Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules

LSST data release processing at CC-IN2P3

status and perspectives

fabio hernandez

Journées LSST-France, Clermont-Ferrand, June 5th, 2019

PEOPLE

- Bastien GOUNON joined LSST team at CC-IN2P3 in December 2018
- Currently working on distributed data management tools

IN2P3

A DISTRIBUTED LABORATORY

in2p3.fr

2500 researchers, engineers and technicians

700 post-docs and PhD students

25 laboratories and research platforms in France, 16 international laboratories

COMPUTING CENTER

IN2P3 COMPUTING CENTER

- CC-IN2P3

  84 people, 80 FTE, 80% permanent positions

  ~15 M€ overall annual budget

  scientific data center, high throughput computing

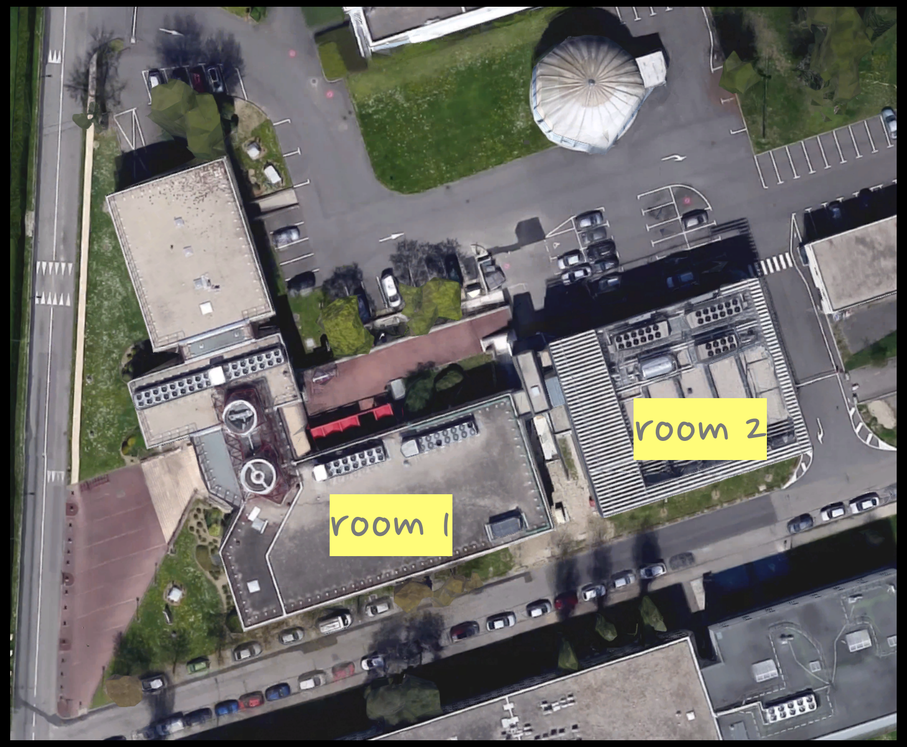
  well connected to national and international networks

- **Shared computing facility** supporting the institute's research program

  ~70 projects in high energy physics, nuclear physics and astroparticle physics

- Operations: 24x7

  unattended during nights and weekends



aerial view

building 2

machine room 2

machine room 1

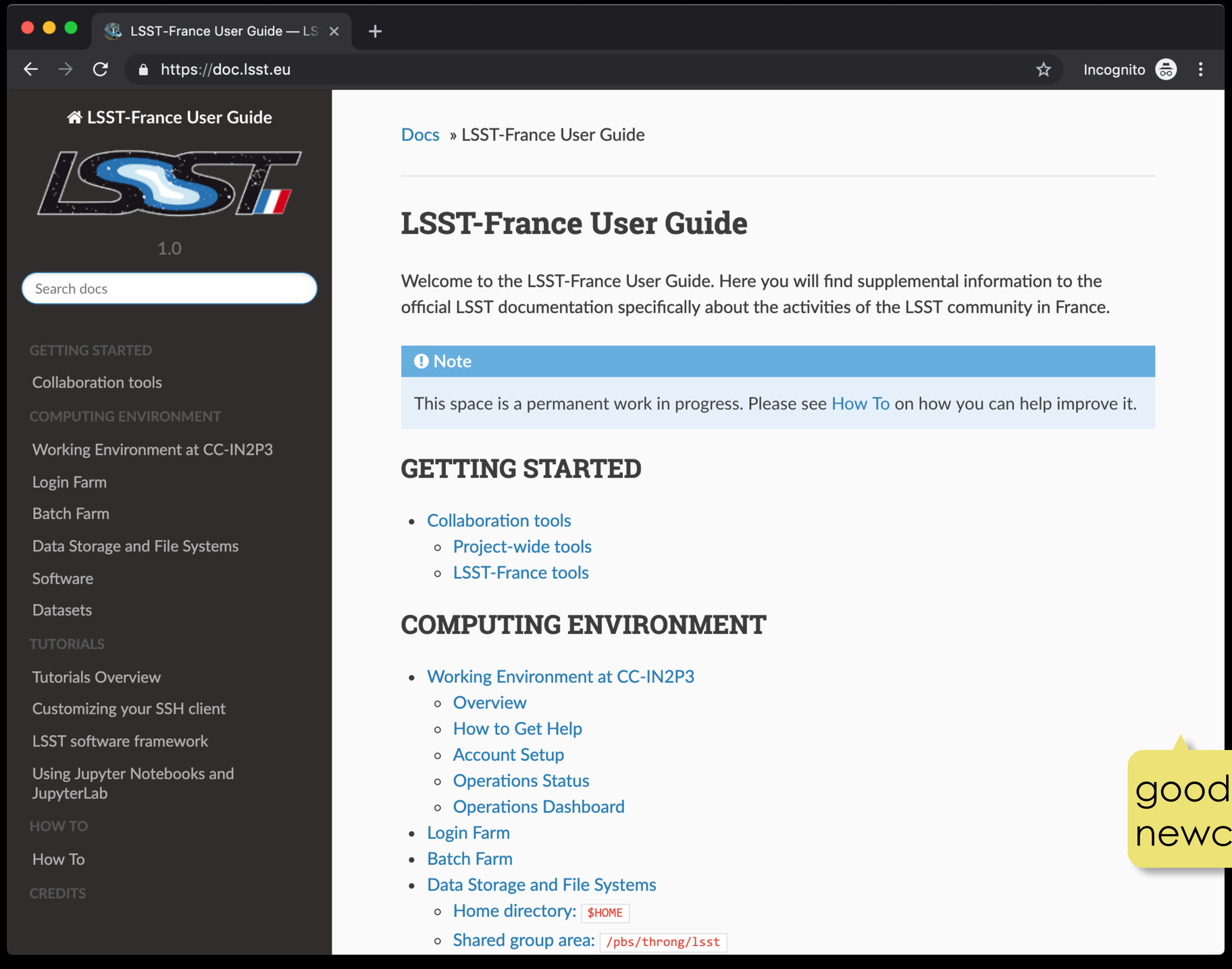
CPU: **34k CPU cores, ~900 nodes**
Disk: **36 PB**
Tape: **64 PB**

2 machine rooms, 1600 m²

click here for a virtual visit

© Cyril FRESILLON / CC IN2P3 / CNRS Photothèque

doc.lsst.eu

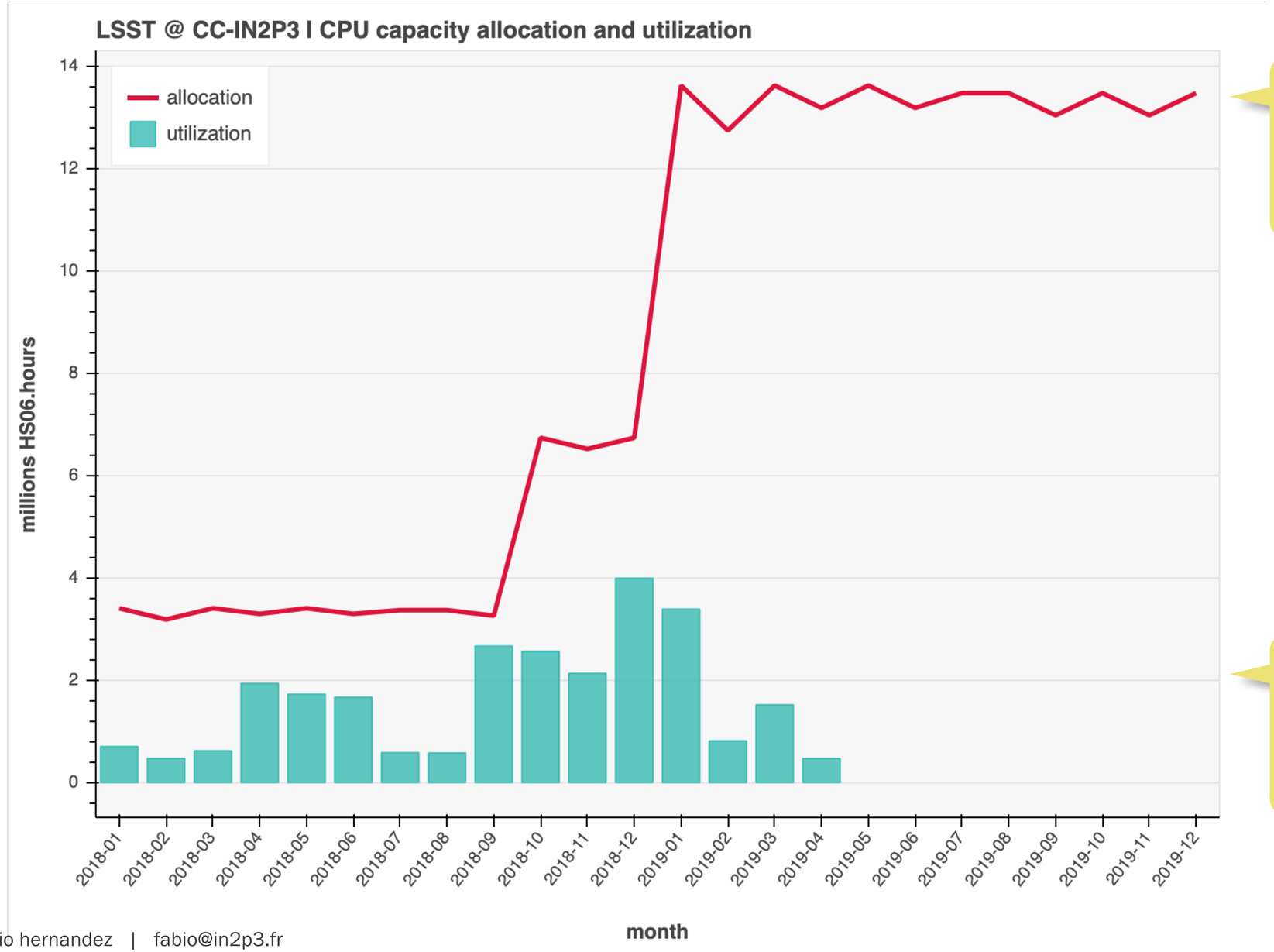

LSST-France User Guide — LS
https://doc.lsst.eu
Incognito
LSST-France User Guide
1.0
Search docs
GETTING STARTED
Collaboration tools
COMPUTING ENVIRONMENT
Working Environment at CC-IN2P3
Login Farm
Batch Farm
Data Storage and File Systems
Software
Datasets
TUTORIALS
Tutorials Overview
Customizing your SSH client
LSST software framework
Using Jupyter Notebooks and JupyterLab
HOW TO
How To
CREDITS
Docs » LSST-France User Guide
LSST-France User Guide
Welcome to the LSST-France User Guide. Here you will find supplemental information to the official LSST documentation specifically about the activities of the LSST community in France.
Note
This space is a permanent work in progress. Please see How To on how you can help improve it.
GETTING STARTED
Collaboration tools
Project-wide tools
LSST-France tools
COMPUTING ENVIRONMENT
Working Environment at CC-IN2P3
Overview
How to Get Help
Account Setup
Operations Status
Operations Dashboard
Login Farm
Batch Farm
Data Storage and File Systems
Home directory: $HOME
Shared group area: /pbs/throng/lsst
good place to start for newcomers

You are kindly invited to subscribe to this Slack channel
(please click on the link below)

#in2p3

There are other channels of potential
interest for LSST-France members.
All of them start by `#in2p3-*`

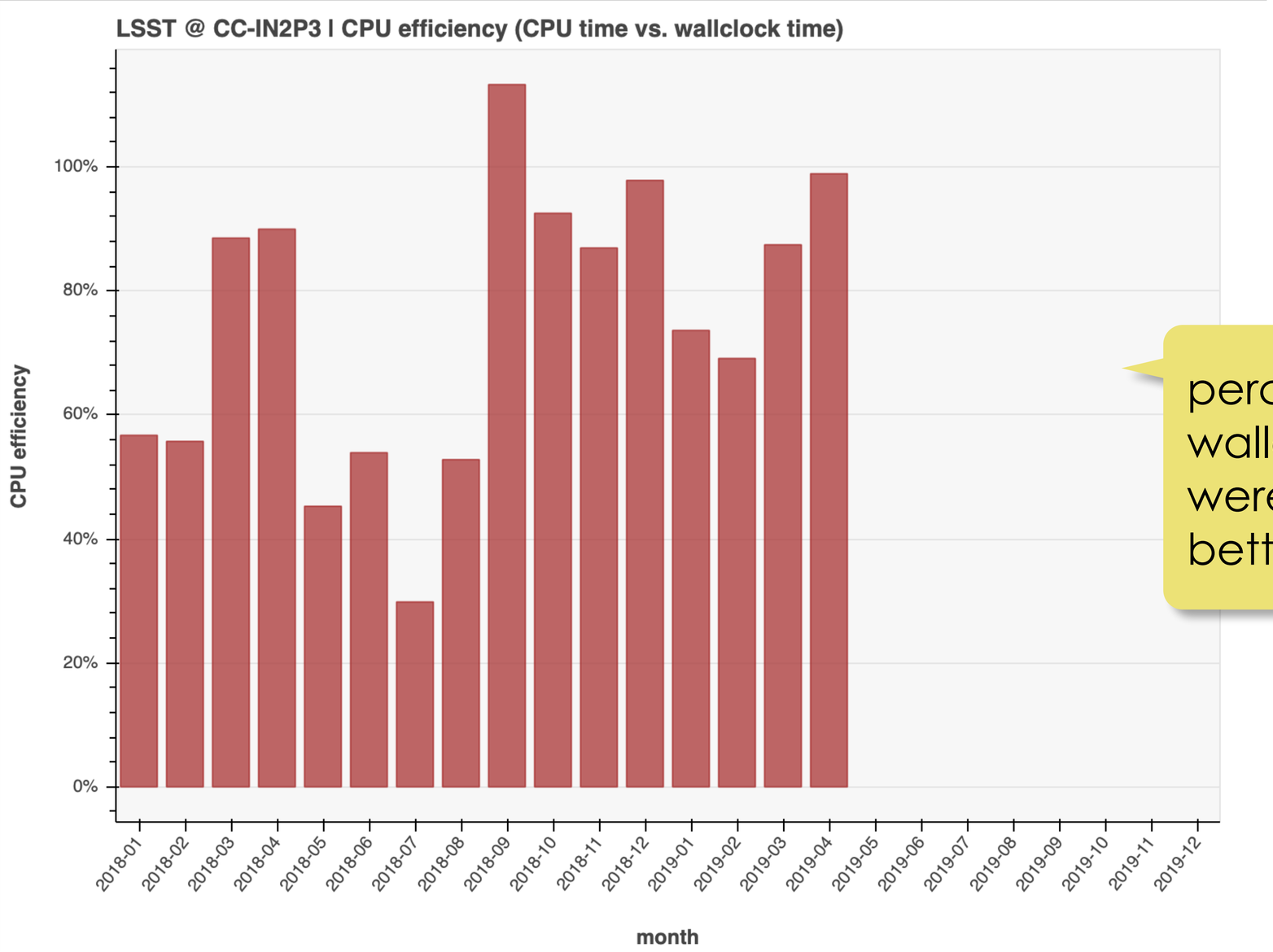
CPU CAPACITY



allocation is equivalent to 1600 CPU cores

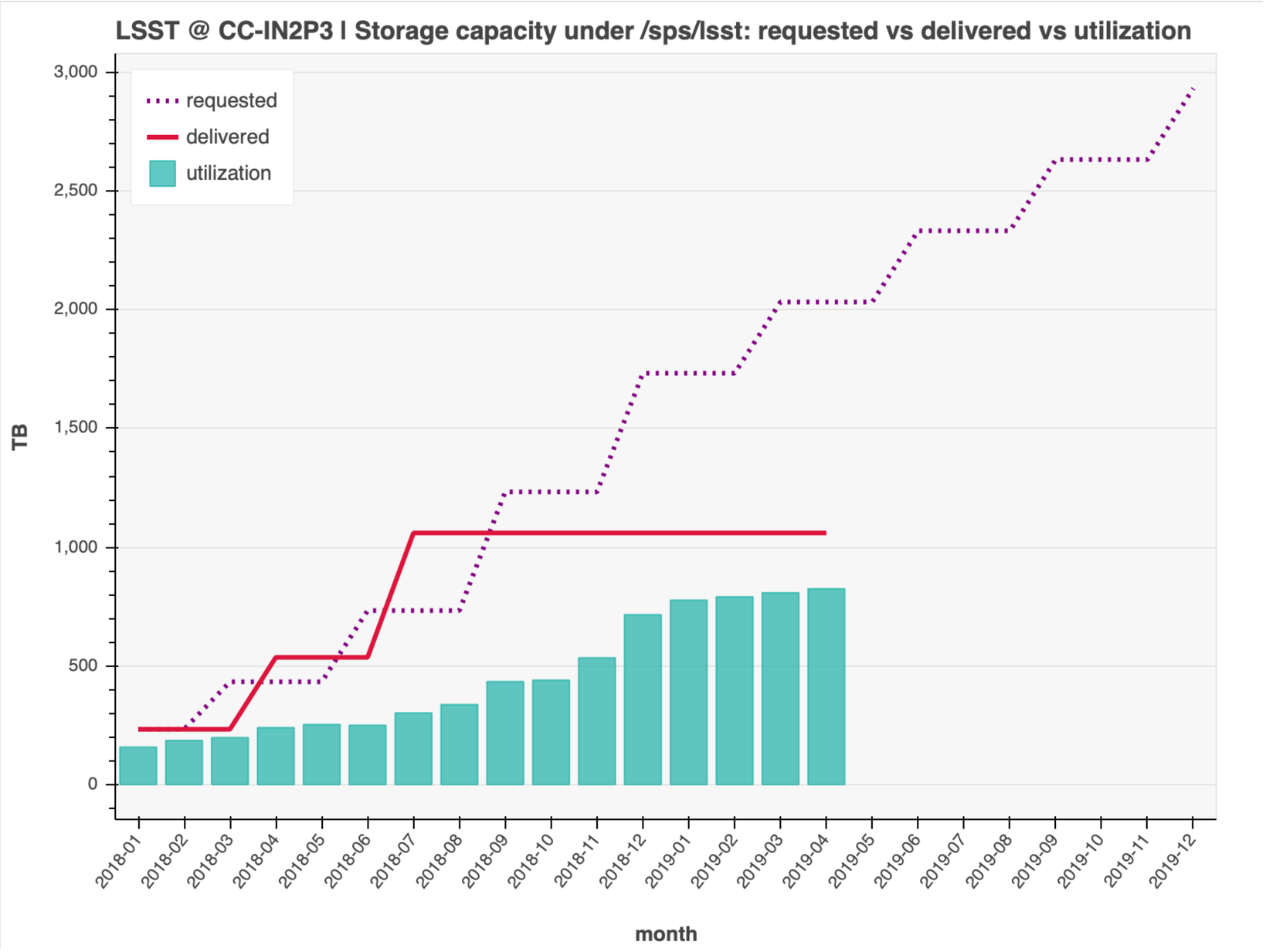
utilization of CPU time by all members of the lsst group

CPU EFFICIENCY

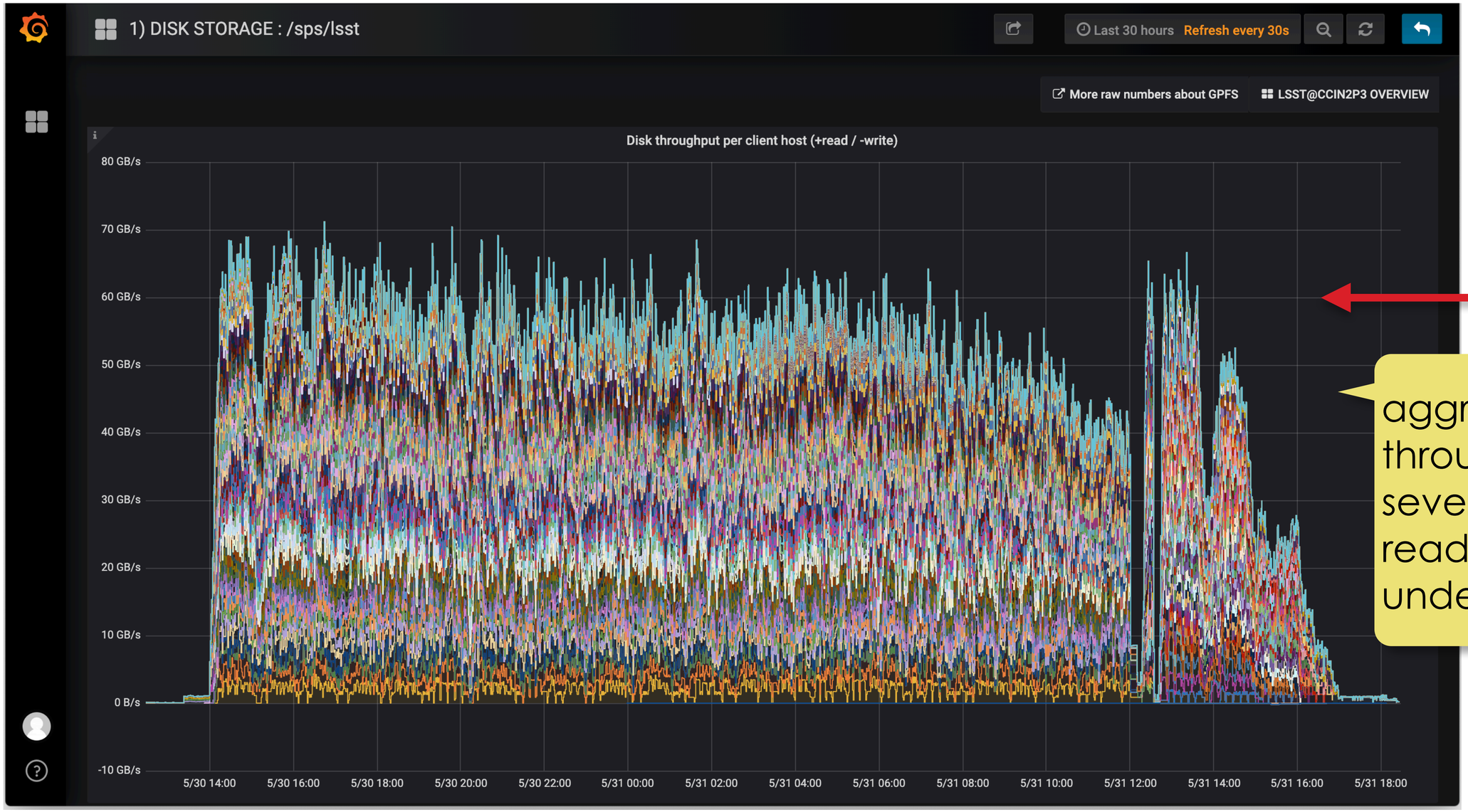


percentage of the **allocated** wallclock time the CPU cores were actually used (higher is better)

DISK STORAGE



DISK STORAGE (CONT.)



60 GB/s

aggregated throughput over several GPFS clients reading data stored under /sps/lsst

GPU RESOURCES AT CC-IN2P3

- Significant GPU resources are available at CC-IN2P3

  40 GPUs Nvidia Tesla K80 GK210, 12 GB DDR5

  24 GPUs Nvidia Tesla V100 GV100, 32 GB HBM2, PCIe

  you can find more details of the installation in this presentation by N. Fournials

- See also the material presented at the Workshop GPU @CC-IN2P3 (April 2019)

- Documentation on how to use that hardware is available here

Grafana - LSST@CCIN2P3 OVE

https://lsstmon.cc.in2p3.fr/d/Xn3y2dYik/lsstatccin2p3-overview?refresh=10s&orgId=1&from=now-6h&to=now

Incognito

LSST@CCIN2P3 OVERVIEW

Last 6 hours | Refresh every 10s

1) DISK STORAGE : /sps/lsst

Write throughput /sps/lsst

Read throughput /sps/lsst

2) INTER-SITE DATA TRANSFER

Reception network throughput

Last 10 minutes

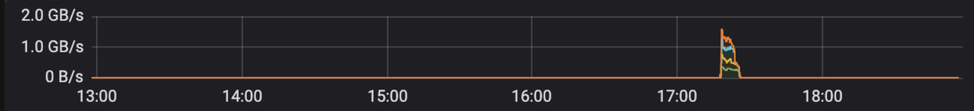
14.74 kB/s

Transmission network throughput

Last 10 minutes

25.91 kB/s

Reception throughput



Transmission throughput

3) BATCH ACTIVITY

Waiting jobs

Last 60 minutes

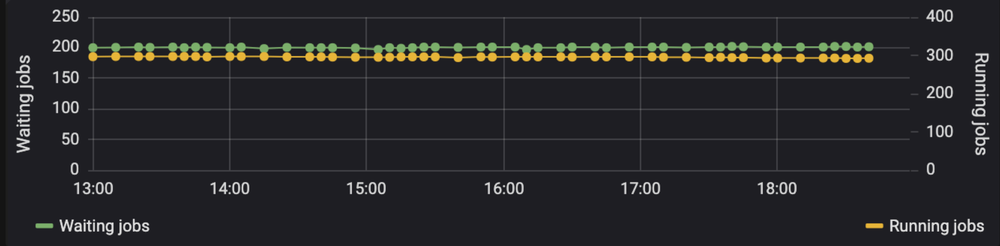
201

Running jobs

Last 60 minutes

293

Waiting & Running jobs



4) CATALOG DATABASE: Qserv

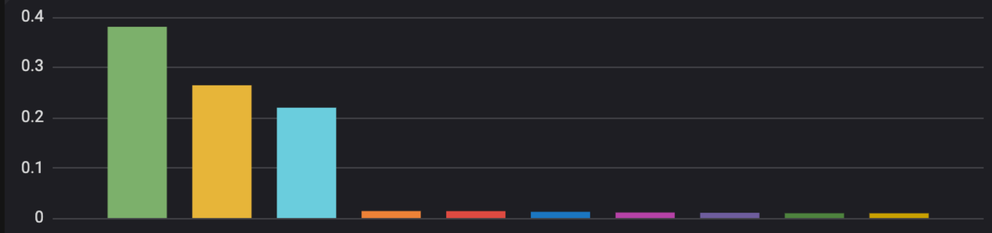
Qserv workers

50

Kubernetes hosts

2

Average load top hosts



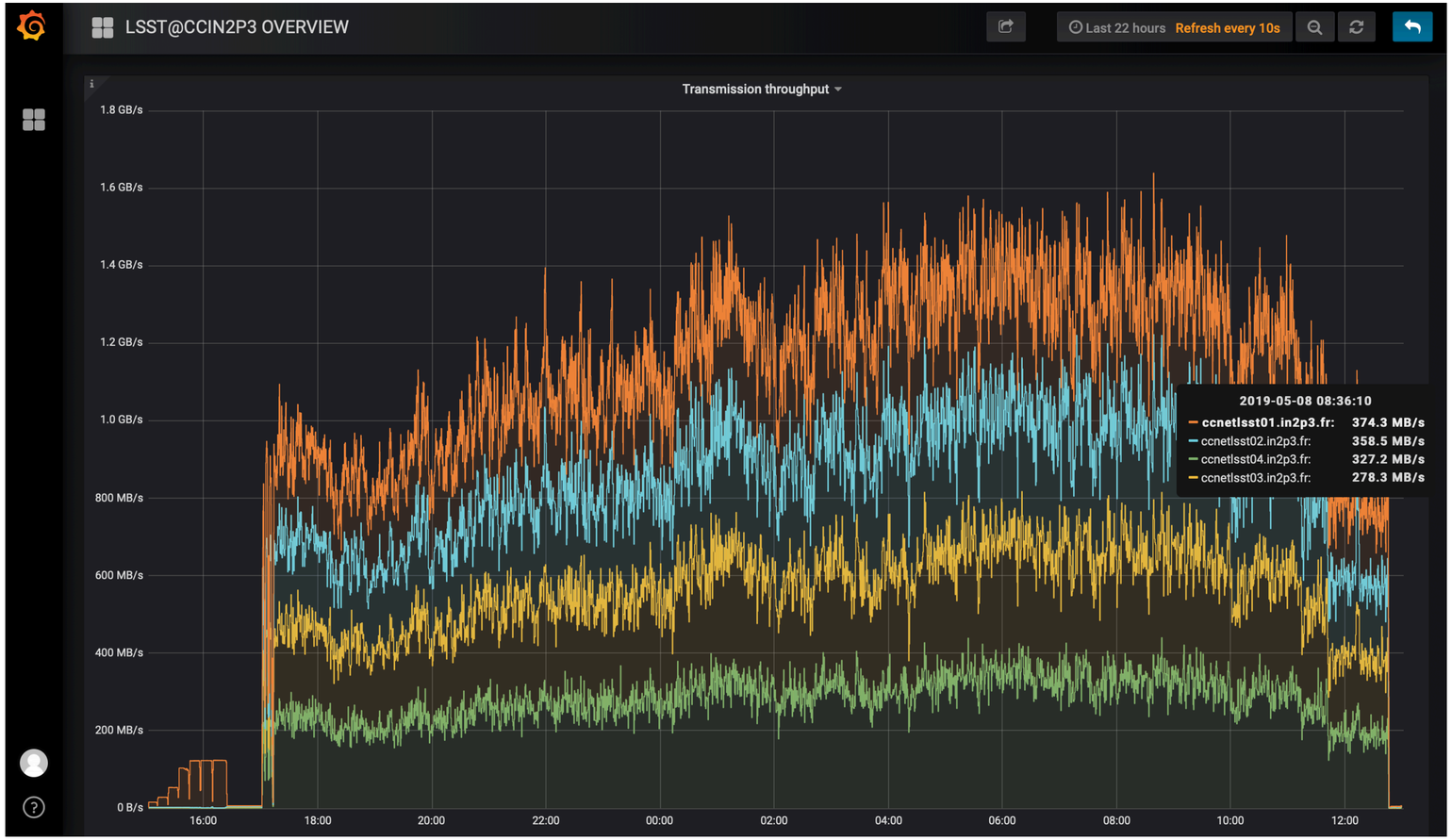
For details select specific dashboards below

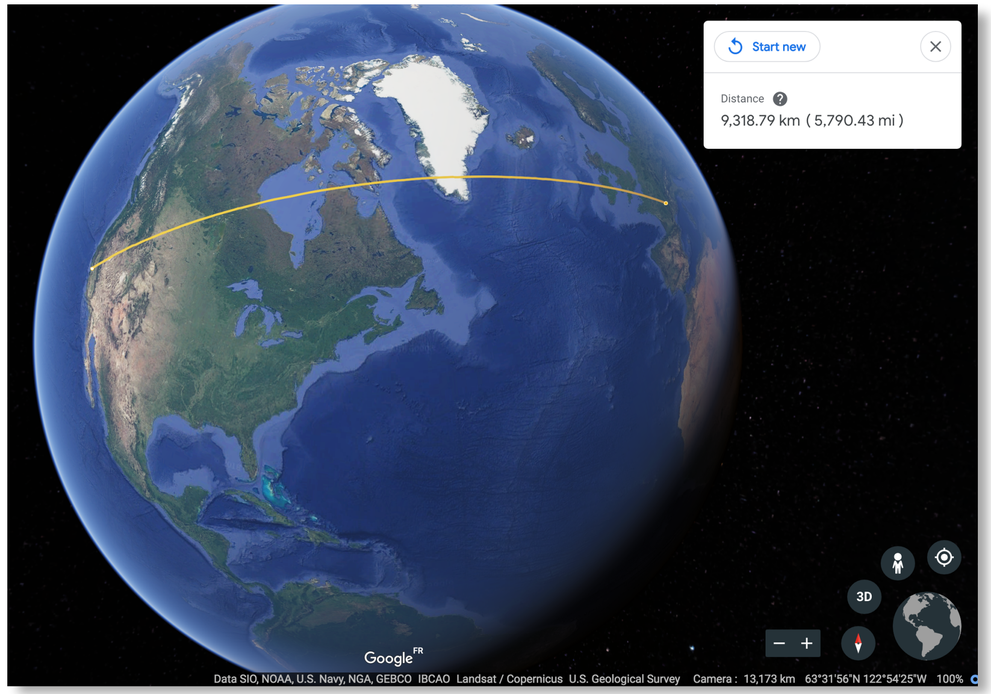
mon.lsst.eu

LSST DATA EXCHANGE WITH NERSC

data flow: **CC-IN2P3 → NERSC**

disk-to-disk, 150ms RTT, 9300 km





Average application-level network throughput over 20 hours from 2019-05-07 17h:

1 GB/s (8 Gbps)

Machine room infrastructure upgraded to host two additional hot corridors, equivalent to **80 additional racks** of computing equipment

ONGOING ACTIVITIES

- Distributed data management

  evaluation of RUCIO (replica catalogue and rule-based replication engine) and FTS (file transfer scheduler) for cataloguing and moving data over distributed sites

  initial deployment complete and ready to perform evaluation in production-like conditions

  FTS used to orchestrate transfer input data for image simulation (a.k.a. instance catalogs) from NERSC to Imperial College's grid site

  close contact with the RUCIO and FTS developers and user community

ONGOING ACTIVITIES (CONT.)

- CC-IN2P3 and LAPP grid interfaces configured to accept LSST image simulation jobs

  contribution by EU grid sites (currently UK + FR) to the simulation effort for DESC

- Renewal of hardware for catalog database (Qserv) in preparation

  hardware configuration better suited for database servers

- Deployment of Firefly for data visualization

  web-based application, integrated with notebook environment but also usable standalone

Demonstration of Firefly for data visualisation inside a Jupyter notebook

Source: B. Rigaud

ONGOING ACTIVITIES (CONT.)

- Stabilising the Kubernetes infrastructure for the catalog database (Qserv)

  joint work between LPC, LAPP, and CC-IN2P3

- Preparation of a web site for LSST members to easily share **static files**

  store the files you want to share under `/sps/lsst/users/yourlogin/web`

  those files will be accessible via `https://me.lsst.eu/yourlogin`

- Exploration of alternative workflow engines for implementing the image processing pipeline

  initial target: Pegasus + HT Condor

Jupyter Notebook Viewer

https://nbviewer.jupyter.org/github/airnandez/alba/blob/master/ALBA.ipynb

Incognito

jupyter nbviewer

JUPYTER FAQ

alba / ALBA.ipynb

Analysis of LSST Batch Activity at CC-IN2P3

Source: https://github.com/airnandez/alba

Author: Fabio Hernandez, CC-IN2P3

Introduction

The purpose of this notebook is to analyse accounting information emitted by GridEngine about the activity of LSST batch jobs executed at CC-IN2P3.

In [1]:
```
import pathlib
import datetime
import sys
import re
import collections

import IPython.display
print_md = IPython.display.Markdown

import pandas as pd
import numpy as np
import bokeh
import bokeh.plotting as bkh
import bokeh.models as bkhmodels
bkh.output_notebook()
```

BokehJS 1.0.1 successfully loaded.

Dependencies

This notebook uses the packages below listed below. The links point to their documentation:

github.com/airnandez/alba

Upcoming work

- Preparation of infrastructure for LSST-France broker demonstrator
- Improve interface for launching notebooks
  based on JupyterHub
  would allow for resuming work sessions
- Instrumentation of jobs for collecting and analysing coarse-grained job profiling data
  used CPU time, memory, I/O activity, stdout/stderr
  error detection by the job itself
- Design and development of tools for monitoring and keeping track of image processing activities
  from the application level perspective

Upcoming work (cont.)

- Contribute to testing Butler generation 3

  when available this summer

- Preparation of a documentation web site for people involved in operations

  for documenting how things work behind the scenes

- Perform real-world exercises to determine if the tools we have for migrating data between disk and tape are suitable for our needs

- Organise storage areas for groups, similar to what we did for individual users

  move the data currently under `/sps/lsst/data` *and organise it under* `/sps/lsst/groups`

  expect this to be disruptive and maybe even painful, both for you and for me, so bear with me

Questions & Comments