

vDC: Virtual Data Center Powered with AS Alliance for Enabling Cost-Effective Business Continuity and Coverage

Yuichiro Hei* Akihiro Nakao†

Tomohiko Ogishi* Toru Hasegawa* Shu Yamamoto‡

KDDI R&D Laboratories, Inc. *

The University of Tokyo †

NICT ‡

INM/WREN'10

27 April, 2010

Requirements for data centers

- In the cloud computing era, more and more applications and data are served from data centers
- Current DCs must be carefully designed to satisfy the requirements such as:
 1. **Business continuity**
 - Providing host mission-critical network services continuously even when catastrophic hardware failures and natural disasters occur
 2. **Coverage and performance**
 - Providing geographically diverse users fast and reliable access to the hosted services
 3. **Cost-effectiveness**
 - Minimizing cost for hosting services

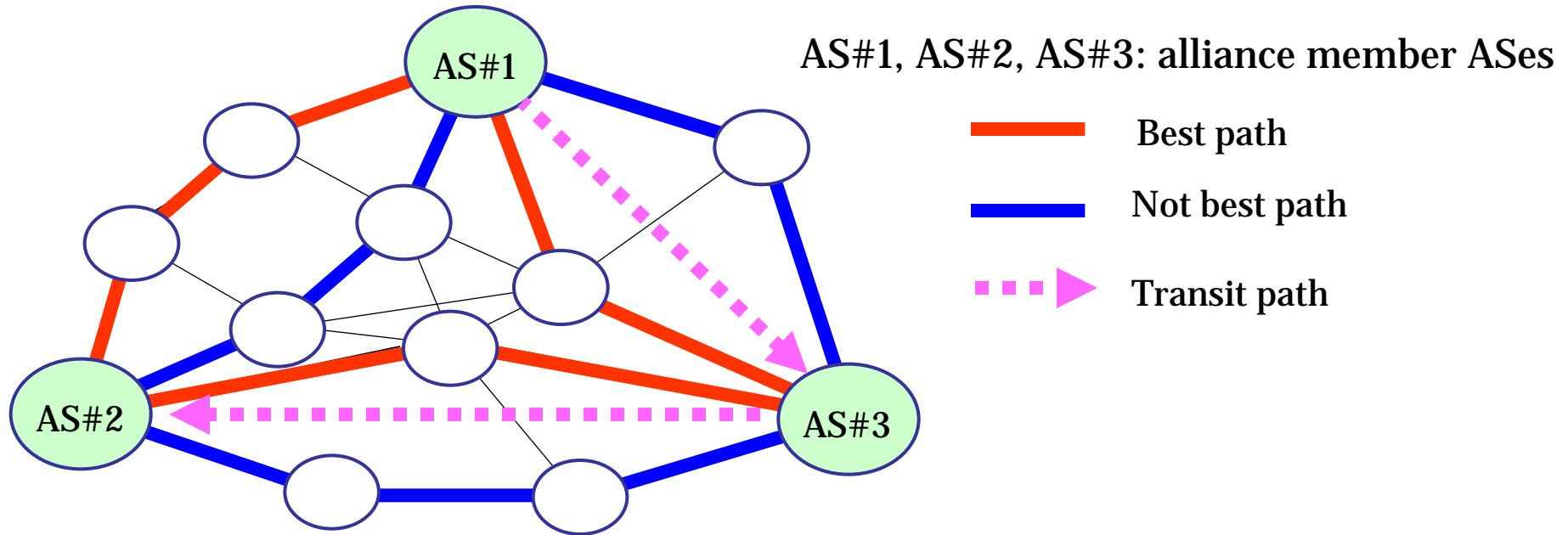
Approaches

- **Elephant providers**
 - **Constructing multiple DCs in different continents**
 - coverage and performance
 - **Provisioning robust backbone network to interconnect DCs on the dedicated/private/closed networks**
 - business continuity
 - **Cost?**
 - Facility and network costs are very high
 - e.g., a new middle-size DC(50K servers): over \$200M
 - laying a trans-oceanic submarine fiber: \$300M
- **Small regional providers**
 - **Almost prohibited from playing the game?**

Our proposal

- Goal: Conducting a cost-effective way for especially **small regional data centers to scale out into a global data center to satisfy the requirements**
- ***Virtual Data Center (vDC)***
 - Like a meta data center consisting of multiple geographically distributed data centers
 - Geo-distribution → **coverage and performance (fast access to DCs)**
 - Using the existing data center infrastructures
→ **cost-effectiveness**
 - Robust communication among DCs over the Internet powered with *AS alliance* (next slide)
→ **cost-effectiveness and business continuity**

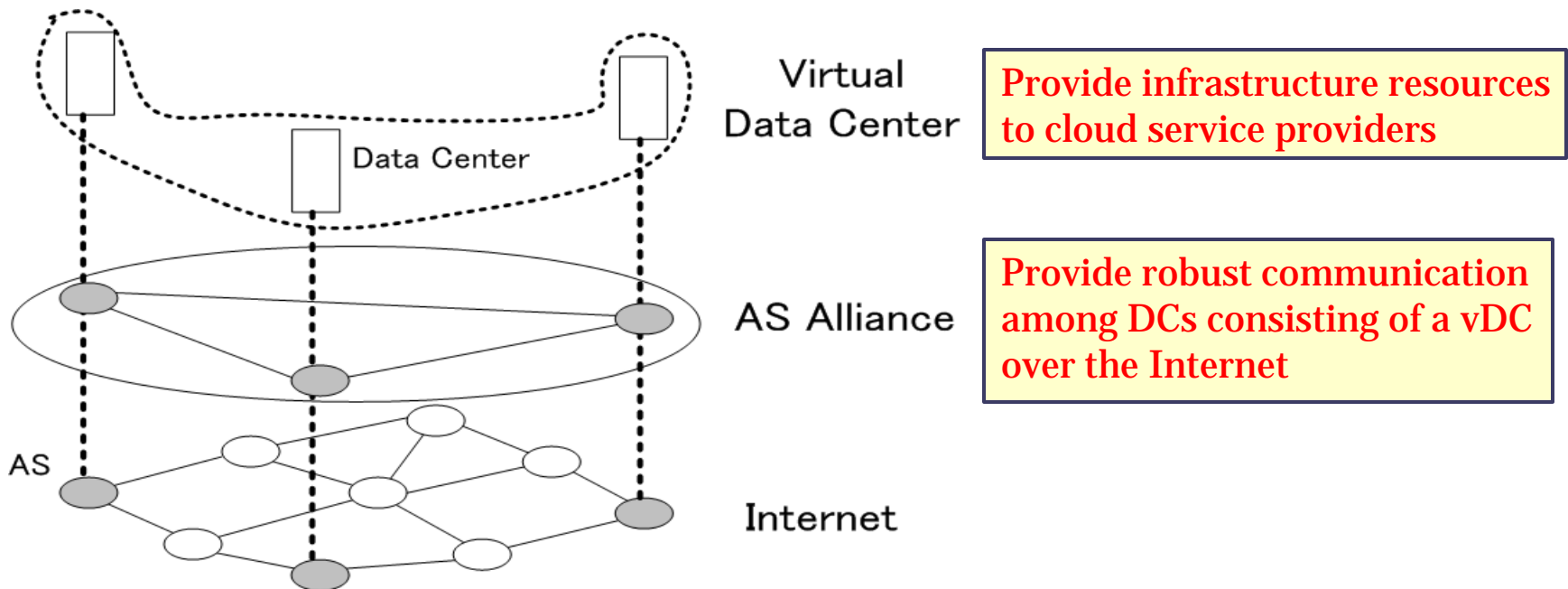
AS alliance



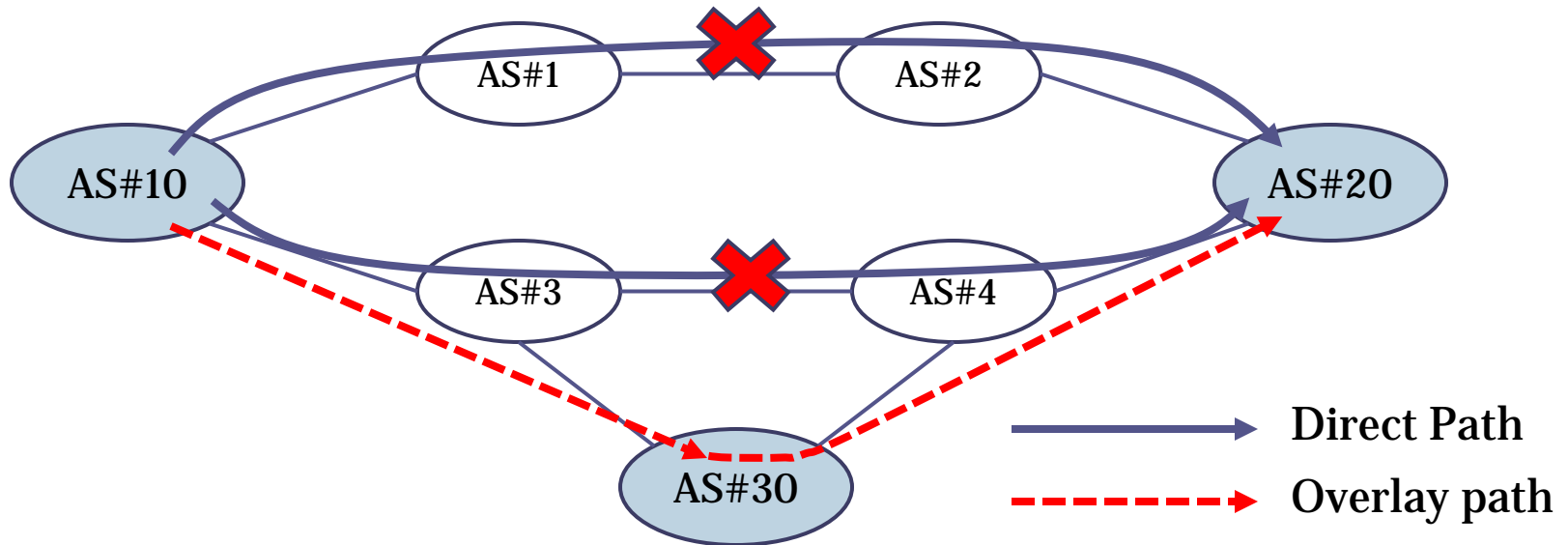
- Each member AS shares BGP routes (not only the best paths, but not the best ones), and computes multiple AS paths among them
- Member AS provides the other members with a transit between them
- Each member AS tries to find AS paths that are as **disjoint** as possible with each other
 - Avoiding a situation that multiple paths become vulnerable to a single failure for ensuring **robust communication over the Internet**

vDC over AS alliance

- vDC is running over AS alliance
 - A cloud service provider purchases resources from multiple data centers
 - Separating cloud service providers and data center providers
 - These data centers are consolidated into a virtual data center over the Internet
 - AS alliance helps to provide robust communication among the alliance member ASes, i.e., data centers consisting of a vDC

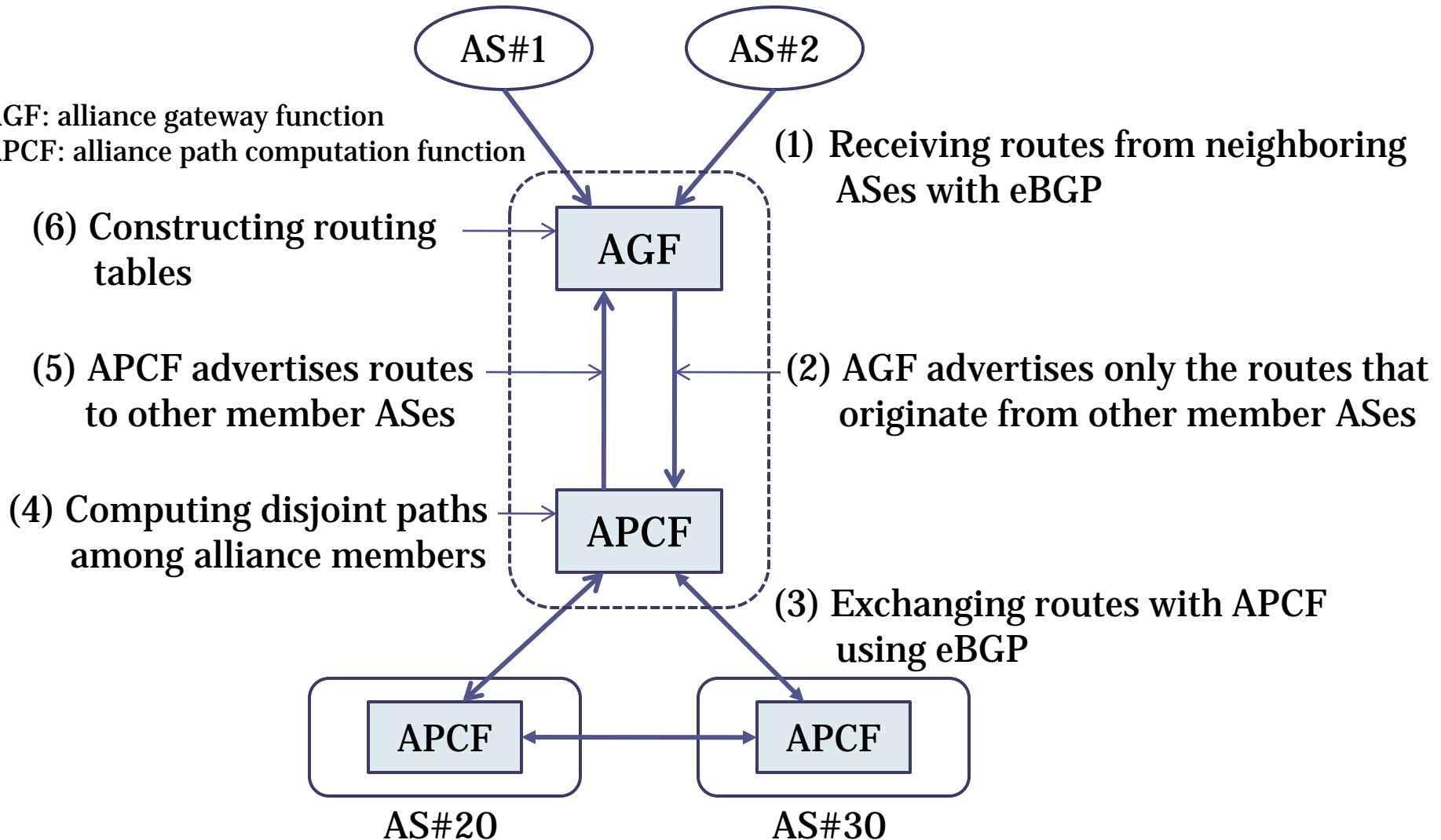


Paths among the AS alliance members



- Path from AS#10 to AS#20 from the viewpoint of AS#10:
 - Direct path: a normal BGP path
 - Overlay path: a path via an other member AS
- Even if all direct paths fail, AS#10 can continue to communicate with AS#20 over the overlay path with help from AS#30

Inside a member AS

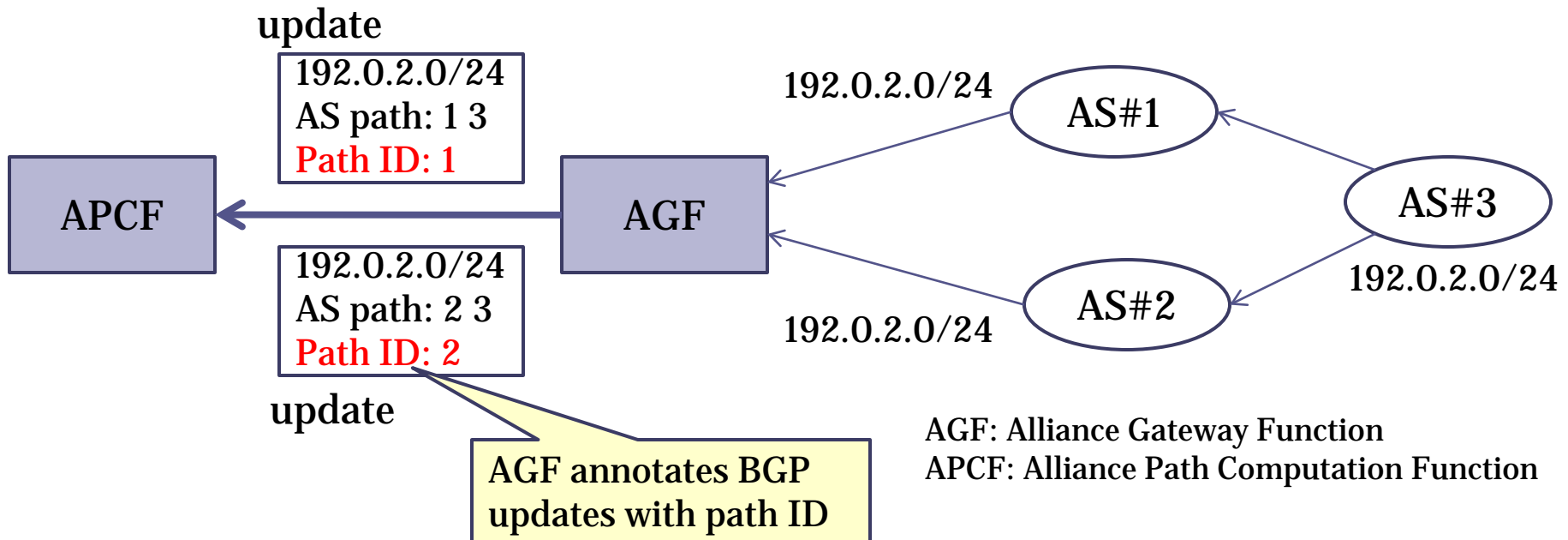


Slightly extension to BGP

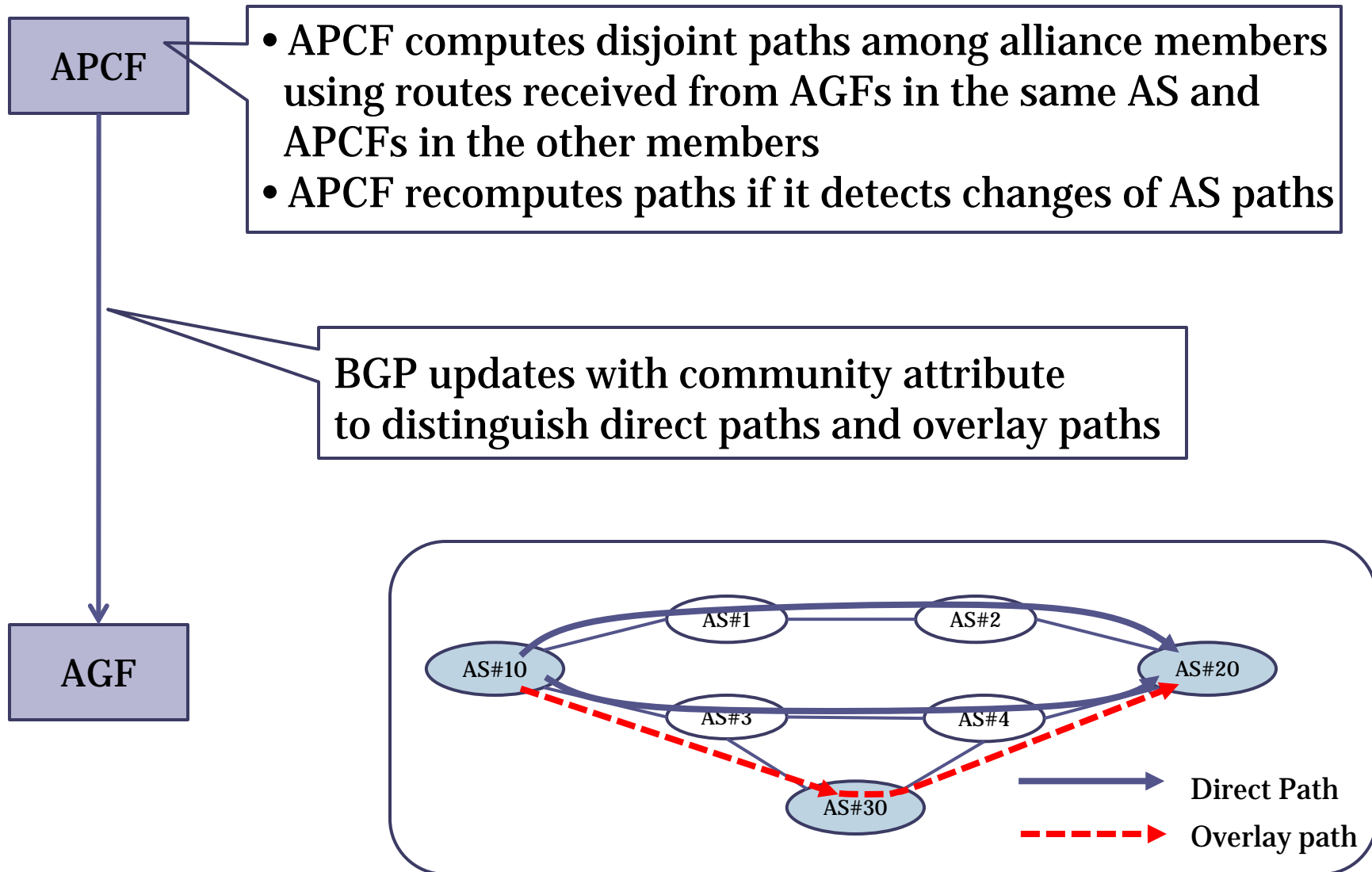
- **Multiple routes**
 - Need to distinguish multiple updates destined to the same prefix
- **Path computation and update**
 - APCF computes paths among alliance members and advertises them to AGFs with BGP updates
- **In network operational perspective, there is not so much difference between the normal BGP operation and the BGP one with AS alliance**

How distinguish multiple routes

- AGF may receive multiple routes to a same prefix of the members from its neighbors
- AGF advertises each routes to APCF without selecting the best routes for APCF to collect as many AS paths as possible
- To distinguish these routes, AGF annotates BGP updates with **path identifiers**

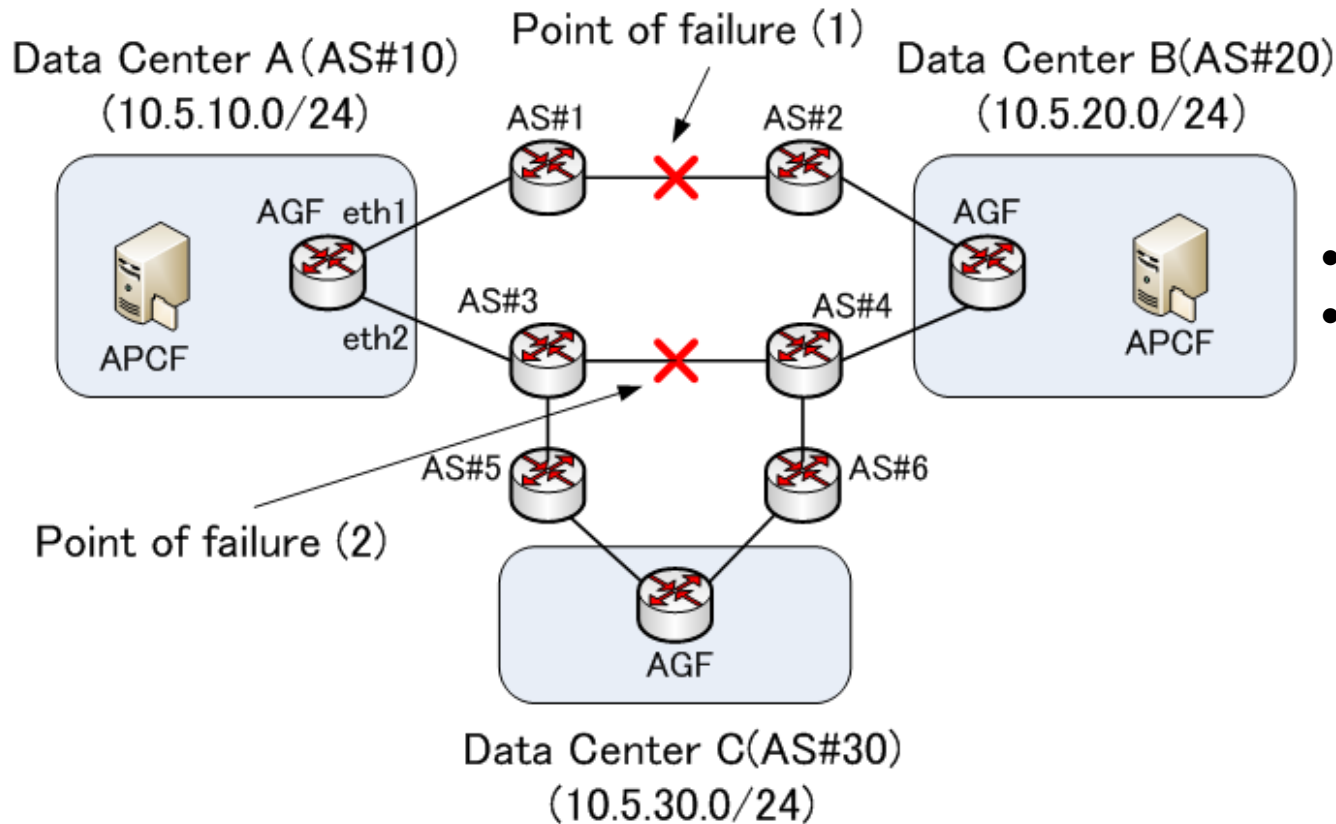


Path computation and update



Prototyping and evaluation

- Implementing a prototype of AS alliance on Linux boxes
- Based on Quagga BGP routing daemon
- Evaluation topology



- DCs are edge ASes
- AS#10, #20, #30 form an AS alliance

Routing table in AS#10

- Normal state

```
root@PC-05-E ~]# ip route show table 10
10.5.20.2 proto zebra metric 10
    nexthop via 172.31.2.3 dev eth1 weight 1
    nexthop via 172.31.3.3 dev eth2 weight 1
10.5.30.2 via 172.31.2.3 dev eth1 proto zebra metric 10
10.5.30.0/24 via 172.31.2.3 dev eth1 proto zebra metric 1
10.5.30.0/24 dev tun10to20 proto zebra metric 10
10.5.20.0/24 proto zebra metric 1
    nexthop via 172.31.2.3 dev eth1 weight 1
    nexthop via 172.31.3.3 dev eth2 weight 1
10.5.20.0/24 dev tun10to30 proto zebra metric 10
default via 10.5.10.6 dev eth0 proto zebra
root@PC-05-E ~]#
```

Direct paths to AS#20

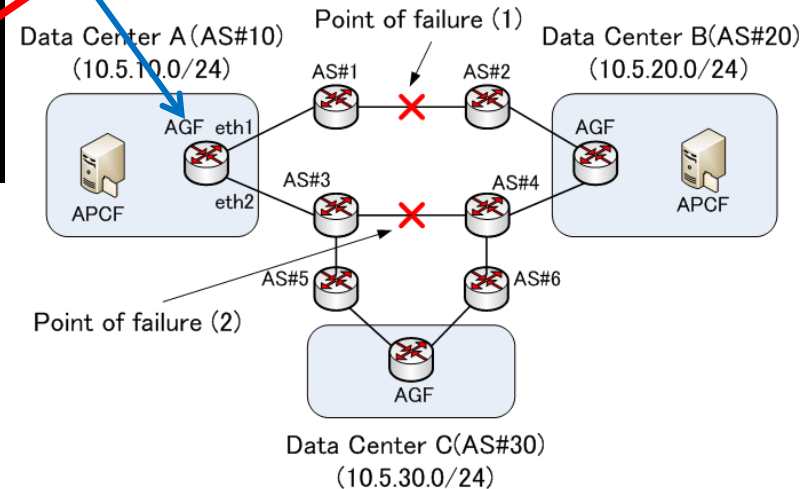
Overlay path to AS#20

- Failure state

```
root@PC-05-E ~]# ip route show table 10
10.5.20.2 via 172.31.2.3 dev eth1 proto zebra metric 10
10.5.30.2 via 172.31.2.3 dev eth1 proto zebra metric 10
10.5.30.0/24 via 172.31.2.3 dev eth1 proto zebra metric 1
10.5.30.0/24 dev tun10to20 proto zebra metric 10
10.5.20.0/24 dev tun10to30 proto zebra metric 10
default via 10.5.10.6 dev eth0 proto zebra
root@PC-05-E ~]#
```

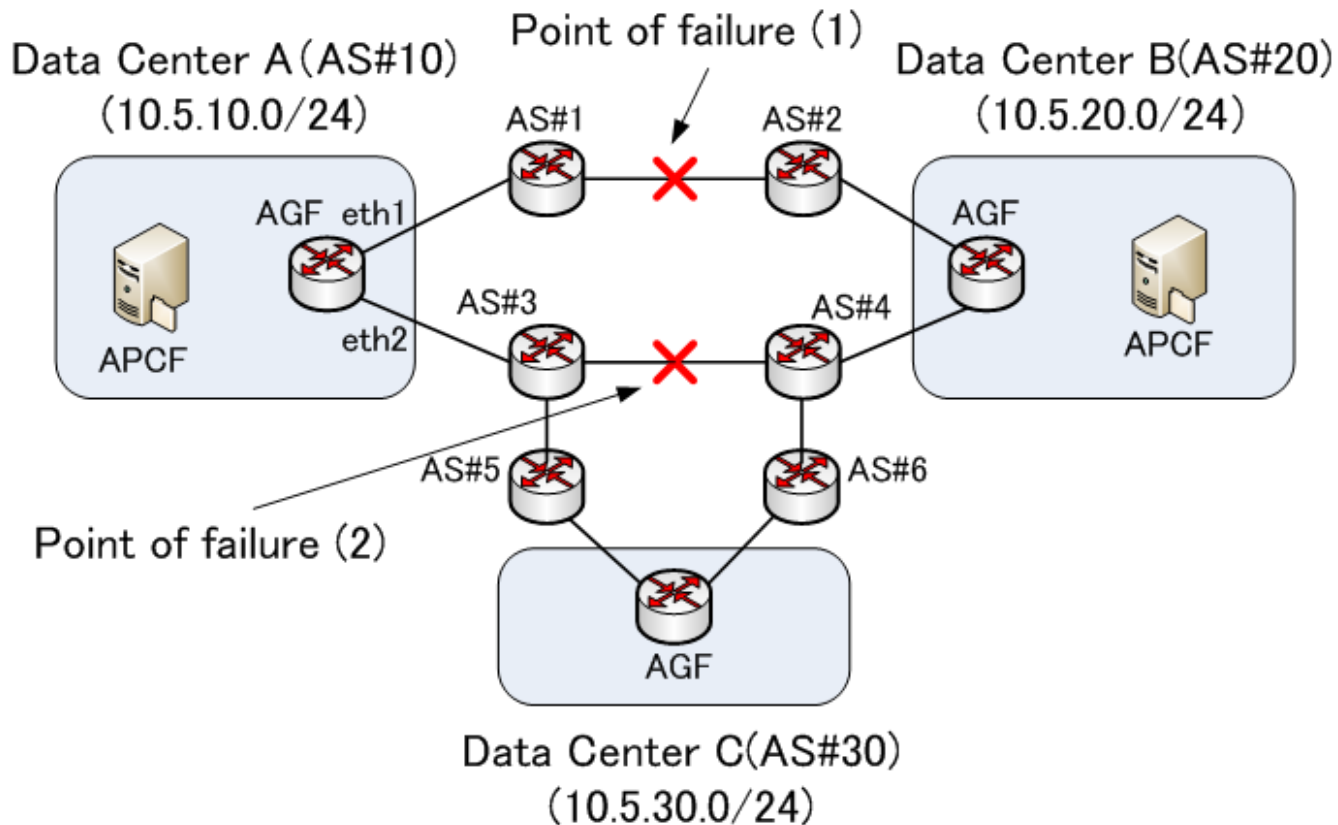
Overlay path to AS#20

Direct paths to AS#20 disappear, but AS#10 can continue to communicate with AS#20 because a overlay path to AS#20 via AS#30 is alive



Demo

- Comparing the cases that AS alliance is formed and not formed
- Showing the *ping* from AS#10 to AS#20 when the link failures occur



Demo

Conclusions

- Proposing a *virtual* data center (vDC) consisting of multiple geographically distributed data centers over the Internet
- Presenting the practical design of an architecture of vDC over AS alliance
- Our feasibility study shows that vDC with AS alliance can provide the robust communication among data centers forming a vDC