

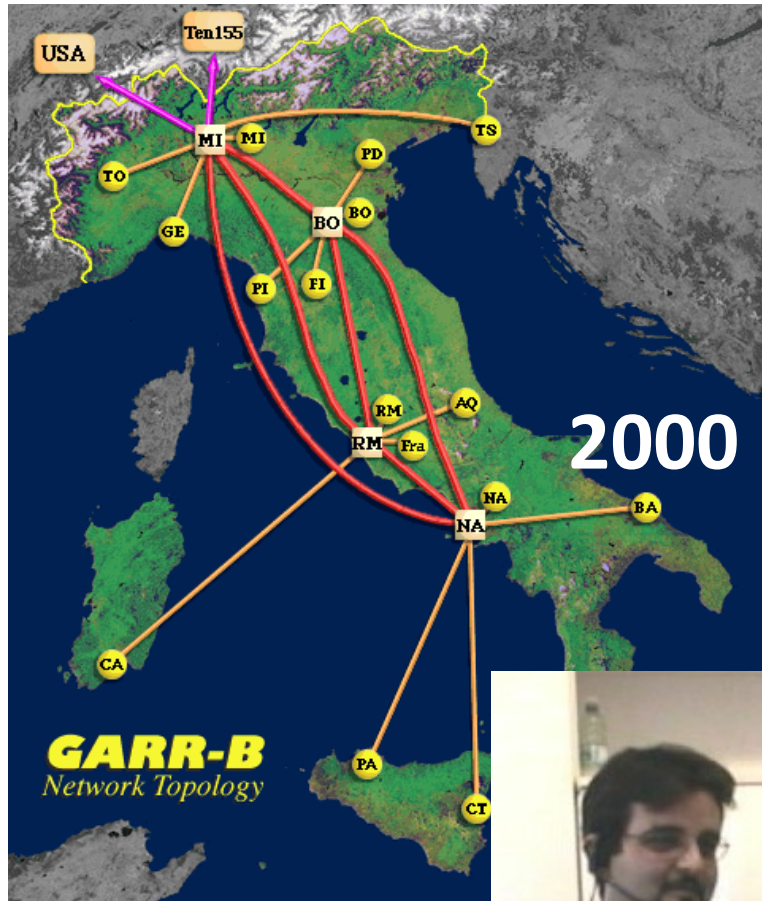
Nuovi scenari della rete IP: come cambiare la rete e vivere felici

MARCO MARLETTA

Roma, 08/10/2019

Workshop GARR 2019

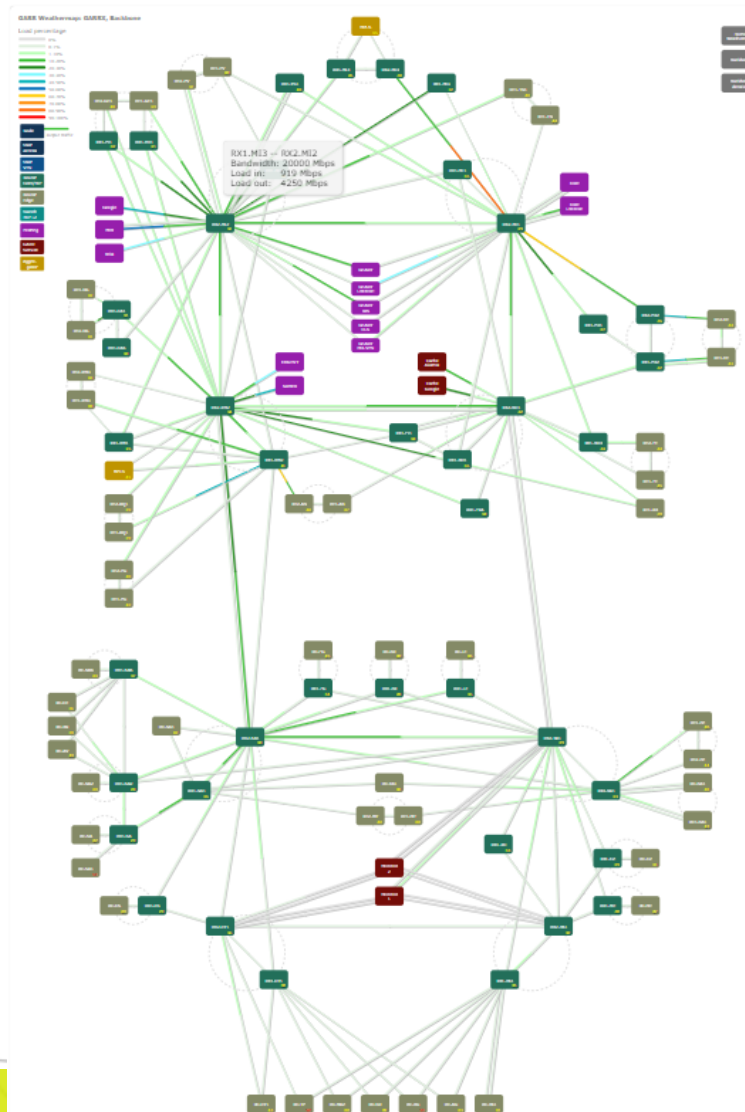
La rete GARR attraverso gli anni



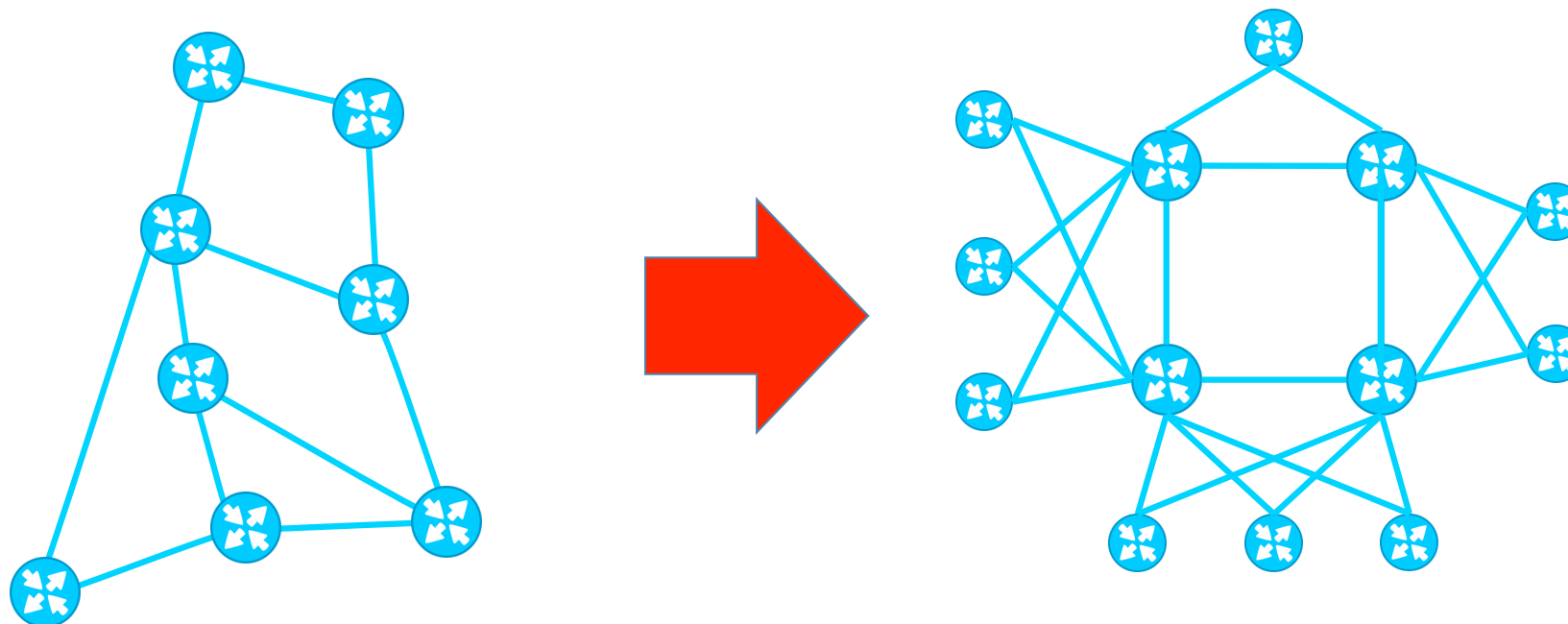
La topologia odierna della rete IP GARR

GARR: this is the actual IP topology

WS19:



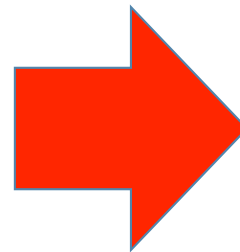
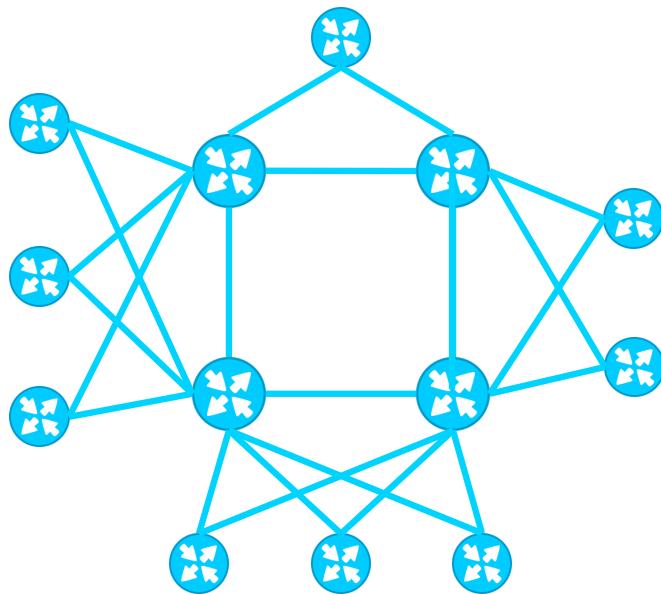
Dalla topologia «geografica» alla topologia «gerarchica»



Passare da una rete basata su circuiti operatore a noleggio a una rete in fibra proprietaria ci ha consentito di cambiare radicalmente topologia e modello di evoluzione:

- I collegamenti di dorsale sono tracciati in funzione delle direttrici di traffico
- L'upgrade di capacità di un PoP non è vincolato all'upgrade di altri PoP
- Il «core» della rete può essere upgradato seguendo la domanda
- Il guasto di un collegamento periferico non causa il reindirizzamento del traffico di altri PoP

Come migliorare ancora la topologia «gerarchica»?



Requisiti della nuova architettura di rete IP

L'infrastruttura fisica sottostante può essere ormai considerata solida, nel tempo potremo aggiungere magliatura, capillarità e (ovviamente) sempre maggiore capacità.

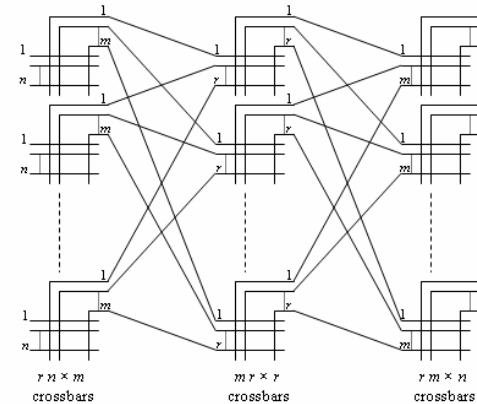
I requisiti che rimangono sul tavolo sono:

- Riuscire a sfruttare al massimo la capacità disponibile
- Aumentare la ridondanza
- Concentrare in maniera efficiente un enorme numero di collegamenti
- Supportare tutti i servizi L2/L3 attuali e anche nuovi servizi
- Aggregare da 1G a 100G, possibilmente anche 400G sul backbone
- Cercare di ridurre MOLTO consumi e spazi rispetto alla situazione odierna

E poi abbiamo bisogno di strumenti per:

- nuove capacità di monitoring (Telemetria, syslog monitoring)
- funzionalità di traffic engineering più evolute
- automatizzare:
 - Operazioni quotidiane
 - Provisioning dei servizi
 - Upgrade di release

Ripartiamo dalle basi: la teoria della commutazione

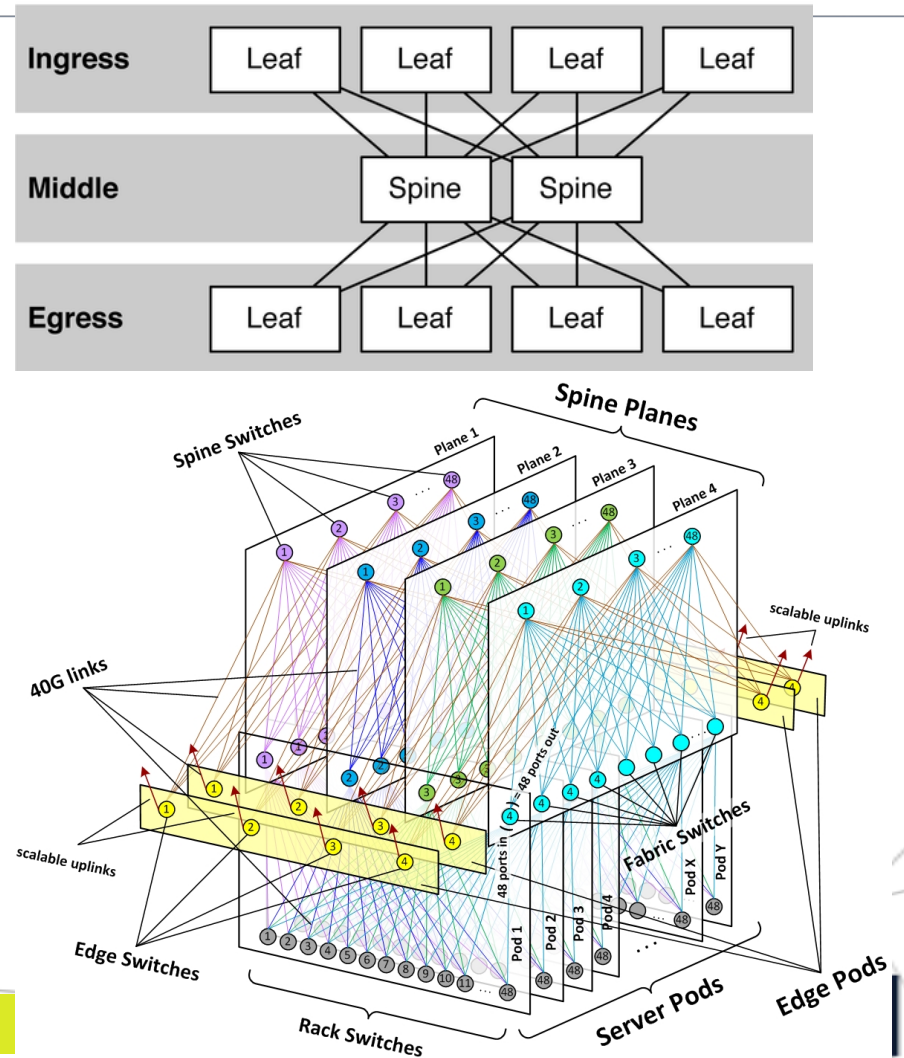


- La matrice di commutazione è realizzata mediante interconnessioni fisiche fra apparati che mettono in comunicazione N ingressi ed M uscite
- E' il sistema comunemente usato nelle reti a commutazione di circuito
- Si riesce a tenere sotto controllo il numero di crossconnessioni realizzando topologie a più stadi
- Ma noi abbiamo un vantaggio: la commutazione di pacchetto...

La topologia CLOS

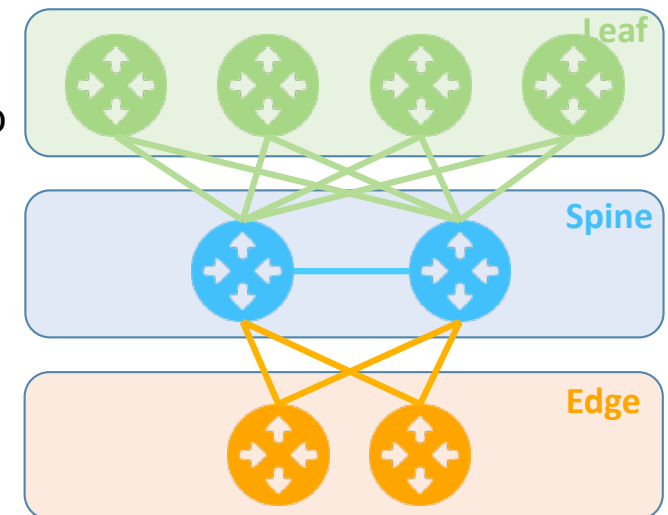
- Ogni stadio di commutazione intermedio permette di minimizzare il numero di crossconnessioni fra N ingressi ed M uscite
- La topologia permette inoltre:
 - l'overbooking ($N > M$)
 - load balancing su più percorsi in condizioni normali
 - resilienza in caso di guasto di un link o di uno stadio di commutazione (il traffico si ridistribuisce sui link rimanenti)
- Si è affermata nel tempo come principio per le reti dei datacenter (complicandola a piacere..)
- Si sta affermando anche per le reti geografiche
- Ma richiede molti più apparati e molti più link (circa 2x) rispetto ad oggi

3-Stage Clos Network

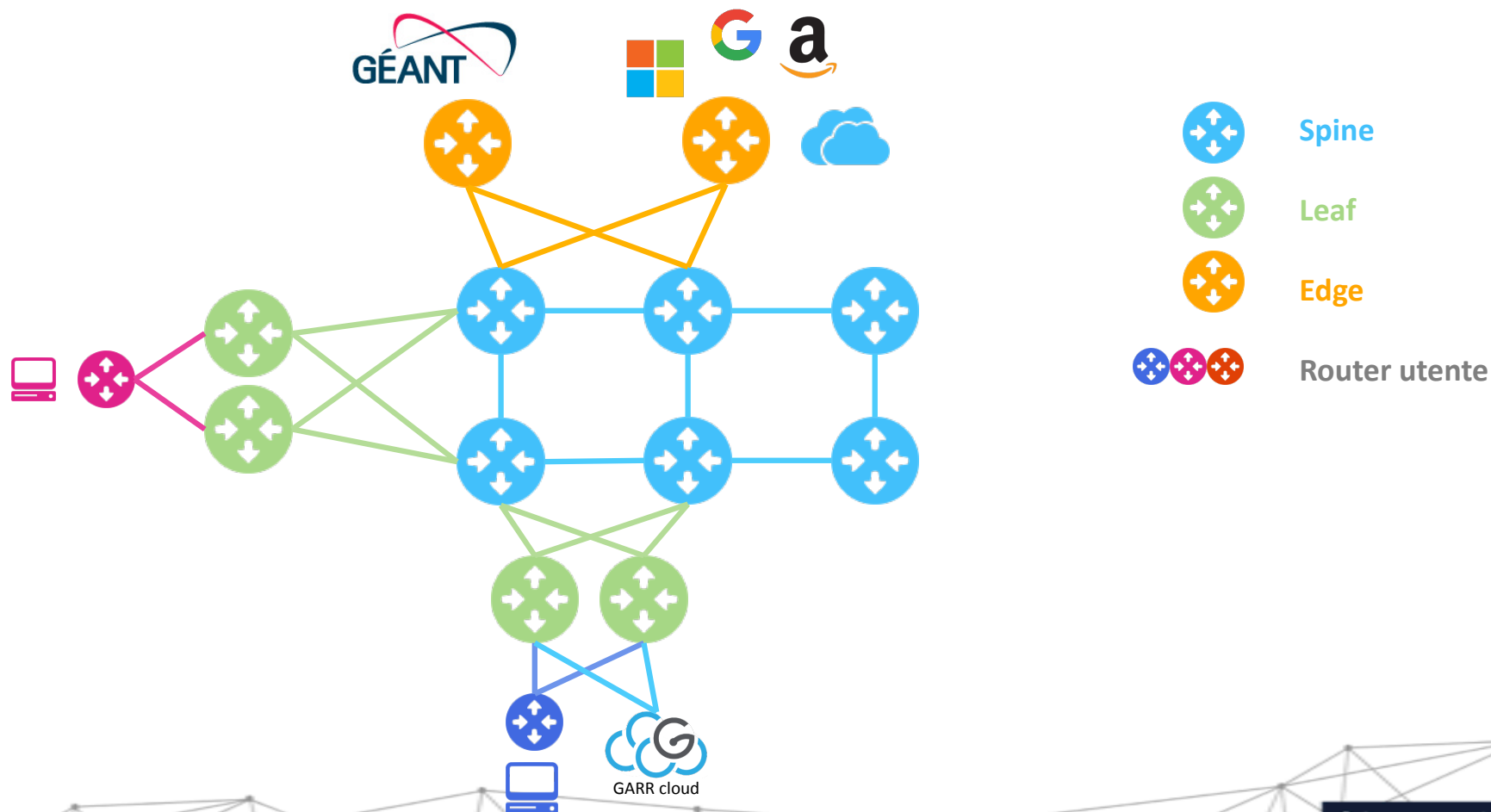


Spine, Leaf, Edge

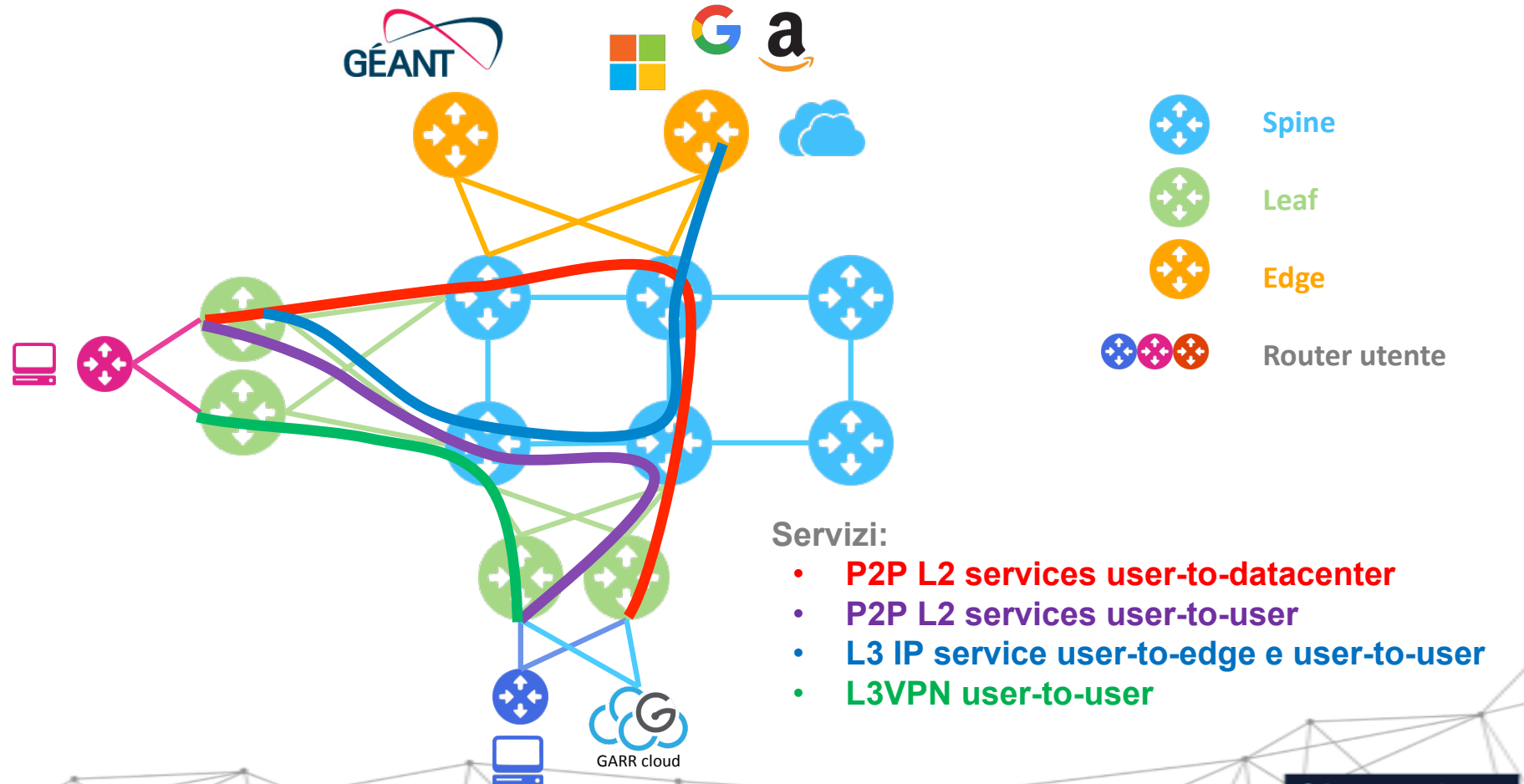
- Gli apparati SPINE realizzano lo stadio di commutazione della topologia CLOS
 - Hanno molte porte ad alta velocità
 - Non serve che abbiano le funzionalità necessarie per:
 - Raccolta utenti e servizi di edge
 - Routing internet e transiti
- Gli apparati LEAF realizzano lo stadio di ingresso della rete (N ingressi)
 - servono a interconnettere gli utenti alla rete
 - «generalmente» sono il punto in cui inizia l'erogazione di un servizio
 - Hanno la gran parte delle feature avanzate
 - Accounting traffico;
 - Passive monitoring (netflow, telemetry);
 - Filtering, policing;
 - QoS;
 - VPN;
 - Bridging;
 - Gestione ridondanze locali per l'accesso
- Gli apparati EDGE realizzano lo stadio di uscita della rete (M uscite)
 - gestiscono le interconnessioni verso l'esterno
 - Funzionalmente sono identici alle LEAF
 - Hanno molta più memoria e molte porte ad alta velocità



Modello topologico

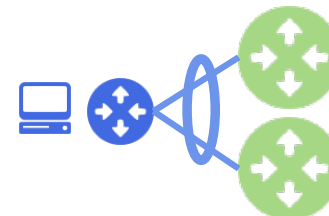
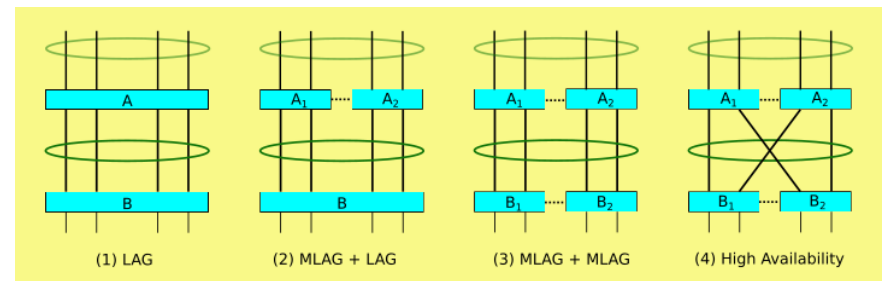
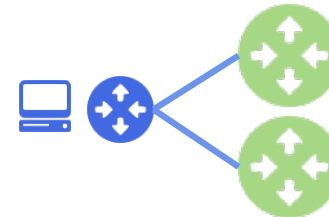


Architettura dei servizi

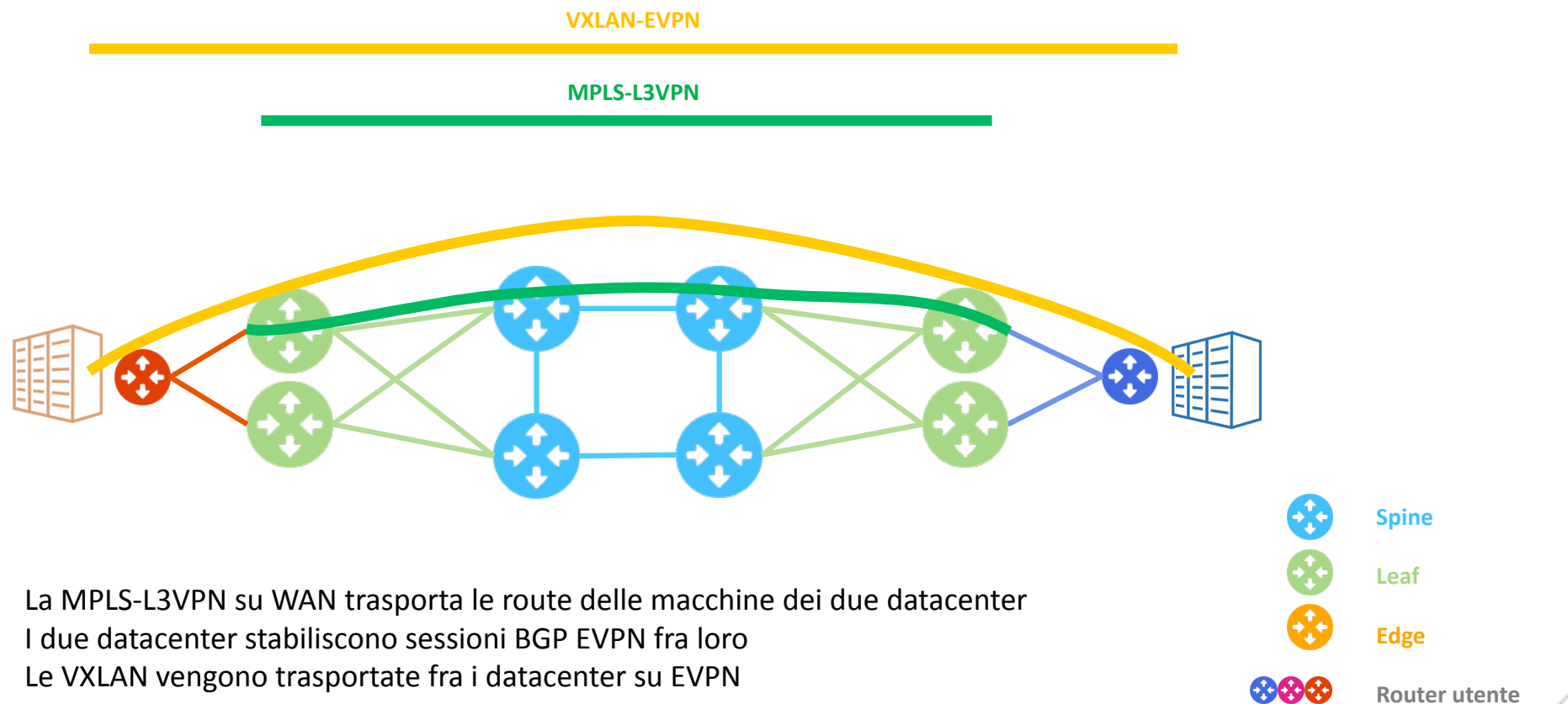


L'accesso degli utenti alla rete

- Tutte le modalità attuali di accesso rimarranno possibili
- Ethernet 1G / 10G / 100G
- Dove sarà realizzabile, permetteremo un doppio accesso su due apparati (leaf) GARR diversi, in modo da ottenere maggiore fault tolerance
- Modalità di gestione del doppio accesso
 - Active/standby tradizionale (BGP/routing statico)
 - Active/active tradizionale (BGP/routing statico)
- LAG / MLAG / MLAG + LAG / MLAG + MLAG
- EVPN multihoming (all-active)

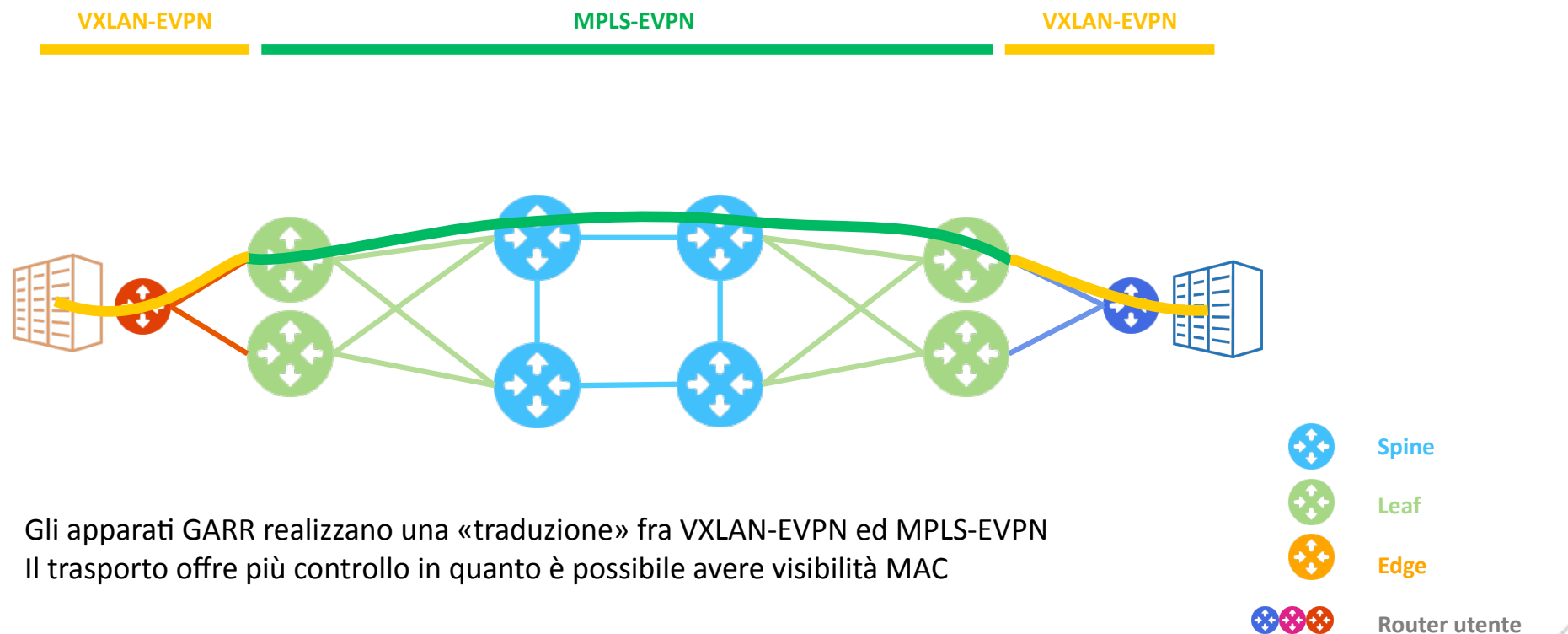


Nuovo servizio DCI – Option 1 - L3VPN su WAN, VXLAN-EVPN fra DC



- La MPLS-L3VPN su WAN trasporta le route delle macchine dei due datacenter
- I due datacenter stabiliscono sessioni BGP EVPN fra loro
- Le VXLAN vengono trasportate fra i datacenter su EVPN

Nuovo servizio DCI – Option 2 - MPLS-EVPN su WAN, VXLAN-EVPN fra DC



- Gli apparati GARR realizzano una «traduzione» fra VXLAN-EVPN ed MPLS-EVPN
- Il trasporto offre più controllo in quanto è possibile avere visibilità MAC

Due parole sulle MAN

- Utilizzare apparati più compatti permette di diffonderli quasi ovunque
- Dove abbiamo usato nodi di aggregazione tramite switch metro (ME3400, ME3600) stiamo valutando di installare nuovi apparati «leaf»
- Il modello di raccolta delle MAN GARR (Firenze, Venezia) inoltre passerà da Layer2 a Layer3
- Questo vuol dire anche un nuovo modello di architettura dei servizi, che potranno essere direttamente erogati sulle leaf, oppure trasportati sul backbone
- Lo stesso meccanismo di trasporto sul backbone potrebbe essere esteso anche alle MAN che non sono sotto gestione diretta GARR, prendendo in carico i servizi sul bordo GARR-MAN e trasportandoli altrove.

Considerazioni finali

1. Il modello di rete a pacchetto IP/MPLS rimane valido e continuerà a rimanere il modello di riferimento ancora per molto tempo
2. L'integrazione con strumenti di automazione, analitica e di provisioning favoriranno lo sviluppo di nuove applicazioni e servizi
3. E' possibile andare verso uno scenario con consumi e spazi ridotti
 - Si potranno aggiungere pop anche in luoghi meno attrezzati
4. Un servizio affidabile è realizzabile anche con apparati più semplici
 - Si investe sulla moltiplicazione di link e di apparati invece che sulla ridondanza del nodo
 - È più semplice aggiungere altri apparati che prevedere apparati espandibili



FINE

Questa attività è stata condotta insieme a:
Massimo Valiante
Giancarlo Viola

ME NOW:



inglip.com