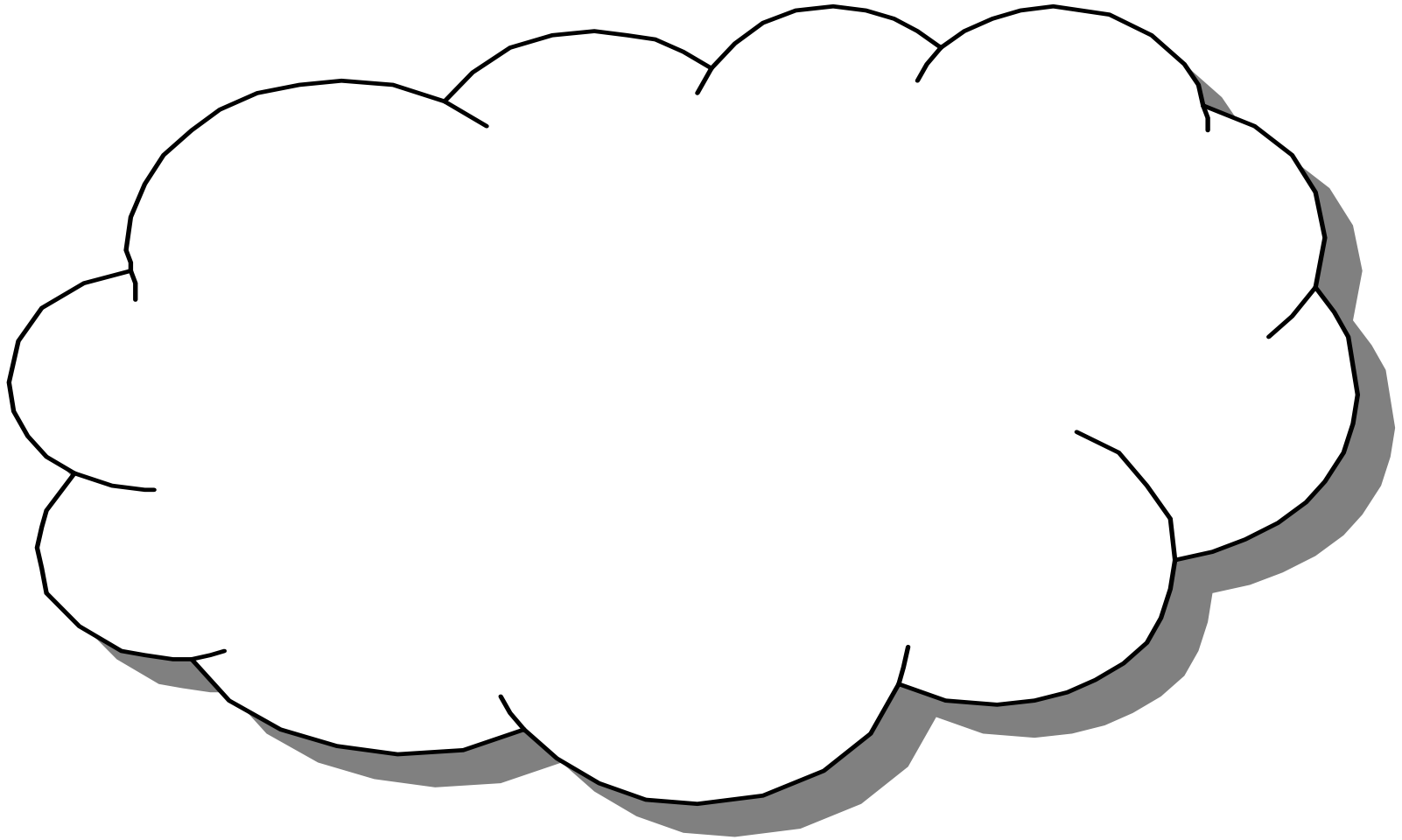


Internet Operation

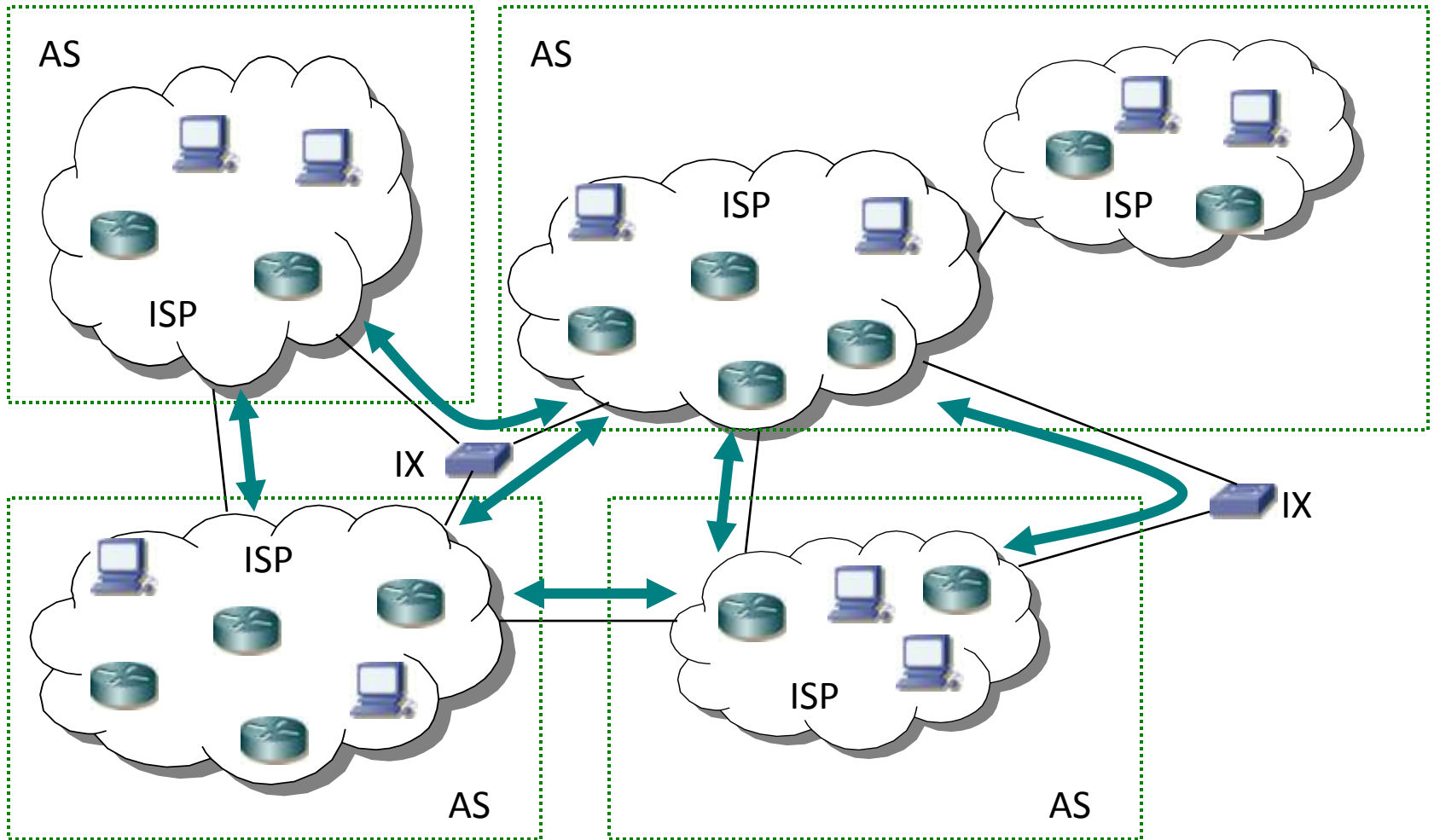
Matsuzaki 'maz' Yoshinobu

<maz@iij.ad.jp>

Internet



Internet



philosophy

Autonomous

Distributed

Cooperative

as a part of the internet

- Your good services
- Your good operations
- Good internet experience

- We are changing our network as the internet is changing
 - technologies
 - operations

strong operation

- Can you persuasively explain your operation?
 - policy
 - design
 - implementation
 - configuration
 - management

IIJ case – IPv6

- pure IP network
 - now dual stacked
 - no MPLS
- IIJ/AS2497 maintains its backbone in Japan and U.S..

- And we have deployed IPv6

History of IJ's IPv6 service

- 1999
 - IPv6 over IPv4 tunnel service (experimental)
- 2000
 - IPv6 native service (leased line, experimental)
- 2001
 - IPv4/IPv6 dual stack service (leased line)
 - IPv6 data center service
 - IPv6 over IPv4 tunnel service for consumers

History of IJ's IPv6 backbone

- Initially we started with dedicated backbone
- 1998
 - PC based router(kame stack)
 - tunnel and ethernet
- 2000
 - cisco c72xx series
 - tunnel, ethernet and T1 line
- 2005
 - we started to migrate to IPv4/IPv6 dual stack backbone

addressing

- /128 for loopback interfaces
- /64 for links
 - /127 is used on several inter-router links
- /48 for customer sites
 - still considering the size
 - possible sizes are: /48, /52, /56, /60, /64

global unicast address

- Inter-router link does not require global address inside AS
 - OSPFv3 uses link-local address to exchange LSAs
 - only loopback interface needs to configure global address
- But we configure global address on every interface
 - as ping destination to check availability

link-local unicast address

- fe80::/64
- AS IS
 - We don't touch
 - Most routers use Modified EUI-64 format address
- A virtual address for vrrp/hsrp is another story.
 - Customers might configure a static route at their equipments with this address, so the address should be assigned statically like fe80::1.

routing protocols

IPv4

- OSPFv2
 - mostly area 0
 - md5 authentication
- BGP4
 - peer through ipv4
 - route-reflector
 - md5 authentication

IPv6

- OSPFv3
 - area 0 only
 - ipsec authentication
- BGP4+
 - peer through ipv6 global
 - route-reflector (same as IPv4)
 - md5 authentication

router ID

- Routing protocols usually require 32bit ID
 - even a routing protocol for IPv6
 - We use IPv4 address of loopback interface as its router ID
- Every routers has IPv4 address in our network

OSPFv3 link cost

- We set the same link cost value as IPv4's.
 - The network topology is almost same.
 - working fine 😊
- When we were using RIPng as IGP (we had no choice at that time 😞), these were so much trouble.

management

- remote access to a router (ssh/telnet)
 - IPv4/IPv6 dual stack
 - out-of-band port has IPv4 address only
- other services
 - IPv4 only
 - AAA, snmp, syslog, ntp, flow export

availability check

- checking by ping
 - both IPv4 and IPv6
 - dual stack routers receive ping twice as much

monitoring

- We have to know about our network. 😊
 - traffic volume and so on

- We are waiting for
 - MIB for IP [RFC4293]
 - NetFlow v9 [RFC3954] for IPv6

hostname rules

- 3 names
 - dual stack, ipv4 only, ipv6 only
 - Users can test reachability by themselves

www	IN A	210.130.137.80
	IN AAAA	2001:240:bb42:b000::1:80
www-v4	IN A	210.130.137.80
www-v6	IN AAAA	2001:240:bb42:b000::1:80

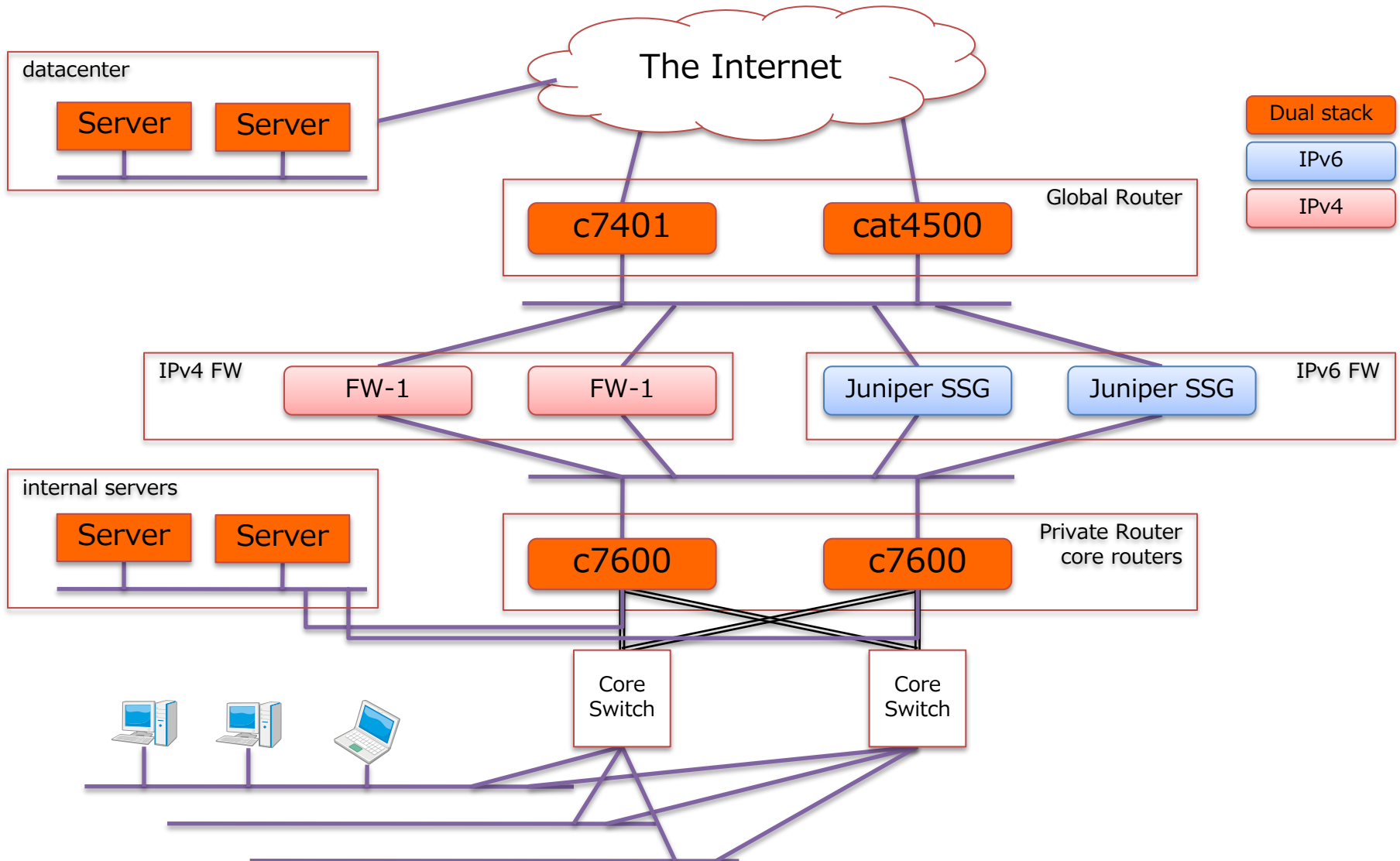
reverse DNS

- For our routers and servers
 - maintain PTR records as possible
 - the same policy as IPv4
- For static address users (such as /48)
 - delegate NS if they ask
- For dynamic address users
 - nothing

IPv6 experience: office network

- 1998
 - We installed IPv6 in network tech divisions
- 2000
 - All of our tech divisions got IPv6 enabled
- 2003
 - We moved to new office, so installed IPv6 to whole office network.
- 2008
 - We performed IPv6 renumbering

IIJ's office network

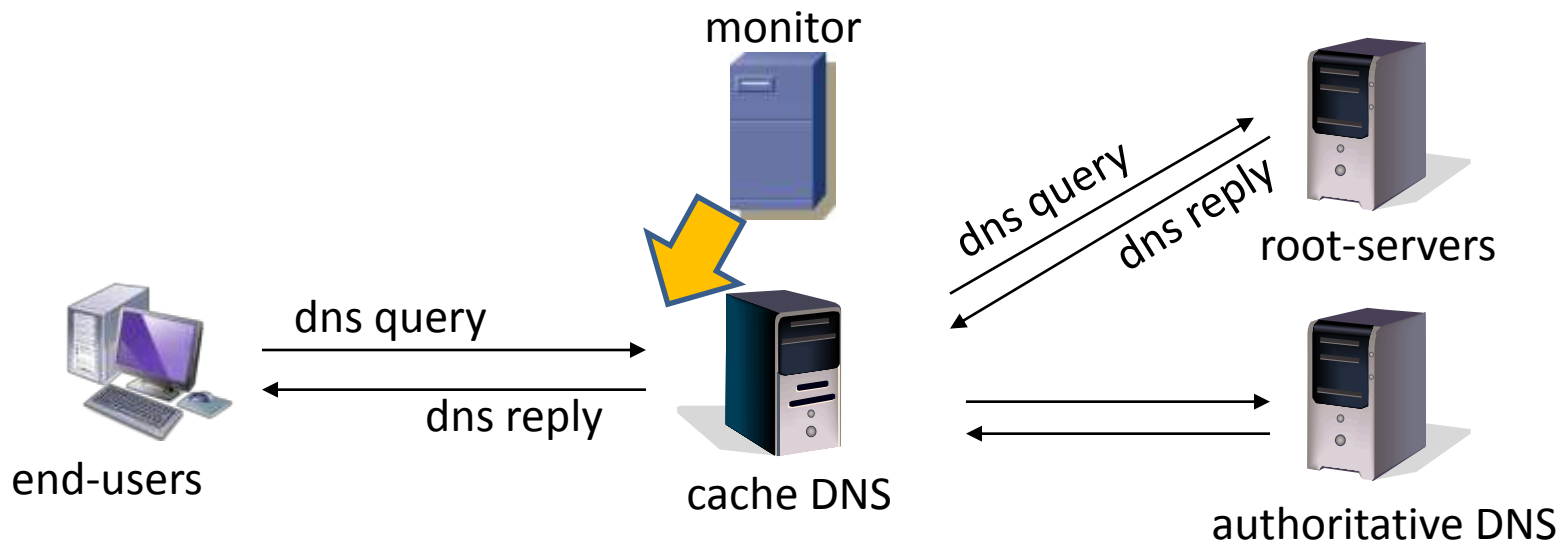


publication

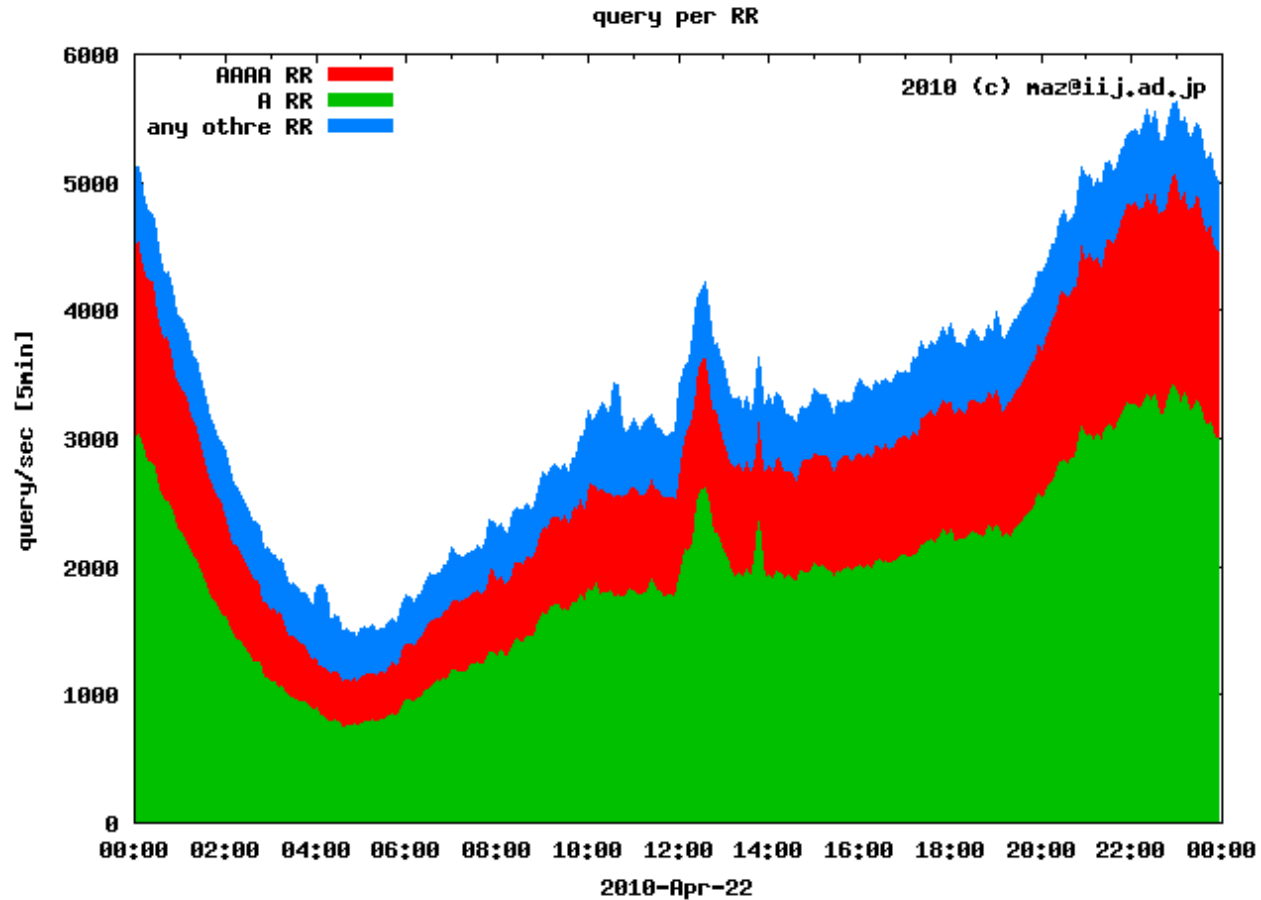
- IJ publishes IPv6 deployment status of its services on its www site.
 - <http://www.ij.ad.jp/service/IPv6schedule/>
 - what we did, what we will do
- This helps our customers to plan their IPv6 deployment.

end-user environments analysis

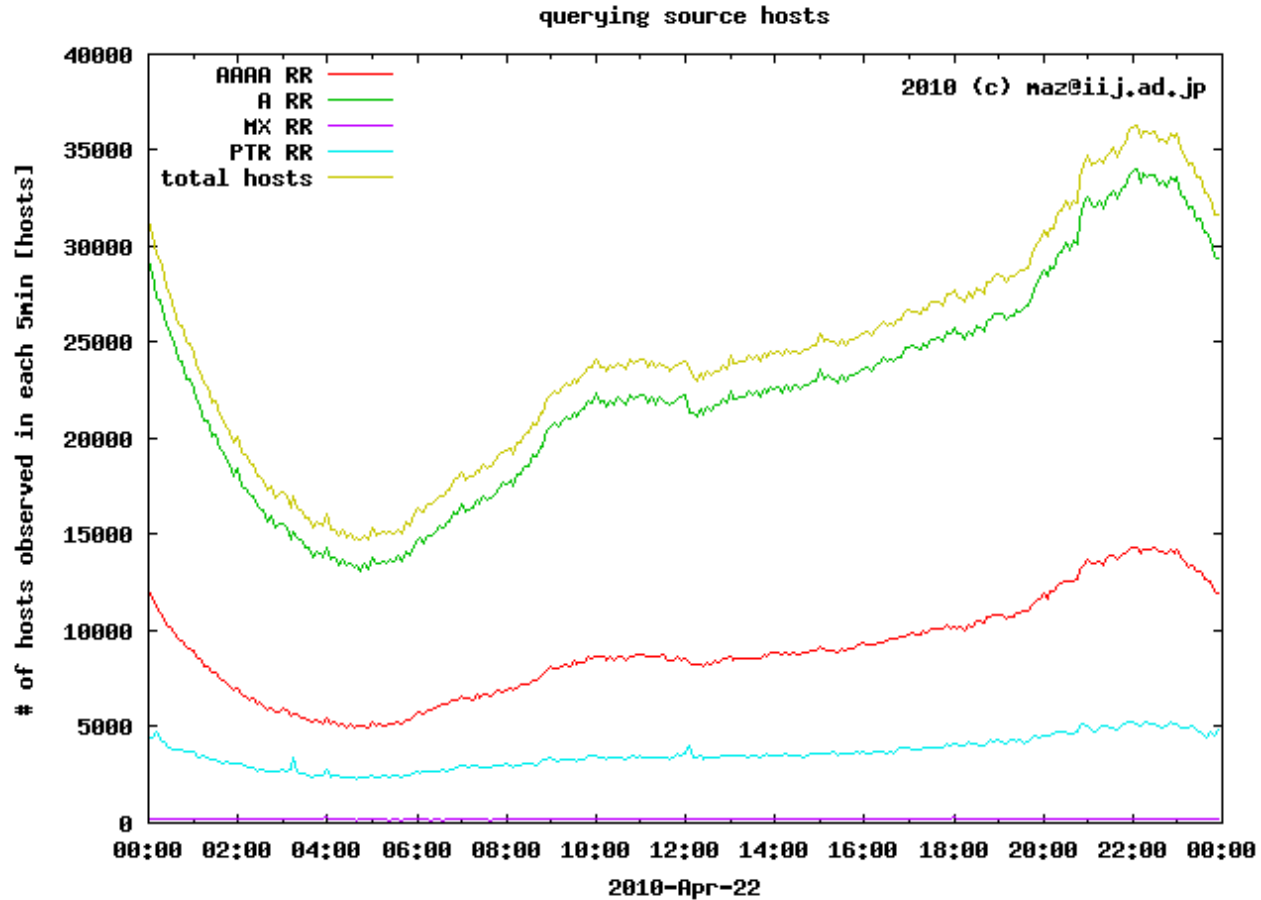
- We gathered data from our cache DNS
 - AAAA query rate



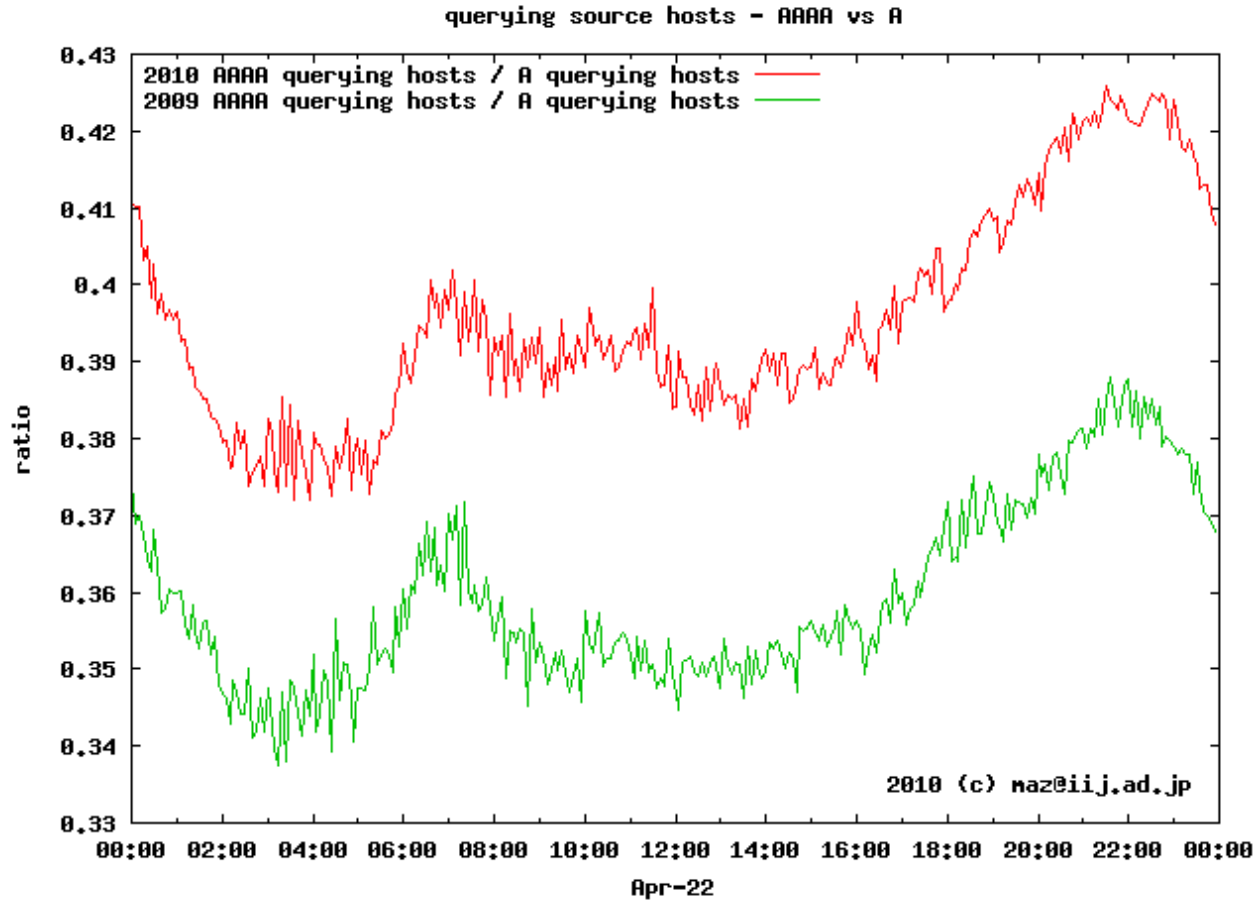
stacked query/sec graph



observed querying end-hosts



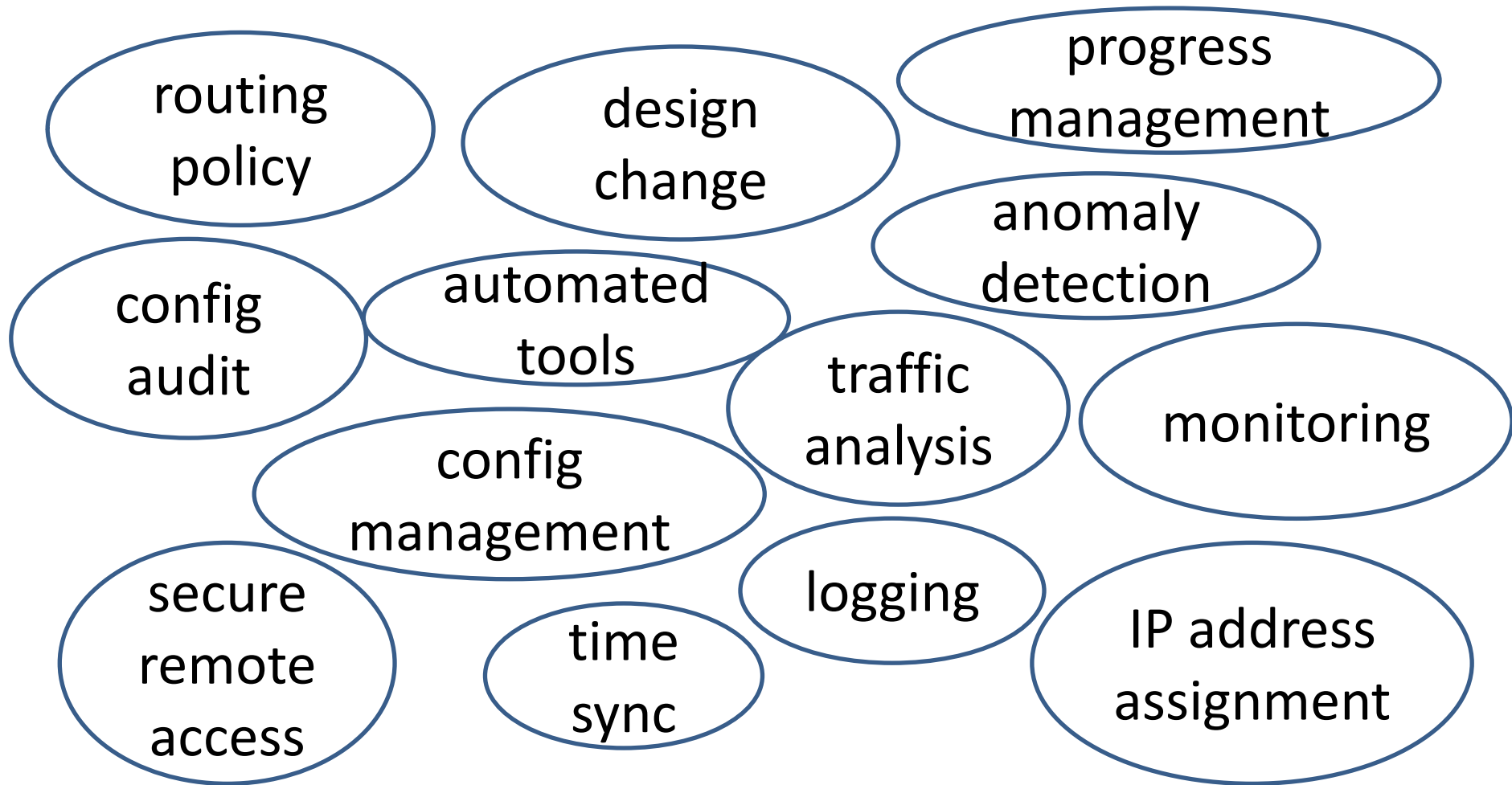
ratio of AAAA capable source



again, as a part of the internet

- Your good services
- Your good operations
- Good internet experience

- operation - environmental consideration



mistakes

- mistakes degrade our services
- In many case, its effects are negligible (hopefully), but sometimes even a small mistake could cause a disaster
- Let's look into our mistakes
 - frequent ones and major outages
- These cases are gathered from Japanese operational community. Thanks!

No.1 cut & paste of configuration

- One person was writing a router configuration by cutting & pasting from other routers’.
- but... forgot to modify IP address setting ☹️
- caused unwanted routing, IP address duplications

No.1 cut & paste of configuration

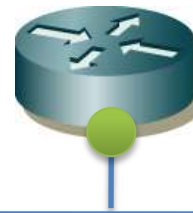
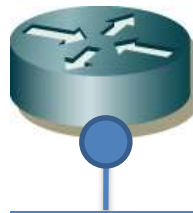
```
Interface Bundle-Ether1
ipv4 address 10.0.0.1 255.255.255.0
ipv6 address 2001:db8::1/64
bundle minimum-active links 2
!
```

cut&paste



```
Interface Bundle-Ether1
ipv4 address 10.0.0.1 255.255.255.0
ipv6 address 2001:db8::1/64
bundle minimum-active links 2
!
```

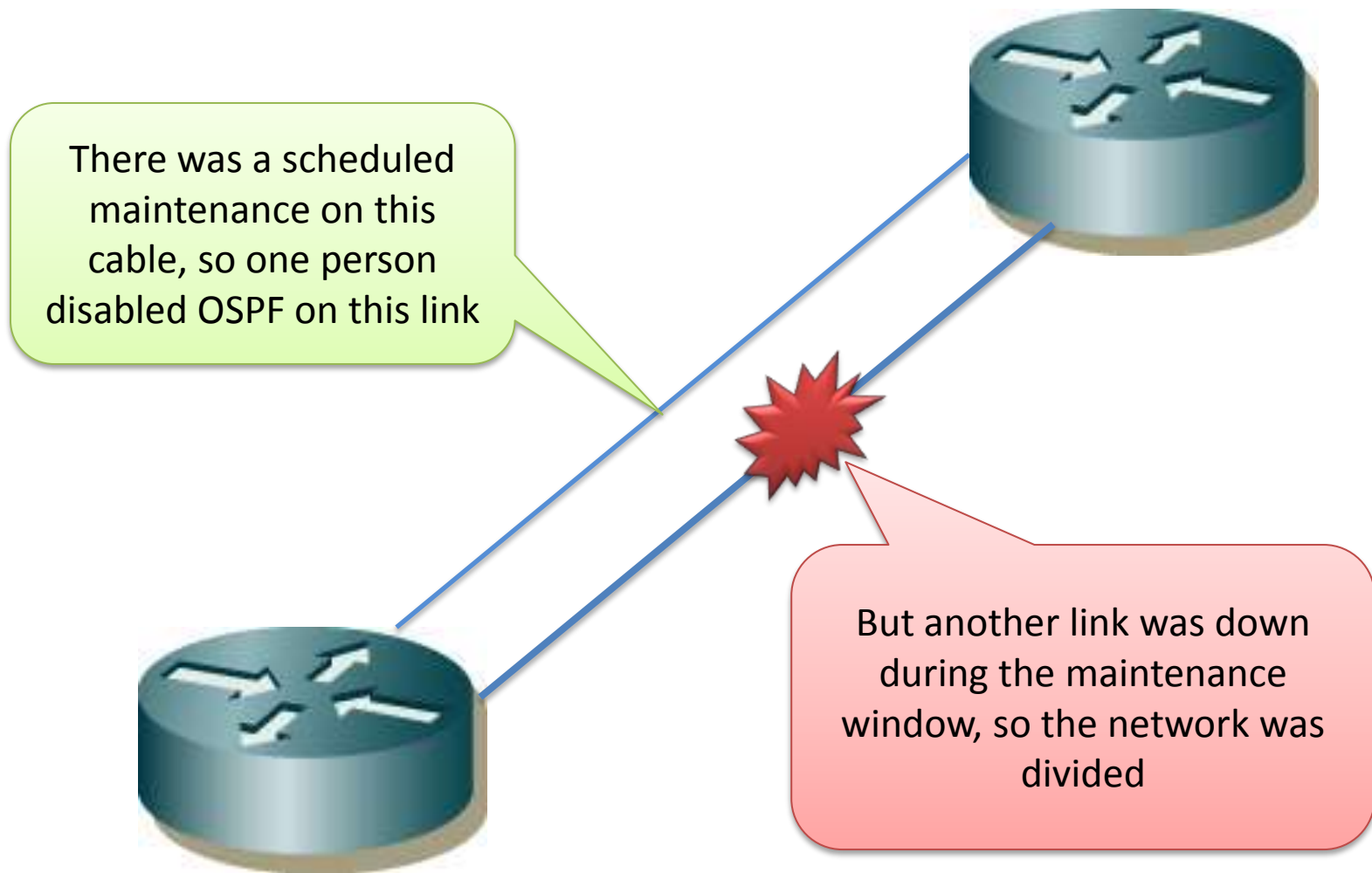
One person forget
to modify IP address



No.2 cable maintenance

- There was a scheduled maintenance at a long distance cable in a network, so one person incremented OSPF link costs on the link
- To avoid flapping on the link, the person disabled OSPF on the link ☹️
- During the maintenance window, another cable was down
- ... the network was divided

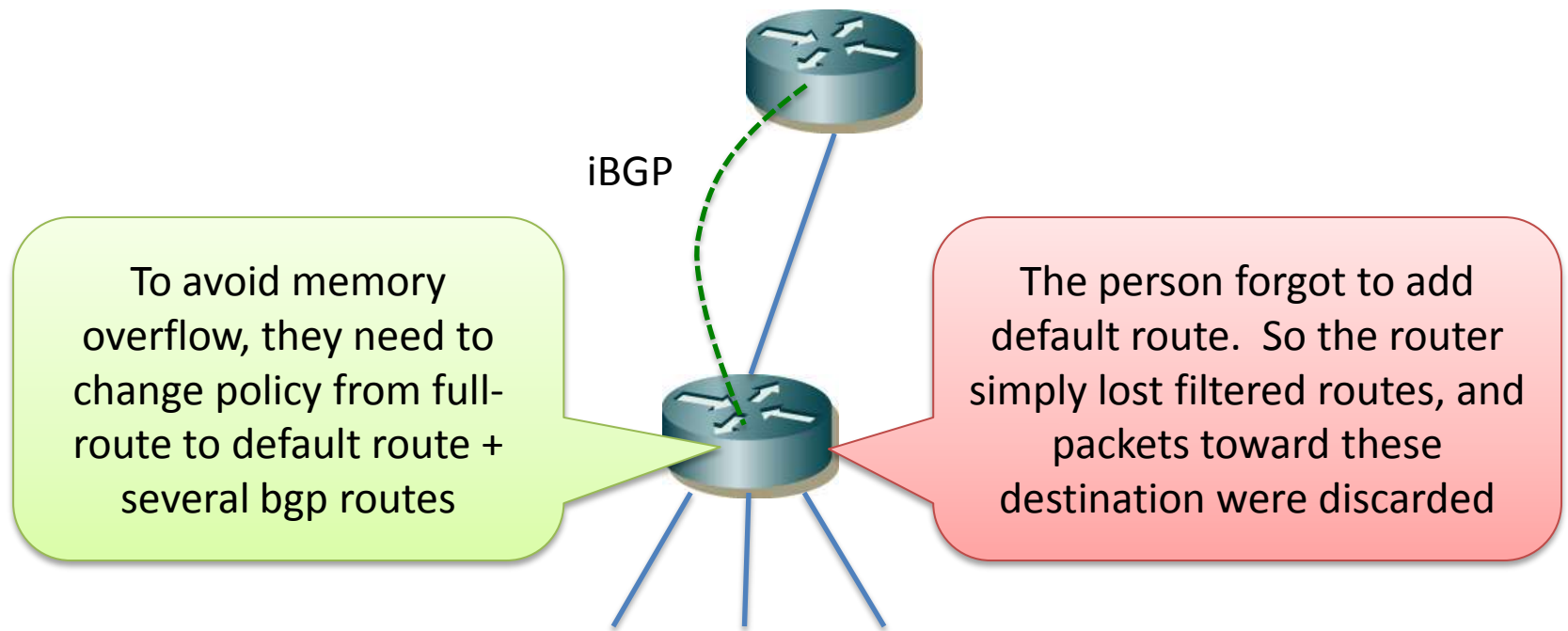
No.2 cable maintenance



No.3 BGP FIB reducing

- One person was asked to filter BGP routes on small routers to avoid memory overflow
- The person forgot to add default route that would cover filtered routes on the routes 😞
- The router lost routes, caused packets discard

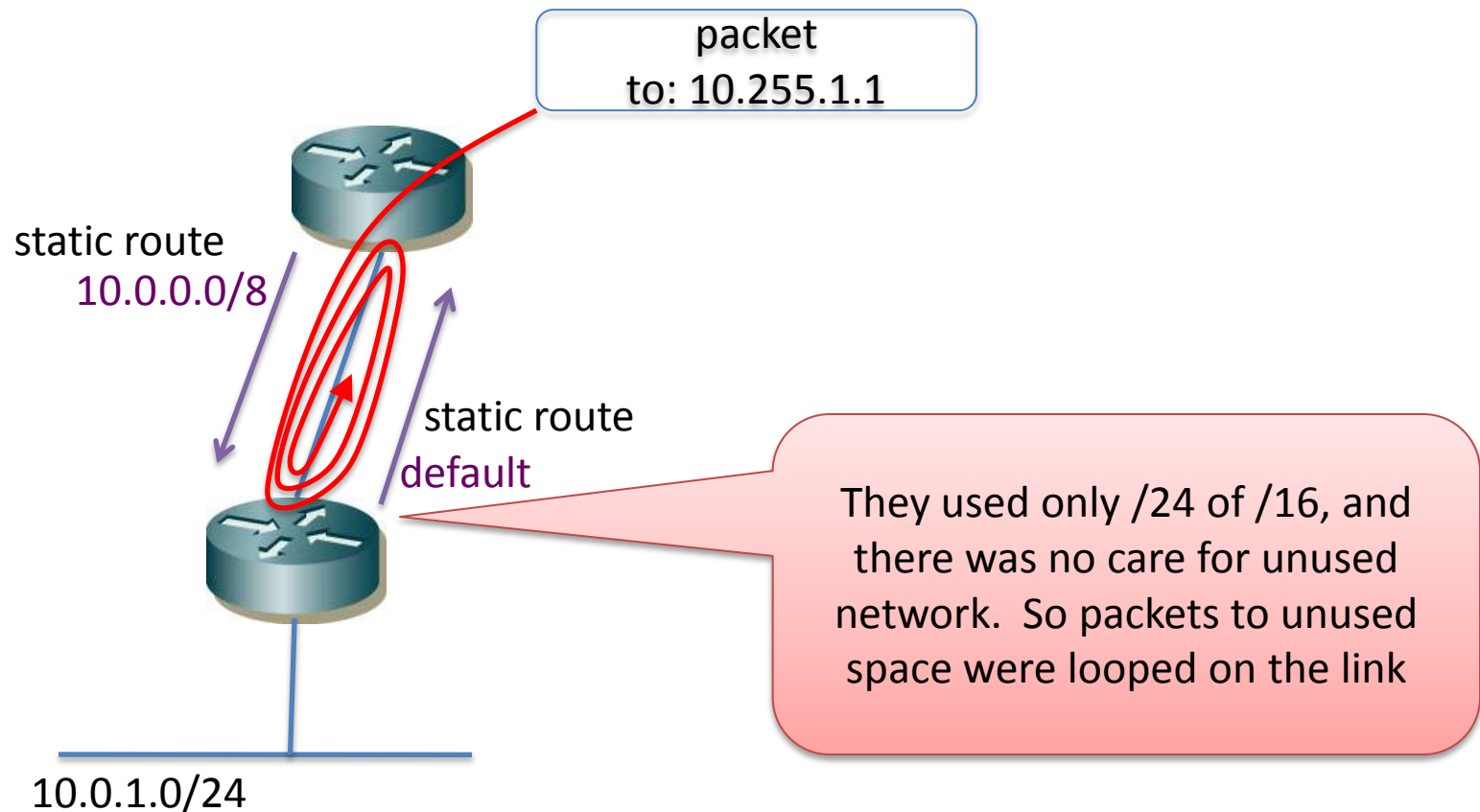
No.3 BGP FIB reducing



No.4 route termination

- One prefix was statically routed to a downstream router, and the downstream router had default route to upstream.
- But the downstream network used only a part of the prefix, and no care for unused space. ☹️
- Packets to unused space like portscan were looped between 2 routers.
 - they should add a null route to terminate route at the downstream router, or
 - they should adjust the routed prefix as needed

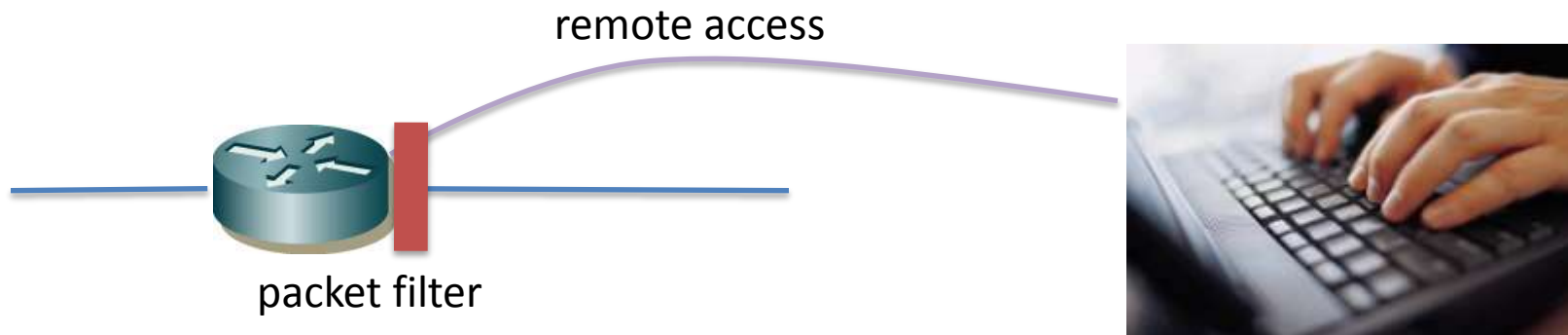
No.4 route termination



No.5 packet filter

- One person was asked to modify packet filter of a remote router
- The person forgot to permit management access 😞
- After modification, the person was also filtered out, and couldn't control anymore

No.5 packet filter



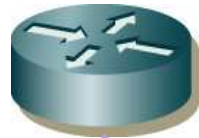
One person tried to modify packet filter remotely but the person forgot to permit management access, and lost his remote connection.

No.6 remote login

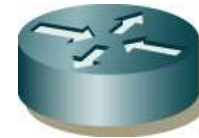
- One person was asked to change configuration on a certain router.
- The person mistyped the remote hostname, but it was still valid one, so the person changed configuration on a wrong router 😞
 - There was a similar case that typing in a wrong terminal among multi terminal windows.
- ..causes an unexpected routing

No.6 remote login

rt010bb101



rt011bb101



remote access

One person
was asked to
change
configuration
HERE!

The person mistyped the hostname,
but it was still valid one.
The person changed configuration
on a wrong router



No.7 configuration cleaning

- One person was asked to delete unused configuration of a router
- The person deleted line by line with leading 'no' keyword, and carelessly deleted a routing process. ☹️
- The router stopped the routing process as the command said so.....

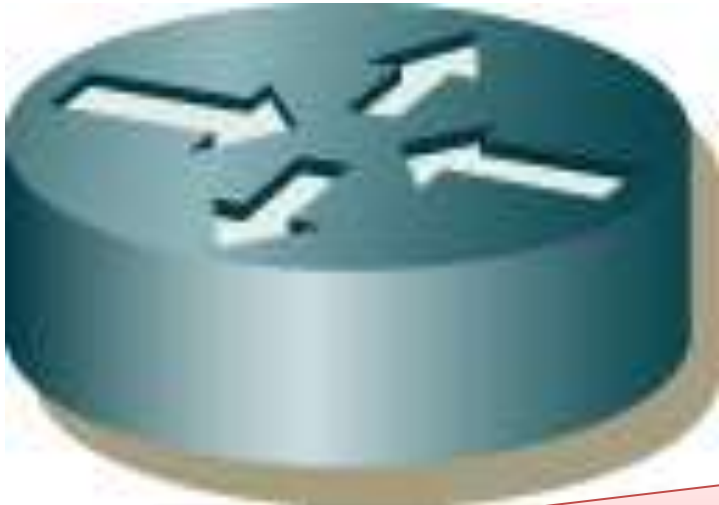
No.7 configuration cleaning

expected configuration

```
router ospf 65535  
no network 10.0.0.0 0.0.0.3 area 0
```

but typed in

```
no router ospf 65535  
no network 10.0.0.0 0.0.0.3 area 0
```



The routing process was stopped,
... caused a routing trouble

No.8 port move

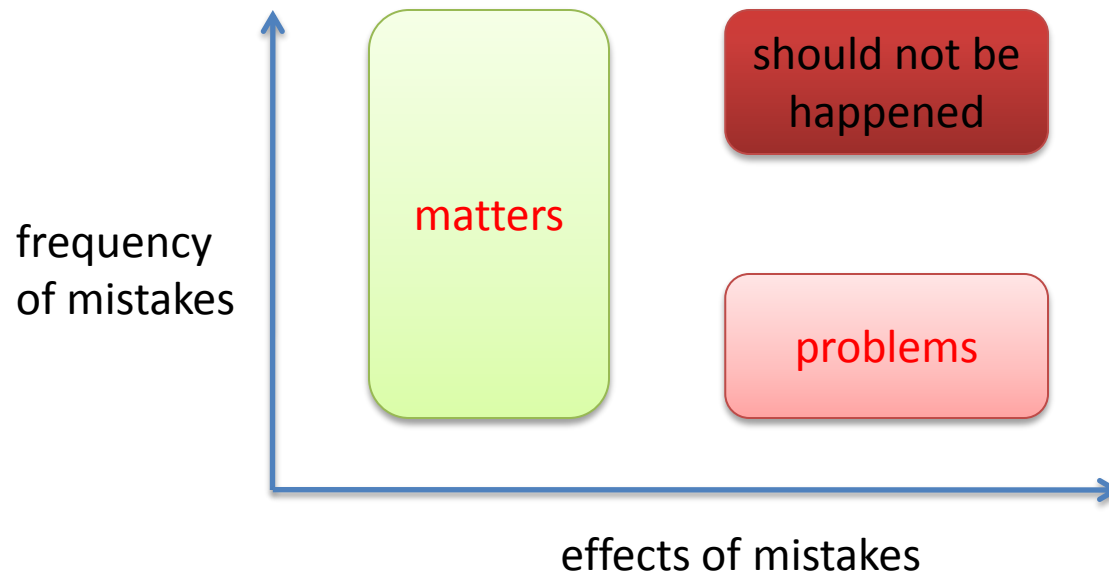
- One team performed a maintenance to move a cable from one switch to another
- One person shutdown a port, then another person unplug the cable checking the port number and the port LED.
- The person at the switch misunderstood the port number and port LED relationship, then unplugged an wrong cable 😞
- Unexpected network down

No.8 port move



react to mistakes

- minimize the effects
- reduce mistakes



minimize the effects of mistakes

- protection
 - straiten the extent of the impact
 - multiple route filters
- restoration
 - detect in early stage
 - monitor of configuration changes, traffic anomaly detection
 - notification properly
 - notify the operator of the mistake
 - recovery from mistake quickly
 - undo the modification, disconnect the wrong part

reduce mistakes

- In many cases, the main reason of mistakes is “careless”
 - more detailed classification would be possible
 - as you might know, to understand others (even yourself) is difficult

mistake, operators and attention

- OK, let's put an assumption
- An attention ability is limited, and it depends on operators and environments
 - If you have enough ability to complete an operation, there would be zero-mistake.

but according to our experiences

- We have made lots of mistakes
 - we don't have enough ability, or have abilities to mistake :P
- So we need a support to reduce mistakes

support for operation

- minimum hand-operation
 - automation
- clear procedure of operation
 - operation sheet
- support for attention
 - better user interface to help operations

again, as a part of the internet

- Your good services
- Your good operations
- Good internet experience

sharing

- These problems are not only for you
 - we have similar issues among the community
- Your experience is important for others
 - how did, problems, actual stats
- We can study from each other

community

- Somehow you can share your experience
 - it (probably) makes internet better
- You can study from others
 - it (probably) makes your network better

We can improve the internet

