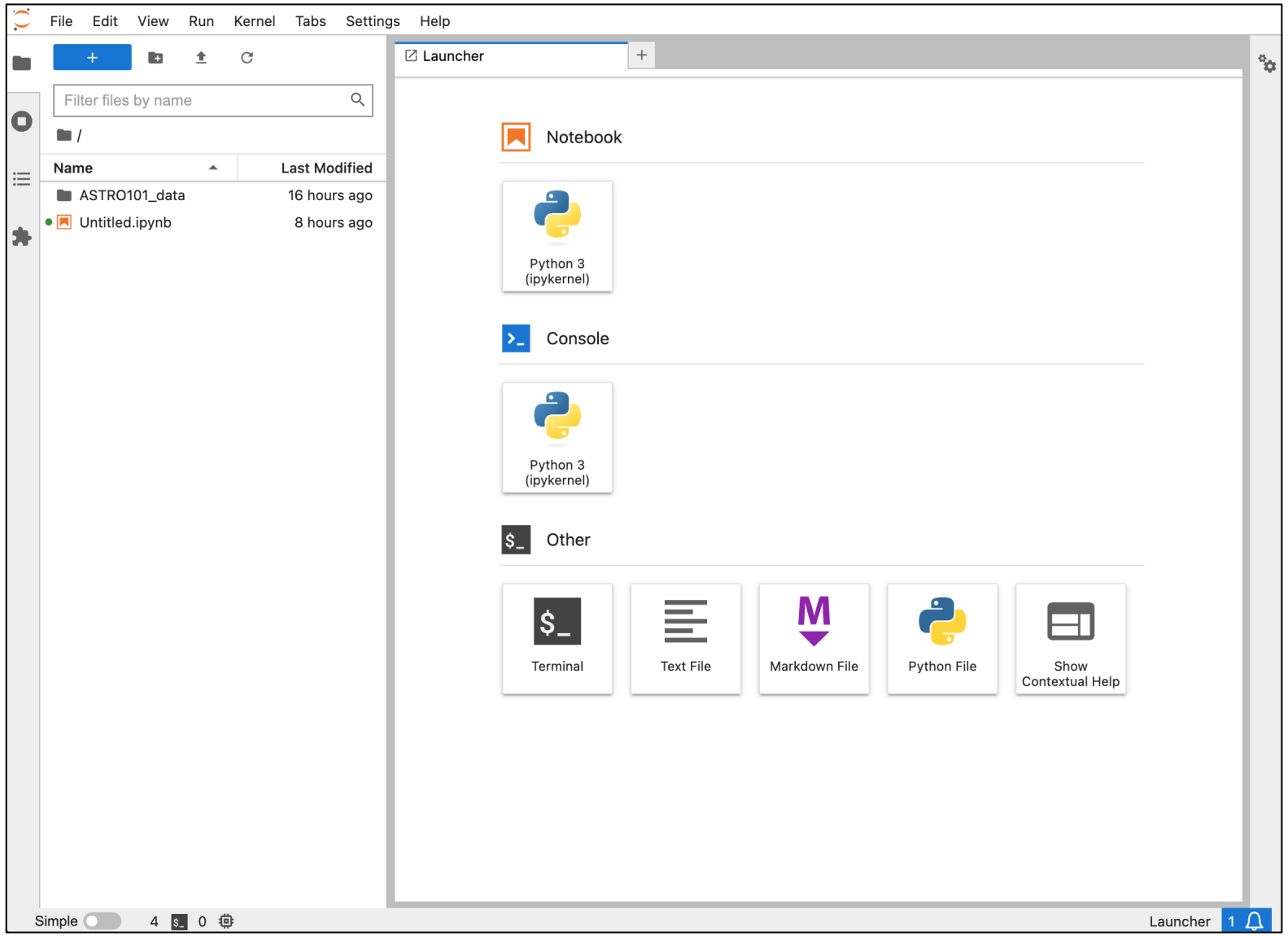
For this workshop

- https://lab1.narit.or.th
- https://lab2.narit.or.th

- Each with 64 CPU cores, 512GB RAM, Lustre Parallel Filesystem



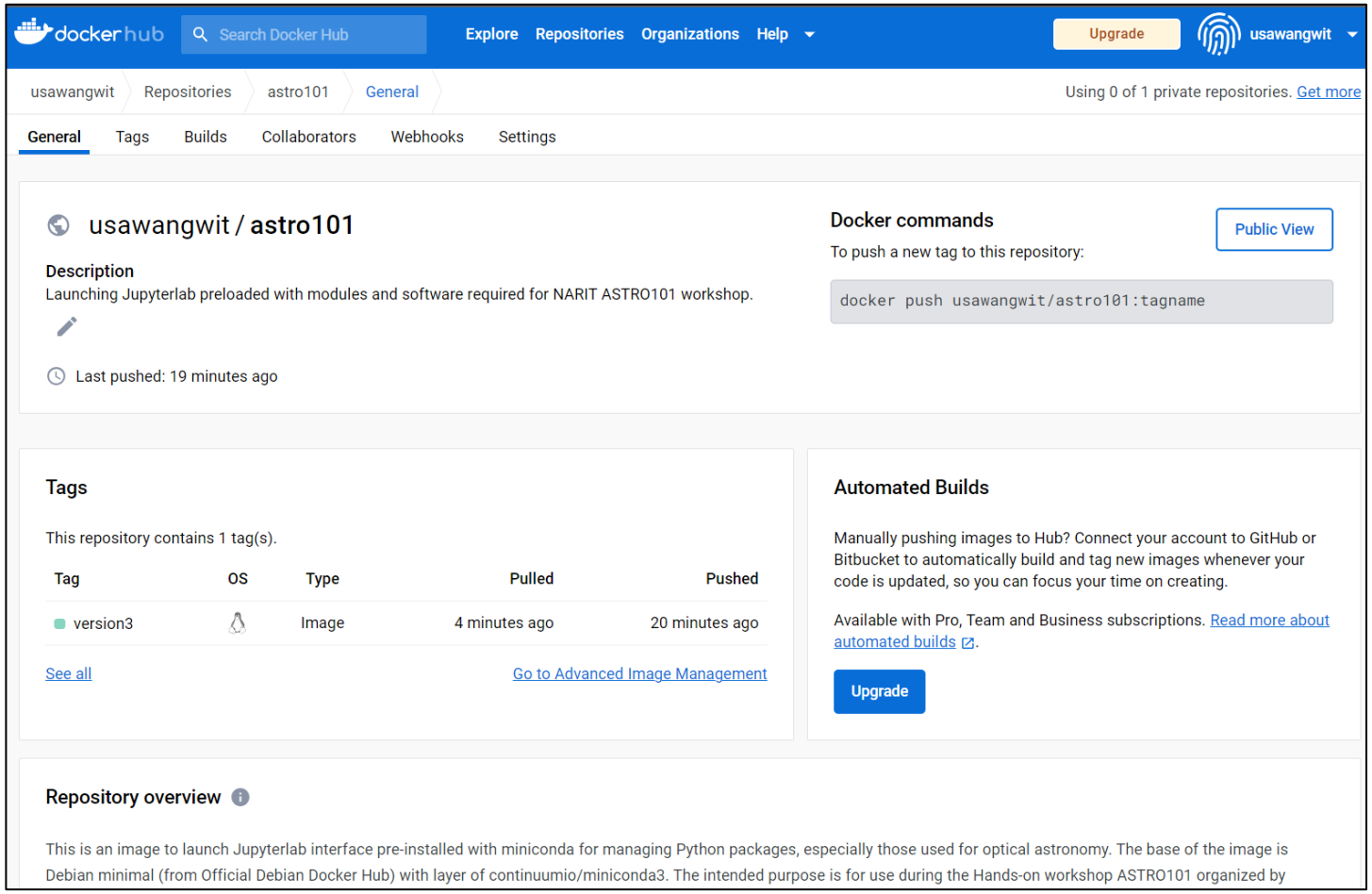
For this workshop



Docker Hub Repository

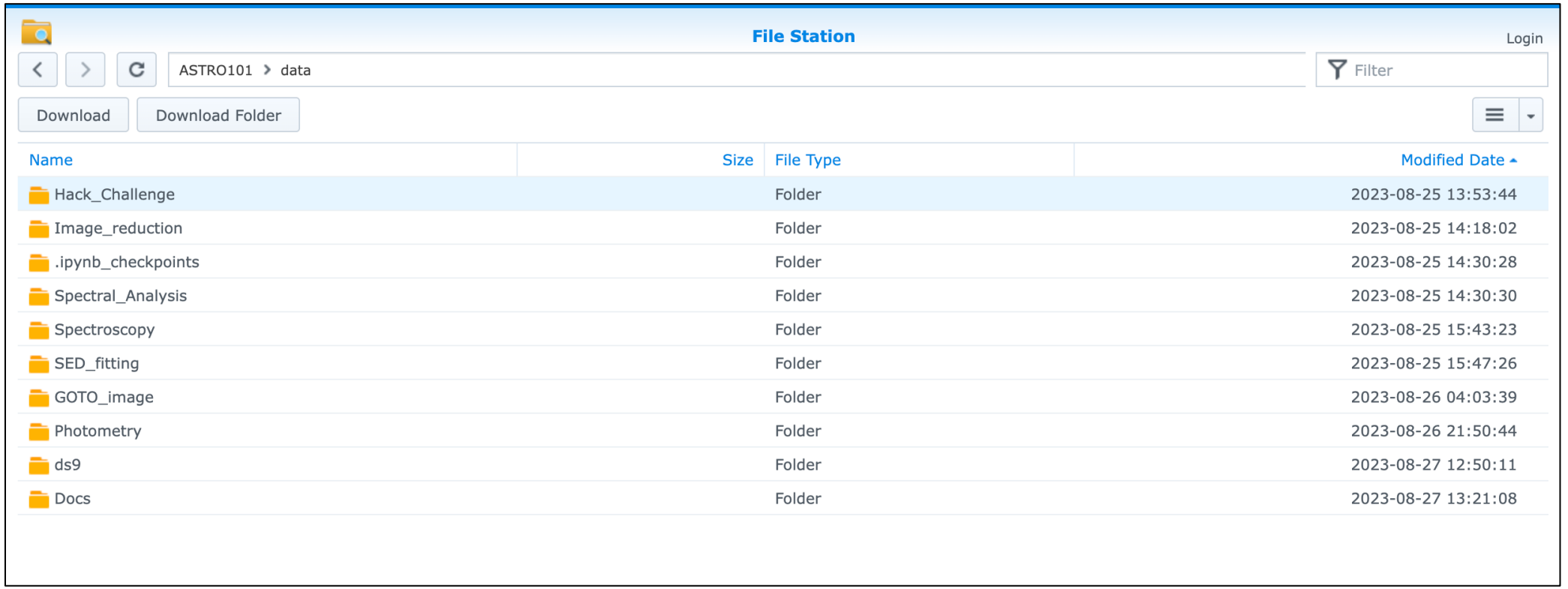
- https://tinyurl.com/usawangwitrepo (Full URL https://hub.docker.com/repository/docker/usawangwit/astro101/general)

- *$docker pull usawangwit/astro101:version3*



Data Access for Online Participants

- *http://gofile.me/4gxKb/QZFL5N7jA*

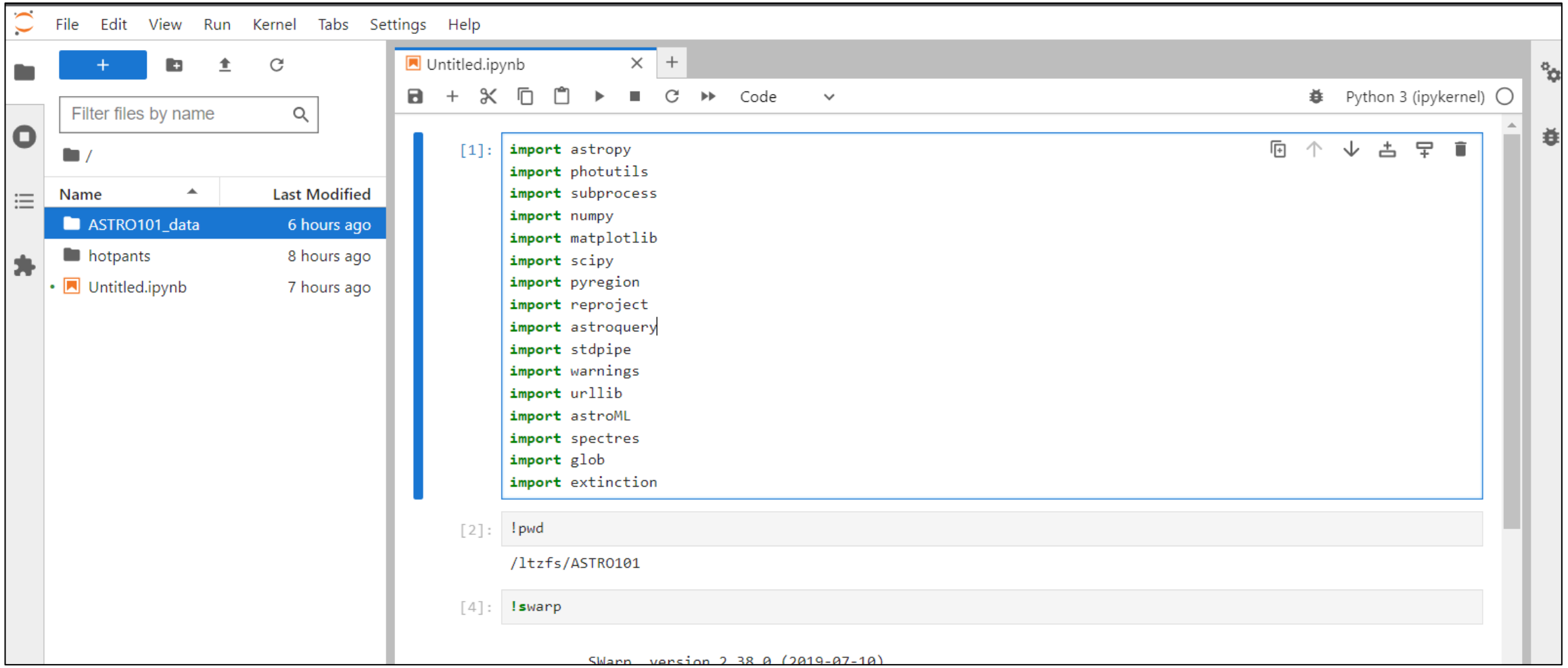


Docker Run

$ docker run -i -t -p 8888:8888 -v /PATH_TO/ASTRO101/data/:/ltzfs/ASTRO101/ASTRO101_data usawangwit/astro101:version3

- Copy and Paste to your browser

```
(base) utane@LAPTOP-RB2EJHT4:~$ docker run -i -t -p 8888:8888 -v /mnt/c/Users/Utane/ASTRO101/ASTRO101/data/:/ltzfs/AS
TRO101/ASTRO101_data usawangwit/astro101:version3
/bin/bash: /opt/conda/lib/libtinfo.so.6: no version information available (required by /bin/bash)
[I 2023-08-28 00:31:52.442 ServerApp] Package jupyterlab took 0.0000s to import
[I 2023-08-28 00:31:52.450 ServerApp] Package jupyter_lsp took 0.0070s to import
[W 2023-08-28 00:31:52.450 ServerApp] A `_jupyter_server_extension_points` function was not found in jupyter_lsp. Ins
tead, a `_jupyter_server_extension_paths` function was found and will be used for now. This function name will be dep
recated in future releases of Jupyter Server.
[I 2023-08-28 00:31:52.453 ServerApp] Package jupyter_server_terminals took 0.0029s to import
[I 2023-08-28 00:31:52.453 ServerApp] Package notebook_shim took 0.0000s to import
[W 2023-08-28 00:31:52.453 ServerApp] A `_jupyter_server_extension_points` function was not found in notebook_shim. I
nstead, a `_jupyter_server_extension_paths` function was found and will be used for now. This function name will be d
eprecated in future releases of Jupyter Server.
[I 2023-08-28 00:31:52.454 ServerApp] jupyter_lsp | extension was successfully linked.
[I 2023-08-28 00:31:52.456 ServerApp] jupyter_server_terminals | extension was successfully linked.
[I 2023-08-28 00:31:52.459 ServerApp] jupyterlab | extension was successfully linked.
[I 2023-08-28 00:31:52.578 ServerApp] notebook_shim | extension was successfully linked.
[W 2023-08-28 00:31:52.587 ServerApp] WARNING: The Jupyter server is listening on all IP addresses and not using encr
yption. This is not recommended.
[I 2023-08-28 00:31:52.587 ServerApp] notebook_shim | extension was successfully loaded.
[I 2023-08-28 00:31:52.589 ServerApp] jupyter_lsp | extension was successfully loaded.
[I 2023-08-28 00:31:52.589 ServerApp] jupyter_server_terminals | extension was successfully loaded.
[I 2023-08-28 00:31:52.590 LabApp] JupyterLab extension loaded from /opt/conda/lib/python3.11/site-packages/jupyterla
b
[I 2023-08-28 00:31:52.590 LabApp] JupyterLab application directory is /opt/conda/share/jupyter/lab
[I 2023-08-28 00:31:52.590 LabApp] Extension Manager is 'pypi'.
[I 2023-08-28 00:31:52.592 ServerApp] jupyterlab | extension was successfully loaded.
[I 2023-08-28 00:31:52.592 ServerApp] Serving notebooks from local directory: /ltzfs/ASTRO101
[I 2023-08-28 00:31:52.592 ServerApp] Jupyter Server 2.7.1 is running at:
[I 2023-08-28 00:31:52.592 ServerApp] http://localhost:8888/lab?token=ddc6305ba44e0f34f6805f9f931e20a20c14fb91604b6a3
3
[I 2023-08-28 00:31:52.592 ServerApp]     http://127.0.0.1:8888/lab?token=ddc6305ba44e0f34f6805f9f931e20a20c14fb91604
b6a33
[I 2023-08-28 00:31:52.592 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip conf
irmation).
[C 2023-08-28 00:31:52.595 ServerApp]
```

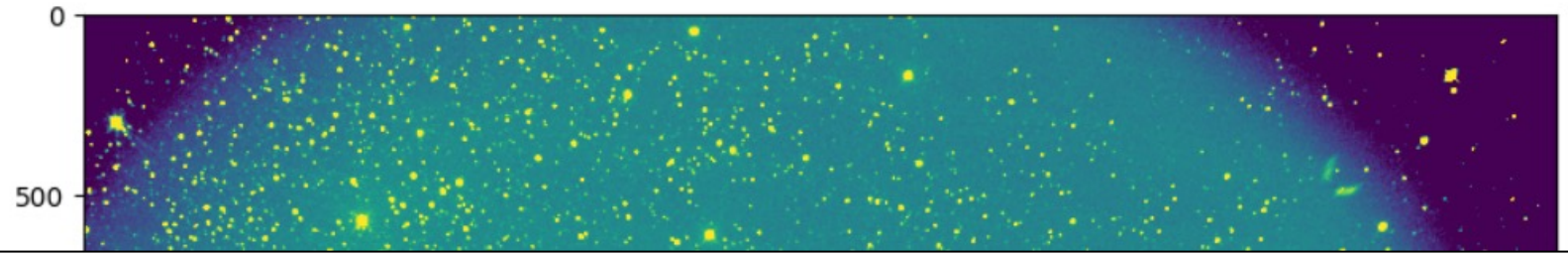

File Edit View Run Kernel Tabs Settings Help
Filter files by name
/
Name
Last Modified
ASTRO101_data 6 hours ago
hotpants 8 hours ago
Untitled.ipynb 7 hours ago
Untitled.ipynb
Code
Python 3 (ipykernel)
[1]:
import astropy
import photutils
import subprocess
import numpy
import matplotlib
import scipy
import pyregion
import reproject
import astroquery
import stdpipe
import warnings
import urllib
import astroML
import spectres
import glob
import extinction
[2]:
!pwd
/ltzfs/ASTRO101
[4]:
!swarp

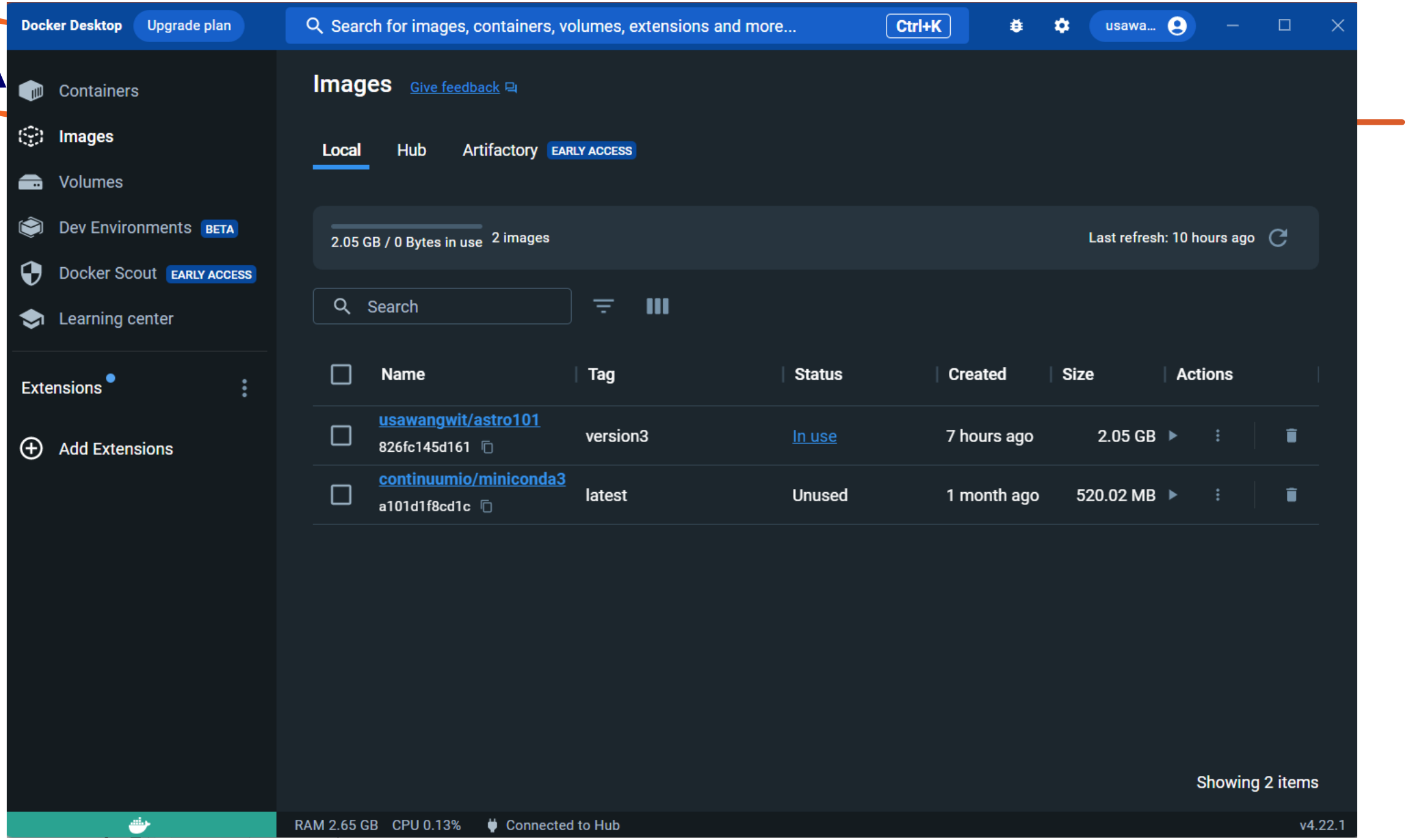
Untitled.ipynb | photometry.ipynb

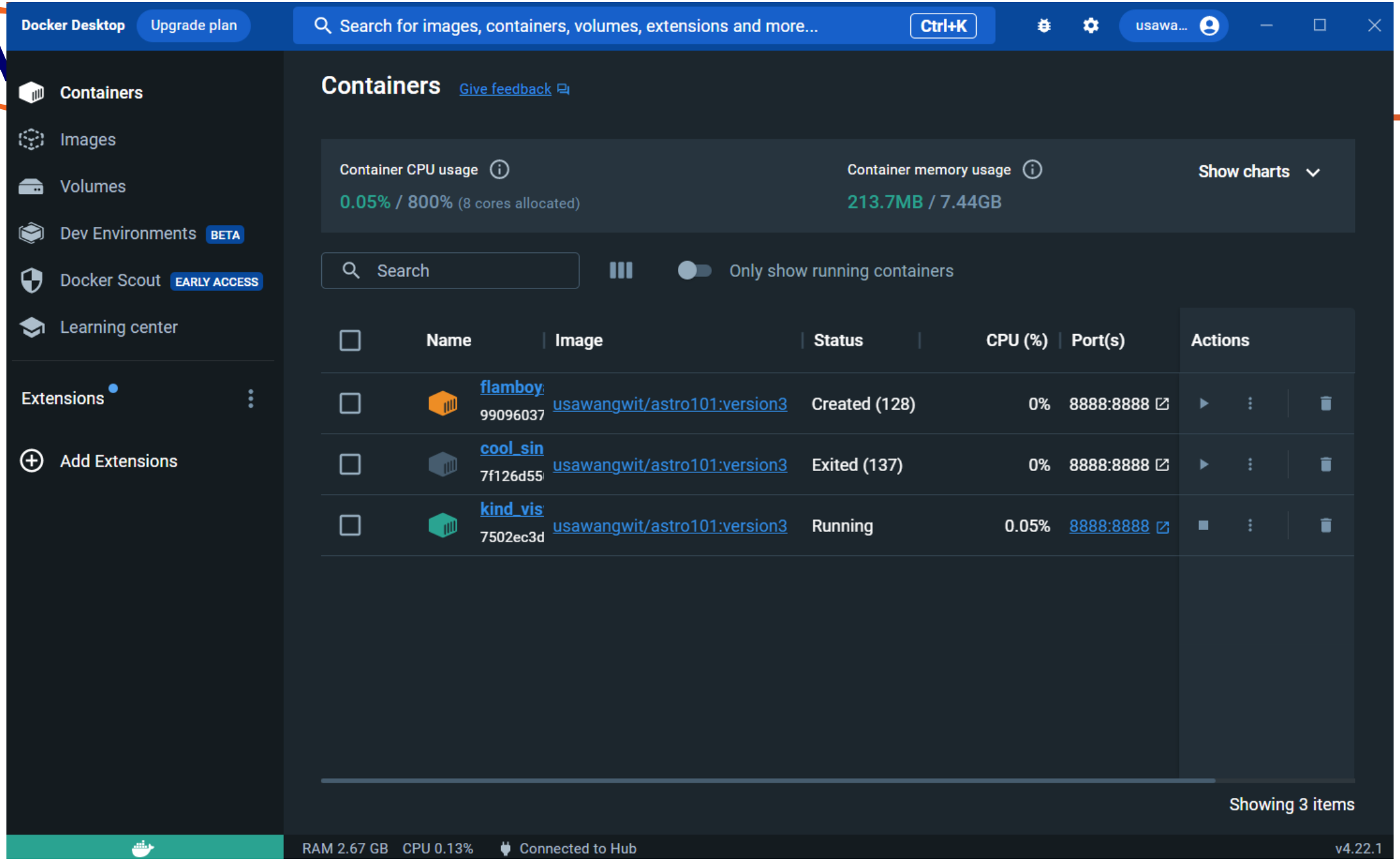
Markdown | Python 3 (ipykernel)

[3]:

```
from astropy.io import fits
import matplotlib.pyplot as plt
%matplotlib inline
import os

# Move to the data directory for our analysis
data_dir = '/ltzfs/ASTRO101/ASTRO101_data/Photometry/data/'
imageName = 'aC0_20181013-174714-557.wcs.fits.proc.cr.fits'
f = fits.open(data_dir + imageName)
data = f[0].data  #This is the image array
header = f[0].header

#Compute some image statistics for scaling the image plot
mean, median, sigma = sigma_clipped_stats(data)

#plot the image with some reasonable scale
plt.figure(figsize=(10,10))
plt.imshow(data, vmin=median-3*sigma, vmax=median+3*sigma)
plt.show()
```




Docker Desktop
Upgrade plan
Search for images, containers, volumes, extensions and more...
Ctrl+K
usawa...
Containers
Images
Volumes
Dev Environments
BETA
Docker Scout
EARLY ACCESS
Learning center
Extensions
Add Extensions
Images
Give feedback
Local
Hub
Artifactory
EARLY ACCESS
2.05 GB / 0 Bytes in use
2 images
Last refresh: 10 hours ago
Search
Name
Tag
Status
Created
Size
Actions
usawangwit/astro101
826fc145d161
version3
In use
7 hours ago
2.05 GB
continuumio/miniconda3
a101d1f8cd1c
latest
Unused
1 month ago
520.02 MB
Showing 2 items
RAM 2.65 GB
CPU 0.13%
Connected to Hub
v4.22.1


Docker Desktop
Upgrade plan
Search for images, containers, volumes, extensions and more...
Ctrl+K
usawa...
Containers
Images
Volumes
Dev Environments
BETA
Docker Scout
EARLY ACCESS
Learning center
Extensions
Add Extensions
Containers
Give feedback
Container CPU usage
0.05% / 800% (8 cores allocated)
Container memory usage
213.7MB / 7.44GB
Show charts
Search
Only show running containers
Name
Image
Status
CPU (%)
Port(s)
Actions
flamboy
99096037
usawangwit/astro101:version3
Created (128)
0%
8888:8888
cool_sin
7f126d55
usawangwit/astro101:version3
Exited (137)
0%
8888:8888
kind_vis
7502ec3d
usawangwit/astro101:version3
Running
0.05%
8888:8888
Showing 3 items
RAM 2.67 GB
CPU 0.13%
Connected to Hub
v4.22.1


WITH THE RIGHT TOOLS YOU CAN BUILD ANYTHING,
WITH THE RIGHT TECHNOLOGY YOU CAN BRING INNOVATIONS TO LIFE
ขอบคุณครับ
THANK YOU