



Infraestrutura de monitoramento da rede LHCONe do grid WLCG do experimento LHC do CERN

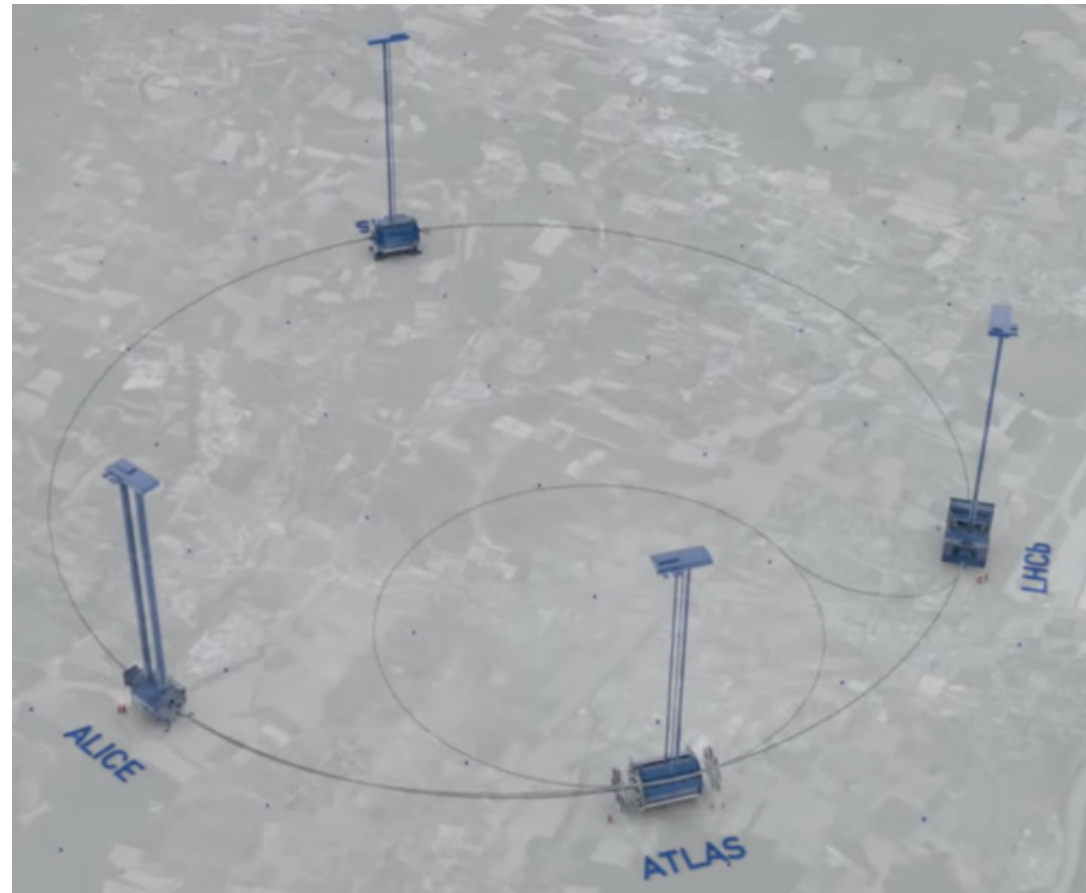
Pedro Diniz, Nilton Alves Jr., Márcio Portes, Renato Santana
phds@cbpf.br, naj@cbpf.br, mpa@cbpf.br, rsantana@cbpf.br

29/05/2015, GTER 39 – CEFET-RJ



LARGE HADRON COLLIDER (LHC)

- **Acelerador de partículas circular**
- **4 grandes experimentos**
 - 50 - 175 metros de profundidade
- **27 km de circunferência**
 - Partículas aceleradas a 99,9999991% da velocidade da luz (11.245 voltas por segundo)





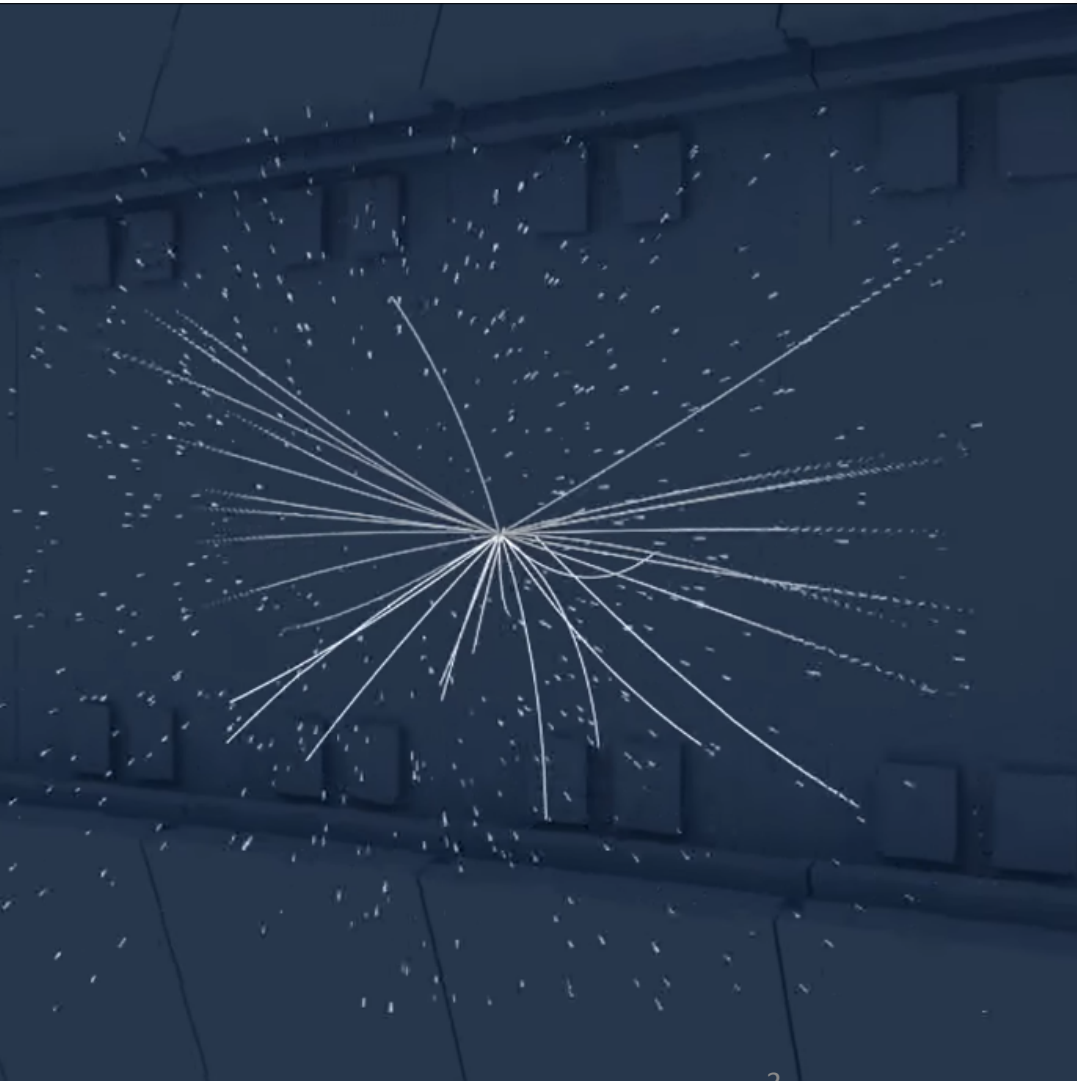
LARGE HADRON COLLIDER (LHC)

600 milhões de colisões a cada segundo

1 em um milhão de colisões é de interesse

Rápida pré-seleção eletrônica aprova 1 de 10.000 eventos e armazena em memória (100 GB/s)

15.000 núcleos de processamento selecionam 1 de cada 100 dos eventos pré-selecionados





WORLDWIDE LHC COMPUTING GRID

- **TIER 0**

- 73.000 cores

- Agregação de dados

- Reconstrução inicial de dados

- Cópia dos dados para fita

- Distribuição de dados

- **TIER 1 (11 centros)**

- Armazenamento permanente

- Reprocessamento

- Análise

- **TIER 2 (140 centros)**

- Simulação

- Análise de usuário final



PARTICIPAÇÃO BRASILEIRA NO LHC

- **Pesquisadores brasileiros participam em todos os experimentos do LHC:**
 - Desenvolvimento de dispositivos eletrônicos
 - Análise de experimentos
 - Física teórica
 - *Sites* Tier2 do WLCG
- **WLCG sites no Brasil:**
 - **CBPF** – LHCb (700 cores / 300 TB)
 - **USP (SAMPA)** – ALICE (1680 cores / 300 TB)
 - **UNESP (SPRACE)** – CMS (1200 cores / 800 TB)
 - **UERJ (HEPGRID)** – CMS (600 cores / 150 TB)



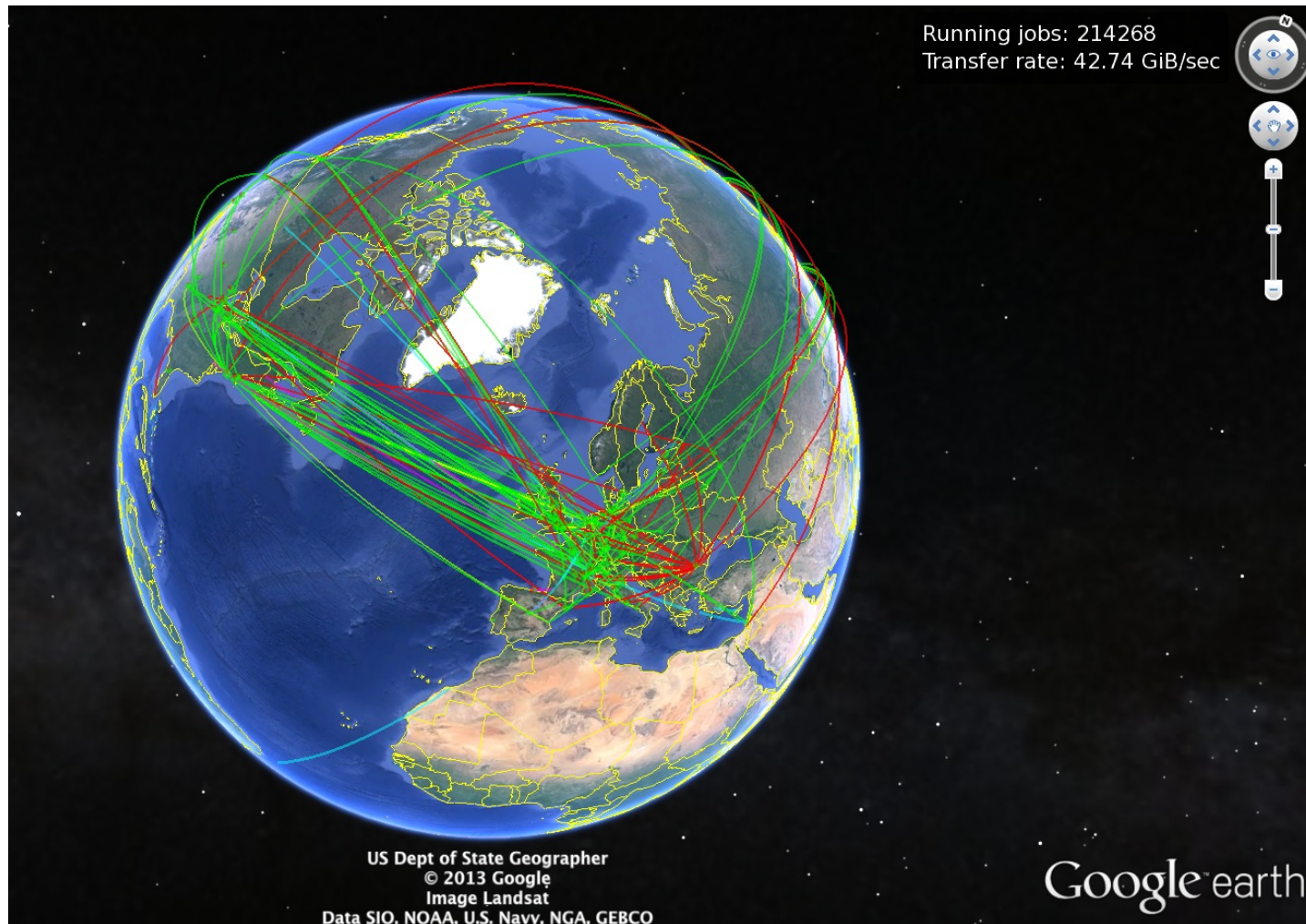
CBPF





PAPEL DAS REDES NO WLCG

- As redes de computadores são um componente essencial do WLCG





REDES NO WLCG

- Duas redes de dados privadas e dedicadas:

LHCOPN

LHC Optical Private Network
(Tier0 - Tier1s)



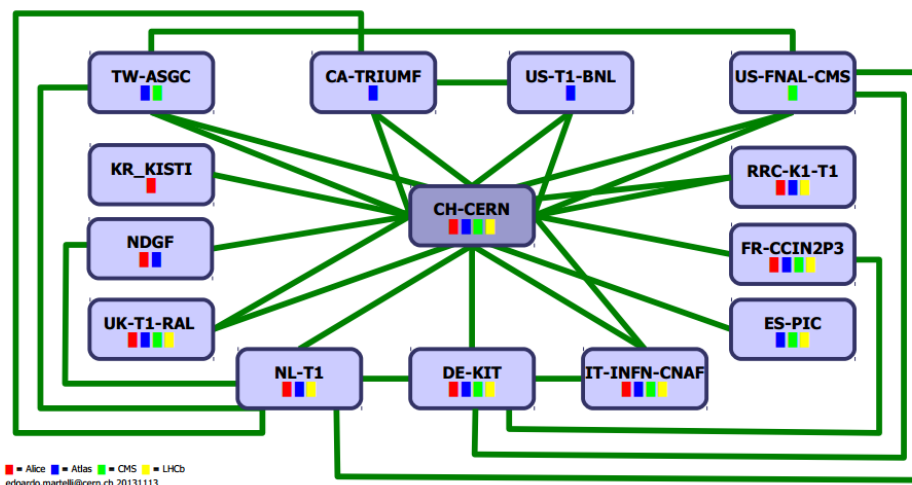
LHC Open Network Environment
(Tier1s - Tier2s)



LHC OPTICAL PRIVATE NETWORK

LHCOPN

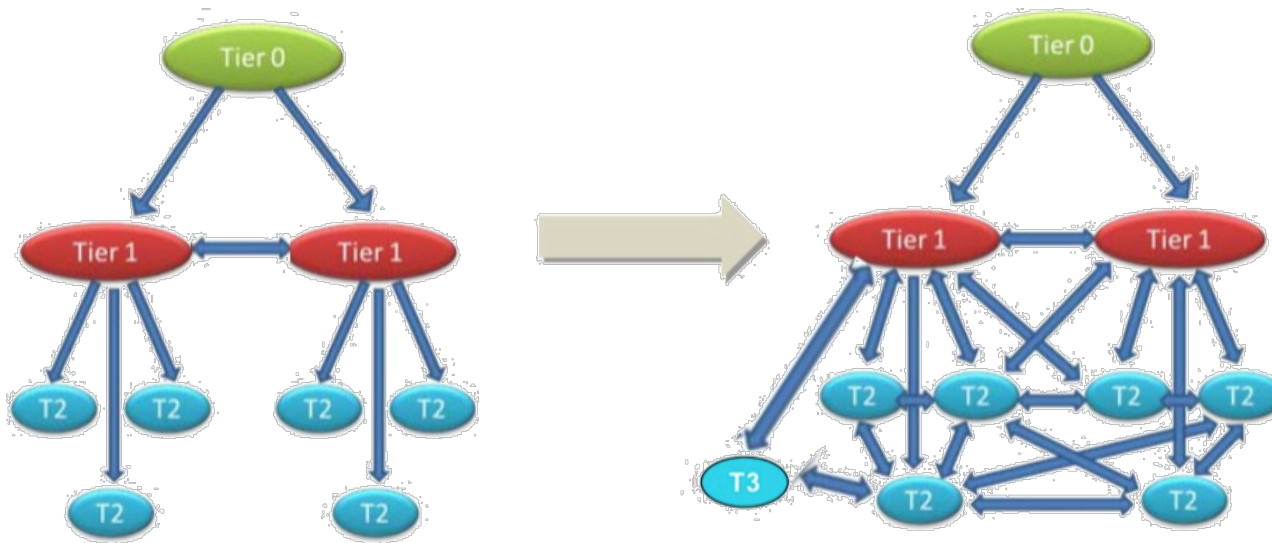
- Rede privada que conecta *Tier0* e *Tier1s*
- Enlace de alta largura de banda dedicada
- Arquitetura altamente resiliente (estrela e *mesh*-parcial)



- Roteamento BGP: *communities* são utilizadas para engenharia de tráfego
- **Segurança:** somente os **prefixos IP declarados** podem **trocar tráfego**

LHC OPEN NETWORK ENVIRONMENT

- Novo modelo de computação



- Uso de *storage* mais dinâmico
- Carga reduzida nos Tiers 1
- Necessidade de rede mais rápidas e previsíveis para conexão entre Tier 1s e Tier 2s



LHC OPEN NETWORK ENVIRONMENT

- **Conceitos Principais:**

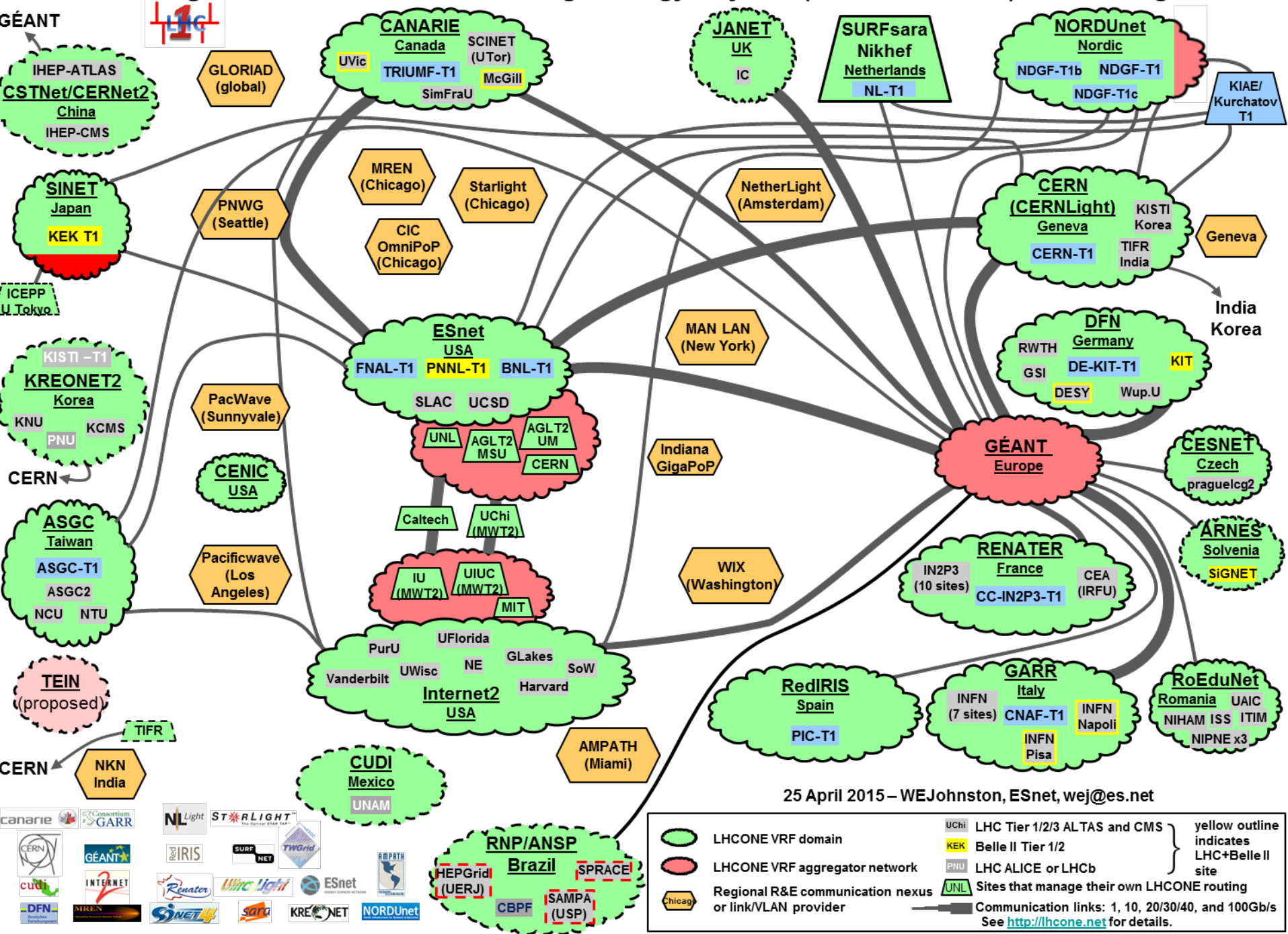
- **Compartilhamento de custo e utilização de recursos caros** (enlaces transoceânicos)
- **Esforço colaborativo** entre diversas **redes de educação e pesquisa** por todo o mundo (e.g. GÈANT, ESnet, Internet2)
- **Segregação de tráfego:** não há disputa com outros tipos de tráfego, recurso **alocado e financiado pela comunidade de física de altas energias**
- Carga reduzida nos Tiers1



LHC OPEN NETWORK ENVIRONMENT

- **O que é?**
 - **Rede privada** sobreposta à Internet através de **VRFs**, suportada por:
 - Mais de 15 **redes de educação e pesquisa nacionais e internacionais**
 - Interconectada através de **Pontos de Troca de Tráfego pelo mundo**, incluindo NetherLight, StarLight, MANLAN, Internet2, ESnet, entre outros
 - **Roteamento multi-domínio** como IP normal, mas com **restrições nos prefixos IP advertidos**, com **enlaces de acesso transatlânticos dedicados**
 - Mais de 50 sites conectados
 - 10 Tier 1s
 - 45 Tier 2s
 - Em operação desde 2012

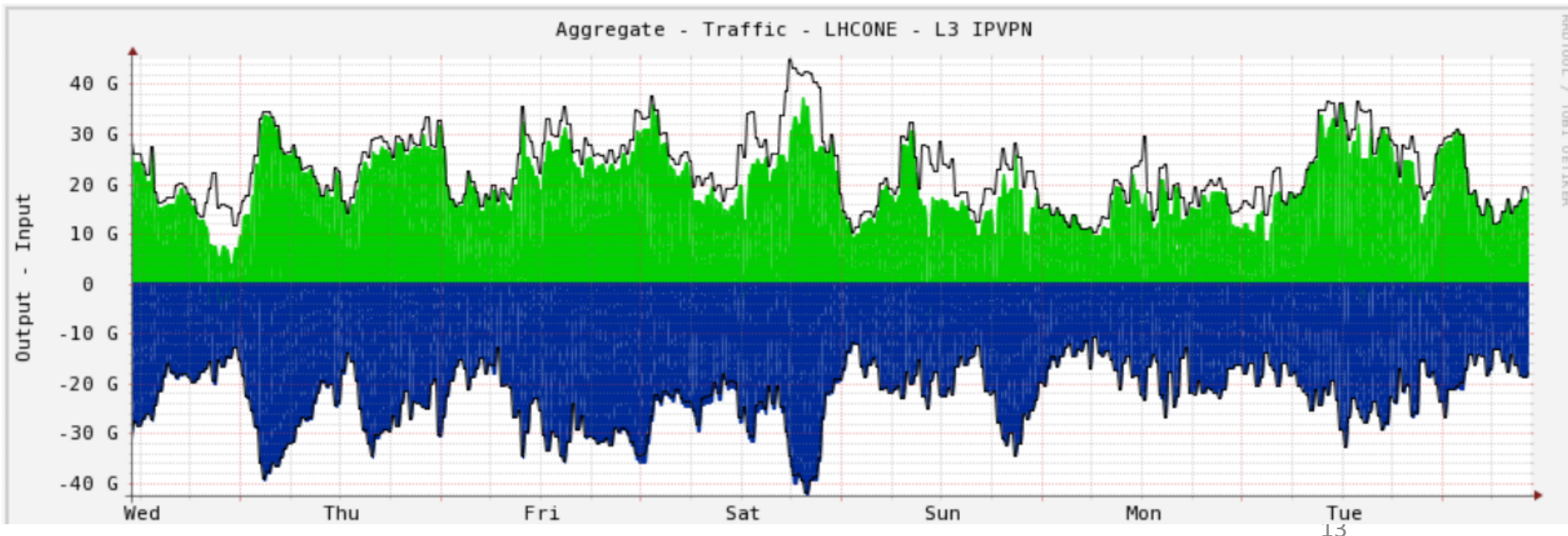
LHCONE: A global infrastructure for the High Energy Physics (LHC and Belle II) data management





LHC OPEN NETWORK ENVIRONMENT

- **Exemplo – LHCONE na Europa:**
 - **Tráfego agregado** de todas as redes de pesquisa europeias e os *peers*:
 - Tráfego médio: ~ 25 Gbps
 - Tráfego de pico sustentado: ~ 35 Gbps
 - Tráfego transatlântico: ~ 25 Gbps





LHC OPEN NETWORK ENVIRONMENT

- **Serviços:**

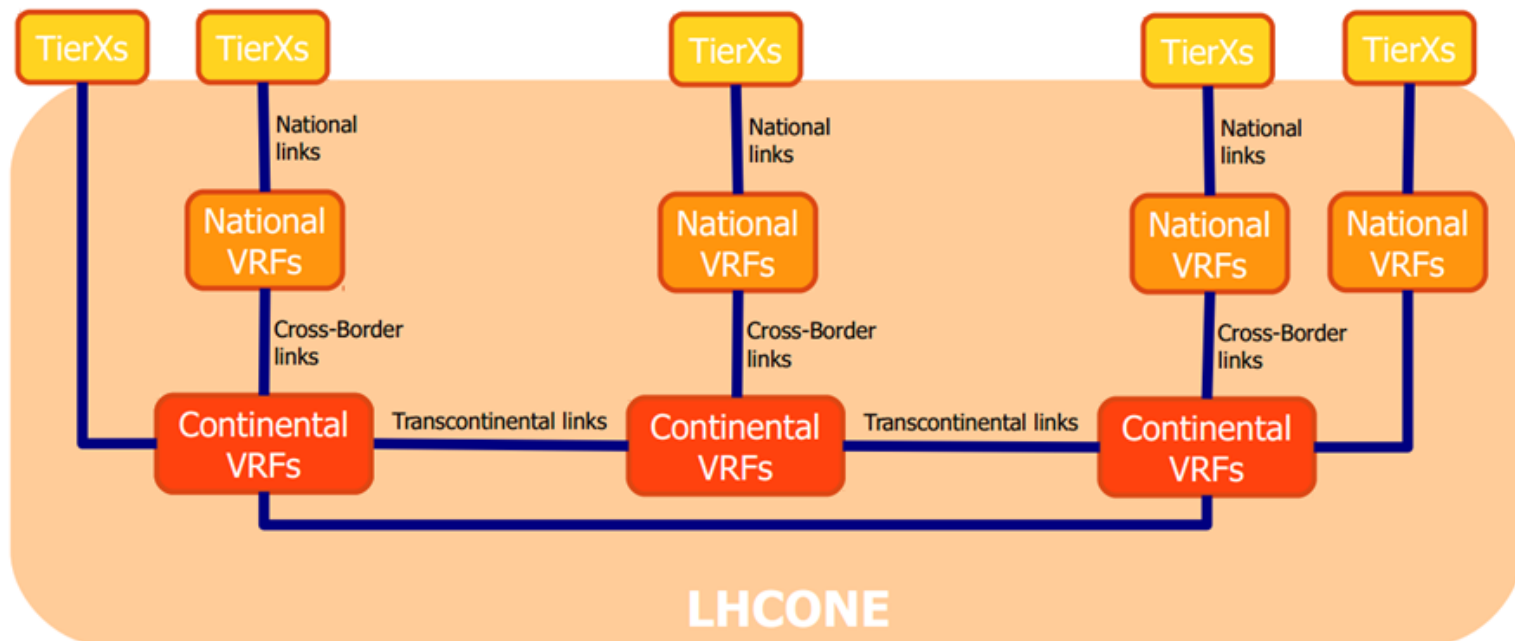
- **L3VPN:** *Virtual Private Network* roteada (em produção)
- **P2P:** enlaces ponto-a-ponto, garantia de banda (em desenvolvimento)
- **PerfSONAR:** infraestrutura de monitoramento (em produção)



LHC OPEN NETWORK ENVIRONMENT

- **Arquitetura L3VPN**

- **TierX** se conectam às **VRFs-Nacionais** ou **VRFs-Continentais**
- **VRFs-Nacionais** são **interconectadas** através das **VRFs-Continentais**
- **VRFs-Continentais** são **interconectadas** através de **enlaces transcontinentais** e/ou **transoceânicos**



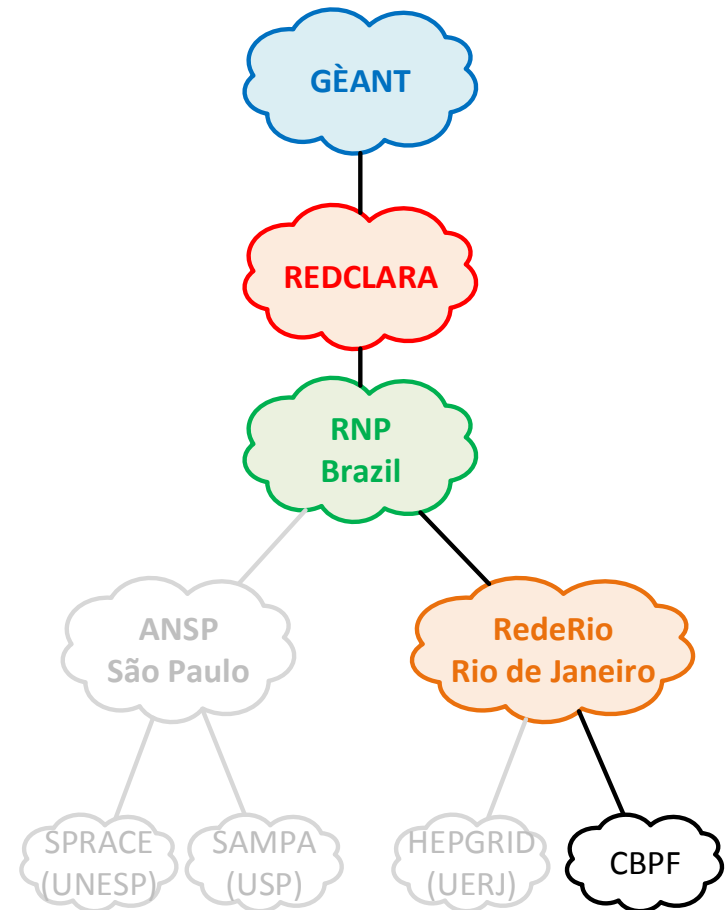


LHC OPEN NETWORK ENVIRONMENT

- **Participação Brasileira:**

- **Projeto piloto de conexão do CBPF**
- **Esforço coordenado** entre diversos grupos:

- GÈANT
- RedCLARA
- RNP
- RedeRio/FAPERJ
- CBPF



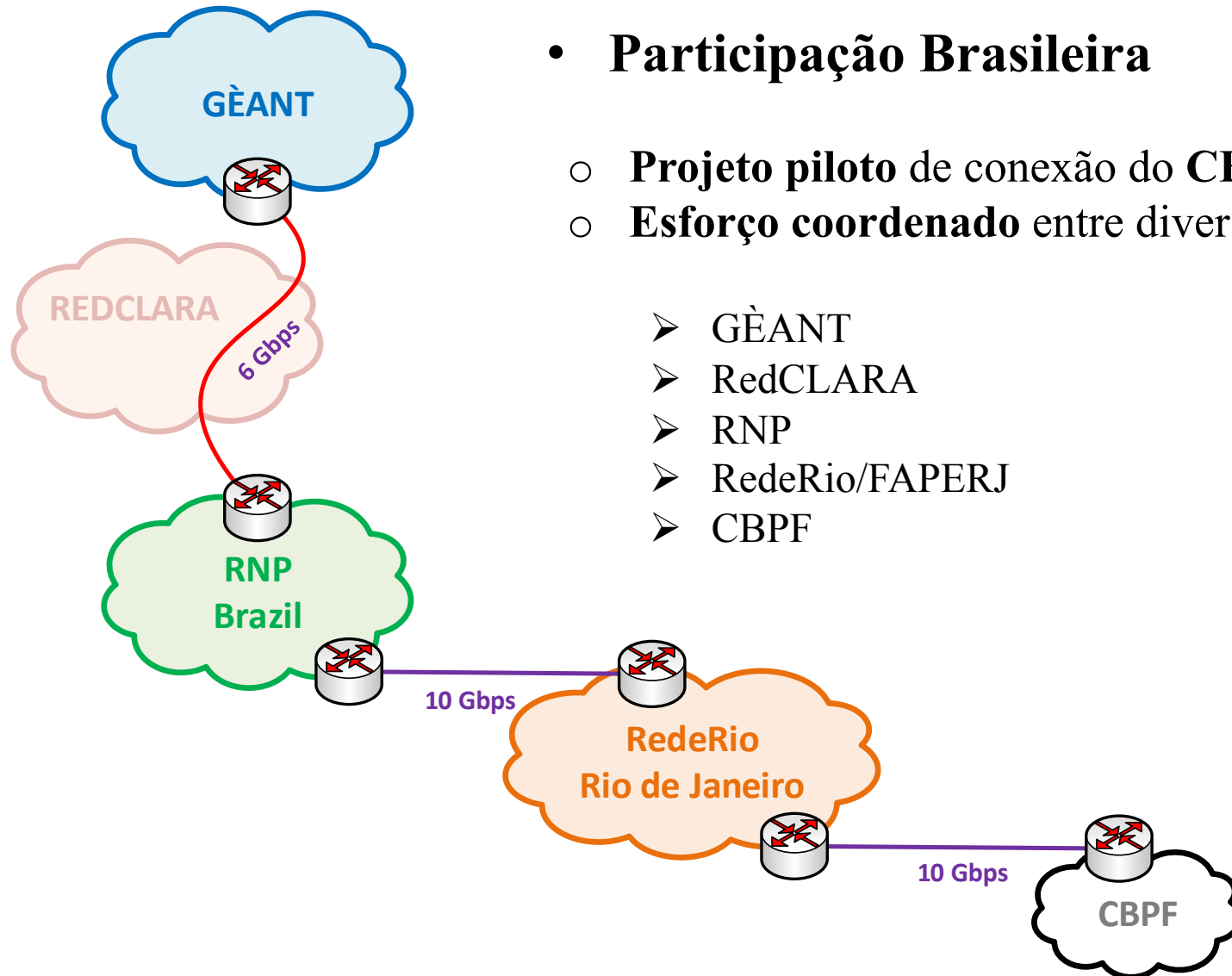


LHC OPEN NETWORK ENVIRONMENT

- **Participação Brasileira**

- **Projeto piloto de conexão do CBPF**
- **Esforço coordenado** entre diversos grupos:

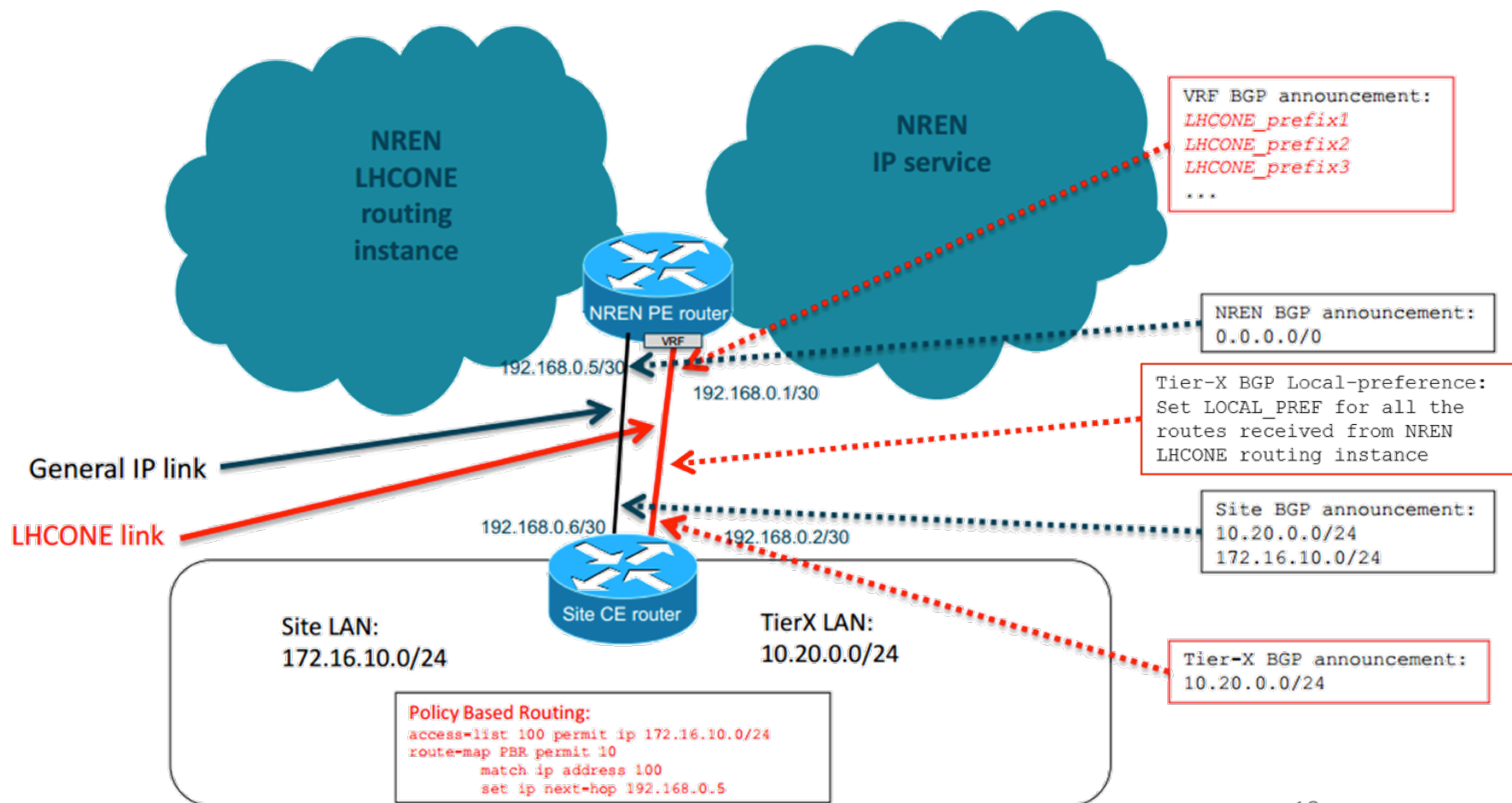
- GÈANT
- RedCLARA
- RNP
- RedeRio/FAPERJ
- CBPF





LHC OPEN NETWORK ENVIRONMENT

- Exemplo de conexão**





LHC OPEN NETWORK ENVIRONMENT

- **Por que testar os sites pertencentes ao LHCONE?**
 - **Identificar problemas** nos *paths* entre os sites
 - Encontrar a **localização** dos problemas
 - **Caracterizar** a utilização da rede
 - **Definir expectativas** sobre o que é possível e esperado
 - **Fornecer métricas** de utilização da rede para **serviços existentes e futuros**



LHC OPEN NETWORK ENVIRONMENT

- **Monitoramento de Desempenho**

- *WLCG Network and Transfer Metrics Working Group*

- **Objetivos**

- ✓ **Encontrar e isolar problemas**
 - ✓ **Caracterizar utilização da rede**
 - ✓ **Prover fontes de métricas de rede para utilização de serviços de mais alto nível**
 - ✓ **Coordenar o comissionamento e manutenção do monitoramento de rede do WLCG**
 - ❖ **Finalizar a implementação do perfSONAR**
 - ❖ **Garantir que todos os enlaces sejam monitorados e corretamente configurados**
 - ❖ **Otimizar os parâmetros de rede**

- **Criação: Setembro de 2014**



LHC OPEN NETWORK ENVIRONMENT

- **Infraestrutura de Monitoramento**

- Por que **perfSONAR** ?

- Amplamente **suportado** em **redes de educação e pesquisa**
- Provê **medições programadas** de **métricas de rede padrão**
- **Simple implantação** ('netinstall', CD-ROM ou pendrive)
- **Suporta testes sob demanda**
- Ferramentas importantes para verificação das métricas
 - ✓ **Banda:** `bwctl -s <source> -c <destination>`
 - ✓ **Latência:** `owping <destination>`
- **Útil para:**
 - ✓ **Debug**
 - ✓ **Verificar mudanças em configuração de firewall**
 - ✓ **Verificar mudanças em serviços**



LHC OPEN NETWORK ENVIRONMENT

- **Infraestrutura de Monitoramento**

- **Histórico**

- **LHCOPN** adotou o **perfSONAR-PS toolkit** em **Junho de 2011**; implementado em setembro de 2011
- **WLCG iniciou** a implementação do **PS toolkit** em todos os *sites* em **Abril de 2014**
- **LHCONE** necessitava de monitoramento específico:
 - ✓ Utilizando o mesmo Sistema do WLCG para o caso específico
 - ✓ Utilizando os sites existentes e instrumentando os PoPs do LHCONE

**Não diretamente
vinculados ao WLCG**



LHC OPEN NETWORK ENVIRONMENT

- Infraestrutura de Monitoramento

http://grid-monitoring.cern.ch/perfsonar_report.txt

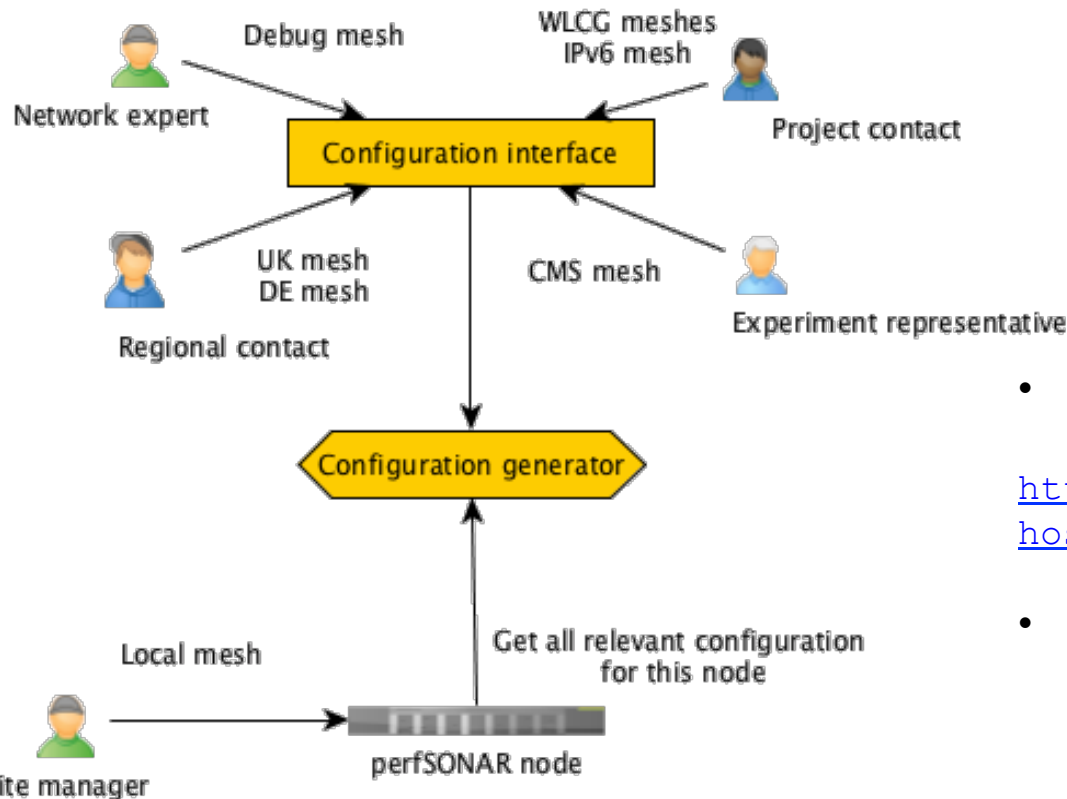




LHC OPEN NETWORK ENVIRONMENT

• Infraestrutura de Monitoramento

○ Interface de Configuração



- perfSONAR basta ser configurado uma única vez, com uma auto-URL
http://myosg.grid.iu.edu/pfmesh/hostname/<node_id>
- Interface de configuração e a auto-URL permite a reconfiguração dinâmica de todas as instâncias perfSONAR



LHC OPEN NETWORK ENVIRONMENT

- **Infraestrutura de Monitoramento**

- Interface de Configuração

 MyOSG ▾ Home Resource Group ▾ Virtual Organization ▾ Status Map Misc ▾

Perfsonar Mesh Configurations

This page is still under development. Please contact GOC for more information.

Auto Mesh Config

Please enter the hostname of your perfsonar toolkit instance to generate a mesh config containing all tests that uses your host as endpoints.

USATLAS Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/us-atlas
USCMS Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/us-cms
CA Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/ca-all
DE Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/de-all
ES Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/es-all
FR Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/fr-all
IT ATLAS Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/it-atlas
IT CMS Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/it-cms
LHCONE Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/lhcone-all
ND Mesh Config	https://myosg.grid.iu.edu/psmesh/json/name/nd-all

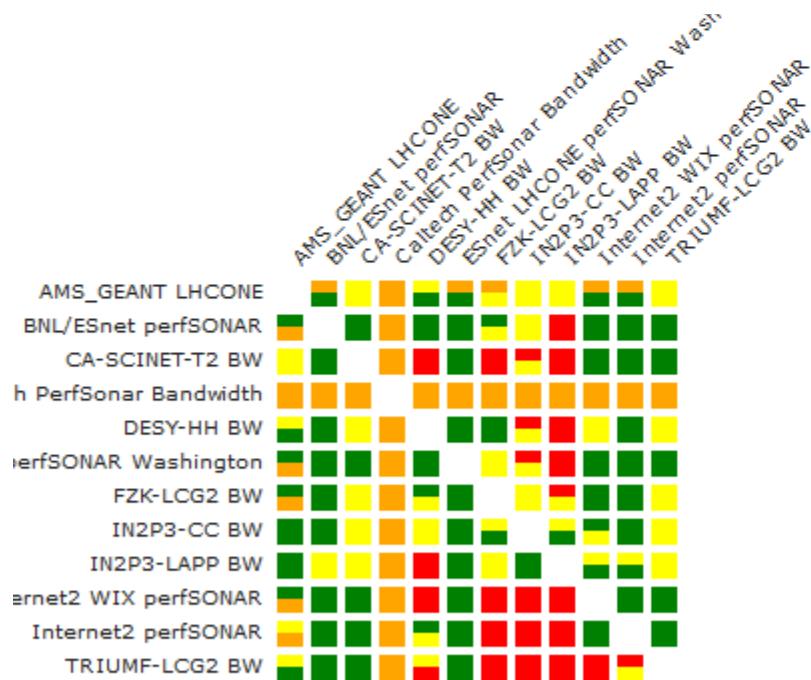


LHC OPEN NETWORK ENVIRONMENT

- **Infraestrutura de Monitoramento**

- Atualmente existem **12 sites** realizando **testes de banda e latência** (<http://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=LHCONE%20Mesh%20Config>)

Banda (BWCTL)



Latência (OWAMP)





LHC OPEN NETWORK ENVIRONMENT

- **Infraestrutura de Monitoramento**

- ❖ **Configuração**

- *Nodes* localizados **mais próximo possível** do *storage* dos sites (garante a medição do mesmo caminho de/para o *storage* destino)
- Duas instâncias de medição e uma verificação são realizadas
 - **Latência**
 - ✓ Mede “*one way delay*” enviando **10 pacotes/s** para todos os alvos
 - **Banda**
 - ✓ Mede a banda alcançável em testes de **25 segundos** a cada **6 horas**
 - **Traceroute**
 - ✓ Testes realizados a todos os parceiros a cada **20 minutos**

Verifica qual caminho está ativo durante os testes

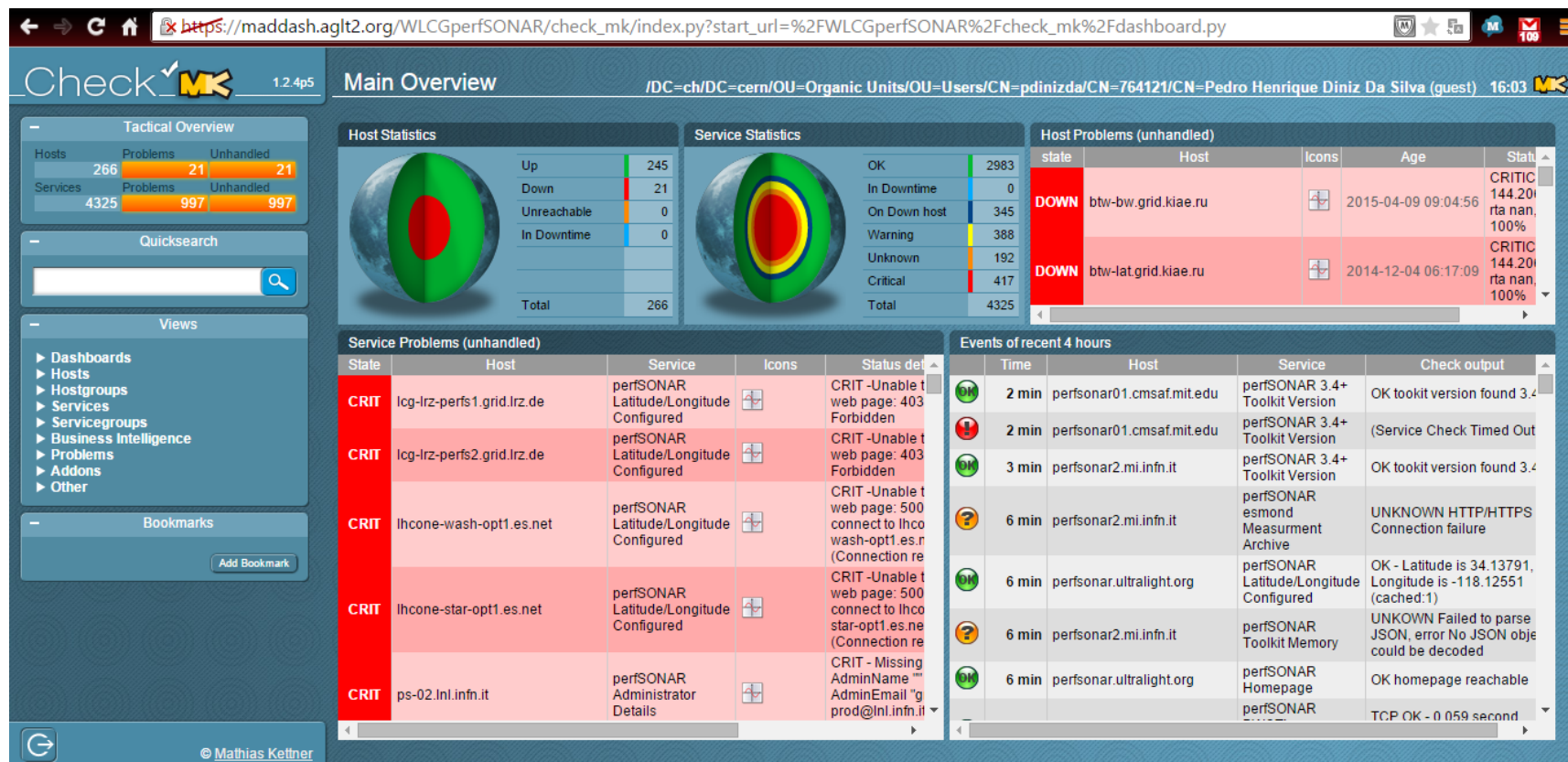


LHC OPEN NETWORK ENVIRONMENT

• Infraestrutura de Monitoramento

❖ Plataforma de Monitoramento

- **Objetivo:** garantir a confiabilidade de todas as métricas





LHC OPEN NETWORK ENVIRONMENT

- Infraestrutura de Monitoramento

- ❖ Plataforma de Monitoramento

- Sumários por *mesh*

The screenshot displays the Check_MK web interface. The top navigation bar shows the URL https://maddash.aglt2.org/WLCGperfSONAR/check_mk/index.py?start_url=%2FWLCGperfSONAR%2Fcheck_mk%2Fview.py%3Fview_name%3Dhostgr. The main header indicates the hostgroup 'LHCONE' with 19 rows of data. The left sidebar contains a 'Tactical Overview' section with a table showing 266 hosts and 4325 services, along with a 'Quicksearch' bar and a 'Views' menu. The main content area displays a table of host status and performance metrics.

state	Host	Icons	Alias	OK	Wa	Un	Cr	Pd
UP	ccperfsnar1.in2p3.fr		ccperfsnar1.in2p3.fr	14	2	0	0	0
UP	ccperfsnar2.in2p3.fr		ccperfsnar2.in2p3.fr	14	2	0	0	0
UP	lapp-ps01.in2p3.fr		lapp-ps01.in2p3.fr	14	2	0	0	0
UP	lapp-ps02.in2p3.fr		lapp-ps02.in2p3.fr	14	2	0	0	0
UP	lhcone-wash-opt1.es.net		lhcone-wash-opt1.es.net	5	2	5	7	0
UP	lhcone.test.manlan.internet2.edu		lhcone.test.manlan.internet2.edu	11	4	1	3	0
UP	lhcone.test.wix.internet2.edu		lhcone.test.wix.internet2.edu	10	4	1	4	0
UP	perfsonar-de-kit.gridka.de		perfsonar-de-kit.gridka.de	13	2	0	1	0
UP	perfsonar-ps-01.desy.de		perfsonar-ps-01.desy.de	14	2	0	0	0
UP	perfsonar-ps-02.desy.de		perfsonar-ps-02.desy.de	12	2	0	2	0
UP	perfsonar.ultralight.org		perfsonar.ultralight.org	14	0	0	2	0
UP	perfsonar2-de-kit.gridka.de		perfsonar2-de-kit.gridka.de	14	2	0	0	0
UP	perfsonar2.ultralight.org		perfsonar2.ultralight.org	14	2	0	0	0
UP	ps-bandwidth.lhcom.triumf.ca		ps-bandwidth.lhcom.triumf.ca	11	2	0	3	0
UP	ps-bandwidth.scinet.utoronto.ca		ps-bandwidth.scinet.utoronto.ca	11	2	0	3	0
UP	ps-latency.lhcom.triumf.ca		ps-latency.lhcom.triumf.ca	14	2	0	0	0
UP	ps-latency.scinet.utoronto.ca		ps-latency.scinet.utoronto.ca	13	1	0	2	0
UP	ps01-nl.geant.net		ps01-nl.geant.net	13	4	1	1	0
UP	quark.es.net		quark.es.net	6	2	5	6	0



LHC OPEN NETWORK ENVIRONMENT

• Infraestrutura de Monitoramento

❖ Plataforma de Monitoramento

- Sumário de serviços

CheckMK 1.2.4p5

Services of Host p... 16 rows /DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=pdinizda/CN=764121/CN=Pedro Henrique Diniz Da Silva (guest) 16:11

Tactical Overview

Hosts	Problems	Unhandled
266	21	21
Services	Problems	Unhandled
4325	998	998

Quicksearch

ps02

Views

- ▶ Dashboards
- ▶ Hosts
- ▶ Hostgroups
- ▶ Services
- ▶ All services
- ▶ Favorite services
- ▶ Recently changed services
- ▶ Serv. by host groups
- ▶ Service search
- ▶ Servicegroups
- ▶ Business Intelligence
- ▶ Problems
- ▶ Addons
- ▶ Other

Bookmarks

Add Bookmark

ps02.cat.cbpf.br

State	Service	Icons	Status detail	Age	Checked	Perf-O-Meter
OK	perfSONAR 3.4+ Toolkit Version		OK toolkit version found 3.4.2	26 hrs	62 sec	
OK	perfSONAR Administrator Details		OK - Administrator is GridLAFEX Team, email gridlafex@cbpf.br (cached:0)	2015-05-18 20:53:11	17 min	
OK	perfSONAR BWCTL Bandwidth Test Controller		TCP OK - 0.178 second response time on 152.84.101.142 port 4823	2015-05-18 15:53:33	17 min	178 277 ms
WARN	perfSONAR esmond Freshness Bandwidth Direct		WARNING Found 30 stale hosts for certain events	2015-05-18 20:56:24	94 min	
WARN	perfSONAR esmond Freshness Bandwidth Reverse		WARNING Found 72 stale hosts for certain events	2015-05-18 20:56:46	93 min	
OK	perfSONAR esmond Measurement Archive		OK esmond reachable	2015-05-18 20:52:07	18 min	
OK	perfSONAR Homepage		OK homepage reachable	2015-05-18 20:53:54	17 min	
OK	perfSONAR Latitude/Longitude Configured		OK - Latitude is -22.954, Longitude is -43.174 (cached:1)	2015-05-18 20:54:16	16 min	
OK	perfSONAR Mesh Configuration		OK auto-URL configured	2015-05-18 20:54:37	16 min	
OK	perfSONAR NDT HTTP Network Diagnostic Tester		Connection refused	9 hrs	35 min	
OK	perfSONAR NDT P Network Diagnostic Tester		Connection refused	9 hrs	34 min	
OK	perfSONAR NPAD Network Path and Application Diagnosis		Connection refused	9 hrs	34 min	
OK	perfSONAR NTP Service		OK NTP synchronized	2015-05-18 15:54:58	16 min	
OK	perfSONAR Regular Testing Service		OK Regular Testing enabled and running	2015-05-18 15:55:20	15 min	
OK	perfSONAR Toolkit Memory		OK memory installed 31GB	2015-05-18 20:55:41	15 min	
OK	perfSONAR Toolkit Version		OK - Version 3.4.2 OK (cached:1)	2015-05-18 20:56:03	14 min	

Mathias Kettner

refresh: 30 secs



LHC OPEN NETWORK ENVIRONMENT

- **Infraestrutura de Monitoramento**

- ❖ **Planos (*WG*)**

- **Integrar** as informações de **latência e *traceroutes*** com as informações de ***throughput*** dos **sistemas de transferência** (FTS, PhEDEx, XRootD...)
- Prover um **mapeamento entre os sites (*storages*)** e os sonares
- Garantir um acesso uniforme ao monitoramento de rede
- Prover uma **plataforma que integre métricas de rede e de transferências**



OBRIGADO!



CBPF

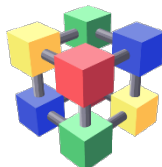


Coordenação de Engenharia
Operacional da Rede Rio



FAPERJ

Fundação Carlos Chagas Filho de Amparo
à Pesquisa do Estado do Rio de Janeiro



WLCG

Worldwide LHC Computing Grid

REDECOMEP RIO  **RNP**



23 anos

RedeRio

de Computadores - FAPERJ

Apoio:



RNP

