WORLD METEOROLOGICAL ORGANIZATION

COMMISSION FOR BASIC SYSTEM
OPAG ON
INFORMATION SYSTEMS & SERVICES
EXPERT TEAM MEETING ON DATA-
COMMUNICATION SYSTEMS AND TECHNIQUES

GENEVA, 27-30 SEPTEMBER 1999

ISS/ET-DCST 99/Doc. 11
(10.IX.1999)

ITEM 3.2

ENGLISH only

FURTHER DEVELOPMENT OF RECOMMENDED PRACTICES AND PROCEDURES FOR THE IMPLEMENTATION AND USE OF TCP/IP ON THE GTS

Practical BGP implementation guide

Submitted by France

Summary and Purpose of Document

The objective of the document is to describe a practical implementation of BGP to be used on the GTS.
AS numbers and IP addresses of hosts follow WMO rules when information was available at time of tests.

ACTION PROPOSED

The session is invited to note the information on the document

Routing on the GTS

1. The dynamic routing protocol adopted (in Attachment II-15 to the manual on the GTS) for TCP/IP on the GTS is Border Gateway Protocol version 4 (BGP4).

2. BGP is a protocol intended for use between an Autonomous System[1] and the Internet at large, or between two or more Autonomous Systems. The TCP/IP based national networks of WMO Members are Autonomous Systems and therefore BGP4 is an appropriate routing protocol to use to exchange routing information between these networks on the GTS.

3. BGP within meteorogical organisation is not at the moment a widely used protocol. Therefore, a test bed was setted up in order to give some practical information and help in the implementation of BGP on the GTS

Description of the tests

1. Phase 1. : a configuration involving six sites was built : Moscow, Beijing, Hong-Kong, Seoul, Melbourne and Tokyo. In each center only one host was involved in the BGP process. As some diverses routes were available between Beijing and Tokyo, the test focuses on configuation technique to allow diverse routing when possible and trafic control when such diverse routing was not desirable

2. Phase 2. : the goal of this second test was to focus on site running BGP on multiple routers. In order to do so, a quite complicated configuration was built at RTH Tokyo. The configuration of RTH Tokyo is not to be used per se, but show how in such configuration to manage BGP.

Methodology and Results

It was decided to add feature step by step. At each significant step, a screenshot of all the configuration was taken.

On the ciscos routers involved, the screenshots consist in :

- wr ter = dump the configuration at the screen
- sh ip rou = dump the IP routing table

and when applicable :

- sh ip bgp = show the BGP view of the router
- sh ip bgp a.b.c.d = where a.b.c.d is the IP address of a site in the BGP table
- sh ip bgp summ = summary of the BGP status

In this paper, it was decided to show quite large dumps when interesting. The key point will be highlighted (red print and underlined). The group is invited to discuss which and how to extract the most relevant information.

[1] An Autonomous System is defined in RFC 1630 as "a set of routers under a single technical administration, using an interior gateway protocol and common metrics to route packets within the AS, and using an exterior gateway protocol to route packets to other ASs."

Phase 1 : BGP configuration in a multi-site, partially redundant sub network

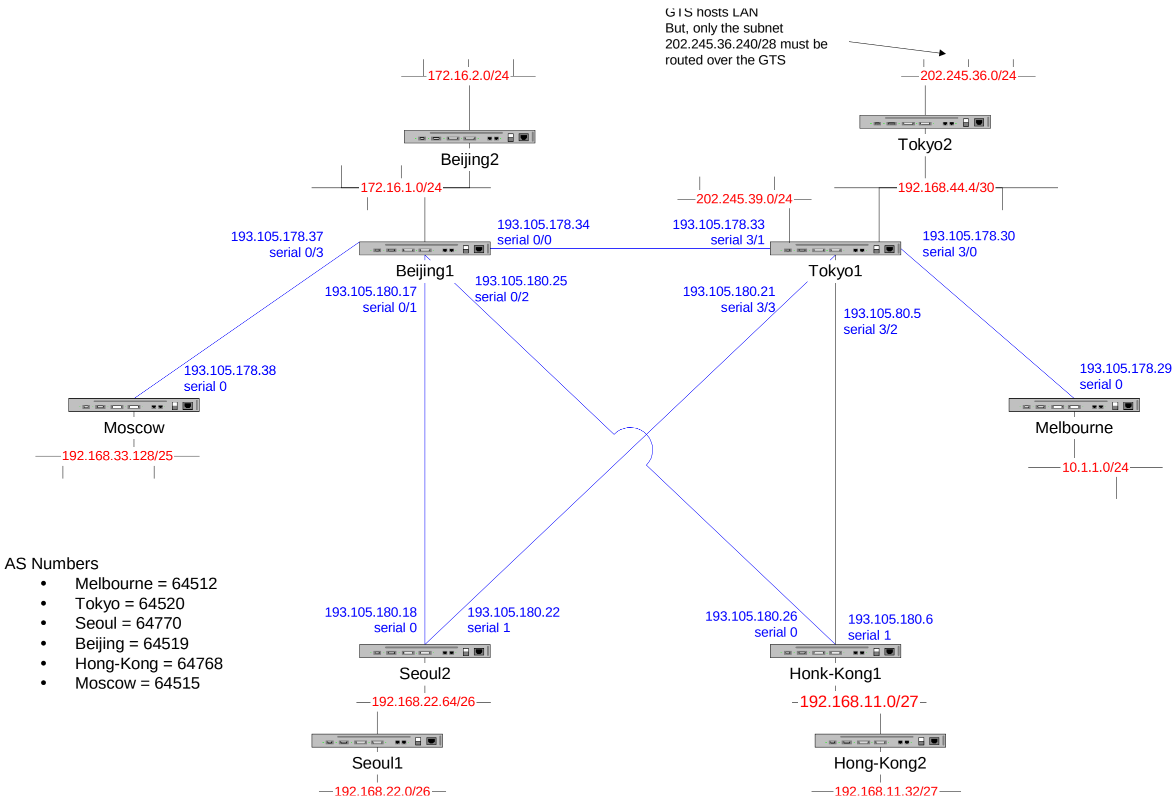
Tests goals :

1. Established basic BGP configuration
2. Redistribution between BGP and IGP (Interior Gateway Protocol)
   - routes are learnt by BGP, how to redistribute them toward the internal network. Check that only subnetworked routes learnt will be redistributed.
   - how to guarantee that only some internal routes will be seen on the GTS[2]
3. Use of backup routes and selection of backup
   - in a meshed or partially meshed environment, verify that backup routes are used and controlled which one are preferred
4. Control redistribution of BGP routes
   - On a peer to peer basis, sites will allow connection through their site for others NMCs. In our example, Beijing and Tokyo agrees to let Melbourne and Moscow to use the link Tokyo-Beijing. But, do not allow Seoul to do so.

Architecture and addressing on the network

When available, addresses and AS numbers used are those provided by WMO. When missing, it was decided to use unofficial IP addresses, which is not conform to WMO policy.

BGP test bed (Phase 1)



Sites involved in each subtest :

1. All sites
2. Beijing and Tokyo
3. Beijing and Tokyo = If the Beijing-Tokyo link fails, use Hong-Kong as the 1st backup, and Seoul as the second. But neither Seoul nor Hong-Kong can reach Beijing via Tokyo or Tokyo via Beijing.
4. Mainly Beijing and Tokyo. More sophisticated feature create more complicated configurations. It was decided to add complexity on RTHs when needed.

---

[2] In the « Guide of TCP/IP over the GTS », it has been decided to allow only operationnal data to be exchanged over the GTS. Therefore, it is necessary to distribute only subnetworked routes on the GTS. E.G. if a site owns a class B network, only a subnet of that one must be seen on the GTS

Step 1 : Basic BGP configurations

Beijing1

```
sh run
Building configuration...

Current configuration:
!
version 11.3
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname Beijing1
!
enable password c
!
!
interface Ethernet0/0
 description interconnection LAN
 ip address 172.16.1.254 255.255.255.0
!
interface Serial0/0
 description link Beijing1 - Tokyo1
 ip address 193.105.178.34 255.255.255.252
 encapsulation ppp
!
interface Serial0/1
 description link Beijing1 - Seoul1
 ip address 193.105.180.17 255.255.255.252
 encapsulation ppp
 clockrate 64000
!
interface Serial0/2
 description link Beijing1 - Hong-Kong1
 ip address 193.105.180.25 255.255.255.252
 encapsulation ppp
 clockrate 64000
!
interface Serial0/3
 description link Beijing1 - Moscow
 ip address 193.105.178.37 255.255.255.252
 clockrate 64000
!
router rip
 version 2
 redistribute bgp 64519 metric 1
 network 172.16.0.0
 no auto-summary
!
```

➨ If "version 2" is not mentioned, RIP used version 1. RIPv1 is quite a stupid protocol. Not only subnets learnt via BGP will be resent but RIPv1 will resend full routes. No auto-summary tells RIP to really send subnets. See Tokyo alternate config.

```
router bgp 64519
 network 172.16.2.0 mask 255.255.255.0
 neighbor 193.105.178.33 remote-as 64520
 neighbor 193.105.180.18 remote-as 64770
 neighbor 193.105.180.26 remote-as 64768
!
```

➨ Tells which (sub)networks must be announced with BGP. BGP must know who are its neighbours.

```
ip classless
!
```

➨ At times the router might receive packets destined for a subnet of a network that has no network default route. To have the Cisco IOS software forward such packets to the best supernet route possible, use the **ip classless** global configuration command.

```
!
line con 0
line aux 0
line vty 0 4
 no login
!
no scheduler allocate
end

Beijing1#sh ip route
Codes: C - connected, S - static, I - IGRP, 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, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

B    202.245.36.0/24 [20/1] via 193.105.178.33, 01:58:52
     193.105.180.0/24 is variably subnetted, 4 subnets, 2 masks
C       193.105.180.24/30 is directly connected, Serial0/2
C       193.105.180.26/32 is directly connected, Serial0/2
C       193.105.180.16/30 is directly connected, Serial0/1
C       193.105.180.18/32 is directly connected, Serial0/1
B    202.245.39.0/24 [20/0] via 193.105.178.33, 01:58:53
     193.105.178.0/24 is variably subnetted, 3 subnets, 2 masks
C       193.105.178.32/30 is directly connected, Serial0/0
C       193.105.178.33/32 is directly connected, Serial0/0
C       193.105.178.36/30 is directly connected, Serial0/3
     172.16.0.0/24 is subnetted, 2 subnets
C       172.16.1.0 is directly connected, Ethernet0/0
R       172.16.2.0 [120/1] via 172.16.1.253, 00:00:22, Ethernet0/0
     192.168.11.0/27 is subnetted, 1 subnets
B       192.168.11.32 [20/0] via 193.105.180.26, 00:07:59
     192.168.22.0/26 is subnetted, 1 subnets
B       192.168.22.0 [20/0] via 193.105.180.18, 00:17:54
```

➨ R are RIP learnt routes, B are BGP ones

```
Beijing1#sh ip bgp
BGP table version is 13, local router ID is 193.105.180.25
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 172.16.2.0/24    0.0.0.0                  1         32768 i
*> 192.168.11.32/27 193.105.180.26           0             0 64768 i
*                   193.105.178.33                         0 64520 64768 i
*  192.168.22.0/26  193.105.178.33                         0 64520 64770 i
*>                  193.105.180.18           0             0 64770 i
*  202.245.36.0     193.105.180.26                         0 64768 64520 i
*                   193.105.180.18                         0 64770 64520 i
*>                  193.105.178.33           1             0 64520 i
*  202.245.39.0     193.105.180.26                         0 64768 64520 i
*                   193.105.180.18                         0 64770 64520 i
*>                  193.105.178.33           0             0 64520 i
```

➨ These are very interesting informations. In the BGP routing table (which is NOT the ip routing table), beijing1 knows three ways to go to Tokyo subnetworks. The third line start by *>; this is the prefered route, which is directly linked to AS 64520 (Tokyo). But alternate paths are available, by Honk-Kong (64768) or Seoul (64770). In case of failure of Tokyo link, one of these alternate paths will be used.

```
Beijing1# sh ip bgp 202.245.36.0
BGP routing table entry for 202.245.36.0/24, version 5
Paths: (3 available, best #3, advertised over EBGP)
  64768 64520
    193.105.180.26 from 193.105.180.26
      Origin IGP, localpref 100, valid, external
  64770 64520
    193.105.180.18 from 193.105.180.18 (193.105.180.22)
      Origin IGP, localpref 100, valid, external
  64520
    193.105.178.33 from 193.105.178.33 (202.245.39.254)
      Origin IGP, metric 1, localpref 100, valid, external, best
Beijing1#
```

Seoul1

```
Seoul1#sh ip route
Codes: C - connected, S - static, I - IGRP, 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, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

B    202.245.39.0/24 [20/0] via 193.105.180.21, 00:22:26
B    202.245.36.0/24 [20/1] via 193.105.180.21, 00:22:26
     193.105.180.0/24 is variably subnetted, 4 subnets, 2 masks
C       193.105.180.20/30 is directly connected, Serial1
C       193.105.180.21/32 is directly connected, Serial1
C       193.105.180.16/30 is directly connected, Serial0
C       193.105.180.17/32 is directly connected, Serial0
     192.168.11.0/27 is subnetted, 1 subnets
```

```
B       192.168.11.32 [20/0] via 193.105.180.17, 00:10:37
     192.168.22.0/26 is subnetted, 2 subnets
C       192.168.22.64 is directly connected, Ethernet0
S       192.168.22.0 [5/0] via 192.168.22.125
     172.16.0.0/24 is subnetted, 1 subnets
B       172.16.2.0 [20/1] via 193.105.180.17, 00:22:23
Seoul1#sh ip bgp
BGP table version is 12, local router ID is 193.105.180.22
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 172.16.2.0/24    193.105.180.17           1             0 64519 i
*                   193.105.180.21                         0 64520 64519 i
*> 192.168.11.32/27 193.105.180.17                         0 64519 64768 i
*                   193.105.180.21                         0 64520 64768 i
*> 192.168.22.0/26  0.0.0.0                  0         32768 i
*  202.245.36.0     193.105.180.17                         0 64519 64520 i
*>                  193.105.180.21           1             0 64520 i
*  202.245.39.0     193.105.180.17                         0 64519 64520 i
*>                  193.105.180.21           0             0 64520 i
```

➨ Seoul at this step knows two ways to reach Tokyo and Beijing, directly and via the other site. This is going to be corrected in the next step.

Step 2 : Redistribution between BGP and IGP and vice versa

Step 2.1 : BGP to IGP

Step 2.1.a = use of RIPv1 locally

```
Tokyo1#wr ter
router rip
 version 1
 redistribute bgp 64520 metric 1
 passive-interface Serial3/0
 passive-interface Serial3/1
 passive-interface Serial3/2
 passive-interface Serial3/3
 network 192.168.44.0
 no auto-summary
!
router bgp 64520
 network 202.245.39.0
 network 202.245.36.0
 neighbor 193.105.178.34 remote-as 64519
 neighbor 193.105.180.6 remote-as 64768
 neighbor 193.105.180.22 remote-as 64770
!
ip classless

Tokyo1#sh ip rou
C    202.245.39.0/24 is directly connected, Loopback0
     193.105.178.0/24 is variably subnetted, 4 subnets, 2 masks
C       193.105.178.28/30 is directly connected, Serial3/0
C       193.105.178.29/32 is directly connected, Serial3/0
C       193.105.178.34/32 is directly connected, Serial3/1
C       193.105.178.32/30 is directly connected, Serial3/1
R    202.245.36.0/24 [120/1] via 192.168.44.6, 00:00:17, Ethernet0/0
     193.105.180.0/24 is variably subnetted, 4 subnets, 2 masks
C       193.105.180.20/30 is directly connected, Serial3/3
C       193.105.180.22/32 is directly connected, Serial3/3
C       193.105.180.4/30 is directly connected, Serial3/2
C       193.105.180.6/32 is directly connected, Serial3/2
C    192.168.44.0/24 is directly connected, Ethernet0/0
     172.16.0.0/24 is subnetted, 1 subnets
B       172.16.2.0 [20/1] via 193.105.178.34, 00:23:02

!!!! Only BGP with Beijing is running at this stage


Tokyo2#sh ip rou
Codes: C - connected, S - static, I - IGRP, 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, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR
       T - traffic engineered route

Gateway of last resort is not set

C    202.245.36.0/24 is directly connected, Loopback0
```

```
C    192.168.44.0/24 is directly connected, Ethernet0
R    172.16.0.0/16 [120/1] via 192.168.44.5, Ethernet0
```

If Tokyo1 really knows the subnet, Tokyo2 received all the Beijing network. This is due to RIPv1. Amongst the IGP (Interior Gateway Protocol), in order to follow the WMO rules, the sites must therefore use IGP with VLSN (Variable SubNetwork Length) support. On cisco routers, RIPv2, EIGRP and OSPF are VLSN compliant (on the opposite, RIP, IGRP are not). Of course, static routes are also possible but not advised.

Step 2.1.b = Use of RIP version 2

```
Tokyo1#wr ter
router rip
 version 2
 redistribute bgp 64520 metric 1
 passive-interface Serial3/0
 passive-interface Serial3/1
 passive-interface Serial3/2
 passive-interface Serial3/3
 network 192.168.44.0
 no auto-summary
!

Tokyo2#sh ip rou
Codes: C - connected, S - static, I - IGRP, 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, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR
       T - traffic engineered route

Gateway of last resort is not set

C    202.245.36.0/24 is directly connected, Loopback0
C    192.168.44.0/24 is directly connected, Ethernet0
     172.16.0.0/24 is subnetted, 1 subnets
R       172.16.2.0 [120/1] via 192.168.44.5, Ethernet0
```

➨ Subnets are know seen on Tokyo2

Step 2.2

The netwok connected to Tokyo 2 (202.245.36.0) is subnetted with 255.255.255.0. But, on this network, only a part is concerned by GTS hosts. The “false” GTS subnet hosts is 202.245.36.240/28. Tokyo1 in order to send BGP updates with this “subsubnetwork” must have a matching entry in its routing table. Two solutions, either put a static route in Tokyo1 (not the best solution) or configure Tokyo2 to force it to send this subsubroute to Tokyo1. This is the solution described below.

Tokyo2

```
Tokyo2#wr ter
interface Ethernet0
 description FTP-GTS hosts LAN
 ip address 202.245.36.254 255.255.255.0
 no ip directed-broadcast
!
interface Ethernet1
 description Interconnection LAN
 ip address 192.168.44.6 255.255.255.0
 ip directed-broadcast
!
interface BRI0
 no ip address
 no ip directed-broadcast
 shutdown
!
router rip
 version 2
 redistribute connected
 redistribute static
 network 192.168.44.0
 no auto-summary
!
no ip classless
ip route 202.245.36.240 255.255.255.240 Ethernet0
```

➨ the last IP route command create a static entry in Tokyo2 routing table. The RIP protocol resend this static route.

Tokyo1

```
Tokyo1#sh ip rou
Codes: C - connected, S - static, I - IGRP, 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, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

     10.0.0.0/24 is subnetted, 1 subnets
B       10.1.1.0 [20/0] via 193.105.178.29, 00:10:35
C    202.245.39.0/24 is directly connected, Loopback0
     193.105.178.0/24 is variably subnetted, 4 subnets, 2 masks
C       193.105.178.28/30 is directly connected, Serial3/0
C       193.105.178.29/32 is directly connected, Serial3/0
C       193.105.178.34/32 is directly connected, Serial3/1
C       193.105.178.32/30 is directly connected, Serial3/1
     202.245.36.0/24 is variably subnetted, 2 subnets, 2 masks
R       202.245.36.0/24 [120/1] via 192.168.44.6, 00:00:26, Ethernet0/0
R       202.245.36.240/28 [120/1] via 192.168.44.6, 00:00:26, Ethernet0/0
     193.105.180.0/24 is variably subnetted, 4 subnets, 2 masks
C       193.105.180.20/30 is directly connected, Serial3/3
C       193.105.180.22/32 is directly connected, Serial3/3
C       193.105.180.4/30 is directly connected, Serial3/2
C       193.105.180.6/32 is directly connected, Serial3/2
C    192.168.44.0/24 is directly connected, Ethernet0/0
     192.168.33.0/25 is subnetted, 1 subnets
B       192.168.33.128 [20/0] via 193.105.178.34, 00:06:20
     192.168.11.0/27 is subnetted, 1 subnets
B       192.168.11.32 [20/100] via 193.105.180.6, 00:06:16
     192.168.22.0/26 is subnetted, 1 subnets
B       192.168.22.0 [20/200] via 193.105.180.22, 00:52:59
     172.16.0.0/24 is subnetted, 1 subnets
B       172.16.2.0 [20/74] via 193.105.178.34, 00:06:20
```

➨ The 202.245.36.240 is in the routing table, therefore, it will be sent by BGP.

Step 3 = Selection of backup routes

When multiple BGP routes are available, the routing processor in the cisco have to decide which route(s) will be inserted in the routing table.
Please find below, a copy of cisco documentation where this process is described :

> Path selection is based on the following:
> 1-If **NextHop is inaccessible** do not consider it.
> 2-Prefer the largest **Weight.**
> 3-If same weight prefer largest **Local Preference.**
> 4-If same Local Preference prefer **the route that the specified router has originated**.
> 5-If no route was originated prefer the **shorter AS path.**
> 6-If all paths are external prefer the **lowest origin code (IGP<EGP<INCOMPLETE)**.
> 7-If origin codes are the same prefer the path with the **lowest MED**.
> 8-If path is the same length prefer the **External path over Internal**.
> 9-If IGP synchronization is disabled and only internal path remain prefer the path through the **closest IGP neighbor.**
> 10-Prefer the route with the **lowest ip address** value for BGP router ID.

By default, when nothing especially configured, in our config, if the Tokyo-Beijing link is out of use, both Seoul and Hong-Kong will be completly equivalent on a BGP point of view for the first 9 steps above. Therefore, the selection will be done by 10. Which is more or less random !
With the IP addresses used on the test bed the « random » choice would be Seoul. It was decided to configure both Tokyo and Beijing to force selection of Hong-Kong has the first choice.
Point 4. And 8. Will therefore be used in Tokyo1 and Beijing1 cisco in order to change the default behaviour. It was decided to to use the “weight” attribute as it is not a BGP standard parameter.

Tokyo1 :

```
Tokyo1#wr ter
router bgp 64520
 network 202.245.39.0
 network 202.245.36.0
 neighbor 193.105.178.34 remote-as 64519
 neighbor 193.105.178.34 route-map beijing-policy-in in
 neighbor 193.105.180.6 remote-as 64768
 neighbor 193.105.180.6 route-map hong-kong-policy-in in
```

```
 neighbor 193.105.180.22 remote-as 64770
!
```

➨ The two new lines means that updates received (the last `in`) by neighbour 193.105.178.34 will be applied the route-map `beijing-policy-in`

```
ip classless
route-map hong-kong-policy-in permit 10
 set local-preference 200
!
route-map beijing-policy-in permit 10
 set local-preference 300
```

➨ The route-map beijing-policy-in set the local preference to 300. The highest the local prefrence is the most « trusted » is the neighbor.

```
!
Tokyo1#sh ip bgp
BGP table version is 40, local router ID is 202.245.39.254
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*  172.16.2.0/24    193.105.180.6              200      0 64768 64519 i
*                   193.105.180.22                      0 64770 64519 i
*>                  193.105.178.34        74    300      0 64519 i
*> 192.168.11.32/27 193.105.180.6          0    200      0 64768 i
*> 192.168.22.0/26  193.105.178.34              300      0 64519 64770 i
*                   193.105.180.6               200      0 64768 64519 64770 i
*                   193.105.180.22         0             0 64770 i
*> 202.245.36.0     192.168.44.6           1         32768 i
*> 202.245.39.0     0.0.0.0                0         32768 i
Tokyo1#sh ip bgp 172.16.2.0
BGP routing table entry for 172.16.2.0/24, version 40
Paths: (3 available, best #3, advertised over EBGP)
  64768 64519
    193.105.180.6 from 193.105.180.6 (193.105.180.26)
      Origin IGP, localpref 200, valid, external
  64770 64519
    193.105.180.22 from 193.105.180.22
      Origin IGP, localpref 100, valid, external
  64519
    193.105.178.34 from 193.105.178.34 (193.105.180.25)
      Origin IGP, metric 74, localpref 300, valid, external, best
```

By default the local preference is set to 100. Therefore, if Hong-Kong received a local preference of 200, and as this parameter as a higher precedence compared to AS path length, even if the Tokyo-Beijing link works all the packet between them will go through Honk-Kong. It is thus necessary to put also local preference at Beijing updates.
In Beijing we decided to use instead the MED option.

Beijing1

```
Beijing1#wr ter
router bgp 64519
 bgp always-compare-med
 network 172.16.2.0 mask 255.255.255.0
 neighbor 193.105.178.33 remote-as 64520
 neighbor 193.105.178.38 remote-as 64515
 neighbor 193.105.180.18 remote-as 64770
 neighbor 193.105.180.18 route-map seoul-policy-in in
 neighbor 193.105.180.26 remote-as 64768
 neighbor 193.105.180.26 route-map hongkong-policy-in in
!
ip classless
!
route-map hongkong-policy-in permit 10
 set metric 100
!
route-map seoul-policy-in permit 10
 set metric 200
!
Beijing1#sh ip bgp
BGP table version is 206, local router ID is 193.105.180.25
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*> 10.1.1.0/24      193.105.178.33                      0 64520 64512 i
*                   193.105.180.26       100             0 64768 64520 64512 i
*                   193.105.180.18       200             0 64770 64520 64512 i
*> 172.16.2.0/24    0.0.0.0               74         32768 i
*  192.168.11.32/27 193.105.178.33                      0 64520 64768 i
```

```
*>                   193.105.180.26       100             0 64768 i
*  192.168.22.0/26  193.105.178.33                       0 64520 64770 i
*>                   193.105.180.18       200             0 64770 i
*> 192.168.33.128/25 193.105.178.38          0             0 64515 i
*> 202.245.36.240/28 193.105.178.33          1             0 64520 i
*                    193.105.180.26       100             0 64768 64520 i
*                    193.105.180.18       200             0 64770 64520 i
*> 202.245.39.0     193.105.178.33         0             0 64520 i
*                    193.105.180.26       100             0 64768 64520 i
*                    193.105.180.18       200             0 64770 64520 i
Beijing1#sh ip bgp 202.245.39.0
BGP routing table entry for 202.245.39.0/24, version 205
Paths: (3 available, best #1, advertised over EBGP)
  64520
    193.105.178.33 from 193.105.178.33 (202.245.39.254)
      Origin IGP, metric 0, localpref 100, valid, external, best
  64768 64520
    193.105.180.26 from 193.105.180.26
      Origin IGP, metric 100, localpref 100, valid, external
  64770 64520
    193.105.180.18 from 193.105.180.18 (193.105.180.22)
      Origin IGP, metric 200, localpref 100, valid, external
```

➨ the result in Tokyo and Beijing will be the same. Honk-Kong will be prefered compared to Seoul as backup. We think that metric is probably a more logical way to establish preference between BGP routes. In Tokyo, we were obliged to put 300 to updates from Beijing in order to keep the “good” path. The best route will be the smallest metric.

Just to check, here is the BGP table and routing table in Tokyo with the Beijing-Tokyo link shut down.

Tokyo1:

```
interface Serial3/1
 description Link Tokyo-Beijing
 ip address 193.105.178.33 255.255.255.252
 encapsulation ppp
 shutdown
 clockrate 64000

Tokyo1#sh ip bgp
BGP table version is 43, local router ID is 202.245.39.254
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*> 172.16.2.0/24    193.105.180.6               200      0 64768 64519 i
*                    193.105.180.22                       0 64770 64519 i
*  192.168.11.32/27 193.105.180.22                       0 64770 64519 64768 i
*>                   193.105.180.6          0    200      0 64768 i
*> 192.168.22.0/26  193.105.180.6               200      0 64768 64519 64770 i
*                    193.105.180.22         0             0 64770 i
*> 202.245.36.0     192.168.44.6           1         32768 i
*> 202.245.39.0     0.0.0.0                0         32768 i
Tokyo1#sh ip bgp 172.16.2.0
BGP routing table entry for 172.16.2.0/24, version 42
Paths: (2 available, best #1, advertised over EBGP)
  64768 64519
    193.105.180.6 from 193.105.180.6 (193.105.180.26)
      Origin IGP, localpref 200, valid, external, best
  64770 64519
    193.105.180.22 from 193.105.180.22
      Origin IGP, localpref 100, valid, external
```

Step 4 = Control of routes redistributions

Up to now, is the previous configuration, we have established BGP connection between all sites. For example, a look at Seoul routing table show that Seoul is able to reach Hong-Kong via Tokyo or Beijing. This is not what we want in fact.

Seoul1:

```
Seoul1#sh ip route
Codes: C - connected, S - static, I - IGRP, 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, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

B    202.245.39.0/24 [20/0] via 193.105.180.21, 00:22:26
B    202.245.36.0/24 [20/1] via 193.105.180.21, 00:22:26
     193.105.180.0/24 is variably subnetted, 4 subnets, 2 masks
C       193.105.180.20/30 is directly connected, Serial1
```

```
C       193.105.180.21/32 is directly connected, Serial1
C       193.105.180.16/30 is directly connected, Serial0
C       193.105.180.17/32 is directly connected, Serial0
     192.168.11.0/27 is subnetted, 1 subnets
B       192.168.11.32 [20/0] via 193.105.180.17, 00:10:37
     192.168.22.0/26 is subnetted, 2 subnets
C       192.168.22.64 is directly connected, Ethernet0
S       192.168.22.0 [5/0] via 192.168.22.125
     172.16.0.0/24 is subnetted, 1 subnets
B       172.16.2.0 [20/1] via 193.105.180.17, 00:22:23
Seoul1#sh ip bgp
BGP table version is 12, local router ID is 193.105.180.22
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*> 172.16.2.0/24    193.105.180.17         1             0 64519 i
*                   193.105.180.21                       0 64520 64519 i
*> 192.168.11.32/27 193.105.180.17                       0 64519 64768 i
*                   193.105.180.21                       0 64520 64768 i
*> 192.168.22.0/26  0.0.0.0                0         32768 i
*  202.245.36.0     193.105.180.17                       0 64519 64520 i
*>                  193.105.180.21         1             0 64520 i
*  202.245.39.0     193.105.180.17                       0 64519 64520 i
*>                  193.105.180.21         0             0 64520 i
```

As it is, in our test bed, Tokyo and Beijing choice to deny access between Hong-Kong and Seoul, it is much more practical solution to put the extra configurations topics in Tokyo and Beijing.
In fact, to do so, there are multiple solutions[3].
First, we might think to filter the IP routes. It is possible to control with access-lists which routes will be accepted in routing updates. But, in this case, we must declare all the IP addresses concerned. Thanks to BGP, there is a more global solution, filters on AS numbers. All routes received from Seoul will be “tagged” with Seoul AS.
Again, two solutions are available. They are technically equivalent but, the syntax is different. Beijing use one, Tokyo the other.

Tokyo1

```
Tokyo1#wr ter
router bgp 64520
 bgp always-compare-med
 network 202.245.39.0
 network 202.245.36.240 mask 255.255.255.240
 neighbor 193.105.178.29 remote-as 64512
 neighbor 193.105.178.29 route-map melbourne-policy-out out
 neighbor 193.105.178.34 remote-as 64519
 neighbor 193.105.180.6 remote-as 64768
 neighbor 193.105.180.6 route-map hong-kong-policy-in in
 neighbor 193.105.180.6 route-map hong-kong-policy-out out
 neighbor 193.105.180.22 remote-as 64770
 neighbor 193.105.180.22 route-map seoul-policy-in in
 neighbor 193.105.180.22 route-map seoul-policy-out out
!
ip classless
ip as-path access-list 1 permit ^$
ip as-path access-list 2 deny 64768$
ip as-path access-list 2 deny 64770$
ip as-path access-list 2 permit .*
route-map melbourne-policy-out permit 10
 match as-path 2
!
route-map seoul-policy-out permit 10
 match as-path 1
!
route-map hong-kong-policy-in permit 10
 set metric 100
!
route-map hong-kong-policy-out permit 10
 match as-path 1
!
route-map seoul-policy-in permit 10
 set metric 200
```

Beijing1:

Instead of neighbor 193.105.180.22 route-map seoul-policy-out out, it is possible to use neighbor 193.105.180.22 filter-list 1 out. This is completly equivalent to Tokyo route-map. The route-map command

---

[3] Due to IOS software richness, this is often the case ! The problem is therefore to choose the “good” way…

includes more possibilities. If only, the filter-list is needed is might be more clear to use this syntax, but it is a matter of choice !

As already seen in previous step, the "set metric"commands are used to control which route will be the preferred backup.
In fact, Seoul and Hong-Kong must only know their closed neighbours, Tokyo and Beijing. Therefore, updates sent by Tokyo or Beijing to these sites must only concern the originator (Tokyo or Beijing). The "ip as-path access-list 1 permit ^$" is the way to control updates. Within BGP updates, the routers exchange route information and the complete AS path used on the way. Each router append in front of as-path list the AS number of the sending neighbour. Local routes (internal routes) are ,therefore , sent with an empty as-path list. The ^$ is an empty access-path list. These are classical regular expression came from Unix world.
Updates sent to Melbourne will concern all as-path (the last permit .*) except those ending by 64768 (deny 64768$) and 64770 (deny 64770$).

Melbourne:
```
Melbourne#sh ip bgp
BGP table version is 31, local router ID is 193.105.178.29
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*> 10.1.1.0/24      0.0.0.0                0         32768 i
*> 172.16.2.0/24    193.105.178.30                       0 64520 64519 i
*> 192.168.33.128/25 193.105.178.30                       0 64520 64519 64515 i
*> 202.245.36.240/28 193.105.178.30         1             0 64520 i
*> 202.245.39.0     193.105.178.30         0             0 64520 i
```

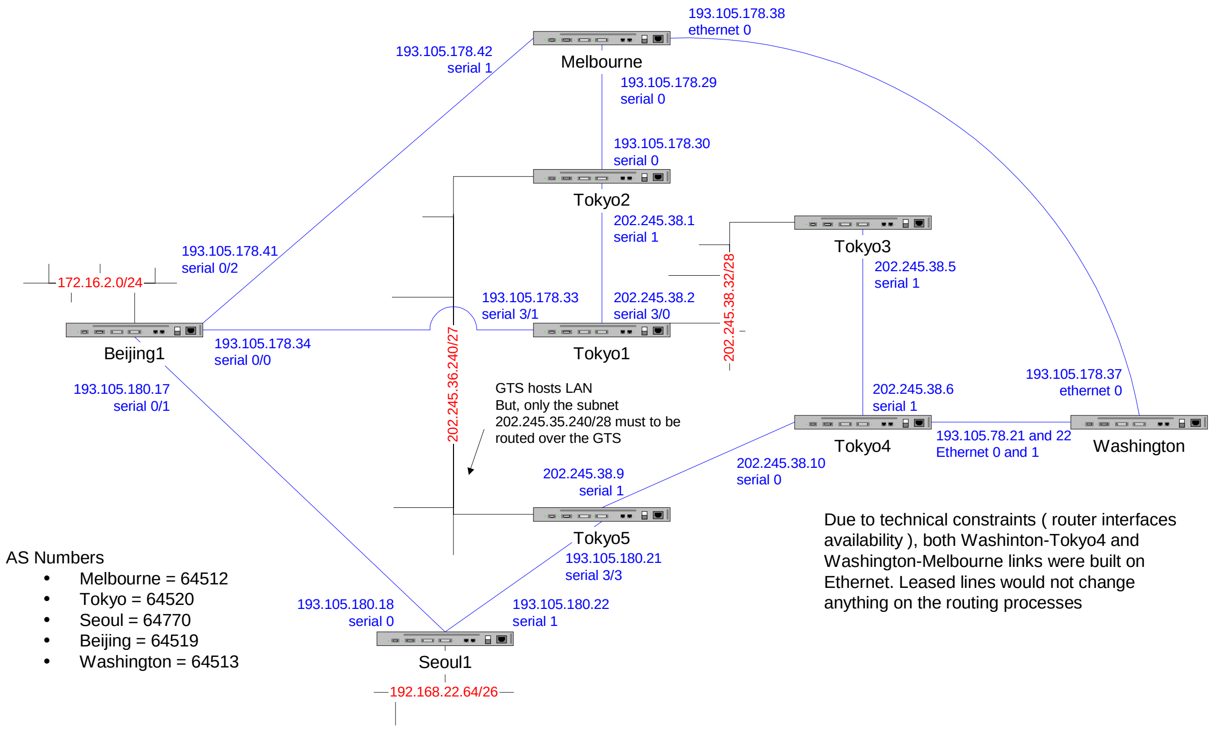
Seoul:
```
Seoul1#sh ip bgp
BGP table version is 312, local router ID is 193.105.180.22
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*> 172.16.2.0/24    193.105.180.17        74             0 64519 i
*> 192.168.22.0/26  0.0.0.0                0         32768 i
*> 202.245.36.240/28 193.105.180.21         1             0 64520 i
*> 202.245.39.0     193.105.180.21         0             0 64520 i
Seoul1#sh ip bgp 202.245.36.240
BGP routing table entry for 202.245.36.240/28, version 309
Paths: (1 available, best #1, advertised over EBGP)
  64520
    193.105.180.21 from 193.105.180.21 (202.245.39.254)
      Origin IGP, metric 1, localpref 100, valid, external, best
```

Phase 2 : Multiple routers running BGP within one AS

In the previous section we show various elements for managing BGP and redistribution between AS. In the second test bed, we focus on the so called IBGP. In our config, Tokyo's connections are established on multiple routers. Therefore, Tokyo's routers must do BGP between them in order to exchange routes with the external AS.
Due to the protocol BGP, in this case, you MUST established a full meshed between all internal BGP routers. As seen above, BGP rely on TCP connexions to work. The full meshed configuration mean that it is mandatory to have TCP connection between all peers in the network. For example, with 5 internal routers with connections out of the AS, the network manager must configure 10 TCP sessions. With a sixth router, six more connection. The nightmare is not that far !
Solutions are available in order to minimize the number of TCP peers to create. The two available solutions are called BGP confederation and route reflectors. The two solutions avoid the full mesh between IBGP.
In the configurations shown below only the second one was tested. The result would be the same with BGP confederation.

BGP test bed (phase 2)



Before going to detailed configuration, a last important point. In the first phase, we do not care about router ID, and IP source address for BGP updates. In most cases, the router ID is not a worthy information, and, BGP update source address is in general the IP address of the link connected with the neighbour. In IBGP, this is not a desirable solution. To avoid any trouble with internal failure, it is recommended to create loopback interfaces on the routers (the IP address will be the router ID) and to use this address as the BGP source address. This give a much more robust configuration. This IP address is used only locally for TCP connection for BGP protocol. This could be privates addresses if wanted.

In the configuration above, Tokyo3, with no outgoing connection do not run BGP. We have decided that Tokyo5 will be the “master” in the route reflectors configuration. We haven’t got the opportunity to teste the situation with a backup “master”.

Tokyo5 :

```
Tokyo5#wr ter
!
interface Loopback0
 ip address 192.168.39.5 255.255.255.255
!
interface Ethernet0
 ip address 202.245.36.253 255.255.255.0
!
interface Serial0
 ip address 202.245.38.9 255.255.255.252
 encapsulation ppp
 no fair-queue
!
interface Serial1
 description Link Tokyo-Seoul
 ip address 193.105.180.21 255.255.255.252
 encapsulation ppp
!
router rip
 version 2
 redistribute connected
 redistribute static
 redistribute bgp 64520 metric 1
 network 202.245.36.0
 network 202.245.38.0
 network 192.168.39.0
 no auto-summary
!
router bgp 64520
 network 202.245.36.240 mask 255.255.255.240
 neighbor 192.168.39.1 remote-as 64520
 neighbor 192.168.39.1 route-reflector-client
 neighbor 192.168.39.1 update-source Loopback0
```

```
 neighbor 192.168.39.2 remote-as 64520
 neighbor 192.168.39.2 route-reflector-client
 neighbor 192.168.39.2 update-source Loopback0
 neighbor 192.168.39.4 remote-as 64520
 neighbor 192.168.39.4 route-reflector-client
 neighbor 192.168.39.4 update-source Loopback0
 neighbor 193.105.180.22 remote-as 64770
 neighbor 193.105.180.22 route-map seoul-policy-in in
!
no ip classless
ip route 202.245.36.240 255.255.255.240 Ethernet0
route-map seoul-policy-in permit 10
 set metric 20
!
!
!
line con 0
line aux 0
line vty 0 4
 no login
!
end

Tokyo5#sh ip rou
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route

Gateway of last resort is not set

     10.0.0.0/8 is subnetted, 1 subnets
R       10.1.1.0 [120/1] via 202.245.36.254, 00:00:09, Ethernet0
R    202.245.39.0/24 [120/2] via 202.245.36.254, 00:00:09, Ethernet0
     202.245.38.0/24 is variably subnetted, 9 subnets, 3 masks
C       202.245.38.8/30 is directly connected, Serial0
C       202.245.38.10/32 is directly connected, Serial0
R       202.245.38.1/32 [120/3] via 202.245.38.10, 00:00:09, Serial0
R       202.245.38.0/30 [120/1] via 202.245.36.254, 00:00:09, Ethernet0
R       202.245.38.2/32 [120/1] via 202.245.36.254, 00:00:09, Ethernet0
R       202.245.38.5/32 [120/1] via 202.245.38.10, 00:00:09, Serial0
R       202.245.38.4/30 [120/1] via 202.245.38.10, 00:00:09, Serial0
R       202.245.38.6/32 [120/3] via 202.245.36.254, 00:00:09, Ethernet0
R       202.245.38.32/28 [120/2] via 202.245.38.10, 00:00:09, Serial0
                         [120/2] via 202.245.36.254, 00:00:09, Ethernet0
     193.105.178.0/24 is variably subnetted, 5 subnets, 2 masks
R       193.105.178.28/30 [120/1] via 202.245.36.254, 00:00:01, Ethernet0
R       193.105.178.29/32 [120/1] via 202.245.36.254, 00:00:01, Ethernet0
R       193.105.178.20/30 [120/1] via 202.245.38.10, 00:00:17, Serial0
R       193.105.178.34/32 [120/2] via 202.245.36.254, 00:00:01, Ethernet0
R       193.105.178.32/30 [120/2] via 202.245.36.254, 00:00:01, Ethernet0
     202.245.36.0/24 is variably subnetted, 2 subnets, 2 masks
C       202.245.36.0/24 is directly connected, Ethernet0
S       202.245.36.240/28 is directly connected, Ethernet0
     193.105.180.0/24 is variably subnetted, 2 subnets, 2 masks
C       193.105.180.20/30 is directly connected, Serial1
C       193.105.180.22/32 is directly connected, Serial1
R    192.168.44.0/24 [120/1] via 202.245.38.10, 00:00:17, Serial0
     192.168.39.0/24 is subnetted, 4 subnets
C       192.168.39.5 is directly connected, Loopback0
R       192.168.39.4 [120/1] via 202.245.38.10, 00:00:17, Serial0
R       192.168.39.2 [120/1] via 202.245.36.254, 00:00:01, Ethernet0
R       192.168.39.1 [120/2] via 202.245.36.254, 00:00:01, Ethernet0
     192.168.22.0/24 is subnetted, 1 subnets
B       192.168.22.0 [20/20] via 193.105.180.22, 00:16:11
     172.16.0.0/16 is subnetted, 1 subnets
R       172.16.2.0 [120/1] via 202.245.36.254, 00:00:01, Ethernet0
Tokyo5#sh ip bgp
BGP table version is 54, local router ID is 192.168.39.5
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*>i10.1.1.0/24      193.105.178.29         0    100      0 64512 i
* i                 193.105.178.34        10    100      0 64519 64512 i
* i172.16.2.0/24    193.105.178.29              100      0 64512 64519 i
*                   193.105.180.22        20             0 64770 64519 i
*>i                 193.105.178.34        10    100      0 64519 i
*> 192.168.22.0/26  193.105.180.22        20             0 64770 i
*>i192.168.44.0     193.105.178.21         0    100      0 64513 i
* i202.245.36.240/28 192.168.39.2          0    100      0 i
*>                  0.0.0.0                0         32768 i
*>i202.245.39.0     192.168.39.1           0    100      0 i
Tokyo5#sh ip bgp 172.16.2.0
BGP routing table entry for 172.16.2.0/24, version 50
Paths: (3 available, best #3, advertised over IBGP, EBGP)
```

```
  64512 64519, (Received from a RR-client)
    193.105.178.29 (metric 1) from 192.168.39.2
      Origin IGP, localpref 100, valid, internal, synchronized
  64770 64519
    193.105.180.22 from 193.105.180.22
      Origin IGP, metric 20, valid, external
  64519, (Received from a RR-client)
    193.105.178.34 (metric 2) from 192.168.39.1 (202.245.39.254)
      Origin IGP, metric 10, localpref 100, valid, internal, synchronized, best
Tokyo5#sh ip bgp 192.168.44.0
BGP routing table entry for 192.168.44.0/24, version 54
Paths: (1 available, best #1, advertised over IBGP, EBGP)
  64513, (Received from a RR-client)
    193.105.178.21 (metric 1) from 192.168.39.4
      Origin IGP, metric 0, localpref 100, valid, internal, synchronized, best
```

Tokyo used RIPv2. It is very important to note that the network of the loopback addresses is 255.255.255.255 subnetted and that is belongs to the RIP process. Every router received BGP update with the loopback address (neighbor 192.168.39.4 update-source Loopback0) and the neighbour address is in the routing table therefore the BGP route is put in the routing table !

Tokyo4

```
Tokyo4#wr ter
!
interface Loopback0
 ip address 192.168.39.4 255.255.255.255
!
interface Ethernet0
 description link tokyo4 - washington
 ip address 193.105.178.22 255.255.255.252
!
interface Serial0
 description link tokyo 4 - 5
 ip address 202.245.38.10 255.255.255.252
 encapsulation ppp
 clockrate 64000
!
interface Serial1
 description link tokyo 4 - 3
 ip address 202.245.38.6 255.255.255.252
 encapsulation ppp
!
interface BRI0
 no ip address
 shutdown
!
router rip
 version 2
 redistribute connected
 redistribute bgp 64520 metric 1
 network 202.245.38.0
 network 192.168.39.0
 no auto-summary
!
router bgp 64520
 neighbor 192.168.39.5 remote-as 64520
 neighbor 192.168.39.5 update-source Loopback0
 neighbor 193.105.178.21 remote-as 64513
!
no ip classless
!
line con 0
line aux 0
line vty 0 4
 no login
!
end

Tokyo4#sh ip rou
Codes: C - connected, S - static, I - IGRP, 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, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate default
       U - per-user static route, o - ODR

Gateway of last resort is not set

     10.0.0.0/24 is subnetted, 1 subnets
R       10.1.1.0 [120/2] via 202.245.38.5, 00:00:13, Serial1
                 [120/2] via 202.245.38.9, 00:00:21, Serial0
R    202.245.39.0/24 [120/2] via 202.245.38.5, 00:00:13, Serial1
     202.245.38.0/24 is variably subnetted, 8 subnets, 3 masks
C       202.245.38.9/32 is directly connected, Serial0
```

```
C       202.245.38.8/30 is directly connected, Serial0
R       202.245.38.1/32 [120/2] via 202.245.38.5, 00:00:13, Serial1
R       202.245.38.0/30 [120/2] via 202.245.38.9, 00:00:21, Serial0
                        [120/2] via 202.245.38.5, 00:00:13, Serial1
R       202.245.38.2/32 [120/2] via 202.245.38.9, 00:00:21, Serial0
C       202.245.38.5/32 is directly connected, Serial1
C       202.245.38.4/30 is directly connected, Serial1
R       202.245.38.32/28 [120/1] via 202.245.38.5, 00:00:13, Serial1
     193.105.178.0/24 is variably subnetted, 5 subnets, 2 masks
R       193.105.178.28/30 [120/2] via 202.245.38.9, 00:00:24, Serial0
R       193.105.178.29/32 [120/2] via 202.245.38.9, 00:00:24, Serial0
C       193.105.178.20/30 is directly connected, Ethernet0
R       193.105.178.34/32 [120/2] via 202.245.38.5, 00:00:15, Serial1
R       193.105.178.32/30 [120/2] via 202.245.38.5, 00:00:15, Serial1
     202.245.36.0/24 is variably subnetted, 2 subnets, 2 masks
R       202.245.36.0/24 [120/1] via 202.245.38.9, 00:00:24, Serial0
R       202.245.36.240/28 [120/1] via 202.245.38.9, 00:00:24, Serial0
     193.105.180.0/24 is variably subnetted, 2 subnets, 2 masks
R       193.105.180.20/30 [120/1] via 202.245.38.9, 00:00:24, Serial0
R       193.105.180.22/32 [120/1] via 202.245.38.9, 00:00:24, Serial0
B    192.168.44.0/24 [20/0] via 193.105.178.21, 00:03:32
     192.168.39.0/32 is subnetted, 4 subnets
R       192.168.39.5 [120/1] via 202.245.38.9, 00:00:24, Serial0
C       192.168.39.4 is directly connected, Loopback0
R       192.168.39.2 [120/2] via 202.245.38.9, 00:00:24, Serial0
R       192.168.39.1 [120/2] via 202.245.38.5, 00:00:16, Serial1
     172.16.0.0/24 is subnetted, 1 subnets
R       172.16.2.0 [120/2] via 202.245.38.5, 00:00:16, Serial1
                   [120/2] via 202.245.38.9, 00:00:24, Serial0

Tokyo4#sh ip bgp
BGP table version is 8, local router ID is 192.168.39.4
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*  10.1.1.0/24      193.105.178.21                       0 64513 64512 i
*>i                 193.105.178.29         0    100      0 64512 i
*  172.16.2.0/24    193.105.178.21                       0 64513 64512 64519 i
*>i                 193.105.178.34        10    100      0 64519 i
*  192.168.22.0/26  193.105.178.21                       0 64513 64512 64519 64770 i
*>i                 193.105.180.22        20    100      0 64770 i
*> 192.168.44.0     193.105.178.21         0             0 64513 i
*>i202.245.36.240/28 192.168.39.5           0    100      0 i
*>i202.245.39.0     192.168.39.1           0    100      0 i

Tokyo4#sh ip bgp summ
BGP table version is 8, main routing table version 8
6 network entries (9/18 paths) using 1272 bytes of memory
9 BGP path attribute entries using 1092 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.39.5    4 64520     198     168        8    0    0 00:04:13        5
193.105.178.21  4 64513     672     676        8    0    0 00:04:06        4

Tokyo4#sh ip bgp 172.16.2.0
BGP routing table entry for 172.16.2.0/24, version 4
Paths: (2 available, best #2)
  64513 64512 64519
    193.105.178.21 from 193.105.178.21 (193.105.178.37)
      Origin IGP, localpref 100, valid, external
  64519
    193.105.178.34 (metric 2) from 192.168.39.5 (202.245.39.254)
      Origin IGP, metric 10, localpref 100, valid, internal, synchronized, best
      Originator : 202.245.39.254, Cluster list: 192.168.39.5
Tokyo4#sh ip bgp 192.168.44.0
BGP routing table entry for 192.168.44.0/24, version 8
Paths: (1 available, best #1, advertised over IBGP)
  64513
    193.105.178.21 from 193.105.178.21 (193.105.178.37)
      Origin IGP, metric 0, localpref 100, valid, external, best
```

The “sh ip bgp 172.16.2.0” above give very interesting information. The preferred route to reach beijing is the direct link between Beijing and Tokyo, which is normal. But, quite surprisingly, there is only one backup solution via Washington and Melbourne. Therefore, for Tokyo4 if the link with Beijing1 is down, Tokyo4 will try to reach Beijing via Washington. This is due to a precedence of External BGP routes over Internal BGP routes. Of course, with metric and/or local preference it is possible to change this default behaviour.
For example, here are briefs outputs for Tokyo2 regarding Seoul.

Tokyo2 (with link Tokyo5-Seoul working)

```
Tokyo2#sh ip bgp 192.168.22.0
```

```
BGP routing table entry for 192.168.22.0/26, version 107
Paths: (2 available, best #1, advertised over EBGP)
  64770
    193.105.180.22 (metric 1) from 192.168.39.5
      Origin IGP, metric 20, localpref 100, valid, internal, synchronized, best
      Originator : 192.168.39.5, Cluster list: 192.168.39.5
  64512 64519 64770
    193.105.178.29 from 193.105.178.29 (10.1.1.254)
      Origin IGP, localpref 100, valid, external
```

Tokyo2 (with link Tokyo5-Seoul shutdown)

```
Tokyo2#sh ip bgp 192.168.22.0
BGP routing table entry for 192.168.22.0/26, version 5
Paths: (2 available, best #2)
  64512 64519 64770
    193.105.178.29 from 193.105.178.29 (10.1.1.254)
      Origin IGP, localpref 100, valid, external
  64519 64770
    193.105.178.34 (metric 1) from 192.168.39.5 (202.245.39.254)
      Origin IGP, metric 10, localpref 100, valid, internal, synchronized, best
      Originator : 202.245.39.254, Cluster list: 192.168.39.5
```

The “best” route for Tokyo to Beijing, for Tokyo2, is through Tokyo1 and Beijing, which is the “human” preferred solution. The “metric 10” is sent via Tokyo1 (connected to Beijing) with a route map policy on Beijing received updates.