



# WLCG networks: LHCOPN and LHCONE

INFN CCR Workshop - 27 May 2021

[edoardo.martelli@cern.ch](mailto:edoardo.martelli@cern.ch)

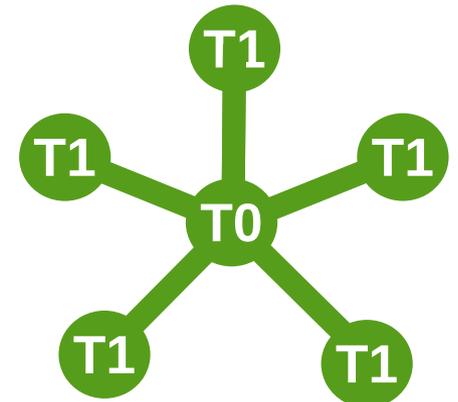
LHCOPN

LHC Optical Private Network

# LHCOPN

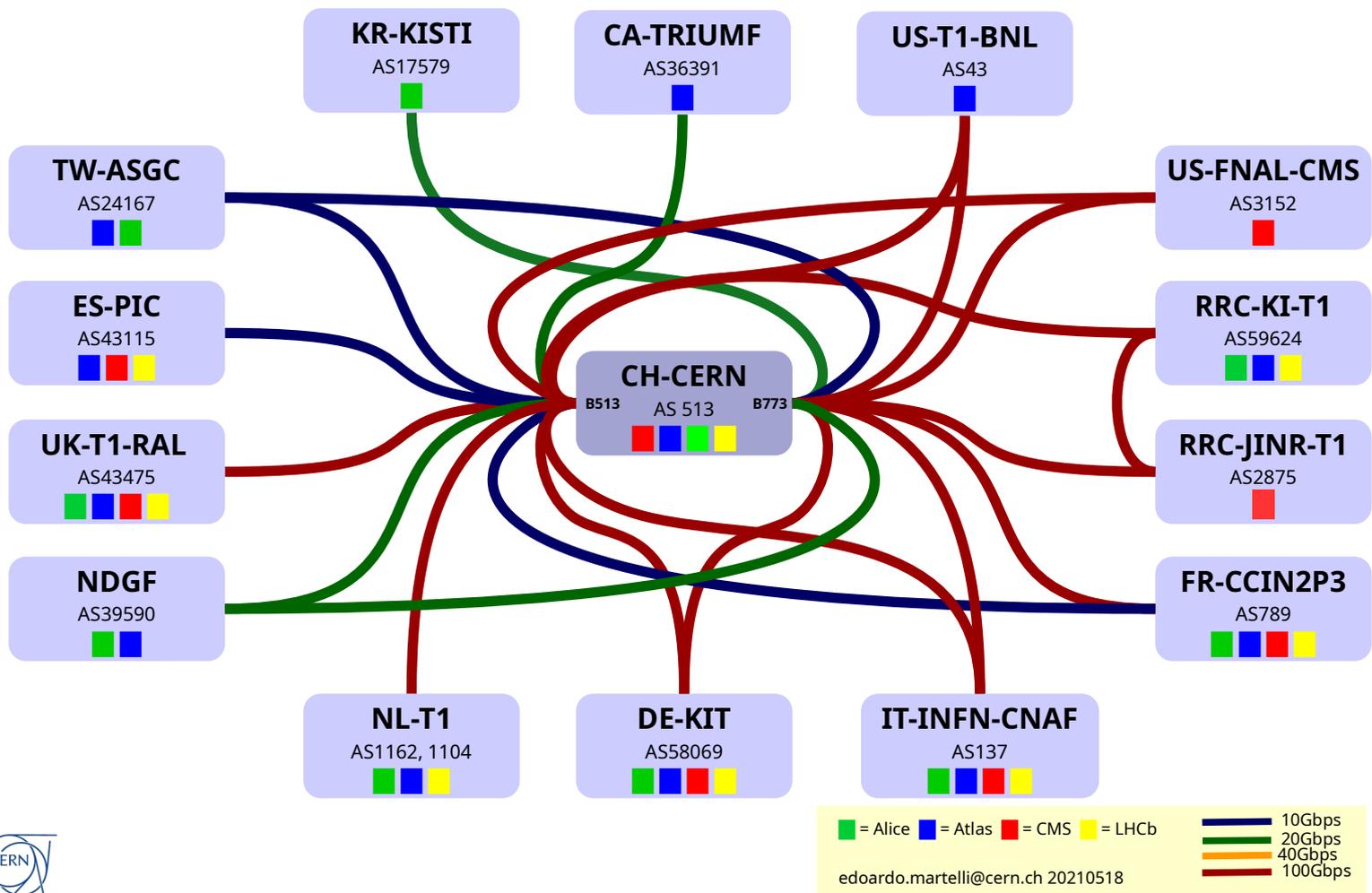
Private network connecting Tier0 to all the Tier1s

- Dedicated to LHC data transfers and analysis
- Star topology: all Tier1s connect to the Tier0
- Single and bundled long distance 10G and 100G Ethernet links
- Secured: only declared IP prefixes can exchange traffic
- BGP routing: communities for traffic engineering, load balancing
- Dual stack IPv4 and IPv6
- Tier1-Tier1 transit via Tier-0



LHCOPN

# LHCOPN Topology

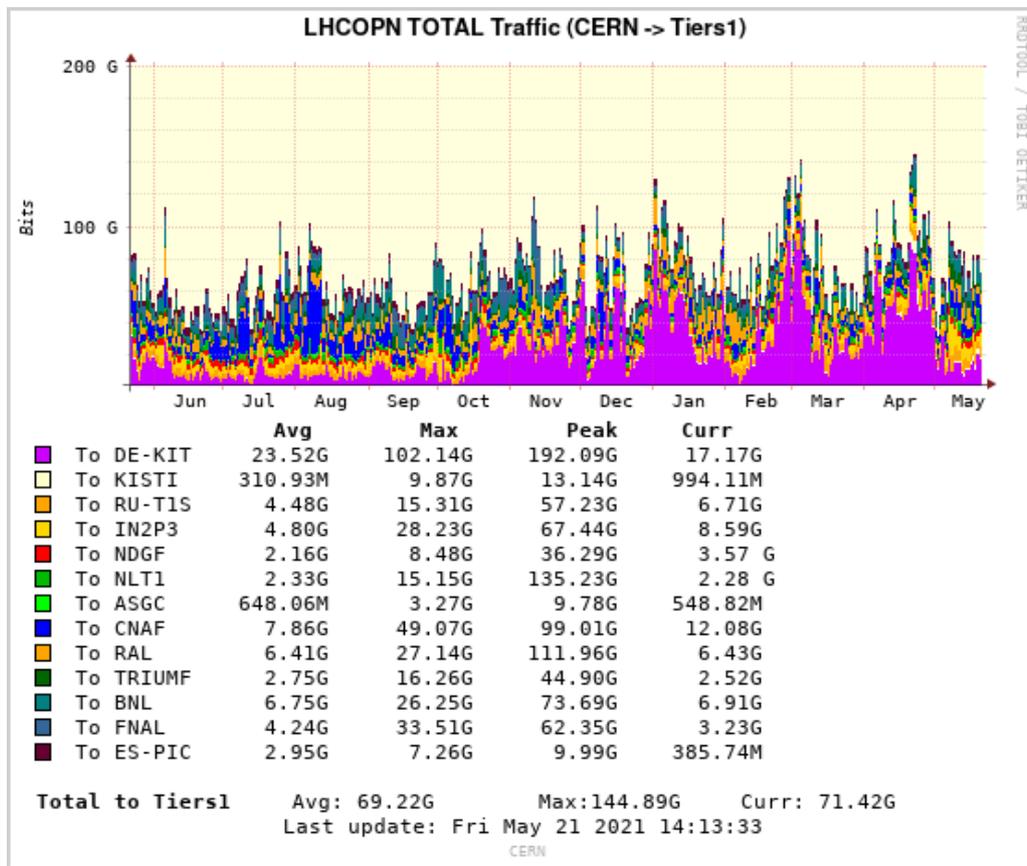


## Numbers

- 14 Tier1s + 1 Tier0
- 12 countries in 3 continents
- Dual stack IPv4 IPv6
- 1.3 Tbps to the Tier0



# LHCOPN Traffic – last 12 months



## Numbers

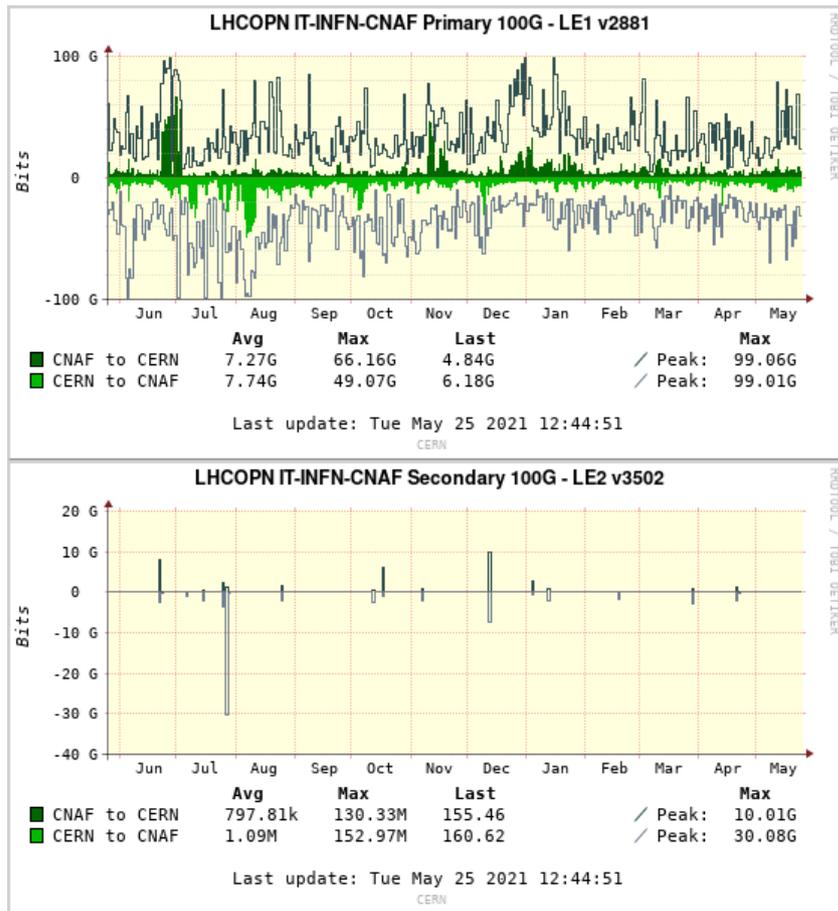
Moved ~273 PB  
in the last 12  
months

~5% less  
compared to  
previous year  
(286PB)

Ref: <https://netstat.cern.ch/monitoring/network-statistics/ext/?q=LHCOPN&p=LHCOPN&mn=00-Total-Traffic&t=Yearly>



# IT-INFN-CNAF LHCOPN links

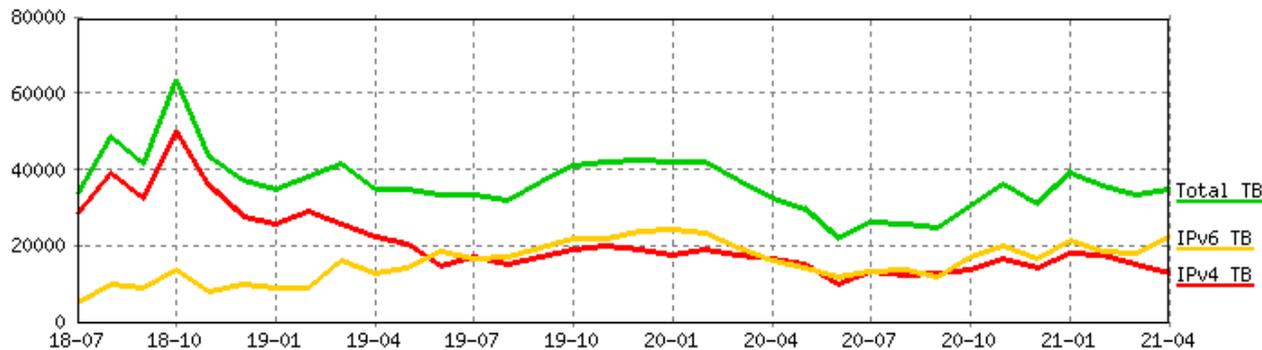


Two 100G links  
provided by  
GARR and  
GEANT

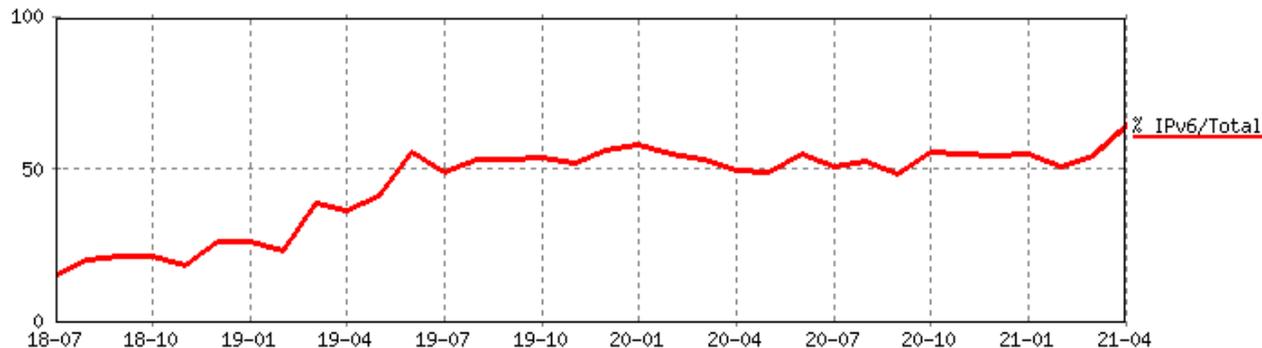


# IPv4 vs IPv6

IPv4 and IPv6 traffic volumes month by month



Percentage of IPv6 traffic over the total



## IPv6 vs IPv4 traffic

(as seen on the CERN LHCOPN/ONE border routers)

IPv6 is stable just above 50% for the last two years



# LHCOPN: latest developments

Majority of Tier1s connected with 100Gbps. More 100Gbps links coming before start of Run3 (ES-PIC, FR-CC-IN2P3).

Few Tier1s already at 2x100G

Testing 400G link between CERN and NL-T1 (SARA-NIKHEF)

LHCOPN capacity will be used for WLCG data challenges in preparation for Run4

LHCONE

LHC Open Network Environment

# LHCONE

Private network connecting Tier1s and Tier2s:

- Serving any LHC sites according to their needs and allowing them to grow
- Model: sharing the cost and use of expensive resources
- A collaborative effort among Research & Education Network Providers



# LHCONE services

**L3VPN (VRF):** routed Virtual Private Network - operational

**P2P:** dedicated, bandwidth guaranteed, point-to-point links  
– in development

**Monitoring:** monitoring infrastructure - operational



# LHCONE L3VPN service

Layer3 (routed) Virtual Private Network

Dedicated worldwide network connecting **Tier0, Tier1s and Tier2s** at high bandwidth

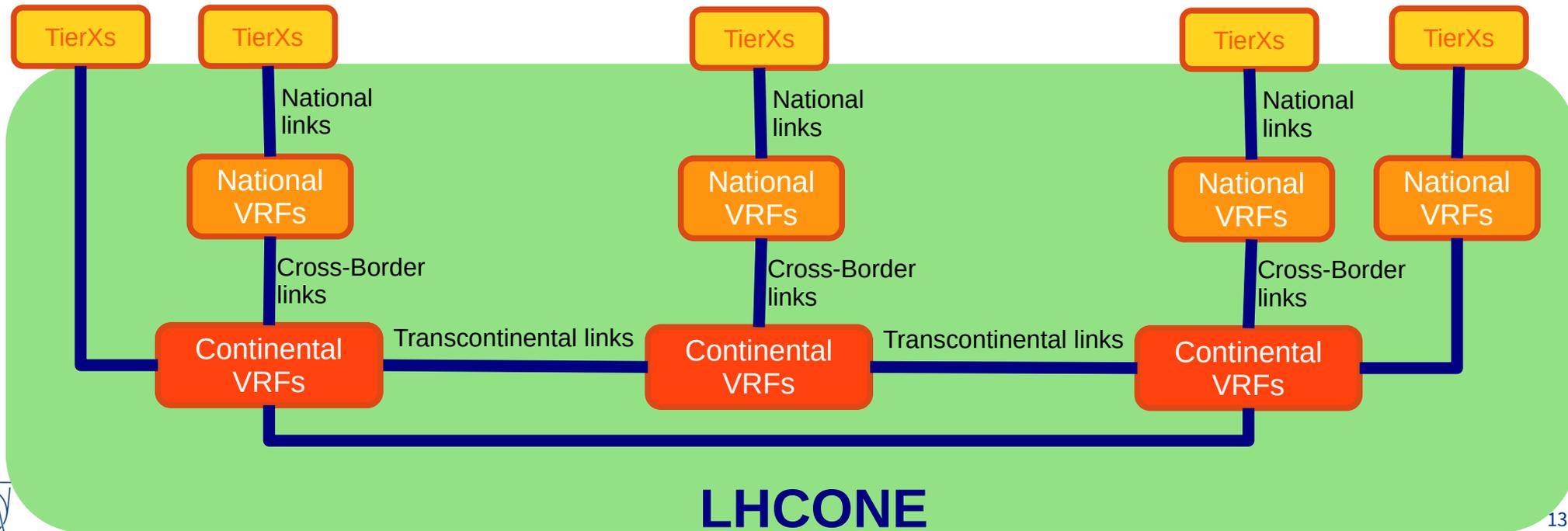
Private network: **Trusted traffic that can bypass perimeter firewalls**



# L3VPN architecture

- TierX sites connected to National-VRFs or Continental-VRFs
- National-VRFs interconnected via Continental-VRFs
- Continental-VRFs interconnected by trans-continental/trans-oceanic links

VRF = Virtual Routing Forwarding (virtual routing instance)

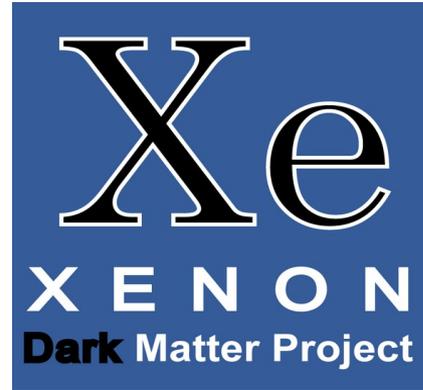


# L3VPN status

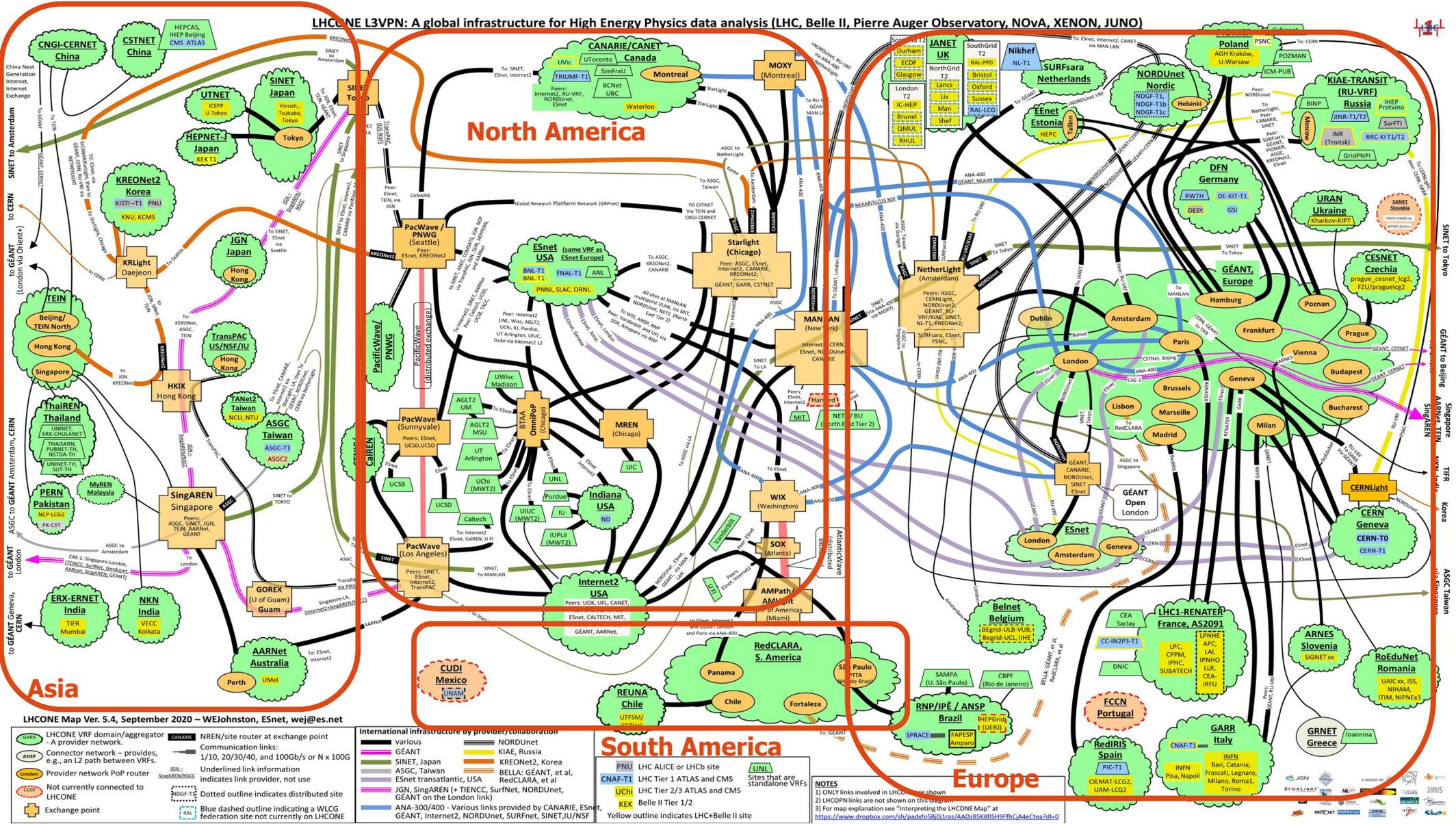
- **VRFs: 30 national and international Research Networks**
- **Connected sites: ~107 (of which 13 Tier1s, RAL coming soon)**
- Trans-Atlantic connectivity provided by ESnet, GEANT, Internet2, NORDUnet and SURFnet
- Trans-Pacific connectivity provided by ASGCnet, KREOnet, SINET, TransPAC
- Interconnections at Open Exchange Points including NetherLight, StarLight, MANLAN, WIX, CERNlight and others



# Open to other collaborations



# LHCONE L3VPN: A global infrastructure for High Energy Physics data analysis (LHC, Belle II, Pierre Auger Observatory, NoVA, XENON, JUNO)



## North America

## South America

## Europe

## Asia

LHCONE Map Ver. 5.4, September 2020 – WJJohnston, ESnet, wej@es.net

- LHCONE VRF domain/aggregator**
  - Green circle: LHCONE VRF domain/aggregator
  - Black circle: A provider network
  - Orange circle: Connector network – provides, e.g., an L2 path between VRFs.
  - Blue circle: Provider network PoP router
  - Red circle: Not currently connected to LHCONE
  - Yellow circle: Exchange point
- NREN/site router at exchange point**
  - Black line: Communication links; 1/10, 20/30/40, and 100Gb/s or N x 100G
  - Underlined link information: Indicates link provider, not use
  - Dotted outline: Indicates distributed site
  - Blue dashed outline: Indicating a WLCG federation site not currently on LHCONE
- International infrastructure by provider/collaboration**
  - Black line: Various
  - Green line: GÉANT
  - Blue line: SINET, Japan
  - Red line: Esnet transatlantic, USA
  - Orange line: JGN, SingAREN (+ TIENCC, SurfNet, NORDUnet, GEANT) on the London link
  - Yellow line: ANA-300/400 - Various links provided by CANARIE, ESnet, GEANT, Internet2, NORDUnet, SURFnet, SINET, IJ/N/NSF
  - Light blue line: NORDUnet
  - Light green line: KIAE, Russia
  - Light orange line: KREONet2, Korea
  - Light purple line: BELLA; GÉANT, et al
  - Light pink line: RedCLARA, et al
  - Light yellow line: JGN, SingAREN (+ TIENCC, SurfNet, NORDUnet, GEANT) on the London link
  - Light blue line: ANA-300/400 - Various links provided by CANARIE, ESnet, GEANT, Internet2, NORDUnet, SURFnet, SINET, IJ/N/NSF

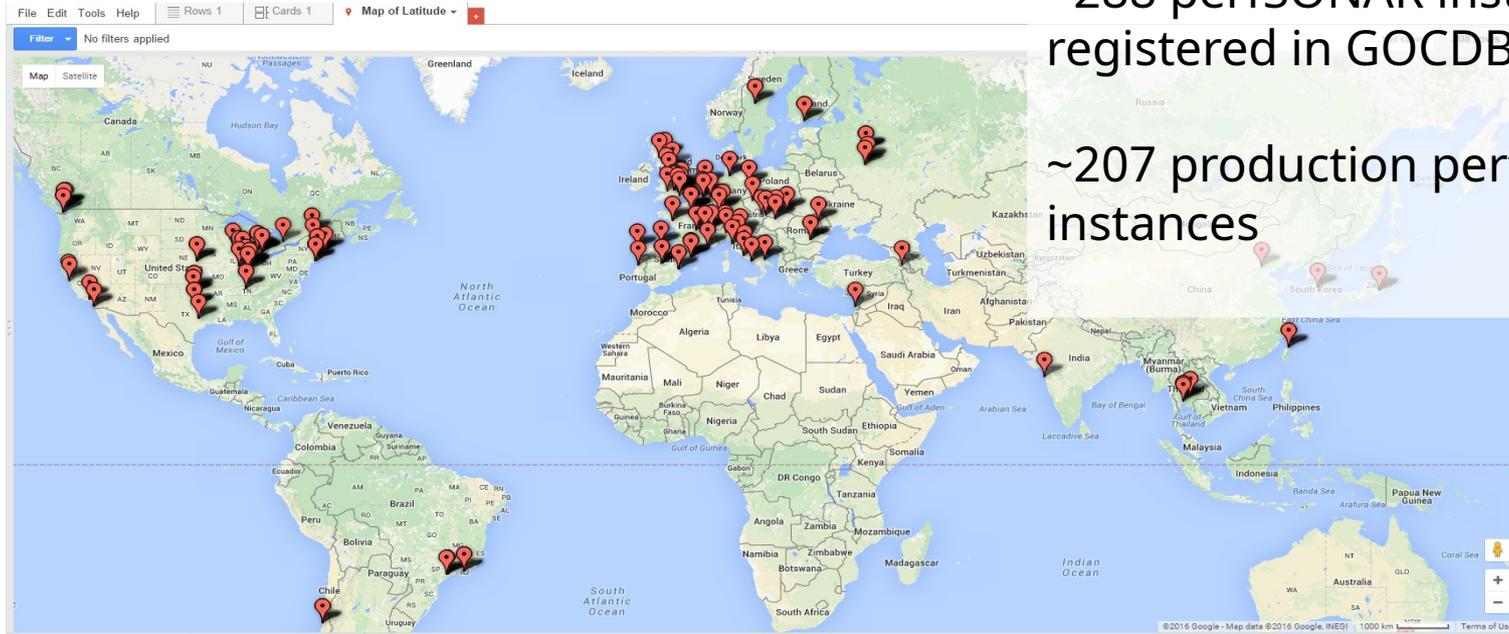
- PNU**: LHC ALICE or LHCb site
- CNAF-T1**: LHC Tier 1 ATLAS and CMS
- UCHI**: LHC Tier 2/3 ATLAS and CMS
- KER**: Belle II Tier 1/2
- UNL**: Sites that are standalone VRFs
- Sites that are standalone VRFs**
- Yellow outline**: Indicates LHC+ Belle II site

**NOTES**

- 1) ONLY links involved in LHCONE are shown
- 2) LHCOPN links are not shown on this diagram
- 3) For map explanation see "Interpreting the LHCONE Map" at <https://www.dropbox.com/sh/pafixf58j01raz/AADa85K8f9H9f9Hf4A4cTea?dl=0>



# Monitoring: perfSONAR



~288 perfSONAR instances registered in GOCDDB/OIM

~207 production perfSONAR instances

- Initial deployment coordinated by WLCG perfSONAR TF
- Commissioning of the network followed by WLCG Network and Transfer Metrics WG

Slide credit: Shawn McKee, University of Michigan

# Monitoring: perfSONAR

Some useful dashboards:

- WLCG MadDASH
- Latency per area
- Throughput
- end2end performance
- IPv4 vs IPv6

# Monitoring: Looking Glass

Looking-glass to analyse the routing tables of the VRFs

Peering with these VRFs:

- ASGC AS24167
- CANARIE AS6509
- CERNlight AS20641
- ESnet AS293
- KREOnet AS17579
- NORDUnet AS2603
- GEANT AS20965 (Geneva and Frankfurt routers)
- RU-VRF AS57484

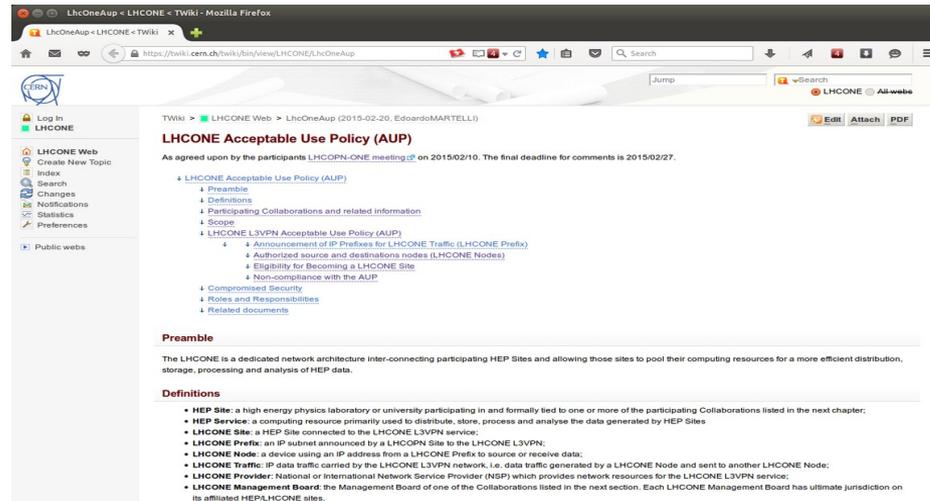
Link: <http://lhcone-lg.cern.ch/>



# LHCONE Acceptable Use Policy

The LHCONE AUP has been defined to regulate the utilization of the L3VPN service.

- Updated at the LHCONE meeting of March 2021
- <https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup>



The screenshot shows a web browser displaying the LHCONE Acceptable Use Policy (AUP) page. The page title is "LHCONE Acceptable Use Policy (AUP)". The content includes a preamble, definitions, and a list of key terms. The definitions section lists:

- **HEP Site:** a high energy physics laboratory or university participating in and formally tied to one or more of the participating Collaborations listed in the next chapter;
- **HEP Service:** a computing resource primarily used to distribute, store, process and analyse the data generated by HEP Sites;
- **LHCONE Site:** a HEP Site connected to the LHCONE L3VPN service;
- **LHCONE Prefix:** an IP subnet announced by a LHCONE Site to the LHCONE L3VPN;
- **LHCONE Node:** a device using an IP address from a LHCONE Prefix to source or receive data;
- **LHCONE Traffic:** IP data traffic carried by the LHCONE L3VPN network, i.e. data traffic generated by a LHCONE Node and sent to another LHCONE Node;
- **LHCONE Provider:** National or International Network Service Provider (NSP) which provides network resources for the LHCONE L3VPN service;
- **LHCONE Management Board:** the Management Board of one of the Collaborations listed in the next section. Each LHCONE Management Board has ultimate jurisdiction on its affiliated HEP/LHCONE sites.



# Database for LHCONE prefixes



- Network information related to LHCOPN, LHCONE and monitoring agreed to be added to CRIC
- A WHOIS route-set RS-LHCONE will be automatically updated from CRIC
- The route-set will be used by NREN and sites to build security routing filters

Next steps:

- implement records
- populate database, from LHCONE twiki page and LHCONE routing tables
- agree with WLCG operations on how to keep records up to date



# LHCONE latest news

- No new sites in the last year
- Traffic decreased 20% in the last year, due to LHC Shut Down
- JUNO collaboration just joined LHCONE
- Completed review of **AUP**: clarified some security aspects, roles and responsibilities
- perfSONAR: improved 100G mesh, using new Kibana dashboard
- NRENs are upgrading their networks to meet Run3's requirements
- Next meeting: 11-12 of October 2021 at CERN (in person, if possible)  
<https://indico.cern.ch/e/LHCOPNE47>



# WLCG data challenges

# LHC requirements for Run4 (2027)

The computing model of the experiments at HL-LHC will be different from what they have today: **ATLAS and CMS will produce**

**350PB/year/experiment**, to be exported in real time to Tier1s

- This would require **4.8Tbps from CERN to the Tier1s**, of which 1.25Tbps over the Atlantic
- **Larger Tier1s are supposed to get connected to CERN and to their Tier2s at 1Tbps** (1Tbps in from Tier0, 1Tbps out to Tier2s)
- Based on these targets, a plan for data challenges is being proposed for the years preceding Run4. First challenges will start at the end of 2021
- Data challenges will use the production infrastructure and will co-exist with production activities.
- Data challenges are being discussed in the DOMA-TPC sub-wg

# Reprocessing at HPCs

HPC will also be used for reprocessing

Tier1s will need to stage the data to be reprocessed at the HPC premises

The use case where an HPC would provide an allocation of 5k nodes (128 cores each) for many days capable to process 10kHz of events, implies **demonstrating the capability to stream 1Tbps of data into a HPC in 2027**

Intermediate targets should be defined for the coming years

# Data and Network challenges

Network challenges could consist in **demonstrating the capability to transfer an increasing volume of data over the next years to reach the production transfer target**, sustained for a few days, by the start of HL-LHC in 2027.

Foreseen **milestones as 15% of the target 2021, 35% in 2023, 60% in 2025 and 100% in 2027.**

This could be adjusted based on the growth plan of the NRENs

R&D activities

# Research Network Technology WG

Working on network research projects, with contributions not only from WLCG and NREN, but also from outside, like RFC editors, Linux kernel developers

Packet marking activity: considered many options: multiple addresses, IPv6 headers, MPLS. IPv6 flowlabel seems to be the more promising

- Proposed a packet marking schema for IPv6 flowlabel field
- Testing already started using iperf3 and other tools
- Targetting implementation in perfSONAR and XRootD

BitPattern	ScienceDomain	Application	Hdr Bit 12	Hdr Bit 13	Hdr Bit 14	Hdr Bit 15	Hdr Bit 16	Hdr Bit 17	Hdr Bit 18	Hdr Bit 23	Hdr Bit 24	Hdr Bit 29	Hdr Bit 30	Hdr Bit 31
xx10000000x000001xx	ATLAS	perfSONAR	x	x	1	0	0	0	0	x	0	1	x	x
xx01000000x000001xx	CMS	perfSONAR	x	x	0	1	0	0	0	x	0	1	x	x
xx11000000x000001xx	LHCb	perfSONAR	x	x	1	1	0	0	0	x	0	1	x	x
xx00100000x000001xx	ALICE	perfSONAR	x	x	0	0	1	0	0	x	0	1	x	x
xx10100000x000001xx	BelleII	perfSONAR	x	x	1	0	1	0	0	x	0	1	x	x
xx01100000x000001xx	SKA	perfSONAR	x	x	0	1	1	0	0	x	0	1	x	x
xx11100000x000001xx	LSSST	perfSONAR	x	x	1	1	1	0	0	x	0	1	x	x
xx00010000x000001xx	DUNE	perfSONAR	x	x	0	0	0	1	0	x	0	1	x	x

# NOTED

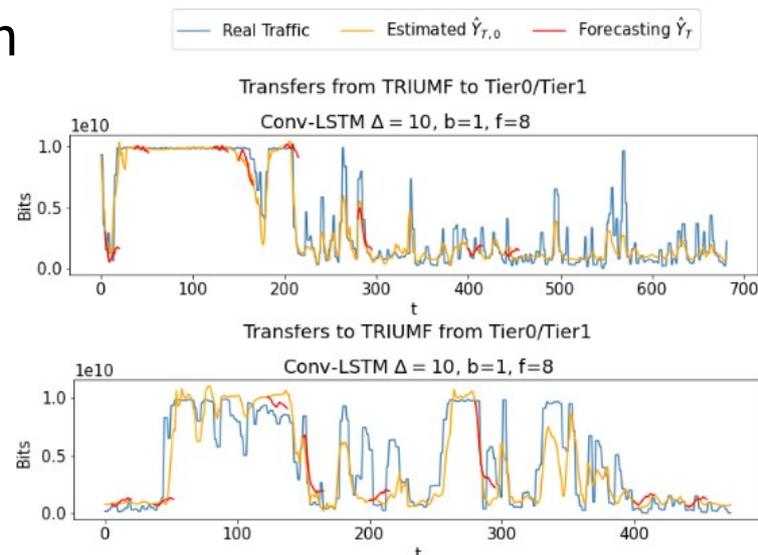
NOTED is a framework that can detect large FTS data transfers and trigger network optimizations to speed up the transfers executions

Two successful tests recently:

- CERN-PIC with LHCOPN-LHCONE load balancing
- CERN-TRIUMF with activation of larger bandwidth circuit

Developed new algorithm using machine learning to predict volume and duration of the transfer

Two papers submitted to vCHEP 2021



# SENSE: SDN for E2E Networked Science at the Exascale

## **End-to-End (network point of view)**

- DTN NIC to DTN NIC, across Science DMZ, WAN(s), Open exchange points (ideally)

## **Multi-domain**

- Multiple administrative domains, independent policies and AUP

## **Provisioning automation**

- Bring-up and management of services without interrupt-driven human involvement

## **Resource orchestration**

- Allocation and reservation of resources including compute, storage and network



**ESnet**  
ENERGY SCIENCES NETWORK

# Conclusions

# Summary

- LHCOPN: upgrading links to CERN to 100Gbps in preparation of Run3
- LHCONE: NRENs upgrading their infrastructure, CRIC for network resources
- WLCG is setting data challenges in preparation of Run4. The LHCONE community will contribute and support them
- Several on-going R&D activities to prepare for Run4: packet marking, NOTED

# Datacentre network architecture workshop

organized by CNAF and CERN

Pre-GDB of the 7-8 of June 2021

<https://indico.cern.ch/event/1028690/>



# References

LHCOPN: <https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome>

LHCONE: <https://twiki.cern.ch/twiki/bin/view/LHCONE/WebHome>

*Questions?*

*edoardo.martelli@cern.ch*

