TPIX

How to connect two IXes?

Konrad Plich, Robert Woźny
Kraków, September 30, 2011

Twój świat. Cały świat.

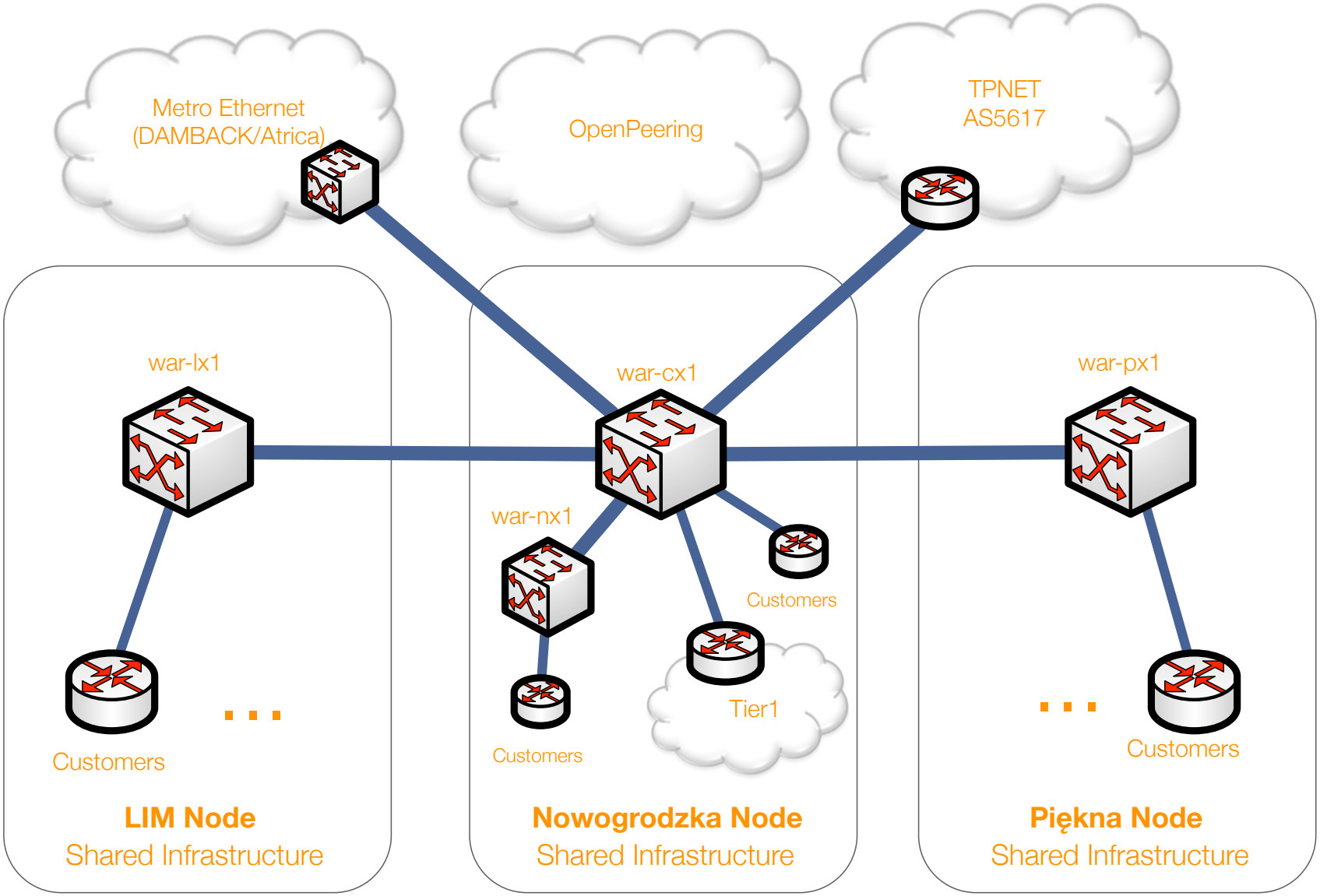
Internet Exchange Point at a glance

- Why to peer? save money, local traffic stays local, better performance...
- Private peerings: does not scale; solution: Internet Exchange Point
- Simple concept: any place where ISP come together to exchange traffic
- Layer2 (Ethernet) or Layer3 (router based) exchange
  - Layer2: ISPs free to set up peering agreements with each other as they wish
  - Layer3: marketing of the transit ISPs instead of IXP
- Each ISP participating in the IXP connect a router...
  - router needs to be able to run BGP! (using public ASN, NOT private one)
- Route Servers: help to scale routing
  - provided as service, usage is NOT mandatory
  - only two eBGP sessions rather than N
  - simplified routing configuration management on ISPs routers
- Community Based Policies: Bi-Lateral Peering, Multi-Lateral Peering, Mandatory Multi-Lateral Peering

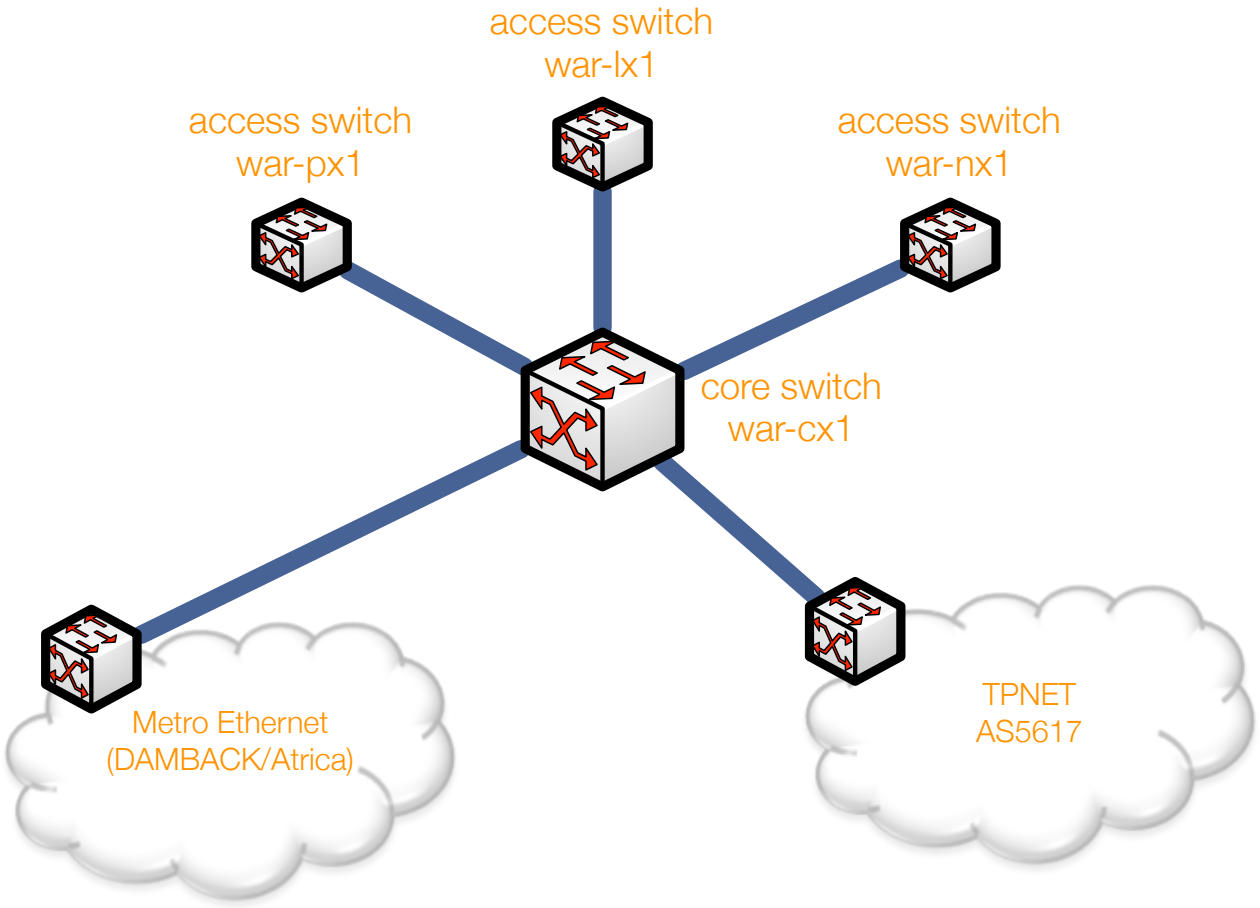
TPIX: integration of services on the Ethernet platform

In the collocation facility Nowogrodzka, Piękna and the LIM building we installed platform based on Ethernet switches...

- cheap access ports - 1GEth and 10GEth (10GEth available only on the core switches)
- possibility to exchange Internet traffic (OpenPeering service)
- possibility to create IP Interconnection with TPNET network
- private VLANs between users of the platform (without bandwidth restrictions)
- possibility to use TP's Metro Ethernet platform to provide services in Poland

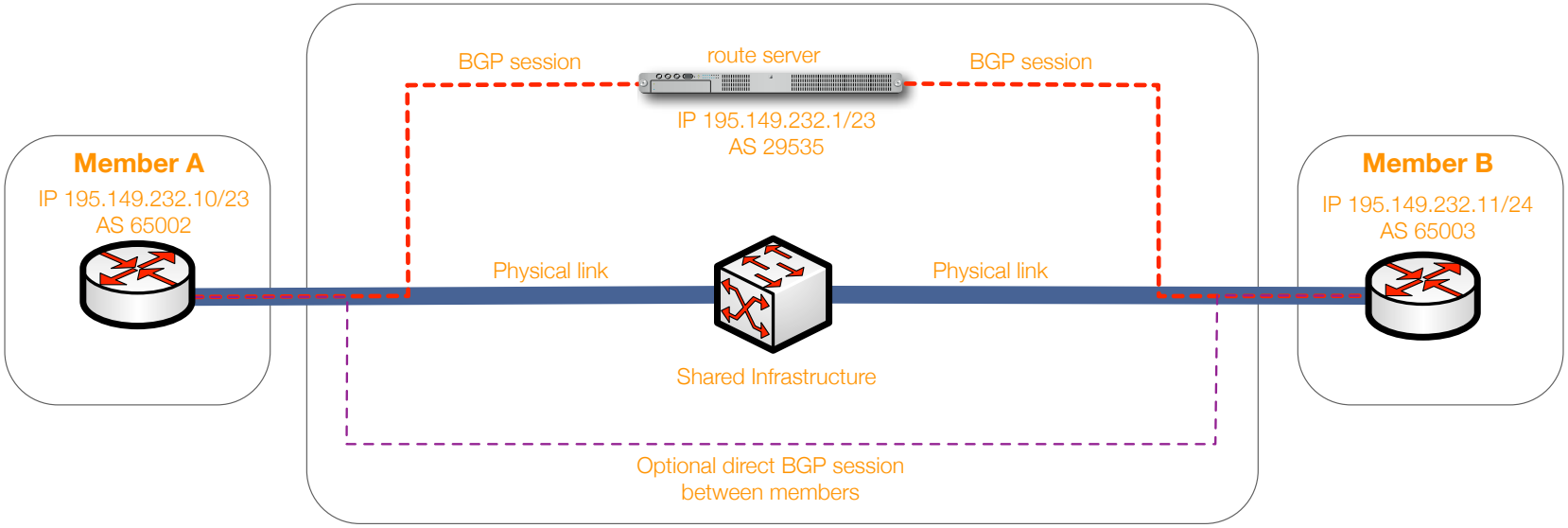

Metro Ethernet
(DAMBACK/Atrica)
OpenPeering
TPNET
AS5617
war-lx1
war-cx1
war-px1
war-nx1
Customers
Customers
Customers
Tier1
Customers
LIM Node
Shared Infrastructure
Nowogrodzka Node
Shared Infrastructure
Piękna Node
Shared Infrastructure

Physical network topology



The access switches gives us cheap 1GEth ports. The whole TPIX's logic is done on the core switch.

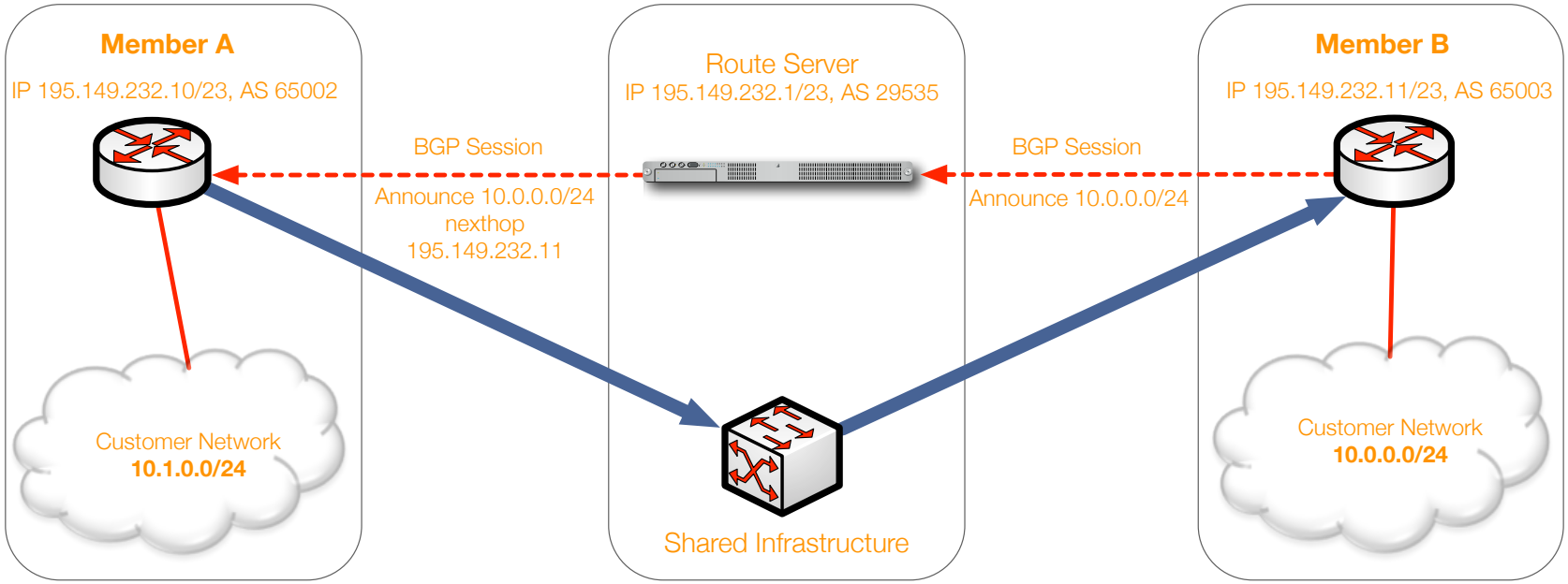
How IXP works!



Every member has physical connection to Platform at a particular speed. Through this connection every member receives a static IP address with the same network mask. The participants have direct visibility towards the IP addresses and are able to establish direct BGP session among each other. The traffic between participants is switched instead of routing. Platform also provides Route Server service for simplicity of establishing BGP sessions

For simplicity on the scheme there is an example of two members...

How IXP works!



Traffic: from **Member A** router goes directly to 195.149.232.11 (**Member B** router)

The scheme illustrates how the Route Server is functioning and how does member A find the network of a client of member B (10.0.0.0/24) and when receives traffic for it sends it through the shared switching infrastructure.

For simplicity on the scheme there is an example of two members...

Connecting IXes: Problems?!

- Political Problem
  - Neutrality or Ownership?
- Sales Problem
  - who will pay for that?
- Technical Problems (see on the next slide)

Connecting IXes?

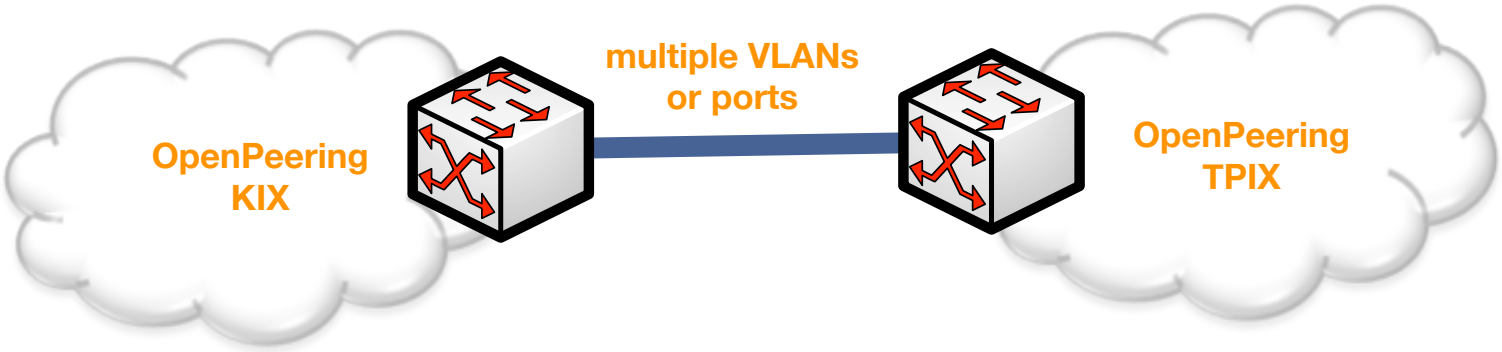
What can we do:

- provide transit between the IXP based on Layer2 transmission
- connect switching fabrics (problems, problems, problems...)
- don't connect (neither the switching fabrics, nor customers)
- use BGP router to provide Layer3 transit between IXP's
- use smart Layer2/Layer3 connectivity

Connecting IXes: Layer 2

The safest way to allow customers of one IXP to access the second IXP...

- port or VLAN
- each customer gets their own port (or VLAN)
- simplicity - it's just layer2 service
- but... it's difficult and expensive to maintain (many cables or vlans)



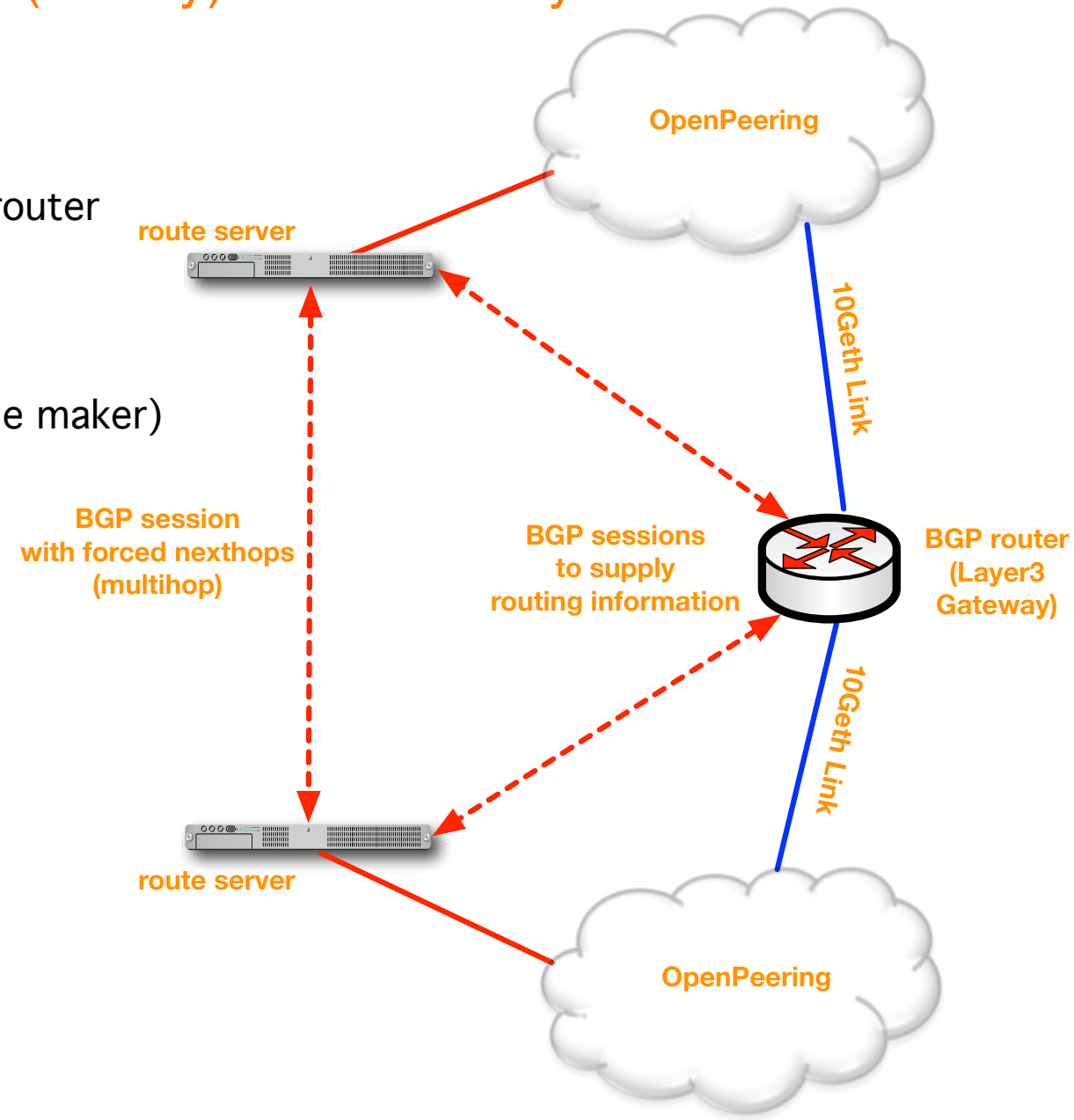
Connecting IXes: (history) “smarter” Layer 2/3

Pros:

- use only one BGP capable router
- separation of Layer2

Cons:

- causing problems (blackhole maker)



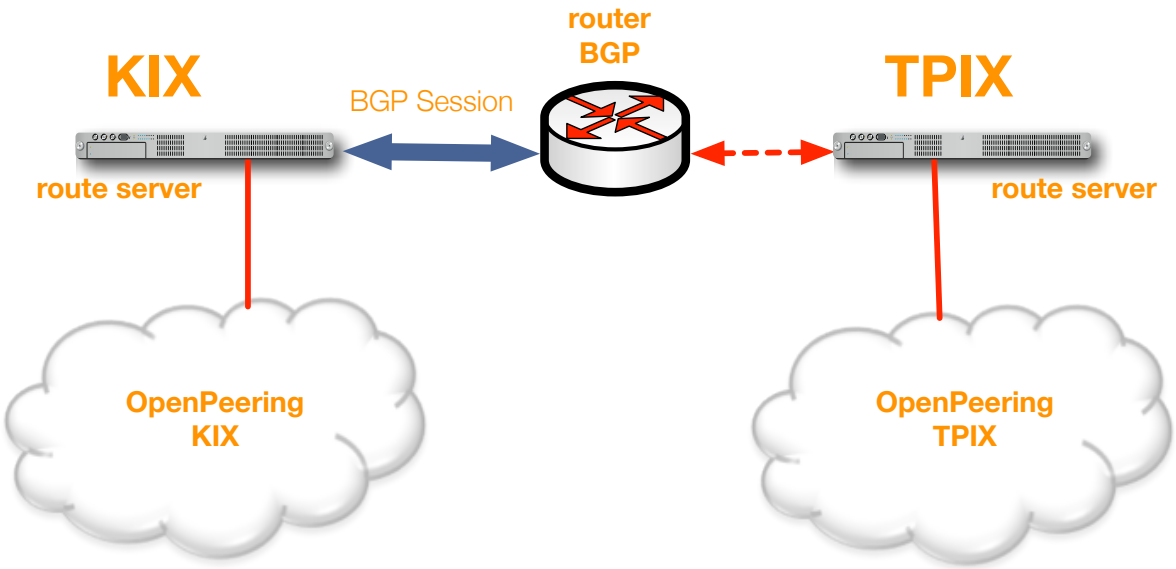
Connecting IXes: Layer 3

Pros:

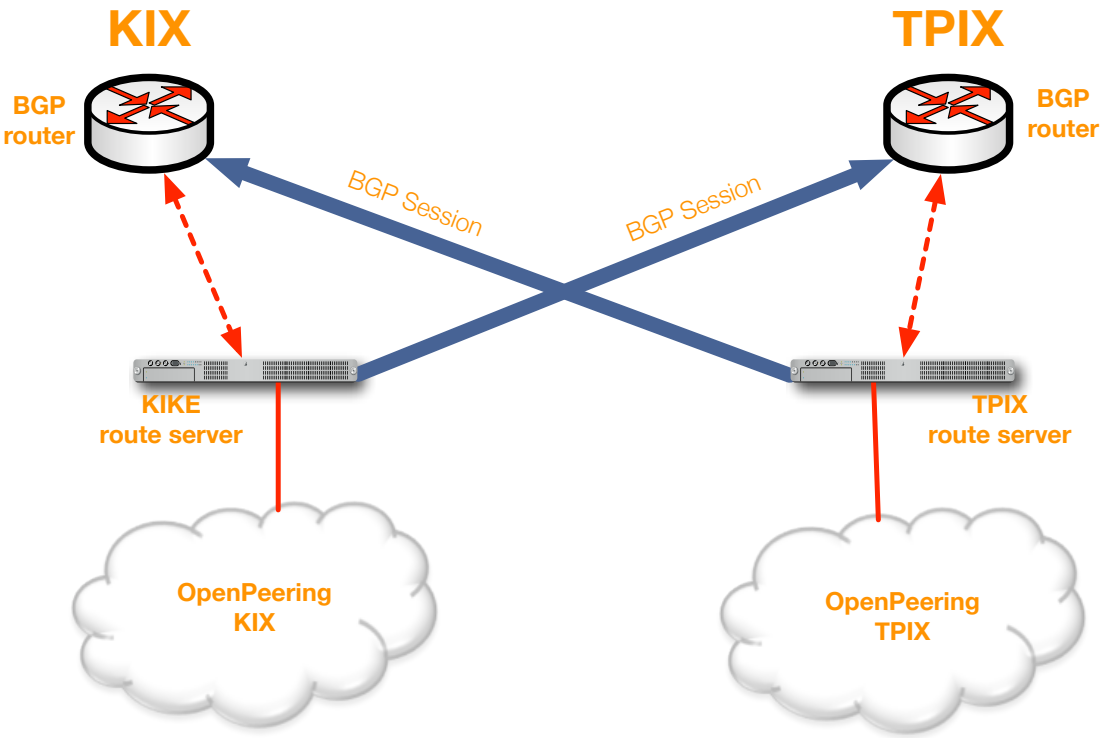
- use only one BGP capable router
- separation of Layer2

Cons:

- discrimination: additional ASN in the path (on one side).



Connecting IXes: “smart” Layer2/Layer3



Pros:

- separation of Layer2
- non discrimination policy: don’t add transit ASN to the path.

Cons:

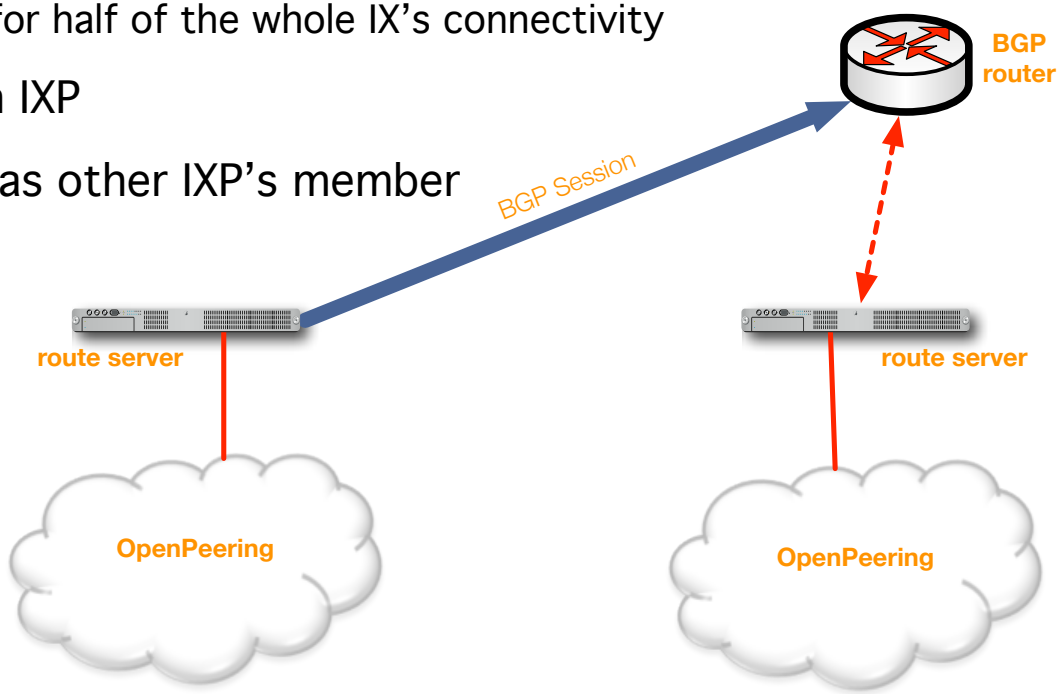
- lack of redundancy: the failure of one IXP causes the unavailability of the customer’s network in both IXPs
- asymmetric traffic

Connecting IXes: “smart” Layer2/Layer3

- two VLANs (or ports)

```
995   IX: OP-KIKE                          active
996   IX: KIKE-RTR                         active
```

- two BGP capable routers
  - routers with the knowledge of routing to the networks available at a given IXP
  - shared ASN between transit router and route server
  - every routers is responsible for half of the whole IX’s connectivity
- route servers service in each IXP
- the transit router is treated as other IXP’s member



Connecting IXes: “smart” Layer2/Layer3

- transit routers don’t announce the reachability information to the other OpenPeering
- transit router announce the reachability information to IXP’s route servers with the forced next-hop self

```
rs> show bgp route protocol R232_5
* 91.221.44.0/23     100      0        >195.149.232.5     34393 i
```

- route server announce the reachability information to the other connected members without adding own ASN

```
 3  mx1.waw.cdp.pl (62.111.157.153)  0.755 ms  0.884 ms  0.715 ms
 4  war-cx1.tpix.pl (195.149.232.5)  0.975 ms  0.951 ms  0.717 ms
 5  kix-waw-sw1.k-ix.net.pl (91.212.9.1)  1.251 ms  1.103 ms  1.016 ms
[..]

213.5.248.0/21     *[BGP/170] 15:15:17, localpref 100
                      AS path: 50188 I
```

Thank You ☺

Twój świat. Cały świat.