

BGP



Introduzione - BGP

void



Introduzione - BGP

ISP – Internet Service Providers

- Tier-1

Sono ISP di dimensioni molto grandi che effettuano peering tra di loro. Tutti gli hosts mondiali sono veicolati dai Tier-1.

es: AT&T – Telecom Italia Sparkle – Deutsche Telekom Global Carrier

- Tier-2

ISP di grandi dimensioni. Acquistano il transit service dai tier-1

es: GARR

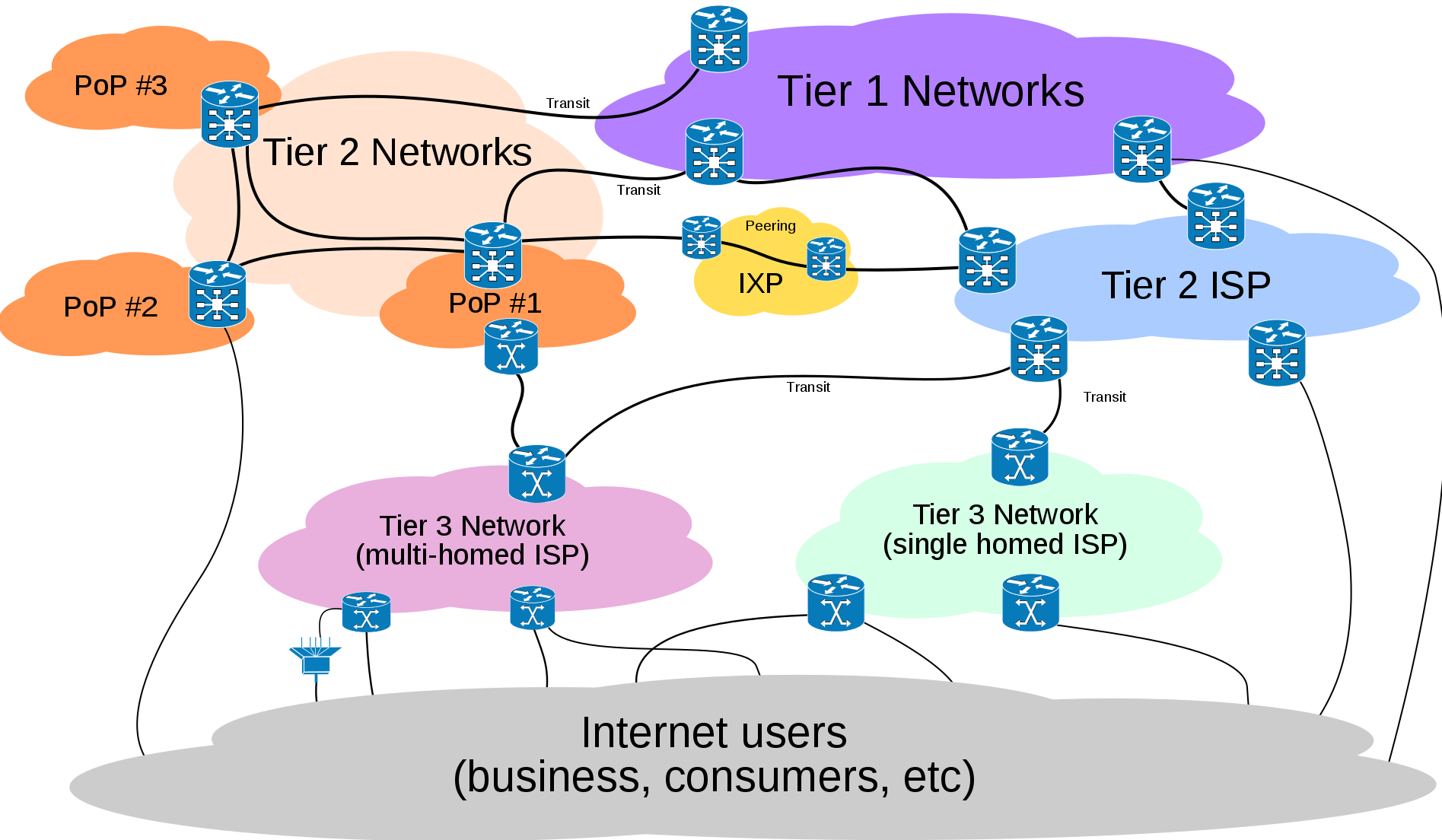
- Tier-3

Piccoli operatori che acquistano servizi di transito da T1 o T2. Se fanno peering con altre reti è di solito molto semplice.

es: Netoip – WispOne – Eolo - ...



Introduzione - BGP





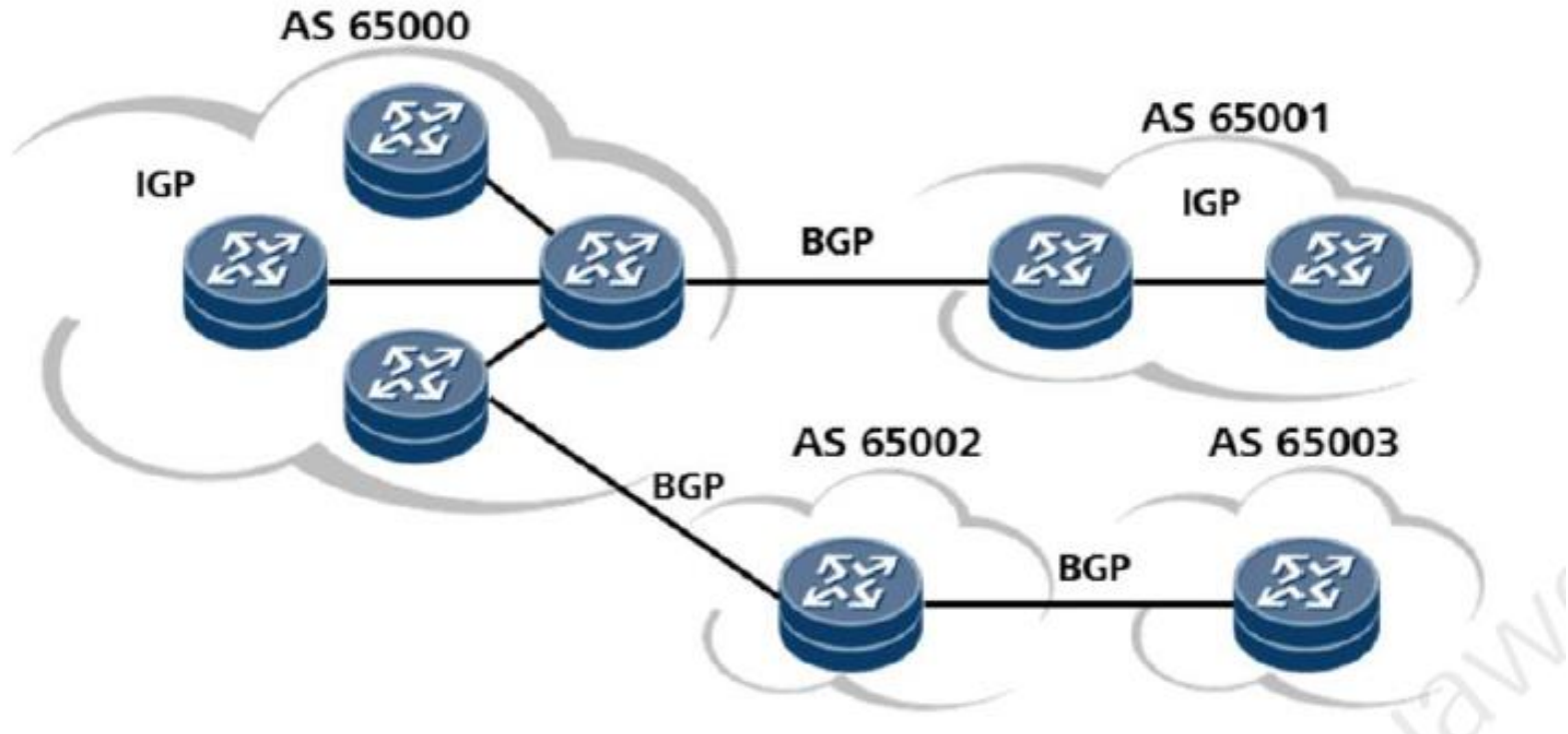
Introduzione - BGP

Protocolli di routing

- IGP – Internal Gateway Protocol
 - sono utilizzati per mantenere aggiornate le tabelle di routing all'interno di una organizzazione;
 - devono essere rapidi nel rispecchiare i cambiamenti topologici;
- EGP – Exterior Gateway Protocol
 - Sono utilizzati per scambiare informazioni tra diverse organizzazioni;
 - Garantiscono che diverse reti possano scambiare informazioni tra loro.



Introduzione - BGP





Introduzione - BGP

BGP

- focalizzato sulla raggiungibilità di blocchi di prefissi;
 - Annunciare ad un ISP i blocchi di indirizzo usati da un customer (in un AS);
- distance-path protocol:
 - non tiene conto della topologia dell'intera rete (come OSPF);
 - non si limita ad un hop-count (come RIP);
- Principio di funzionamento:
 - ogni router riceve informazioni di raggiungibilità dai suoi peer;
 - calcola la best path per includerla nella tabella di routing;
 - se le **politiche di routing** lo consentono, annuncia la rotta ai suoi peer;
 - il “path” è un elenco di AS che devono essere attraversati per raggiungere un blocco di indirizzi.

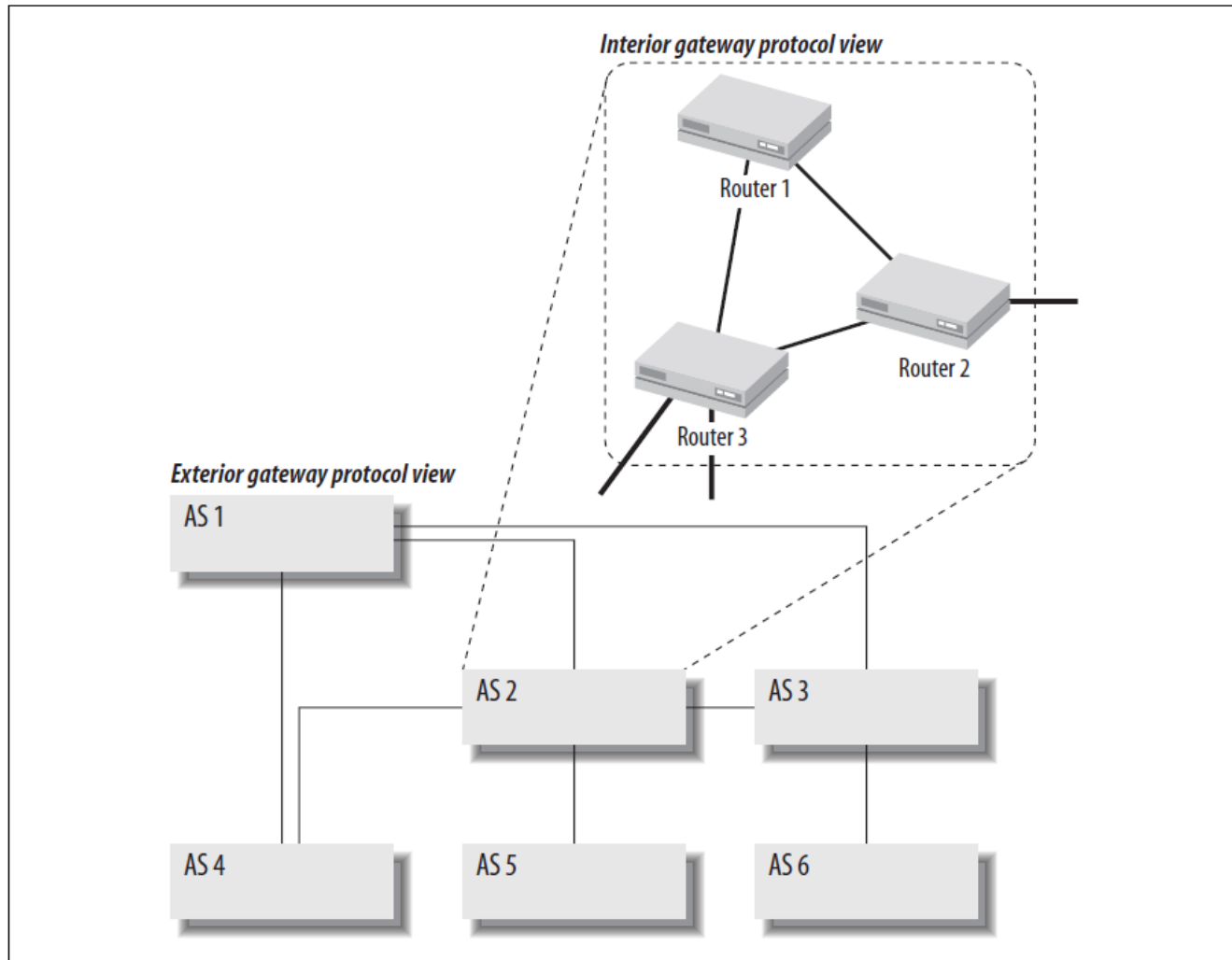


Introduzione - BGP

OSPF/EIGRP	BGP
Forms neighbor relationship before sending routing information	Same
Neighbors typically discovered using multicast packets on the connected subnets	Neighbor IP address is explicitly configured and may not be on common subnet
Does not use TCP	Uses a TCP connection between neighbors (port 179)
Advertises prefix/length	Advertises prefix/length, called <i>Network Layer Reachability Information (NLRI)</i>
Advertises metric information	Advertises a variety of path attributes (PA) that BGP uses instead of a metric to choose the best path
Emphasis on fast convergence to the truly most efficient route	Emphasis on scalability; might not always choose the most efficient route
Link-state (OSPF) or distance-vector (EIGRP) logic	Path-vector logic (similar to distance-vector)



Introduzione - BGP





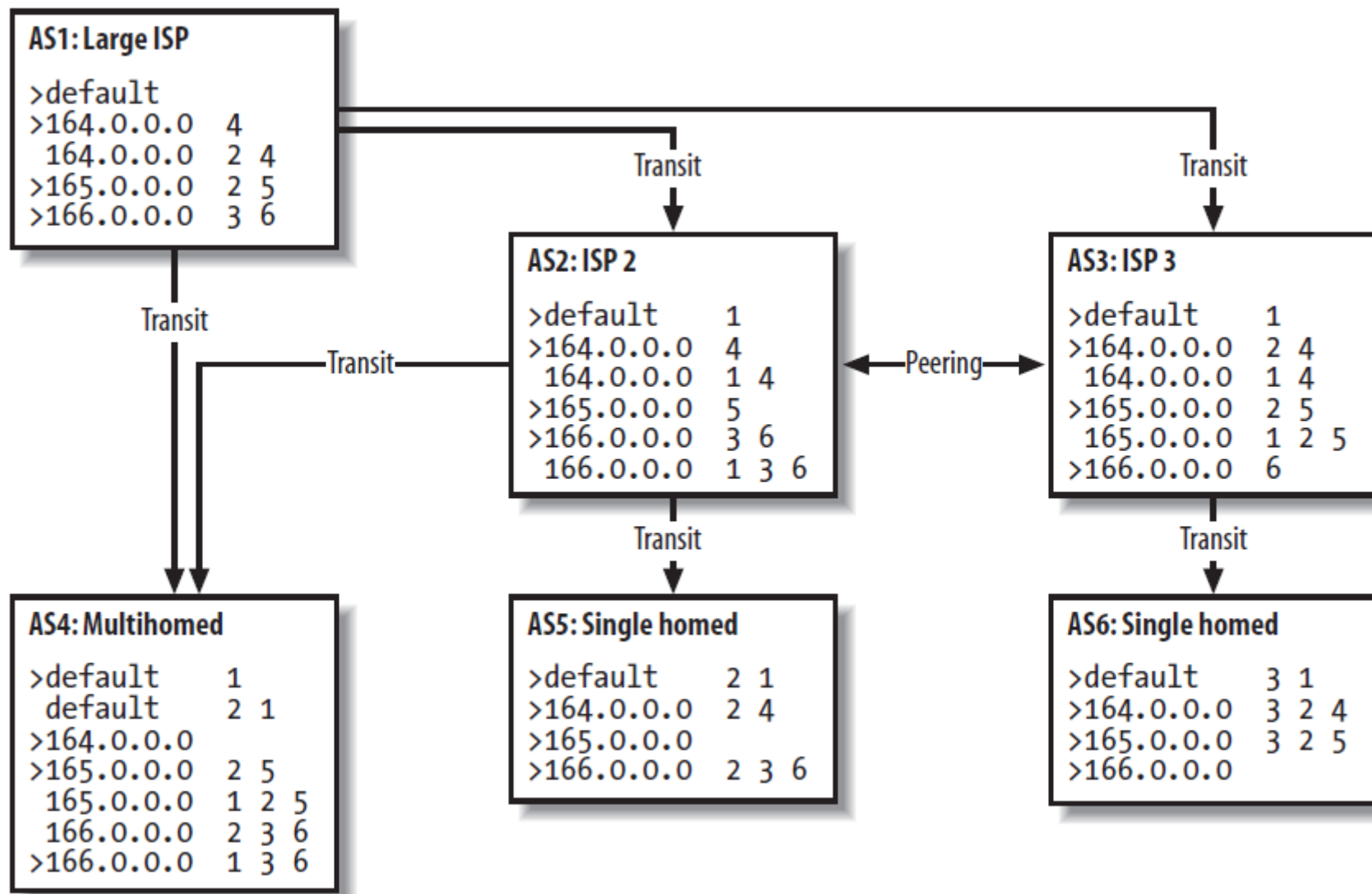
Introduzione - BGP

Policies

- Annunciare una rotta ad un router, significa invitarlo ad inviare del traffico;
- Le routing policies hanno precedenza sulle informazioni di raggiungibilità:
 - ISP ricevono le rotte dai loro upstream ISP e le annunciano ai customers;
 - in questo modo l'ISP è un transit ISP per il customer;
 - se un customer è collegato a due ISP deve evitare di annunciare le rotte di ISP1 ad ISP2 (diventerebbe transit);



Introduzione - BGP





Introduzione - BGP

Configurazioni tipiche:

- A prescindere da quanto stabilito dal protocollo usato per la connessione con ISP, il traffico passa per la stessa interfaccia;
- Può essere realizzato:
 - static routes;
 - BGP per lo scambio di una default route (ISP->Enterprise) e per annunciare il prefisso della enterprise (Enterprise -> ISP).
- Richiede attenzione nella configurazione del router di frontiera per evitare spreco di banda o routing-holes.



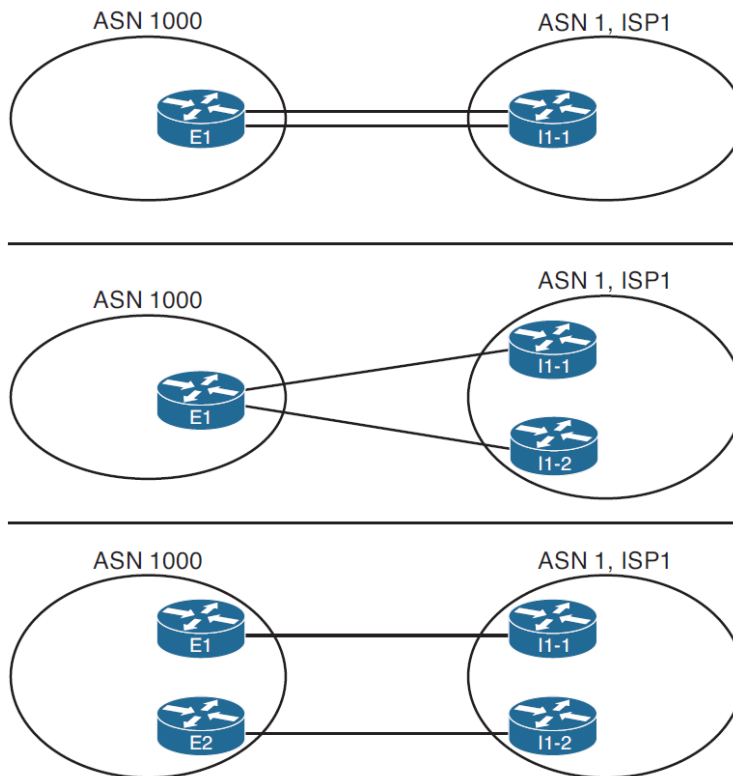


Introduzione - BGP

Configurazioni tipiche:



- Fisicamente può essere realizzato con 1,3 o 4 routers;



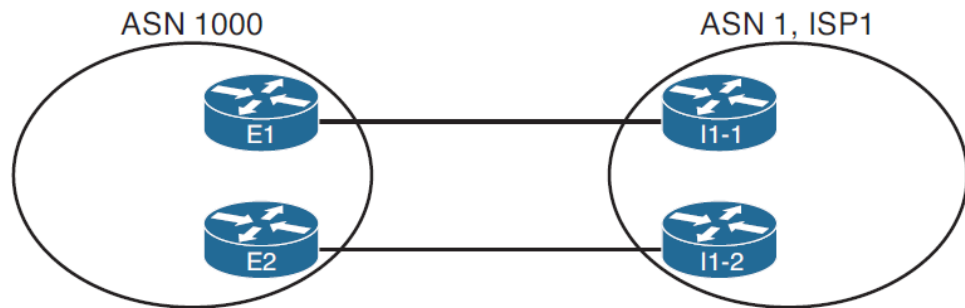


Introduzione - BGP

Configurazioni tipiche:

Dual Homed

- In questo caso è possibile influenzare il modo in cui il traffico esce dalla rete aziendale:
 - privilegiare uno dei link ed utilizzare l'altro come backup nel caso di link failure;
 - Annunciare in enterprise network due default routes, aventi metrica differente;
 - Si usano i comandi di redistribuzione delle rotte;
 - differenziare il link utilizzato sulla base della destinazione da raggiungere;

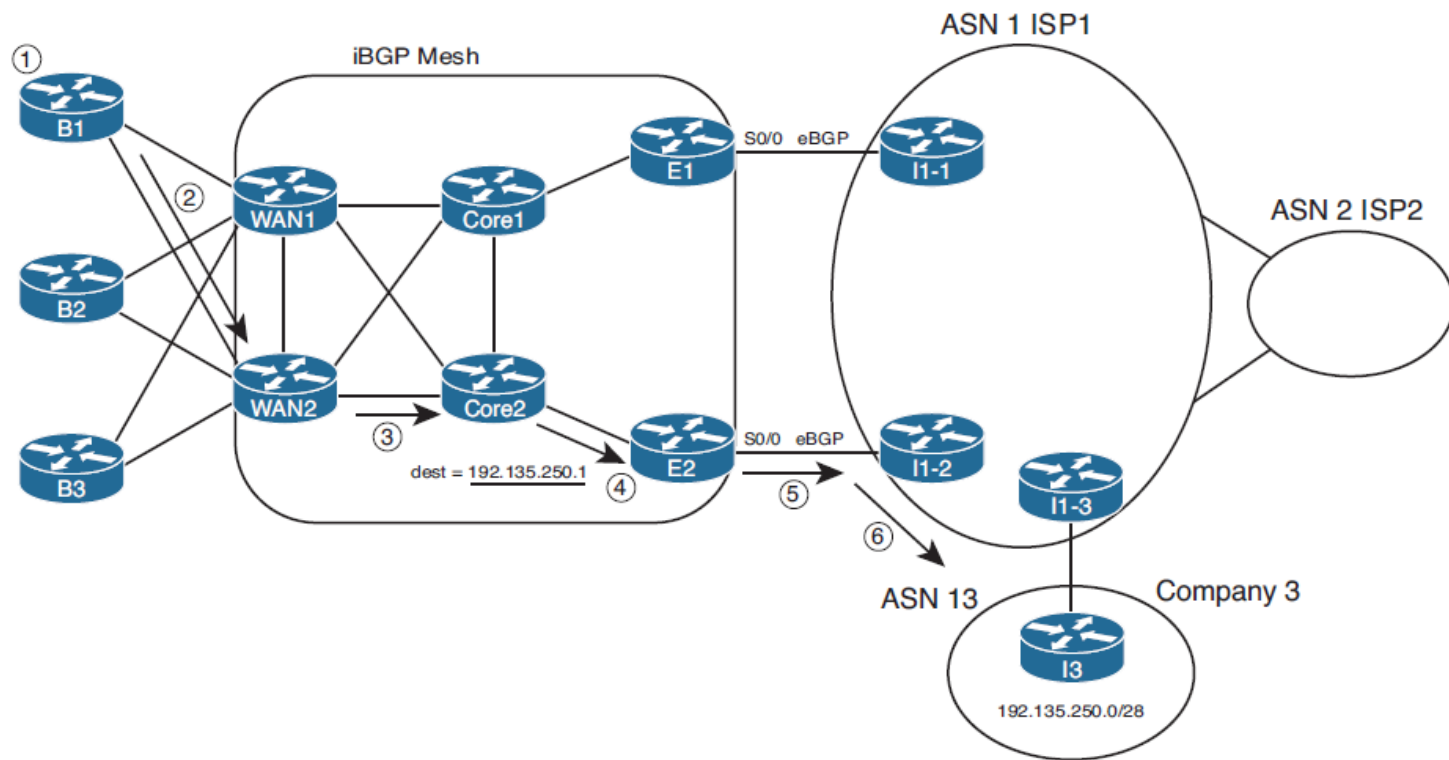




Introduzione - BGP

Configurazioni tipiche:

Dual Homed

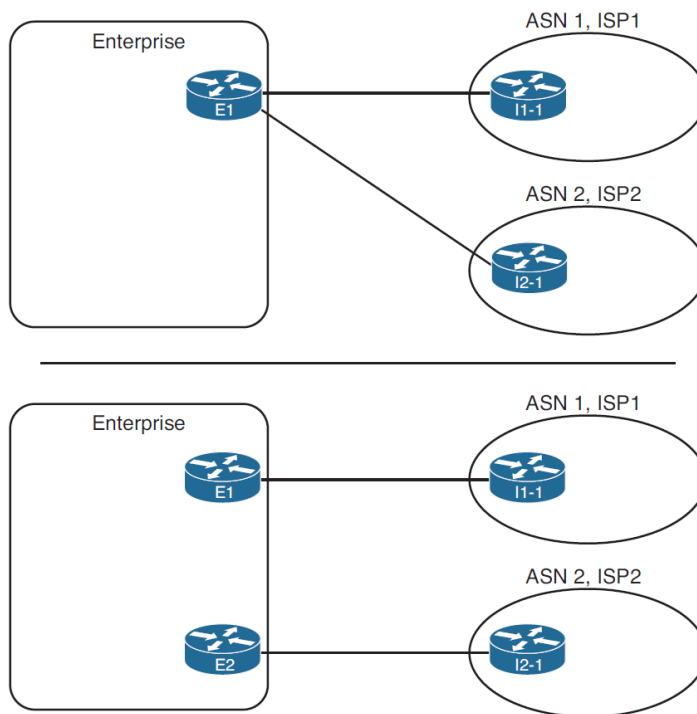
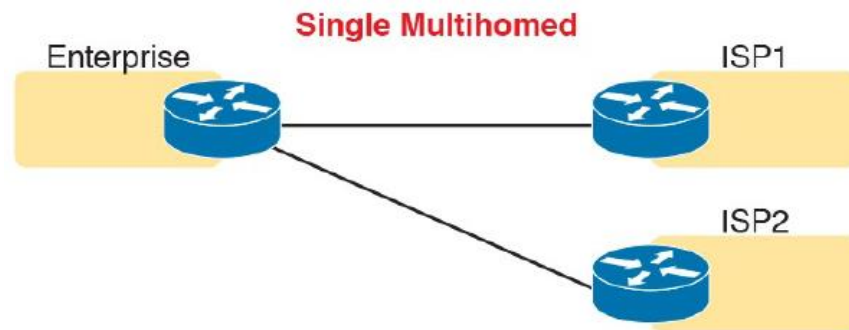




Introduzione - BGP

Configurazioni tipiche:

- Connessione a più di un ISP attraverso un singolo link.

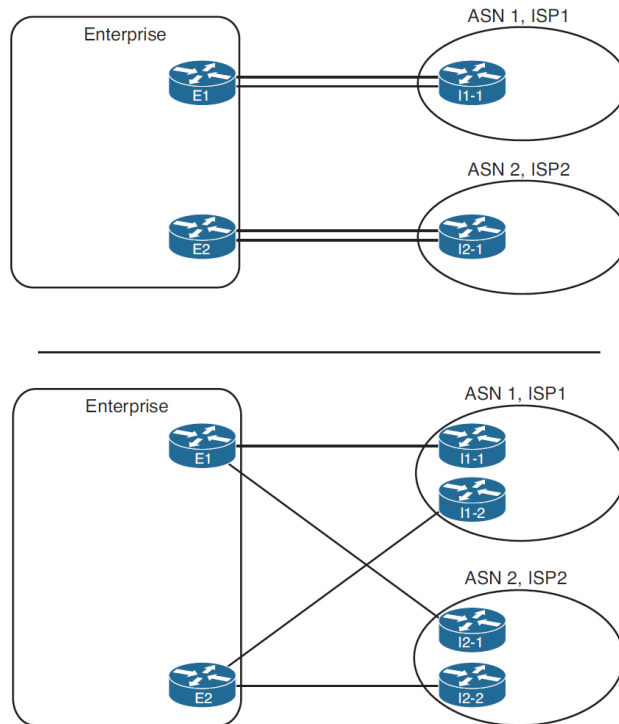
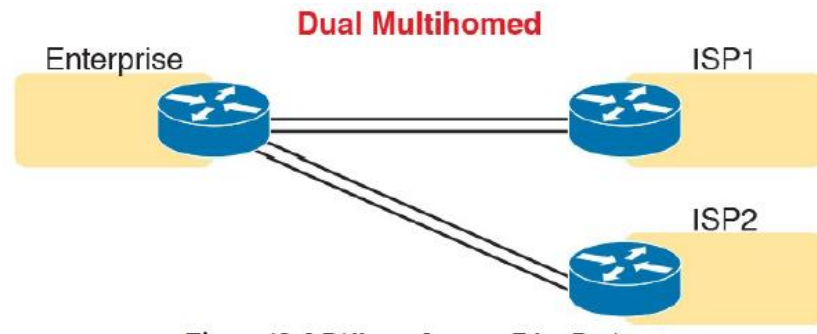




Introduzione - BGP

Configurazioni tipiche:

- Due o più ISP connessi ognuno con due o più link.





Introduzione - BGP

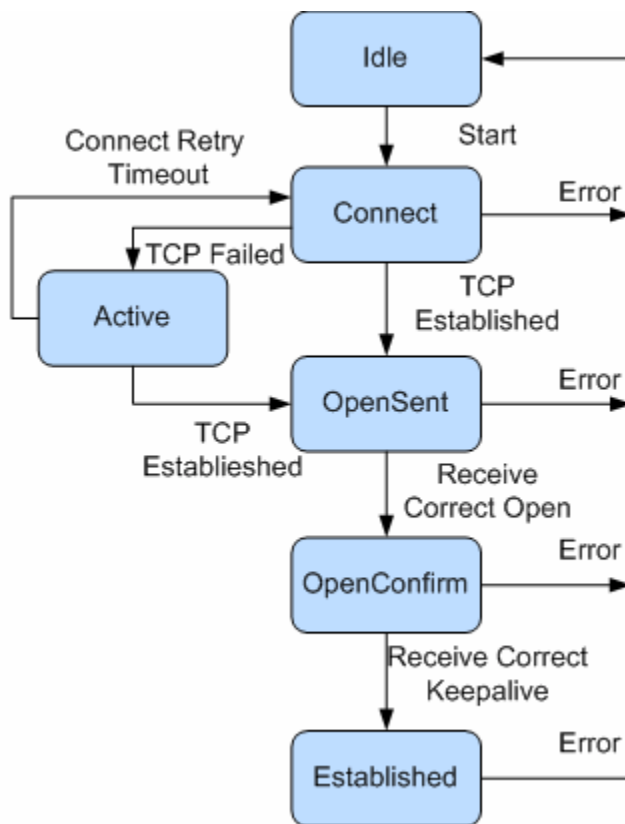
Relazione con ISP

- I router enterprise debbono conoscere quante più informazioni possibili per potere correttamente instradare il traffico;
- ISP offrono delle alternative per popolare le tabelle di routing:
 - Default route only;
 - Full Updates
 - Tabella BGP completa;
 - Partial Updates
 - Sono scambiati solo i prefissi che possono essere raggiunti meglio attraverso lo specific ISP.
 - Viene aggiunta una default route per gestire il traffico restante.



BGP - Peers

Finite State Machine





BGP

Messaggi BGP

Message	Purpose
Open	Used to establish a neighbor relationship and exchange basic parameters, including ASN and authentication values.
Keepalive	Sent on a periodic basis to maintain the neighbor relationship. The lack of receipt of a Keepalive message within the negotiated Hold timer causes BGP to bring down the neighbor connection.
Update	Used to exchange RAs and the associated prefix/length (NLRI) that use those attributes.
Notification	Used to signal a BGP error; typically results in a reset to the neighbor relationship.



Caratteristiche principali

- Può essere usato come
 - EBGp: si scambiano informazioni di raggiungibilità tra router che appartengono ad AS differenti;
 - IBGP: si scambiano informazioni di raggiungibilità tra router che sono nello stesso AS.
- Non usa una METRICA ma dei “path attributes” che sono annunciate con ogni prefisso;
 - I path-attributes consentono di prendere delle decisioni molto più accurate sulla rotta migliore verso un prefisso;
 - Sono utilizzati nella “best path selection”, ovvero la scelta della rotta migliore tra due che sono in competizione.



BGP

Router con BGP attivo:

- non forma relazioni di vicinato con i router adiacenti;
- I router vicini non vengono scoperti in modo dinamico;
- non viene abilitato sulle “interfacce” come nel caso di OSPF;
- Richiede:
 - configurazione manuale del peer;
 - configurazione manuale dei prefissi da annunciare;

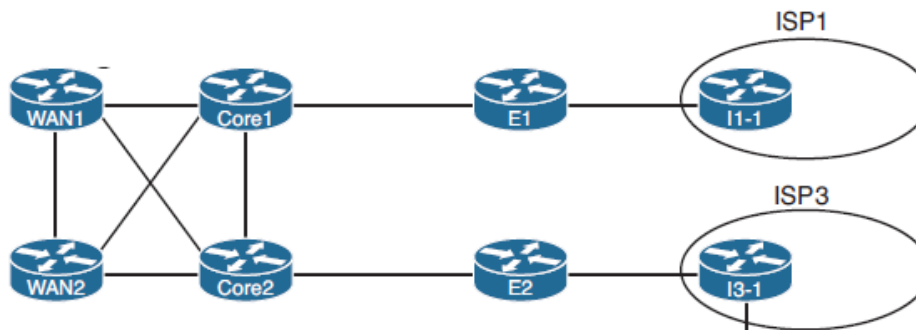
Caratteristiche salienti:

- utilizza TCP per lo scambio dati tra i router (porta179);
- viene popolata la BGP table del router;



BGP

- Le caratteristiche avanzate sono utili quando un customer ha una connessione verso due o più ISP;
- Gli ISP decidono il tipo di relazione che vogliono instaurare con il customer:
 - invio di una default route;
 - invio di partial updates;
 - invio di full update;
 - combinazione delle precedenti (partial + default);
- In queste topologie è importante che i router di frontiera siano configurati con iBGP;

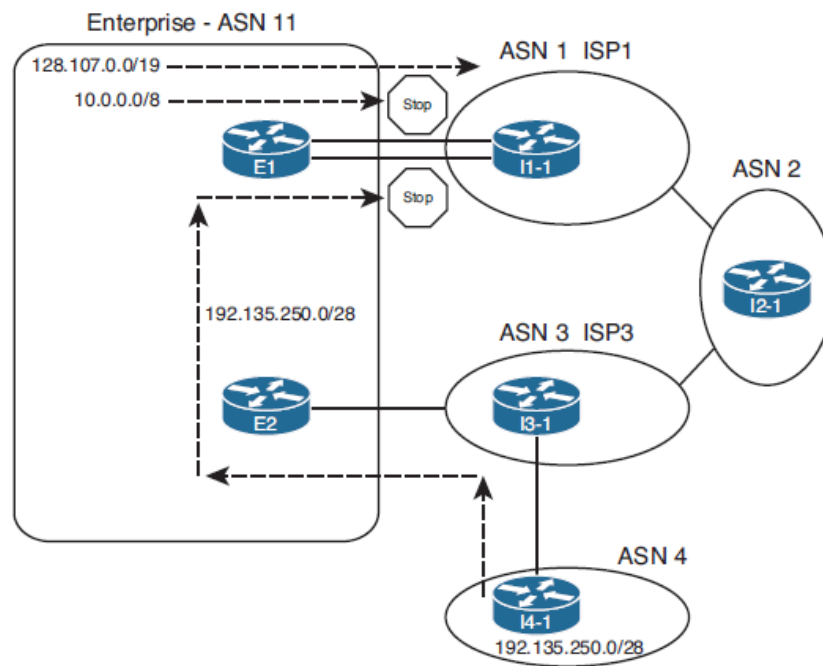




BGP – Route Filtering

Perchè dovrei filtrare?

- Controllo dei routing-updates per evitare di divenire un transit-AS per un ISP;
- Limitare le rotte esportate ad un ISP a quelle effettivamente necessarie;





BGP – Route Filtering

- I messaggi BGP update possono essere filtrati su ogni router;
 - I filtri consentono di esaminare gli update di ogni router;
 - Vengono esaminati il prefix ed il prefix length;
 - Possono essere differenti per la direzione “in” ed “out”;
- Si possono realizzare con:
 - ACL;
 - ip-prefix lists;
 - AS_Path Filter;
 - Community Filter;
 - ExtCommunity Filter;
 - Route-Policy.
- Supportano regular-expression!



BGP – Route Filtering

- Se viene applicato/modificato un filtro è necessario effettuare
 - route-refresh (se il peer lo supporta);
 - reset bgp



BGP – Path Attributes (PA)

BGP Best Path

- il processo attraverso il quale BGP esamina percorsi differenti verso lo stesso spazio di indirizzi e resolve le contese, scegliendo una rotta come migliore per quell singolo prefisso.

BGP supporta una serie piuttosto nutrita di PA:

- alcuni sono usati solo per la definizione di un BGP best-path;
- altri non hanno nulla a che fare nella scelta del BGP best-path;
- altri influenzano la scelta del best path



BGP – Path Attributes (PA)

- **Autonomous System Path (AS_Path)**

Elenca tutti gli AS nel percorso end-to-end. Viene utilizzato per la loop-prevention e per la determinazione della path-length.

- **Next-Hop IP address di un prefisso**

BGP può annunciare uno dei tanti next-hop definiti per una rete;

- **Local Preference**

Valore numerico impostato e comunicato in uno stesso AS per influenzare la scelta della rotta migliore tra tutti i router dello stesso AS.

- **Origin**

Indica come una rotta è stata inserita in BGP.



BGP – Path Attributes (PA)

- **Multi-Exit-Discriminator (MED)**

Impostato dai router di un AS. Influenza le decisioni dei router che ricevono il MED. Valori minori hanno una maggiore preferibilità.

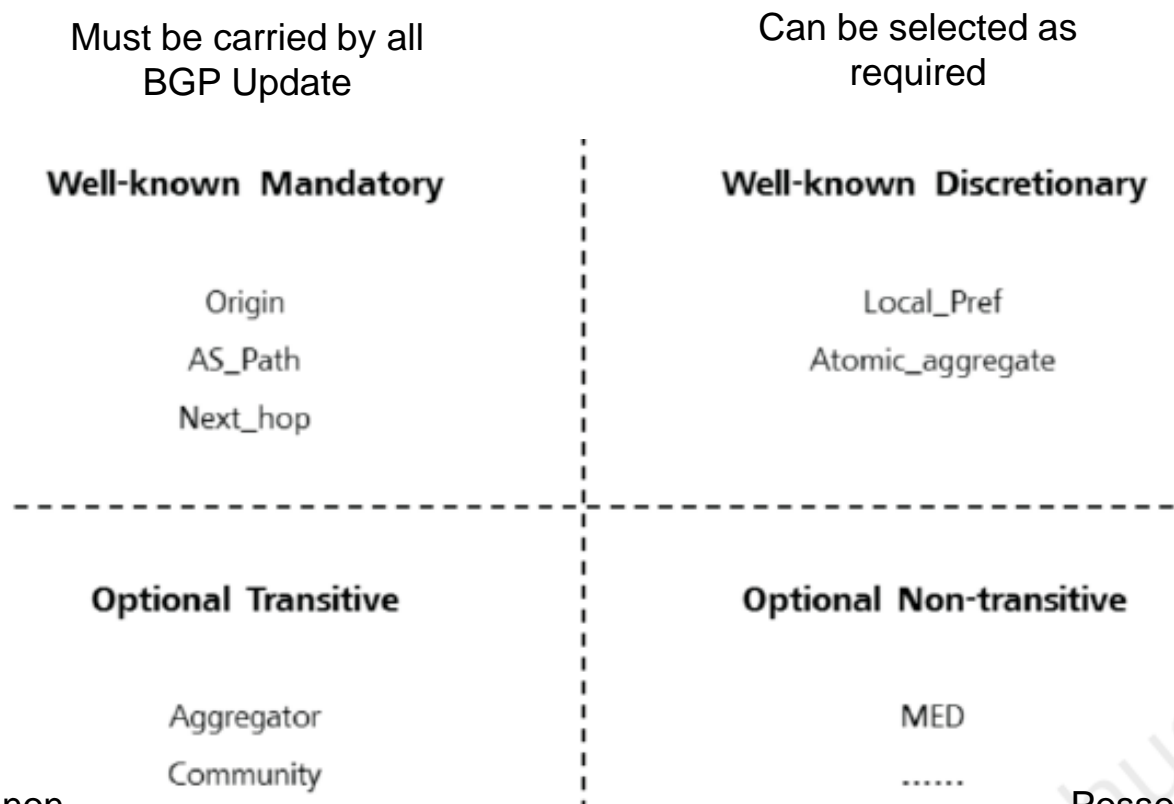
- **Communities**

Valore 32 bit espresso tipicamente nella forma 701:120 in cui 701 è il numero di AS e 120 è una label interna all'AS. Possono essere usate per triggerare un comportamento sulle rotte.

Attribute	Type
Origin	Well-known mandatory
AS_Path	Well-known mandatory
Next_Hop	Well-known mandatory
Local_Pref	Well-known discretionary
Community	Optional transitive
MED	Optional non-transitive
Originator_ID	Optional non-transitive
Cluster_List	Optional non-transitive



BGP – Path Attributes (PA)



Potrebbero essere non interpretati da tutti i router. Possono essere trasportati nei messaggi BGP e annunciati ai vicini.

Possono essere ignorati da un router. Non sono annunciati ai vicini.



BGP – Route Selection Process

- **Prefers the route with the largest PrefVal value**

The PrefVal attribute is a Huawei proprietary attribute and is valid only on the device where it is configured. Cisco usa un Sistema analogo: weight.

- **Prefers the route with the highest Local_Pref**

If a route does not have the Local_Pref attribute, the Local_Pref attribute of the route uses the default value 100.

- Prefers the **manually summarized route, automatically summarized route, route imported using the network command, route imported using the import-route command, and route learned from peers**. These routes are in descending order of priority.
- **Prefers the route with the shortest AS_Path.**



BGP – Route Selection Process

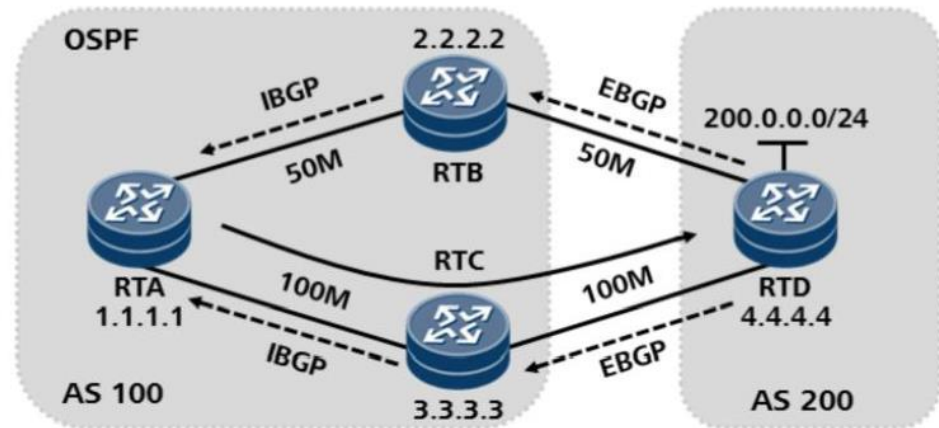
- **Prefers the route with the lowest origin type.**
IGP is lower than EGP, and EGP is lower than Incomplete.
- **Prefers the route with the lowest MED** if routes are received from the same AS.
- Prefers EBGP routes, IBGP routes, LocalCross routes, and RemoteCross routes, which are listed in descending order of priority.
- **Prefers the route with the lowest IGP metric to the BGP next hop.**



BGP – Influencing Outbound Routes

PrefVal:

- Customer ha una connessione doppia verso un ISP;
- Si vuole privilegiare un percorso rispetto ad un altro per motivi di costo o prestazioni;



- RTA riceve le informazioni sulla rete 200.0.0.0/24 da RTC ed RTB;
 - si effettua un matching della rotta verso 200.0.0.0/24 negli update da RTC;
 - si impone la PrefVal al valore 100



BGP – Influencing Outbound Routes

PrefVal

- con questo sistema abbiamo modificato il modo in cui un singolo router raggiunge un prefisso;



BGP – Influencing Outbound Routes

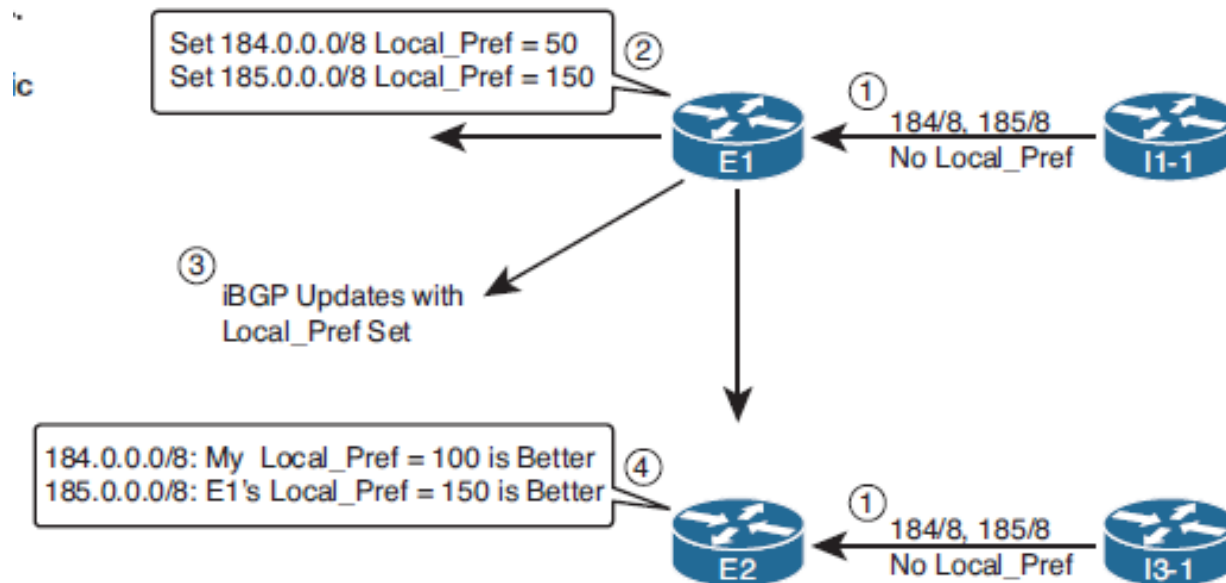
Local Preference:

- è un path attribute che comunica ai router di un AS un valore che può essere impostato per ogni rotta e può essere annunciato a tutti i router iBGP nell'AS. Tutti i router nell'AS concordano su quale sia il migliore router di uscita da AS per raggiungere il prefisso annunciato.
- Local Preference non viene annunciato a peer eBGP;
- Viene utilizzato nel caso di una rotta che preveda due punti di uscita, per rompere l'incertezza.
- Un valore più alto “vince”.



BGP – Influencing Outbound Routes

Local Preference:





BGP – Influencing Inbound Routes

Risulta più difficile controllare le rotte in ingresso:

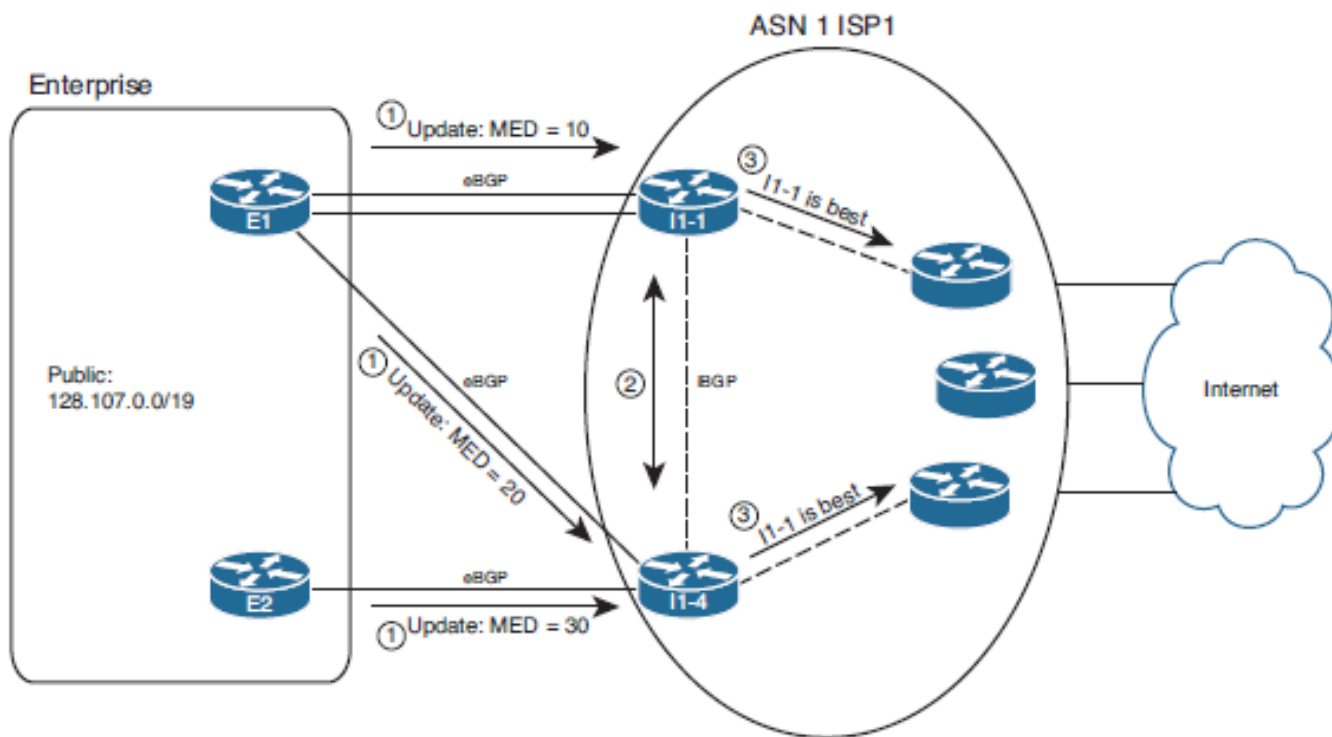
- sono gestite da router che non sono sotto il controllo del customer (ISP);
- ci possono essere vincoli tecnici che limitano la possibilità di controllo;

MED – Multi-Exit-Discriminator

- Il customer può annunciare all'ISP un valore di MED che dice all'ISP quale sia il percorso migliore per l'ingresso nell'AS del customer.
- In questo modo ISP può scegliere il migliore tra i diversi percorsi che collegano ISP e customer.
- Logica di MED: smallest is best!



BGP – Influencing Inbound Routes





Propagazione delle rotte BGP

- Controllo dei filtri (policies) in ingresso;
- Inserimento della rotta nella tabella BGP;
- Confronto della rotta con analoghe che abbiano lo stesso prefisso di destinazione (NLRI) ed esegue il BGP route-selection algorithm
- Se la rotta inserita è la migliore, viene inserita nella tabella di routing al posto di quella sub ottima;
- La rotta sub-ottima viene revocata da tutti i peer;
- Viene propagata la nuova rotta a tutti i peer esterni (sempre che i filtri lo consentano);
- Propaga la rotta a tutti i peer interni.



Scelta delle rotte in BGP

- Per garantire robustezza alla rete le destinazioni sono spesso raggiungibili attraverso molti vicini BGP.
- Questo rende necessario un meccanismo di selezione delle rotte migliori tra quelle disponibili per la stessa destinazione.
- La scelta della rotta migliore avviene confrontando gli attribute che sono propagate tra un peer e l'altro.



Route-selection Algorithm

Ci sono numerose implementazioni personalizzate dell'algoritmo. Generalmente sono preferite le rotte con:

- Local Preference maggiore;
- AS-Path più corta;
- Valore di MED inferiore;
- Ci sono poi dei passi ulteriori in caso di ulteriore contesa:
 - next-hop address con costo o metrica IGP minori;
 - rotte annunciate da BGP peer con indirizzo IP minore;



Route Generation

Prima di popolare una tabella di routing bgp è necessario impostare il «family-type:

- `ipv4-family <type>`

Before performing BGP configurations in an IPv4 address family, you need to run the `ipv4-family` command in the BGP view to enable the IPv4 address family, and then enter the address family view. By default, BGP uses the IPv4 unicast address family.

Possibili famiglie:

- ▣ `unicast`
- ▣ `multicast`
- ▣ `vpnvp4`
- ▣ `vpn-instance`
- ▣ `mdt`



BGP - Route Generation

Come è possibile popolare le tabelle BGP?

- Network Command

The network command configures BGP to statically add routes in the IP routing table to the BGP routing table and advertise these routes to peers.

The network command is **used to import exactly-matching routes**. This means that only the routes in the local IP routing table that **exactly match** the specified destination address and prefix length can be added to the BGP routing table. If mask is not specified, routes are exactly matched against the **natural network mask**.



BGP – Route Generation

Come è possibile popolare le tabelle BGP?

- Import Command

```
import-route protocol [ process-id ] [ med med | route-policy route-policy-name ]
```

Protocollo:

- ▣ direct, isis, ospf, rip, static, unr.
- ▣ MED Multi-Exit-Discriminator ;



BGP – Route Generation

L'importazione di rotte in BGP avviene attraverso l'applicazione di una o più route policies.

- route-policy

A routing policy is used to filter routes and set route attributes for the routes that match the routing policy. A routing policy consists of multiple nodes. One node can be configured with multiple if-match and apply clauses.

The if-match clauses define matching rules for this node, and the apply clauses define behaviors for the routes that match the rules.

- node

Specifies the index of the node in the routing policy. When the routing policy is used to filter routes, the node with the smaller value of node is matched first.



BGP – Route Generation

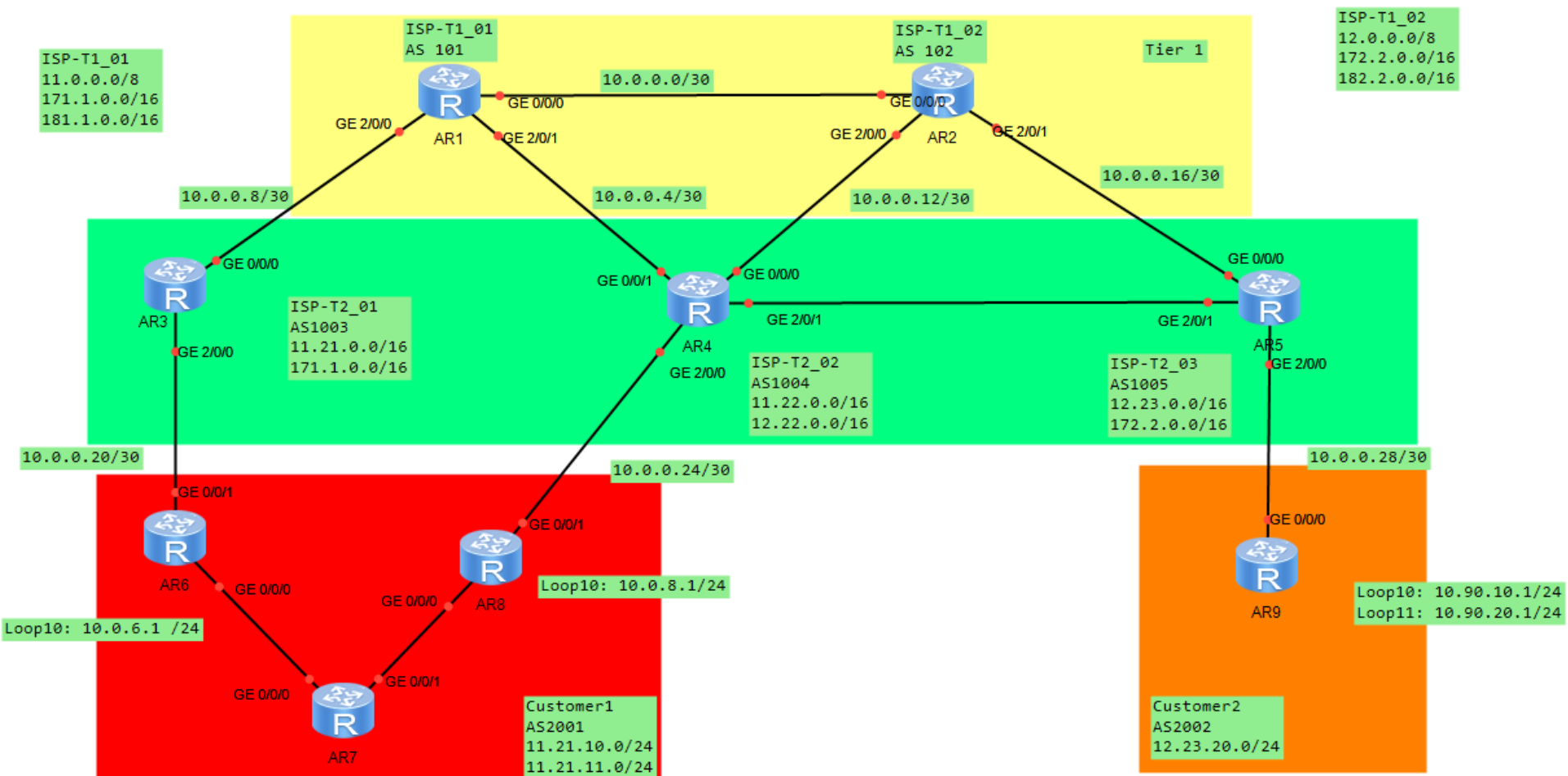
L'importazione di rotte in BGP avviene attraverso l'applicazione di una o più route policies.

- `ip ip-prefix`

This command is used to configure an IPv4 prefix list. An IP prefix can be used as a filter by various protocols or used together with a routing policy.

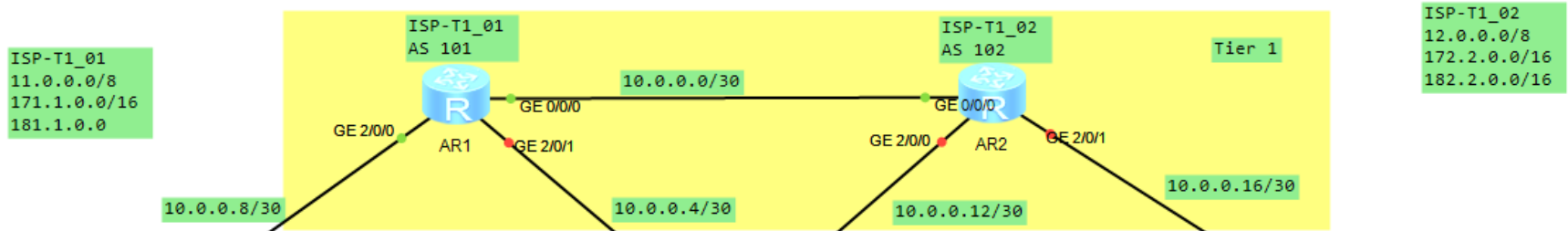


BGP – Esempio





BGP – Esempio

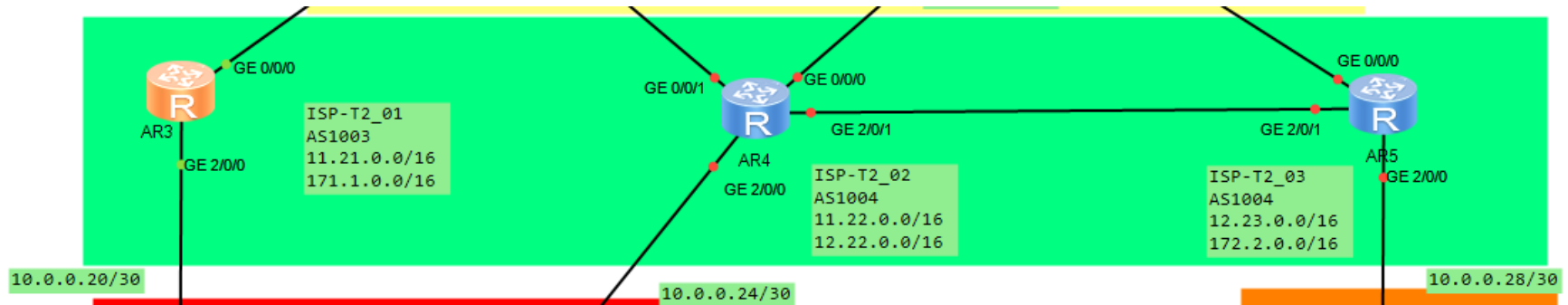


Struttura tier-1

- Due ISP in peering BGP;
- Gli IP assegnati agli ISP sono:
 - Redistribuiti agli ISP-tier2;
 - Assegnati ai customers.



BGP – Esempio

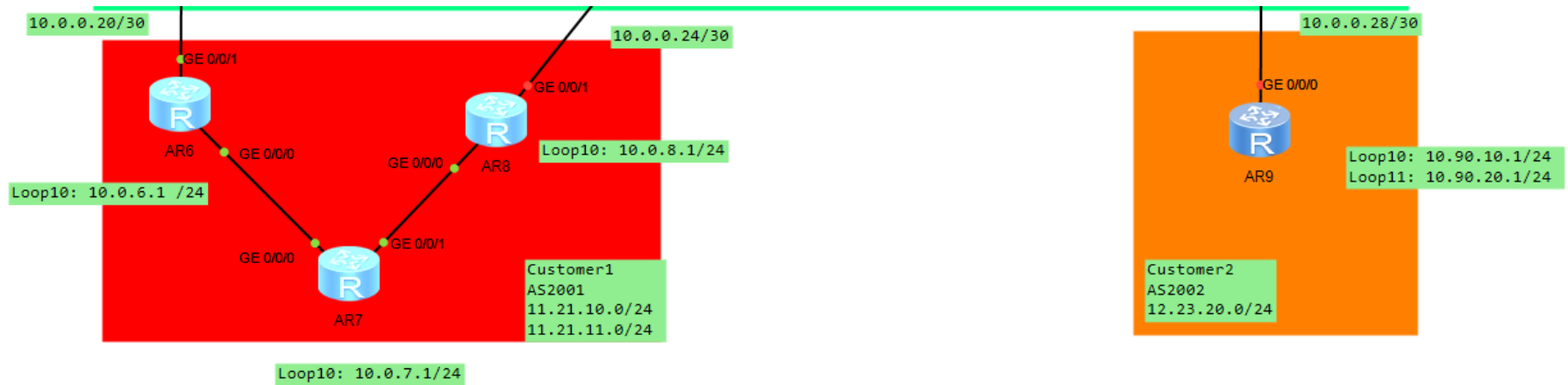


Struttura tier-2

- 3 ISP di cui due effettuano peering diretto;
 - ISP-T2_02 e ISP-T2_03 scambiano rotte direttamente tra loro;
- Gli indirizzi IP assegnati agli ISP vengono dati ai customers;



BGP – Esempio



Customers

- clienti di grandi dimensioni che hanno un numero AS assegnato;
- utilizzano gli IP assegnati dal ISP;
- utilizzano una numerazione interna per la gestione delle reti;



BGP – Analisi della configurazione

Tier-1 ISP

- ISP-T1_01 – AS 101 e ISP-T1_02

```
#ISP-T1_01
interface LoopBack0
 ip address 172.16.31.1 255.255.255.255
#
interface LoopBack1
 ip address 181.1.0.1 255.255.0.0
#
bgp 101
 peer 172.16.31.2 as-number 102
 peer 172.16.31.2 ebgp-max-hop 255
 peer 172.16.31.2 connect-interface LoopBack0
 peer 172.16.31.3 as-number 1003
 peer 172.16.31.3 ebgp-max-hop 255
 peer 172.16.31.3 connect-interface LoopBack0
#
ipv4-family unicast
 undo synchronization
 network 181.1.0.0
 peer 172.16.31.2 enable
 peer 172.16.31.3 enable
#
ip route-static 172.16.31.2 255.255.255.255 10.0.0.2
ip route-static 172.16.31.3 255.255.255.255 10.0.0.10
#
```

placeholder



BGP – Analisi della configurazione

Tier-2 ISP

- ISP-T2_01

```
#ISP-T2_01
#
bgp 1003
 peer 172.16.31.1 as-number 101
 peer 172.16.31.1 ebgp-max-hop 255
 peer 172.16.31.1 connect-interface LoopBack0
 peer 172.16.31.4 as-number 2001
 peer 172.16.31.4 ebgp-max-hop 255
 peer 172.16.31.4 connect-interface LoopBack0
#
ipv4-family unicast
 undo synchronization
 network 171.1.0.0
 peer 172.16.31.1 enable
 peer 172.16.31.4 enable
#
ip route-static 172.16.31.1 255.255.255.255 10.0.0.9
ip route-static 172.16.31.4 255.255.255.255 10.0.0.22
#
```



BGP – Analisi della configurazione

Customer AR6

- CUS_R6 – single-homed (step 1)

```
#CUS_R6
#
bgp 2001
 peer 172.16.31.3 as-number 1003
 peer 172.16.31.3 ebgp-max-hop 255
 peer 172.16.31.3 connect-interface LoopBack0
#
ipv4-family unicast
 undo synchronization
 import-route ospf 1 route-policy isp_t2_01
 peer 172.16.31.3 enable
#
#
ospf 1 router-id 6.6.6.6
 area 0.0.0.0
 network 10.0.6.1 0.0.0.0
 network 11.21.10.6 0.0.0.0
```

```
#
route-policy isp_t2_01 permit node 10
 if-match ip-prefix isp_t2_01
#
ip ip-prefix isp_t2_01 index 10 permit 11.21.10.0 24
ip ip-prefix isp_t2_01 index 20 permit 11.21.11.0 24
#
ip route-static 172.16.31.3 255.255.255.255 10.0.0.21
#
```

<ISP-T2_01>dis bgp routing-table

BGP Local router ID is 10.0.0.10

Status codes: * - valid, > - best, d - damped,

h - history, i - internal, s - suppressed, S - Stale

Origin : i - IGP, e - EGP, ? - incomplete

Total Number of Routes: 5

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	11.21.10.0/24	172.16.31.4	0		0	2001?
*>	11.21.11.0/24	172.16.31.4	3		0	2001?
*>	171.1.0.0	0.0.0.0	0		0	i
*>	181.1.0.0	172.16.31.1	0		0	101i
*>	182.2.0.0	172.16.31.1			0	101 102i

<ISP-T2_01>



BGP – Analisi della configurazione

Customer AR6

- CUS_R6 – dual-homed (step 2 – configurazione IBGP)

```
#
bgp 2001
peer 172.16.31.3 as-number 1003
peer 172.16.31.3 ebgp-max-hop 255
peer 172.16.31.3 connect-interface LoopBack0
peer 172.16.31.5 as-number 2001
peer 172.16.31.5 connect-interface LoopBack0
#
ipv4-family unicast
undo synchronization
import-route ospf 1 route-policy isp_t2_01
peer 172.16.31.3 enable
peer 172.16.31.5 enable
#
ospf 1 router-id 6.6.6.6
area 0.0.0.0
network 10.0.6.1 0.0.0.0
network 11.21.10.6 0.0.0.0
network 172.16.31.4 0.0.0.0
#
```

```
#
bgp 2001
peer 172.16.31.4 as-number 2001
peer 172.16.31.4 connect-interface LoopBack0
#
ipv4-family unicast
undo synchronization
peer 172.16.31.4 enable
#
ospf 1 router-id 8.8.8.8
area 0.0.0.0
network 10.0.8.1 0.0.0.0
network 11.21.11.8 0.0.0.0
network 172.16.31.5 0.0.0.0
#
```



BGP – Analisi della configurazione

Customer AR6

- CUS_R6 – dual-homed (step 3 – configurazione dual-home)
- A seguito dell'espansione aziendale si decide per una connessione verso ISP-T2_02 in modo da:
 - ridondare la connettività verso internet;
 - potere accedere allo spazio di indirizzi di ISP-T2_02;
- Operativamente:
 - Aggiunte due loopback su CUS_R7;
 - Aggiunte le nuove rotte ad IGP;
 - Aggiunte le nuove rotte a BGP.



BGP – Analisi della configurazione

Analizziamo la situazione su ISP-T1_01

```
[ISP-T1_01]dis bgp routing-table
```

```
BGP Local router ID is 1.1.1.1
```

```
Status codes: * - valid, > - best, d - damped,
```

```
h - history, i - internal, s - suppressed, S - Stale
```

```
Origin : i - IGP, e - EGP, ? - incomplete
```

```
Total Number of Routes: 18
```

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*> 11.21.10.0/24	172.16.31.3			0	1003 2001?
*	172.16.31.6			0	1004 2001?
*	172.16.31.2			0	102 1004 2001?
*> 11.21.11.0/24	172.16.31.3			0	1003 2001?
*	172.16.31.6			0	1004 2001?
*	172.16.31.2			0	102 1004 2001?
*> 11.22.0.0/16	172.16.31.6	0		0	1004i
*	172.16.31.2			0	102 1004i
*> 12.22.1.0/24	172.16.31.3			0	1003 2001?
*	172.16.31.6			0	1004 2001?
*	172.16.31.2			0	102 1004 2001?
*> 12.22.2.0/24	172.16.31.3			0	1003 2001?
*	172.16.31.6			0	1004 2001?
*	172.16.31.2			0	102 1004 2001?
*> 171.1.0.0	172.16.31.3	0		0	1003i
*> 181.1.0.0	0.0.0.0	0		0	i
*> 182.2.0.0	172.16.31.2	0		0	102i
*	172.16.31.6			0	1004 102i

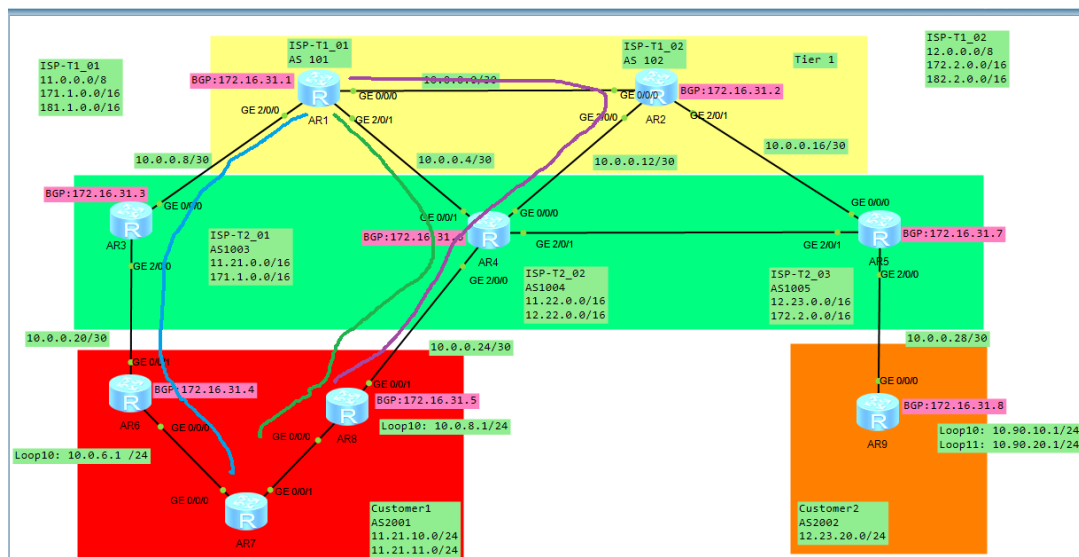


BGP – Analisi della configurazione

Analizziamo la situazione su ISP-T1_01

- Troppi path verso la stessa rotta. Necessario effettuare una operazione di selezione e filtraggio:

```
*> 12.22.1.0/24    172.16.31.3      0    1003 2001?
*                172.16.31.6      0    1004 2001?
*                172.16.31.2      0    102 1004 2001?
*> 12.22.2.0/24    172.16.31.3      0    1003 2001?
*                172.16.31.6      0    1004 2001?
*                172.16.31.2      0    102 1004 2
```





Thank you
www.huawei.com