Seminario
Utilizzo dei protocolli di routing nelle reti per dati a supporto del calcolo ad alte prestazioni

Università di Bologna
DEIS – Laboratorio di Reti di Telecomunicazione
25 Maggio 2009

Ing. Vincenzo Vaccarino
Dipartimento Sistemi e Tecnologie - Settore Operazioni
CINECA

Agenda

- Cosa è il CINECA
- Protocolli di routing
- La struttura di rete CINECA
- Il routing nella struttura di rete CINECA

Agenda

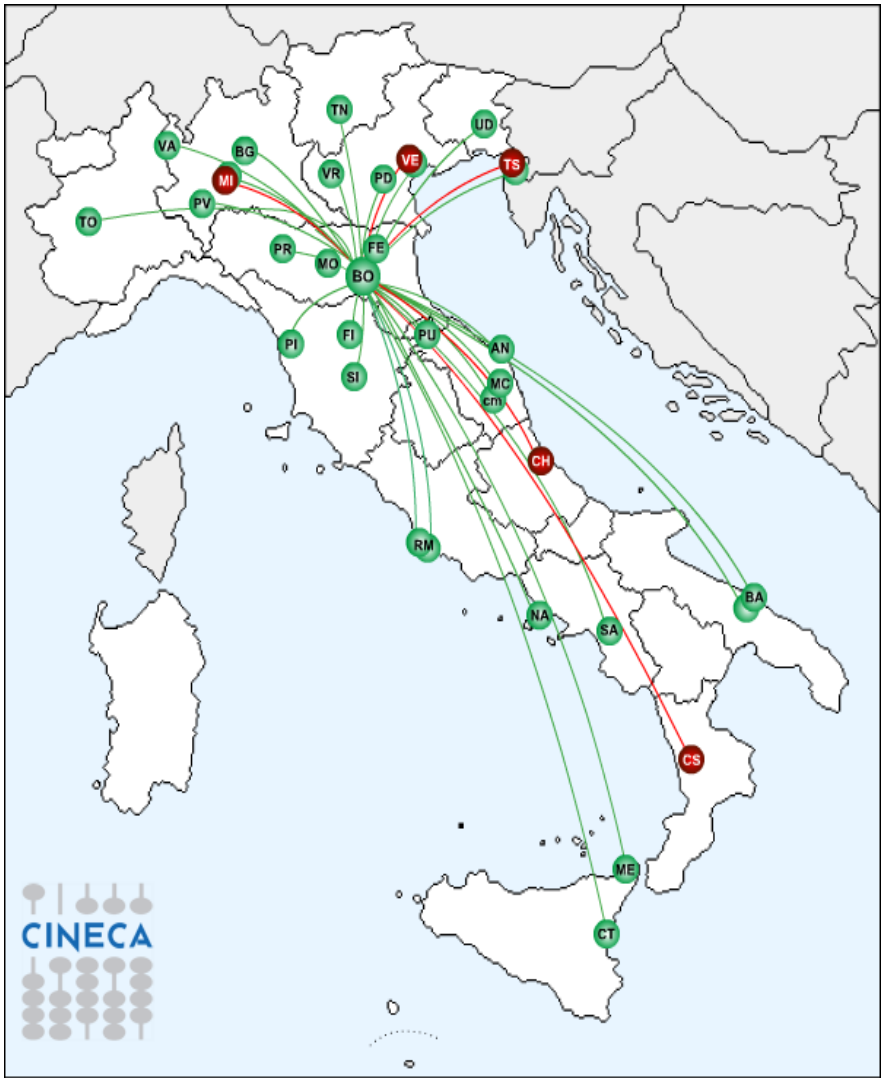
- Cosa è il CINECA
- Protocolli di routing
- La struttura di rete CINECA
- Il routing nella struttura di rete CINECA

CINECA

- Nasce nel 1969
- La prima funzione è la condivisione delle risorse di calcolo fra gli atenei consorziati
- Negli anni successivi si sviluppano ulteriori attività di supporto ai consorziati, il cui numero aumenta progressivamente

36 atenei:
Bari, Bari Politecnico,
Bergamo, Bologna, Calabria,
Camerino, Catania, Cassino, Chieti,
Ferrara, Firenze, Insubria,
Messina, Macerata,
Milano Biccocca, Milano Politecnico,
Modena, Napoli Federico II, Padova,
Parma, Pavia, Perugia, Pisa, Politecnico delle Marche, Mediterranea di Reggio Calabria, Salerno, Siena,
Torino, Politecnico di Torino, Trento, Trieste, Udine,
Urbino, Venezia, Venezia IUAV, Verona

il CNR
l'OGS
Il MiUR



Attività istituzionali

- Calcolo ad alte prestazioni, per utenze pubbliche (accademiche) e private
- Servizi gestionali a supporto delle Università
- Servizi gestionali a supporto del MIUR
- Partecipazione al progetto DEISA

Attività istituzionali

Trasferimento tecnologico verso:

- Pubblica Amministrazione ed Enti Locali
- Sanità
- Industrie
- Unione Europea

Altre attività

Servizi commerciali:

- Hosting e housing per clienti privati (Datacenter commerciale)

Laboratori

- TV Digitale, servizi multimediali, E-learning

Qualche numero

Risorse umane : 370 persone circa

Nella classifica dei 500 centri di calcolo a livello mondiale, CINECA occupa il 46. posto (novembre 2008)

- 2 Cluster Bladecenter HS21 da 5120 processori
- Cluster BCX, da 10240 processori

Infrastruttura di rete

CINECA ha recentemente intrapreso un progetto di miglioramento della propria infrastruttura di rete, che ha visto una completa riprogettazione della stessa

Il progetto ha ricevuto uno dei premi del Computerworld Honors Program per il 2008

In questo momento è in fase di progressivo completamento la migrazione di tutta l'infrastruttura al nuovo modello

Riferimenti online

Sito istituzionale

http://www.cineca.it

Riferimenti sulle strutture di supercalcolo e di rete

http://www.top500.org
http://www.top500.org/sites/273
http://www.cwhonors.org/laureates/2008laureates.htm
http://www.cwhonors.org/viewCaseStudy2008.asp?NominationID=722

Agenda

- Cosa è il CINECA
- Protocolli di routing
- La struttura di rete CINECA
- Il routing nella struttura di rete CINECA

Protocolli di routing

- OSPF (Open Shortest Path First)
- BGP (Border Gateway Protocol)

OSPF – Basi

Il protocollo OSPF nasce per supplire ad alcuni problemi dei protocolli di routing esistenti:

- Protocollo *link-state*
- Utilizzo dell'algoritmo di Dijkstra, detto anche SPF
- Standard non proprietario (OSPF = Open SPF)

OSPF – Protocolli Link-state

I protocolli *link-state* hanno una serie di caratteristiche comuni, che si riflettono anche in OSPF:

- Ogni router ha una mappa completa della topologia di rete, da cui ricava le informazioni per il routing
- Buona velocità di reazione ai cambiamenti topologici
- Invio di aggiornamenti solo all'apparire di cambiamenti e riduzione dei refresh

OSPF – Principio di funzionamento link-state

- Ogni router riceve dagli altri router della rete le informazioni relative alle reti di cui fanno parte
- Crea una tabella con tutte le informazioni ricevute
- Mediante SPF, genera un albero con i costi di ogni percorso
- Seleziona per la tabella di routing soltanto le rotte con i costi inferiori

OSPF – Strutture dati

- Neighbor table: contiene le informazioni relative ai router con i quali ci si scambiano le informazioni sulle rotte (vedremo dopo in che modo)

- Tabella topologia (LSDB – Link State DataBase): contiene tutti i record ricevuti dagli altri router, in modo da costruire la topologia completa della rete

- Routing table: sulla base dei costi calcolati con SPF, vengono inserite le rotte per le singole reti

OSPF – Pacchetti di comunicazione (Link-state)

La comunicazione fra i router dipende dallo scambio di pacchetti, definiti LSA (Link-state advertisement):

- Gli LSA vengono trasmessi via multicast, per raggiungere tutti i router dell'area
- Gli LSA vengono accettati, ed i dati contenuti sono inseriti nel DB della topologia (LSDB)
- Vengono poi inviati a tutti i propri vicini (neighbor), in modo che in breve tempo tutti i router abbiano lo stesso DB
- Ognuno di essi poi aggiorna il proprio albero dei path ed eventualmente la tabella di routing

OSPF – Pacchetti di comunicazione

- Vista l'importanza degli LSA, esiste un criterio per l'acknowledge
- Vengono diffusi, via multicast, in tutta l'area o dominio
- Hanno un numero di sequenza ed un tempo di vita
- Non devono mai essere rispediti al router che li ha inviati (*split horizon*)

Link-state – Pro e contro

Vantaggi

- Ogni router costruisce la topologia di rete e la tabella di routing basandosi su una conoscenza completa della rete
- Il routing non viene basato sulle informazioni dei vicini (routing by rumours) ma viene calcolato direttamente

Svantaggi

- Le strutture dati sono contenute in ogni router, per cui il consumo di risorse è spesso non ottimizzato
- Alcuni criteri di aggiornamento comportano un traffico eccessivo sulla rete, e le correzioni portano a compromessi (come vedremo)

Rimedi e compromessi

Rimedi introdotti per sopperire ai problemi evidenziati in precedenza:

- Architettura a due livelli
- Introduzione di una divisione gerarchica e topologica
- Introduzione di una gerarchia fra apparati comunicanti

Architettura a livelli – aree

Gli apparati che costituiscono una rete OSPF vengono divisi in gruppi, denominati aree. Le tipologie sono due:

- Area di Transito (*Transit Area*) – ha il solo scopo di interconnettere altre aree e di farne passare il traffico. Non ha di solito utenze collegate

- Area Regolare (*Regular Area*) – è un area che contiene le utenze, e fa passare solo traffico ad esse destinato

Aree OSPF

- Esiste sempre una area di transito, denominata Area 0, detta anche *backbone area*
- Le altre aree comunicano tra di loro attraverso l'Area 0
- Il numero di router per ogni area è altamente variabile, ma le linee guida consigliano un massimo di 50 router
- I router di collegamento fra Area 0 e aree normali sono detti ABR (Area Border Router)

OSPF – Gli ABR

Gli ABR hanno diverse funzioni particolari:

- Aggregano le tabelle degli indirizzi della propria area, rendendola disponibile all'esterno (altre aree)
- Solitamente contengono le rotte di default
- Inviano LSA su due aree, la propria e l'area 0, ed hanno strutture dati doppie, una per ogni area a cui sono collegati

Aree OSPF – Pro e contro

- Le aree OSPF permettono di avere routing table ridotte per i diversi dispositivi
- Le reti scalano su dimensioni maggiori in maniera efficiente
- I cambi di topologia avvengono in una sola area, all'esterno sono diffusi per mezzo di *summary*
- Si perde la precisione legata alla conoscenza della topologia completa della rete

OSPF – Calcolo della metrica

- In ogni router viene eseguito il calcolo del costo di ogni percorso possibile, mediante l'algoritmo di Dijkstra, prendendo il router stesso come "radice" dell'albero risultante

- Per ogni destinazione, si inserisce in routing table solo la rotta con il costo minore

- Per default, il costo dipende dalla larghezza di banda calcolata per ogni interfaccia, ma questo comportamento si può modificare

Gerarchia dei router

- Su un link WAN, i router sono due e possono scambiarsi le informazioni senza duplicazioni

- Su un mezzo condiviso, ad ogni update, tutti i router spedirebbero lo stesso aggiornamento a tutti gli altri router presenti

- Vengono quindi eletti un DR (*designated router*) e un BDR (*backup DR*)

- Gli update vengono inviati **solo** al DR, che provvede poi ad informare gli altri dispositivi

OSPF in ambiente condiviso

La comunicazione in OSPF avviene tramite 5 tipi di pacchetti

- I pacchetti di tipo *hello* permettono di stabilire le adiacenze e di eleggere, ove necessario, DR e BDR, e verificare la connessione fra neighbor

- Solo dopo la loro identificazione, si può procedere allo scambio di informazioni, per popolare i DB

- Gli altri pacchetti: DBD, LSR, LSU, LSAck

OSPF in ambiente condiviso

- I primi a sincronizzarsi sono DR e BDR, vengono poi anche tutti gli altri router,verso cui DR diffonde gli annunci

- Solo alla fine di questo processo, i router sono in stato “full” e cominciano a ruotare il traffico

- In questa fase, i database della topologia dei router sono tutti uguali

OSPF – Uso di LSA e LSU

- La criticità dei pacchetti di informazione (LSA) e aggiornamento (LSU) è legata alla loro consistenza con la situazione reale

- Quando un router riceve una LSA, segue uno schema preciso:

- Se non la conosce, invia un ACK e la inserisce nel DB
- Se la conosce, ne verifica l'aggiornamento: se è vecchia, manda una versione aggiornata al mittente, se è nuova la inserisce nel DB e manda ACK, altrimenti la ignora

OSPF – Uso di LSA e LSU

- Quando un router vede che un record del DB è alla fine del suo tempo di vita, ne rispedisce le informazioni, cambiando solo il numero di sequenza (default è di 30 minuti)

- Le LSU sono usate per le variazioni

- Il router responsabile la invia ai soli DR e BDR, mediante multicast

- DR diffonde il cambiamento in tutta l'area, sempre in multicast, e attende ACK da tutti

OSPF – Uso di LSA e LSU

- I router connessi a più aree, inviano le LSU ai DR di tutte le aree interessate

- I router che ricevono la LSU, aggiornano il DB, ricalcolano l'albero SPF ed eventualmente aggiornano le route

- Una entry nel DB viene rimossa dopo che per 60 minuti non si ricevano refresh

- Ogni tipologia di rete (P2P, condiviso) ha intervalli tipici diversi

OSPF – Problemi di dimensione

Le reti di dimensioni troppo grandi presentano problemi particolari:

- SPF gira troppo spesso
- La Routing Table diviene troppo grande
- Anche LSDB diventa troppo grande, perché contiene anche route non strettamente “utili”

BGP – Nozioni base

- BGP è acronimo di Border Gateway Protocol
- E’ un protocollo dedicato al routing fra Autonomous System
- Si può considerare un Autonomous System come un’insieme di apparati di rete sottoposti ad un’unica amministrazione, tecnica e amministrativa
- Gli AS sono assegnati da una ente apposito (IANA)

BGP – Natura del protocollo

- E’ un protocollo *path vector*
- Le informazioni sulla topologia sono tratte dai router vicini
- Le decisioni sono basate su *policy routing*
- Il controllo delle politiche sono basate sugli *attributi* BGP

BGP – Particolarità

- Path vector: le informazioni scambiate sono una lista di AS da attraversare per raggiungere le varie reti presenti negli AS

- Ogni AS può “annunciare” solo le reti che contiene, non influenza il routing degli AS vicini

- Non si accettano update per il proprio AS, escludendo così i loop

BGP – Caratteristiche

- Si basa su TCP per affidabilità della consegna dei messaggi
- Scala meglio su grandi dimensioni
- Performance: deve garantire buone prestazioni quando cambia la topologia di rete
- Permette il multihoming (es. utilizzo diversificato del collegamento attraverso più ISP)

BGP – Quando usarlo

L'uso del BGP è indicato in alcune situazioni particolari:

- Necessità di far transitare traffico
- Multihoming
- Necessità di manipolare l'andamento del traffico in ingresso/uscita

BGP – Quando non usarlo

Non è il caso di usarlo quando:

- Scarsa conoscenza dei meccanismi di filtraggio e definizione delle rotte
- Connettività singola
- Hardware poco potente come processore e/o memoria

BGP – IBGP e EBGP

BGP può essere usato anche come protocollo di routing all'interno di una rete, o di un AS

In tal caso, si distinguono IBGP (*Internal BGP*) e EBGP (*External BGP*)

Usandoli entrambi, il rischio maggiore è trasformare il proprio AS in un AS di transito, per traffico non voluto

BGP – Altre proprietà

- La versione corrente, adottata su tutti gli apparati presenti in Internet, è la 4 (BGP4)
- Permette di usare notazione CIDR, abbassando il numero di rotte di un ISP da 2 milioni a circa 170mila
- Gli update sono affidabili perché fanno uso di connessioni TCP
- Oltre a questo, si inviano messaggi di *keep alive*

BGP – Altre proprietà

- Inizialmente, i router si scambiano tutta la tabella di routing
- Successivamente, vengono scambiati solo gli *update* incrementali
- La presenza di attributi permette di alterare le metriche di default
- E' in grado di scalare a dimensioni "importanti" (es. ISP)

BGP – Strutture dati

- Tabella dei vicini (neighbor) – per ogni processo BGP, i neighbor devono essere configurati esplicitamente, con cui viene instaurata una sessione

- Tabella BGP – contiene le informazioni reperite dai vari neighbor, compresi gli attributi. Sono quindi possibili path multipli

- IP Routing Table – contiene i migliori percorsi per le varie reti. La scelta avviene in base alla distanza amministrativa

BGP – Pacchetti di comunicazione

- Vengono sempre scambiati solo a sessione TCP avviata
- Sono chiamati *open, keepalive, update, notification*
- Il *keepalive* serve a non far scadere i time-out relativi alla sessione, anche in assenza di aggiornamenti
- L'*update* fornisce informazioni aggiornate su un singolo path

BGP – Adiacenza

- In BGP i termini *peer, speaker* e *neighbor* identificano orientativamente gli stessi apparati, ossia i router abilitati a scambiarsi informazioni di routing con i vicini

- Non essendo possibile comunicare con tutti i router esistenti, i path verso reti remote sono conosciute tramite scambio di informazioni

- I router di AS diversi sono spesso connessi fra loro direttamente, o con una route statica

BGP – Uso interno all'AS

- Ove sia usato IBGP, è sufficiente che i router che dialogano siano tra di loro raggiungibili

- In un AS di transito, si usa IBGP in quanto la redistribuzione completa delle rotte fatta da EBGP genererebbe troppo traffico e sfrutterebbe troppe risorse

- In un AS "normale" è solitamente necessario prevedere la non propagazione delle rotte interne, ed una rete "full meshed"

BGP – Uso del TCP

L'utilizzo del TCP come supporto per la gestione delle sessioni ha vantaggi e svantaggi:

- Affidabilità delle connessioni
- Tramite acknowledge e trasferimenti di pacchetti di dimensioni variabili, è possibile inviare anche quantità di dati consistenti
- Non è possibile usare broadcast e multicast

Agenda

- Cosa è il CINECA
- Protocolli di routing
- La struttura di rete CINECA
- Il routing nella struttura di rete CINECA

Linee guida del progetto

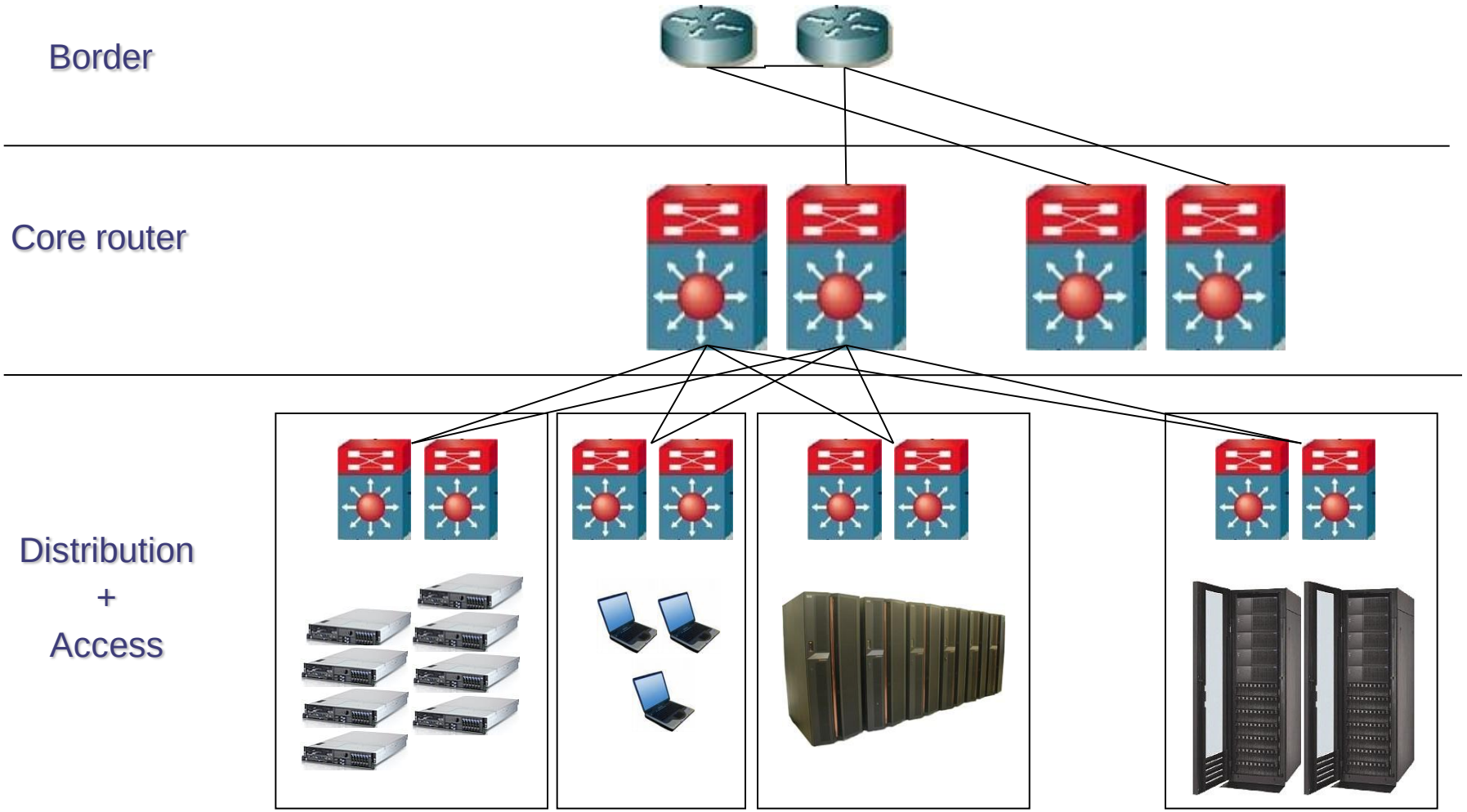
- Architettura multilivello
- Virtualizzazione domini di routing
- Scalabilità e prestazioni
- Affidabilità e gestibilità

Architettura Multilivello

- Internet Access
- Core (ad alte prestazioni)
- Distribution
- Access

I protocolli di routing si applicano ai primi 3 punti

Architettura multilivello



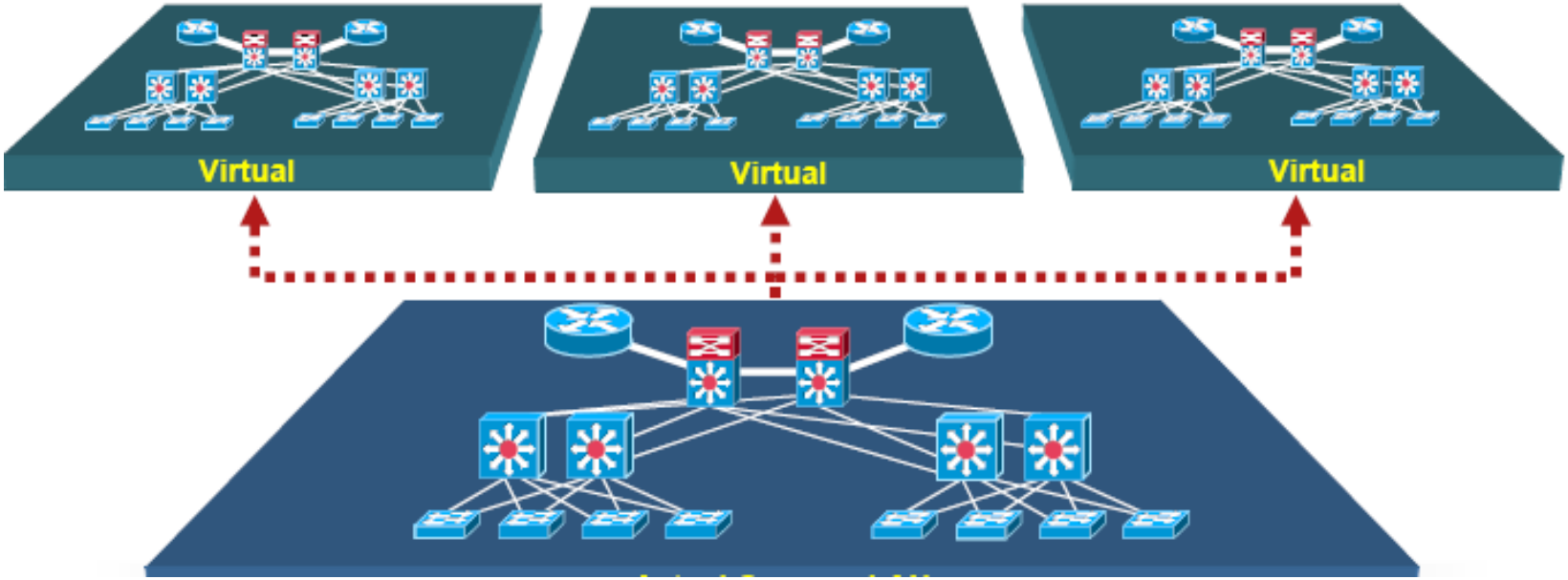
Virtualizzazione domini di routing

- Permette la separazione di domini di routing, pur essendo questi gestiti all'interno dello stesso apparato fisico
- Diminuzione del numero di apparati, consolidati in pochi apparati ad alte prestazioni

VRF

Virtualizzazione domini di routing

- Semplificazione dell'infrastruttura di rete



Prestazioni (Velocità)

- Collegamenti ad alta velocità verso Internet
- Backbone a 10 Gbps per la LAN
- Capacità elevata di routing

Cluster per il supercalcolo

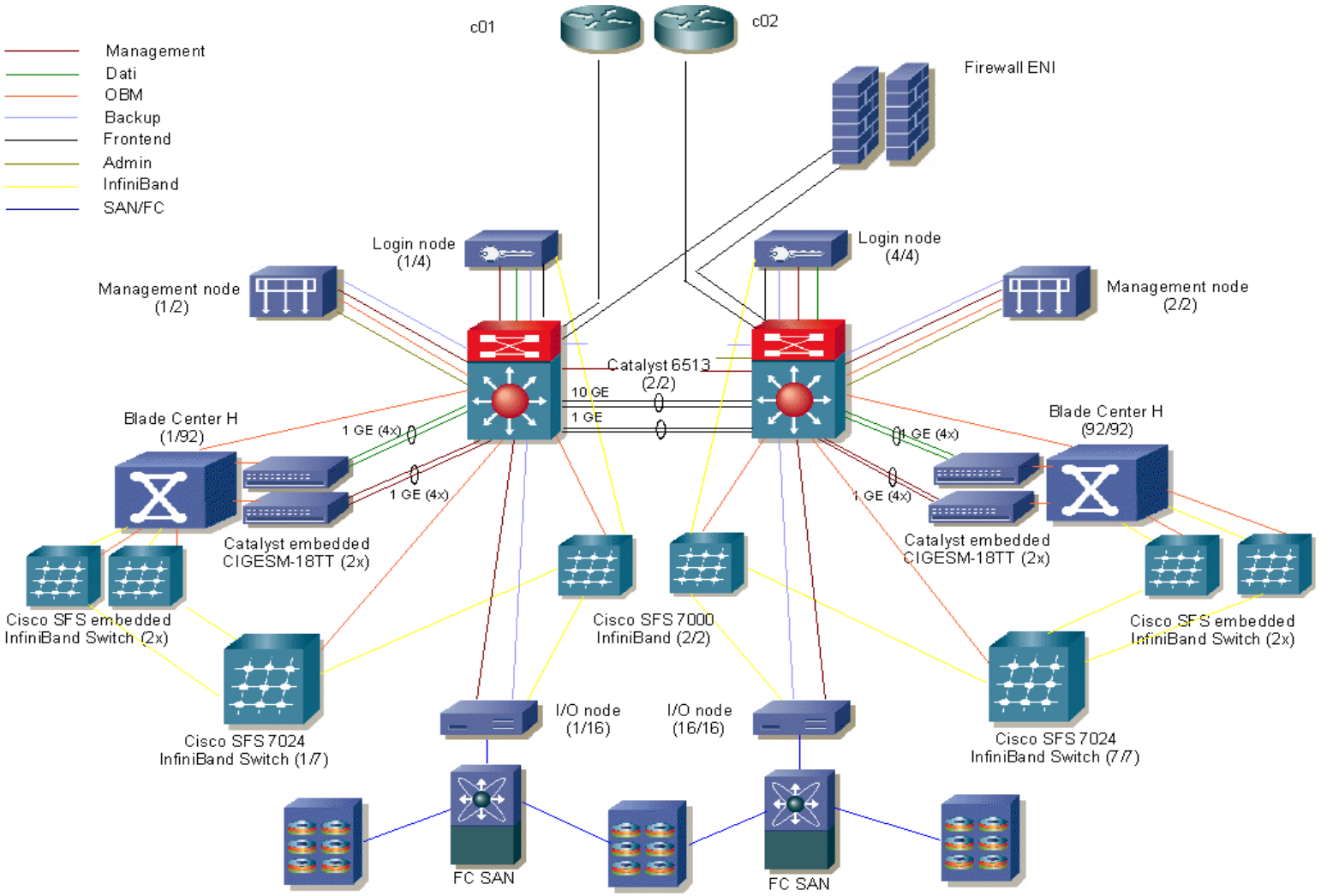
Definizione di una “interfaccia” standard verso la rete

Schema interno basato sui livelli core, distribution e access

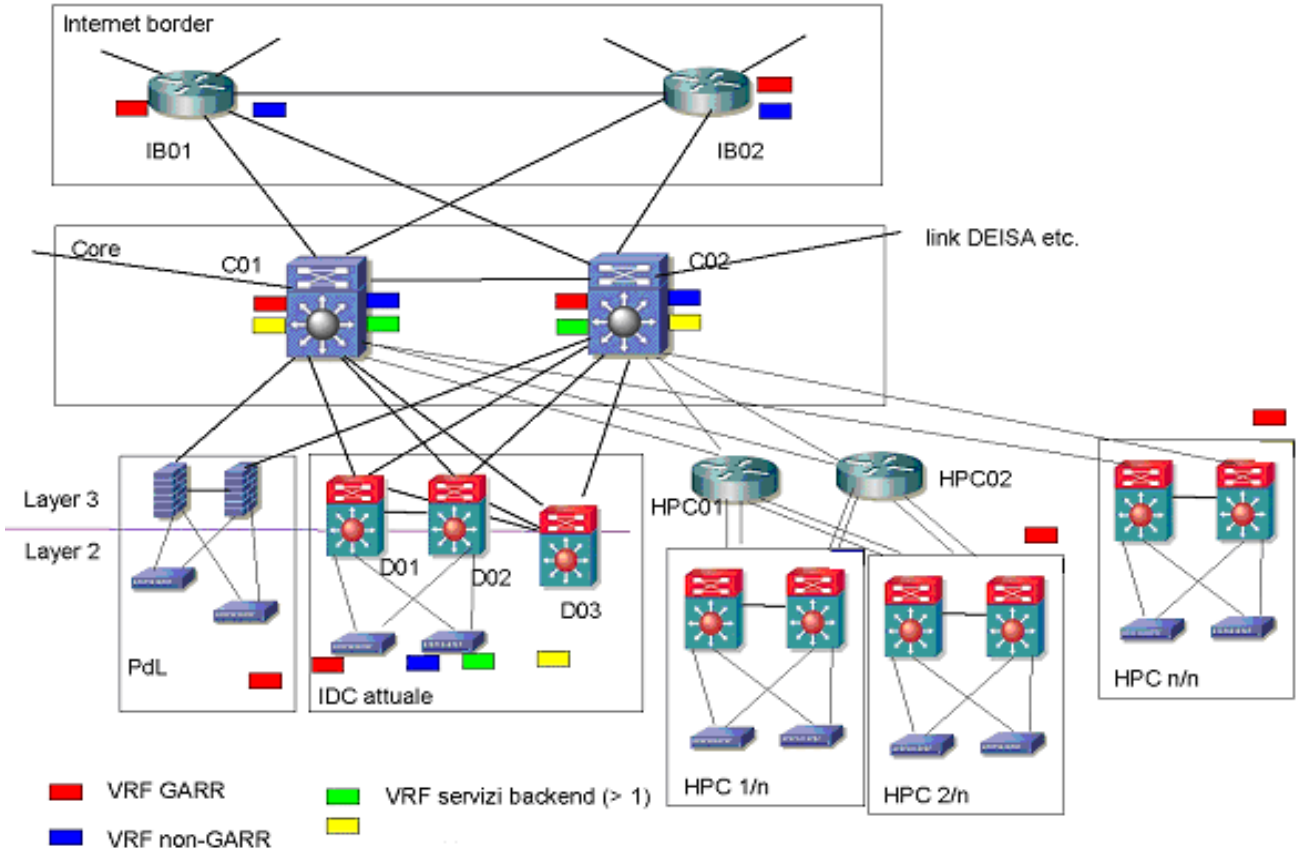
Connessioni ad alta velocità con i core router

Aggiornamento a questa struttura dei cluster esistenti

Cluster per il supercalcolo



Schema di principio



Evoluzione connessioni WAN

- Verso Internet
  - Link GARR a 1 Gbps, con backup a 100 Mbps
  - Link “Non GARR”: 2 link da 100 Mbps
- Link WAN dedicati
  - DEISA a 10 Gbps, Eni 2 link a 1 Gbps
  - Passaggio a MPLS, altrimenti linee punto-punto

Conseguenze

- Asimmetria connettività GARR
  - Le due connettività GARR non sono “equivalenti”
- Link WAN dedicati
  - Per ENI e DEISA connettività “diretta” sui core router
  - Apparato dedicato per le linee punto-punto
  - Connettività separata per le linee MPLS

Connessione per rete DEISA

- CINECA partecipa al progetto DEISA con:
  - risorse di calcolo attestate su due cluster
  - linea dati ad alta velocità
- Il link in questo caso è unico ed è attestato su uno dei due core router (su VRF dedicato)
- La connessione è a 10Gbps, tramite rete Lepida/GARR/GEANT

Connessione per rete ENI

- Le due connessioni sono attestate su apparati di proprietà ENI
- Ognuno di essi è collegato ad uno dei core router, garantendo quindi percorsi alternativi
- I core router provvedono poi a trasferire il traffico sugli apparati di produzione interessati

Altre reti dedicate

- Un border gateway opera come terminatore delle linee dedicate ancora utilizzate, con connessione diretta al core router
- Su un apparato della rete COMMERCIAL è attestato il router MPLS di Fastweb, su cui passa il traffico di queste connettività

Agenda

- Cosa è il CINECA
- Protocolli di routing
- Interscambio di informazioni di routing
- La struttura di rete CINECA
- Il routing nella struttura di rete CINECA

I protocolli utilizzati

- La nuova infrastruttura è stata studiata in funzione dei protocolli di routing utilizzati
- La comunicazione con l'esterno avviene tramite sessioni BGP
- Il routing interno alla struttura avviene tramite OSPF
- E' evidente la possibilità di individuare un'Area 0 che raccorda le altre

BGP

- Le connessioni verso l'esterno da parte di CINECA sono dirette verso peer accademici e commerciali

- Nel caso del VRF ACADEMIC, i peer sono i terminali della rete GARR

- Nel caso del VRF COMMERCIAL, sono invece i terminali dei fornitori di connettività commerciali (Fastweb, Tiscali)

BGP

- I comandi più comuni, per verificare la situazione del protocollo BGP:

```
# show ip bgp vpnv4 vrf <nomevrf> [summary|neighbors]

# show ip bgp all
```

Questo secondo comando fornisce il database topologico completo del protocollo BGP

BGP

- Le connessioni verso l'esterno da parte di CINECA sono dirette verso peer accademici e commerciali

- Nel caso del VRF ACADEMIC, i peer sono i terminali della rete GARR

- Nel caso del VRF COMMERCIAL, sono invece i terminali dei fornitori di connettività commerciali (Fastweb, Tiscali)

BGP

```
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
130.186.23.52   4  3275  434440  455079 31756756    0    0 28w3d           75
193.204.120.22  4  3228  299668  329345 31756756    0    0 29w4d           14
193.206.128.137 4   137  345701  301498 31756756    0    0 13w4d          784
```

- Questo è il “bgp summary” per il border gateway su vrf accademico

- I neighbor sono, nell’ordine, il secondo border gateway, un peering con Tiscali ancora attivo, ed il peering con il GARR

BGP

```
BGP router identifier 10.255.2.6, local AS number 3275
BGP table version is 31756756, main routing table version 31756756
981 network entries using 138321 bytes of memory
1056 path entries using 71808 bytes of memory
44858/22086 BGP path/bestpath attribute entries using 3409208 bytes of memory
16065 BGP AS-PATH entries using 385724 bytes of memory
216 BGP community entries using 7608 bytes of memory
```

- Qui sono presenti altre informazioni, quali l'AS di riferimento, e le dimensioni di alcune tabelle utilizzate dal protocollo

BGP

```
BGP neighbor is 130.186.23.52,  vrf ACADEMIC,  remote AS 3275, internal link
 Description: : I-BGP con I02
  BGP version 4, remote router ID 10.255.2.7
  BGP state = Established, up for 28w3d
  Last read 00:00:24, last write 00:00:16, hold time is 180, keepalive interval
is 60 seconds
```

Dettaglio di uno dei neighbors, quello interno alla rete CINECA, mentre il successivo è esterno alla rete CINECA

```
BGP neighbor is 193.206.128.137,  vrf ACADEMIC,  remote AS 137,  local AS 64770
no-prepend replace-as, external link
 Description: : E-BGP con POP-GARR-BO
  BGP version 4, remote router ID 193.206.128.252
  BGP state = Established, up for 13w4d
  Last read 00:00:10, last write 00:00:07, hold time is 90, keepalive interval
is 30 seconds
```

BGP

```
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
130.186.95.52   4  3275  341812 6143044 31756861    0    0 28w3d         5458
130.186.248.249 4  3228   48862   53966 31756861    0    0 4w5d            96
193.206.128.137 4   137  300428  301593 31756861    0    0 13w4d           88
213.200.87.14   4  3257 19469673  452636 31756861    0    0 20w3d       146171
```

Passando al lato commerciale, le differenze non sono molte:

```
BGP neighbor is 130.186.95.52,  vrf COMMERCIAL,  remote AS 3275, internal link
 Description: : I-BGP con I02
  BGP version 4, remote router ID 10.255.2.7
  BGP state = Established, up for 28w3d
  Last read 00:00:36, last write 00:00:01, hold time is 180, keepalive interval
is 60 seconds
```

L'altro border gateway, i02, è sempre neighbor, ma con indirizzi diversi

BGP – Tabella Topologica

```
BGP table version is 31757130, local router ID is 10.255.2.6
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 3275:20 (default for vrf COMMERCIAL)
*> 4.0.0.0/9        213.200.87.14          165             0 3257 3356 i
*> 4.0.0.0          213.200.87.14          165             0 3257 3356 i
*> 4.21.103.0/24    213.200.87.14            0             0 3257 3549 46133 i
*> 4.23.88.0/24     213.200.87.14          930             0 3257 7018 46164 i
*> 4.23.88.0/23     213.200.87.14            0             0 3257 7018 46164 i
*> 4.23.89.0/24     213.200.87.14          930             0 3257 7018 46164 i
*> 4.23.92.0/23     213.200.87.14          930             0 3257 7018 46164 i
*> 4.23.92.0/22     213.200.87.14            0             0 3257 7018 46164 i
*> 4.23.94.0/23     213.200.87.14          930             0 3257 7018 46164 i
*> 4.23.112.0/24    213.200.87.14           85             0 3257 174 21889 i
*> 4.23.113.0/24    213.200.87.14           85             0 3257 174 21889 i
```

OSPF

- L'architettura scelta si presta molto bene ad una definizione per aree OSPF

- In particolare, l'Area 0 contiene i Core router

- All'interfaccia fra Area 0 ed altre aree OSPF, oppure sul confine fra OSPF e BGP, troviamo rispettivamente i router di distribuzione ed i border gateway.

- Entrambi questi gruppi di dispositivi fanno parte dell'Area 0, come ABR

OSPF - LSDB

Il comando `# show ip ospf database` fornisce tutto il contenuto del LSDB, suddiviso per processi OSPF:

```
OSPF Router with ID (130.186.23.1) (Process ID 1)

                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
130.186.23.1    130.186.23.1    979         0x8000458D 0x00BDE8 18
130.186.23.2    130.186.23.2    1305        0x80004D48 0x00C99B 17
130.186.23.51   130.186.23.51   1592        0x800022F4 0x004DA2 4
[..]
Summary Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
10.100.2.0      130.186.91.72   1069        0x80000201 0x005A32
10.100.2.0      130.186.91.73   305         0x80000201 0x005437
```

OSPF - LSDB

Comprende anche le route arrivate dall'esterno:

```
 Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
0.0.0.0         130.186.91.4    1928        0x80004520 0x004D62 3
0.0.0.0         130.186.91.5    86          0x80004CE7 0x00A23D 3
10.201.0.0      130.186.91.72   1074        0x800008E3 0x00748C 0
10.201.0.0      130.186.91.73   309         0x800008E0 0x00748E 0
```

Il tipo di pacchetto LSA definisce la provenienza dell'informazione, e permette al database, e poi al processo di riduzione e scelta delle route, di comprendere anche questo criterio nel processo stesso.

OSPF - Route

Il comando `# sh ip route vrf <nomevrf> ospf` permette di conoscere in dettaglio le caratteristiche delle route:

```
O E2 192.168.132.0/24 [110/20] via 130.186.22.166, 2d21h, Vlan1211
     193.204.120.0/24 is variably subnetted, 4 subnets, 3 masks
O E2    193.204.120.32/30
           [110/20] via 130.186.22.65, 2d21h, GigabitEthernet4/48
O E2    193.204.120.1/32
           [110/20] via 130.186.22.65, 2d21h, GigabitEthernet4/48
O IA    193.204.120.192/28 [110/20] via 130.186.22.166, 2d21h, Vlan1211
     193.205.127.0/26 is subnetted, 1 subnets
```

E’ un esempio preso dal core router, vrf ACADEMIC

OSPF - Route

L'output mostra anche come leggere i tag:

```
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route
```

Quindi anche fra le stesse route legate ai processi OSPF, ci sono differenze di classificazione,a seconda della provenienza

OSPF – Configurazione

La denominazione dell'area è contenuta nella configurazione dei router fuori dall'area 0:

```
area 105 authentication message-digest
[..]
network 130.186.2.64 0.0.0.63 area 105
network 130.186.3.0 0.0.0.255 area 105
network 130.186.4.64 0.0.0.63 area 105
network 130.186.4.208 0.0.0.15 area 105
network 130.186.4.224 0.0.0.15 area 105
[..]
```

In corrispondenza, nel core router si trova:

```
O IA    130.186.3.0/24 [110/110] via 130.186.22.166, 2d21h, Vlan1211
```

IA significa proveniente da una diversa area, ma dallo stesso AS

Riferimenti

Ing. Vincenzo Vaccarino

CINECA
Dipartimento Sistemi e Tecnologie
Settore Operazioni

Tel. 051/6171411 (centralino)
Email: v.vaccarino@cineca.it