

Il Protocollo Bgp & Internet

Davide Giuffrida
d.giuffrida@libero.it

25/02/2004

Davide Giuffrida

1

Contenuti

- I - Il Routing - Struttura di Internet
- II - Il Protocollo BGP
- III - Esempi e pratica
- IV - Strumenti di Diagnostica

25/02/2004

Davide Giuffrida

2

Terminologia

Documento RFC:
specifiche d'implementazione applicativi tcp/ip

Routing/Router:
Periferica di rete che si presenta almeno su 2 reti < >
capace di inoltrare dati tra le 2 o più reti

IPv4:
Protocollo TCP/ip versione 4.0 (next ipv6)

BGP:
Border Gateway Protocol = Routers "Intelligenti"

25/02/2004

Davide Giuffrida

3

Parte I La Struttura di Internet

Come Funziona la Rete?

25/02/2004

Davide Giuffrida

4

Il Routing

Il Routing dei pacchetti

- In un sistema **packets switching** quale il TCP/IP, il **routing** rappresenta il processo di scelta del percorso su cui inoltrare i pacchetti ed il **router** è un computer che effettua tale instradamento.
- Quando un programma applicativo su un host tenta di instaurare una comunicazione con un host remoto, sia l'host locale che i routers partecipano all'instradamento dei datagrammi IP fino alla loro destinazione.

Il Routing

Routing Diretto ed Indiretto

- Si può parlare di due tipi di routing:
- **Diretto**: se l'host locale ed il remoto appartengono alla stessa rete fisica (es. una singola *Ethernet*); in questo caso non sarà necessario l'impiego di routers.
- **Indiretto**: se l'utente destinazione è connesso ad una rete fisica diversa ed è necessario instradare il datagramma sorgente attraverso un router. Per sapere se una destinazione appartiene alla propria rete, l'utente sorgente estrae dall'indirizzo IP di destinazione la parte relativa alla rete, la cosiddetta **netid**, e la confronta con la propria: se differisce, evidentemente la destinazione del datagramma è esterna.

Il Routing

Routing in Internet

- I routers in Internet formano una struttura interconnessa in continuo contatto fra loro nella quale il datagramma scorre finché non raggiunge quel particolare router che gli permette di giungere direttamente a destinazione. L'algoritmo che svolge queste funzioni (IP routing algorithm) utilizza su ogni macchina una **tabella di routing** (*IP routing table*), che contiene informazioni circa alcune possibili destinazioni e su come raggiungerle. Se comprendesse tutte le destinazioni possibili diventerebbe troppo ingombrante e sarebbe impossibile tenerla aggiornata. Allora di solito ci si limita a mantenere le informazioni degli utenti sulla stessa rete o quelle più frequentemente usate lasciando un indirizzo di default per tutti gli altri.

Indirizzamento IP

<http://www.iana.org/assignments/ipv4-address-space>

- INTERNET PROTOCOL V4 ADDRESS SPACE (last updated 2002-12-10)
- The allocation of Internet Protocol version 4 (IPv4) address space to various registries is listed here.
- Originally, all the IPv4 address spaces was managed directly by the IANA.
- Later parts of the address space were allocated to various other registries to manage for particular purposes or regional areas of the world. RFC 1466 documents most of these allocations.

Suddivisione della Rete

Block	Date	Registry - Purpose	Notes or Reference
-----	-----	-----	-----
.....			
192/8	May 93	Various Registries	
213/8	Mar 99	RIPE NCC	(whois.ripe.net)
.....			
217/8	Jun 00	RIPE NCC	(whois.ripe.net)
.....			
Reference [RFC1466] [RFC1918] [RFC3330]			

25/02/2004

Davide Giuffrida

9

Struttura di Internet: organismi ufficiali

- Coordinamento Mondiale: IANA
www.iana.org



Dedicated to preserving the central coordinating functions of the global Internet for the public good.

25/02/2004

Davide Giuffrida

10

Struttura di Internet: organismi ufficiali

Coordinamento Continentale - Le 4 RIR:

There are currently four Regional Internet Registries:

- RIPE NCC**
Réseaux IP Européens Network Coordination Centre
<http://www.ripe.net>
- ARIN**
American Registry for Internet Numbers
<http://www.arin.net>
- APNIC**
Asia Pacific Network Information Centre
<http://www.apnic.net>
- LACNIC**
Latin American and Caribbean IP address Regional Registry
<http://lacnic.net>



25/02/2004

Davide Giuffrida

11

Struttura di Internet: organismi ufficiali

★ Local Internet Registries (Autonomous Systems (AS)):

GARR - LIR

GARR-NIC | GARR-NOG | GARR-CERT | GARR

<ul style="list-style-type: none"> ▪ Descrizione del servizio ▪ Documentazione ▪ Richiesta Nuovi indirizzi IP ▪ Modifica delle informazioni ▪ Whois Client ▪ Richiesta Informazioni 	<p>Il GARR-LIR e' un servizio riservato alle organizzazioni GARR ed alle organizzazioni autorizzate ad accedere alla rete GARR.</p> <p>Compiti del GARR-LIR sono :</p> <ul style="list-style-type: none"> • la assegnazione di nuovi indirizzi IP agli utenti GARR • il mantenimento delle relative informazioni presso il Regional Internet Registry europeo (RIPE) e presso il Regional Internet Registry americano (ARIN) <p>L'uso della rete GARR e' soggetto a regole e deve essere richiesto secondo la procedura per il collegamento.</p> <p>Il GARR-LIR e' composto da:</p> <p>Gabriella Paolini (Responsabile del servizio) Marco Gallo Bruno Melideo Vincenzo Puglia</p> <p>E-mail: lir@garr.it</p>
---	--

25/02/2004

Davide Giuffrida

12

Struttura di Internet: organismi ufficiali

- ★ Provider (connette ad Internet le proprie “utenze”):

Provider IP (ISP) :

**Università degli studi
di Brescia**



Piazza Mercato ▲
◀ San Faustino

25/02/2004

Davide Giuffrida

13

La realizzazione...

- ★ Peering tra providers:

```
aut-num: AS137
as-name: ASGARR
descr: GARR Italian academic and research network
import: from AS20965 action pref=100;
import: from AS1299 action pref=100;
import: from AS1299 accept ANY action pref=100;
import: from AS3549 accept ANY action pref=100;
export: to AS20965 announce AS137
export: to AS1299 announce AS137
export: to AS3549 announce AS137
admin-c: EV182-RIPE
tech-c: GL965-RIPE
tech-c: GN450-RIPE
mnt-by: GARR-LIR
changed: noc@garr.it 20000302
changed: noc@garr.it 20000830
changed: paolini@garr.it 20020617
changed: paolini@garr.it 20020821
source: RIPE
```

25/02/2004

Davide Giuffrida

14

La realizzazione...

- ★ Neutral Access Point (NAP):



Milan Internet eXchange

25/02/2004

Davide Giuffrida

15

La realizzazione...

- ★ Internet Routing Registry (IRR):



25/02/2004

Davide Giuffrida

16

Parte II il protocollo BGP

Cos'è il BGP ?

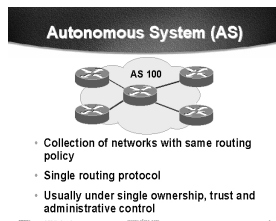
Border Gateway Protocol

Border Gateway Protocol

- Protocollo di routing utilizzato per scambiare informazioni di routing tra le reti
- protocollo "esterno" di routing
- RFC1163 & RFC1267 (bgp v 2/3)
- RFC1771 (bgp v 4)
- lavori in corso....
- draft-ietf-idr-bgp4-13.txt

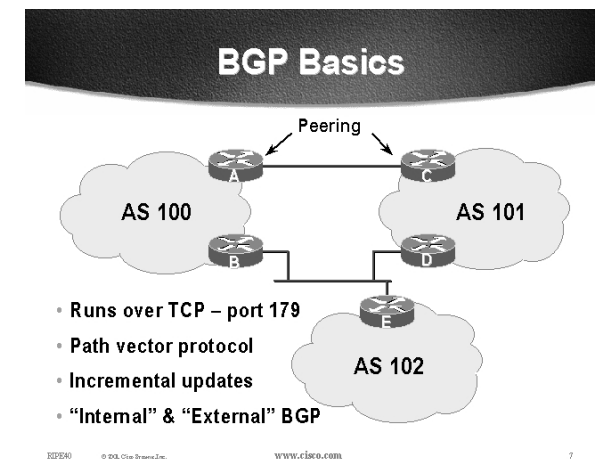
Autonomous System (AS)

- Insieme di reti con le stesse politiche di routing
- Unico protocollo di routing
- Solitamente gestito da un'unica entità, in modalità collaborativa e controllata

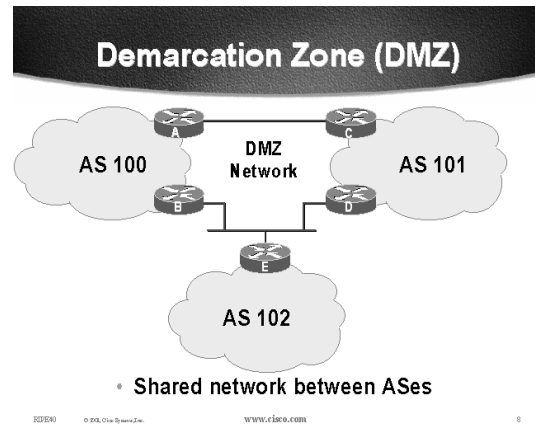


Basi del BGP

- Lavora sulla porta 179 TCP
- Aggiornamenti incrementali tabelle
- BGP interno ed esterno



Zona di Demarcazione (DMZ)



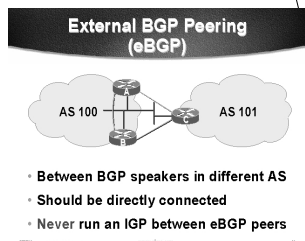
Reti condivise tra AS differenti

Comuni operazioni BGP

- Impara instradamenti multipli attraverso interlocutori BGP interni ed esterni
- Sceglie l'instradamento migliore e lo applica alla tabella di forwarding
- Invia il percorso migliore ai "neighbours" BGP esterni
- La scelta dell'instradamento migliore è condizionata dalle politiche adottate

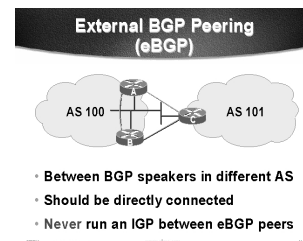
Peering BGP esterno (eBGP)

- Tra interlocutori BGP di AS differenti
- Devono essere connessi direttamente
- Tra peers esterni (eBGP) non utilizzare iBGP !



Schema di configurazione

- Router BGP xxx
- Network X.Y.Z.0
- Neighbor *addr* route map xx in/out
- Neighbor *addr* remote AS yyy
- Route map xx *permit prog*
Match *condition* 111
[set commands]
- acl 111 permit/deny

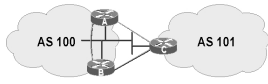


Configurare il BGP esterno-A

Router A in AS100

```
interface ethernet 5/0
ip address 222.222.10.2 255.255.255.240
router bgp 100
network 220.220.8.0 mask 255.255.252.0
neighbor 222.222.10.1 remote-as 101
neighbor 222.222.10.1 prefix-list RouterC in
neighbor 222.222.10.1 prefix-list RouterC out
```

External BGP Peering (eBGP)



- Between BGP speakers in different AS
- Should be directly connected
- Never run an IGP between eBGP peers

25/02/2004

Davide Giuffrida

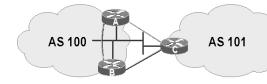
25

Configurare il BGP esterno-C

Router C in AS101

```
interface ethernet 1/0/0
ip address 222.222.10.1 255.255.255.240
router bgp 101
network 220.220.16.0 mask 255.255.240.0
neighbor 222.222.10.2 remote-as 100
neighbor 222.222.10.2 prefix-list RouterA in
neighbor 222.222.10.2 prefix-list RouterA out
```

External BGP Peering (eBGP)



- Between BGP speakers in different AS
- Should be directly connected
- Never run an IGP between eBGP peers

25/02/2004

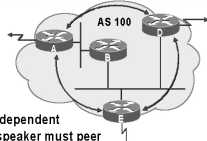
Davide Giuffrida

26

BGP interno (iBGP)

- Sessioni BGP all'interno dello stesso AS
- Non è necessario che siano connessi direttamente
- Interlocutori iBGP danno luogo a reti connesse.

Internal BGP Peering (iBGP)



- Topology independent
- Each iBGP speaker must peer with every other iBGP speaker in the AS

25/02/2004

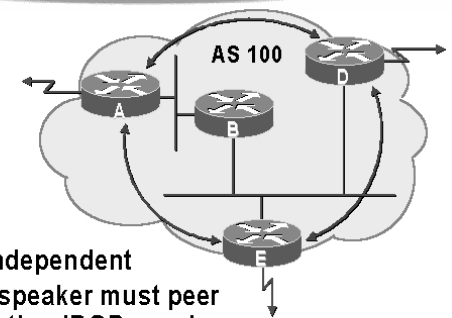
Davide Giuffrida

27

Peering BGP interno (iBGP)

- La tipologia delle reti è indipendente
- Ciascun interlocutore iBGP deve colloquiare con ogni altro peer iBGP presente nell'AS

Internal BGP Peering (iBGP)



- Topology independent
- Each iBGP speaker must peer with every other iBGP speaker in the AS

81E940

© 2001, Cisco Systems, Inc.

www.cisco.com

13

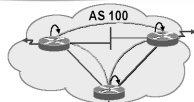
Davide Giuffrida

28

Peering con l'indirizzo loop-back

- Per il peering iBGP e' meglio utilizzare l'indirizzo di loop-back
L'interfaccia loop-back non cade mai!
- La sessione iBGP non dipende dallo stato di una singola interfaccia
- La sessione iBGP non dipende dalla topologia fisica della rete sottostante

Peering to Loop-Back Address



- Peer with loop-back address
Loop-back interface does not go down - ever!
- iBGP session is not dependent on state of a single interface
- iBGP session is not dependent on physical topology

25/02/2004

Davide Giuffrida

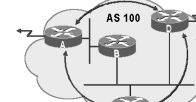
29

Configurare il BGP interno-A

Router A

```
interface loopback 0
ip address 215.10.7.1 255.255.255.255
router bgp 100
network 220.220.1.0
neighbor 215.10.7.2 remote-as 100
neighbor 215.10.7.2 update-source loopback0
neighbor 215.10.7.3 remote-as 100
neighbor 215.10.7.3 update-source loopback0
```

Internal BGP Peering (iBGP)



- Topology independent
- Each iBGP speaker must peer with every other iBGP speaker in the AS

25/02/2004

Davide Giuffrida

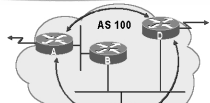
30

Configurare il BGP interno-B

Router B

```
interface loopback 0
ip address 215.10.7.2 255.255.255.255
router bgp 100
network 220.220.5.0
neighbor 215.10.7.1 remote-as 100
neighbor 215.10.7.1 update-source loopback0
neighbor 215.10.7.3 remote-as 100
neighbor 215.10.7.3 update-source loopback0
```

Internal BGP Peering (iBGP)



- Topology independent
- Each iBGP speaker must peer with every other iBGP speaker in the AS

25/02/2004

Davide Giuffrida

31

Parte III In Pratica ...

- Esempio: Connessione di 3 As
- Diagnostica: Raggiungibilità rete

25/02/2004

Davide Giuffrida

32

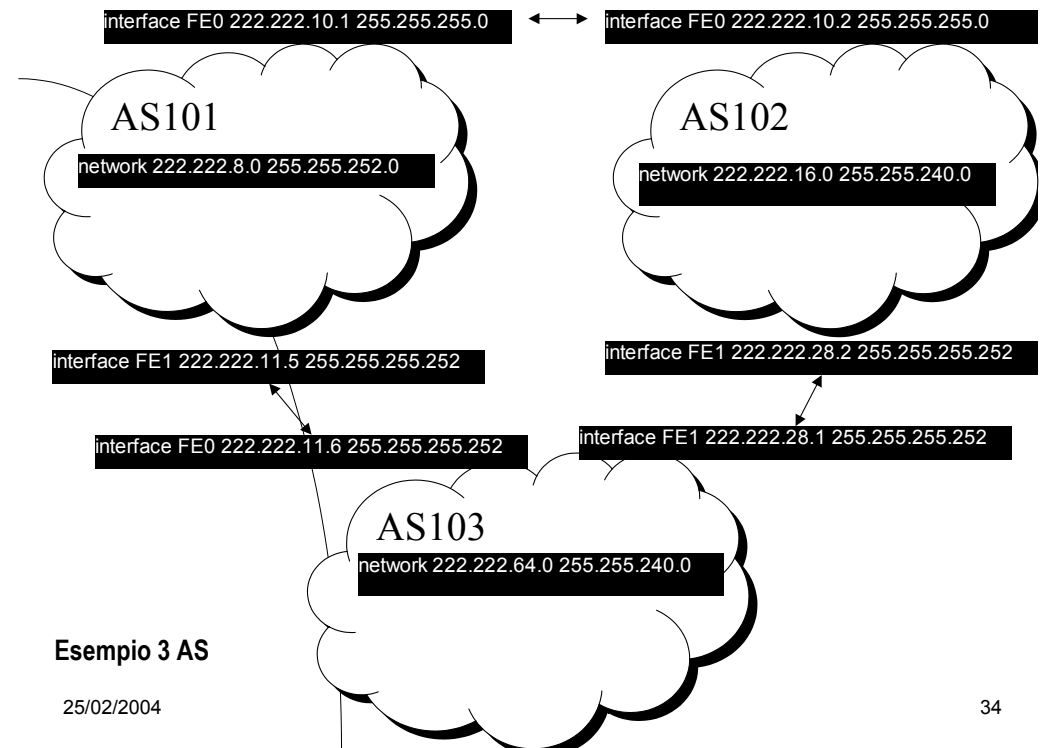
Esempio 3 AS

- AS 101
222.222.8.0 mask 255.255.252.0
- AS 102
222.222.16.0 mask 255.255.240.0
- AS 103
222.222.64.0 mask 255.255.240.0

25/02/2004

Davide Giuffrida

33



25/02/2004

34

Selezione del percorso

Qual'è il path migliore ?

25/02/2004

Davide Giuffrida

35

Algoritmo di scelta del Path 1/3

- Ignora il path se manca la route verso il gateway (next hop)
- Il peso maggiore vince (locale nel router)
- La local preference vince (all'interno dell'AS)
- Preferisce le route originate localmente
- L'AS path piu' breve è preferito

25/02/2004

Davide Giuffrida

36

Algoritmo di scelta del Path 2/3

- Origin code migliore:
IGP < EGP < incomplete
- Multi-Exit Discriminator (MED) inferiore:
se usato bgp deterministic-med, ordinare i paths prima della comparazione
- se usato bgp always-compare-med, comparare per ogni percorso
- Altrimenti MED viene considerato solo se i paths provengono dal medesimo AS. Default

25/02/2004

Davide Giuffrida

37

Algoritmo di scelta del Path 3/3

- Meglio path eBGP di path iBGP
 - Meglio path con metrica IGP inferiore verso il next-hop
 - Meglio il router-id inferiore (originator-id per routes riflesse)
 - Preferire Cluster-List più corte
- Non preoccuparsi degli attributi di Route Reflector ! (usati da provider mondiali)
- Meglio IP address di neighbor inferiore

25/02/2004

Davide Giuffrida

38

Applicare le politiche al BGP

Controllare e correggere i controlli BGP

25/02/2004

Davide Giuffrida

39

Applicare le politiche al BGP

- Applicare le Politiche:
Decisioni in base agli AS da attraversare, le community o il prefix
Rifiutare o accettare le route selezionate
Impostazione degli attributi per influenzare la selezione del path
- Strumenti:
Prefix-list (filtra i prefissi)
Filter-list (filtra gli AS)
Route-maps e communities

25/02/2004

Davide Giuffrida

40

Politiche BGP – Prefix-list:

- Filtrare le routes in base ai prefissi
- Inbound and Outbound

```
router bgp 200
  neighbor 220.200.1.1 remote-as 210
  neighbor 220.200.1.1 prefix-list PEER-IN in
  neighbor 220.200.1.1 prefix-list PEER-OUT out
  ip prefix-list PEER-IN deny 218.10.0.0/16
  ip prefix-list PEER-IN permit 0.0.0.0/0 le 32
  ip prefix-list PEER-OUT permit 215.7.0.0/16
```

25/02/2004

Davide Giuffrida

41

Politiche BGP – Filter-list:

- Filtrare le routes in base al percorso AS
- Inbound and Outbound

```
router bgp 100
  neighbor 220.200.1.1 remote-as 210
  neighbor 220.200.1.1 filter-list 5 out
  neighbor 220.200.1.1 filter-list 6 in
  !
  ip as-path access-list 5 permit ^200$
  ip as-path access-list 6 permit ^150$
```

25/02/2004

Davide Giuffrida

42

Politiche BGP – Route Maps:

- Una route-map è assimilabile ad un “programma” per IOS
- Ha numeri di riga, come un programma
- Ogni riga è un’istruzione (condizione/azione)
- Concetto:
se condizione soddisfatta,
eseguo ed esco, altrimenti
se condizione soddisfatta,
eseguo ed esco, altrimenti....

25/02/2004

Davide Giuffrida

43

Politiche BGP – Communities

```
Router bgp 100
  neighbor 220.200.1.1 remote-as 200
  neighbor 220.200.1.1 send-community
  neighbor 220.200.1.1 route-map set-community out
  !
  Route-map set-community permit 10
  Match Ip address prefix-list NO-ANNOUNCE
  Set community no-export
  !
  Route-map set-community permit 20
  !
  Ip prefix-list NO-ANNOUNCE permit 172.168.0.0/16 ge 17
```

25/02/2004

Davide Giuffrida

44

Politiche BGP – Matching Communities

```
Router bgp 100
neighbor 220.200.1.2 remote-as 200
neighbor 220.200.1.2 route-map filter-on-community in
!
Route-map filter-on-community permit 10
Match community 1
Set local-preference 50
!
Route-map filter-on-community permit 20
Match community 2 exact-match
Set local-preference 200
!
Ip community-list 1 permit 150:3 200:5
Ip community-list 2 permit 88:6
```

25/02/2004

Davide Giuffrida

45

Parte IV Verificare il BGP

- Looking Glass
<http://www.nat.bg/look/>
<http://www.inetcomm.net/lg.shtml>
<http://www.traceroute.org>
- Progetto Zebra (linux router)
- Progetto Ris (route info service)
- Proteggere il BGP (S-BGP)
- Cisco cmd: Sh ip bgp ?

25/02/2004

Davide Giuffrida

46

Looking Glass:

Address <http://www.traceroute.org/>

...traceroute.org...

Maintained by [Thomas Kernen](#)
Please feel free to send updates, links, corrections, extra info
Note that I don't provide support for the linked web pages

hosted by
DECKPOINT

By country:

Argentina	Armenia	Australia	Austria	Bangladesh	Belgium	Bolivia	Brasil
Bulgaria	Canada	Chile	China	Colombia	Costa Rica	Croatia	Cyprus
Czech Republic	Denmark	Estonia	Faroe Islands	Finland	France	Germany	Greece
Grenada	Hong Kong	Iceland	India	Indonesia	Israel	Italy	Japan
Korea	Kyrgyzstan	Latvia	Luxembourg	Malaysia	Malta	Mexico	The Netherlands
New Zealand	Paraguay	Philippines	Poland	Portugal	Romania	Russia	Saint Kitts
Saint Vincent	Singapore	Slovakia	South Africa	Spain	Sweden	Switzerland	Taiwan
Thailand	Turkey	Ukraine	United Kingdom	USA			

Or:

Looking Glass	Route Servers	BGP Info	Traceroute & Looking Glass source code	Extras
-------------------------------	-------------------------------	--------------------------	--	------------------------

25/02/2004

Davide Giuffrida

47

Progetto Zebra



- www.zebra.org
- Protocolli di routing supportati
bgpd: BGP v4, BGP v4+
ripd: RIP v1, v2
ripngd: RIPng
ospfd: OSPFv2
ospf6d: OSPFv3

25/02/2004

Davide Giuffrida

48

Progetto Zebra



- www.zebra.org
- RFC Supportati
 - 2453 RIP and RIPv2
 - 2080 RIPng
 - 2328 OSPF
 - 2460, 2373 2463, 2464 IPv6
 - 2236 IGMP and IGMPv2
 - 1812 Router Requirements
 - 1771 BGP4
 - 1157 SNMP
 - 2011, 2012 2013 MIB2
 - 1493 Bridge MIB
 - 1643 Ethernet MIB
 - 1757 RMON(4groups)
 - 1724 RIPv2 MIB
 - 1850 OSPF MIB
 - 1657 BGP4 MIB
 - 2037 Entity MIB
 - 2096 IP Forwarding Table MIBs

25/02/2004

Davide Giuffrida

49

Progetto Ris



- <http://www.ripe.net/ripencncc/pub-services/np/ris/index.html>
- The Routing Information Service (RIS) provides information about BGP routing much like "Looking Glass" services offered elsewhere on the Internet. However, the RIS is much more than a conventional "Looking Glass", because it can provide historical information about Internet routing without being bound to the perspective of a particular autonomous system. The service collects routing information by using Remote Route Collectors at different locations around the world and integrates this information into a comprehensive view.

Davide Giuffrida

50

Progetto Ris



- <http://www.ripe.net/ripencncc/pub-services/np/ris/index.html>
- For example, the RIS allows you to see how the RIPE NCC web server has been reachable through the LINX on January 1st, 2002. You can change the fields in the form to try your own queries.
- You can also see in which way a specific AS has been reachable as seen from our RIPE NCC collector over the last few months. Note that the latter question produces a concise output but takes several minutes to process several months worth of BGP data.

25/02/2004

Davide Giuffrida

51

Protocollo a rischio ?

- Che i router possano essere aggrediti non è certo una novità. Ma da qualche tempo esperti IT prefigurano rischi ancora maggiori per provider e imprese
- Il router, vero e proprio snodo del traffico internet, è da lungo tempo nel mirino di un certo tipo di attacchi informatici ma ora, stando alle rilevazioni di alcuni esperti, la "questione router" in materia di sicurezza sta diventando centrale. E fa preoccupare provider ed imprese.
- Stando alle dichiarazioni riportate dalla Reuters, esperti come Carlos Recalde, direttore TLC presso KPMG, ritengono che chi si dedica ad attività aggressive via internet stia sempre più spesso focalizzando le proprie attenzioni sulle vulnerabilità dei router. Falle che, a sentire Recalde e gli altri esperti, consentirebbero agli aggressori di boicottare grandi quantità di traffico dati sulla rete delle reti.
- A segnalare un interesse "nuovo" per vecchi problemi che da anni affliggono, per esempio, il Border Gateway Protocol (BGP) (che traduce le tabelle di routing da device di produttori diversi), sarebbero gli scambi di tool di attacco scambiati sui network chat IRC e che sarebbero pensati proprio per aggredire i router.
- Secondo Recalde uno dei problemi più importanti per gli Internet Service Provider oggi è proprio quello di lavorare sulla sicurezza dei propri router ma sia le imprese che gli operatori non sarebbero pronti. Proprio in KPMG uno script capace di mappare le configurazioni di routing e verificare in qualsiasi momento che non siano state modificate dall'esterno, non viene considerato sufficiente perché consente analisi "a posteriori" di un attacco e non "a priori".

25/02/2004

Davide Giuffrida

52

Sicurezza: Filtri e SBGP

Dal CERT, centro sulla sicurezza informatica finanziata dal governo americano, arrivano in questo senso le "solite" raccomandazioni, come quella di non usare mai password di default per l'amministrazione dei router o di aggiungere strutture di chiave pubblica per una maggiore sicurezza basata sull'autenticazione.

Ma le responsabilità maggiori relativamente ad attacchi che potrebbero compromettere il traffico su porzioni molto ampie della rete, secondo Recalde, le condividono soprattutto i grandi operatori. E Jim Lippard, direttore della sicurezza presso l'operatore americano Global Crossing, sostiene che oltre a sistemi di autenticazione, i provider dovrebbero inserire filtri di gestione del traffico e tool di monitoraggio e individuazione delle origini di eventuali attacchi.

Ma, nonostante tutto questo, insiste Lippard, il problema è che deve essere aumentata la sicurezza di protocolli come il BGP. E va detto che in più occasioni, importanti produttori di sistemi di networking, come Cisco, hanno dovuto affrontare problemi legati proprio a falle di sicurezza.

Le soluzioni? La bacchetta magica ancora non s'è vista, però si lavora, come nel caso del Secure BGP Project (<http://www.net-tech.bbn.com/sbgp/sbgp-index.html>), proprio per sostituire o aggiornare quegli aspetti del routing che si presentano fallati. Lo sviluppo di quello che viene chiamato S-BGP sembra aprire molte speranze, perché "aggancia" l'aspetto dell'autenticazione al protocollo. Ma la strada è lunga anche per S-BGP che, per funzionare a dovere, deve essere adottato contestualmente da provider, produttori di router e internet registrar.

Riferimenti:

<http://www.ripe.net/ripe/meetings/archive/ripe-40/tutorials/bgp-tutorial/index.html>

<http://www.academ.com/nanog/feb1997/BGPTutorial/>

<http://www.cisco.com/warp/public/459/18.html>

<http://www.netaxs.com/~freedman/bgp.html>

<http://joe.lindsay.net/bgp.html>

http://www.avici.com/documentation/HTMLDocs/02223-04_revAA/index.html

<http://info.connect.com.au/docs/routing/general/multi-faq.shtml>

<http://www.bgpbook.com/>

<http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm>

http://www.cisco.com/warp/public/759/ipj_4-4/ipj_4-4_regional.html La vera storia di Internet

<http://www.ietf.org> (Riferimento per tutti gli standard RFC)