

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

Belle II's preparation for Data Challenge '24

Aresh VEDAEE (29.01.24)

- Data Challenge 2024
- BELLE II plans
- BELLE II pre-DC24 tests

- **What?**

- Adding transfers of real data to the normal activity (which will remain effective during DC24 period) according to the data distribution model of each experiment

- **Why?**

- Validate readiness for high-lumi throughput & high concurrency:
 - E.g. 25% of the nominal rate of HL-LHC (HL-LHC throughput estimated to be 1.2Tb/s or 4.8Tb/s for overprovisioning)
 - 6 experiments will participate: ATLAS, CMS, LHCb, ALICE, BELLE II, DUNE.
- Validate new technologies/solutions for data transfers

- **When?**

- February 12 – February 23

- Link:

- <https://indico.cern.ch/event/1307338/timetable/#20231109.detailed> (9-10 Nov 2023)

- **3 themes:**

- The test plans and throughput targets by experiment
- New network features (e.g. packet marking, packet pacing, jumbo frames, Noted, Sense).
- Storage software stack (XRootD and dCache).

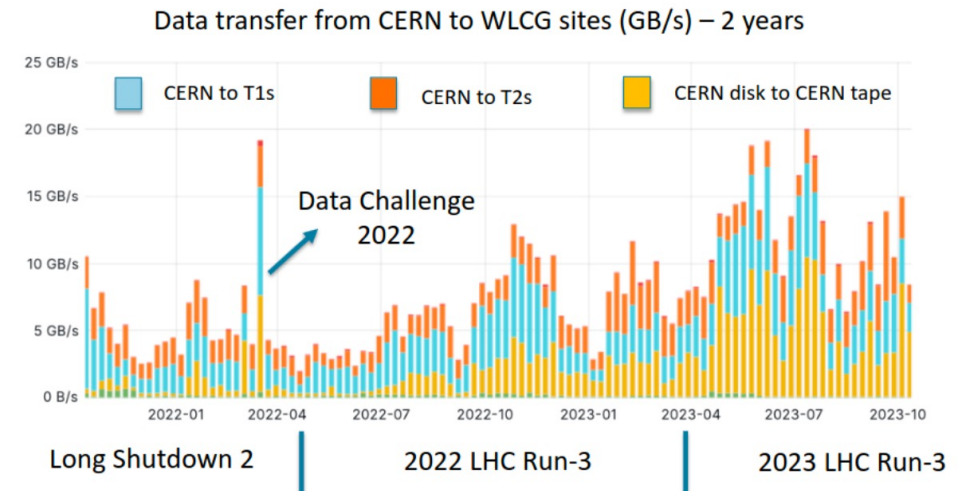
- **DC24 vs DC22:**

- Disk storage -> network-> Disk storage.
 - Only LHCb wants to validate the transfers to tape.
- 2 Scenarios:
 - Minimal and flexible
 - Flexible = 2 * minimal => more realistic
- Each test should last at least 48h

- **Muttermost channel:**

- <https://mattermost.web.cern.ch/wlcg-gdb/channels/wlcg-data-challenges>

TO EXPORT



Data Challenge 2024: Packet Marking

- Objective: to identify the different flows at the network level (e.g. by experiment and activity)

- Two possible implementations:

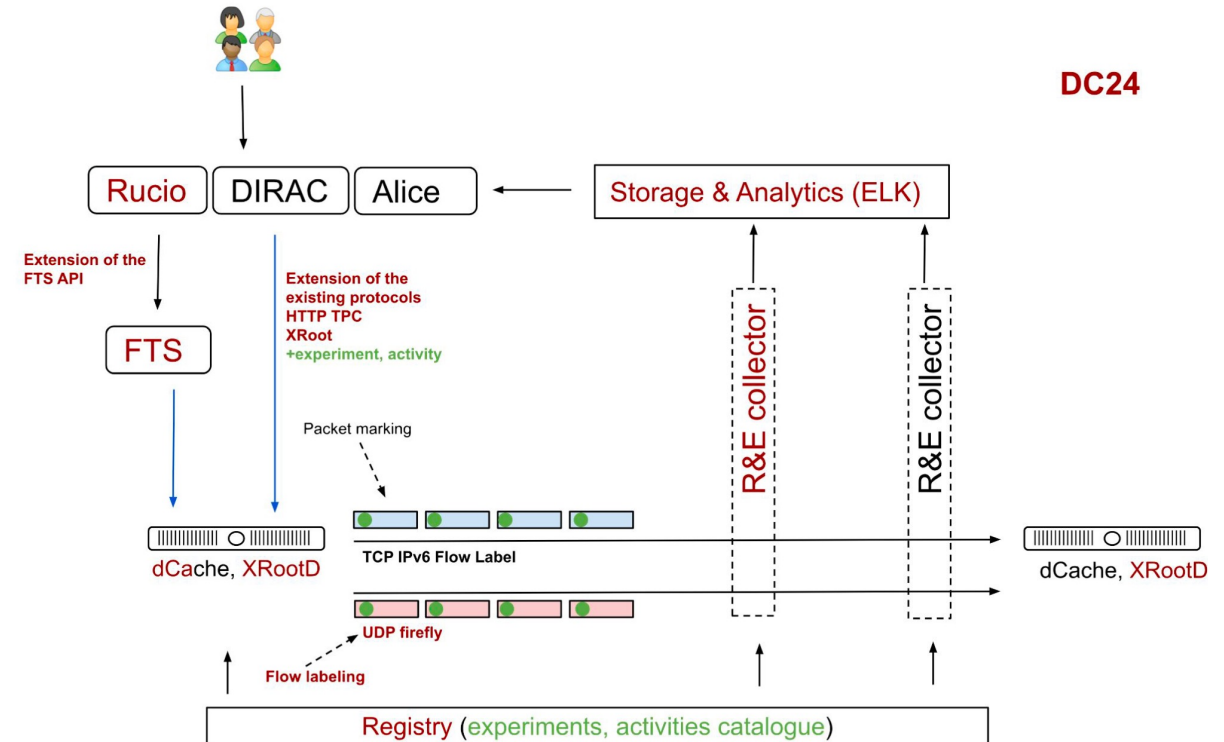
- Insertion of Scitag in the packet header (IPv6 flow label)
 - IPV6 only
 - Perfect method, we miss nothing
- Insertion of UDP fireflies into transfer flow (UDP packets in syslog format and port 10514):
 - Works with IPv4 and IPv6
 - Possible UDP packet loss

- Both implementations require compatible middleware:

- Rucio: support for SciTags since 32.4.0
- FTS/gfal2: support for SciTags since FTS 3.12.11 and GFAL2 2.22.0 (HTTP-TPC headers)
- dCache: prototype available for SciTags & testing at AGLT2

- Link:

- https://indico.cern.ch/event/1307338/contributions/5657197/attachments/2749118/4785253/DC24_%20Packet%20marking.pdf



- **OIDC Tokens (Scitoken and WLCG JWT):**
 - dCache: since version 7.2
 - HTCondor 9.0.20 (+ SSL mapping)
 - <https://confluence.egi.eu/display/EGIBG/HTCondor+and+SSL+authentication>

- HTCondor:
 - 9.0.20 (while HTCondorCE 5.4) => still problems with the SSL mapping configuration
 - https://ggus.eu/index.php?mode=ticket_info&ticket_id=163985
- dCache (2 instances):
 - 9.2 for LCG => capability based authentication (e.g. storage.{read,create,modify}) successfully tested with ATLAS & CMS over webdav
 - 7.2.27 for EGEE, which includes Belle II => upgrade prior to DC24
- Network:
 - LHCOPN upgraded from 100 Gb/s to 200 Gb/s (15 Jan 2024)
 - LHCONE upgraded from 100 Gb/s to 200 Gb/s (29 Jan 2024) => see Benoit's presentation!

- BELLE2 presentation:

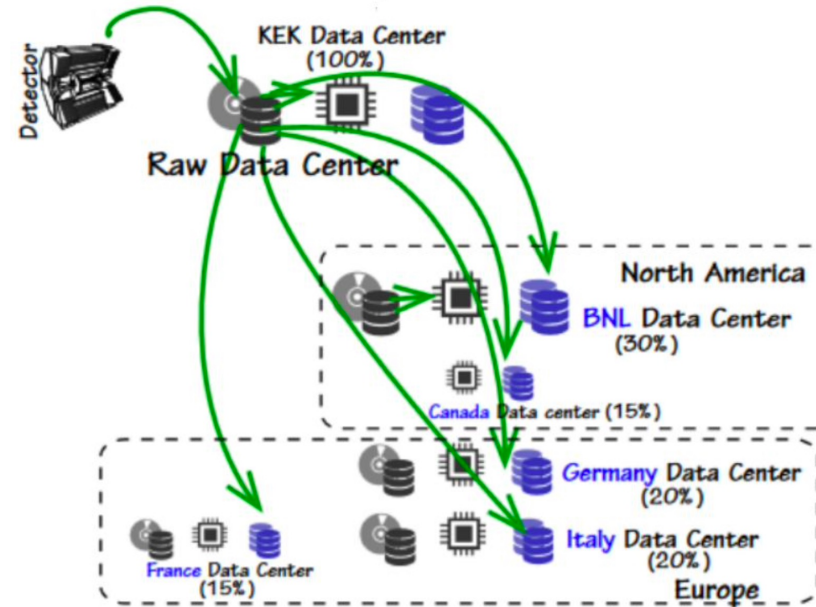
- <https://indico.cern.ch/event/1307338/contributions/5657149/attachments/2749230/4784743/DC24%20Workshop%2009-11-2023-1.pdf>

- What?

- Main test:
 - KEK traffic to RAW data centers
 - Aiming at **avg rate of 18.5 Gbit/s** or **200TB/day**
- Additional tests:
 - Site-to-site transfers (LHCONE, LHCOPN, SINET)
 - OIDC Tokens
 - Packet marking (python-flowd library for UDP fireflies)

- Why?

- Validate Belle II infrastructure and services (Rucio, FTS, SEs, Network) at max luminosity:
 - Belle II max luminosity is estimated to be 40TB per day (avg 3.7Gb/s) distributed by share (CC-IN2P3 is 15%)
 - For DC24, 5 times high-lumi throughput to absorb any possible bursts from data acquisition



- How?

A primary dataset of 8.000 of 5GB files for a total of 40TB has been created at KEK and registered in RUCIO.

A secondary data test with 10GB file could be created if needed.

The idea is to send and delete the same dataset multiple time up to 5.

SITE	Nomial Share %	Share (TB)	#5Gfiles	#10Gfiles
BNL	30	12.000	2.400	1.200
CNAF	20	8.000	1.600	800
KIT	10	4.000	800	400
DESY	10	4.000	800	400
IN2P3CC	15	6.000	1.200	600
Uvic	15	6.000	1.200	600
	100	40.000	8.000	4.000

6 TB for CC-IN2P3

BELLE II plans: Target rate by Raw Data Center

DC24 target throughput for BELLE II

Site	Country	#5G Files	Replica Factor	Total TB	Ingress (Gb/s)	Egress (Gb/s)
KEK	JP	8000	5,0	200	0,0	18,5
BNL	US	2400	5,0	60	5,6	0
CNAF	IT	1600	5,0	40	3,7	0
DESY	DE	800	5,0	20	1,9	0
KIT	DE	800	5,0	20	1,9	0
IN2P3CC	FR	1200	5,0	30	2,8	0
UIVc	CA	1200	5,0	30	2,8	0
Napoli	IT	TBD	TBD	TBD	TBD	TBD
SIGNET	SL	TBD	TBD	TBD	TBD	TBD

DC24 target throughput for CC-IN2P3

	Ingress (Gb/s)	Egress (Gb/s)
ALICE	3,2	0
ATLAS	96	66
CMS	45	36
LHCB	12	0
BELLE 2	2,8	??
TOTAL	159	102

Sur LHCOPN, sur LHCONE, sur LHCONE/LHCOPN

- **Pre-DC24 tests:**
 - *"Some extra traffic on your Network from KEK"*
 - 2 days since 16th Jan
 - For CC-IN2P3, the rate was $\sim 377\text{MB/s} = \sim 3\text{Gb/s}$ (target rate = 2.8Gb/s)
 - *"BELLE II reached its target rate"*
 - Setup (network/endpoints config, monitoring tools) reusable for DC24
- **Additional tests => BELLE II not ready yet**
 - Site-to-site transfers (BNL/IN2P3CC/CNAF/KIT in LHCOPN)
 - Tokens WLCG
 - Packet marking (python-flowd library for UDP fireflies)

BELLE II pre-DC24 tests: Network configuration

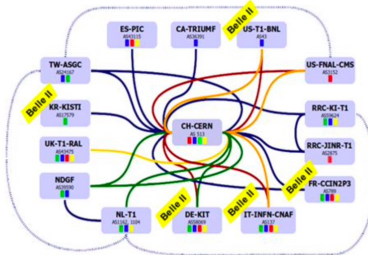
• Links:

- <https://docs.google.com/spreadsheets/d/1LqA2UBAuzyGpGINzu7T7vWX6Y9tiVzAA4nhi-u6Ht49w/edit#gid=0>
- <https://docs.google.com/spreadsheets/d/1LqA2UBAuzyGpGINzu7T7vWX6Y9tiVzAA4nhi-u6Ht49w/edit#gid=0>

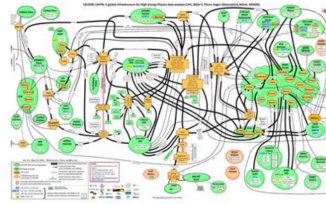
100G Global Ring via SINET



LHCOPN Optical infrastructure that can be used without jeopardizing resources



LHCONE L3 VPN Connecting all the major Data Centres



Storage Name	Site	Country	IPv4	IPv6	Storage Net IPv4	Storage Net IPv6	Note
1 KEK-TMP-SE	KEK	JP	202.13.206.0/23	2001:2f8:3e:cc21::/64	belle2-webdav-raw-data.cc.kek.jp (40 Gbps/node x4 nodes): 202.13.207.173/32 202.13.207.174/32 202.13.207.193/32 202.13.207.194/32	belle2-webdav-raw-data.cc.kek.jp (40 Gbps/node x4 nodes): 2001:2f8:3e:cc21::/64	SOURCE
2					belle2-webdav-analysis-data.cc.kek.jp (10 Gbps x2/node x4 nodes): 202.13.207.164/32 202.13.207.165/32 202.13.207.139/32 202.13.207.140/32	belle2-webdav-analysis-data.cc.kek.jp (10 Gbps x2/node x4 nodes): 2001:2f8:3e:cc21::29/128 2001:2f8:3e:cc21::2a/128 2001:2f8:3e:cc21::10/128 2001:2f8:3e:cc21::11/128	
3 BNL-TMP-SE	BNL	US	130.199.48.0/23, 130.199.185.0/24, 192.12.15.0/24, 192.33.128.0/24	2620:0:210::/48			DEST
4 CNAF-TMP-SE	CNAF	IT	131.154.128.0/17	2001:760:4205::/48			DEST
5 DESY-TMP-SE	DESY	DE	131.169.160.0/21, 131.169.191.0/24, 131.169.192.0/24, 131.169.80.0/24, 131.169.98.0/24	2001:638:700:1050::/64, 2001:638:700:1062::/64, 2001:638:700:10a0::/64, 2001:638:700:10bf::/64, 2001:638:700:10c0::/64			DEST
6 KIT-TMP-SE	KIT	DE	157.180.228.0/22, 157.180.232.0/22, 192.108.45.0/24, 192.108.46.0/23, 192.108.68.0/24	2a00:139c::/45			DEST
7 IN2P3CC-TMP-SI	IN2P3CC	FR	134.158.209.0/24	2001:660:5009:9::/64			DEST
8 UIVC-RAW-SE	UIVc	CA	206.12.154.0/24, (in 2024 142.104.133.0/24)	2607:f8f0:c10:70f3::/64, (in 2024 2607:f8f0:c10:104 /64)	rdc-redirector.belle.uvic.ca: with 10Gbps nodes behind (will be upgraded early next year)		DEST
9 Napoli-TMP-SE	Napoli	IT	90.147.67.0/24	2001:760:422a:137::/64			2001:760:422a:137::202:000/112 DEST
10 SIGNET-TMP-SE	SIGNET	SL	194.249.156.0/24	2001:1470:ff8a::/48			DEST

NETWORK

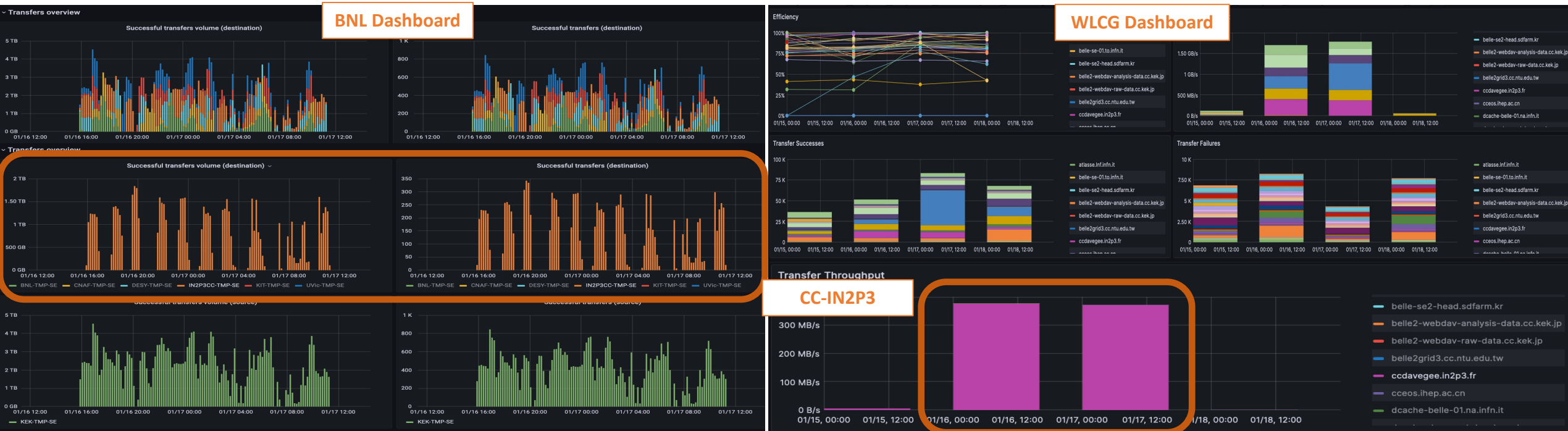
Storage Name	Site	Country	Technology	Storage Connection	WAN	PacketMarking	TOKEN Authentic	DAVS Full Test Path	DIR	People involved
1 KEK-TMP-SE	KEK	JP	STORM	Multiple WebDAV doors			TBD	davs://belle2-webdav-raw-data.cc.kek.jp:8443/belle/TMP/belle/test/DC24/	/belle/test/DC24/	
2										
3 BNL-TMP-SE	BNL	US	DCACHE	3 WebDAV doors. Each door 2x25Gbps	300Gb/s (site report)	Interested	TBD	davs://dcbldoor.sdcc.bnl.gov:443/pnfs/sdcc.bnl.gov/data/belldiskdata/TMP/belle/test/DC24/	/belle/test/DC24/	
4 CNAF-TMP-SE	CNAF	IT	STORM	WebDAV doors, each 50 Gb/s	240Gb/s (shared)	Interested	Testing IAM configured	davs://xfer-archive.cr.cnaf.infn.it:8443/webdav/belle/TMP/test/DC24	/belle/test/DC24/	Lucia Morganti, Daniele Lattanzio
5 DESY-TMP-SE	DESY	DE	DCACHE	Multiple WebDAV doors. Each door 2x10Gbps	100G to LHCONE	TBD	Testing IAM configured prod in configuration	davs://dcache-desy-webdav.desy.de:2880/pnfs/desy.de/belle/belle2/TMP/belle/test/DC24/	/belle/test/DC24/	Andreas Gelrich, Christian Voss
6 KIT-TMP-SE	KIT	DE	DCACHE	2 WebDAV doors. Each door 80Gbps	2x100G to GEANT 2x100G to CERN	Interested	TBD	davs://dcachewebdav-kit.gridka.de:2880/pnfs/gridka.de/belle/disk-only/TMP/belle/test/DC24/	/belle/test/DC24/	Matthias Schnepf
7 IN2P3CC-TMP-SE	IN2P3CC	FR	DCACHE	2 WebDAV doors. Each door 10Gb/s	LHCOPN: 100Gb/s LHCONE: 100Gb/s	Interested	Testing IAM configured	davs://ccdavegee.in2p3.fr:2880/belle2/disk/TMP/belle/test/DC24	/belle/test/DC24/	Aresh
8 UIVC-RAW-SE	UIVc	CA	xrootd	1 door	10Gb/s (site report)	Interested	TBD	davs://dcache-belle-01.na.infn.it/dpm/na.infn.it/home/belle/TMP/belle/test/DC24	/belle/test/DC24/	Marcus Ebert, Tristan Sullivan, Randy
9 Napoli-TMP-SE	Napoli	IT	DCACHE	4 WebDAV doors. Each doors 25Gbps	100G to LHCONE 20G to General IP	Interested	TBD	davs://dcache-ijc:2880/pnfs/ijc.si/belle/TMP/belle/test/DC24	/belle/test/DC24/	Silvio Pardi
10 SIGNET-TMP-SE	SIGNET	SL	DCACHE	TBD	25Gb/s SIGNET 100Gb/s EuroHPC	TBD	TBD	davs://dcache.ijc:2880/pnfs/ijc.si/belle/TMP/belle/test/DC24	/belle/test/DC24/	Marko Bracko

ENDPOINTS

BELLE II pre-DC24 tests: Monitoring tools

- Dashboards:

- https://monitoring.sdcc.bnl.gov/pub/grafana/d/belle2xfers/belle-ii-transfers-and-deletions?orgId=1&var-src_rse=KEK-TMP-SE&var-dst_rse=BNL-TMP-SE&var-dst_rse=CNAF-TMP-SE&var-dst_rse=DESY-TMP-SE&var-dst_rse=IN2P3CC-TMP-SE&var-dst_rse=KIT-TMP-SE&var-dst_rse=UVic-TMP-SE&var-activity=All&var-binning=10m&from=1705352005923&to=1705537964271
- https://monit-grafana-open.cern.ch/d/000000759/fts-transfers?orgId=16&var-group_by=dst_hostname&var-bin=1d&var-vo=belle&var-src_country=All&var-dst_country=All&var-src_site=All&var-dst_site=All&var-fts_server=All&var-staging=All&var-include=&from=1705276800000&to=1705622399000



Merci!