Source: LSST

LSST data processing at CC-IN2P3: status and perspectives

fabio hernandez

LSST-France, Marseille, Jan 17th, 2018

Background: LSST at CC-IN2P3

○ Formal agreement signed between LSST Corp., NCSA and IN2P3 in March 2015

○ Main roles of CC-IN2P3

satellite data release production under NCSA leadership

CC-IN2P3 to **process 50% of the data** *and store the full dataset, both raw and derived data*

both NCSA and CC-IN2P3 will exchange and validate the data produced by the other party

each site to host an **entire copy of every annual data release**

deployment and operations of a **data access center** *are not part of the agreement*

Contents

- Planning

- Documentation

- R&D activities

- Software

- Data storage, file systems, datasets

- Training and events

- Resource utilisation and requests

- Perspectives

- Summary

CCIN2P3   3

Planning

Planning

◦ Required equipment, estimate budget and deployment plan developed

covers period 2018-2031

*mostly **for data release processing** at CC-IN2P3: it includes equipment for catalogue database*

◦ Delivered in October an reviewed during 2017Q4

many thanks to M. Betoule (LPNHE), F. Chollet (LAPP) and G. Rahal (CC-IN2P3) for their detailed work and very useful feedback

◦ Documents

model and companion explanation: ATRIUM-215611

review report: ATRIUM-280394

the current model does not yet include the reviewers' recommendations

Planning (cont.)

LSST Documents Map
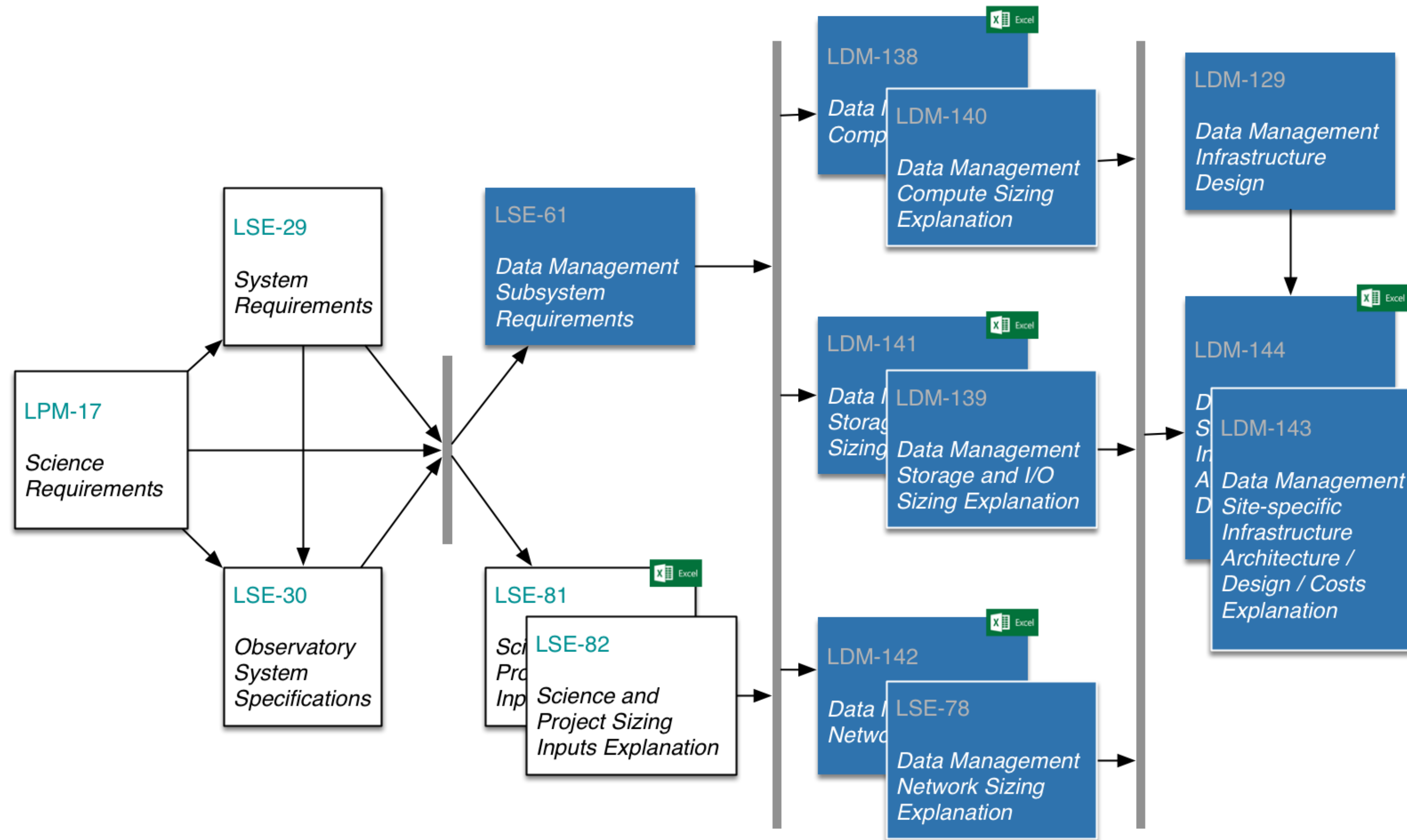
Subset of the LSST official documents used for estimating the computing requirements

NOTE: to retrieve e.g. LDM-144 use the URL : https://docushare.lsstcorp.org/docushare/dsweb/Get/LDM-144

Author: Fabio Hernandez
Last update: 2017-11

Planning (cont.)

Planning (cont.)

LSST — estimate budget for data release processing at CC-IN2P3
(contingency included)

■ Equipment & Travel ■ Operations
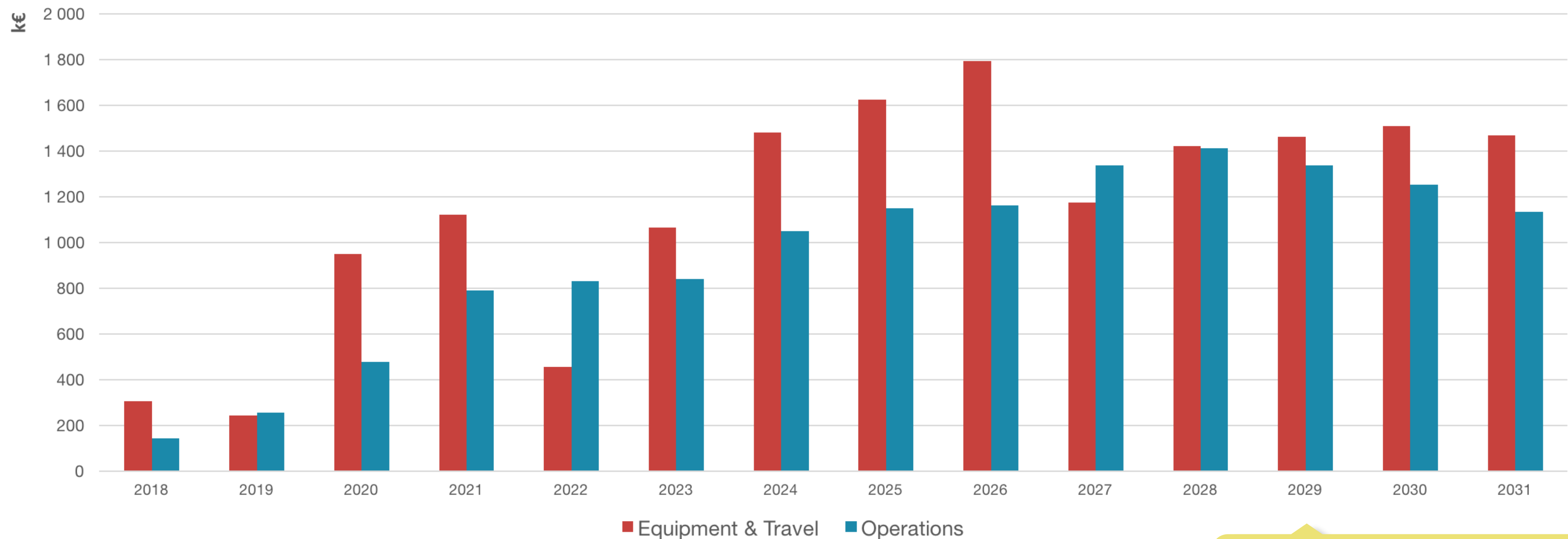
Total budget:
- Equipment: 16M€
- Operations: 13M€

Planning (cont.)

Number of racks on the floor
LSST data release processing at CC-IN2P3
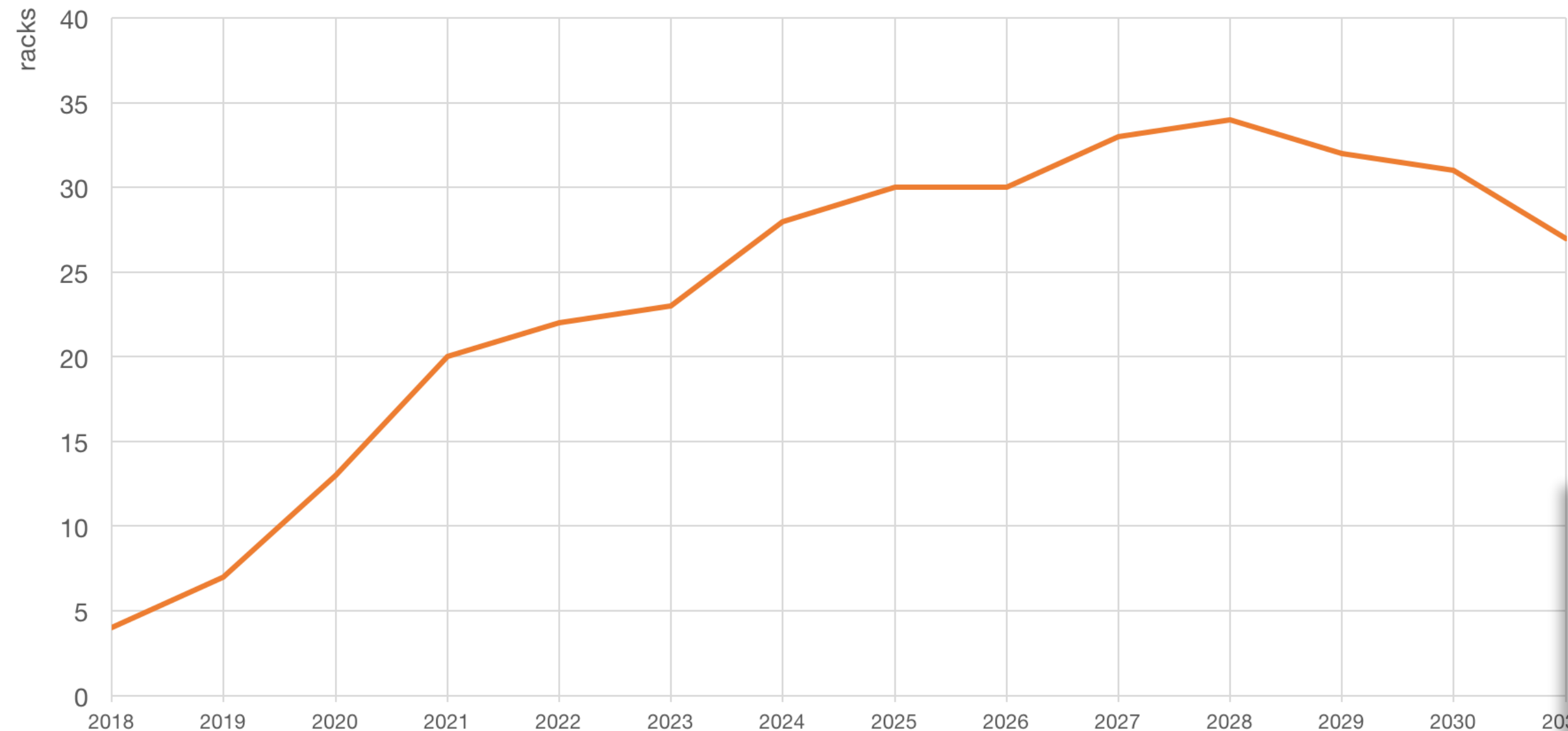
Racks

Power

Required Power
LSST data release processing at CC-IN2P3
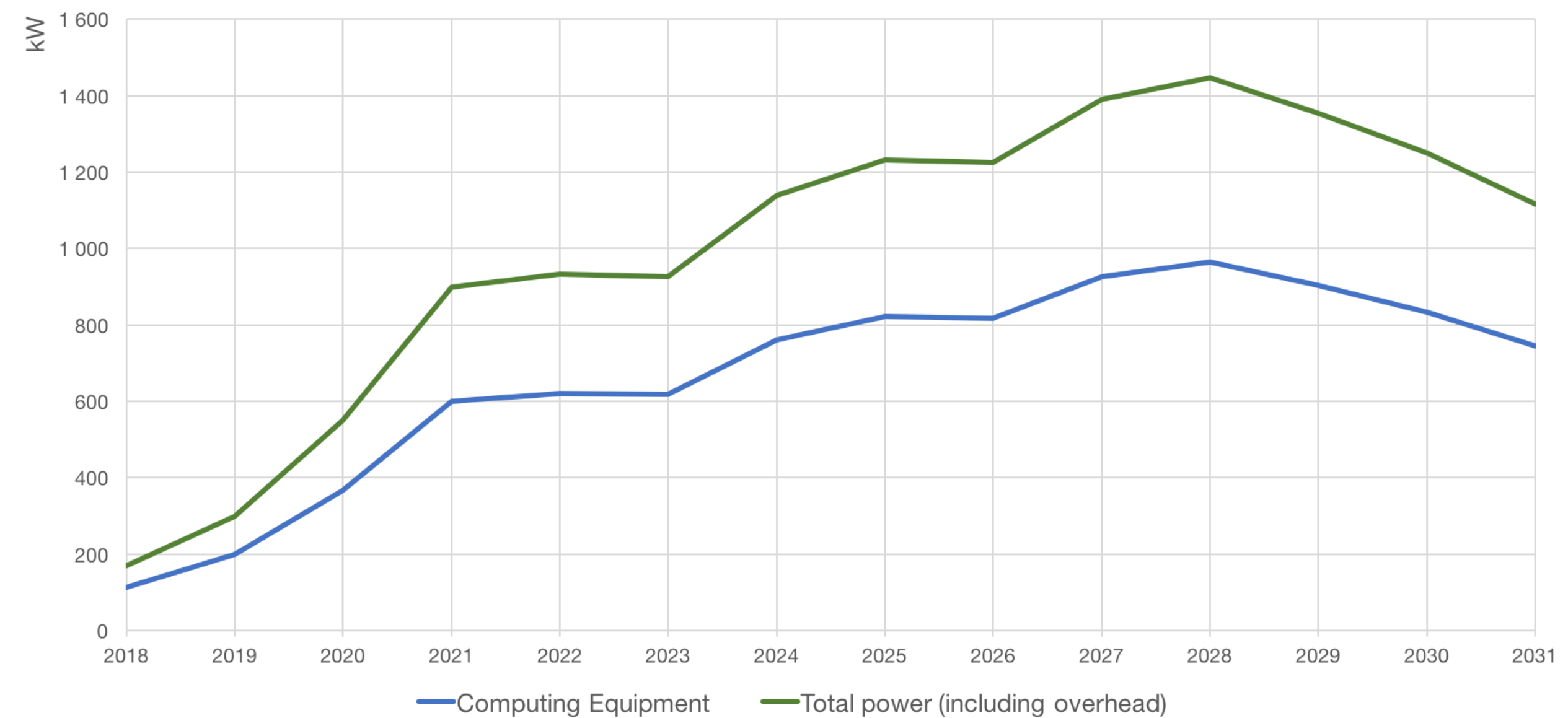
Planning (cont.)

Disk storage — required capacity
LSST data release processing at CC-IN2P3
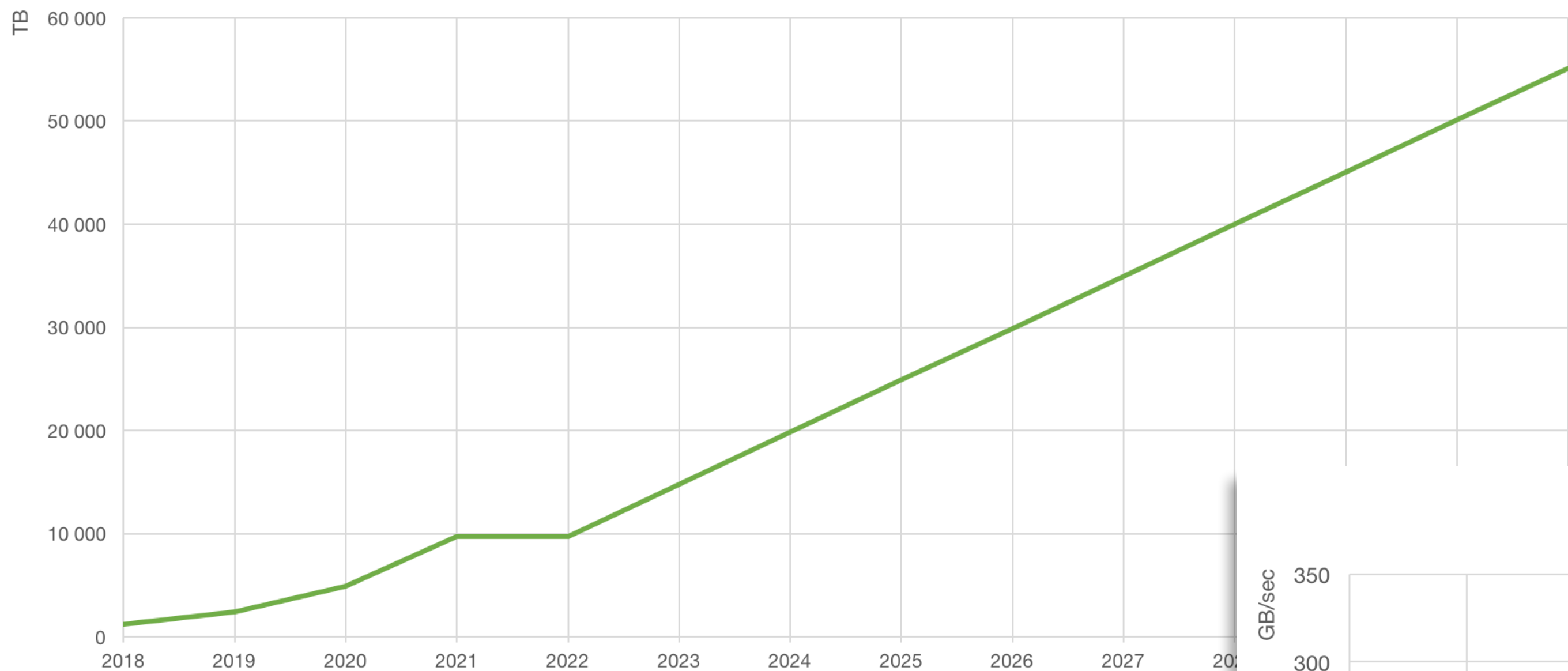
Disk capacity

Disk bandwidth

Disk storage - required I/O bandwidth
LSST data release processing at CC-IN2P3
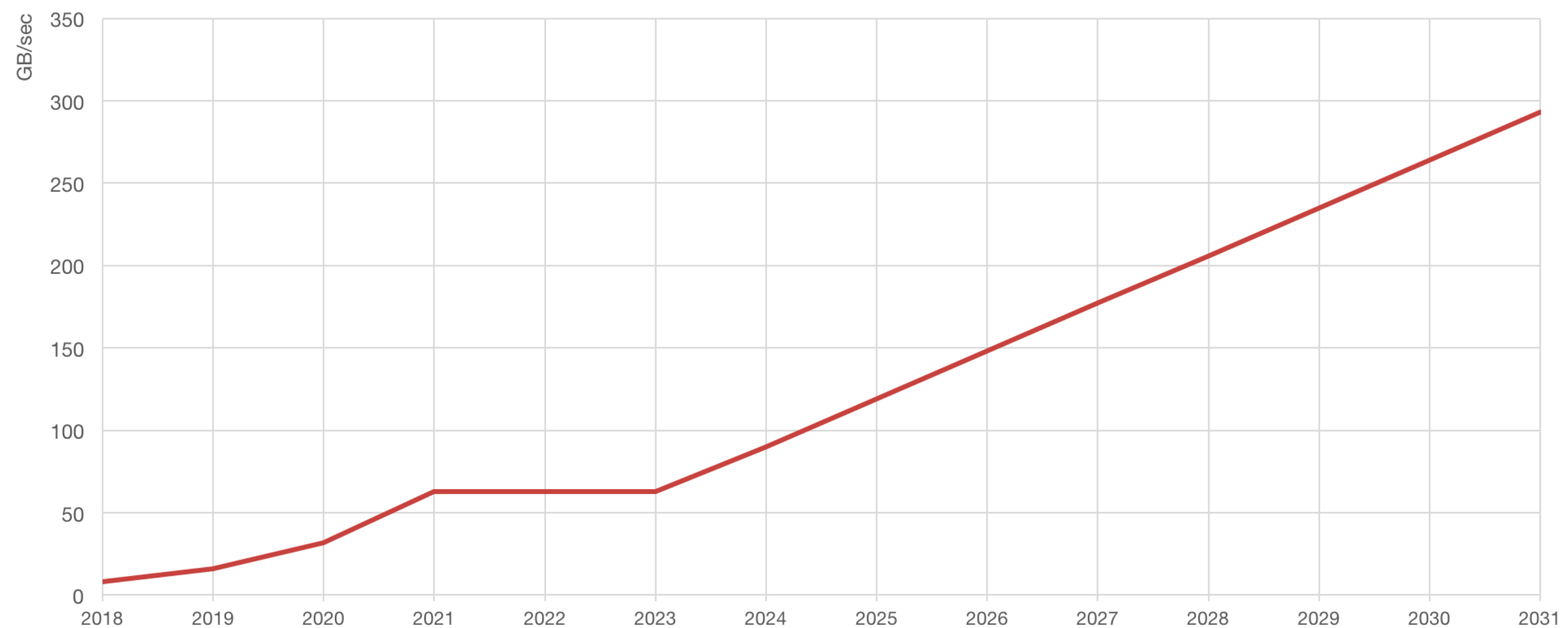
Planning (cont.)

Catalog database — required capacity
LSST data release processing at CC-IN2P3
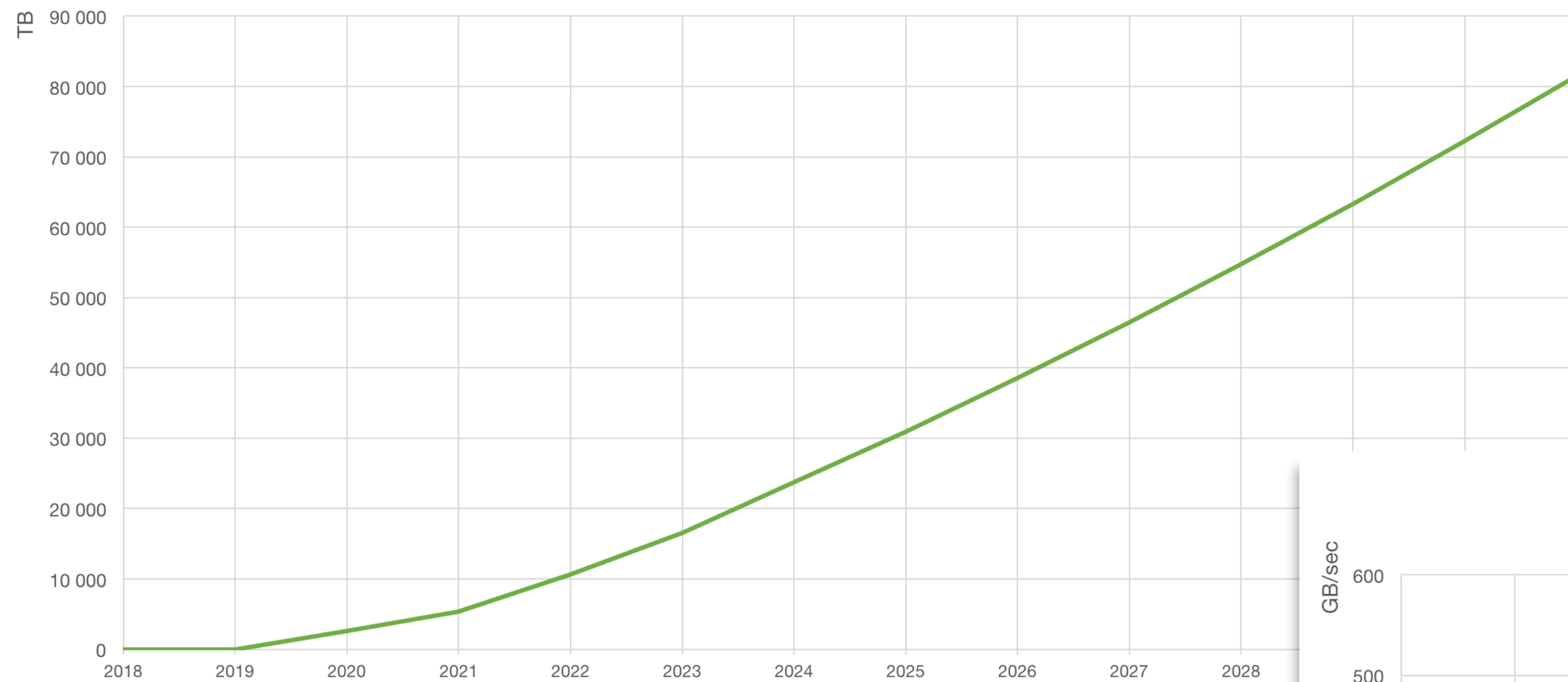
Catalog capacity

Catalog database — required I/O bandwidth
LSST data release processing at CC-IN2P3

Catalog bandwidth
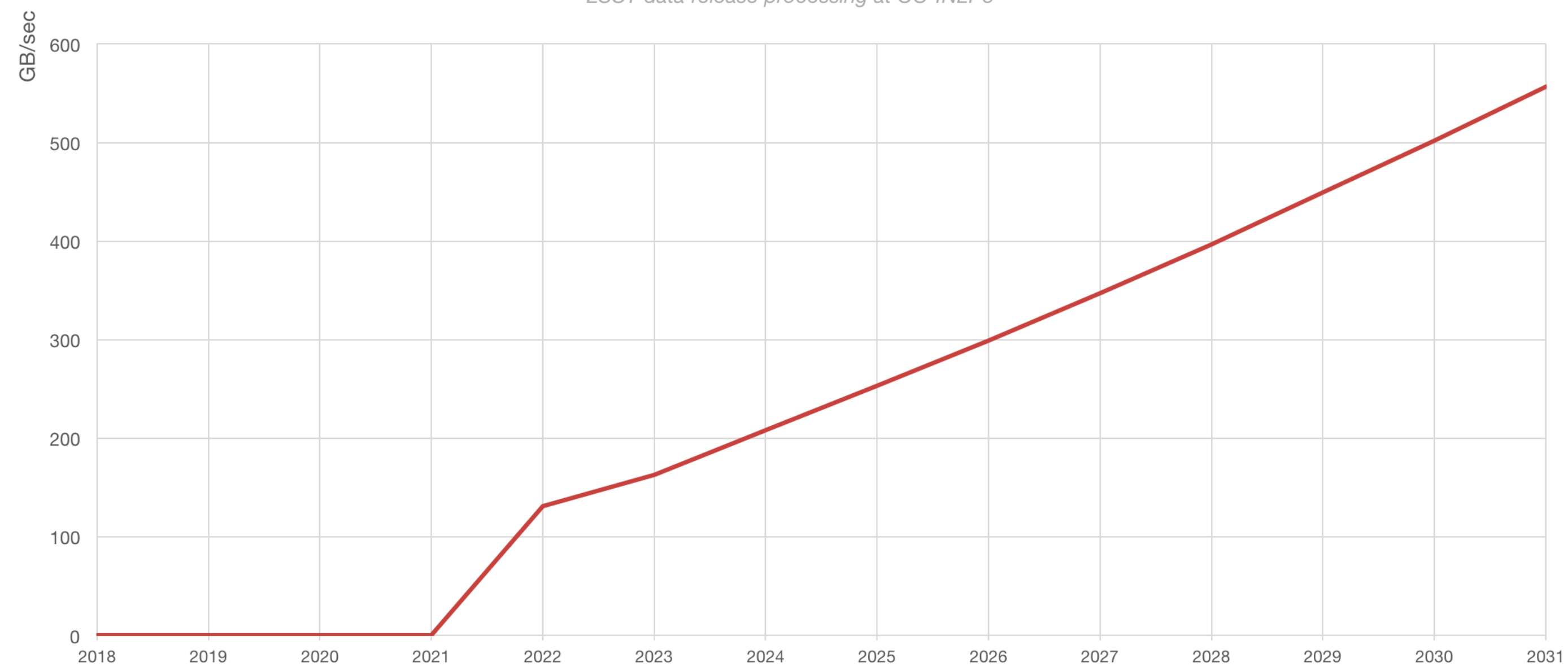
Planning (cont.)

Mass storage — required capacity
LSST data release processing at CC-IN2P3

Mass storage bandwidth
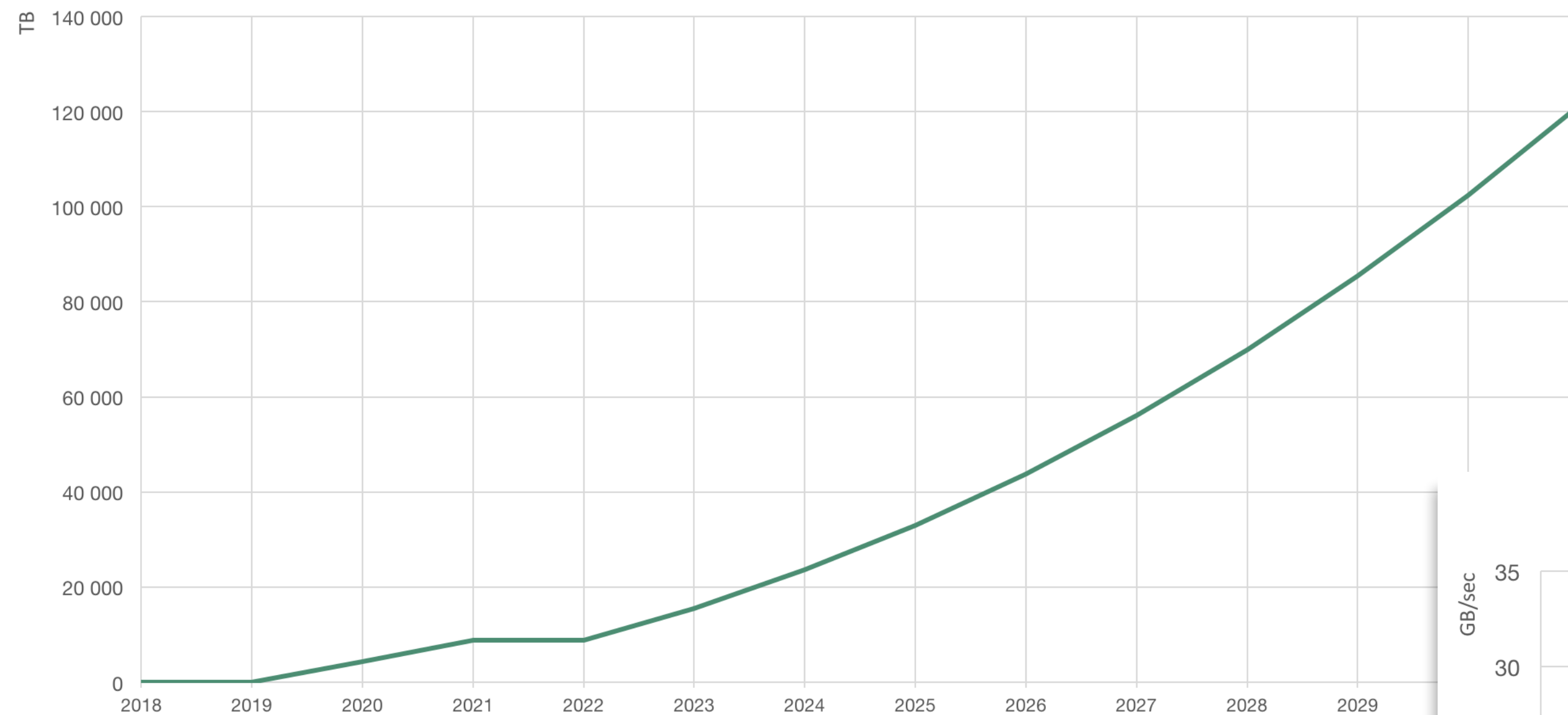
Mass storage capacity

Mass storage — required I/O bandwidth
LSST data release processing at CC-IN2P3

Planning (cont.)

CPU — required capacity

LSST data release processing at CC-IN2P3
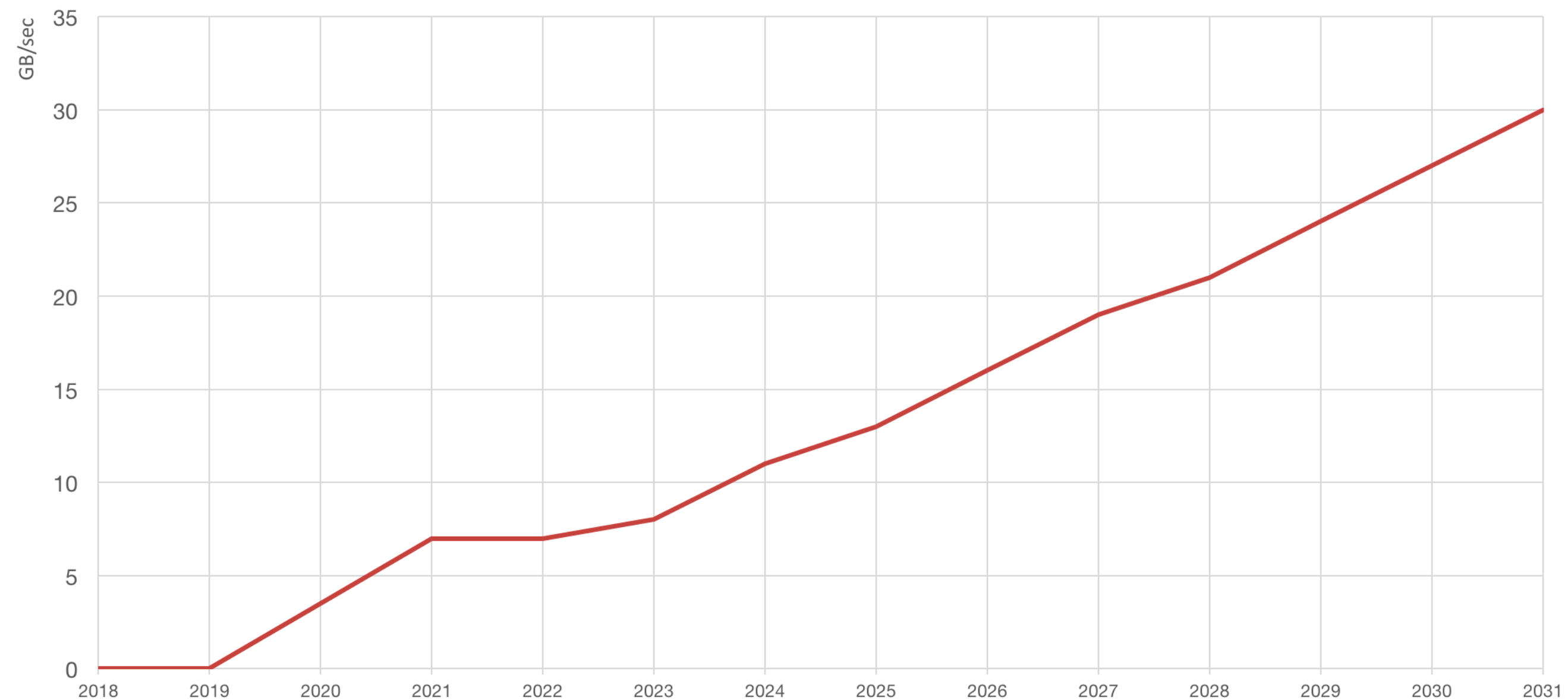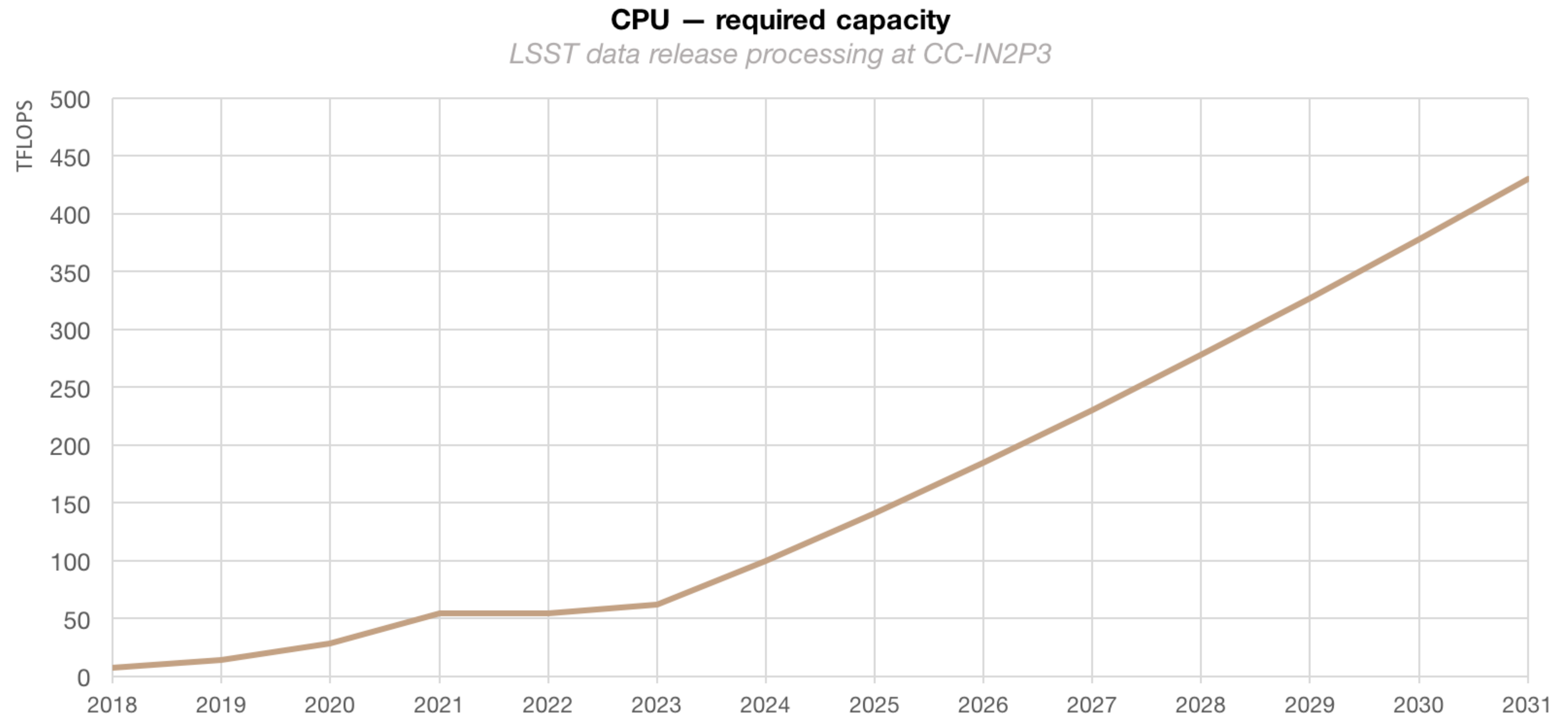
Documentation

◦ Special effort devoted to documentation: doc.lsst.eu

◦ Permanent ongoing work to improve contents, but we have set up the foundation to build up on

follows LSST documentation good practices and visual guide

contents written in reStructuredText markup and managed via git

uses IN2P3's gitlab instance for continuous deployment

◦ You can contribute too !

tell us what topics need improvements / clarifications

what additional topics need to be included in the documentation

R&D activities

◦ Started work aiming at understanding I/O activity induced by LSST image processing software

includes image ingestion and single frame processing

produced technical note DMTN-053

◦ Tools to bechmark generation of Butler repositories as produced by `ingestImages.py`

encouraging preliminary results

pending completion and publication

R&D activities (cont.)

○ Building blocks for bulk data import / export

secure HTTP for transport, object stores as emission / reception buffers

used for importing HSC data release from NCSA: 800K+ files, 17 TB
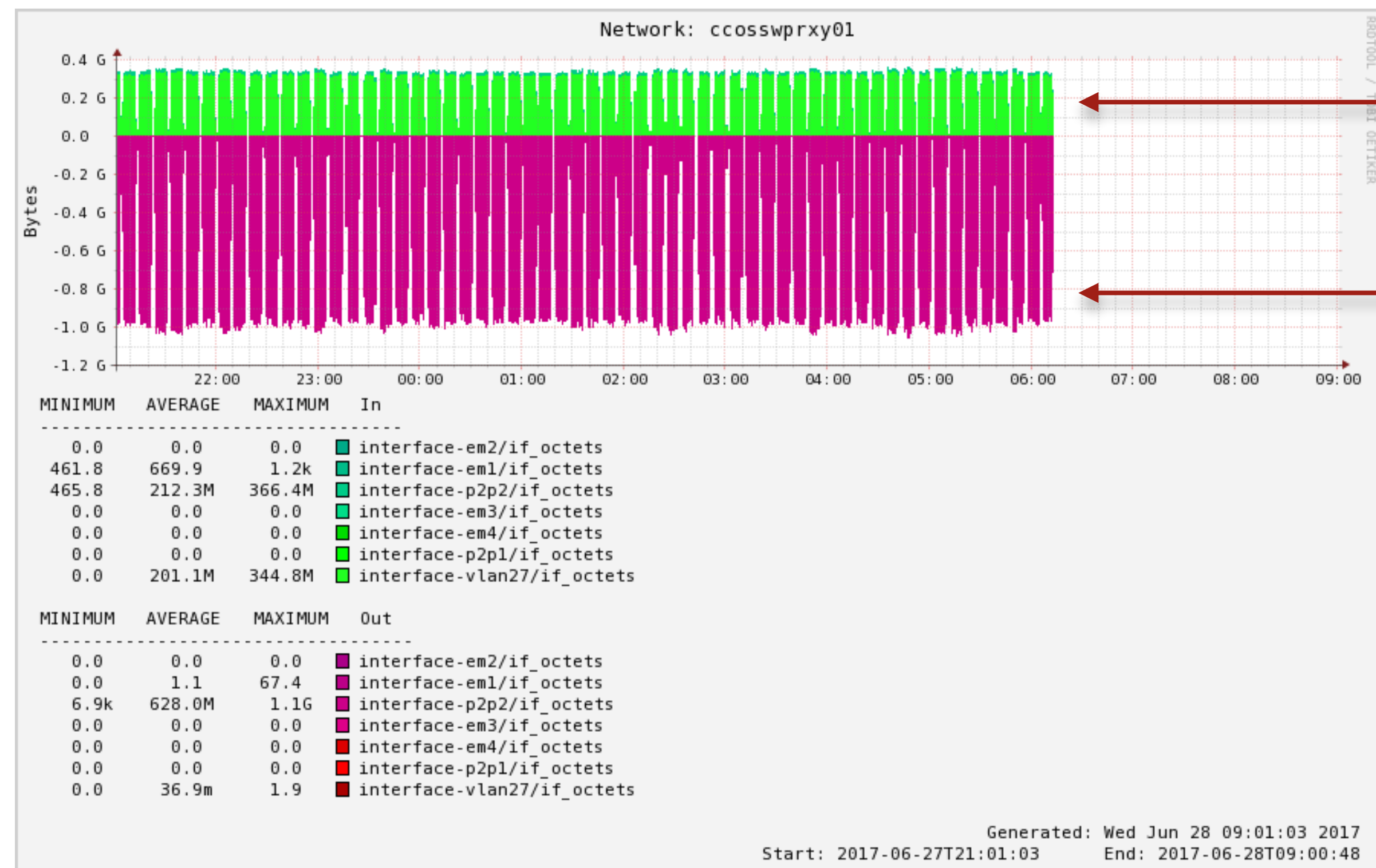
NCSA → CC-IN2P3 throughput using HTTPS: ~300 MB/sec

Data replication to 3 different file servers in object store

Results obtained without tuning all the components

Some of the limiting factors identified

R&D activities (cont.)

○ Building blocks for bulk data import / export (cont.)

development of benchmarking tools: `tlsping`*,* `netperf`*,* `chasqui`

publication of white paper as a contribution to the European project PRACE 4IP

○ Exploration of application containers as packaging mechanism for LSST processing pipelines

explored Docker and Singularity

○ Qserv team routinely using the two testbed clusters at CC-IN2P3 for their development work

e.g. benchmarking (see Fall 2016 report), big memory machines for Qserv secondary indexes (see report)

R&D activities (cont.)

○ One of the two big memory machines routinely used by LPNHE group for processing data from Subaru

○ Enabled mechanism for establishing secure, passwordless connections to CC-IN2P3 login farm

convenient also for launching notebook servers

see the documentation for details

Software

○ Consistent installation of LSST software framework at CC-IN2P3 in shared area

both stable versions and 12 most recent weeklies

Python 3 only

automatic configuration of C++ runtime dependencies: from version `w_2017_50` *on the stack requires GCC 6.3 (devtoolset-6)*

includes additional Python packages, not included in the official LSST distribution (e.g. `ipython`, `jupyter`, `ginga`, *…)*

details in the [documentation](#)

ongoing work for preparing systematic deployment of weeklies in the cloud, accessible via CernVM-FS from both the hosts at CC-IN2P3 and your laptop

once this work if finished, the installation local to CC-IN2P3 will then be phased out

Software (cont.)

◦ The CC-IN2P3 environment for LSST can be used via Python Notebooks

notebook-based interface is one of the components of the LSST science platform

builds up on Nicolas Chotard's `stackyter` *tool to conveniently launch notebook servers at CC-IN2P3 from your laptop*

configurable to use any version of the LSST software available at CC-IN2P3

see the documentation for details

Data storage and file systems

○ Ongoing campaigns to organise the storage areas usable by members of LSST-France

shared image datasets and reference catalogs: `/sps/lsst/datasets`

area for individual user's datasets: `/sps/lsst/users/you`

backed-up area for individual user's software and documents: `/pbs/throng/lsst/users/you`

intended usage of each one of the available areas is *documented*

○ The process of reorganisation has been disruptive for some of you

apologies, but we are preparing the field for playing together during the next 15 years

more disruptions to come !

Data storage and file systems (cont.)

○ Creation of areas for sharing data among members of a group within LSST

end goal: make shared data discoverable, reduce documentation effort for things that should be intuitive, make data owner obvious, avoid data duplication, etc.

example of potential groups: photoz, camera, bao, calibration, clusters, …

existing data would be progressively moved under `/sps/lsst/groups`

each group area to be owned by and under the responsibility of one member of the group

feedback welcome

Data storage and file systems (cont.)

○ Bottom line: **don't hardcode paths to data**

current paths very likely won't survive the next 15 years, you have been warned !

○ Deletion of obsolete data

*without action by the person responsible for LSST at each IN2P3 site, **all the files owned by expired accounts will be deleted***

relevant people will be notified by e-mail: your prompt response (in fact, any response) will be appreciated

Datasets

- Datasets available in shared area at CC-IN2P3

  Hyper Suprime-Cam Subaru Strategic Program public data release 1 (DR1)

  other relevant datasets to be included? (e.g CFHT, DES)

- Reference catalogs

  currently Gaia, Pan-STARRS and SDSS

- See documentation

Training and events

- Training programme targeted at the LSST-France community on the usage of LSST software

  Python language (12 attendees): agenda

  Python libraries for data analysis (numpy, matplotlib, …) and LSST stack (9 attendees): agenda

  2 days-long, notebook-based, hands-on focused

  support material attached to the agendas

- International event: lyon2017.lsst.eu

  school, workshop and hackathon

  24 speakers, 60 attendees from 13 countries

  all presentations video-recorded and available online

Resource utilisation

Resource utilisation: disk
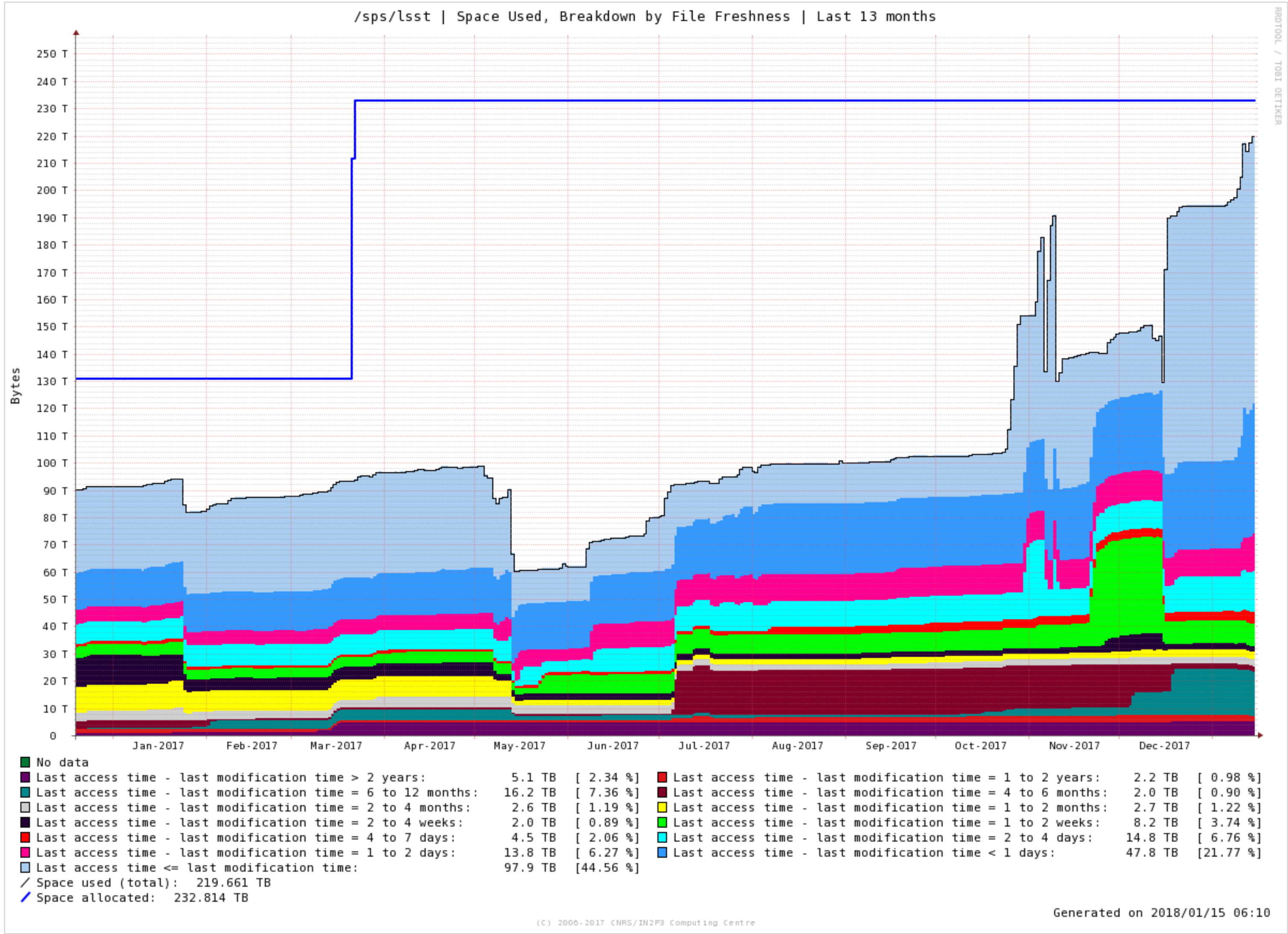
Resource requests

○ Requested for 2018

	2018 Q1	2018 Q2	2018 Q3	2018 Q4
CPU [HS06 hours]	10M	10M	10M	20M
Disk /sps/lsst [TB]	+200 TB	+300 TB	+500 TB	+500 TB

• CPU: same request than 2017 allocation. Equivalent to 200 to 400 recent CPU cores devoted to LSST. In 2017 LSST used 10% of its allocation

• Disk: current disk allocation is 230 TB (95% used)

Perspectives

Perspectives

○ Activities on several fronts foreseen in 2018

○ LSST data release processing

to prototype integration of CC-IN2P3's **authentication and authorisation** *mechanisms to LSST's*

to demonstrate CC-IN2P3 able to **execute LSST tasks** *submitted from NCSA and report the results back*

to understand what capability do we need to provide to support **workflows**

to investigate mechanisms for **central logging facility** *for LSST jobs executed at CC-IN2P3, with advanced query capabilities*

to finalise systematic deployment of LSST stack **weeklies** *on CernVM FS*

to **benchmark the existing storage platform** *(both disk- and tape-based) in order to establish a baseline*

○ LSST data analysis center

to make further progress on the evaluation of Kubernetes for dynamically managing the resources devoted to Jupyter servers

○ DESC data challenge 2

to understand what are the goals of this exercise, the foreseen contribution by IN2P3 and the mechanisms required at CC-IN2P3 to participate

Questions & Comments

Backup slides

LSST data management

◦ Archival

*to **record, transport** and permanently **store raw data** issued by camera*

◦ Processing

*to **detect transients** and **emit alerts** within 60 seconds after observation*

*once per year, to **release a self-consistent, immutable dataset**, composed of processed data since the beginning of the survey ("Data Release Processing")*

*to **develop the software** necessary for processing the data: image processing algorithms (calibration, point spread function, co-addition of images, characterization of objects, processing pipelines, …), catalogue database, middleware (workload management, orchestration, …), data transfer, etc.*

◦ Publication

*to **deliver the reduced data** (images + catalogs)*

to facilitate custom data reduction and individual data analysis

Data products

Nightly

Stream of 10M time-domain **events** per night, detected and transmitted to event distribution networks within 60 seconds of observation

Catalog of **orbits** for 6M bodies in the Solar System

Annual

Catalog of 37B objects (20B galaxies, 17B stars), 7T observations, 30T measurements, produced annually, accessible through databases

Deep co-added **images**

On demand

Services and computing **resources** to enable user-specified custom processing and analysis

Software and APIs enabling development of analysis code

Source: LSST

LSST DATA CENTERS

SATELLITE RELEASE PRODUCTION SITE

Archive center
data release production
long-term storage (copy 3)

HEADQUARTERS SITE

HQ facility
observatory management
science operations
education & public outreach

ARCHIVE SITE

Archive center
alert production
data release production
calibration products production
long-term storage (copy 2)
education & public outreach
infrastructure

Data access center
data access and user services

BASE SITE

Base facility
long-term storage (copy 1)

Data access center
data access and user services

SUMMIT SITE

Summit facility
telescope & camera
data acquisition
crosstalk correction

Sources: LSST, Google