T8-1. Broadband Network Design

< ISP Access NW Design & Operation>

SoftbankBB Corporation

Hiroki Muramasu <hmuramas@bb.softbank.co.jp>

Prologue

Purpose of this session

- Introduce how we should construct a broadband access NW by own-provided line
  (Supplying DF, NW Topology, Routing Design, etc..)

- Consider the resource we need for broadband access NW
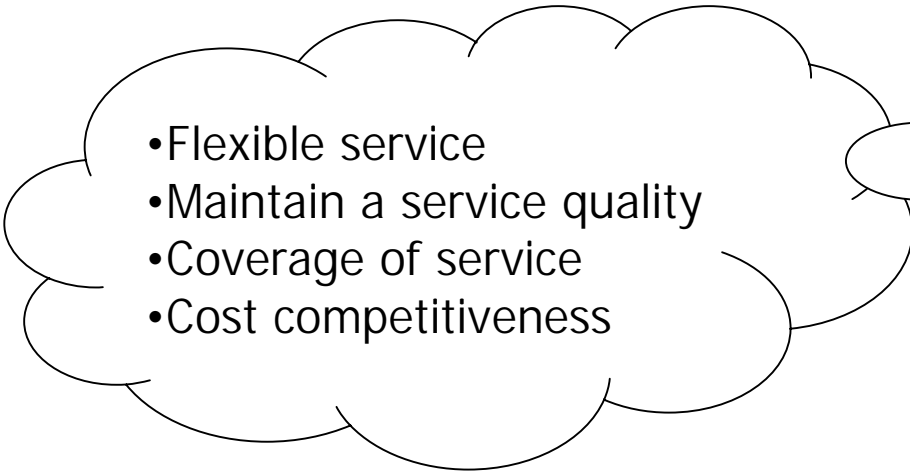
- Introduce some useful tools for broadband NW operation

Table of Contents

- **Own-provided access NW**
  v.s   Wholesale access NW
- **Composition of**
  own-provided access NW
- **How to supply a Dark Fiber**
- **Network Topology (Ring? or Star?)**
- **Equipment Selection**
- **Routing Design**
- **Useful traffic monitoring tool**
- **Summary**

Own-provided access NW
v.s　Wholesale access NW

- **Selection of Access NW in ISP's perspective**

  - **Select the own-provided access NW?**
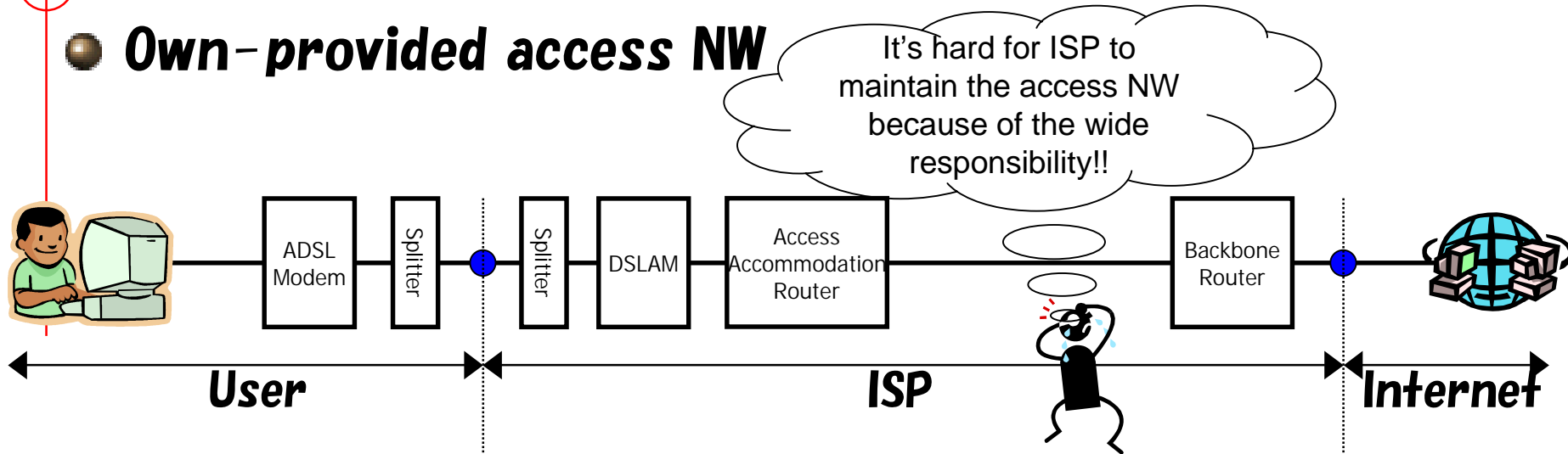 - **Select the wholesale access NW ?**

•Investment
•Operation
•Business
　　scheme

•Flexible service
•Maintain a service quality
•Coverage of service
•Cost competitiveness

Compare each access NW

● Own-provided access NW

It's hard for ISP to maintain the access NW because of the wide responsibility!!

User	ISP	Internet

ADSL Modem — Splitter — Splitter — DSLAM — Access Accommodation Router — Backbone Router

● Wholesale access NW

Condition of the access NW?

ADSL modem — Splitter — **Black Box?** — Access Accommodation Router — Backbone Router

User	Wholesaler	ISP	Internet

Compare each access NW

Q. **Why is the construction**

of each access NW different ?

A. **Because**

" *Function ISP requires for access NW is different* "

In the case of Wholesale access NW, there are some
limitations on NW design to satisfy ISP's required function,

But in the case of own-provided access NW,
ISP can design their desired service suitable NW.

(ISP's Requirement against Wholesaler)
- Provides an authentication function(PPP & Radius)
- Distributes ISP's IP address to their Users(Virtual Router)
- Do the best to save ISP's IP address

Compare each access NW

- ● **Each access NW in ISP's perspective**

Item	Own-provided Access NW	Wholesale Access NW
Initial Cost	(Very)BIG!	Small
Coverage	ISP can judge	Depend on Wholesaler
Service/Quality	ISP can judge	Depend on Wholesale NW and their NW function
Operation /Maintenance	Very Hard	Leave to Wholesaler
General Comment	Flexible business deployment, however big initial cost & operation hardship	Small initial cost & operation easiness, however business deployment depends on Wholesaler.

Coffee Break ⟨No.1⟩

● Access Provider's Distress

Struggle with circuit noise!
(individual investigation for user have unstable circuit)

Hard competition of development
(1.5Mbps→8Mbps →12Mbps→Double →Quad→?)

Launch new location to expand coverage

Upgrade circuit & facility to maintain the quality for several millions of port

Increase the access bandwidth per user rapidly by the change of application

Etc…

Compare each access NW (Multicast)

● Own-provided access NW

Can save the bandwidth, because of only once sending

There is no severity on Spec, because small number of copy.

| ADSL Modem | Splitter | | Splitter | DSLAM | Access Accommodation Router | | Backbone Router | |
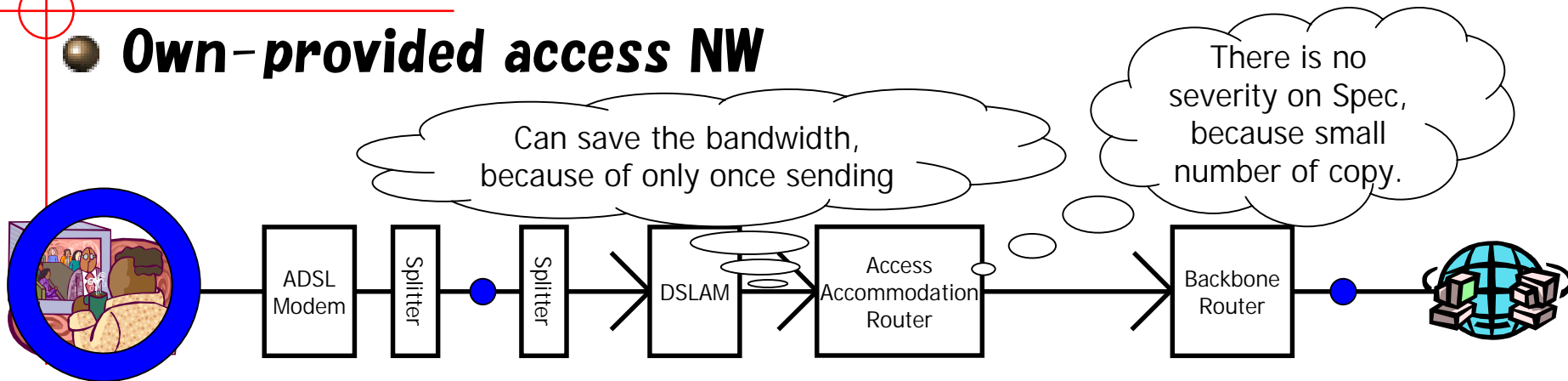
- ● It's easy to realize multicast, because of small number of copy per node.
- ● The bandwidth on Trunk NW is saved, if the topology is designed better.

● Wholesale access NW

| ADSL modem | Splitter | | Splitter | DSLAM | | | Router | |

The possibility of the huge number of copy on BRAS terminates several thousands of session?

If multicast packet flows per ATM-VC, it will be saturated. (Though we use multicast, the bandwidth will not be saved.)
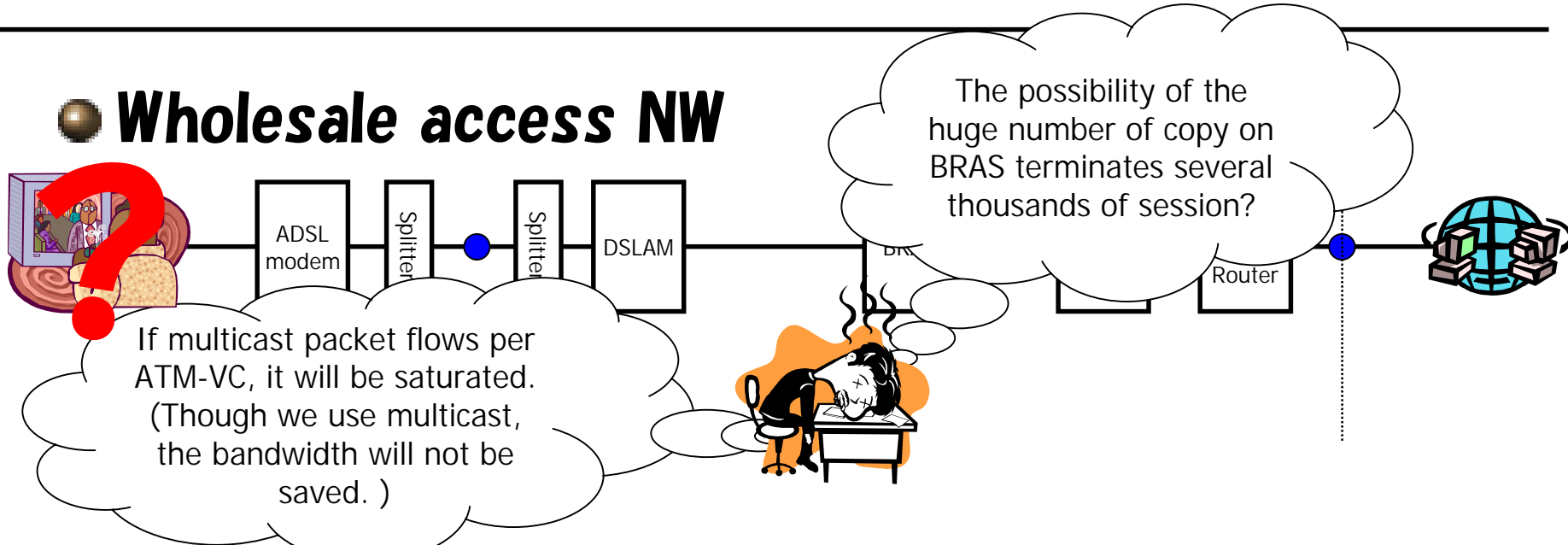
Selection of Access NW in ISP's perspective
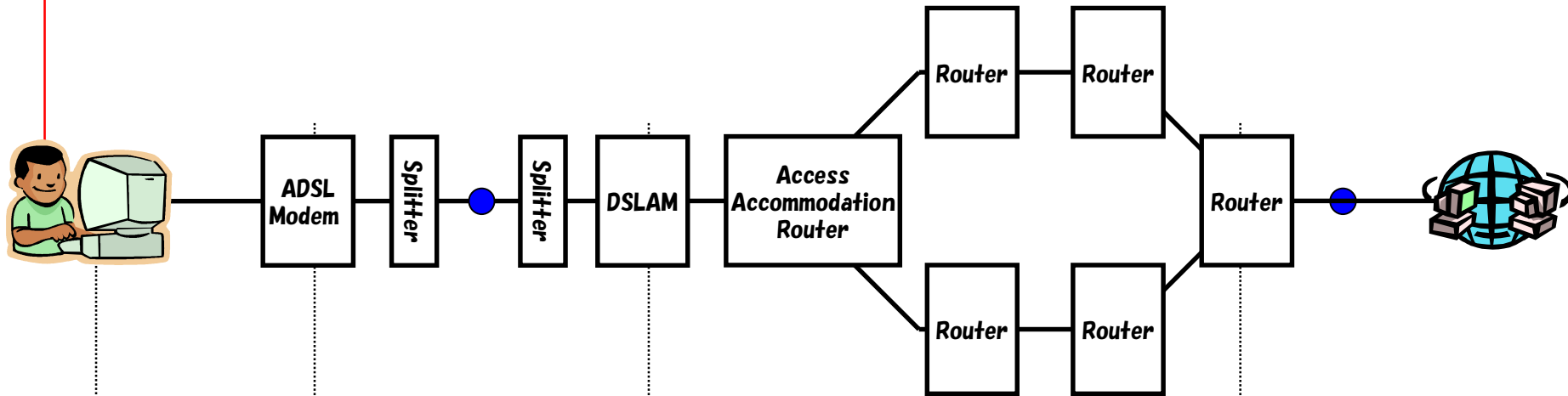
- **Own-provided access NW**

  v.s   Wholesale access NW
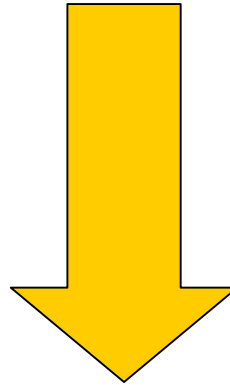
- **In the result, <u>it depends on the ISP's Business model</u>.**
- **The own-provided access NW is very flexible for service/technology deployment, but initial cost is huge and operation is hard.**
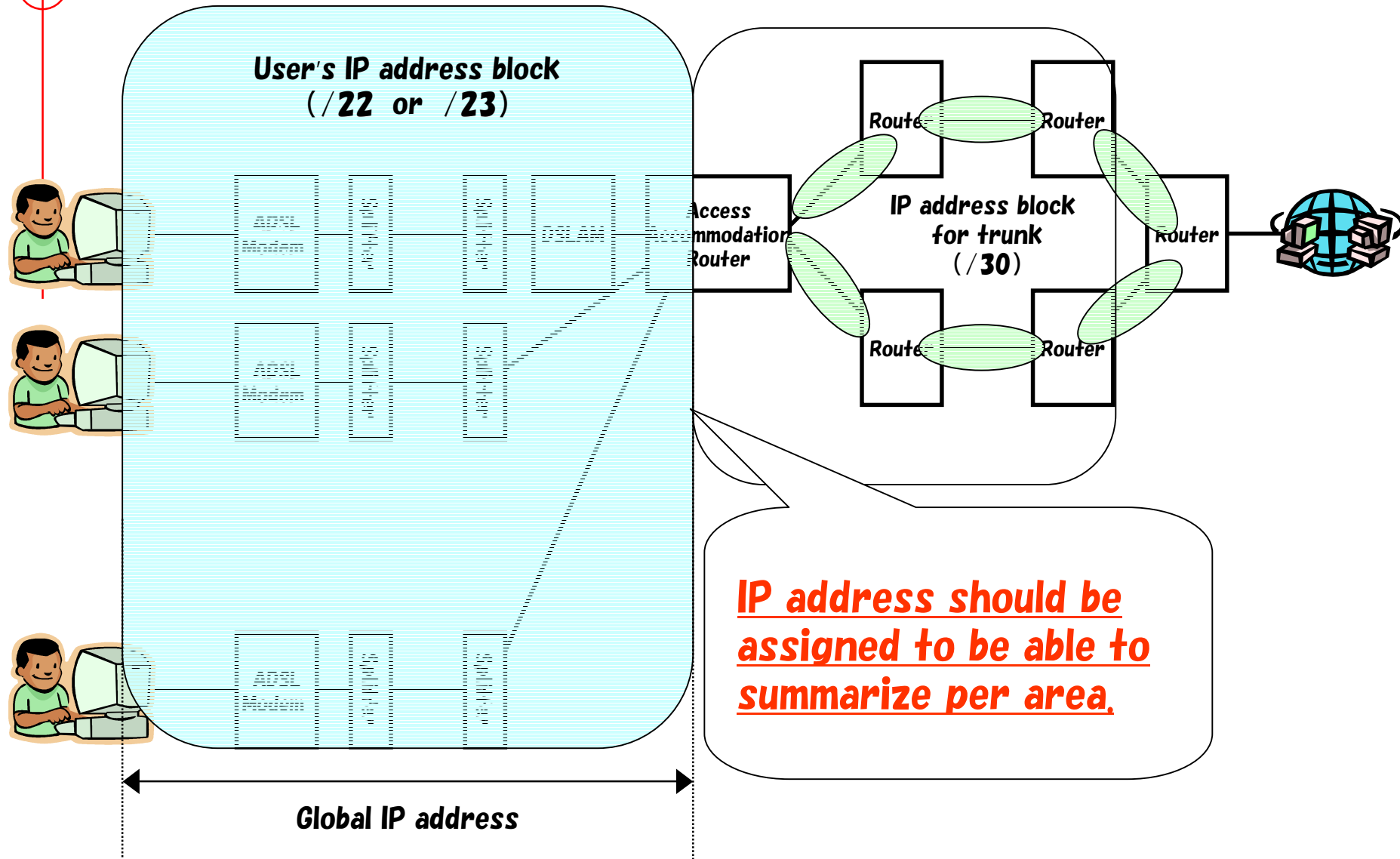
Composition of own-provided access NW

● **Protocol Structure of own-provided access NW**

TCP/IP	TCP/IP		TCP/IP	Layer 3
Ethernet / WLAN	Ethernet		Ethernet	Layer 2
	ATM			
	ADSL			
UTP / Radio	Copper		UTP / Fiber	Layer 1

IP address assignment on access NW

User's IP address block
(/22 or /23)

ADSL Modem ... **DSLAM** ... **Access Accommodation Router**

Access Accommodation Router

Router — **Router**

IP address block for trunk (/30)

Router

Router — **Router**

IP address should be assigned to be able to summarize per area.

Global IP address

Supply the trunk circuit

- ## Let's use the Dark Fiber (DF) !!

  - ### Because SONET/SDH, ATM circuit service depends on the carrier.
  - ### But if there is no Dark Fiber on a certain section, we should use the carrier's circuit service.

Can we rent the DF easily?

- **Confirm the fiber space on the public information** (Judge Level: A〜D)

Judge Level	Fiber Space
A	Large Space(can rent DF)
B	Middle Space(can rent DF)
C	Small Space
D	No Space

- **If the judge level is "A" or "B", we can rent the DF easily. Therefore we should design NW on the assumption that we rent them.**

How much is DF?

● **The price of Trunk DF depends on the distance.**

(2004/6)

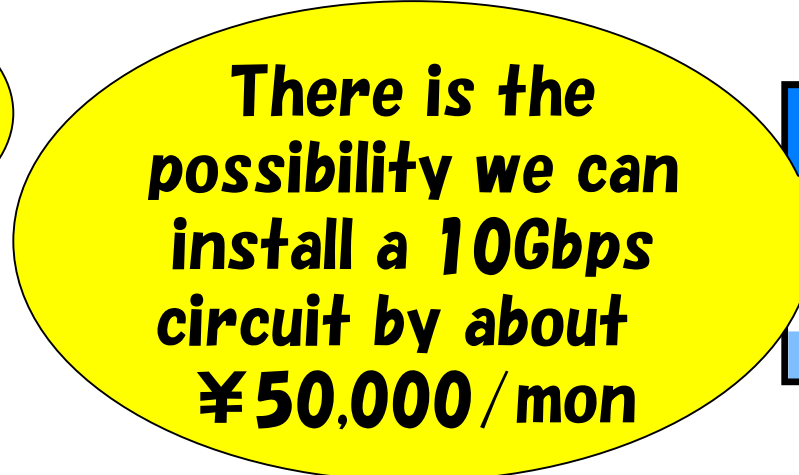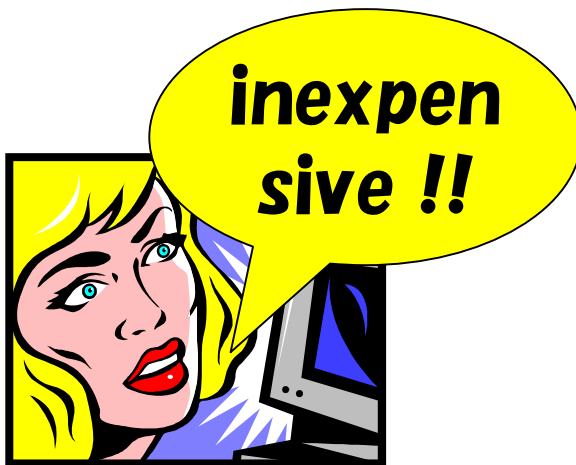| Item	Trunk DF	Subscribe DF
Purpose	Use between NTT station	Use between NTT station and Subscriber
Circuit Price	(E) ￥2.627/**m**/Fiber/month (W) ￥2.751/**m**/Fiber/month	(E)(W) ￥4,879/Fiber/month
Maintenance Fee	(E) ￥139/Fiber/month (W) ￥143/Fiber/month	(E) ￥129/Fiber/month (W) ￥135/Fiber/month
Additional Fee	—	(E)(W) ￥471/Fiber/month

(E) NTT East   (W) NTT West

Coffee Break ⟨No.2⟩

🔘 **Dark Fiber's price is high? or low?**

ex), If we use the trunk DF to connect
between NTT station is 10km,
the circuit price is

$$¥2.627(/m/Fiber/mon) × 10,000(m) × 2(Fiber)$$

$$= ¥52,540/mon \quad (NTT East's Case)$$

inexpen sive !!

There is the possibility we can install a 10Gbps circuit by about ¥50,000/mon

Let's rent DF!

- **Apply "POI investigation" to East/West NTT**

  - **Dark Fiber**
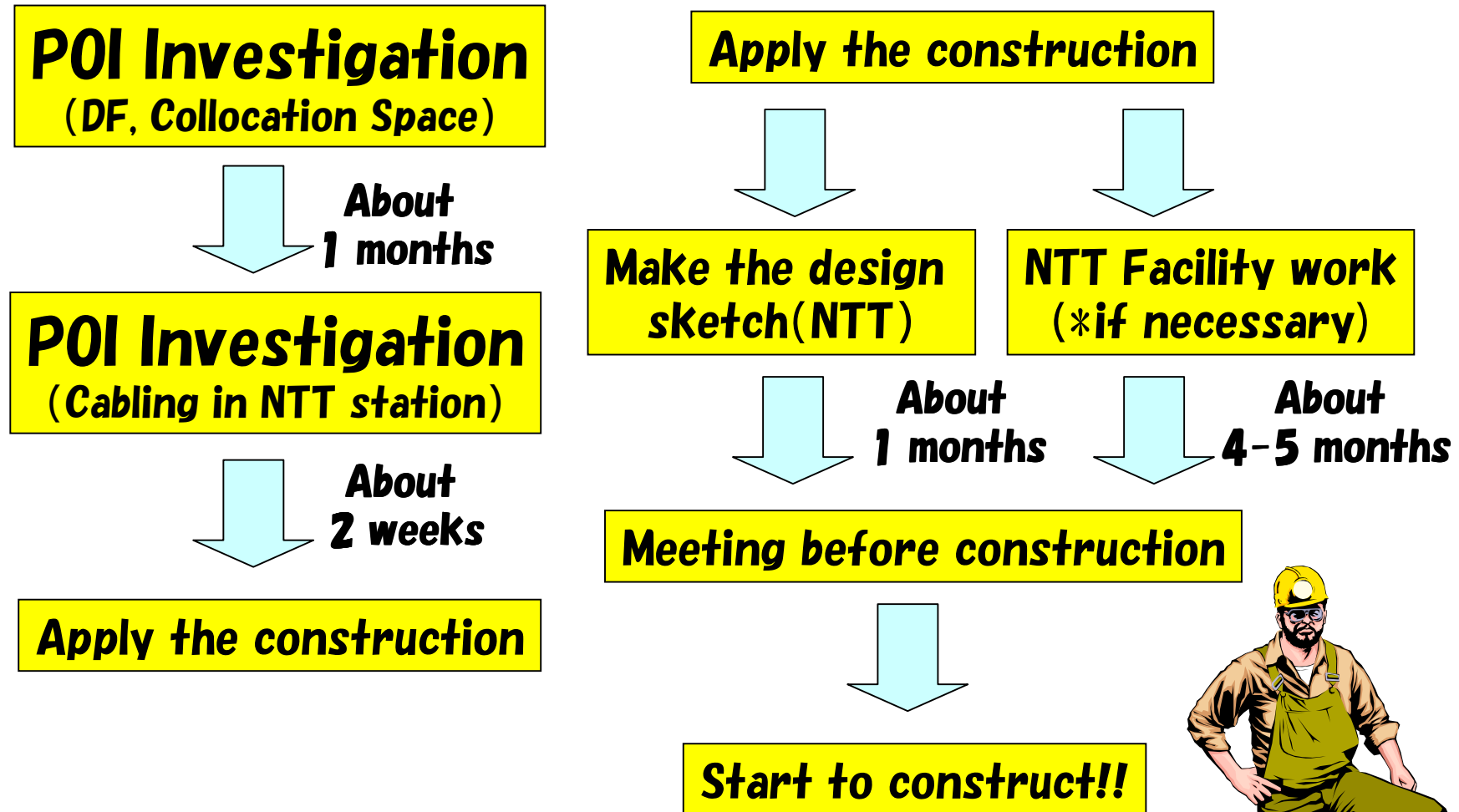 - **Collocation Space**
  - **Power**
  - **Cabling in NTT Station**
    (FTM ～ Rack, Rack ～ Rack)

Process from POI investigation to DF construction

POI Investigation
(DF, Collocation Space)

↓ About 1 months

POI Investigation
(Cabling in NTT station)

↓ About 2 weeks

Apply the construction

Apply the construction

↓

Make the design sketch(NTT) | **NTT Facility work (*if necessary)**

↓ About 1 months | ↓ About 4-5 months

Meeting before construction

↓

Start to construct!!

Attention after supplying DF

- ## It's very important to manage facilities

  - ### Adjust the management method beforehand

  - ### Don't forget to manage the cable number on FTM
    - Make clear which port on the router a fiber connects to.
    - Attach tags on all cables in NTT station

- ## It's better to manage facilities on DB
  - Be able to manage many DFs efficiently

NW Topology（Ring? or Star？）

- If you select the DF for trunk part,

    ## *Ring topology is practical*

    - It's easy to make the ring topology by connecting between adjacent NTT station
    - On the other hand, it's difficult to make the star topology by NTT's trunk DF

Let's select "Ring" topology this time!

How to make a Ring Topology

- **Connect between adjacent NTT station as much as you can**

- **Pay attention that large ring may cause you the difficulty to control traffic**

- **Also pay attention that multistage small ring may cause you difficulty to control traffic**

- **Ring hierarchy is necessary (up to 2 tier)**

  - **Major Ring(high Speed Ring)**

    ：Summarize all traffic in certain prefecture

  - **Minor Ring(low Speed Ring)**

    ：Summarize all traffic in small area.

Major Ring Design

All Traffic in certain prefecture is summarized in this Major Ring. We should plan to upgrade the bandwidth here mainly !

: Router in large station (in Major Ring)
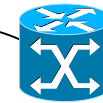
: Router in middle or small station (in Minor Ring)

Backbone
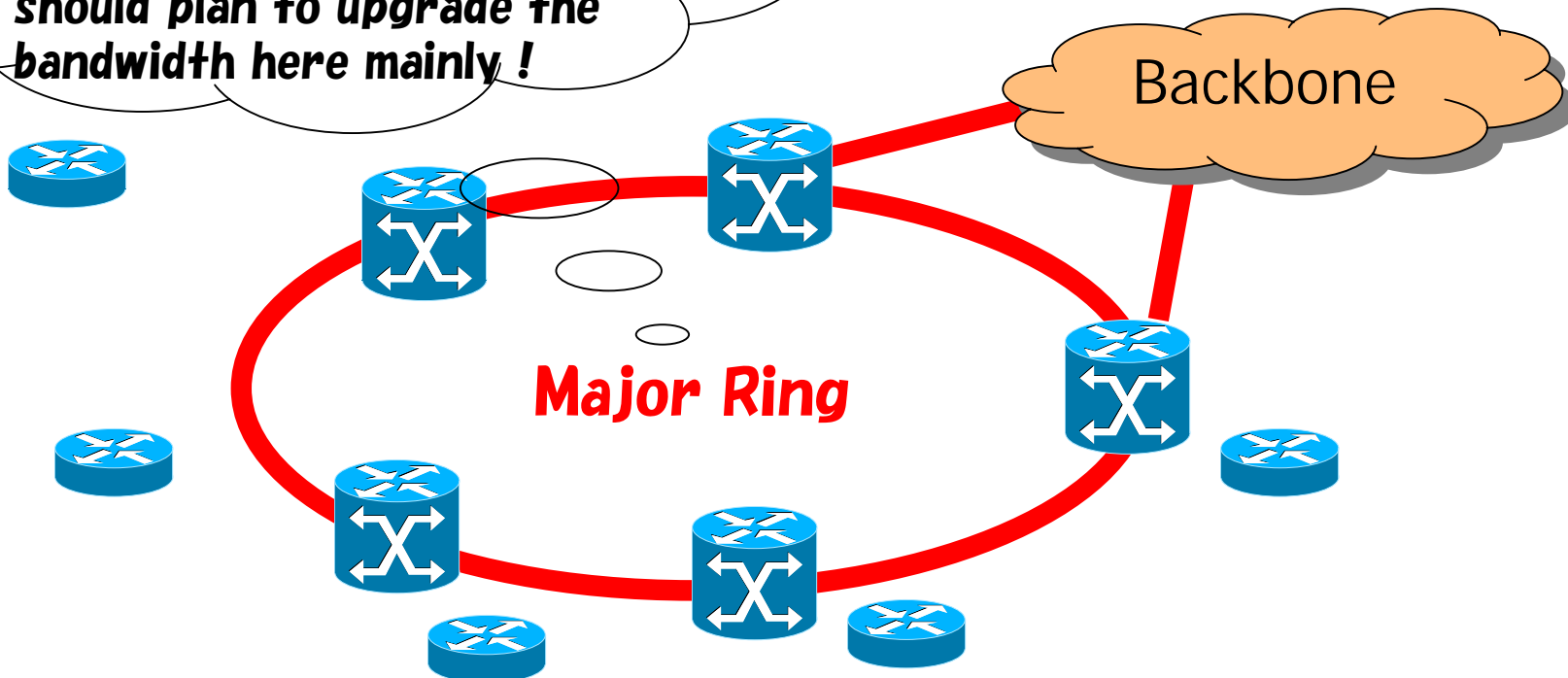
Major Ring

We should select the station accommodates large scaled user as the major ring station (includes one connects to Backbone.

And this major ring stations are better if it's easy to upgrade the facility(collocation space or power) and supply trunk DF.
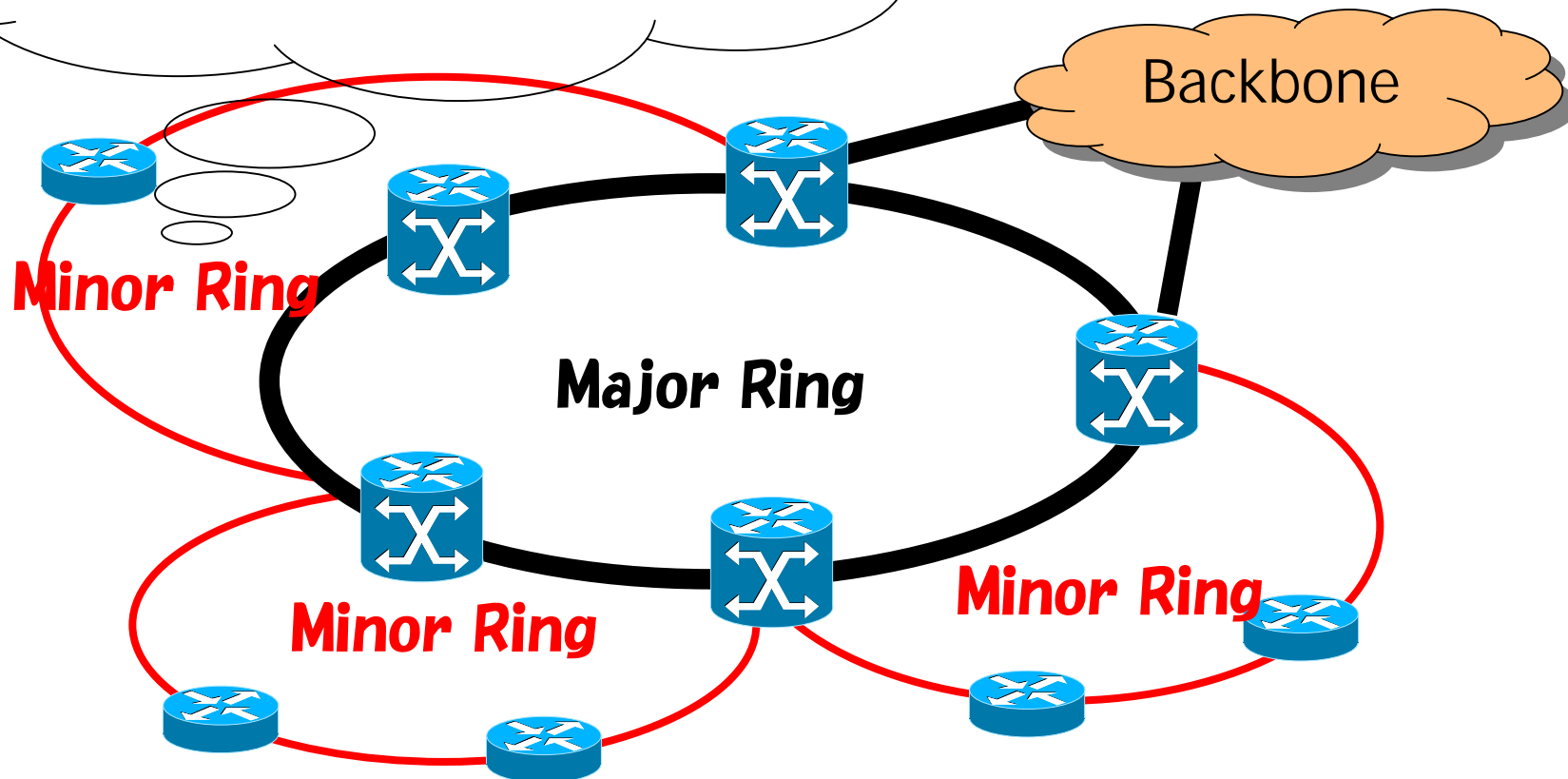
Minor Ring Design

All Traffic in middle/small area is summarized in this Minor Ring. This should be connected at two points to Major Ring!

: Router in large station
  (in Major Ring)

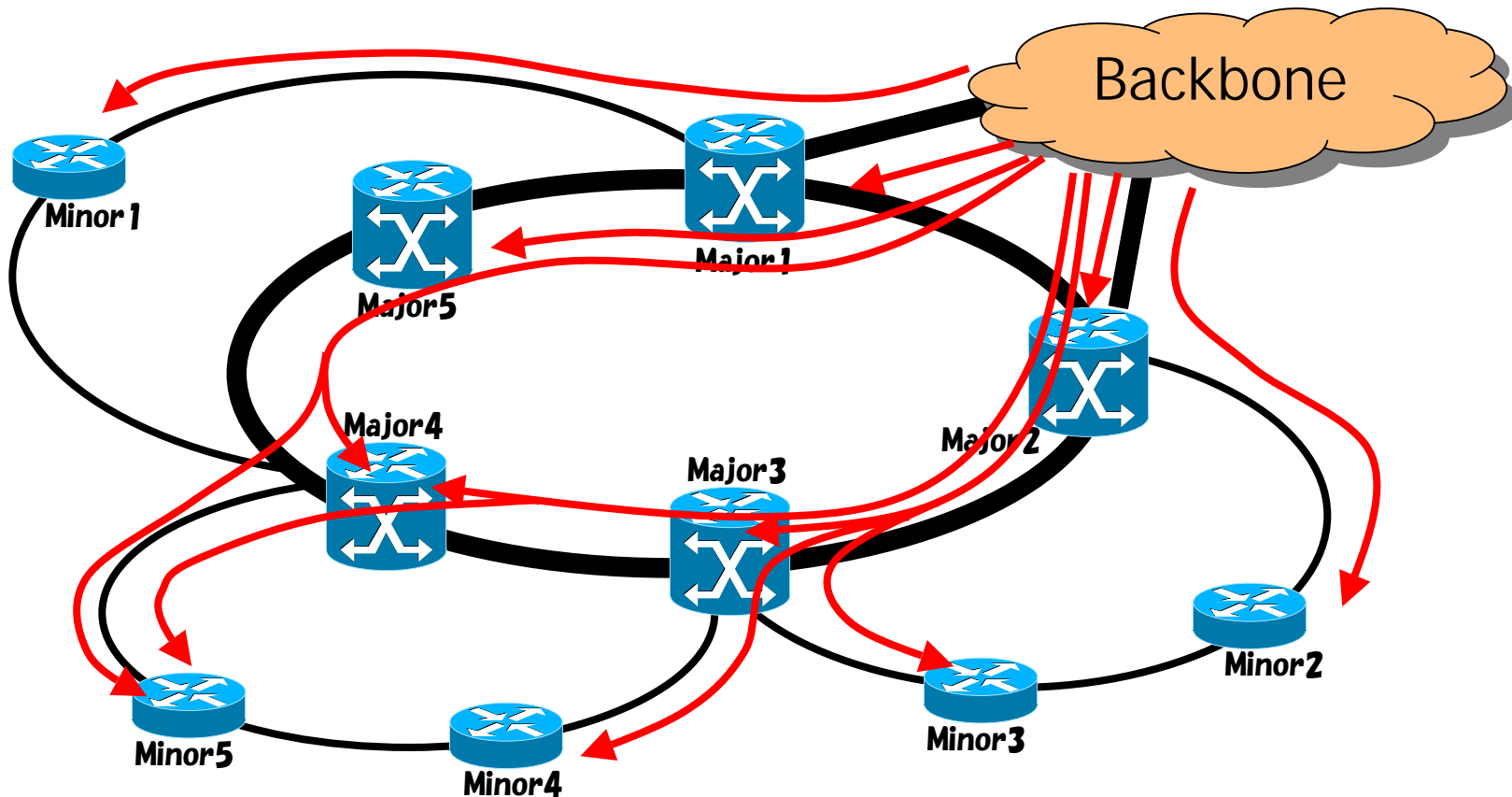: Router in middle or small station
  (in Minor Ring)

Backbone

Minor Ring

Major Ring

Minor Ring

Minor Ring

Traffic Flow on Ring Topology

● **Traffic flows as below**
   on Ring Topology

: *Router in large station*
   (in Major Ring)

: *Router in middle or small station*
   (in Minor Ring)

Backbone

Minor1

Major5

Major1

Major4

Major3

Major2

Minor5

Minor4

Minor3

Minor2

Equipment Selection(Router vs. L3 Switch)

- **What's appropriate equipment for broadband NW design?**

  - Router ?
  - L3 Switch ?

First of all,
What's the difference
between Router and
L3 SW

Compare Router with L3 Switch

Item	Router	L3 Switch
Cost	High	(Very) Low
Supported IF Type	E3, ATM, SONET/SDH, Ethernet,	Ethernet (Sometimes, other IF is supported)
IF accommodation density	Low	High
IF buffer Capacity	Several hundreds [Mbyte]	Several ten [Mbyte]
Routing Function	RIP/OSPF/IS-IS/BGP	←
Routing table volume	Over a million	150 thousands ~250 thousands
MPLS Support	Yes!(exists experience)	No (in future?)
Switching Capacity	Several hundreds [Gbps] ~Several [Tbps]	Several hundreds [Gbps]

Compare Router with L3 Switch

● Occurrence Factor of Congestion

It's caused by Speed inequality or Aggregation, not IF speed or performance of equipment.

Seed inequality	Aggregation
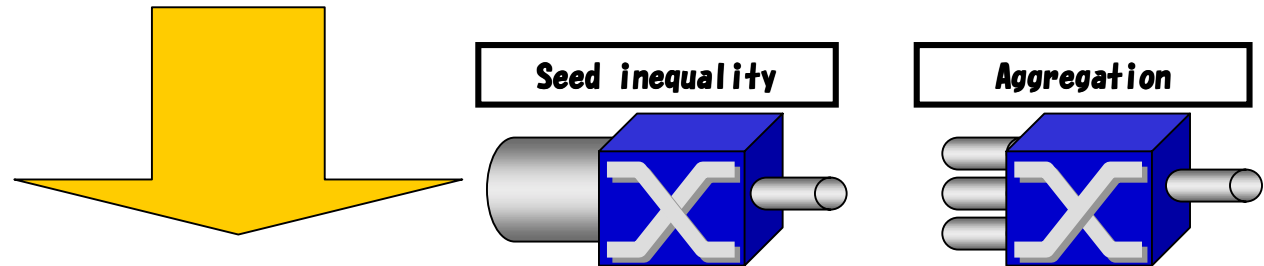
● Problems by congestion

The congestion causes large latency, zitter, and packet loss. Although it may not affect Web or Mail service, Application service such as Voice or VoD must be affected.

We can't ignore that.

Compare Router with L3 Switch

- **Traffic changes as saw-shape by TCP's flow-control, when it congests**
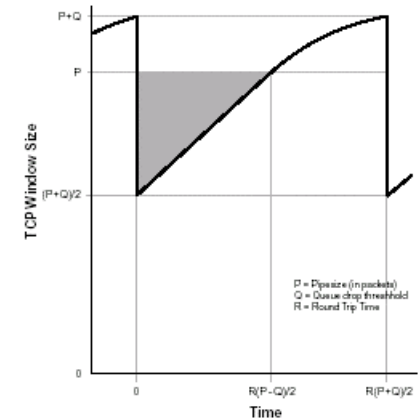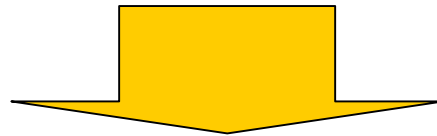

Figure 7: Detail of one TCP congestion avoidance cycle

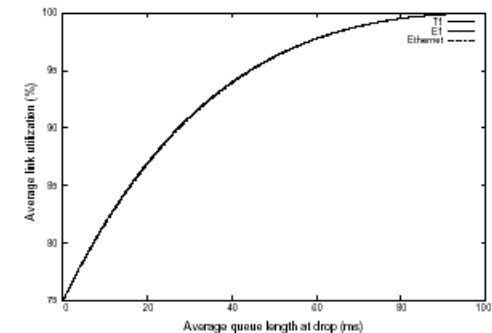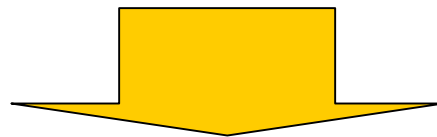- **Effective bandwidth usage rate is affected by the buffer volume**

Figure 8: Effect of threshold on average utilization

- **In equipment selection, we should consider a design policy of circuit bandwidth, service policy, etc.**

Point on equipment comparison and NW construction

- Basically, ISP can select both
- In result, you should select appropriate one by the circuit type you use or services you want to provide

Item	Router	L3 Switch
Cost	High	Very Low
Supported IF	E3, ATM, SONET/SDH, Ethernet,	Ethernet (Sometimes, other IF is supported)
IF accommodation density	Low	High
IF buffer Capacity	Several hundreds [Mbyte]	Several ten [Mbyte]
Routing Function	RIP/OSPF/IS-IS/BGP	←
Routing Table volume	Over a million	150 thousands ~250 thousands
MPLS Support	Yes! (exists experience)	No (in future?)
Switching Capacity	Several hundreds [Gbps]~Several [Tbps]	Several hundreds[Gbps]

How much do you allow as a cost for equipment?

Which do you use DF or leased circuit?

How is the speed inequality on a circuit? What's the bandwidth design policy? What's the topology at aggregation point?

Do you provide a MPLS VPN?

Routing Design

- ● **Let's use OSPF as IGP(Popular in Japan)**
  - ● **Define Backbone as Area0**
  - ● **Define own-provided NW in each prefecture as sub-Area**
      - ⇒**Define as one area if LSA number is about 200.**
- ● **Let's use BGP as EGP(for scalability in future)**
  - ● **IP address block assigned to user is distributed by BGP**
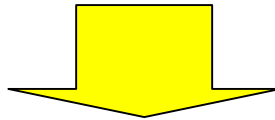
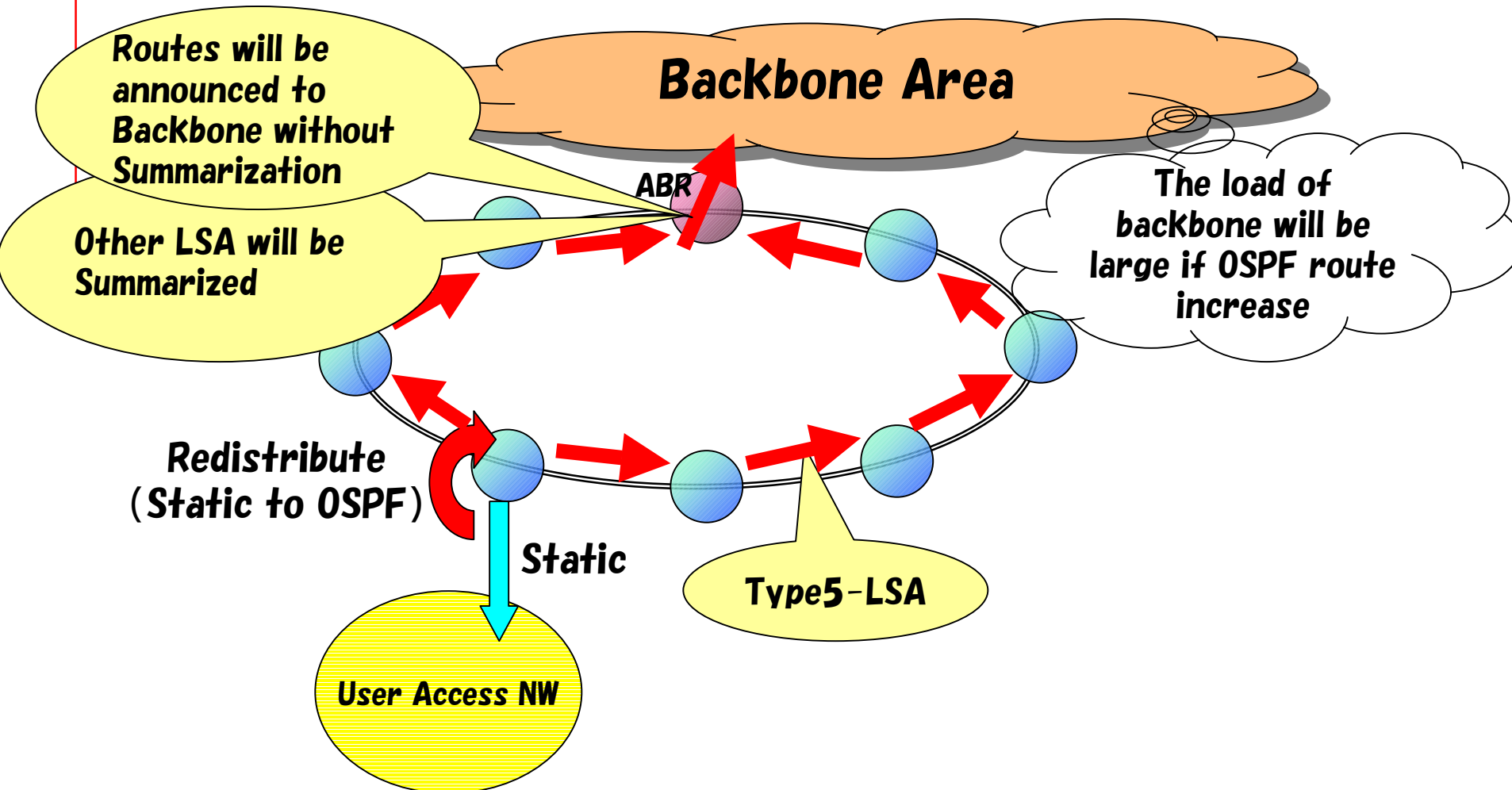**We should reduce the route volume for the cost reduction and route increase in the future.**

- ● **Reduce the OSPF route volume by BGP (Load by OSPF is high)**
- ● **Reduce the route volume flows into backbone area**
- ● **Use the default Route effectively**
- ● **Use functions of OSPF or BGP effectively**
- ● **Consider a countermeasure when trouble occurs**
- ● **IP address design is important**

Routing Design (IGP)

● **If we use OSPF for user's IP address block**.........

Backbone Area

Routes will be announced to Backbone without Summarization

Other LSA will be Summarized

ABR

The load of backbone will be large if OSPF route increase

Redistribute (Static to OSPF)

Static

Type5-LSA

User Access NW

Routing Design (IGP)

- If we use BGP for user's IP address block.........

Backbone Area

OSPF route volume is small as compared with BGP

Summarized route by BGP

Route Reflector is summarized BGP routes

Route Reflector

Originate

Static

User Access NW
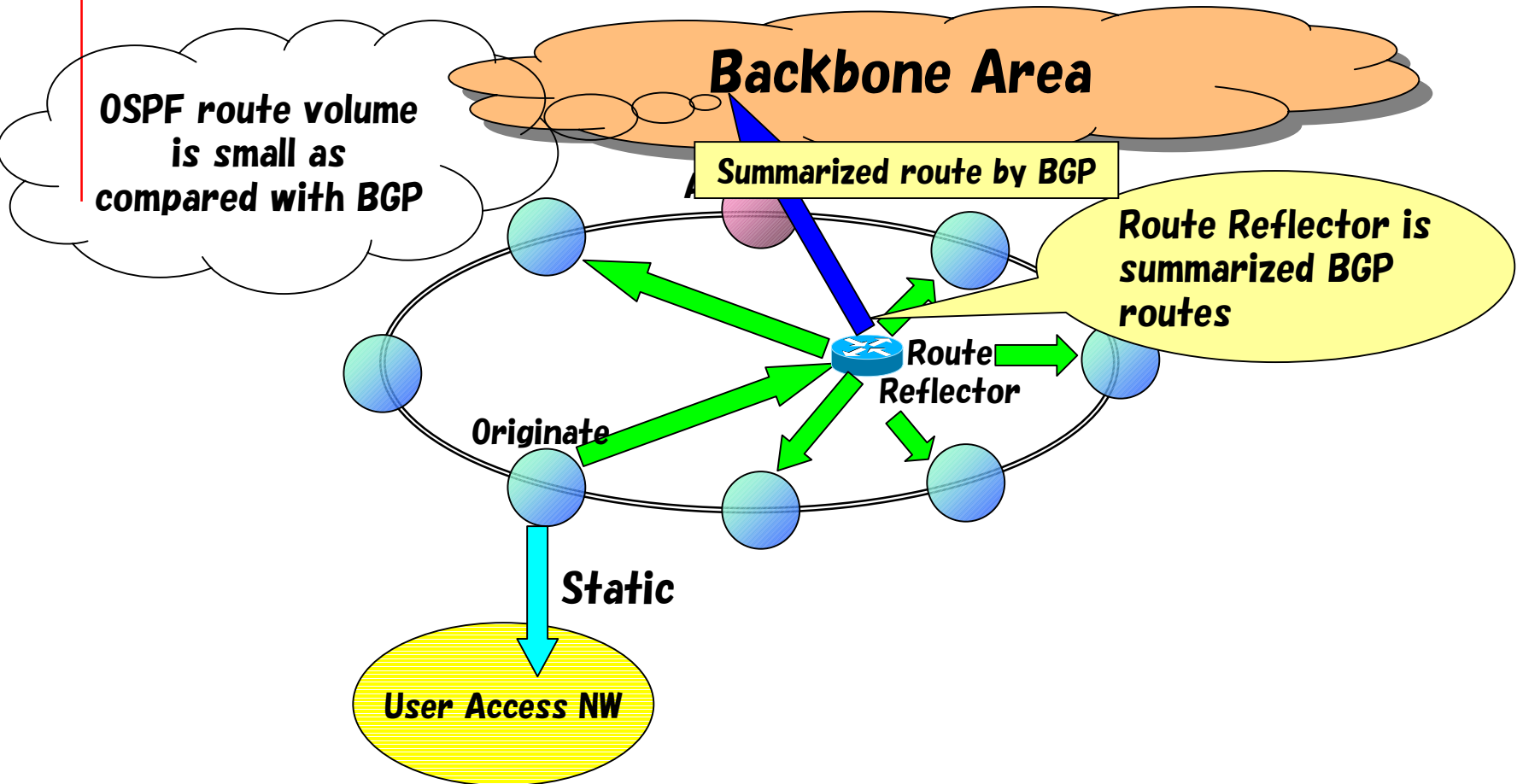
Routing Design (IGP)

- OSPF Cost design: Very important on especially Ring topology

Backbone Area

When a incident occurs

Normal

Trouble!!

Large diversion traffic flows into the Low speed Minor Ring. **Dangerous!!**
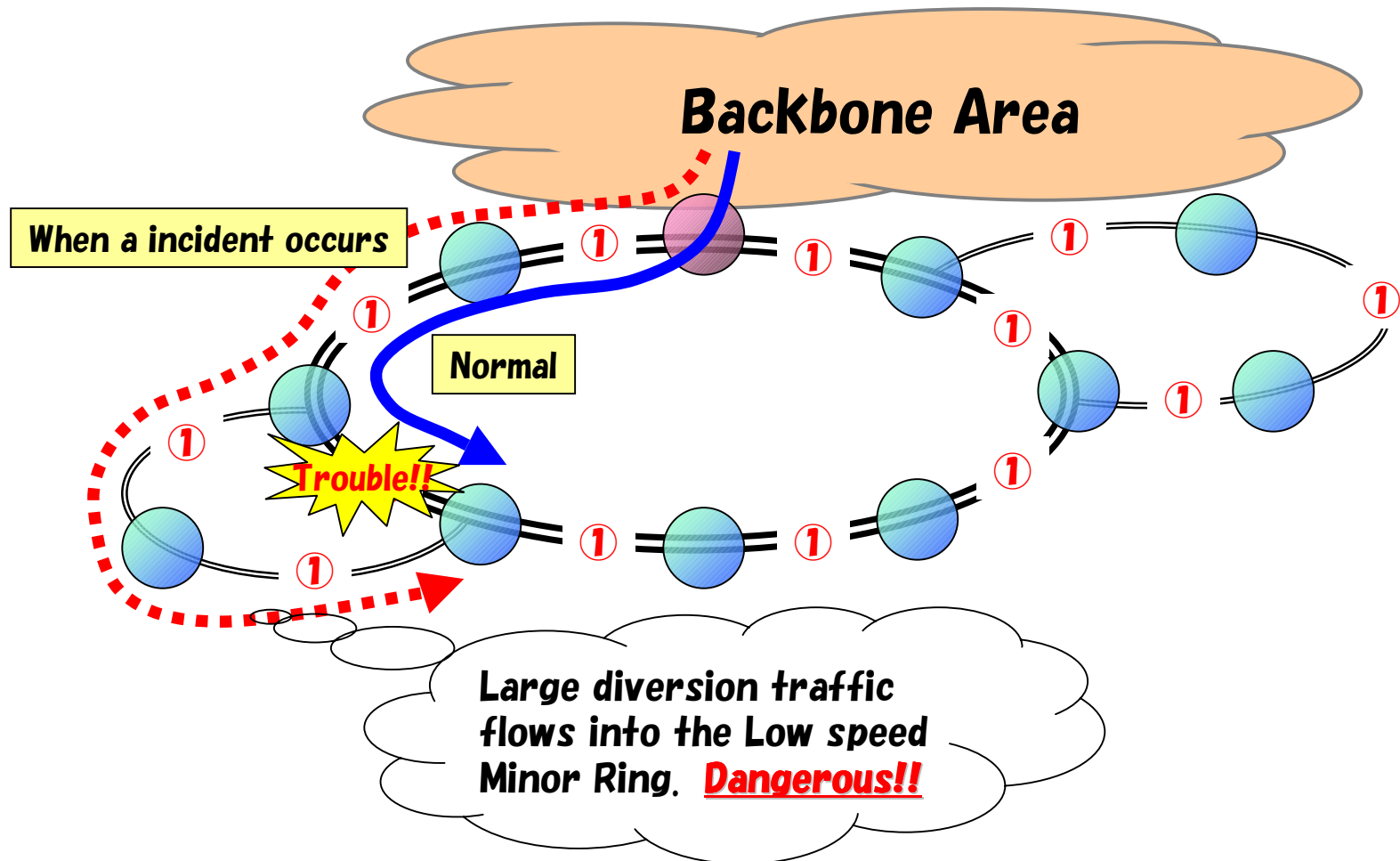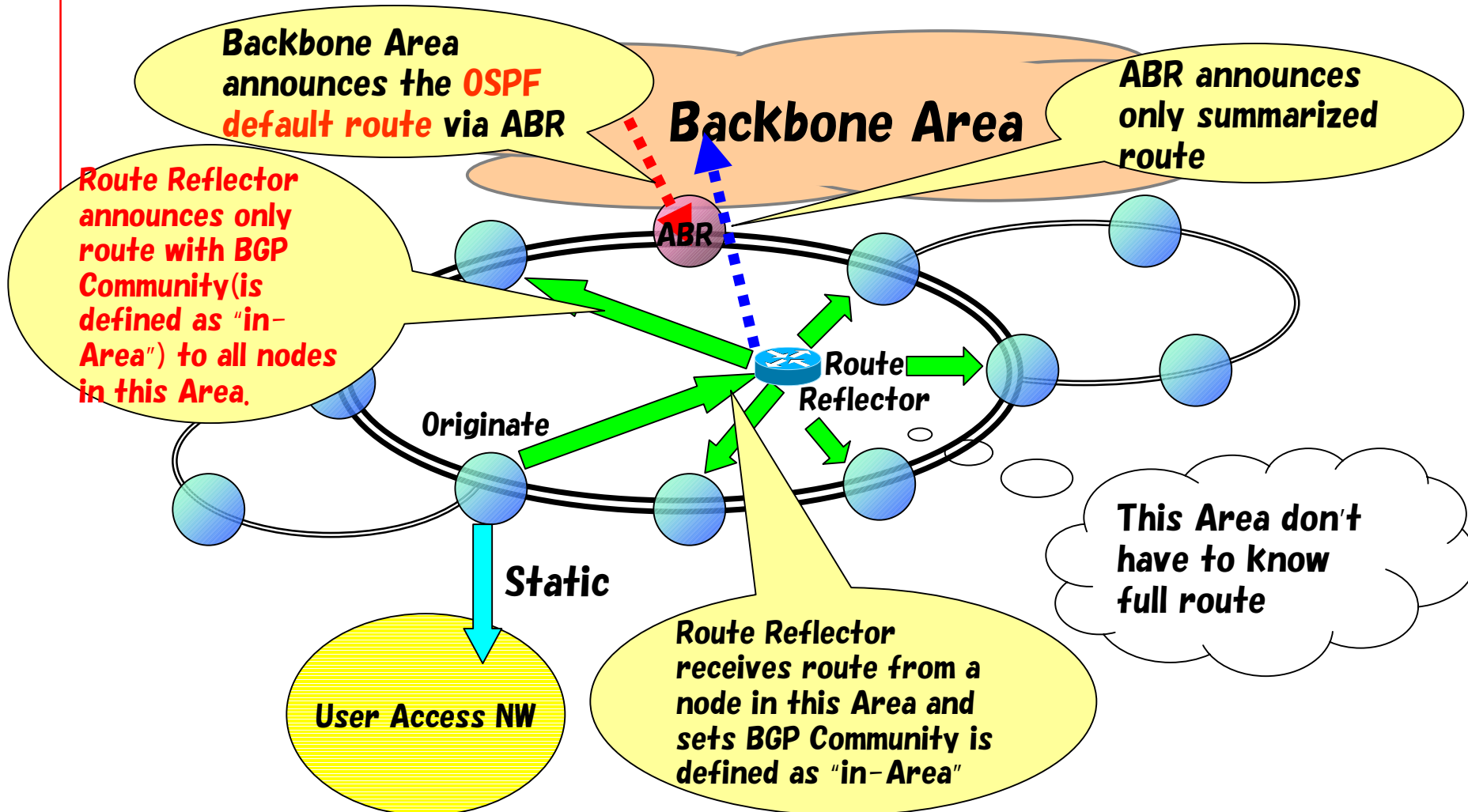
Routing Design (IGP)

- OSPF Cost design:

  Cost volume on Minor Ring should be assigned larger volume than Major Ring. (The Cost is defined by Hop number)
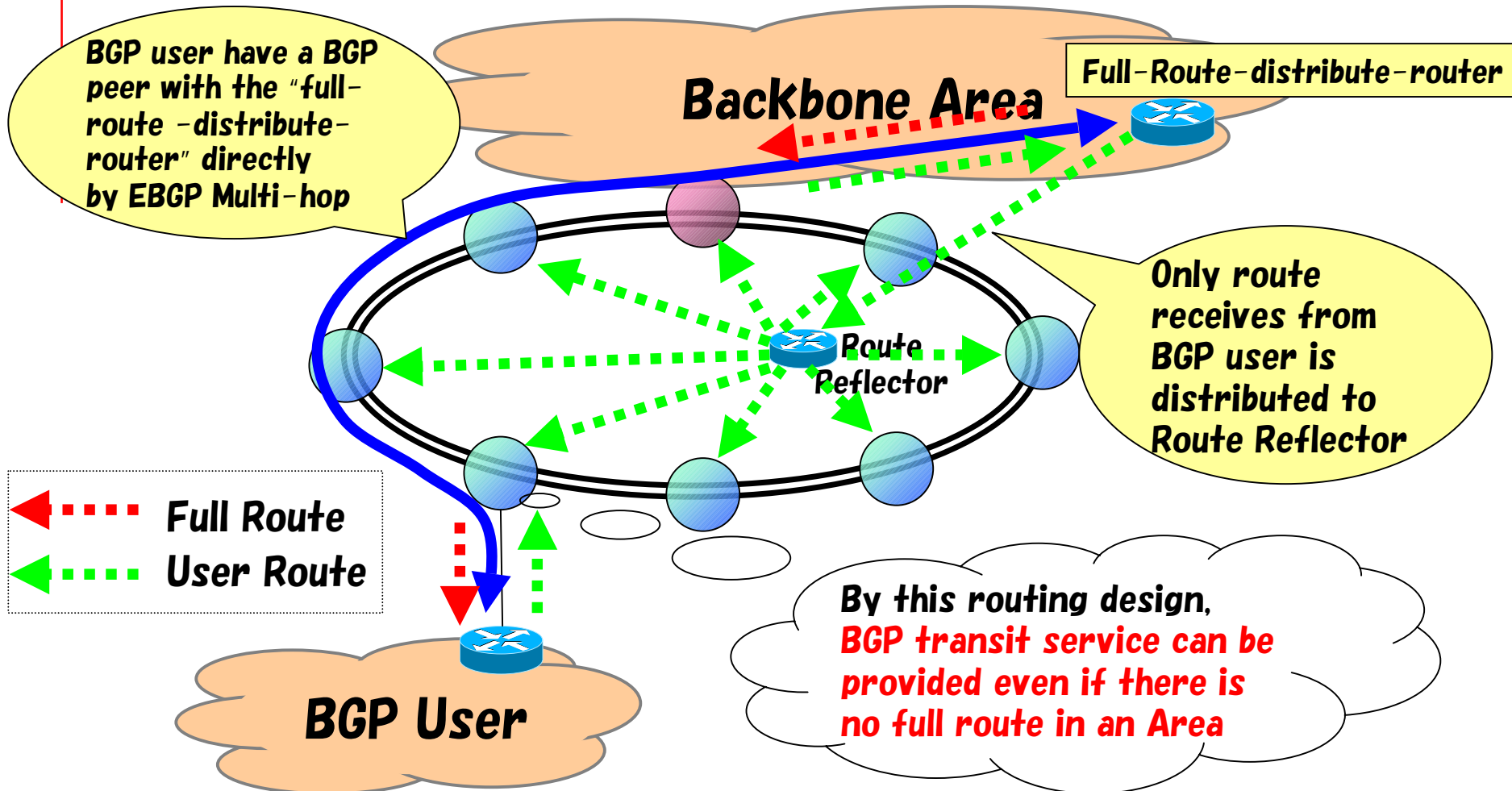
Backbone Area

Normal

When a incident occurs

Trouble!!

Large diversion traffic doesn't flow into the Low speed Minor Ring.

Safe! Great!!
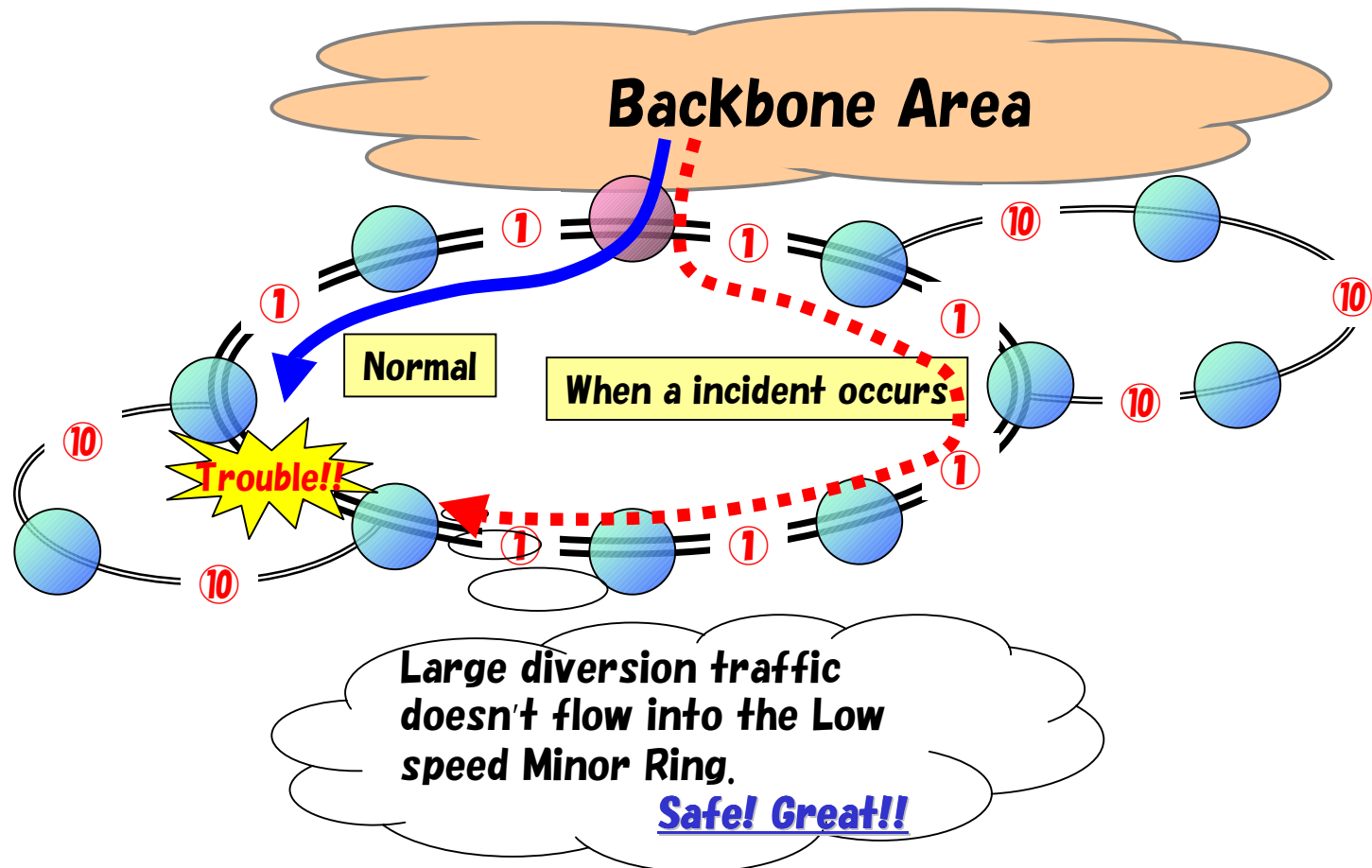
Routing Design (BGP)

- **Route Reflector Keeps only BGP routes in an Area**

Backbone Area announces the OSPF default route via ABR

Backbone Area

ABR announces only summarized route

Route Reflector announces only route with BGP Community(is defined as "in-Area") to all nodes in this Area.

ABR

Route Reflector

Originate

Static

This Area don't have to know full route

User Access NW

Route Reflector receives route from a node in this Area and sets BGP Community is defined as "in-Area"

Routing Design (BGP)

🟤 **Transit service can be provided anywhere**

BGP user have a BGP peer with the "full-route -distribute-router" directly by EBGP Multi-hop

Full-Route-distribute-router

Backbone Area

Route Reflector

Only route receives from BGP user is distributed to Route Reflector

- - - - ◄ **Full Route**
- - - - ◄ **User Route**

BGP User

By this routing design, BGP transit service can be provided even if there is no full route in an Area

Routing Design (BGP)

- Transit service for only local route in a certain Area can be provided

Backbone Area

The route receives from BGP user isn't announced to Backbone Area

Route Reflector

◄···· Area Route

◄···· User Route

BGP User
(for only local route)

It's easy to provide the transit service for only local area anywhere in this area.
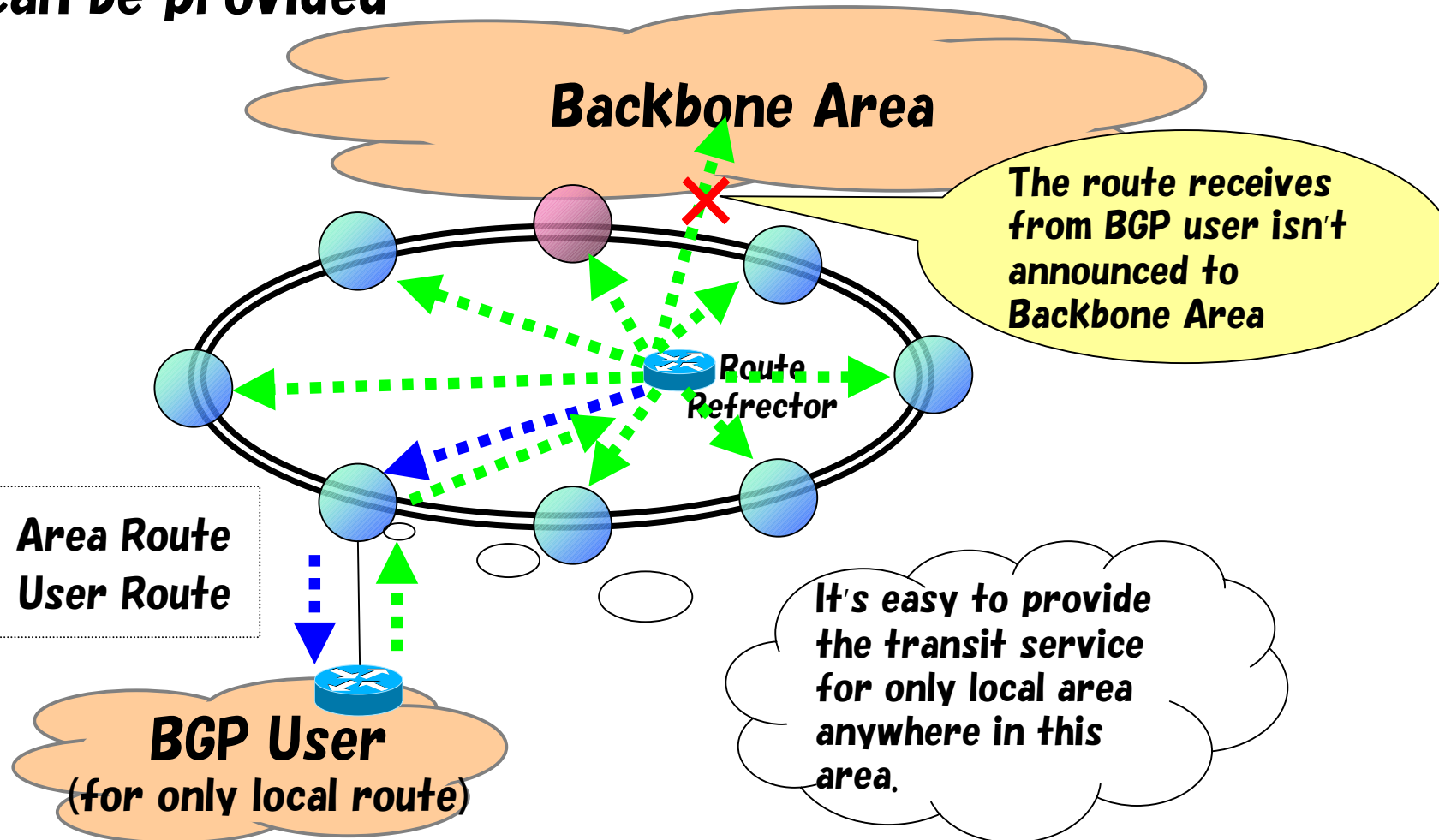
Let's flow various service traffic on Access NW

You can provide various service traffic on customer needs

⬤ **Multicast** ⬤ **VoIP** ⬤ **VoD** ⬤ **Streaming** etc…

⬇

⬤ **Pay attention to some service requires "Real Time" such as VoIP/Multicast !!**
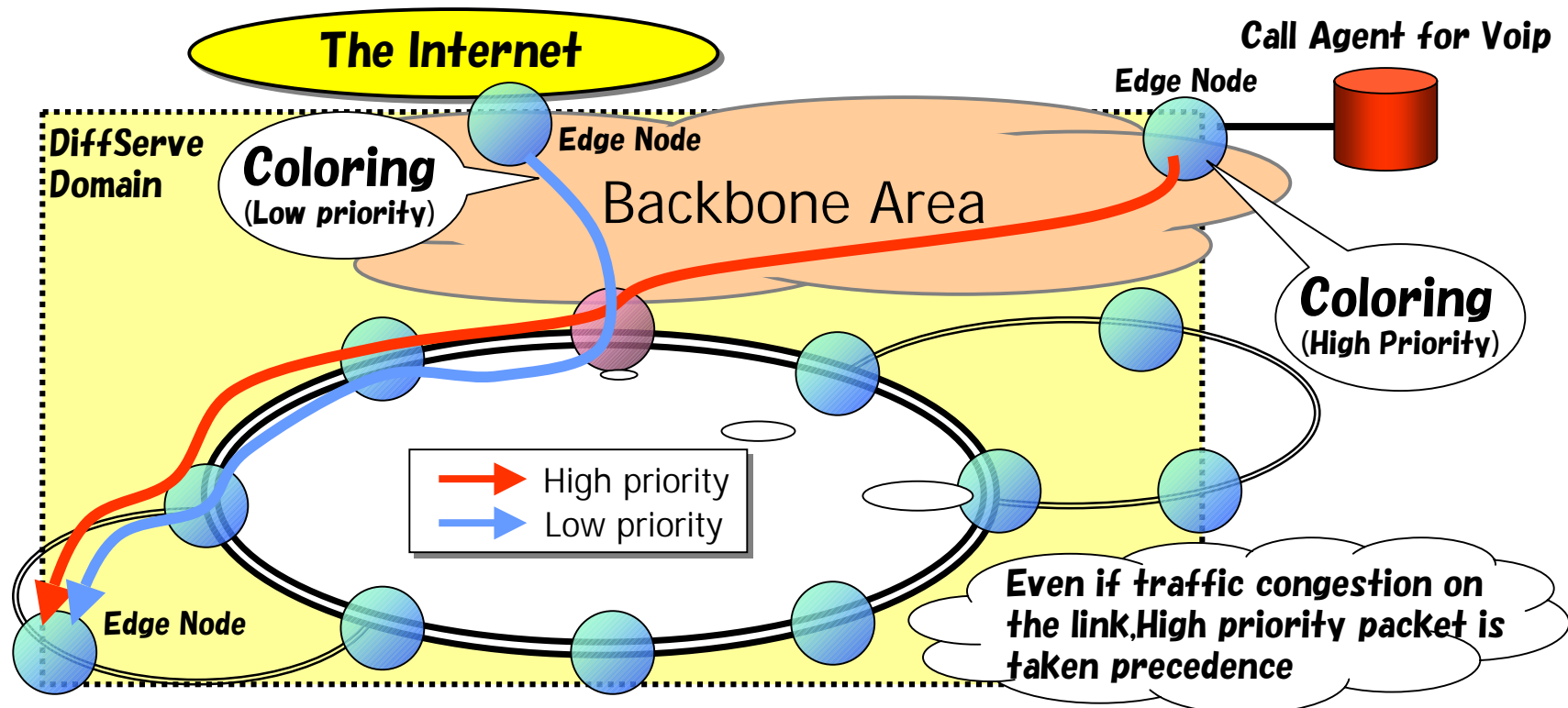
➡ **"QoS" may be the solution**

⬤ **Pay attention to large traffic by big contents such as VoD/Streaming !!**

➡ **"CDN" may be the solution**

Let's flow various service traffic on Access NW
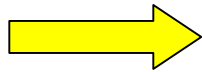
🔴 DiffServe (QoS)

The network is divided into "Domain" & "Edge", and the edge node adds a weight for service packet (is called "Coloring") and distributes the packet with the weight, all nodes on the domain control the service packet based on the weight
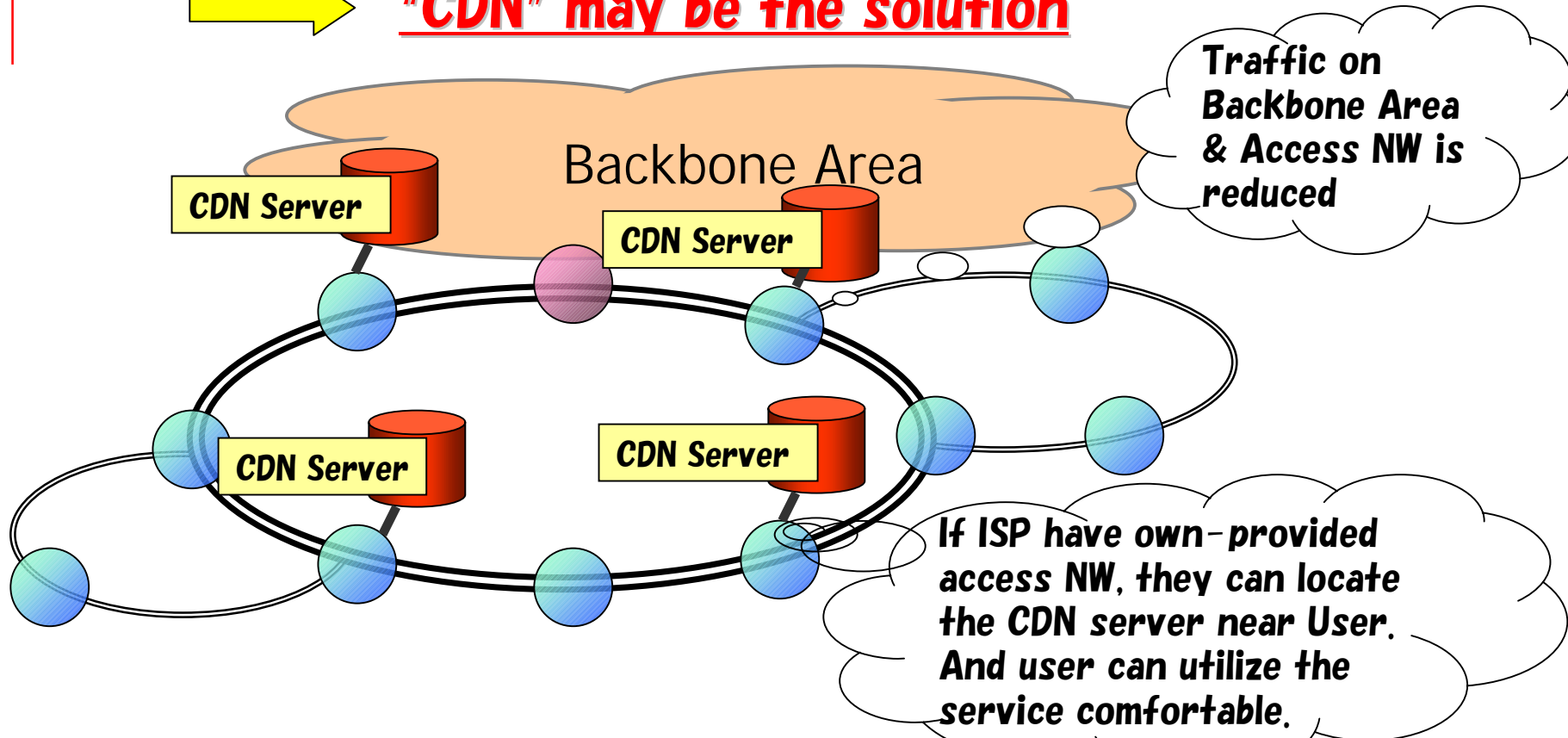

The Internet

Call Agent for Voip

Edge Node

DiffServe Domain

Coloring (Low priority)

Backbone Area

Edge Node

Coloring (High Priority)

High priority
Low priority

Edge Node

Even if traffic congestion on the link, High priority packet is taken precedence

Let's flow various service traffic on Access NW

● CDN

User continues to require more rich contents and traffic volume on Backbone and Access NW increase rapidly.
In result, ISP's cost also increases.

➡ **"CDN" may be the solution**

Backbone Area

CDN Server

CDN Server

CDN Server

CDN Server

Traffic on Backbone Area & Access NW is reduced

If ISP have own-provided access NW, they can locate the CDN server near User. And user can utilize the service comfortable.

Traffic Monitoring tool on Access NW

- **Very important to detect any NW trouble *without delay*, especially "Traffic monitoring"**

It's extremely hard to check all diversion traffic when NW trouble/maintenance, since there are many nodes and circuits on the own-provided Access NW

If there is the tool to monitor all traffic real time and alarm the congestion to operator before traffic saturation,
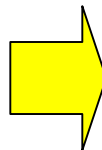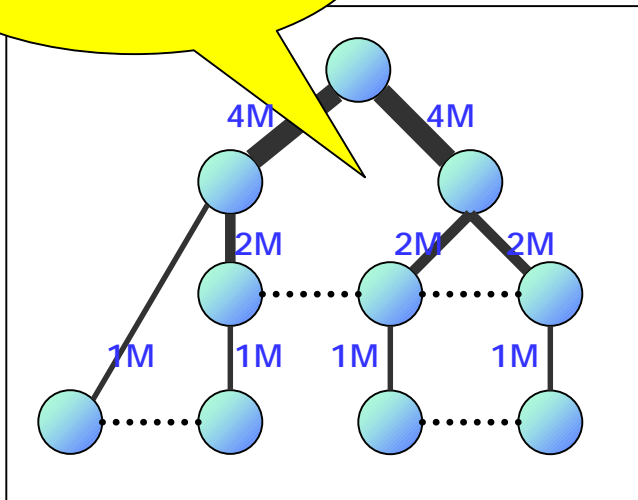Very Very Useful !! Terrific !!

Dangerous !

Detect and inform the traffic which Usage rate is 80% over
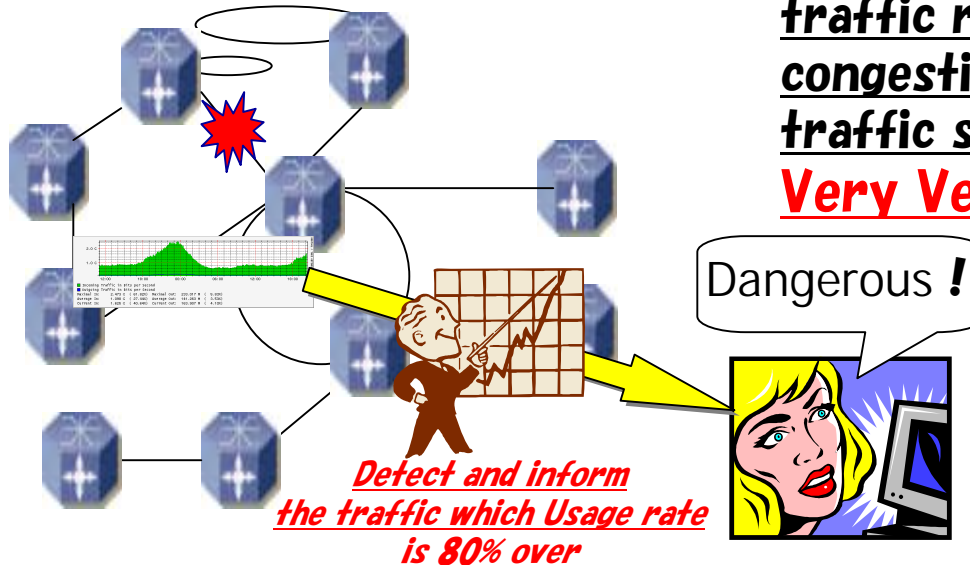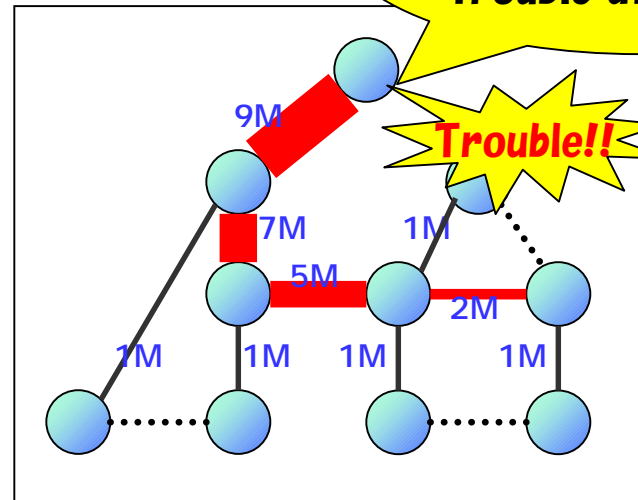
Traffic Analysis tool on Access NW

● **Very useful if there's "Traffic analysis tool" to simulate a diversion traffic before traffic congestion**

Can know the traffic flow by analysis tool can simulate a logical topology on demand

Normal Case

Can confirm the traffic flow when a trouble at a glance

4M 4M

2M 2M 2M

1M 1M 1M 1M

9M

Trouble!!

7M 1M

5M

2M

1M 1M 1M 1M

Multicast Route Analysis tool

- **Very useful if there's "Multicast route analysis tool" has the following function**

  ▶ Can know the end-to-end route from multicast source
  to a certain station

  ▶ Can know all affected stations easily before and when trouble

Multicast
Source

End-to-End
multicast route

Affected stations
by multicast trouble

Multicast
Source

Trouble
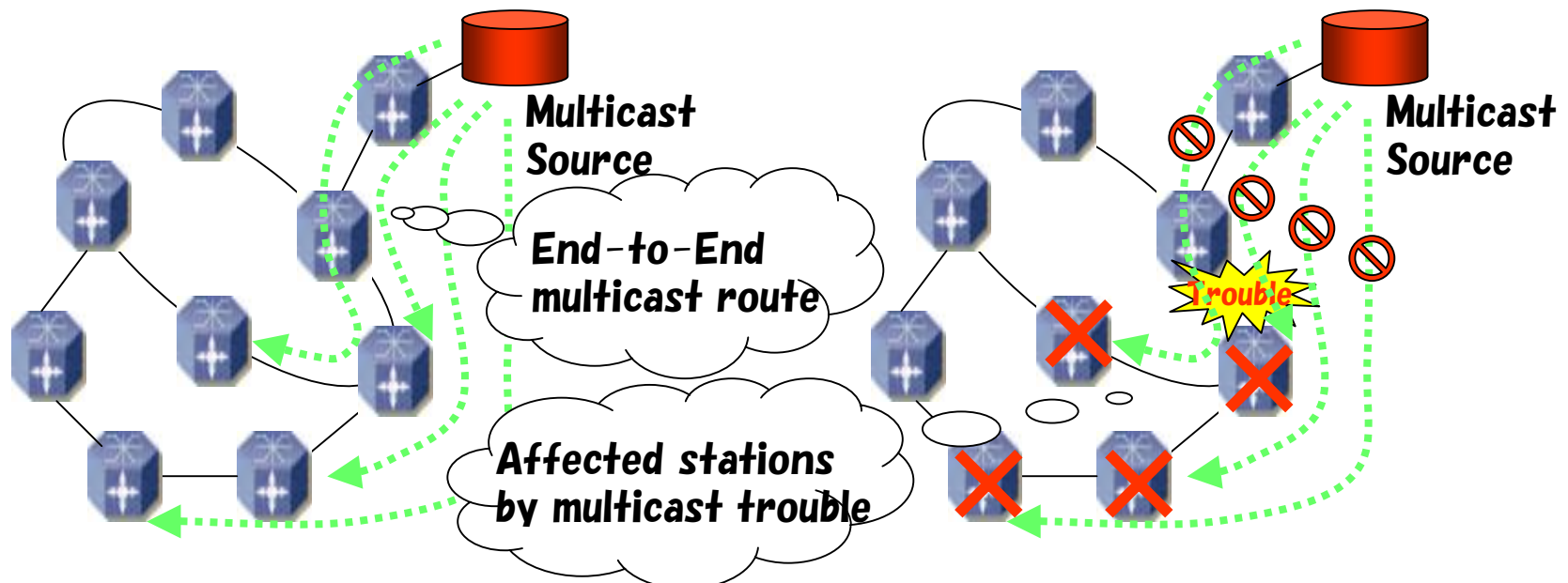
Summary

NW Design
- Supply Dark Fiber
- Ring Topology
  (2 Tier : Major/Miner)
- Selection hint of
  L3SW or Router

Route volume Reduction
- Effective routing policy by
  OSFP/BGP
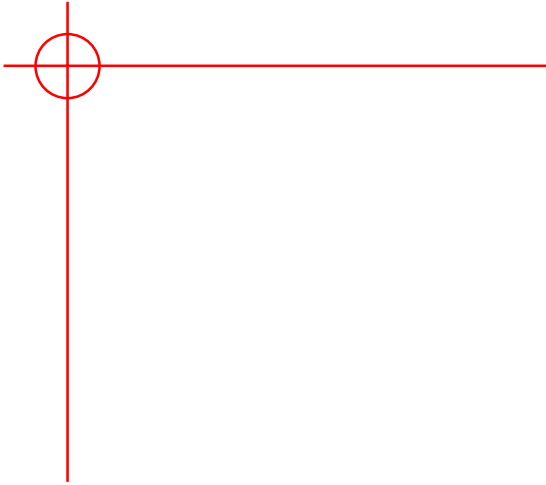- IP address assignment policy

Backbone NW

Access NW

ISP can provide various services by the own-provided Access without restriction
- Anyway Transit Service
- Multicast
- CDN
- QoS(DiffServe)

Operation hardship & Some Useful Tool
- Facility Management
- Traffic Monitoring
- Traffic Analysis
- Multicast Route Analysis

Thank you!