

Pre-scrubbing follow-up

June 30, 2021

Pre-scrubbing WBS 2.3(.1) follow up #2

Today's pre-scrubbing follow-up primarily triggered by Paolo's email and disk storage table

Tuesday, June 29, 2021 at 11:31:14 Eastern Daylight Time

Subject: BNL disk storage

Date: Wednesday, June 23, 2021 at 6:03:51 PM Eastern Daylight Time

From: Paolo Calafiura

To: Lancon, Eric, Benjamin, Douglas, McKee, Shawn, Rob Gardner

CC: Kaushik De

Hi everyone,

We had a look at the pledges for 2021 and 2022, at the information Eric provided us last week, and made some assumptions on what we may expect for 2023 pledges and retired disk space (or more accurately repurposed disk, if we want to use it for CEPH testing).

This is the result of our exercise:

<https://docs.google.com/spreadsheets/d/1yDhhaEmK1MbvuiFFf3vUjPQhgICKnkFo4ZnmCMrt0mA/edit?usp=sharing>

Unless we got some of the assumptions wrong, it would look like if we keep the disk purchases constant at 7.6 PB/year, we should be OK this year and 2022 and 2023.

Please have a look and let us know if you see anything unclear/wrong.

Kaushik and Paolo

-

-

| Ph:1-510-4866717 Fax:4864004

|

| Lawrence Berkeley National Lab

| Mailstop 50F-1606

| 1 Cyclotron Road

| Berkeley CA 94720-8147 - USA

|

| <http://crd.lbl.gov/departments/computational-science/phax/staff/paolo-calafiura/>

Paolo's table

Available/purchase disk storage volume is different from usable storage volume because of several effects: File system overhead (~30%), data replication factor (2x-1.25x), operational overhead (~15%) etc...

Purchase year	K\$	TB	TB/K\$		\$/TB (usable)
2015		322	1046	3.2 <2023	307.8
2016		579	4064	7.0 <2023	142.5
2017		823	6096	7.4 <2023	135.0
2018		126	0	0.0	
2019		1796	12099	6.7	148.4
2020		63	0	0.0	
2021		686	7635	11.1	89.8
Average		628			

Storage go by quanta represented by a single JBOD + head node pair (not continuous decimal values)

Year	Retire --> CEPH (TB)	Purchase TB	Available TB	Pledge TB	LOCALGROUP DISK	Needed TB	Extra (TB)	Notes
Now			23305					
2021	1046	7635	29894	24000	2000	26000	3894	Retire 2015
2022	4064	7600	33430	27800	3000	30800	2630	Pledge recommendation +10%=26.4 PB. Retire 2016
2023	6096	7600	34934	30580	4000	34580	354	Hypothesis pledge +10%, Retire 2017

'old' (passed life time) are planned to be retired at once in FY23 (not over 3 years as indicated here) as part of transition plan to new data center

ATLAS requests

- ATLAS needs are known up to 2022 only
 - They have not been approved by CRSG
 - 2022 is +15% over 2021 agreed by CRSG resources
- Extrapolations are required for forward planning
 - Atlas computing model (Conservative R&D)
 - 2023 wrt 2022: +5% CPU, +10% Disk, +30% Tape
 - Conservative R&D departure from baseline in 2025

In summary, after careful evaluation, we have updated our resource request, lowering the requested CPU thanks to the possible improvements in Full Simulation and reviewing the contingency on the number of MC simulated events. As noted in the October 2020 C-RSG report, the request assumes 270 kHS06 will be provided by the HLT farm during 2022. The final resource request for 2022 is summarised in Table 5.

	2021 Agreed @ April 2020 RRB	2022 Request @ Oct 2020 RRB	2022 Request @ March 2021 RRB	Balance 2022 wrt 2021
T0 CPU (kHS06)	525	550	550	5%
T1 CPU (kHS06)	1170	1415	1356	16%
T2 CPU (kHS06)	1430	1730	1656	16%
SUM CPU	3125	3695	3562	14%
T0 DISK (PB)	29	32	32	10%
T1 DISK (PB)	105	121	121	15%
T2 DISK (PB)	130	148	148	14%
SUM DISK	264	301	301	14%
T0 TAPE (PB)	95	118	120	26%
T1 TAPE (PB)	235	272	272	16%
SUM TAPE	330	390	390	18%

Table 5: Summary of the final ATLAS requests for computing resources in 2022.

Disk storage plan

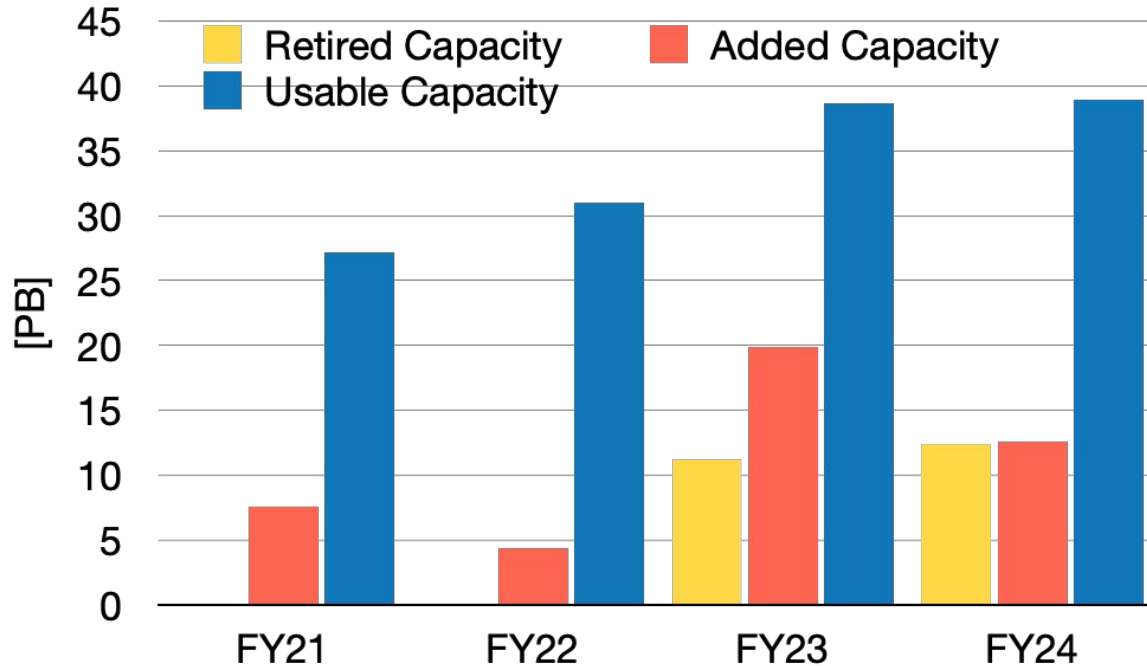
Ceph or no Ceph in FY23

- Scenario 1 (S1):
 - Ceph is validated and deployed for main storage solution in FY23
 - dCache (1+1) replicated storage is purchased in FY22
- Scenario 2 (S2):
 - Ceph is deployed for main storage solution in FY24
 - dCache (1+1) replicated storage is purchased in FY22 and FY23
- In both scenarios
 - No disk storage equipment is retired before FY23
 - The initial purchase of new Ceph cluster capacity contains 16 JBOD + head node pairs (due to the desire to reach 1.25 replication factor with EC formula 12+3, and at the same time allowing one OSD server to be absent at any time due to a HW failure without compromising the recoverability of the cluster from)

(S1) Retired, added, usable disk space

Scenario 1

Ceph in FY23

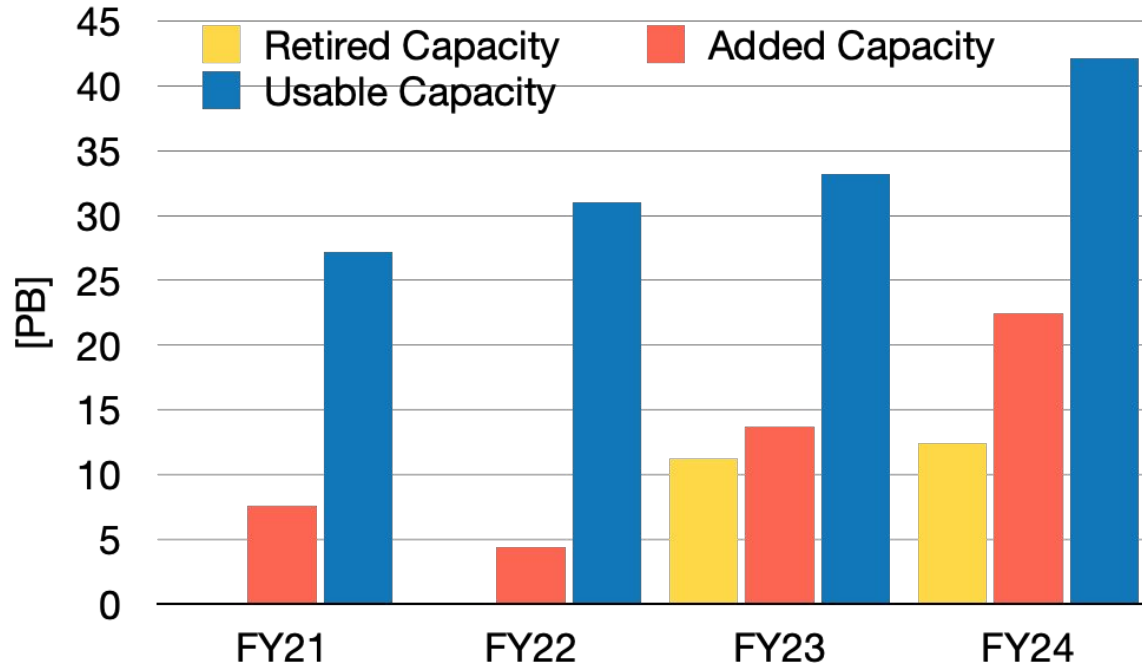


Usable
includes
Added
capacity

(S2) Retired, added, usable disk space

Ceph in FY24

Scenario 2



Usable
includes
Added
capacity

Additional items



Proposed New 2.3.1 Organization

❖ Why?

- More understandable organization with clearer roles
- Reflects more closely the current allocation of effort

❖ 2.3.1.1: Admin

❖ 2.3.1.2: (shared) Facility infrastructure

❖ 2.3.1.3: Linux farm (including T3)

❖ 2.3.1.4: All storage effort (Hardware & Software)

❖ 2.3.1.5: Services of (US)ATLAS

Do we agree on this proposed reorganization of the WBS?

❖ **Total SDCC effort across WBS 2.3/2.4 remains unchanged**

Cross-cutting WBS 2.3

- Change of Ofer's effort from 0.65 to 0.85 across WBS 2.3
 - BNL Analysis Facility POC, 0.55 FTE (was 0.2)
 - CIOPS contribution 0.3 FTE (was 0.45), complemented by Qiulang Huang (0.25 FTE)