



**NET
MAKERS**

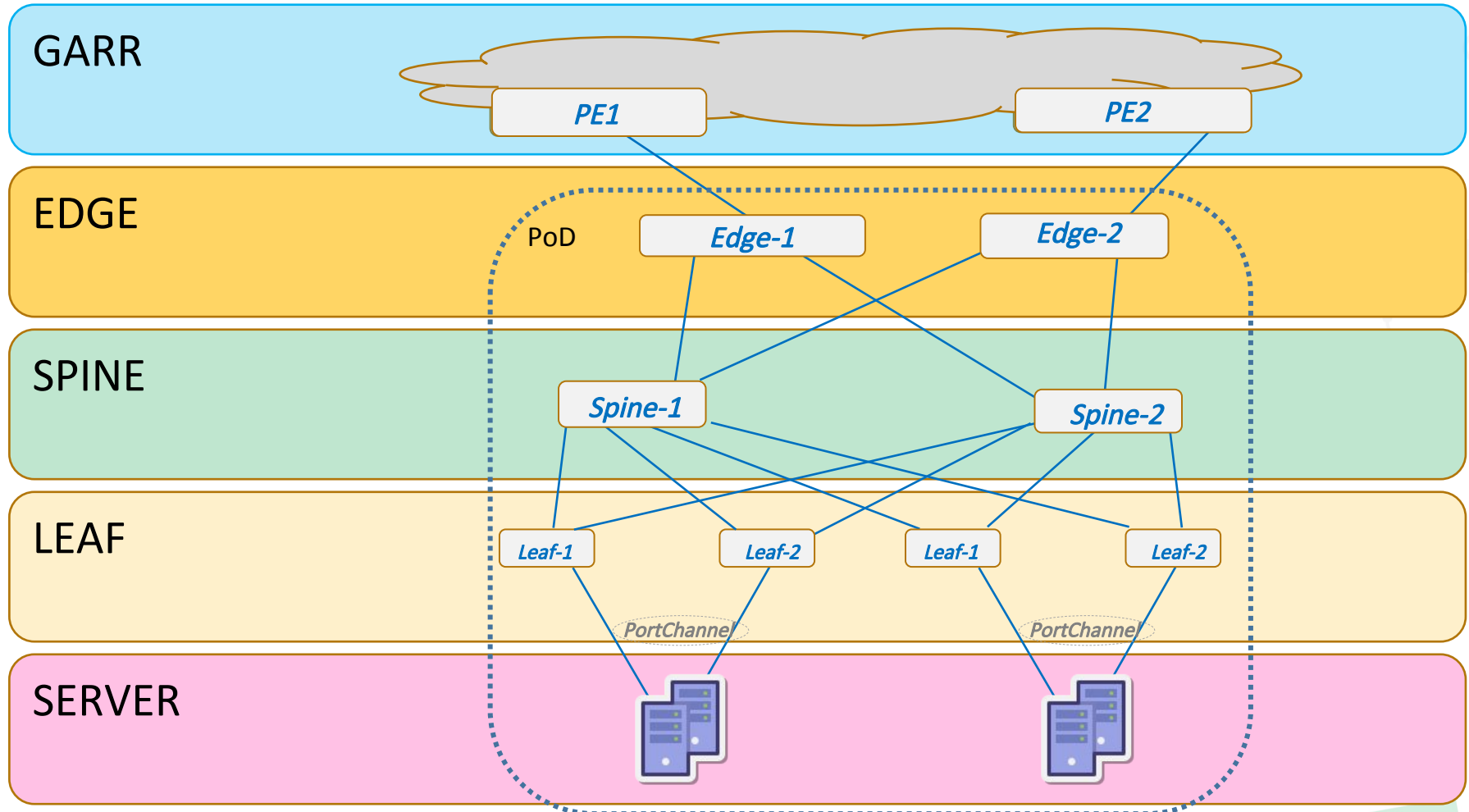
Programmabilità della rete Datacenter

*Nino Ciurleo
Consortium GARR*

Agenda

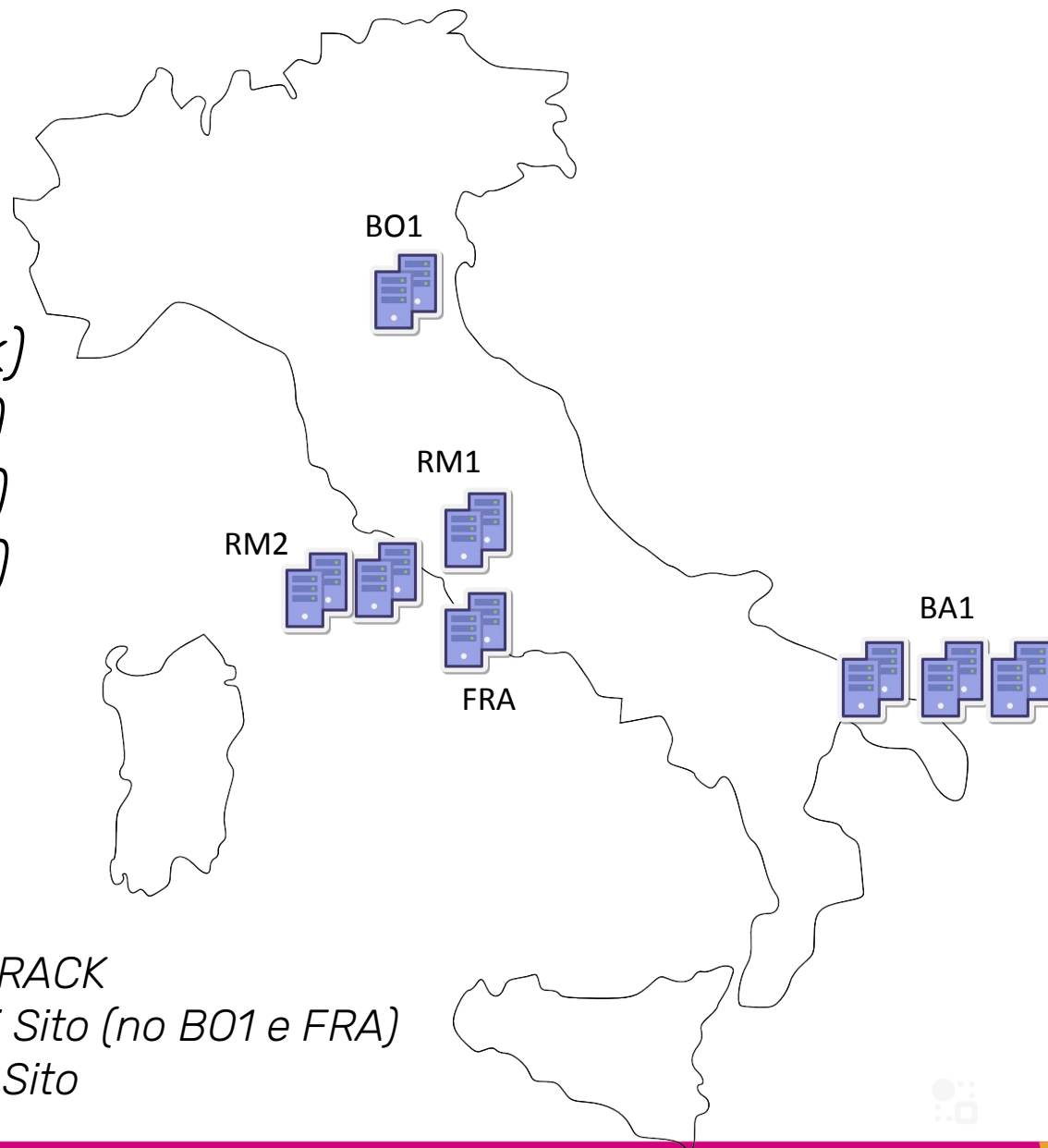
- *Come e' fatta la nuova rete DC GARR Infra*
 - *Architettura*
 - *Siti*
 - *Routing*
 - *Requisiti servizi di rete*
- *Gestire la complessita'*
 - *Automation*
 - *Cloud Vision Portal (CVP)*
 - *Ansible playbooks*
 - *Arista Validated Design (AVD)*
 - *Documentazione automatica*
 - *Statistiche*
 - *Alerting*

Architettura nuova rete DC



Siti DC

- *RM2 (2 rack)*
- *RM1 (1 rack)*
- *FRA (1 rack)*
- *BA1 (3 rack)*
- *BO1 (1 rack)*



2xLEAF per ogni RACK
2xSPINE per ogni Sito (no BO1 e FRA)
2xEDGE per ogni Sito



Componenti di un sito DC

MX-204

(4x 100G + 8x 10/1G)



BL

JUNIPER
NETWORKS

7050CX3-32S-F

(32x 100G)



SPINE

ARISTA

7050SX3-48YC8-F

(8x 100G + 48 25/10/1G)



LEAF



Network

Mini DC

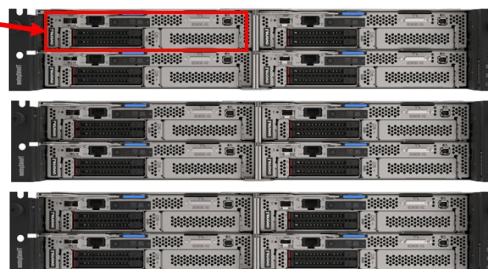
SD630 v2

64 CORE

1T RAM

2x 25GE

Lenovo



CPU

3x enclosure DA240

DE6000H

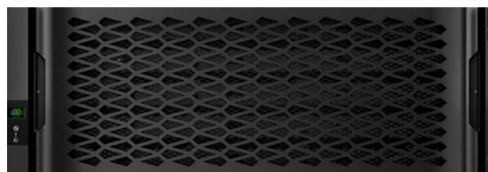
96 TByte SSD



DE600S

1 PByte meccanico

8x 25GE

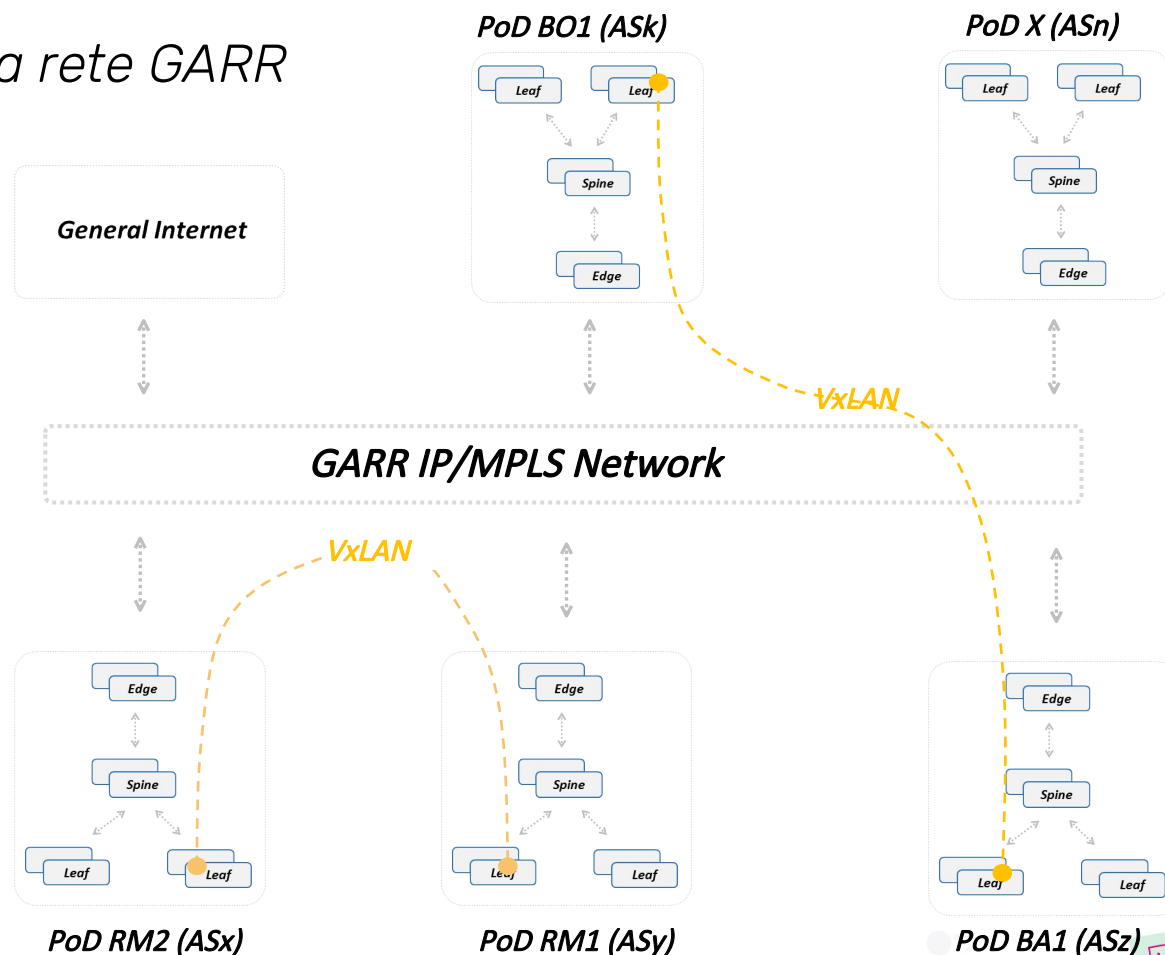


Storage

Architettura a PoD

I Datacenter fanno parte di un unico sistema:

- ogni PoP un PoD
- ogni PoD un AS
- PoD connessi alla rete GARR



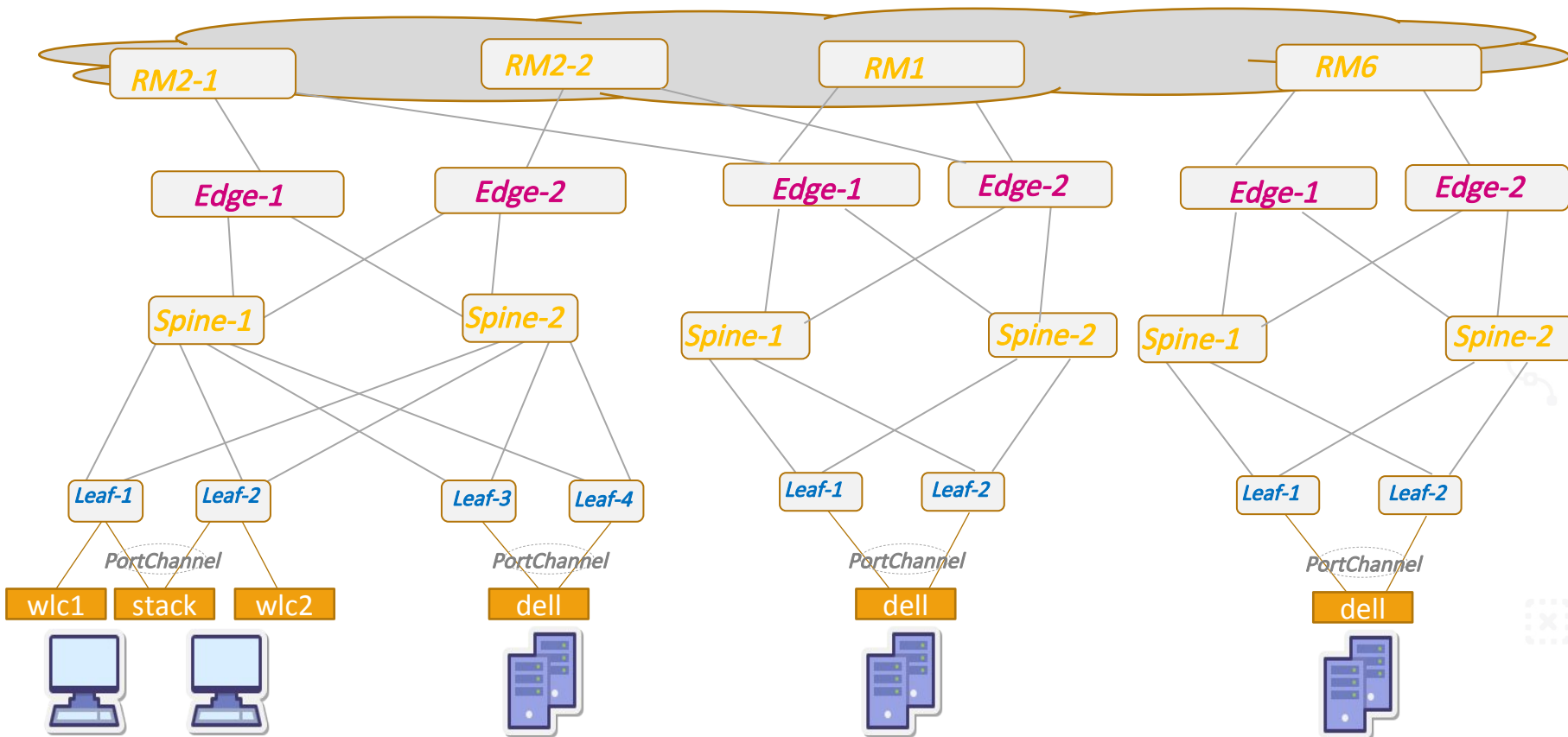
Protocolli di routing

- *ISIS all'interno dei PoD*
- *Segnalazione BGP EVPN per MAC e IP*
- *VARP (tra leaf e server): GW distribuito su due leaf*
 - *Anche in caso di punto-punto*
- *VXLAN tunnel terminati su VTEP (solo su LEAF/EDGE):*
- *Gli SPINE trasportano solo traffico IP*
- *Anycast gateway*
 - *Indirizzo IP gateway duplicato su piu' nodi contemporaneamente*
- *full-mesh BGP EVPN tra EDGE*
- *BGP EVPN sulla rete GARR non necessario*
 - *solo trasporto IP / VXLAN*

Requisiti dei servizi di rete

- *Segregazione delle LAN*
 - *Le LAN devono essere piu' isolate possibile*
- *Aggregazione LAN in VRF per categorie di traffico*
 - *Alcune LAN devono scambiare traffico tra loro, indipendentemente dalla loro locazione geografica.*
- *Anycast Gateway (possibilita' di spostare le VM tra PoD)*
- *Sicurezza*
 - *Solo i servizi esplicitamente esposti devono essere raggiungibili*
- *Possibilita' di ospitare LAN estranee nei Data Center GARR*
 - *Garantendone la separazione con le altre LAN*

Molteplicita'



pc

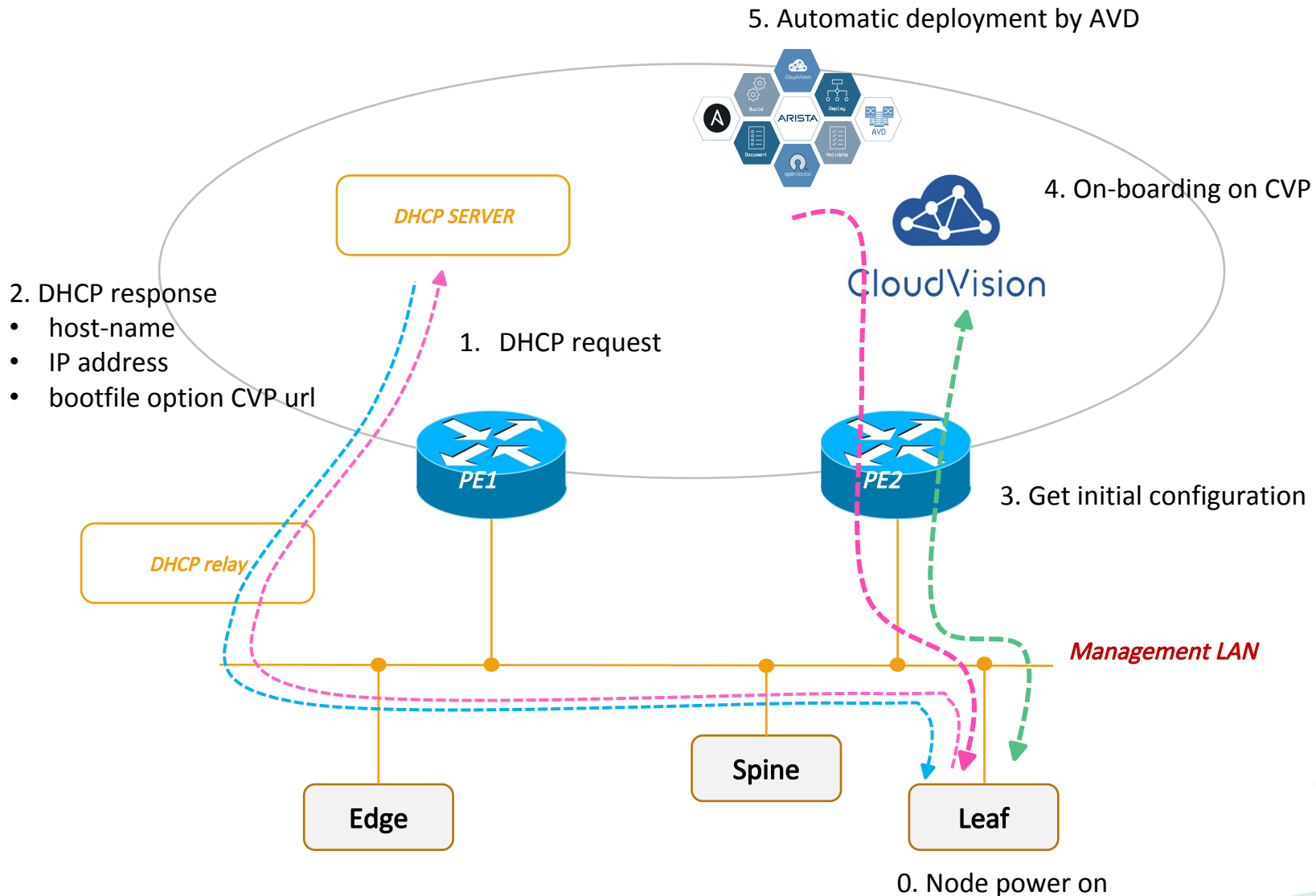


Server

Gestire la complessita'

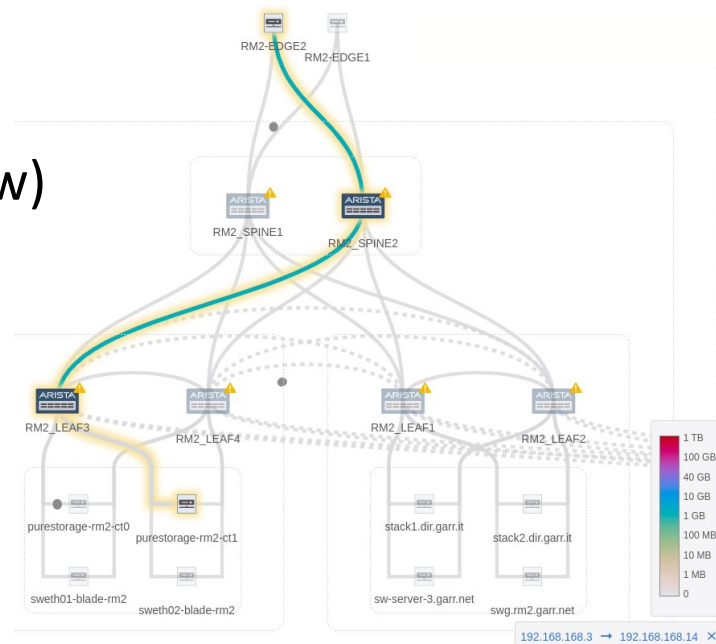
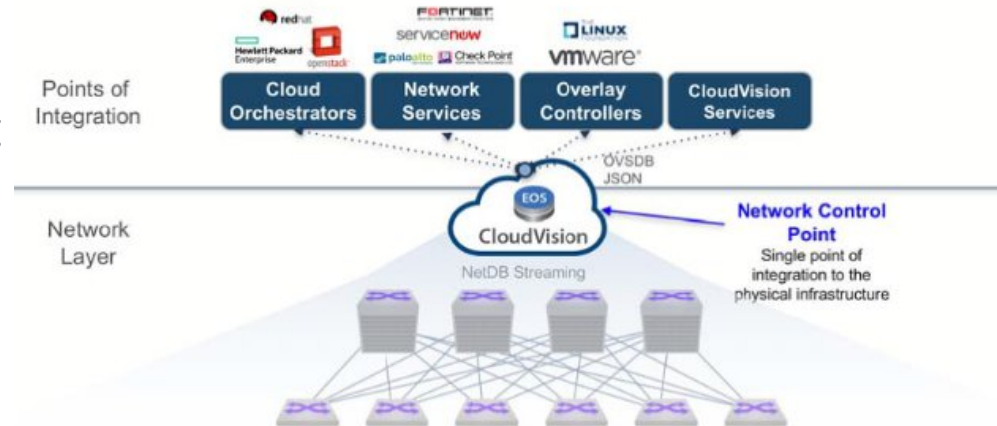
- *Zero touch provisioning*
- *Cloud Vision Portal*
 - *System upgrade*
 - *Configlets*
- *Arista Validated Design*
- *Ansible playbook*
 - *Juniper Firewall*
- *Bash script per la migrazione*
- *Documentazione automatica*
 - *Via CI/CD*
 - *Integrazione con RtD*
- *Telemetria*

Procedura zerotouch provisioning



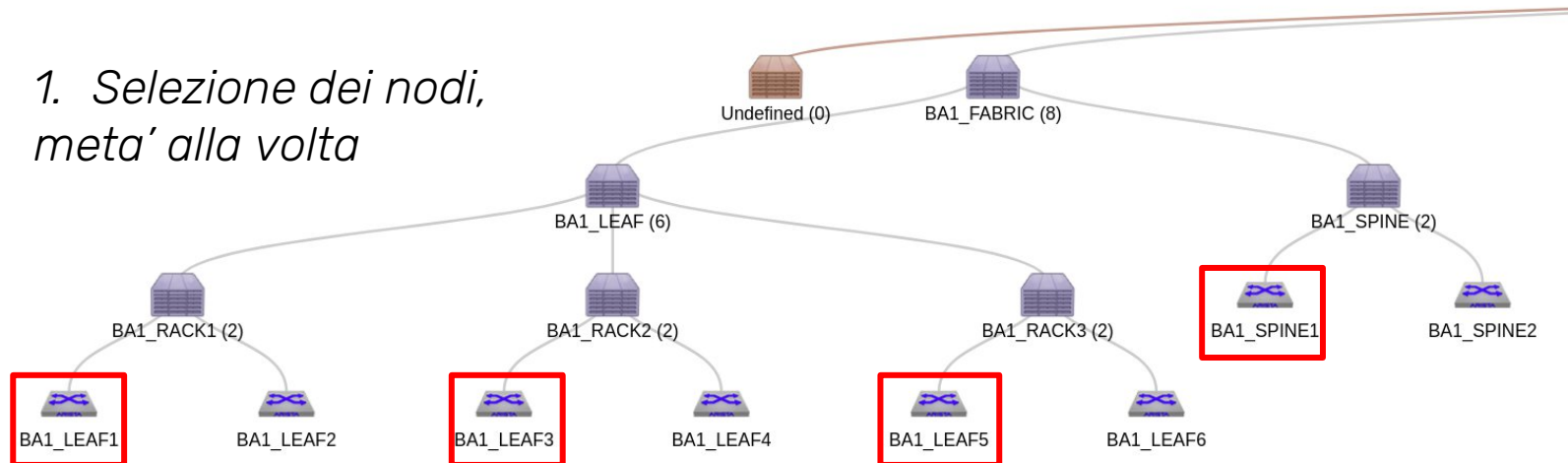
Management System (Cloud Vision Portal)

- Provisioning
 - zerotouch
 - configurazione via configlet
 - system upgrade massivo
- Inventory hardware
- Monitoring integrato
 - Dashboards
 - Telemetry
 - Logs
 - Flows (sflow)



Procedura system upgrade

1. Selezione dei nodi, meta' alla volta



2. Selezione della release

3. Lancio del'upgrade

Name	
<input type="checkbox"/>	EOS-4.27.2F
<input type="checkbox"/>	4.28.0F
<input type="checkbox"/>	4.27.3F
<input type="checkbox"/>	bootstrap
<input type="checkbox"/>	4.27.0F
<input type="checkbox"/>	EOS-4.25.4M
<input type="checkbox"/>	4.28.1F
<input checked="" type="checkbox"/>	4.28.2F

4. rete aggiornata

Device ↑	Issues	Model	Software
Filter	Filter	Filter	Filter
BA1_LEAF1	✓ ⚠	7050SX3-48YC8	4.28.2F
BA1_LEAF2	✓ ⚠	7050SX3-48YC8	4.28.2F
BA1_LEAF3	✓ ⚠	7050SX3-48YC8	4.28.2F
BA1_LEAF4	✓ ⚠	7050SX3-48YC8	4.28.2F
BA1_LEAF5	✓ ⚠	7050SX3-48YC8	4.28.2F
BA1_LEAF6	✓ ⚠	7050SX3-48YC8	4.28.2F

Senza disservizio!

Configlet

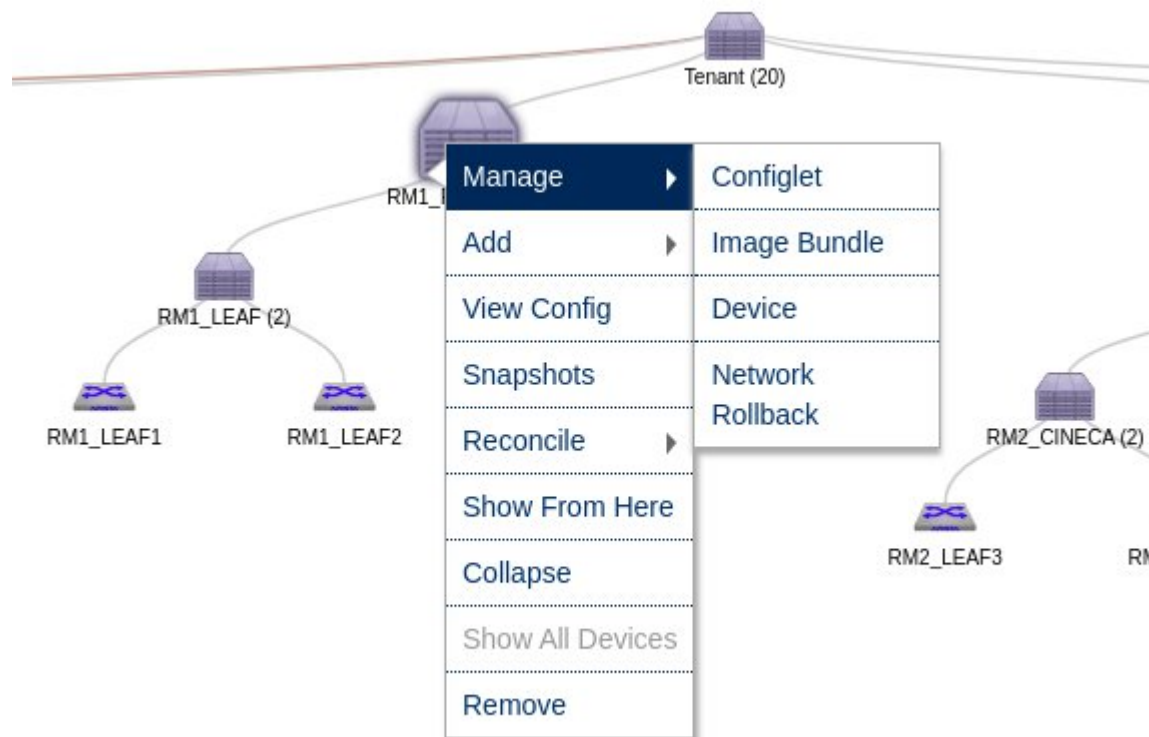
Configlets > GNMI

Summary Logs Change History

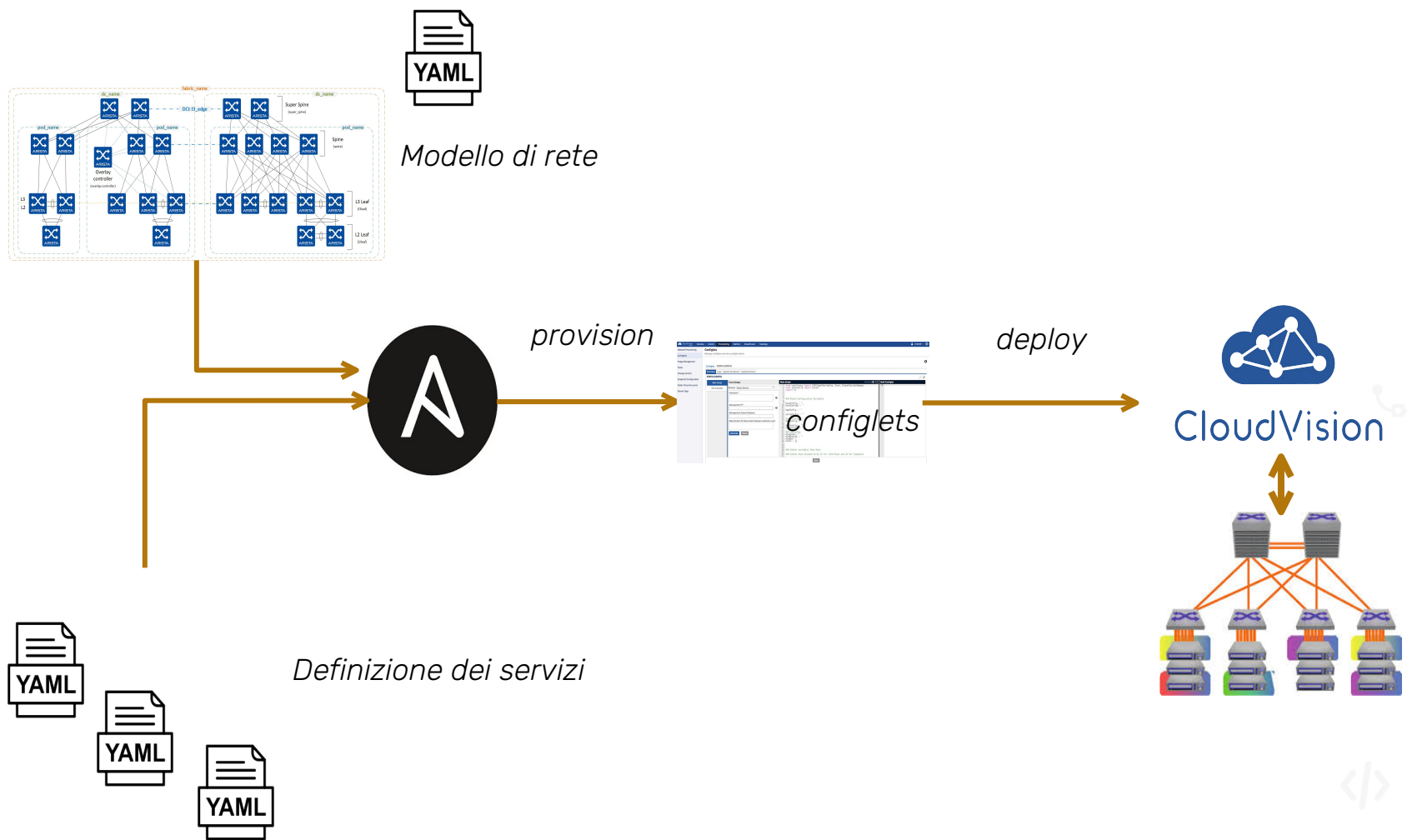
GNMI

Configuration

```
1 management api gnmi
2   transport grpc def
3     vrf MGMT
4   provider eos-native
5
```

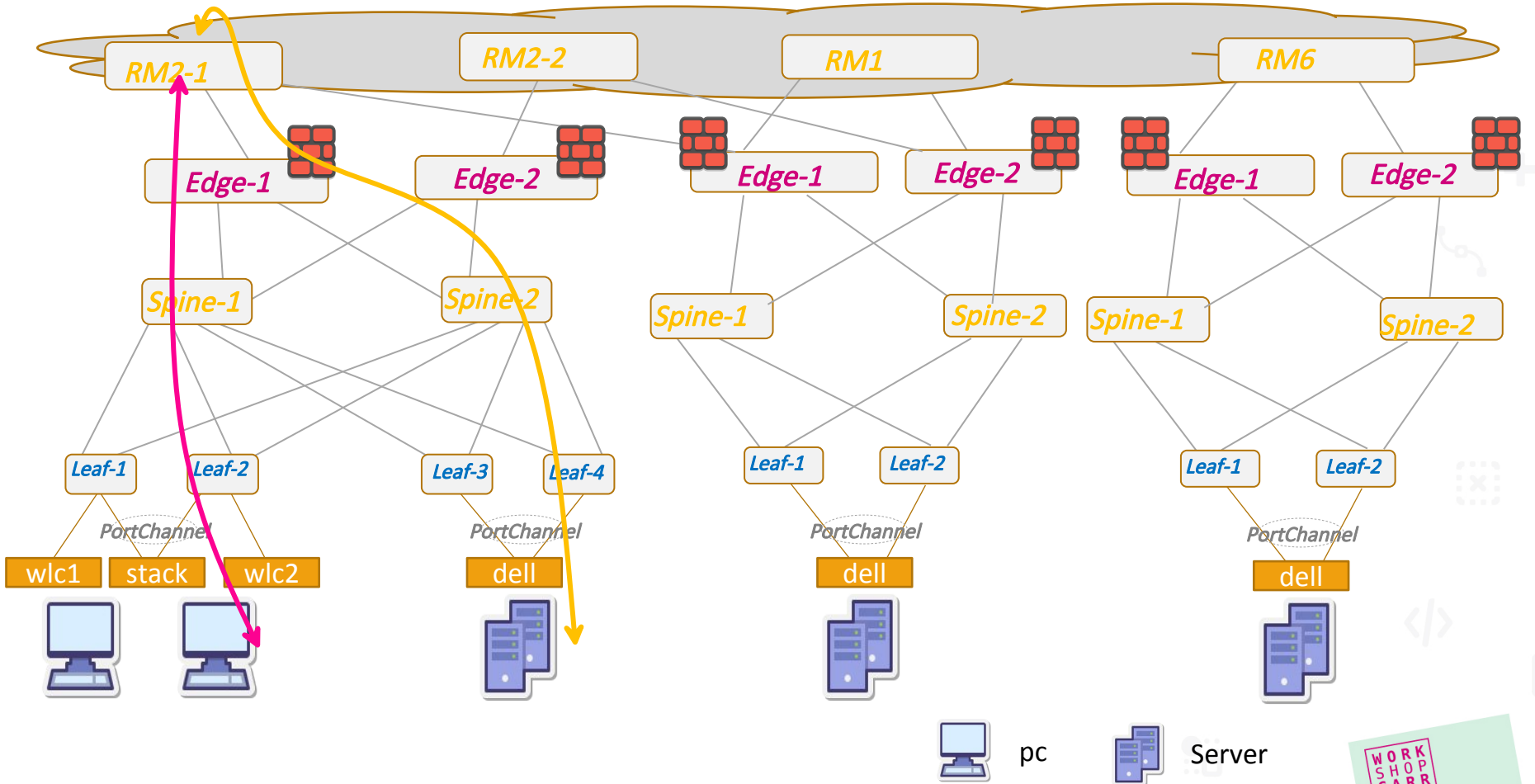


Arista Validatad Design



Gestione policy di sicurezza

- Firewall stateless su tutti gli EDGE per il traffico esterno
- E' permesso solo il traffico dei servizi da esporre
- Anche il traffico tra VRF diverse dello stesso DC passa attraverso il Firewall



Playbook per le policy

INVENTORY

```
EDGES:
hosts:
  RM2_EDGE1:
    ansible_host: 10.1.100.26
  RM2_EDGE2:
    ansible_host: 10.1.100.27
vars:
  site: RM
```

```
EDGES:
hosts:
  RM1_EDGE1:
    ansible_host: 10.2.100.26
  RM1_EDGE2:
    ansible_host: 10.2.100.27
vars:
  site: RM
```

FIREWALL/TERMS

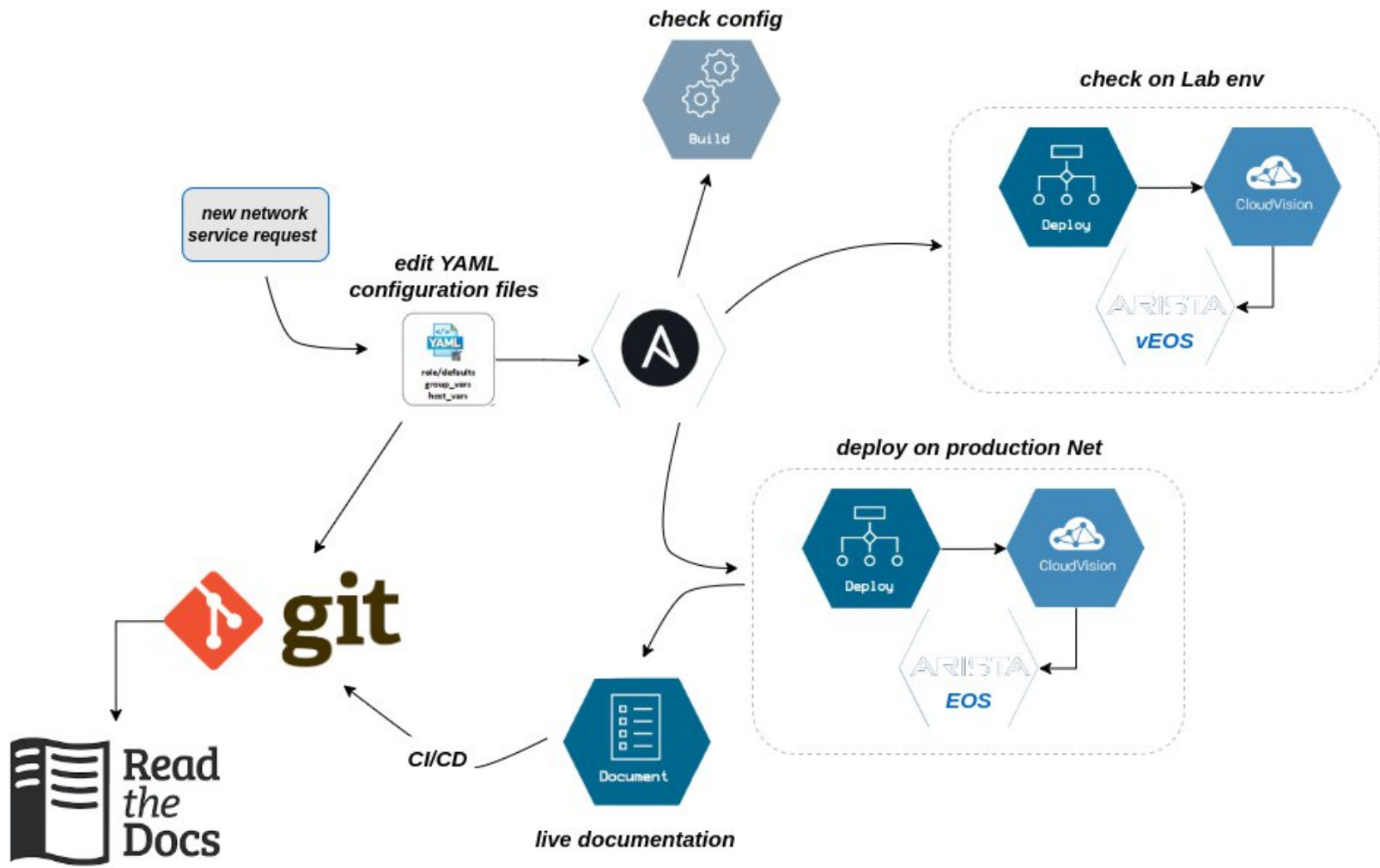
```
firewall_filter:
  RM:
    - WEB
    - SMTP
    - ...
terms:
  WEB:
    comment: Siti WEB
    match-conditions:
      destination-prefix-list:
        web-server
      destination-port:
        - 80
        - 443
    protocol:
      - tcp
    actions:
      count: web
      traffic-control: accept
```

PREFIX LIST

```
prefix_lists:
  web-server:
    RM:
      - xxx.xxx.xxx.xxx/32
      - xxx.xxx.xxx.xxx/32
      - ...
      - xxxx:xxx:x:xxx::x/128
      - xxxx:xxx:x:xxx::xx/128

  smtp-server:
    RM:
      - xxx.xxx.xxx.xxx/32
      - xxx.xxx.xxx.xxx/32
      - xxx.xxx.xxx.xxx/32
      - xxxx:xxx:x:xxx::x/128
```

Procedura di deploy

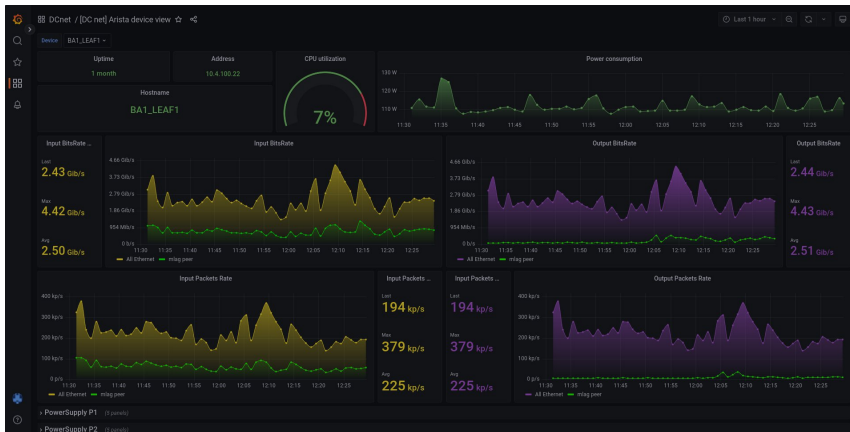


Flusso di lavoro

- 1) *Chi sviluppa l'applicazione decide:*
 - 1) *In quale rete deve stare*
 - 2) *Con quali eventuali altre reti deve parlare*
 - 3) *Che servizi devono essere esposti*
- 2) *Clone del repo GIT e modifica dei file yaml:*
 - 1) *File del servizio*
 - 2) *File delle firewall policy*
- 3) *Merge request*
- 4) *Validazione del codice da parte degli amministratori di rete:*
 - 1) *Eventuale test su laboratorio*
- 5) *Merge*
- 6) *Lancio dei playbook sulla rete di produzione*

Statistiche del DC: Telegraf InfluxDB Grafana

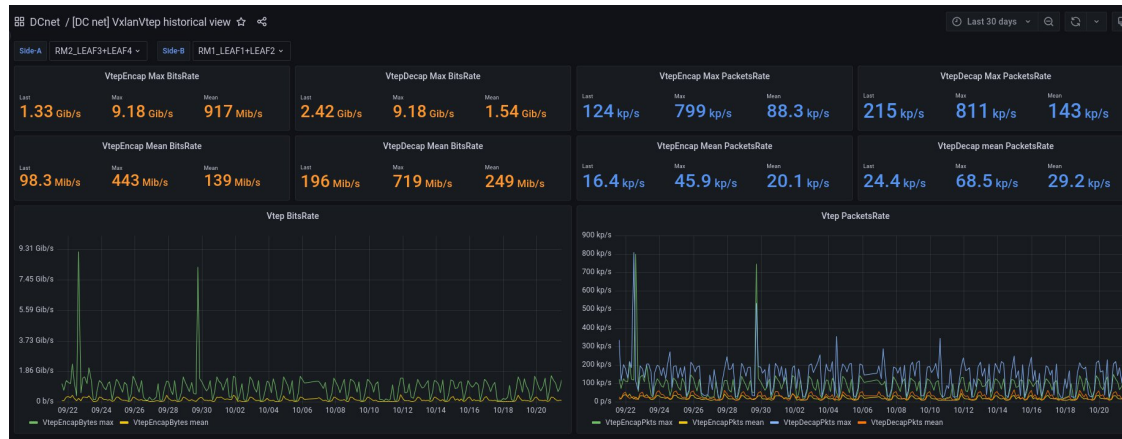
Device view



PoD view



VxLAN counters



Sflow

Dettagli

CloudVision
ARISTA

Devices Events Provisioning Dashboards Topology

Topology Flows

[Back to Flows List](#)

The following traffic flow information is sourced from sFlow and IPFIX.

Source Host:

Destination Host:

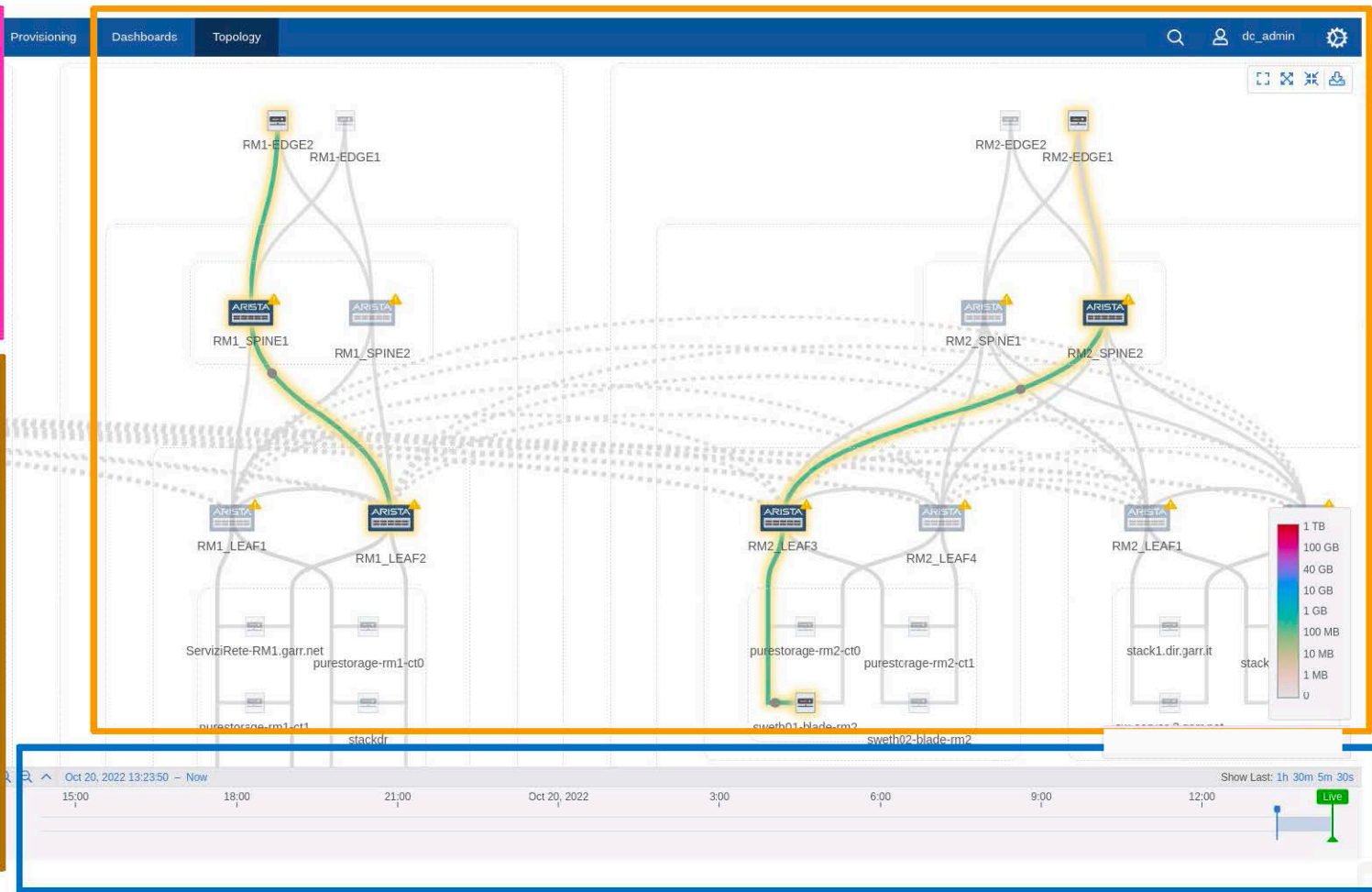
Source Port: 2049 **Destination Port:** 805 **Protocol:** TCP

Color links by: Bytes

Devices Reporting Matching Flows:

- RM1_SPINE1** 14:23:00.000
 - Ingress Interface: Ethernet2/1
 - Egress Interface: Ethernet31/1
 - Packets: 1M packets
 - Bytes: 364.2 MB
 - [Explore](#)
- RM2_LEAF3** 14:23:09.000
 - Ingress Interface: Ethernet50/1
 - Egress Interface: Ethernet4
 - Packets: 983k packets
 - Bytes: 340.8 MB
 - [Explore](#)
- RM2_SPINE2** 14:22:52.000
 - Ingress Interface: Ethernet33
 - Egress Interface: Ethernet3/1
 - Packets: 1M packets
 - Bytes: 336.2 MB
 - [Explore](#)

Percorso del flusso

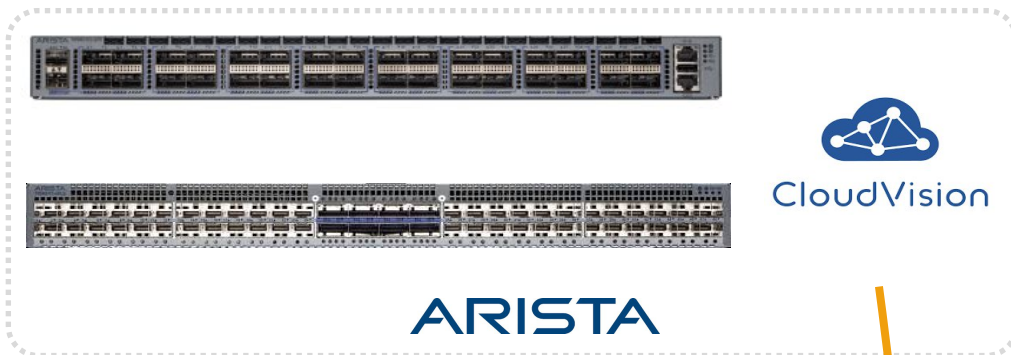


Volume di traffico per nodo/interfacce

Selezione temporale

Alerts

JUNIPER
NETWORKS



ARISTA

Poll

ZABBIX

TIG stack



Telemetry

slack

Check routing via Telemetry (in sviluppo)

ISIS adjacency state

BGP peering state

ISIS Adjacency State			
Router ↑	Interface	system_id	Status
10.1.100.22	Ethernet49/1	0001.0000.0001	UP
10.1.100.22	Ethernet50/1	0001.0000.0002	UP
10.1.100.22	Vlan4093	0001.0001.0002	UP
10.1.100.23	Ethernet49/1	0001.0000.0001	UP
10.1.100.23	Ethernet50/1	0001.0000.0002	UP
10.1.100.23	Vlan4093	0001.0001.0001	UP
10.1.100.24	Ethernet49/1	0001.0000.0001	UP
10.1.100.24	Ethernet50/1	0001.0000.0002	UP
10.1.100.24	Vlan4093	0001.0001.0004	UP
10.1.100.25	Ethernet49/1	0001.0000.0001	UP
10.1.100.25	Ethernet50/1	0001.0000.0002	UP
10.1.100.25	Vlan4093	0001.0001.0003	UP
10.2.100.22	Ethernet49/1	0001.0000.0001	UP
10.2.100.22	Ethernet50/1	0001.0000.0002	UP
10.2.100.22	Vlan4093	0001.0001.0002	UP
10.2.100.23	Ethernet49/1	0001.0000.0001	UP
10.2.100.23	Ethernet50/1	0001.0000.0002	UP
10.2.100.23	Vlan4093	0001.0001.0001	UP
10.2.100.23	Ethernet49/1	0001.0000.0001	UP

BGP Peering State		
router ↑	neighbor_address	State
0.1.100.22	10.1.20.1	ESTABLISHED
0.1.100.22	10.1.20.2	ESTABLISHED
0.1.100.23	10.1.20.1	ESTABLISHED
0.1.100.23	10.1.20.2	ESTABLISHED
0.1.100.24	10.1.20.1	ESTABLISHED
0.1.100.24	10.1.20.2	ESTABLISHED
0.1.100.25	10.1.20.1	ESTABLISHED
0.1.100.25	10.1.20.2	ESTABLISHED
0.2.100.22	10.2.20.1	ESTABLISHED
0.2.100.22	10.2.20.2	ESTABLISHED
0.2.100.23	10.2.20.1	ESTABLISHED
0.2.100.23	10.2.20.2	ESTABLISHED
0.3.100.22	10.3.20.64	ESTABLISHED
0.3.100.22	10.3.20.65	ESTABLISHED
0.3.100.23	10.3.20.64	ESTABLISHED
0.3.100.23	10.3.20.65	ESTABLISHED
0.4.100.22	10.4.20.1	ESTABLISHED
0.4.100.22	10.4.20.2	ESTABLISHED
0.4.100.22	10.4.254.254	ESTABLISHED

Segnalazione dello stato del routing in tempo reale

Rovescio della medaglia

- *Servono molte piu' competenze informatiche*
- *Non si puo' tornare indietro alle procedure manuali*
- *Troubleshooting su CLI*

FINE

- *nino.ciurleo@garr.it*