

Dalla LAN alla WAN

l'evoluzione del networking nei
DataCenter

Giancarlo Viola
INFRA - GARR

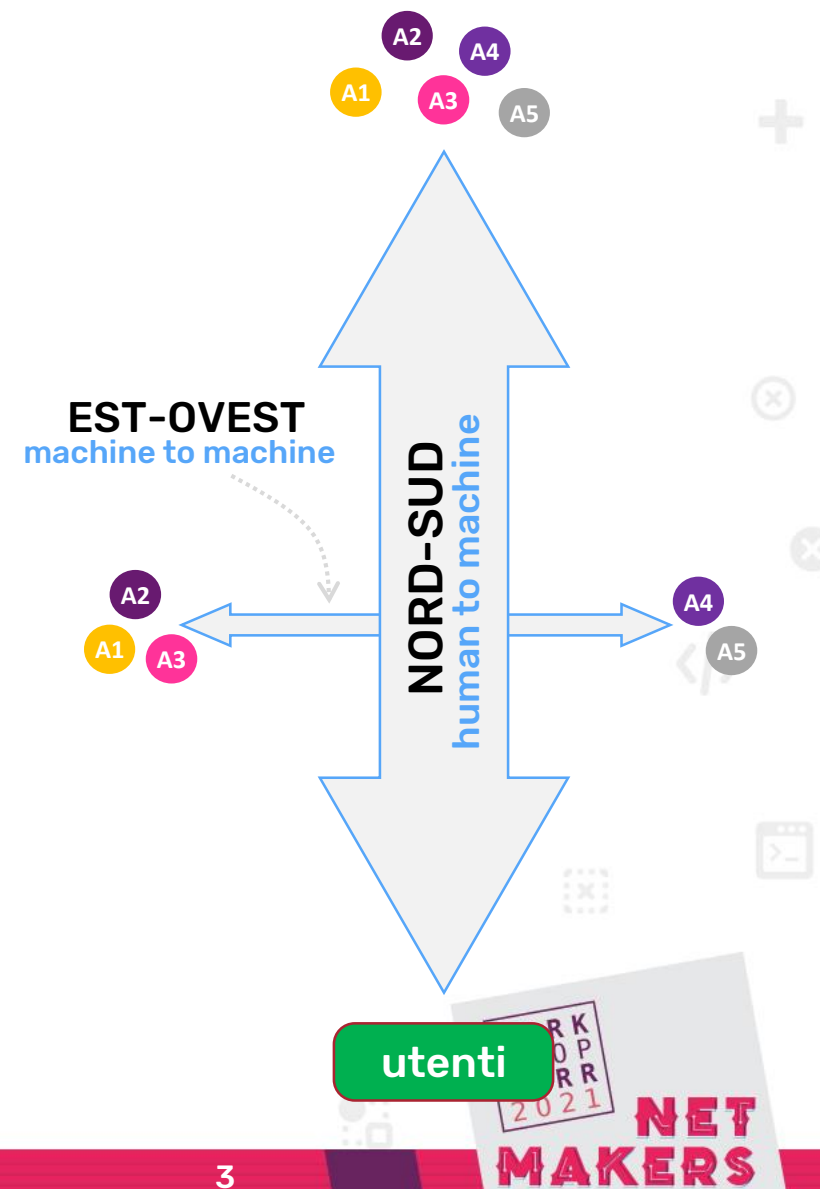
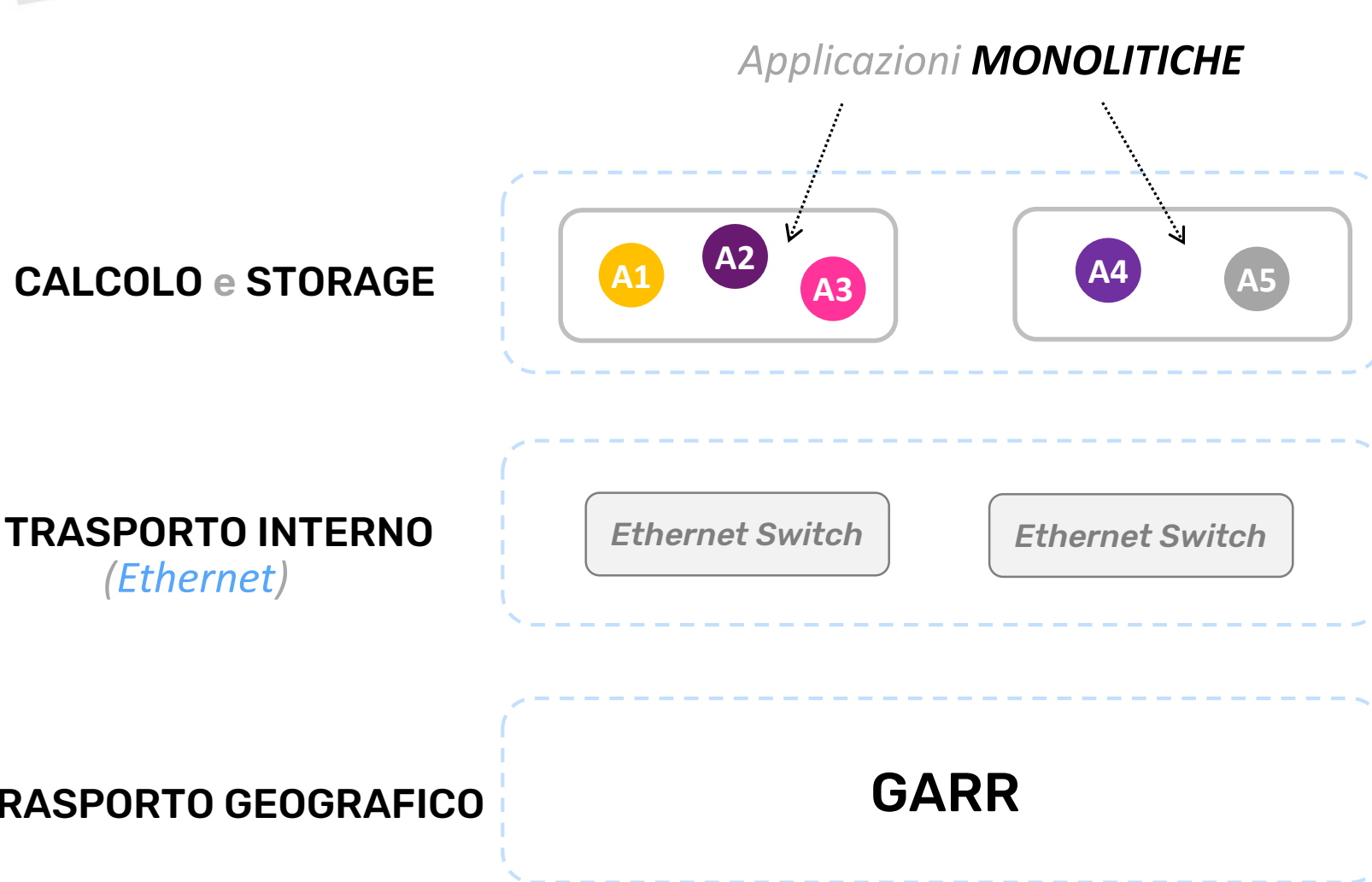


**NET
MAKERS**

Indice

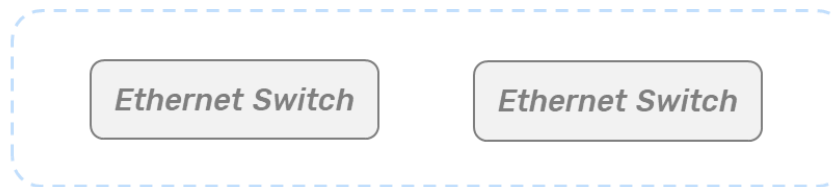
- Rivoluzione applicativa
- Effetti sul Data Center Network
- Nuovo modello di Data Center Network
- Cenni EVPN/VxLAN
- Distributed Data Center
- La soluzione adottata
- Conclusioni

Old Style Data Centers



Old Style Data Centers

TRASPORTO INTERNO
(*Ethernet*)



EDGE

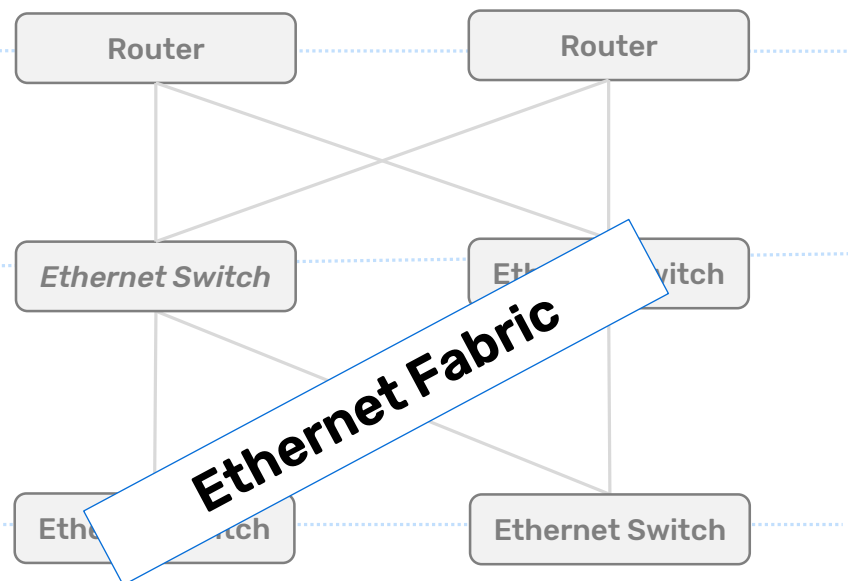
interfacciamento con la rete geografica

SPINE

interconnessione ad alta capacita'

LEAF

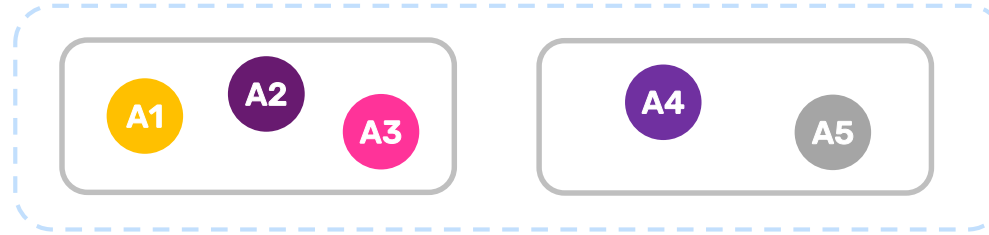
aggregazione Calcolo e Storage



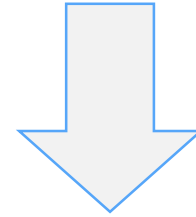
Topologia CLOS

Nuovo paradigma per le applicazioni

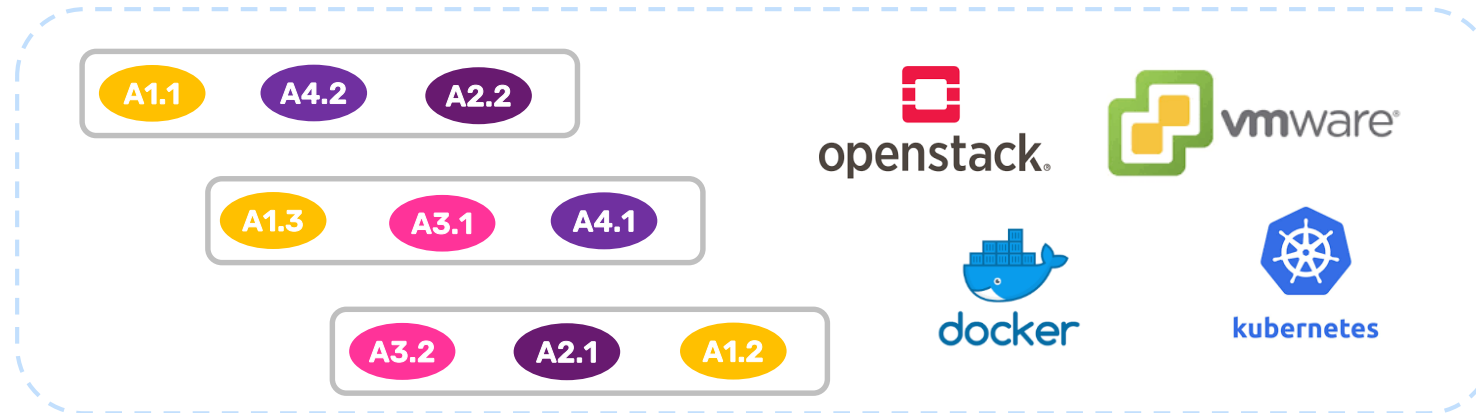
CALCOLO e STORAGE



Approccio **MONOLITICO**

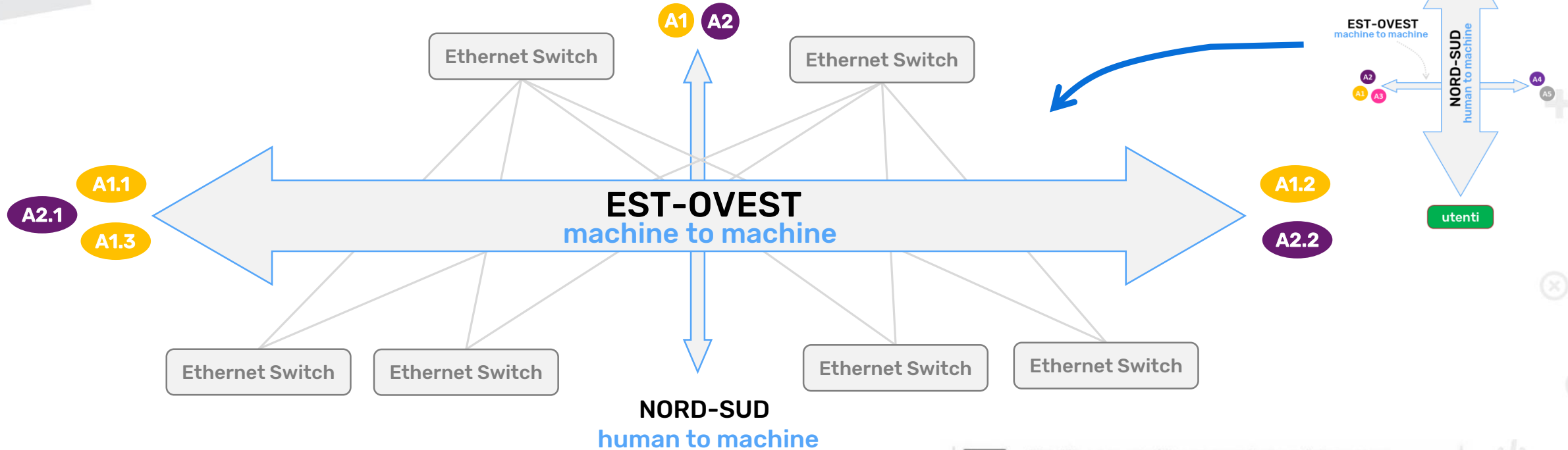


CALCOLO e STORAGE



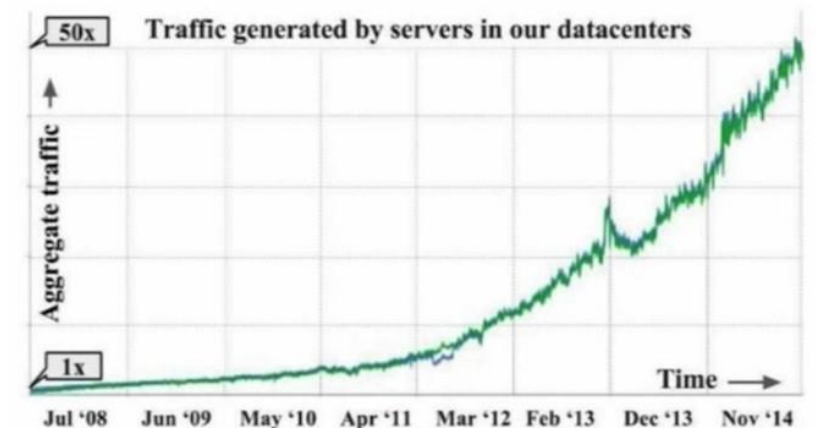
Approccio **DISTRIBUITO**

Nuovo paradigma per le applicazioni



Facebook: traffico MtM e MtU

$$MtT \approx 1000 \times MtU$$

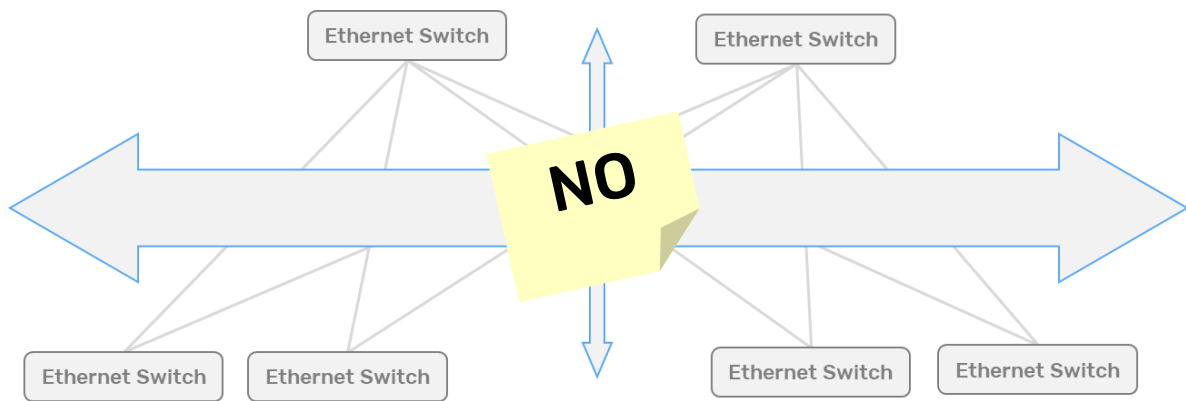


Google: traffico MtM

DOMANDA

Il modello di rete «Ethernet Fabric», è ancora idoneo a supportare le esigenze delle applicazioni sviluppate secondo il nuovo modello disaggregato?

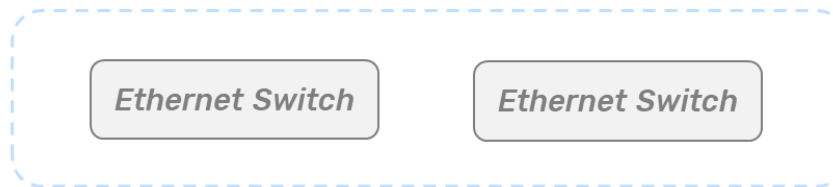
Limiti del modello «Ethernet Fabric»



- Utilizzo non ottimo della capacità di rete (STP mantiene i link in blocking)
- Ridotta scalabilità in termini di aumento della capacità di forwarding;
- Tempi di convergenza lunghi a seguito di guasti (STP converge in ~10sec)
- Estensione geografica L2 (come conseguenza dei μ Servizi) è vincolata ai servizi L2VPN, VPLS, o λ dedicate sulla rete DWDM, offerti dall' operatore (GARR);
- La disaggregazione delle applicazioni porta ad un aumento del numero di domini segregati (max VLAN = 4096)

Il modello «IP-Fabric»

TRASPORTO INTERNO
(*Ethernet*)



EDGE

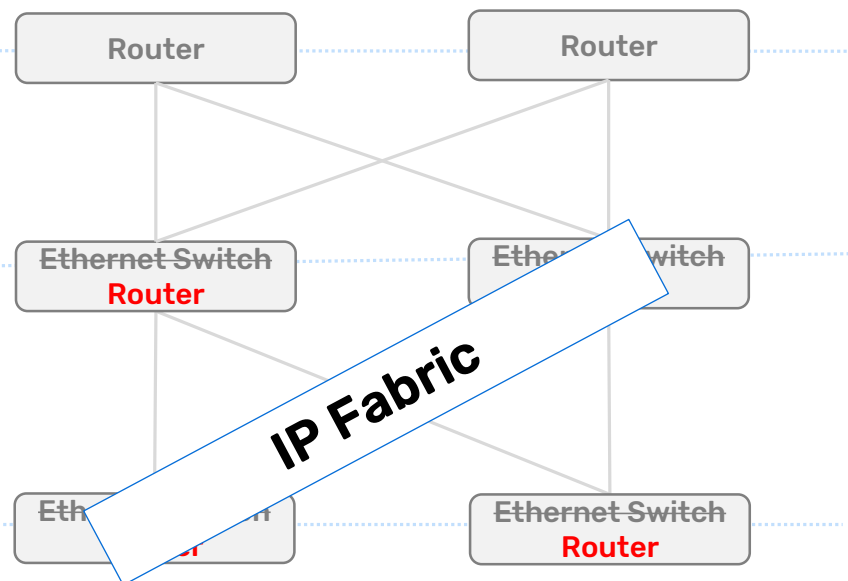
interfacciamento con la rete geografica

SPINE

interconnessione ad alta capacità

LEAF

aggregazione Calcolo e Storage



Topologia CLOS

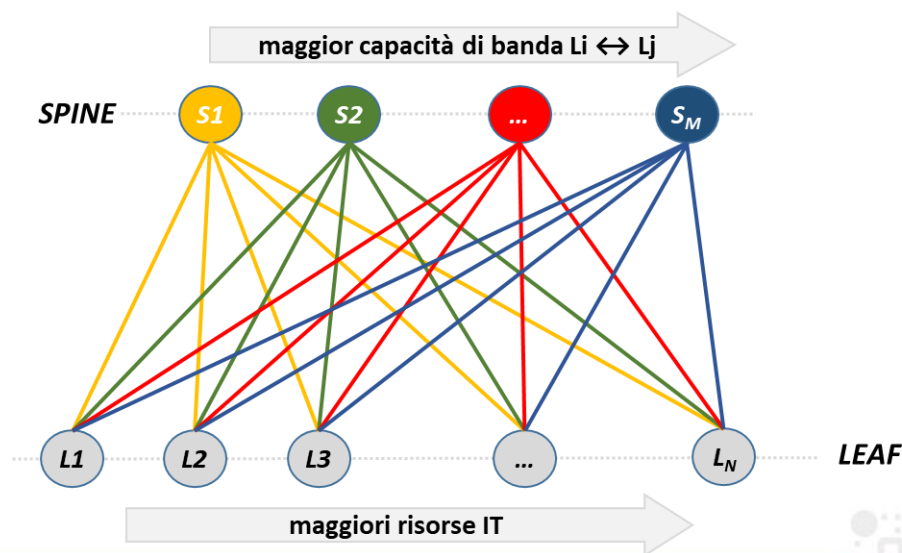
Il modello «IP-Fabric»

Vantaggi

- I protocolli di routing (BGP, IS-IS, OSPF) implementano nativamente meccanismi di impegno dei link Equal Cost (TUTTI i link della Fabric sono utilizzati);
- L3 rende disponibili dei protocolli che implementano funzionalità di Fast Failure Detection e Local Repair che velocizzano i tempi di convergenza (da 10s a 50 ms);
- Più efficace interazione protocollare tra la rete del DataCenter e rete di trasporto nazionale;
- Architettura maggiormente scalabile in termini di capacità di forwarding che di capacità di calcolo e storage

Svantaggi

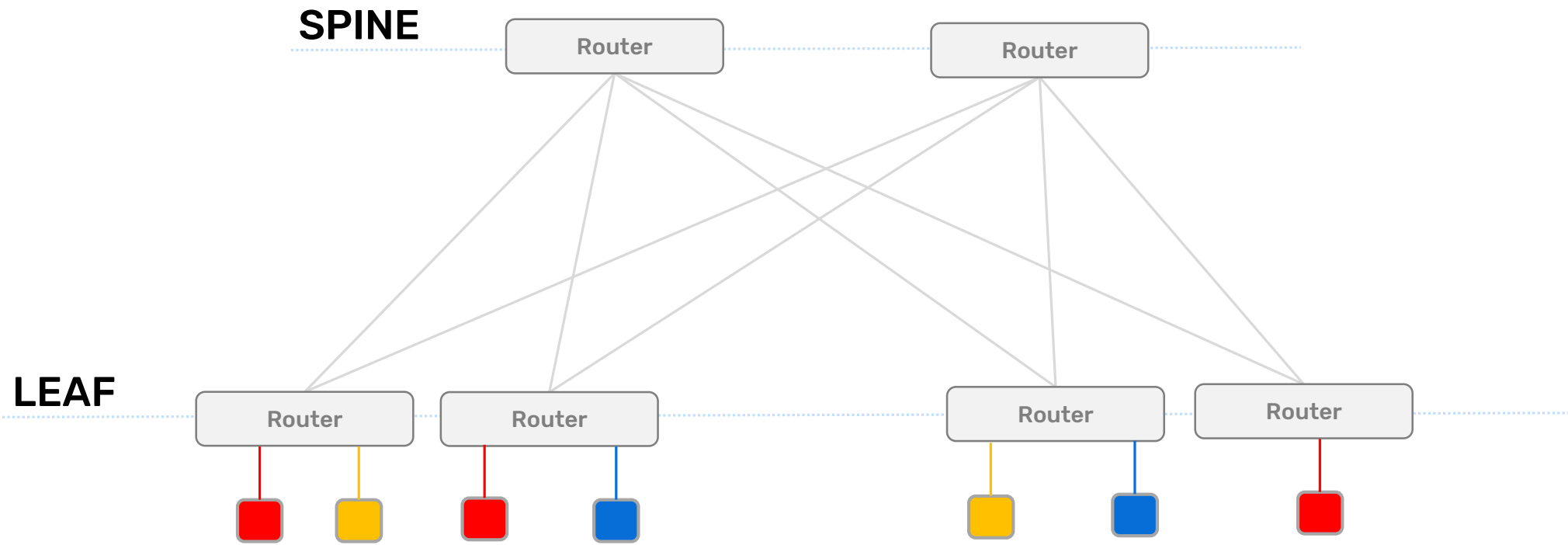
- La Fabric è realizzata con ROUTER, che hanno più funzioni rispetto ad un ethernet switch e, di conseguenza, sono più costosi;
- La Fabric richiede l'abilitazione di diversi protocolli di Livello 3 (es. BGP, IGP, EVPN, VxLAN, ecc.), è quindi ha una configurazione più complessa;
- Una configurazione più complessa richiede una maggiore competenza nella gestione operativa della Fabric;



DOMANDA

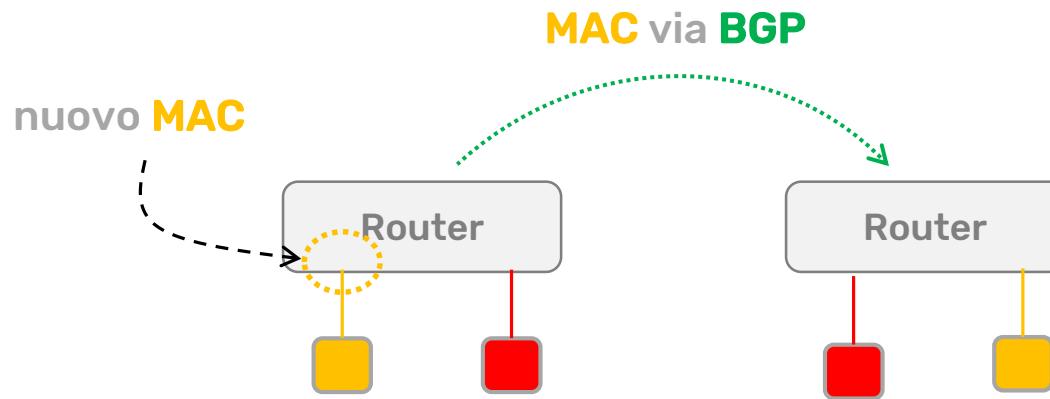
Su una rete fatta da Router, come è possibile estendere i domini di Livello 2 (VLAN) tra tutti i nodi di rete del Data Center?

Modello «IP Fabric» – L2 stretching



Modello «IP Fabric» – L2 stretching

EVPN
Ethernet VPN

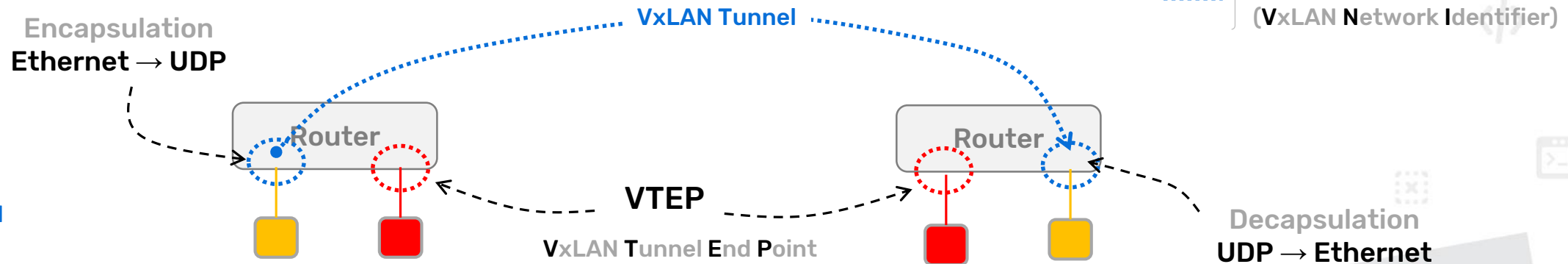


- Estensione del protocollo BGP (**family evpn**);
- BGP per la segnalazione di MAC, IP degli host, VNI per ciascun dominio VxLAN;
- Utilizzato sia in ambito locale (Data Center) che geografico (DCI);
- Introduce nuove funzionalita', utili in ambito DC (MAC mobility, multi homing A/A, ecc.)

↑ **Control Plane**

↓ **Forwarding Plane**

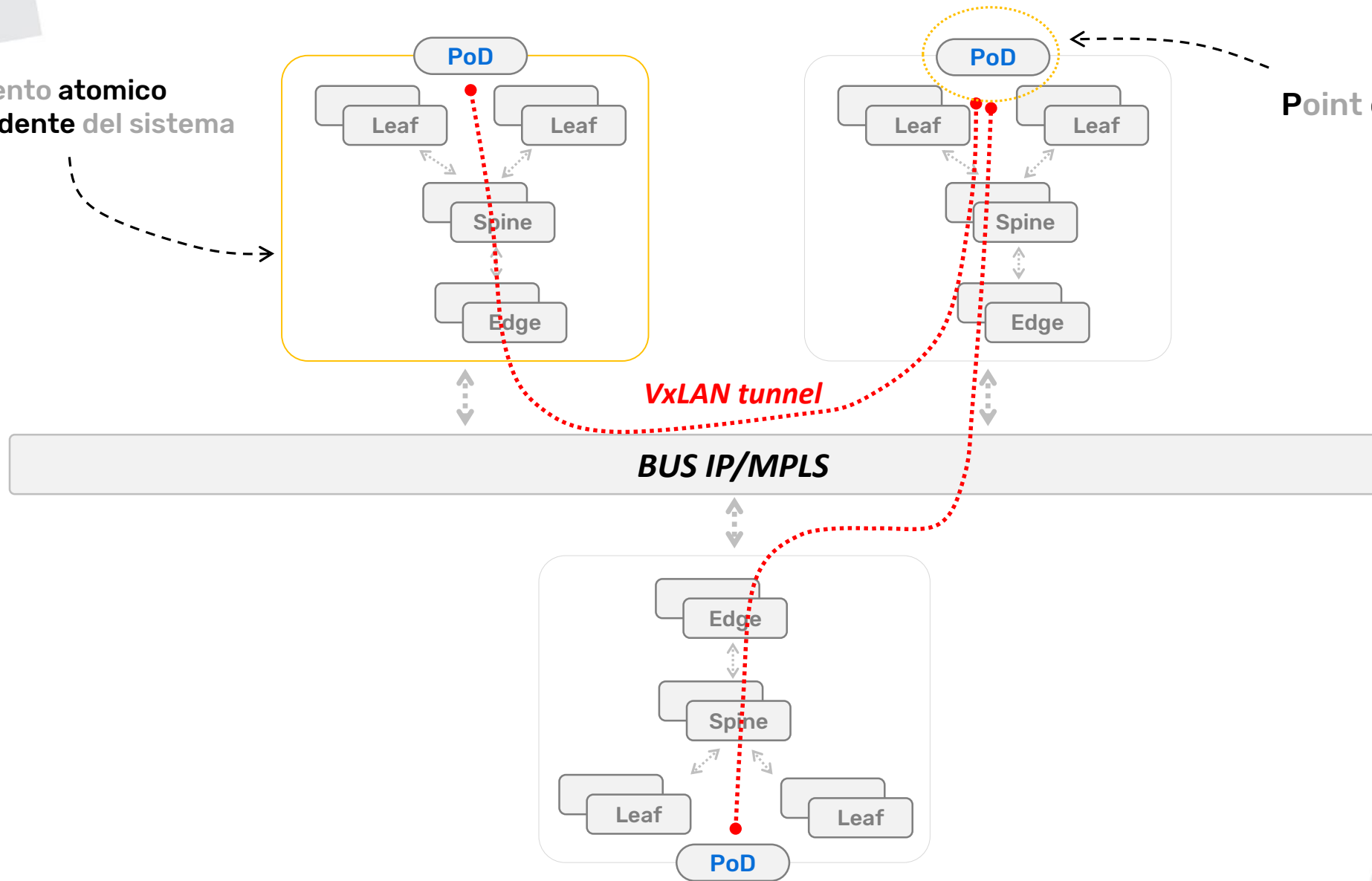
VxLAN
Virtual eXtensible LAN



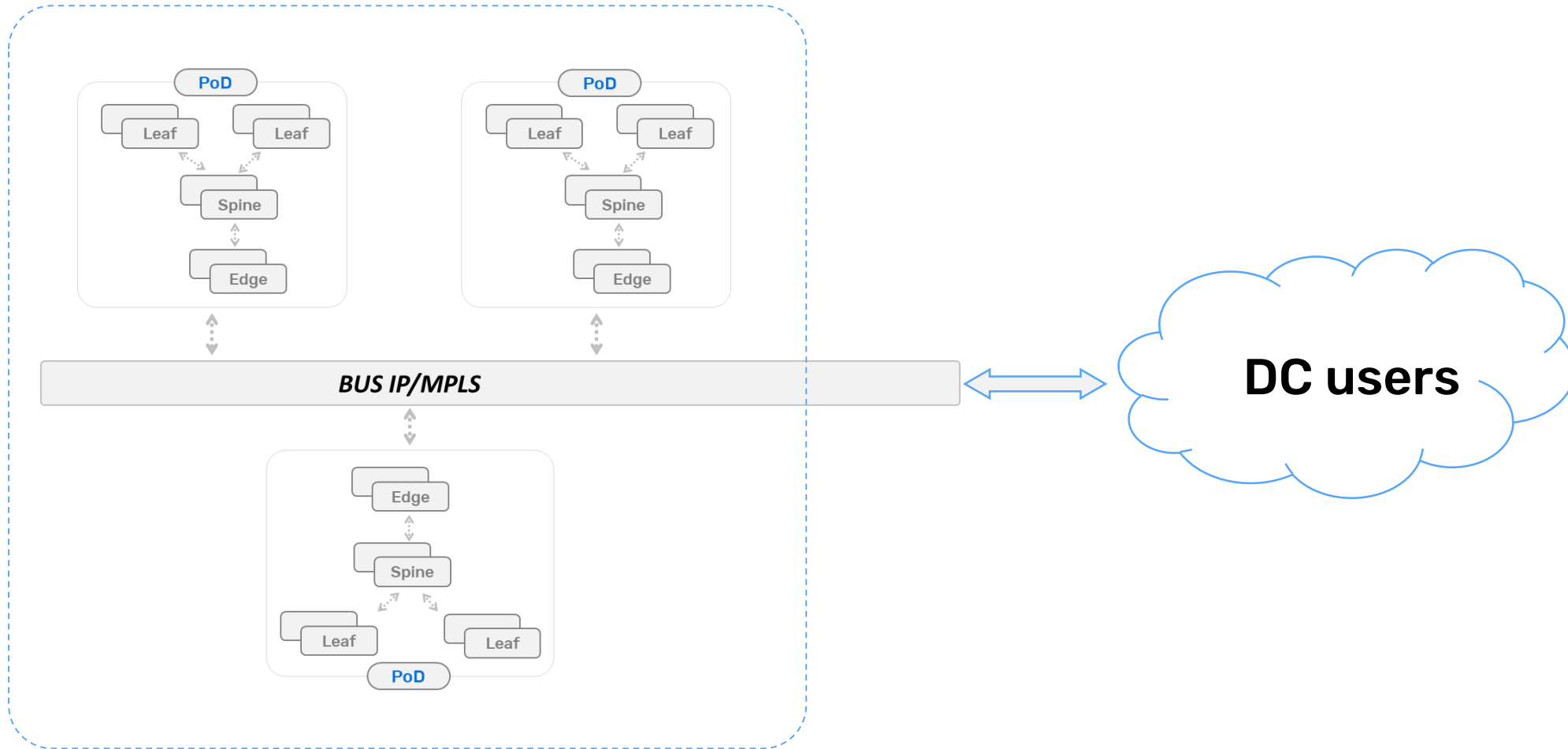
dalla LAN alla WAN: Distributed Data Center

elemento atomico
ed indipendente del sistema

Point of Delivery



dalla LAN alla WAN: Distributed Data Center



Data Center → Sistema Data Center

Evoluzione DC INFRA GARR

Data Center INFRA GARR

Data Center GARR coinvolti: [Roma Tizii](#), [Roma Sapienza](#), [Bari](#), [Bologna](#).

[400 Gbps] **MX-204**
(4x 100G + 8x 10/1G)



EDGE

JUNIPER
NETWORKS

[6.4 Tbps] **7050CX3-32S-F**
(32x 100G)



SPINE

ARISTA

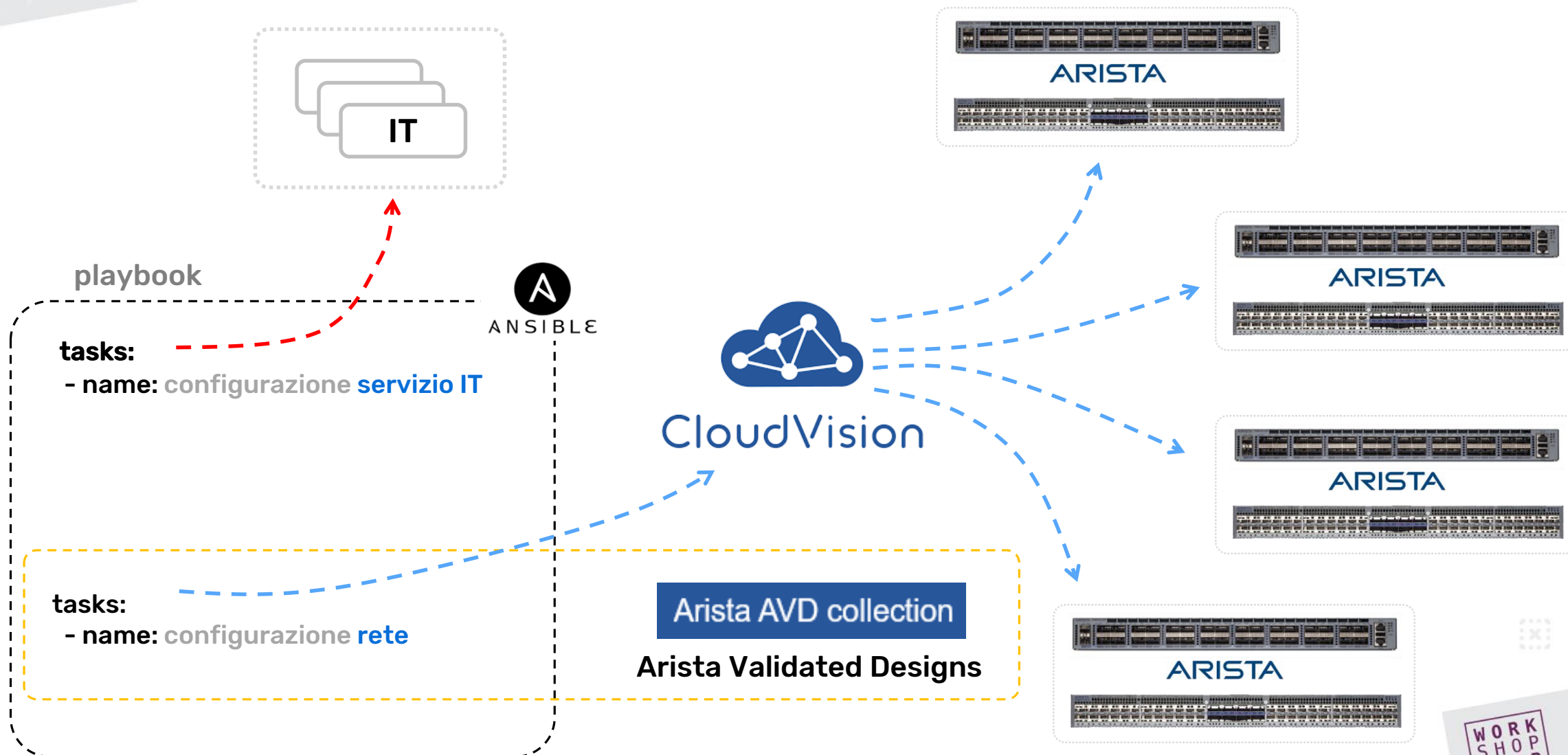
[4 Tbps] **7050SX3-48YC8-F**
(8x 100G + 48 25/10/1G)



LEAF



Data Center INFRA GARR



Conclusioni

Maggiore efficienza del network (1/2)

- **APPARATI PIU' COMPLESSI E COSTOSI:** la maggiore complessita' nell'elaborazione del pacchetto (Ethernet $\leftrightarrow$ UDP), produce una maggiore efficienza con cui avviene il trasporto su una rete IP
- **ETHERNET OVER UDP** consente:
 - una piu' semplice estensione dei domini di L2 tra Data Center remoti (DCI)

Conclusioni

Maggiore efficienza del network (2/2)

- **EVPN E' BGP BASED** e, nell'ottica DCI, questo garantisce piena interoperabilità tra il piano di controllo della rete del Data Center ed il piano di controllo della rete GARR di produzione
- **NON PIU' SPANNING TREE PROTOCOL**, sulla Fabric, con tutti i vantaggi che ne conseguono

Conclusioni

Flessibilit  e nuove opportunit 

- **LA RETE SEMPRE DI PIU' VERSO LE APPLICAZIONI.** La funzione VTEP, che di norma   collocata all'interno di un apparato di rete (Leaf), puo' essere posizionata anche all'interno del sistema di virtualizzazione, e gestita con gli stessi strumenti che operano le applicazioni.
- **4096 DOMINI VLAN → 16M DOMINI VxLAN.** Le applicazioni sviluppate con il paradigma dei μ Servizi, richiedono una maggiore granularita' in termini di segregazione. Con VLAN (VLAN-ID 12 bit) si possono definire 4096 domini, con VxLAN (VNI 24 bit) i domini sono 16M.
- **ANSIBLE.** L'adozione di questo strumento, consente la realizzazione di un modello in cui   lo sviluppo ed il deployment dell'applicazione che pilota direttamente la propria configurazione di rete.

Conclusioni

Punti di attenzione

- La transizione dal modello **Ethernet Fabric** al modello **IP Fabric** richiede:
 - sviluppare maggiori competenze (si passa dalle VLAN+STP al ventaglio di protocolli di Livello3 utilizzati in una IP Fabric)
 - TEMPO per familiarizzare con i nuovi concetti
 - TEMPO per acquisire e padroneggiare le nuove metodologie per gestire ed effettuare troubleshooting su un'infrastruttura piu' complessa



Grazie
(giancarlo.viola@garr.it)

