



Tata Institute of Fundamental Research
टाटा मूलभूत अनुसंधान संस्थान

CMS Grid Computing Facility at TIFR, India
Site Report of T2_IN_TIFR

The 6th Asian Tier Center Forum
Krabi, Thailand
Nov 21 - 24, 2022

Puneet Kumar Patel, Brij Kishor Jashal, Kajari Mazumdar, Gobinda Majumder

Facility



Compute node



Storage node



Firewall and switches

- 14 Racks, more than 300 servers
- Servers are fully compatible with both IPv4 and IPv6
- 8 + 8 Gbps WAN Link (recently)
- 150KVA UPS + 20 min. of power backup and 220KVA Isolation transformer
- UPS panel is common place to control and distribute power supply
- Dedicated Fire Suppression System
- 17TR centralized + **new 11TR in-Row cooling**
- 2U Twin configuration compute nodes
- 4U - 36 drives storage nodes with RAID 6
- GPU nodes with 8 x Nvidia v100 card and additional 16 different cards



Federation

Experiment Site Name	State	Country	VO name	Tier level	Status
T2_IN_TIFR	ACTIVE	India	cms	2	production
T3_IN_TIFRCloud	ACTIVE	India	cms	3	production
Experiment Site Name	State	Country	VO name	Tier level	Status

state	Host	Icons	OK	Wa	Un	Cr	Pd
UP	condor-ce01.indiacms.res.in		13	0	0	0	0
UP	cream-ce02.indiacms.res.in		13	0	0	0	0
UP	se01.indiacms.res.in		19	0	0	0	0

T2 storage Head Node (HN)

T3 computing HN

T2 computing HN

T2_IN_TIFR

- CVMFS and grid middleware on all the components

Storage:

- Software – DPM 1.13.0 on CentOS7, one head node with 53 disk nodes
- Head node and disk node (Apache, DPNS, DPM, SRM, XrootD)
- Protocol supported WebDAV, GridFTP, HTTPS, XrootD
- Storage : 11PB

Computing:

- HTCondorCE and Htcondor local batch on CentOS7
- Total cores reached up to 14K in production
- Production cores have been successfully tested
- Could not keep all of the cores available, due to the limited cooling
- Nearly 3.3K cores are active
- Arranging extra space and cooling – [work under progress](#)

GPU:

- Nvidia GPU with latest LHC software stack via CVMFS and CUDA
- Accessible via Jupyter notebook
- DNS, Accounting, Transfers, Monitoring services
- Commissioning and orchestration via Puppet

TIFR T3 for India-CMS Users

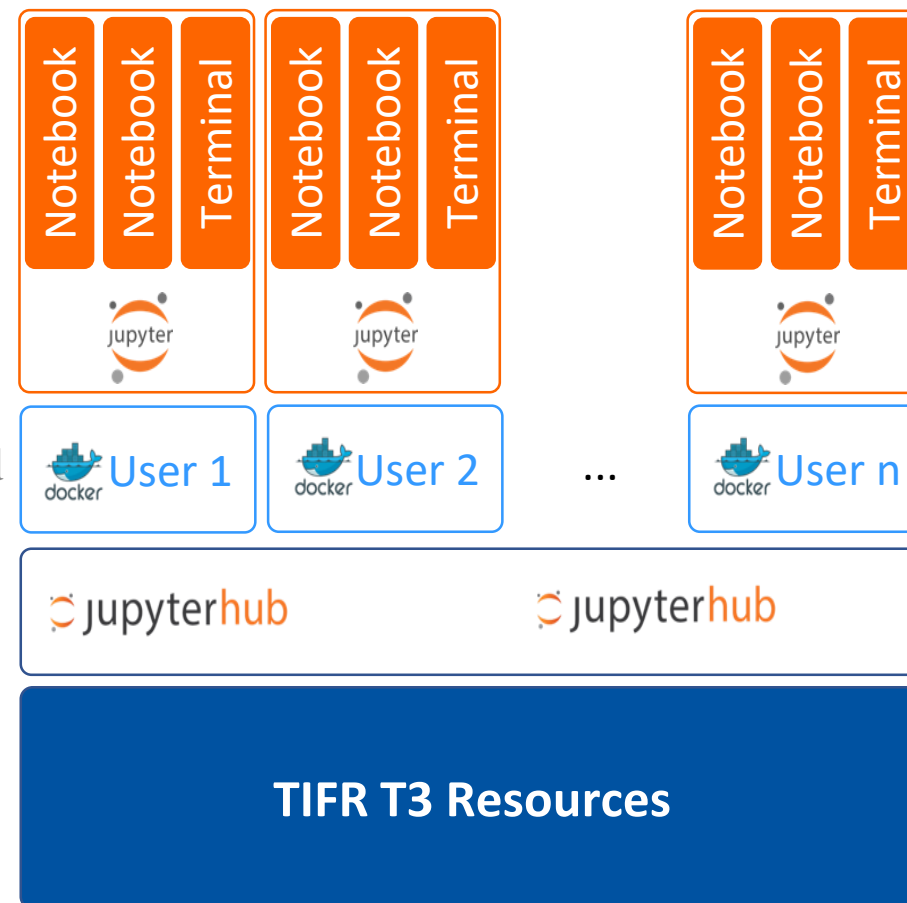
T3_IN_TIFRCloud

(detailed information can be referred from earlier ATCF meetings)

- This endpoint allowed us to scale the CPU as opportunistic resources including public clouds
- Dynamic Resources for WLCG commissioned in May 2017
- Collaboration with Microsoft Azure: Azure infrastructure with Grant of USD \$100,000 in terms of resources
- Development of tools and technologies for interfacing WLCG Grid with Azure (Grid ASCII Helper Protocol and Condor Annex)

Local T3

- Dedicated T3 condor cluster for user analysis
- Total number of Job slots ~ 370
- Multiple login machines
- Dedicated storage (separate to the T2 storage quota) ~300 TB and growing
- Support for JupyterHub and HTTPs based access to data
- Dedicated GPU resources and Local Gitlab instance.



Hardware Status

Year	Number of nodes	Size (rack unit)	CPU and frequency (dual socket)	CPU cores (HT enabled)	Memory	Types of storage	Storage capacity	Power Utilization
Storage node								
Till 2019	32	4U each	8C*@(2.4GHz and 2.53GHz), 12C@2.4GHz, 20C@2.2GHz, 32C@2.8GHz	~800	4 to 8GB per core	SSD + NL-SAS (3, 6, 8 and 10TB)	3PB	~30KVA
2020-22	54			~2488			11PB	~40KVA
Compute Nodes								
Till 2019	75	2U4N	16C, 48C, 56C, 72C	3500	2GB per core	SSD	10 GB per core	~60KVA
2020-22	179		56C, 72C, 128C	14000 (active 3384)				~80KVA

*C: CPU Cores

Pledge and Requirements

Federation	Year	Storage		CPU			Available Cooling (TR)**
		Pledge (TB)*	CMS % req.	Pledge (CPU cores)	CMS % req.	Active CPU	
T2	2022	11000	11.22%	14000	11.67%	3384	17TR
T2	2021	11000	11.96%	14000	13.08%	3384	17TR
T2	2020	3000	6.41%	3500	3.50%		17TR
T2	2019	3000	3.85%	3000	3.50%		17TR

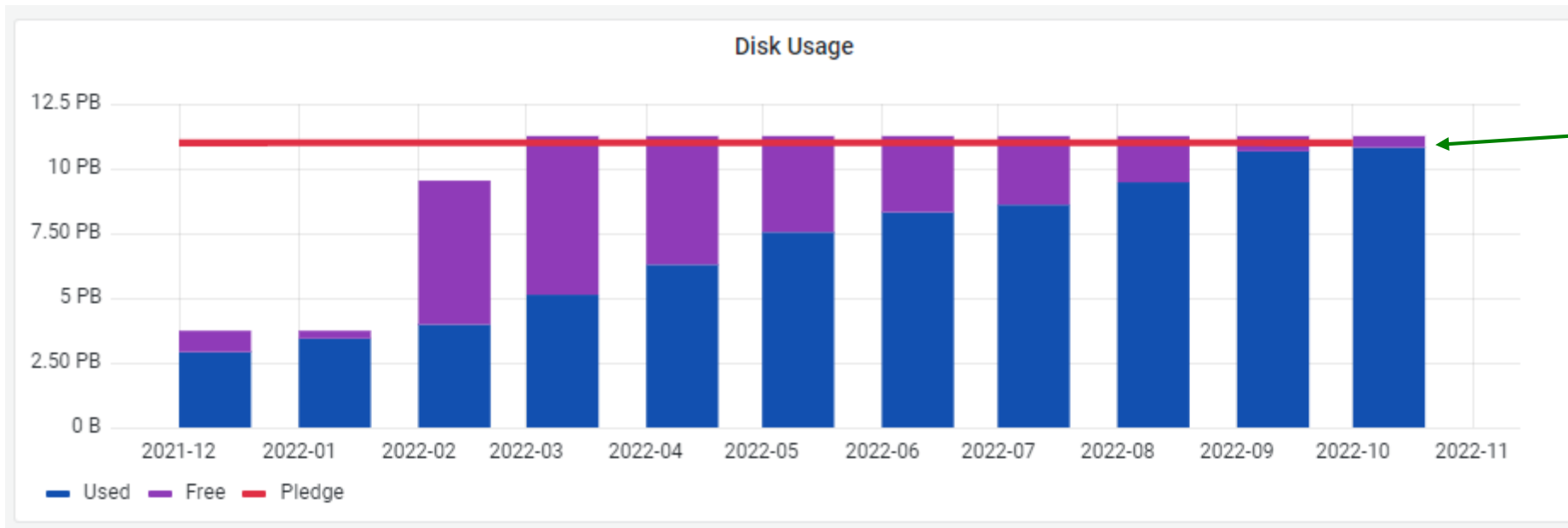
*TB: Tera Byte

**TR: Ton of Refrigeration

- **Required cooling is 35TR**
- Installation and commissioning of in-Row cooling units (11TR) are under progress

Storage Capacity Status

Federation	Country	Tier	Pledge Type	Year	CMS	CMS % of req
IN-INDIACMS-TIFR	India	2	CPU	2022	140000	11.67 %
IN-INDIACMS-TIFR	India	2	Disk	2022	11000	11.22 %
Federation	Country	Tier	Pledge Type	Year	CMS	CMS % of req



Almost 100% utilization

- Reminders are being circulated to the users to clear the data that is no more required

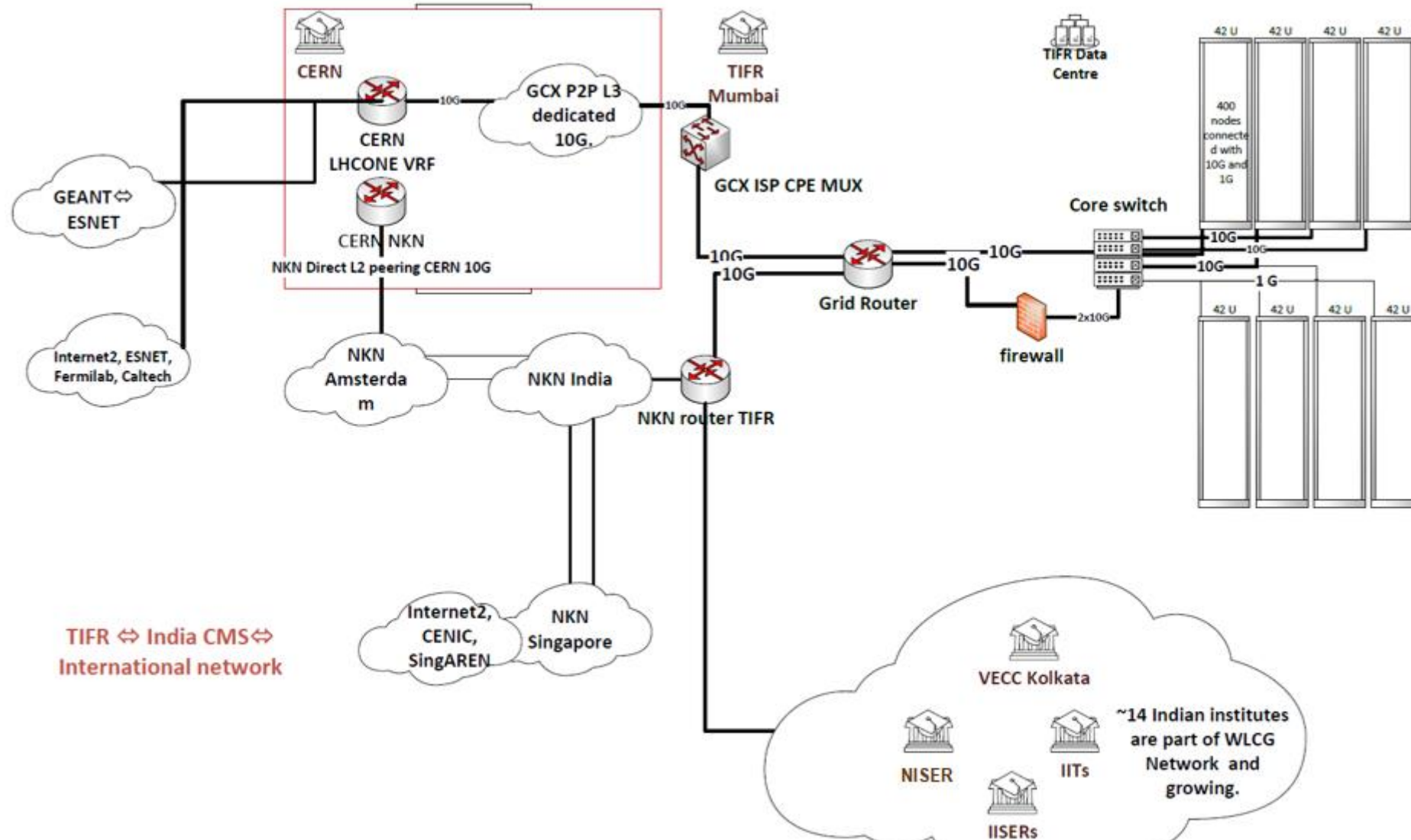
Source: <https://monit-grafana.cern.ch/>

WLCG Connectivity

- Major force behind the development are NKN and Indian R&E Network
- 2009: 1 G dedicated P2P link from TIFR \Leftrightarrow CERN
- 2012: Upgraded to 2Gbps
- 2014: Upgraded to 4Gbps
- 2015: Implemented fall back path using 10G shared TEIN link to Amsterdam
- 2015: CERN P2P link Upgraded to 8G
- 2015-16: Implemented LHCONE peering and L3VRF over NKN, all collaborating Indian institutes
- 2017: Upgraded to full 10G dedicated circuit till CERN
- 2018-19: NKN implemented CERN PoP with 10G link
- 2022: At present (8G + 8G) active links to LHC network.
- NKN L3 peering to US via Singapore and Amsterdam
- Network for Run III - Requested / Expected from NKN ~40 G, Mumbai to Geneva.

WLCG Connectivity cont..

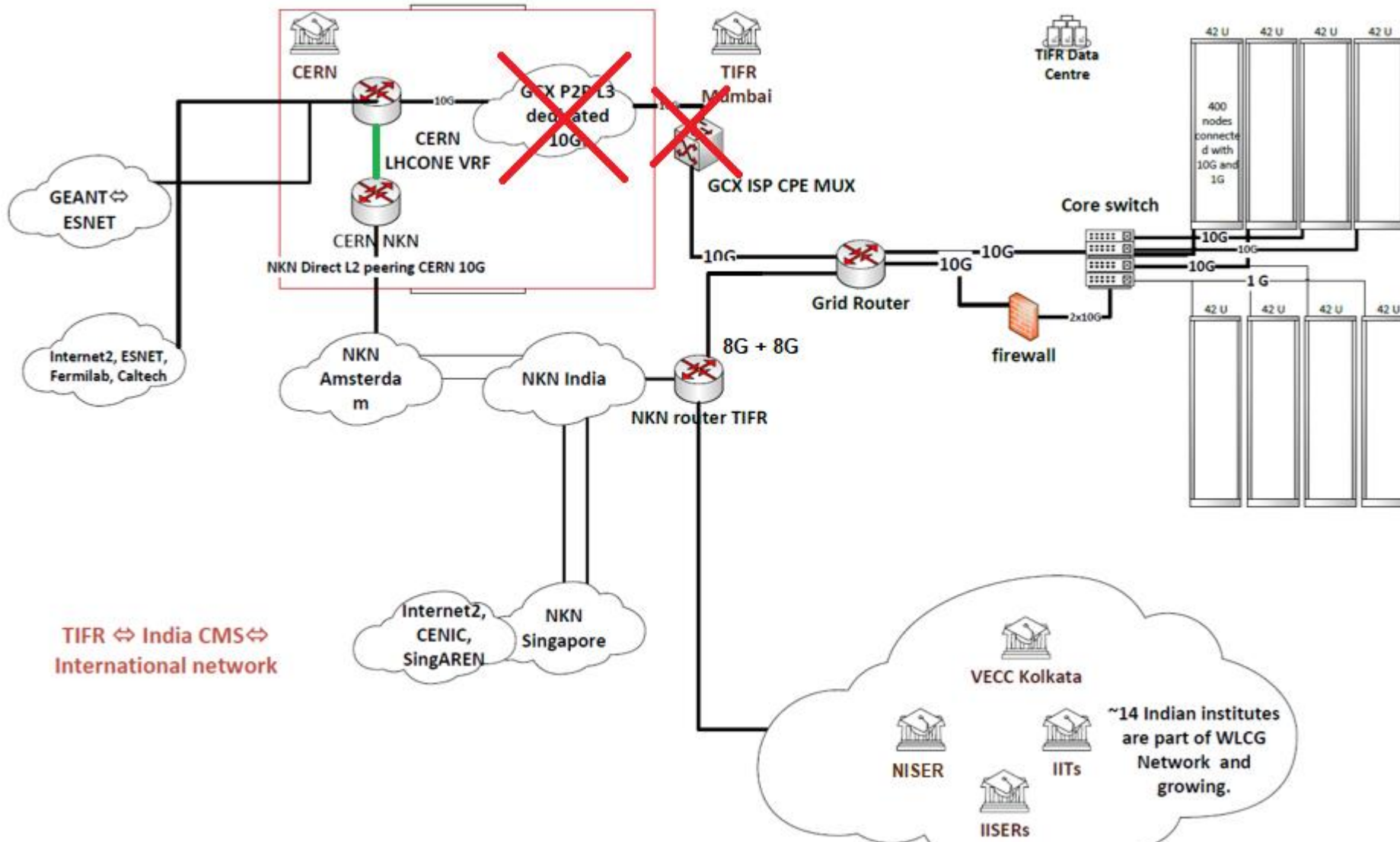
Connectivity till 2019



TIFR ↔ India CMS ↔
International network

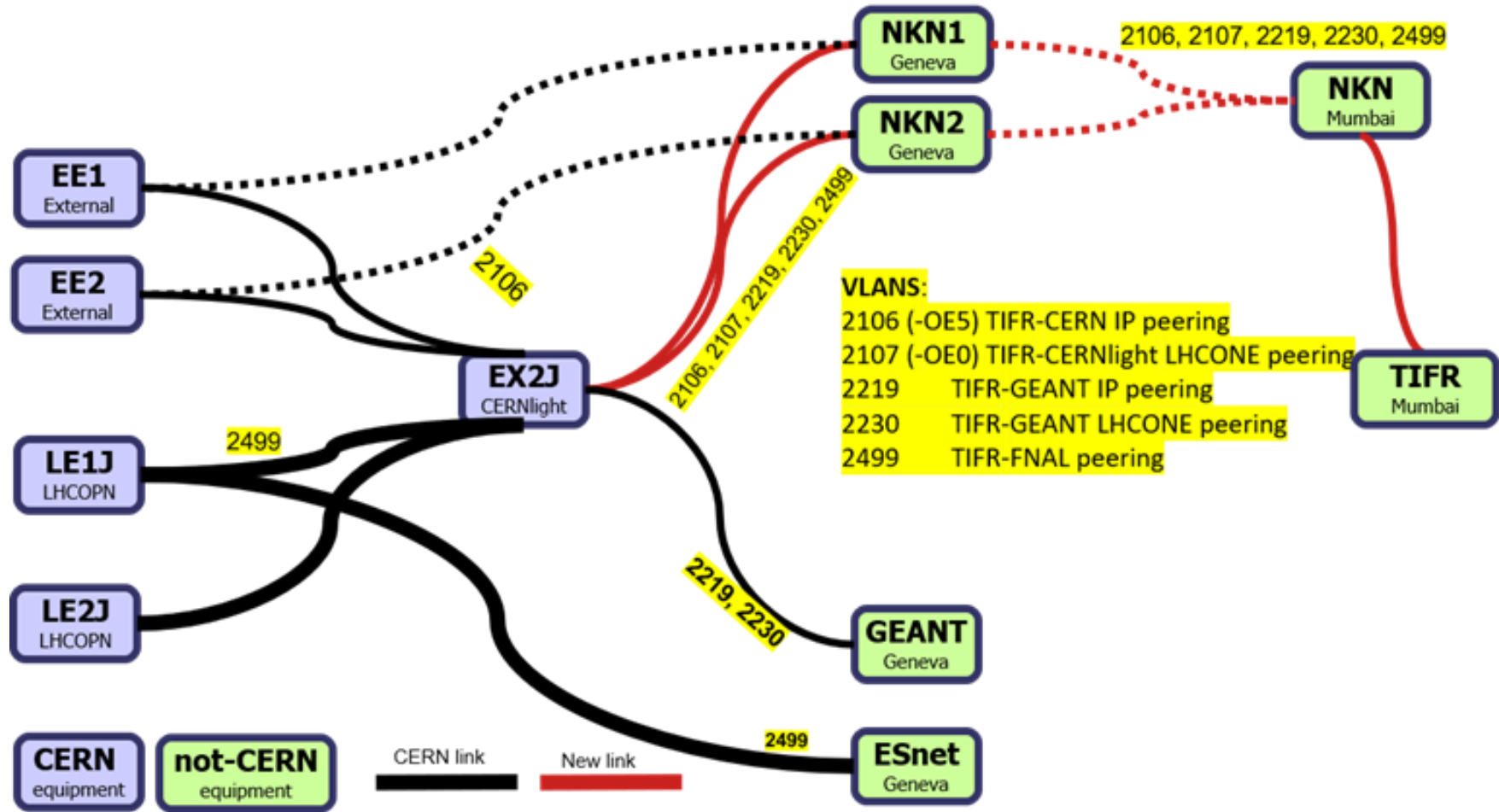
WLCG Connectivity cont..

Removed dedicated link, shifted to NKN in 2020



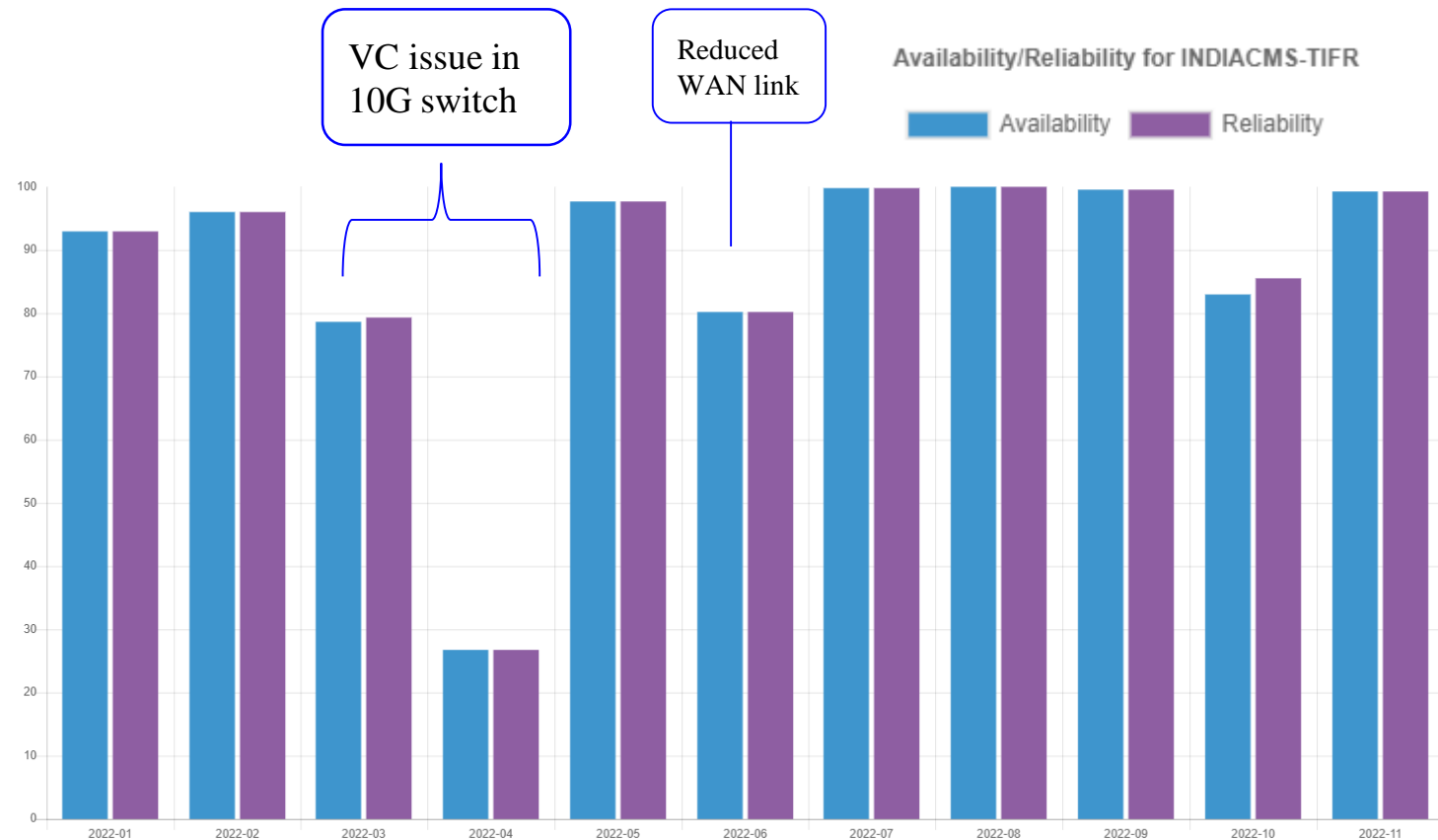
WLCG Connectivity cont..

Removed dedicated link, shifted to NKN in 2020



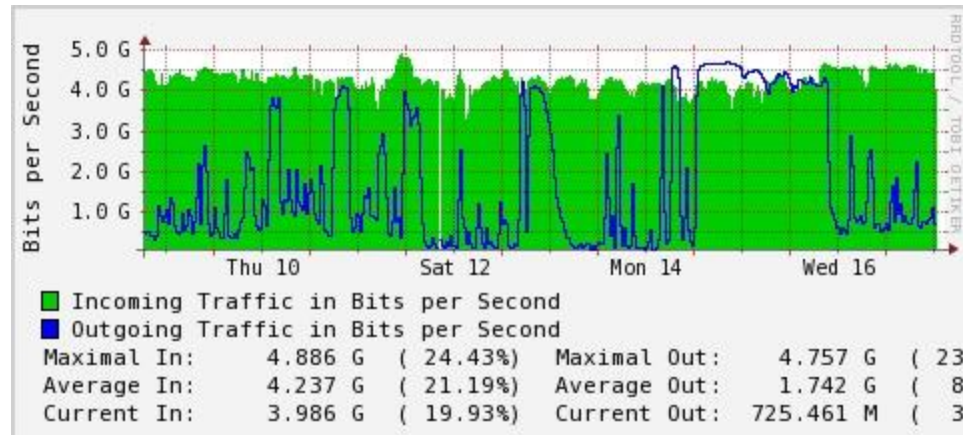
Challenges: Switch issue and Resolutions

- Virtual Chassis (VC) issue on new 10G switches, resolution received as new firmware after 13 months
- 19-Mar-2021: Juniper switch ports down issue, Case ID: 2021-0318-0193
- 27-Apr-2021: Server connectivity issue from outside, Case ID: 2021-0425-0131
- 25-Oct-2021: Server connectivity issue from outside (new case), CaseID: 2021-1022-346614 [previous case was closed without proper verification and acknowledgement]
- 28-Dec-2021: Packet drop issue, Case ID: 2021-1227-388391 [previous configuration is applied]
- **New firmware has been received and installed during Apr. 2022. Stable version - 19.1R3-S8**

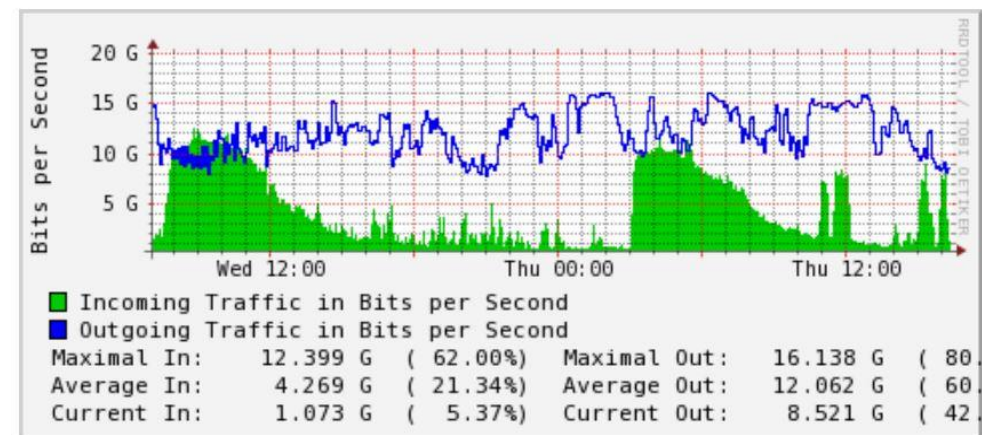


Challenges: WAN link issue from ISP

- Site suffered some power and cooling outages which reduced the performance
- Name resolution issue on DNS from ERNET
- CPU core count reduced because of current insufficient cooling issue
- It was informed to us on May 11, 2022 that out of two 5G, one undersea fibre cut incidence happened and Link have been reduced to 5Gbps. It was one of the link to Amsterdam and fiber cut happened near Jeddah



Link was limited to 5G



Link upgraded to 16G

Challenges and Resolutions

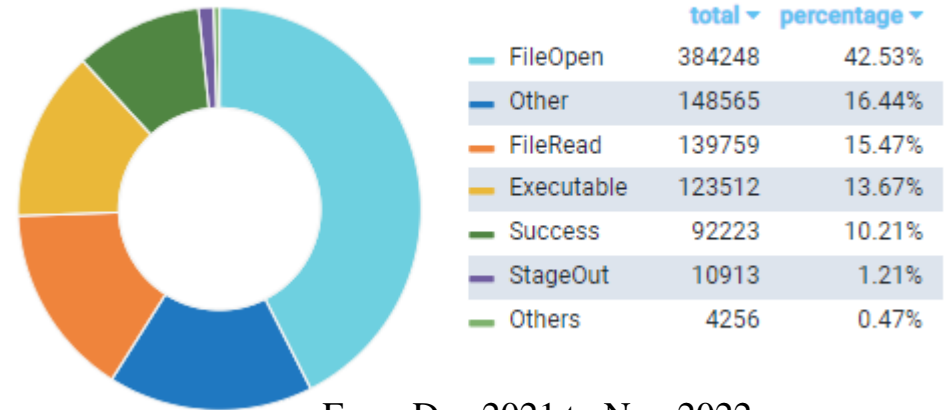
- ISP downgrade was one of the reason for transfer failure related issues
- Site went to waiting room several times due to high transfer failures
- Experienced multiple disk failures in storage servers due to heating issue
- Issue in host certificate

e.g "ERROR: Copy failed with mode 3rd push, with error: Transfer failed: failure: problem sending data: Certificate for doesn't match any of the subject alternative names: [puneet.patel@tifr.res.in]\n"

Solution: alternative name should be same as the host name

- Lost RAID controller card on dpms27, in 2021. **Half of the data have been recovered, out of 30TB**

Error types of failed jobs



From Dec 2021 to Nov 2022
<https://monit-grafana.cern.ch/>

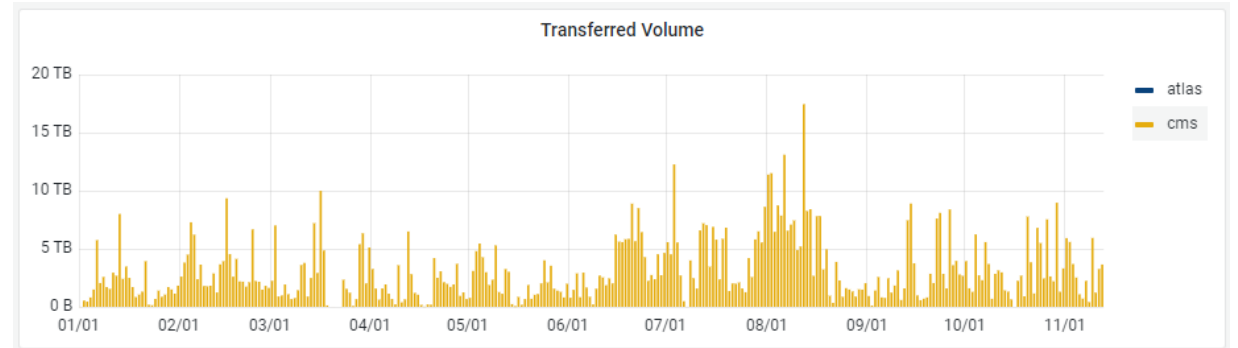
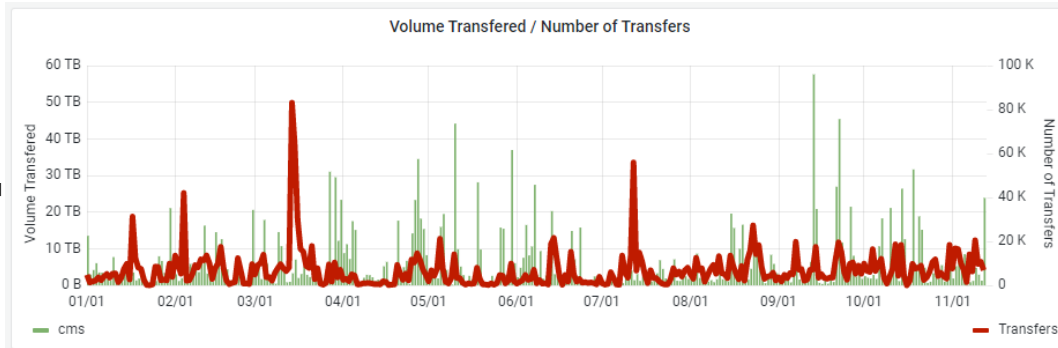
- Network congestion were responsible for around 57% of total error type of failed jobs

Transfer Status

FTS Transfers

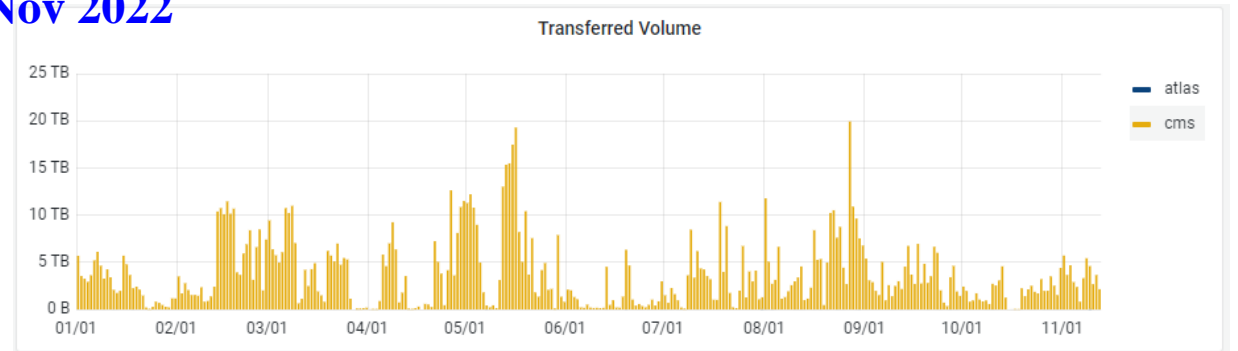
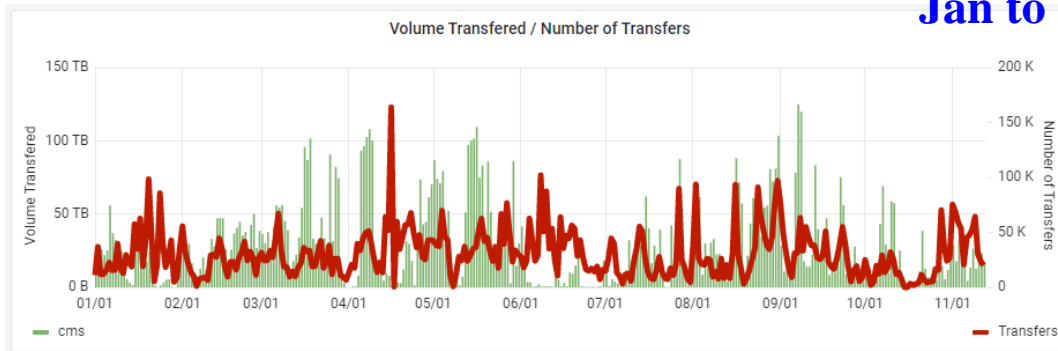
Xrootd Transfers

Uploads



Jan to Nov 2022

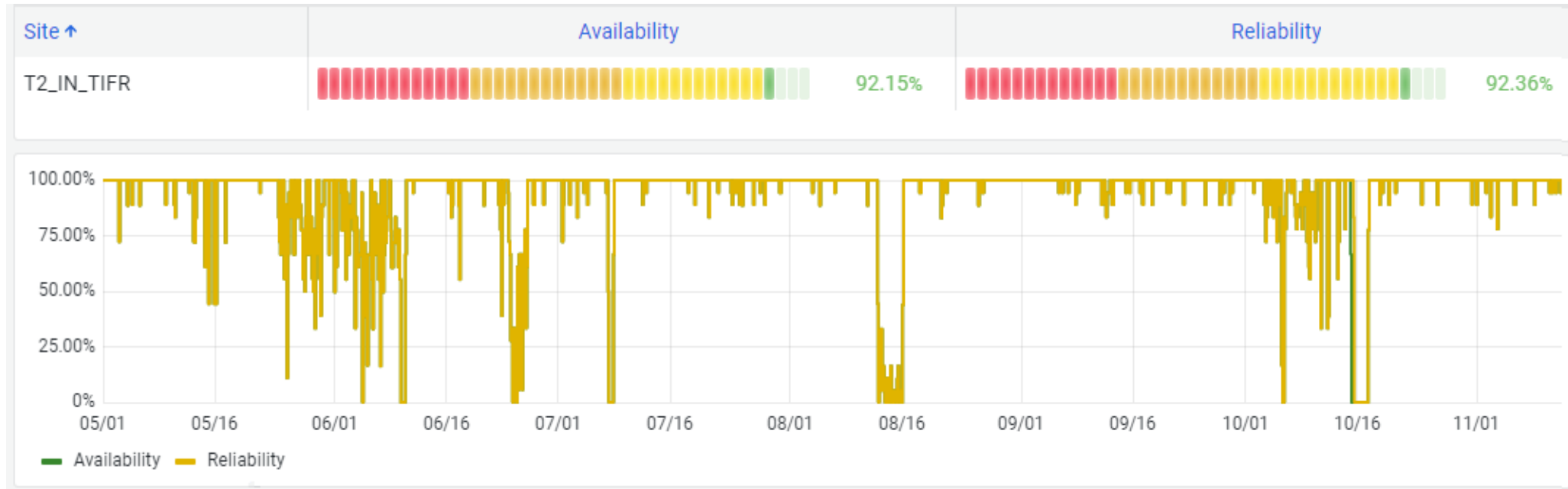
Downloads



- Significant amount of data transfers
- Bandwidth enhancement request is pending with ISP

A/R Report

Last 6 month status (May to Nov 2022)



- Stable Availability and Reliability
- Best efforts to keep above 90%
- Issues resolved whenever they occur

Summary

- TIFR, Mumbai hosts the Indian Tier2 for the CMS experiment under the aegis of WLCG
- It is part of the global CMS computing resources, active since 2009
- It provides the crucial support for the Indian participation in computing and analysis efforts of CMS
 - ~ 100 active users from collaborating Indian Institutes
 - 30% of storage earmarked for Indian community
 - Dedicated resources for T3 for analysis of Indian users
- The international connectivity is the backbone of grid and needs to upgrade as per the requirement
- The facility maintains average availability and reliability above 90% level via large number of tests and probes monitoring the health of system
- Presently there are two commissioned sites for CMS central computing (Production and analysis)
- Dynamic resources site - TIFR HEP Cloud commissioned in 2017

Thank You !