Atlas Release Env & Analysis Support

Shuwei Ye

Intro Talk for NPPS group meeting

Sept 2019

Brief about myself

- Prior to joining Atlas, I worked on

  BES, AMS and BaBar (discovered two puzzling particles, Y(4260) and Y(4350))

- In 2007, I joined Atlas group at BNL
  - 50% on USAtlas analysis support (Omega)
  - 50% on Software development & librarian (PAS=>NPPS): LXR & AtlasSetup

Atlas Source Code Cross-Reference

LXR is a general-purpose source code indexer and cross-reference,
providing web-based browsing of source code,
with links to the definition and usage of any identifier.

And Atlas code is mainly in C++ and python.

At BNL, we (Alex and me) provide 2 LXR servers:

- Nightlies releases: https://acode-browser**1**.usatlas.bnl.gov/lxr/source
- Stable releases: https://acode-browser**2**.usatlas.bnl.gov/lxr/source/

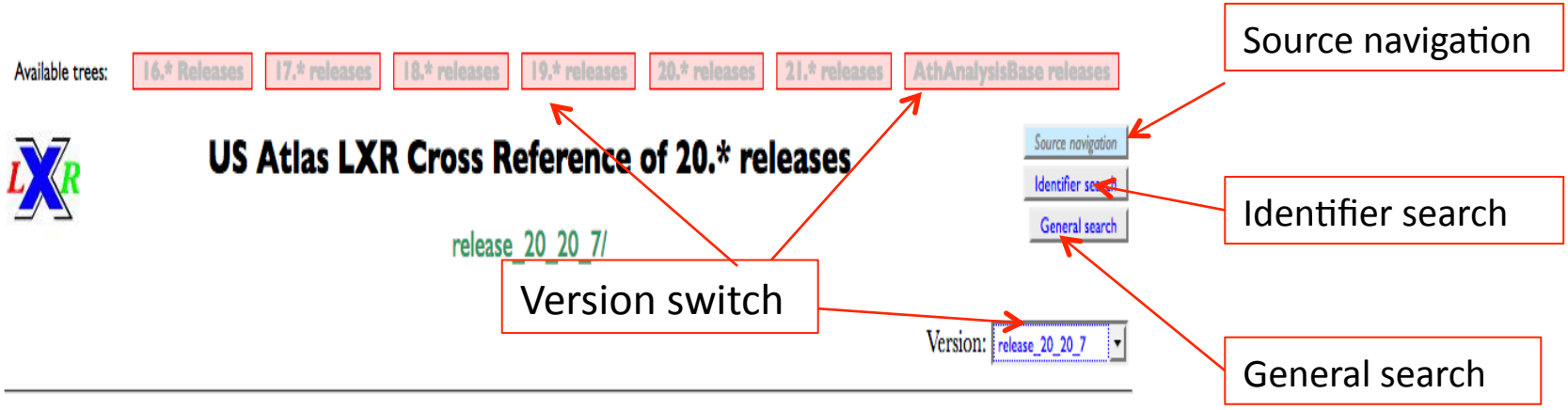
Nightlies releases are indexed everyday, the indexing takes a few hours.

To provide non-interruption service, we introduced symlink switch pending on the indexing job success.

Atlas LXR: Tree-like Version Change

Convenient version change;

Browse source, search for identifier & keyword



	Name	Size	Date (UTC)	Last indexed	Description
	GAUDI_v26r2p3-lcg81f/	-	2016-11-23 21:33:02		
	atlas/	-	2016-11-23 21:32:27		
	external/	-	2016-11-23 21:33:08		
	Name	Size	Date (UTC)	Last indexed	Description

[Source navigation] [Identifier search] [general search]

This page was automatically generated by the 1.2.0 LXR engine
on ATLAS STABLE RELEASES LXR server.

Atlas LXR: General Search

General search is conducted by glimpse.

Files named: ☐ Advanced (allows usage of perl regex)

Or containing: GetBranch

Characters ^ $ \ [ ] ( ) | , ; ~ have special meaning for Glimpse and can be used for building simple regex.

☐ Case-sensitive Search

Powered by Glimpse. (Tips for search syntax.)

---

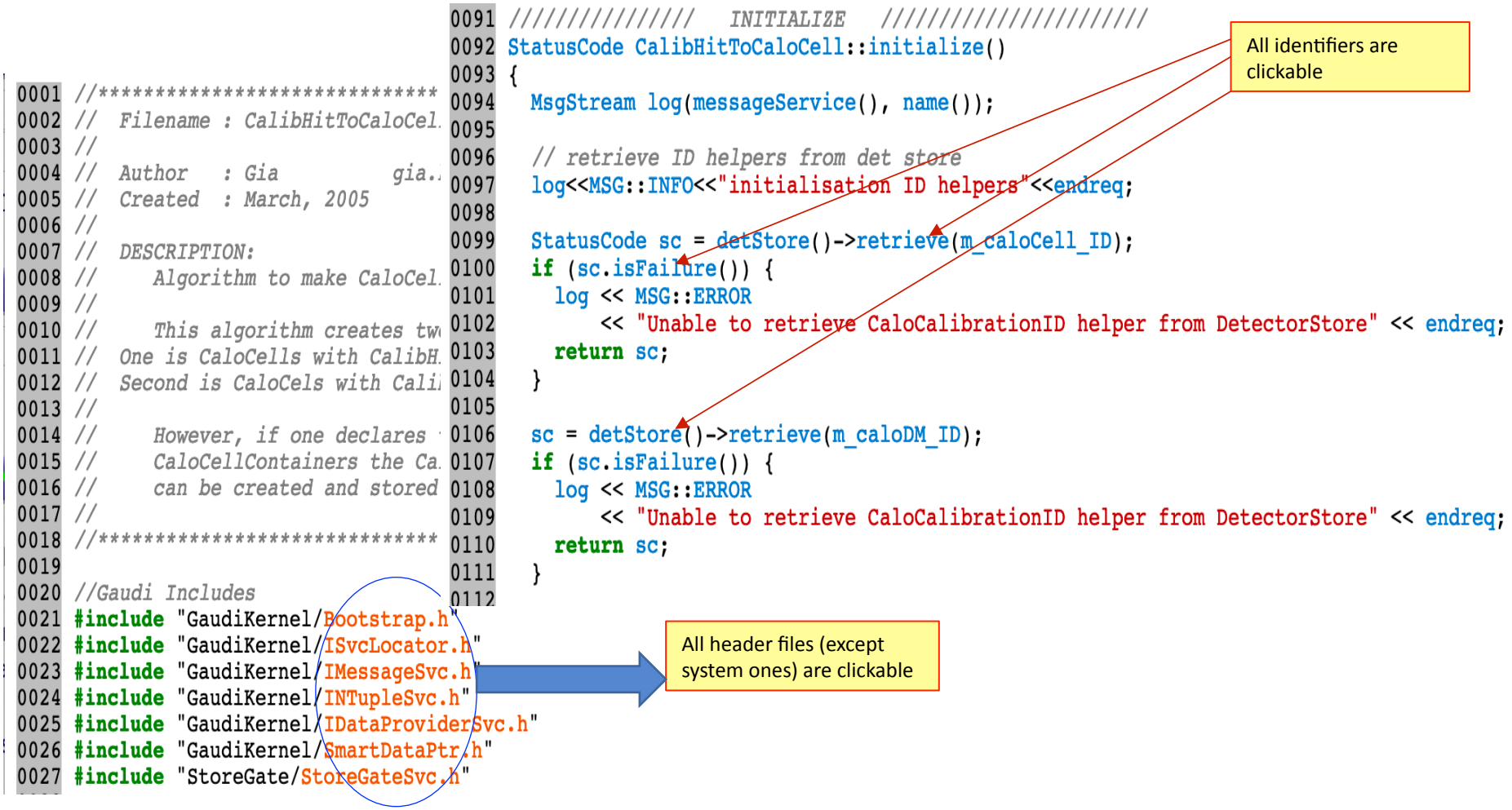
979 occurences found.

Results for GetBranch

File	Line	Text
Gaudi/RootCnv/RootCnv /RootDataConnection.h	220	`virtual TBranch* getBranch( std::string_view section, std::string_view n ) = 0;`
	303	`TBranch* getBranch( std::string_view section, std::string_view branch_name ) {`
	304	`return m_tool->getBranch( section, branch_name );`
	307	`TBranch* getBranch( std::string_view section, std::string_view branch_name, TClass* cl, void* ptr, int buff_siz,`
Gaudi/RootCnv/merge/extractEvt.C	265	`br_in = m_ref_in->GetBranch( "Databases" );`
	267	`br_out = m_ref_out->GetBranch( "Databases" );`
	275	`br_in = m_ref_in->GetBranch( "Containers" );`
	277	`br_out = m_ref_out->GetBranch( "Containers" );`
	284	`br_in = m_ref_in->GetBranch( "Links" );`
	286	`br_out = m_ref_out->GetBranch( "Links" );`
	293	`br_in = m_ref_in->GetBranch( "Params" );`
	295	`br_out = m_ref_out->GetBranch( "Params" );`
	324	`br_out = m_evt_out->GetBranch( name );`

LXR: Clickable Identifiers & Header Files

All identifiers and header files (except system ones) are clickable.



AtlasSetup: Release Env Setup

Atlas release env is set up by the **AtlasSetup** tool (via **asetup** command).

The purpose of **asetup** is to help users to find Atlas releases and set up all the necessary env in a **convenient** and **fast** way. And the users need not know much about the release location.

It provides:

- For users, look up the location of a release, and set up the required compiler and cmake, then the release env. Something like:
  asetup Athena,22.0.0 # where project=Athena, release=22.0.0
- For release builders, set up the specified compiler (gcc/clang) and cmake. Something like:
  asetup gcc8,cmakesetup,none

The setup is fast: usually takes **4s~7s**.
It is widely used, from grid(Panda)/batch to interactive machines, from T0 to T3.

Atlas Releases

Each Atlas release is composed of hundreds of packages. For example of Athena,21.5.8, it requires 2053 (sub-)packages (internal only), contains 2568 shared libs.

Atlas releases used to be managed by CMT tool (also used by LHCb), and were moved to use cmake in 2018.

- **CMT**: Configuration Management Tool, a home grown HEP tool, developed by Christian Arnault (LAL), development stopped.
- **cmake**: open source, widely used, a tool to manage the build process of software using a compiler-independent method

And there are 2 types of releases, usually distributed on cvmfs:

- Stable releases: release name is numeric like *22.0.1* or *21.2.62.0*.
- Nightlies releases: release name is timestamp like *2019-09-04T2133* (cmake) or *rel_3* (cmt).

To locate a specific release, it requires a project name (such as *Athena* or *AtlasOffline*), a release name, a branch name (nightlies releases), and platform (x86_64-slc6-gcc8-opt).

asetup Migration from CMT to cmake

Because the old cmt asetup, designed for cmt releases, is complicated, unnecessary for cmake releases, and difficult to maintain/improve.

Thus a **new simplified asetup** was **re-designed** for cmake releases.

Here are some comparison numbers: cmake vs cmt

- 9 (17) files, 8 (20) classes, 94 (224) functions,

  4.2K (**15K**) lines (excluding tests)

- Compiler setup: 1 (3) function, **129** (**850**) lines,
- Release locator: 1 (7) function, **208** (**1350**) lines.
- Full option help printout: 176 (410) lines; **210 options removed** in cmake asetup.

The above comparison demonstrates:
how complicated in the old asetup to locate a release and to set up a compiler.

asetup Deployment

Because asetup is **widely used**, we must be very careful on a new version deployment:

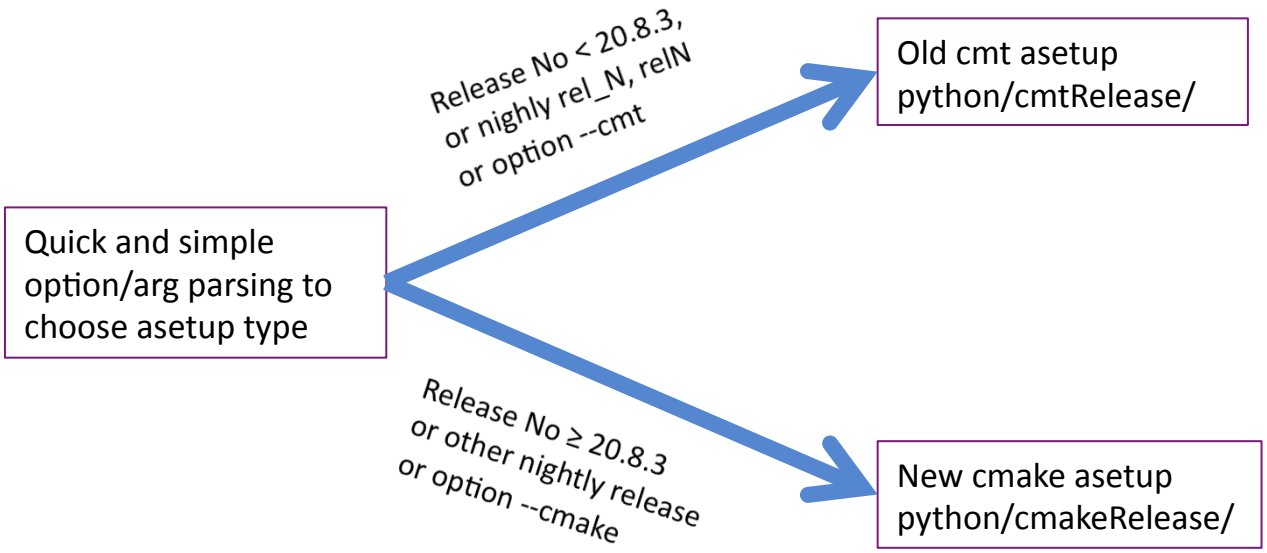
- StressTest (composed of >70 different tests) under native SL7,
  SL6/SL7 containers (many grid jobs run within containers).
  and comparison with previous version.
- Tests on ALRB
- Also HammerCloud tests during cmt=>cmake migration

A new asetup version is usually deployed as a beta (testing) version first.

Then it will become a production version if no problem is found.

Determination of Release Type

The old cmt releases will co-exist with new cmake releases for a while, and asetup is widely used.
Make the change **transparent** to users/panda as much as possible.



Then pass the remaining option/arg to cmt/cmake asetup.
Issue warning msg for **258** old deprecated **tags** and **210** deprecated **long options**.

asetup Command Options

There are hundreds of options in cmt asetup, and about 60 in cmake asetup.

The help printout with all options would be very len…gthy, and not helpful.

Solution: divide the options into a few groups and print live help per group or brief help fitting on one screen.

% asetup –h

```
Usage: asetup [options] [tags]

Options:
  -h, --help            show this help message and exit
  --version             Print out the asetup version
  --printLast, --printLastSession
                        Print out the last saved session info under PWD
  --listRemovedTags     Print out the list of removed old asetup tags
  --listRemovedOpts, --listRemovedOptions
                        Print out the list of removed old asetup options
  --releasebase=<dir>, --releasepath=<dir>
                        Specify an exact path to the release bypassing the area search
  --helpMoreOn=<group>  The group name for help printout, valid groups: [CMAKE,
                        RELLOC,ASETUP,DB,PLATFORM,RELNAME,ENVVAR,COMPILER,ALL],
                        where 'All' will print all groups options

Examples: asetup Athena,22.0.0         # set up stable release 22.0.0
          asetup Athena,master,latest  # set up the latest master nightlies
          asetup Athena,master,r08     # set up the master nightlies of date 08

 or asetup source otherScript    # source additional user script

For further details, look at
  https://twiki.cern.ch/twiki/bin/view/AtlasComputing/AtlasSetupReference
```

asetup Configuration and Debug

There are multiple layers of configuration for sites/users:

- ✓ Default configuration in asetup
- ✓ $AtlasSetupSiteCMake, $AtlasSetupSite
- ✓ $AtlasSetup/../asetupSite/.asetup
- ✓ $HOME/.asetup, $PWD/.asetup
- ✓ Options --siteconfigfile=, --inputfile=

And one option --dumpconfig to print out all configuration parameters together with their source origin.

There are 2 options for debugging purpse:

- ✓ --debugprint: Indented output and with both function and caller name (for cmake asetup, automatical name lookup)
- ✓ --simulate: keep and do not source the generated temporary shell.

Convenient Ways to Use asetup

It is not very convenient to use options, like:

Asetup --project=Athena --release=22.0.0 --os=slc6 --gccversion=6.2

With the automatically translated **tags** from most long options, you can run instead:

asetup Athena,22.0.0,slc6,gcc62

Fast/convenient to setup the **latest** nightlies release, and shortcut to timestamp nightly release (using prefix-**r**, says r04 for 2019-09-04T*, otherwise it would be difficult to remember/look up, there are ~100 nightlies under branch master/)

asetup master,Athena,latest

*asetup master,Athena,**r04***

Restore the last session env under one work directory, just run without arg or option --restore

It would print out the following msg and set up the previously used release under that directory:

% **asetup**

```
Set up the following release as last used under the current directory
Using Athena/22.0.0 [cmake] with platform x86_64-slc6-gcc62-opt
        at /cvmfs/atlas.cern.ch/repo/sw/software/22.0
```

An option to print out the last saved release info:

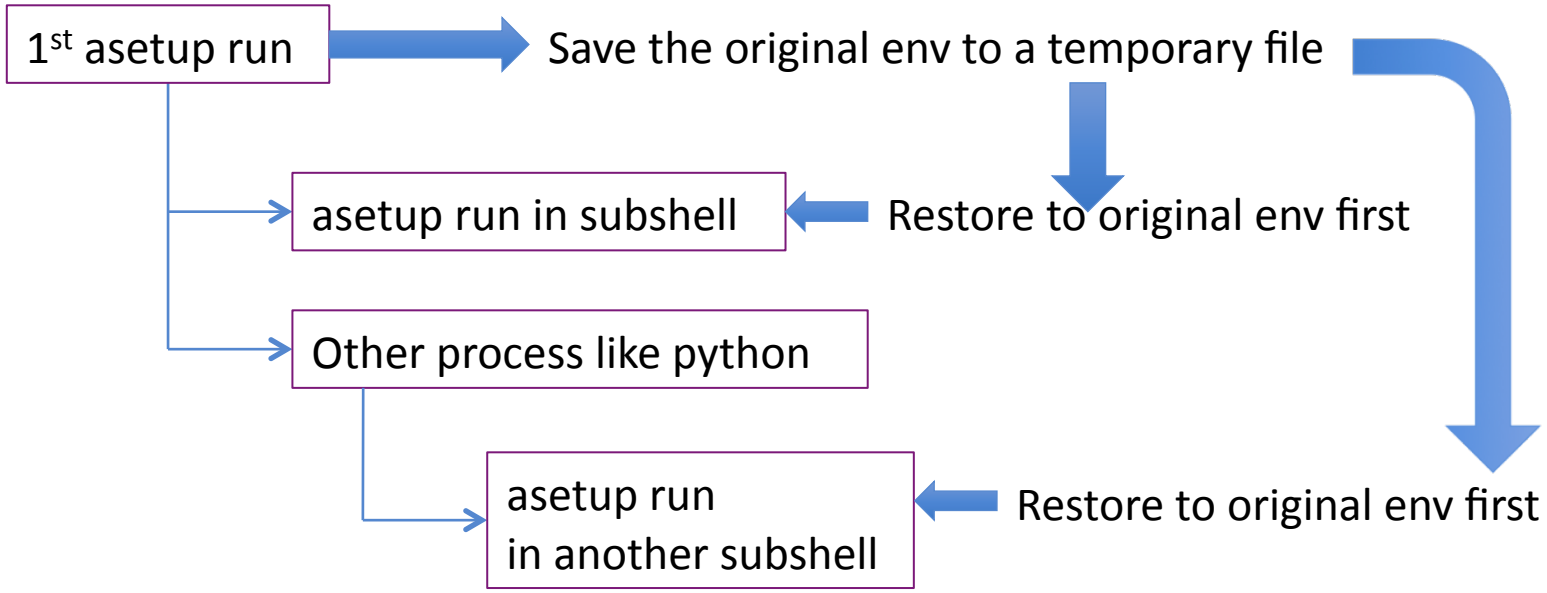
% **asetup --printLast**

```
The last saved release under the current directory is:
        Athena/22.0.0 [cmake] with platform x86_64-slc6-gcc62-opt
        at /cvmfs/atlas.cern.ch/repo/sw/software/22.0

Afterwards, the following epilog script(s) will also be sourced
        /cvmfs/atlas.cern.ch/repo/ATLASLocalRootBase/swConfig/asetup/asetupEpilog.sh
```

Shell Env Restoration in asetup

Sourcing the **new** release setup script **on** an **existing** release env will mess up the env.

Solution: save the shell env at the 1st time asetup run.



The subshell can be any levels below the shell of the 1st asetup run.

There is also a mechanism to clean up those saved env files during each new asetup run.

Resolving Conflicts b/w Releases & System

Due to alternate system libraries, some Atlas releases break system commands such as emacs, git or expr.

For example, new manipulation instruction set in libgmp.so is not available in some old CPUs, breaking the command expr.

asetup will check those conflict, and provides alias/wrapper with no/little performance loss:

alias expr='/lib64/ld-linux-x86-64.so.2 --library-path /usr/lib64 /bin/expr'

Wrapper need varies from release to release, is generated on the fly, and its path is added in $PATH for usage in subshell.

The wrapper in work directory is also cleaned up automatically.

In addition, if users have already aliased the command, the re-aliased command will carry over user option, says,

% alias emacs="emacs -nw"

% alias_sys_exe emacs

% alias emacs
emacs='/lib64/ld-linux-x86-64.so.2 --library-path /usr/lib64 /bin/emacs -nw'

Where alias_sys_exe is introduced in asetup, and aliasing emacs is done automatically in release env setup if necessary.

USAtlas Shared Tier-3s at BNL/SLAC

There are 2 shared Tier-3s for USAtlas users:

- BNL (introduced first)
- SLAC (added later)

There are a lot of CPU/Space available at both shared Tier-3 sites:

	BNL	SLAC
Interactive-CPU	8 machines (8 cores each)	88 cores
Batch-CPU	~2300 slots (25K in shared pool)	~3000 cores
Private space	20GB (home), 500GB (data), 5TB (dCache)	100 GB (home), 2TB (data, could increase up to 10TB)
Shared space	Rucio RSEs (localgroupdisk)	Rucio RSEs (SLACXRD*)
Remote X-Window	NoMachine NX	FastX
Batch system	Condor	LSF

Usage at Shared Tier-3's

There are many users using the shared Tier-3 sites:

	BNL	SLAC
Users	130 (30 institutes)	50+

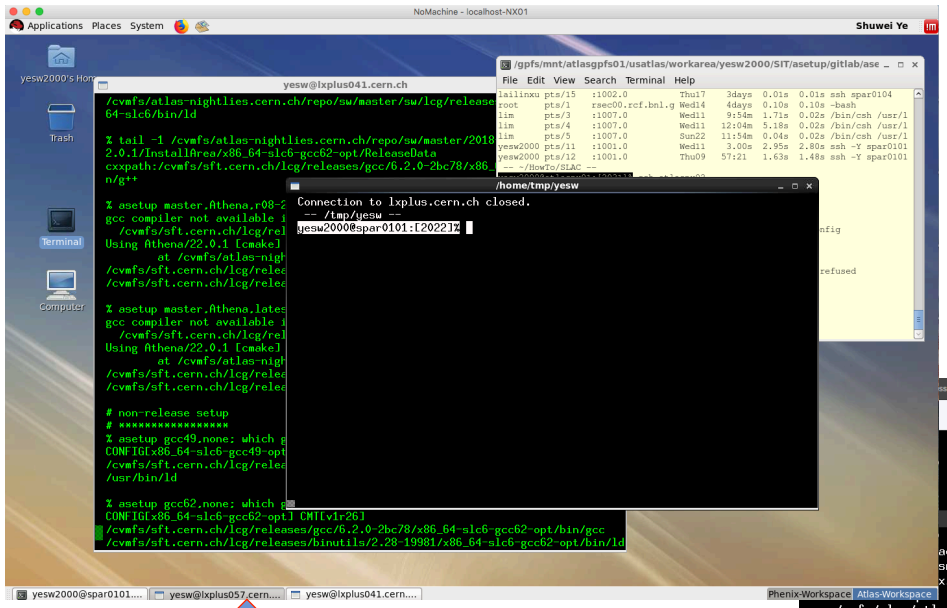
Space/batch usage at BNL:

	Users
Private space (20GB)	64 (>50% usage)
Data space (500GB)	14 (>50% usage)
dCache space (5TB)	30 (>1TB)

Batch	No (users)	No (institutes)	No (jobs)	Ave time
Last 6 months	50	15	2.84M	2800s
Last 12 months	57	16	7.96M	2137s

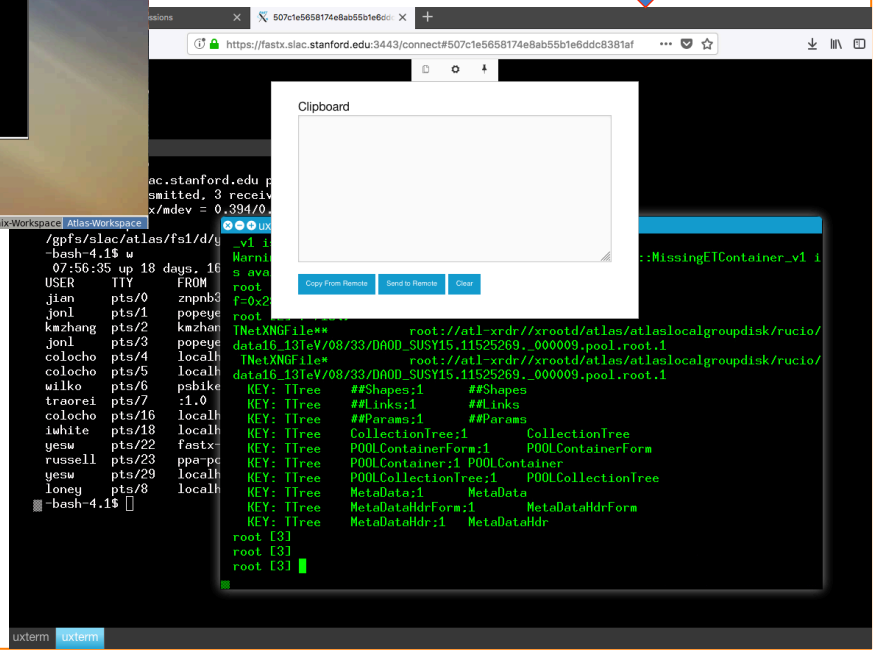
Remote X-Window at BNL/SLAC

BNL/SLAC provide different Remote X Window



NoMachine at BNL

FastX at SLAC



dCache at BNL: Dataset Listing

Developed a tool to generate clist file for both official/private datasets on dCache.

% **pnfs_ls.py –h**

```
Usage:
     pnfs_ls.py [options] dsetListFile
  or
     pnfs_ls.py [options] dsetNamePattern[,dsetNamePattern2[,more namePatterns]]
  or
     pnfs_ls.py -o clistFilename /pnfs/FilePathPattern [morePaths]
  or
     pnfs_ls.py -p -o clistFilename [pnfsFilePath | pnfsDirPath] [morePaths]

  This script generates pfn (physical file name), pnfs-path,
or xrootd-path of files on BNL dcache for given datasets or files on PNFS,
where wildcard and symlink are supported in pnfsFilePath and pnfsDirPath

  Options:
  -h, --help            show this help message and exit
  -v                    Verbose
  -V, --version         print my version
  -p, --privateFiles    List private non-dataset files on dCache
  -i, --incomplete      Use incomplete sites if complete not available
  -u, --usersDCache     Use datasets under users private dCache
  -l, --listOnly        list only matched datasets under users dCache, no pfn output
  -o OUTPFNFILE, --outPfnFile=OUTPFNFILE
                        write pfn list into a file instead of printing to the screen
  -d OUTPFNDIR, --dirForPfn=OUTPFNDIR
                        write pfn list into a directory with a file per dataset
  -N, --usePNFS         using pNFS access, default is xrootd within BNL
  --useXRootdOutside    using xroot from outside BNL: access, default is
                        xrootd within BNL
  -L LOCALBNLSITE, --localBNLSite=LOCALBNLSITE
                        specify a BNL site, overriding the one choosen by the script
```

Hiro developed a tool rucio-bnl-get for fast dataset replication onto private dCache.

Users had problem with this tool and dCache access from to time.

Copying Files/Dirs to dCache at BNL

Developed a tool to help copy files to BNL dCache in multiple threads.

% **pnfs-copy.py –h**

```
Usage: pnfs-copy.py [options] files pnfs_directory
   pnfs-copy.py [options] dirs pnfs_directory

    or using quote-enclosed wildcard such as
       pnfs-copy.py [options] "*.root" pnfs_directory

   This script uses gfal-copy to conduct the file copy, thus requires the
  setup of RucioClients and a valid grid proxy.

Options:
  -h, --help  show this help message and exit
  --verbose   Print verbose info
  --version   Print the script version then exit
```

On default it starts 3 threads with 3 retries, and will report the copy status: numbers of successfully copied files and failed files after 4 tries.

Survey on SLAC T3 Users & Improvements

Since SLAC T3 was just added, we conducted a 19-question survey.

- There are 8 responses in total
- User support is very good
- Enough space to most users
- 6/8 could use SLAC as local T3 alternative, and would recommend to collaborators
- Quick batch jobs start
- Documentation needs improvement
- Most users have never used FastX
- Batch job failure rate is not very low, but many of them should be able to be fixed easily

We (Wei & I) have improved the documentation (easy to follow). SLAC would think about FastX and possible replacement.

I plan to development a tool similar to pnfs_ls.py for SLAC.

Xcache at SLAC/BNL

Xcache enables to access data remotely and also to cache them (**fully** or **partially**) locally for faster access in future.

SLAC/BNL provide the Xcache server to try:

- SLAC Xcache server: **root://atlfax.slac.stanford.edu/**
- BNL Xcache server: **root://xrootd03.usatlas.bnl.gov:1094/**

There are two ways of Xcache:

1. Specify both Xcache server and remote xrootd server, hence contain **TWO root:/**

root://atlfax.slac.stanford.edu//**root:**//dcgftp.usatlas.bnl.gov:1094/pnfs/usatlas.bnl.gov/LOCALGROUPDISK/rucio/data18_13TeV/da/ea/DAOD_EXOT12.14278917._000001.pool.root.1

root://xrootd03.usatlas.bnl.gov:1094//**root:**//griddev03.slac.stanford.edu:2094//xrootd/atlas/atlaslocalgroupdisk/rucio/data16_13TeV/f9/bd/DAOD_SUSY15.11525262._000003.pool.root.1

2. gLFN (global Logical File Name) access, **without knowing file location**

root://atlfax.slac.stanford.edu/**/atlas/rucio/***data18_13TeV:DAOD_EXOT12.14278917._000001.pool.root.1*

root://xrootd03.usatlas.bnl.gov:1094/**/atlas/rucio/***data16_13TeV:DAOD_SUSY15.11525262._000003.pool.root.1*

We need test and study the performance.

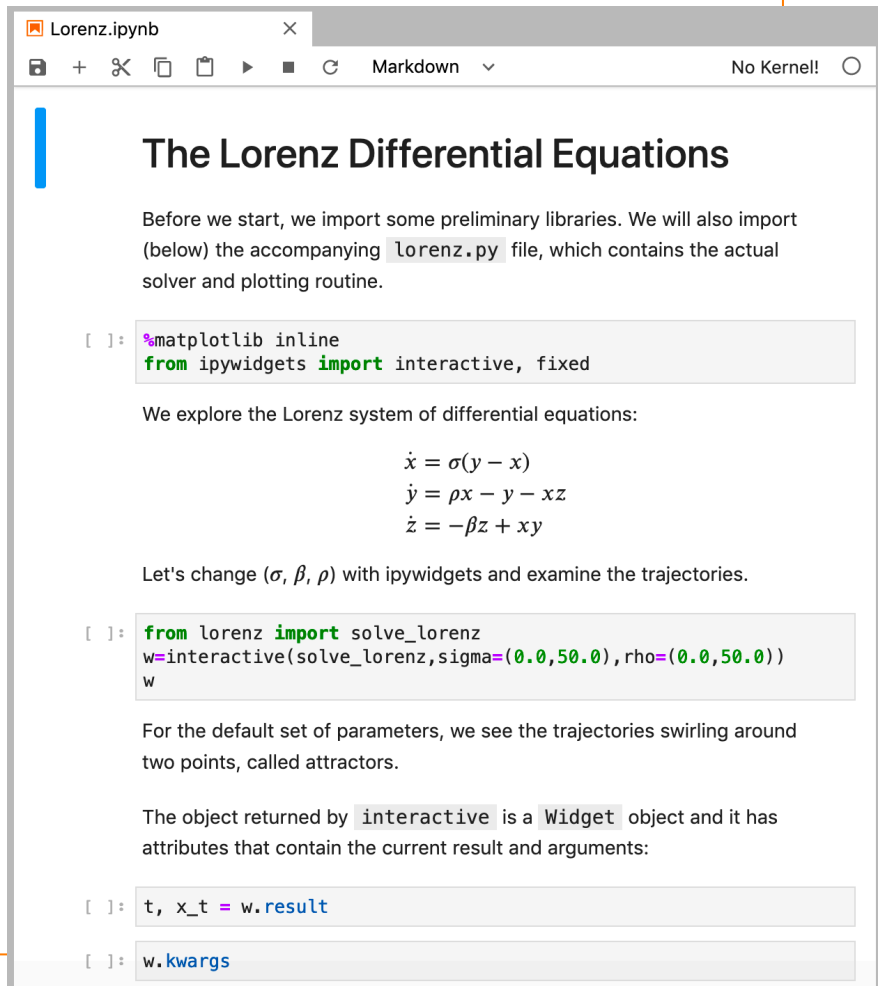
Jupyter at BNL/SLAC

BNL/SLAC provide experimental Jupyter Notebook service (powerful w/ GPUs).

Jupyter Notebook provides the following features:

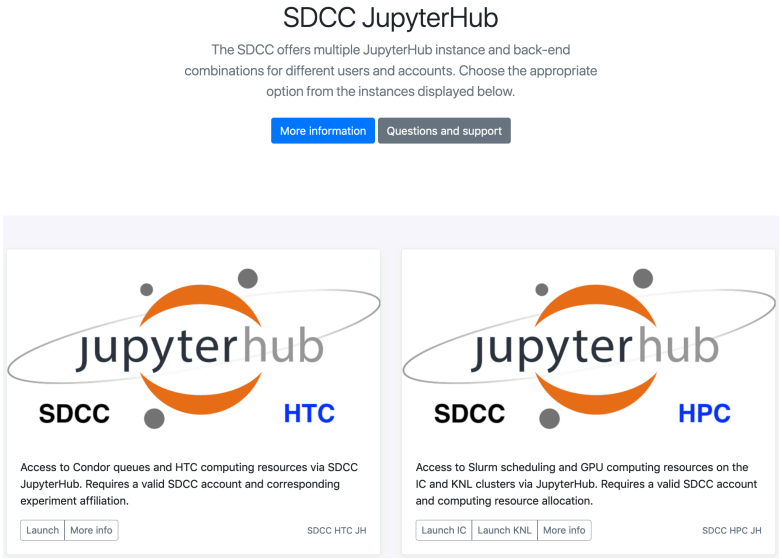
- Interactive on a web client (remotely or locally)
- Explanatory rich text (shown on the right)
- Intersperse code with its result (charts/figures)
- Instant result after code change
- Work stored on a notebook document
- Easy to share/reproduce
- Able to convert to other formats (script, PDF, latex, html, …)

Available kernels are python (default), R, Ruby, Julia, etc.



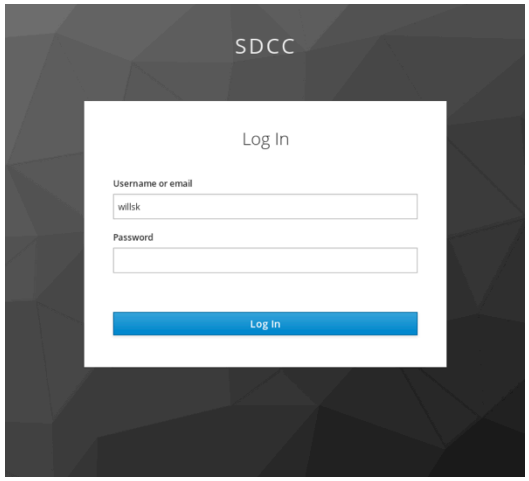
Logging In to BNL Jupyter

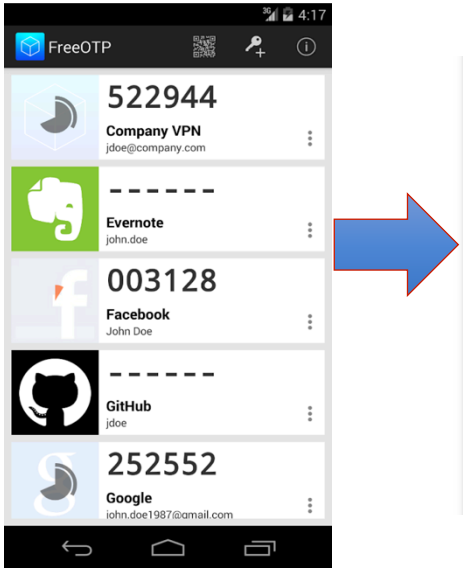
Visit Webpage: https://jupyter.sdcc.bnl.gov/
providing 2 types: HTC and HPC (GPU)

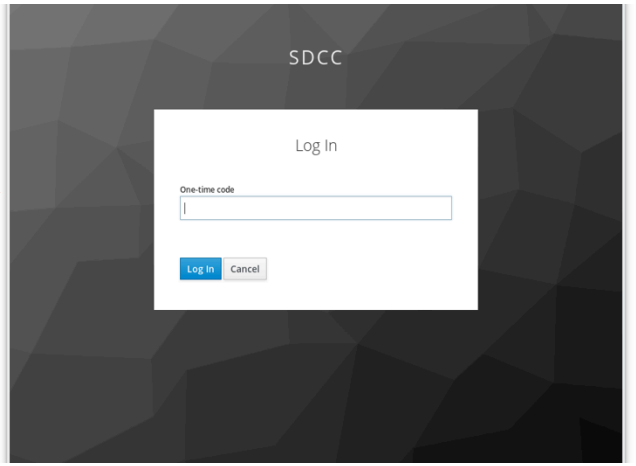


Two-step authentication:

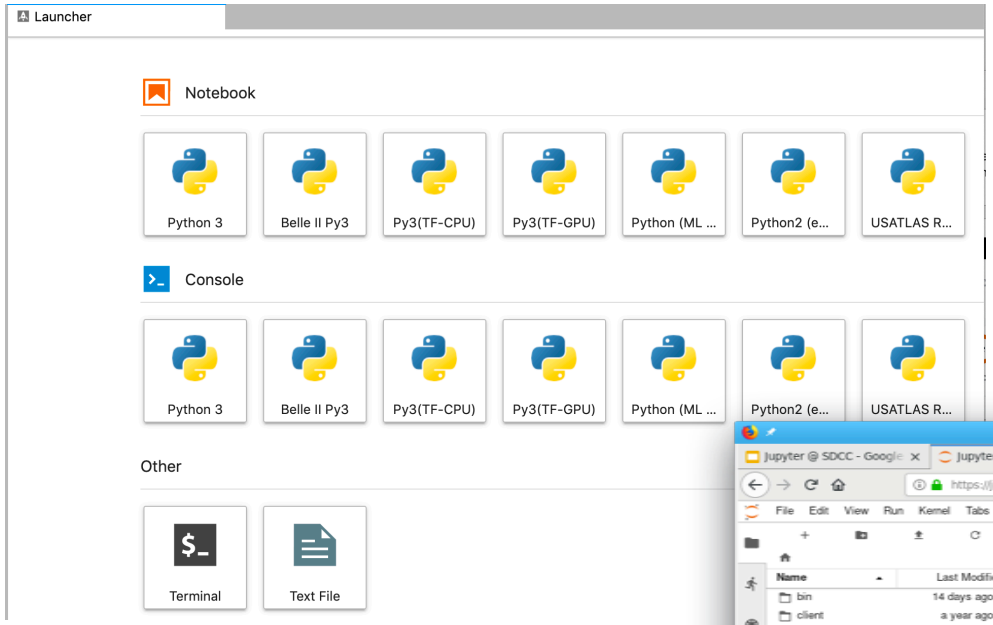
1. SDCC account login
2. QR Code → Google Authenticator or FreeOTP app
   Easy setup by scanning QR code first time



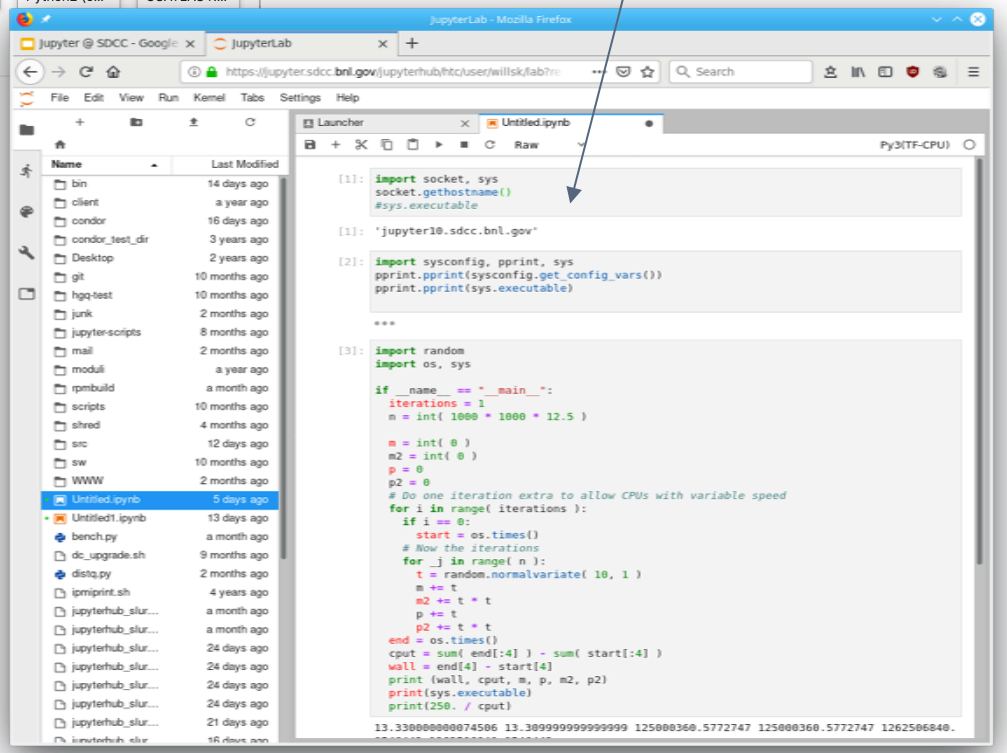




Jupyter at BNL



Notebook Files



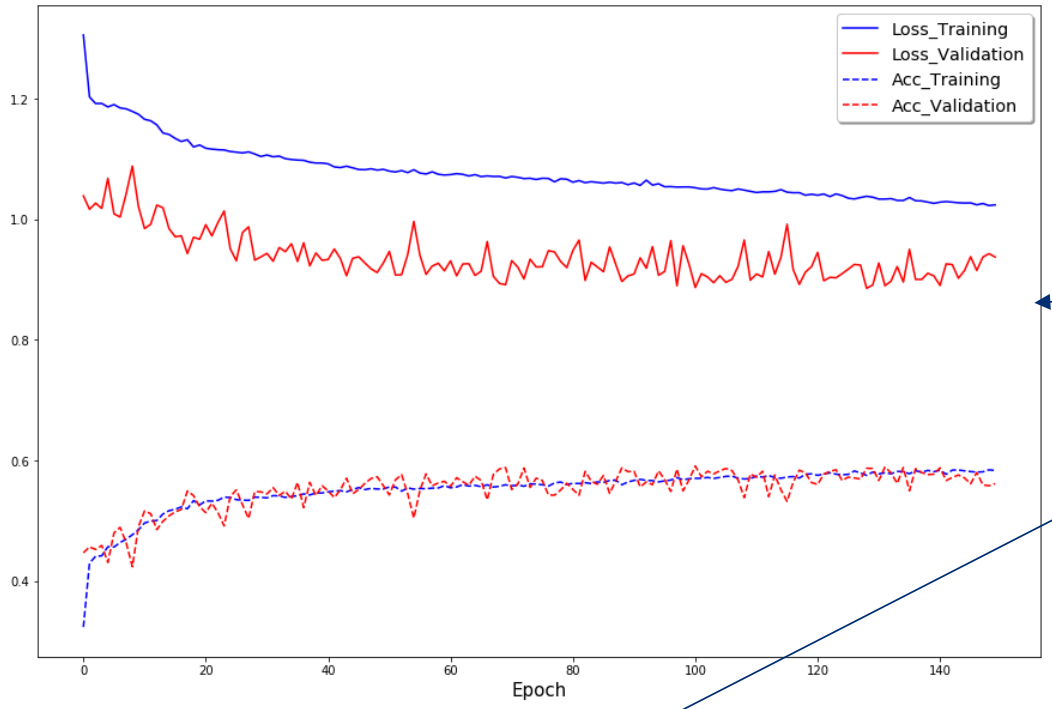
Some kernels are provided

Jupyter Example at BNL

SoftMax_Check.ipynb

Code | Python 3

```
[25]: plt.figure(figsize=(15,10))
loss_history=history.history["loss"]
vloss_history=history.history["val_loss"]
acc_history=history.history["acc"]
vacc_history=history.history["val_acc"]
plt.plot(range(len(loss_history)),loss_history,'b', label="Loss_Training")
plt.plot(range(len(vloss_history)),vloss_history,'r', label="Loss_Validation")
plt.plot(range(len(acc_history)),acc_history,'b--', label="Acc_Training")
plt.plot(range(len(vacc_history)),vacc_history,'r--', label="Acc_Validation")
plt.xlabel('Epoch', fontsize=15)
plt.legend(loc='upper right', shadow=True, fontsize='x-large')
plt.show()
```



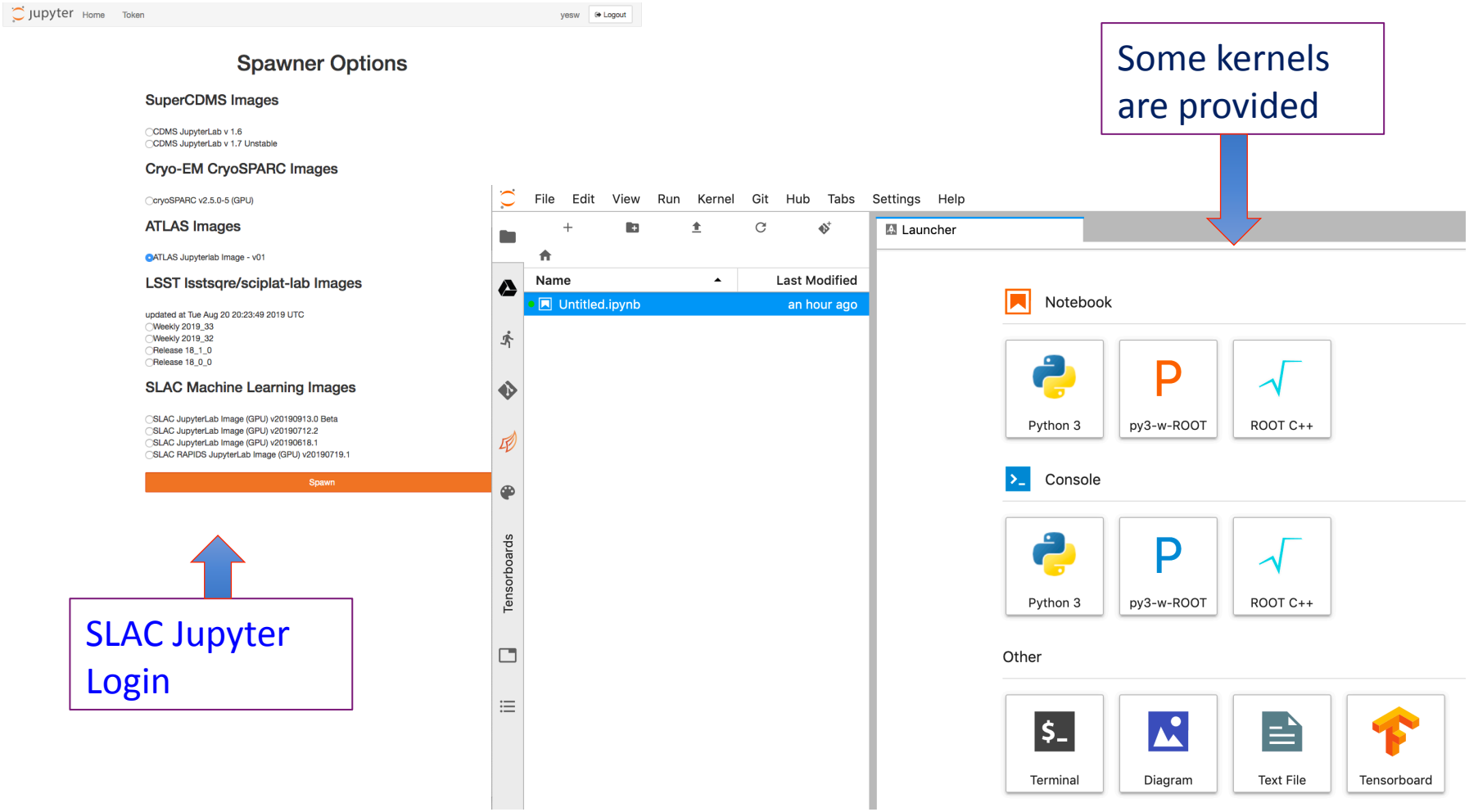

```
[26]: model.predict(X_test)[0:10]

[26]: array([[0.9464846 , 0.02703056, 0.02648488],
             [0.13972566, 0.13177393, 0.7285004 ],
```

Code input

Result right after code

Jupyter at SLAC

SLAC Jupyter Login: https://jupyter.slac.stanford.edu



Jupyter Example at SLAC