Distributed lambda-based internet exchange (IX) using GMPLS controlled photonic cross connects

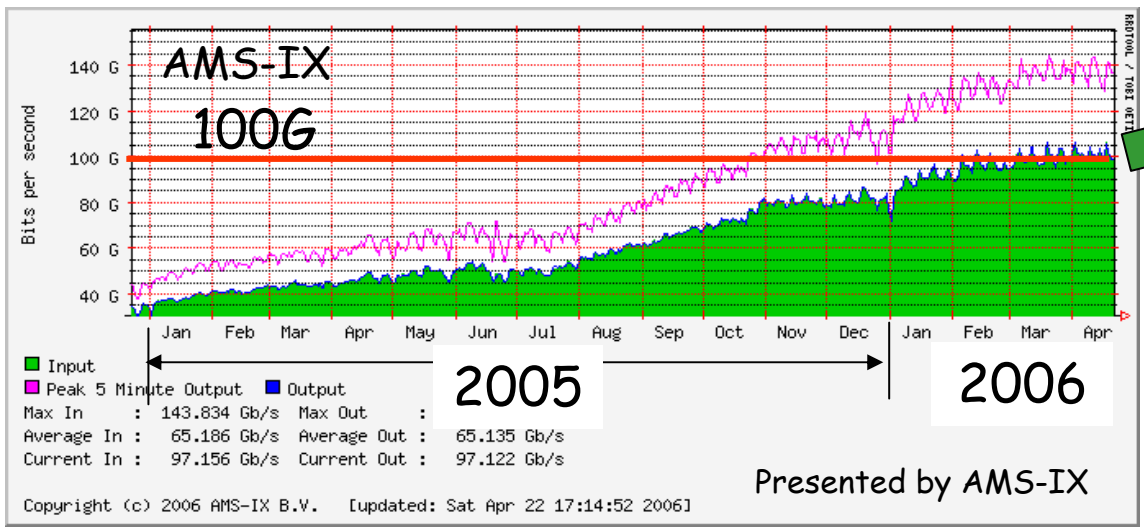
Takehiro Tsuritani, Masanori Miyazawa,
Tomohiro Otani, Hedeaki Tanaka

KDDI R&D Laboratories Inc.

2006 iPOP IP+Optical Network

Background

- Traffic load via major Internet exchange (IX) points
  - Has doubled every year.
  - Has reached around 100Gbps.

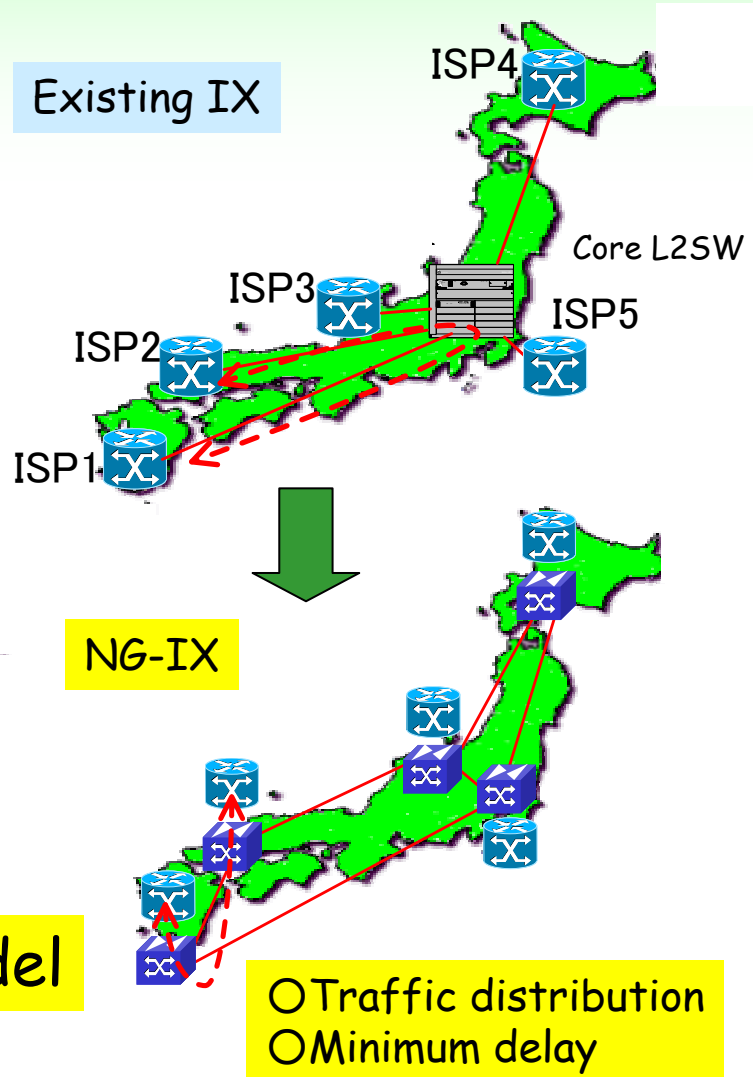


Terabit at 2010 !?

- Internet exchange (IX) is required to transact more than Terabit/s traffic in the near future !!

Requirements for next generation IX

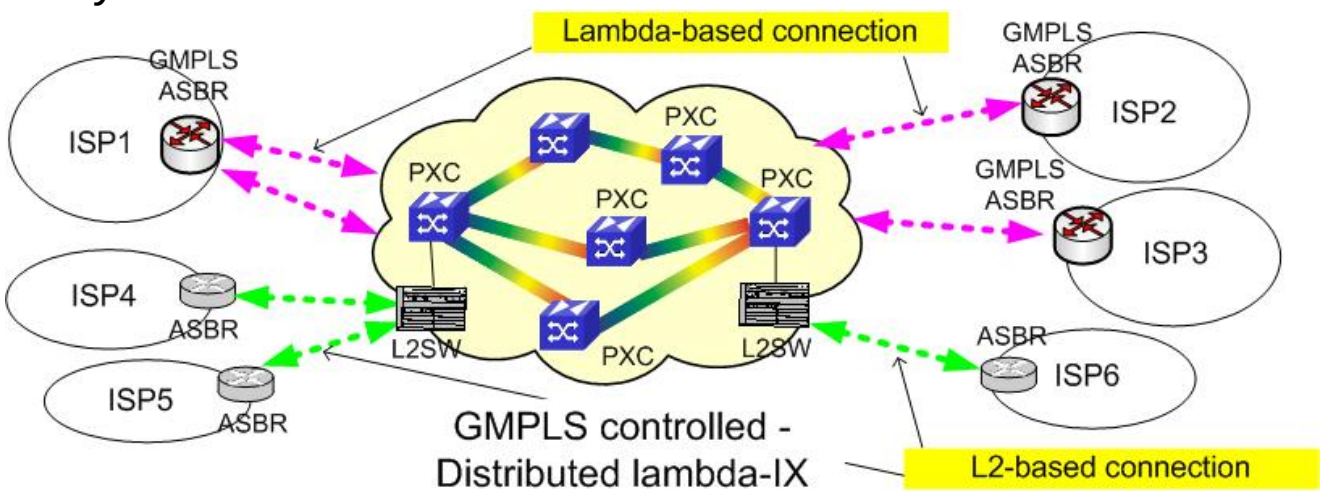
- Existing Layer 2 (L2) -based IX
  - has core L2 switches (L2SWs) as IX nodes only in major cities, sometimes referred to “concentrated model”
- Issues of existing IX
  - Traffic concentration into a core node
  - Large transmission delay
- Requirements for next generation IX
  - Traffic control and management of high-capacity (OC48/192/768/10GbE/….) traffic exchange



Distributed lambda-IX model

Distributed lambda-IX (1)

- Architecture
  - Basically consists of some GMPLS-enabled PXC equipments and some GMPLS-enabled L2 switches
    - Connected by using transparent transport equipments such as D(C)WDM.
  - Enables to directly interconnect between ISPs with lambda-based connection
    - With any capacity (10Gbps/10GbE/OC48/GbE..) between ISPs
  - Enables to also provide L2-based connections by connecting ASBRs to a L2SW like existing L2-IX so as to offer the interconnection with the finer granularity than one of lambda-based interconnection between ISPs



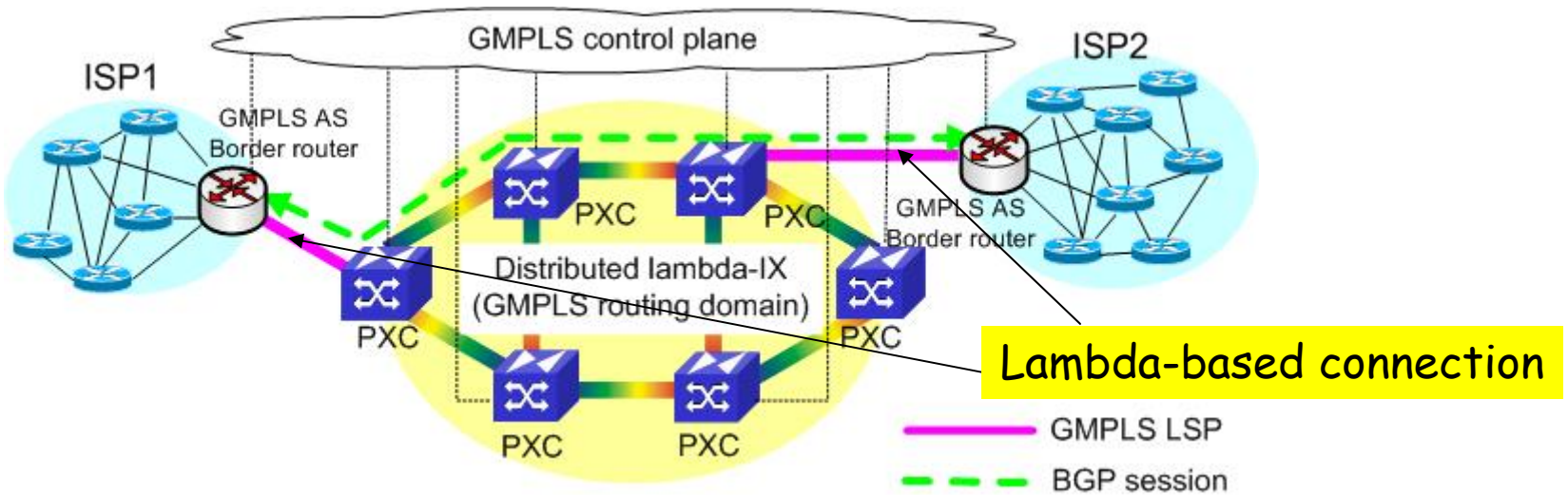
Distributed lambda-IX (2)

- Distributed lambda-IX model enables
  - To control the traffic with high capacity via lambda-IX depend on IX policy
    - by utilizing GMPLS Traffic Engineering techniques
  - To Improve the reliability and resiliency for IX
    - By using GMPLS fault recovery mechanism over optical layer
  - To manage the traffic via lambda-IX with L2-based connections
    - By interworking between GMPLS control plane and traffic monitoring of L2SW

Lambda-IX can
-achieve the enhancement of existing L2-IX
-be migrated smoothly from existing IX

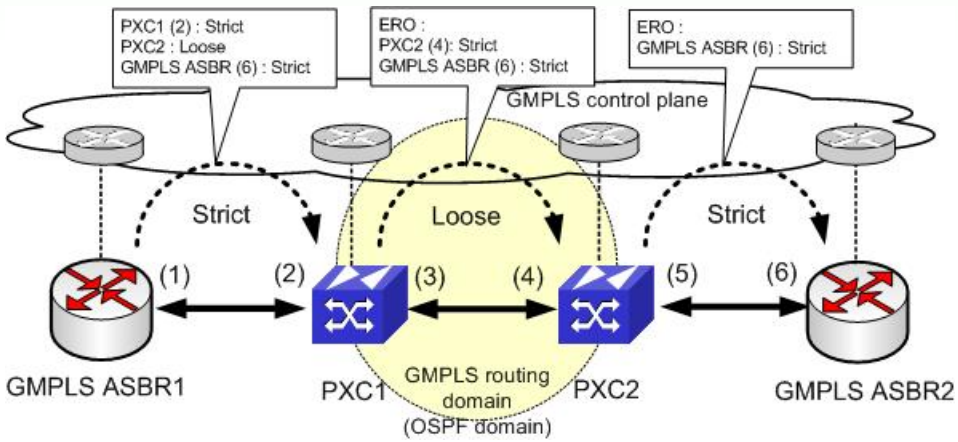
Distributed lambda-IX with direct lambda connection

- Requirements for lambda-IX with direct lambda connection
  - Preparation of GMPLS implemented ASBR and interconnection between ASBR and PXCs over GMPLS control plane for ISPs
  - Routing separation between the IX domain and ISP's domains
    - Core PXC nodes hold only topological information without any routing information exchanged between ISPs
  - BGP session establishment by using head-end and tail-end IP addresses of Logical interfaces of GMPLS-LSP tunnel setup between ASBRs



Loosely routed GMPLS-LSP (1)

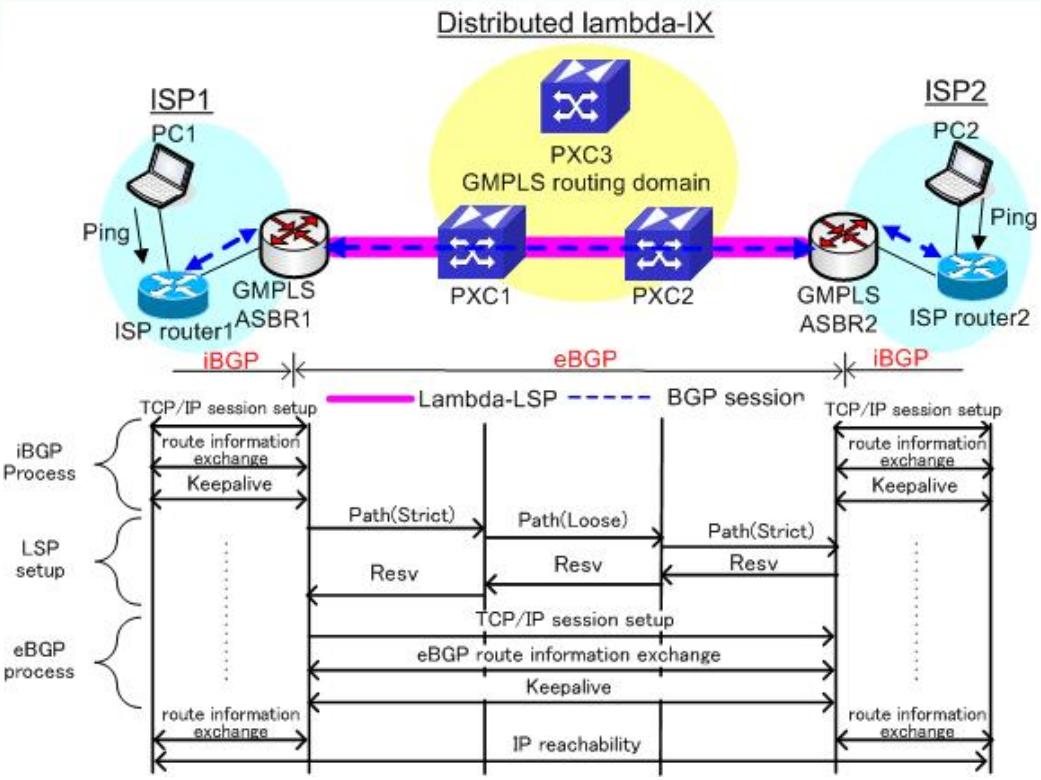
- GMPLS routing
  - Use OSPF-TE just only in lambda-IX domain
  - No routing information exchange between ASBRs and Lambda-IXs
    - Ingress ASBR statically configures routing information of adjacent PXC connected
- GMPLS signaling (RSVP-TE)
  - Ingress ASBR designates strict route for adjacent PXC
  - Ingress ASBR designates loose route in lambda-IX

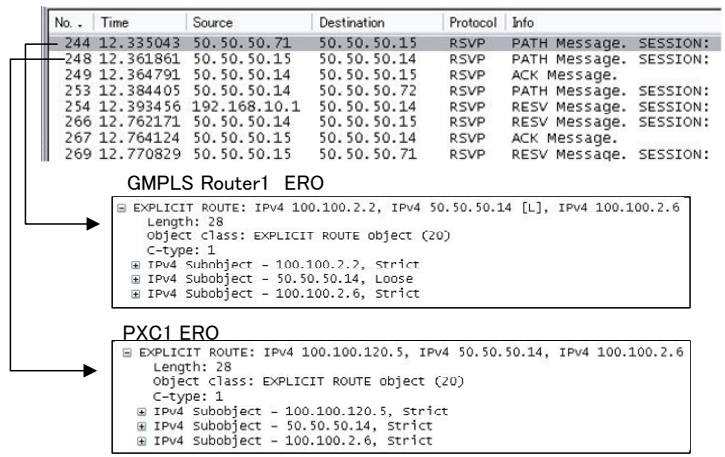


-Can hide topology and routing information of IX toward ISPs
-Can dynamically create suitable routes over lambda-IX

Loosely routed GMPLS-LSP (2)

- Evaluation results

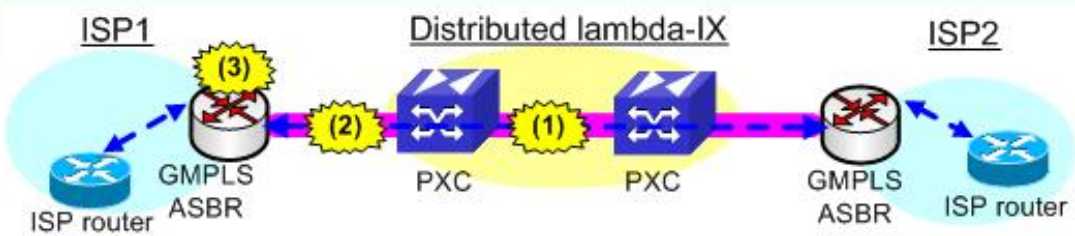




We successfully demonstrated the establishment of
-the eBGP session over the loosely routed GMPLS-LSP
-inter-domain IP connectivity between two peer ISPs

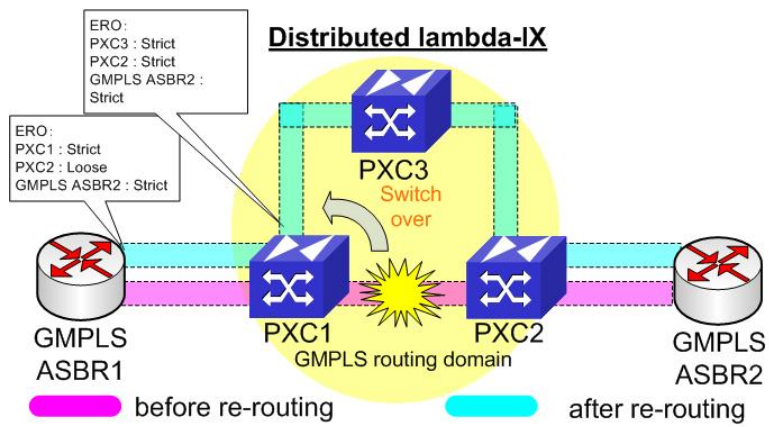
Lambda-based fault recovery (1)

- Assumed failure points
  - ***(1)In distributed lambda-IX***
  - (2)Between ASBR – PXC
  - (3)In GMPLS ASBR



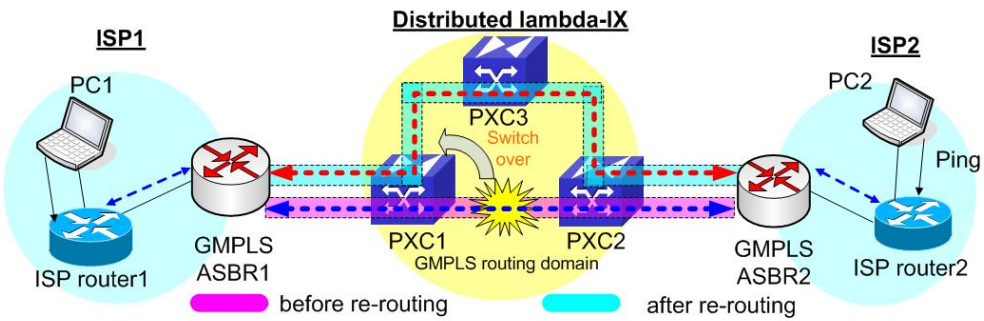
- L1 restoration sequence in the case of (1)
  1) PXC2 detects Loss of Light (LOL)
  2) PXC2 notifies GMPLS ASBR1 of a failure of the TE link by PathError messages
  3) GMPLS ASBR1 initiates re-signaling

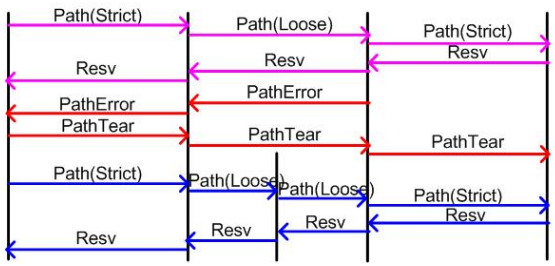
     -PXC2 recalculates CSPF to dynamically re-route in lambda-IX

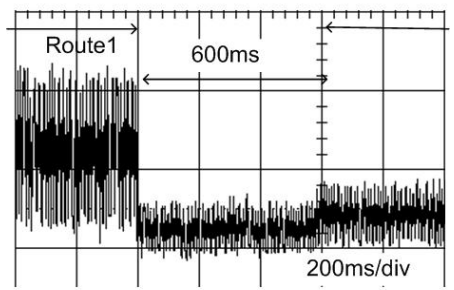


Lambda-based fault recovery (2)

- Experimental results
  - The LSP was successfully and dynamically restored to the back up route
    - Less than 600ms over optical layer
    - The number of packet loss transmitted between PCs : 1 packet
    - Without impact of the BGP session between GMPLS ASBRs on a failure





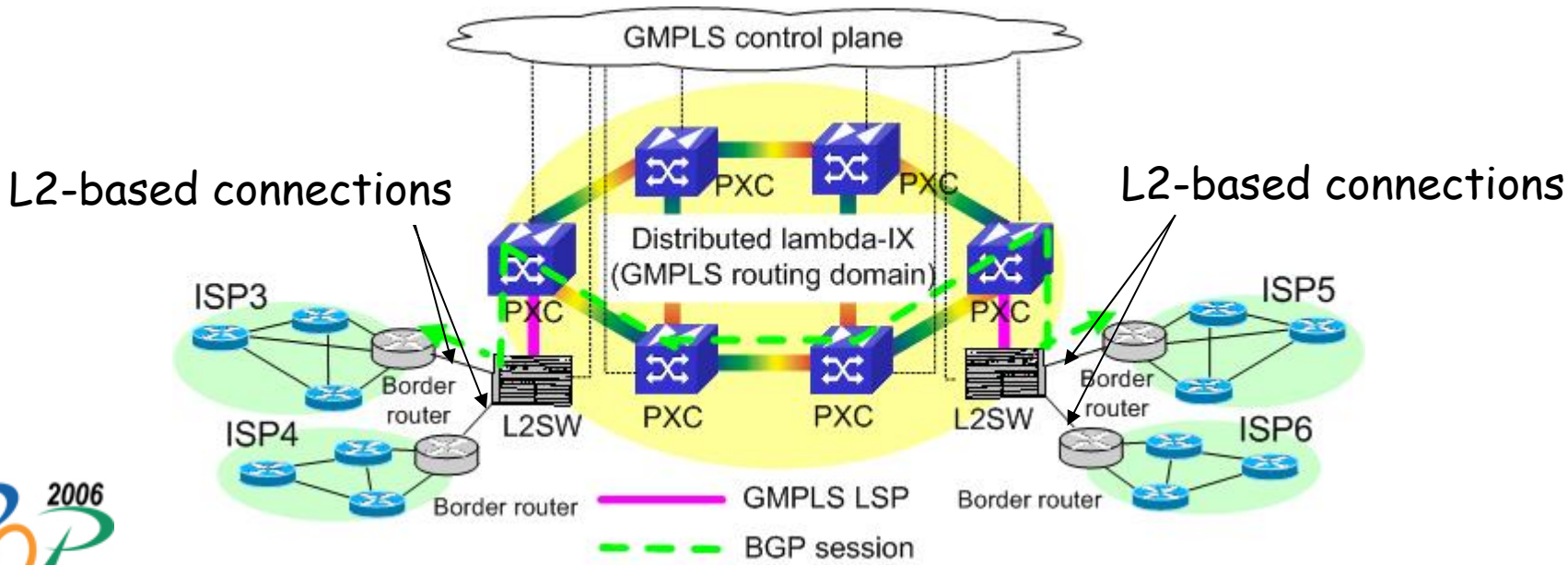


We demonstrated the feasibility of reliable lambda-IX
by using the GMPLS fast fault recovery mechanism and PXCs

Distributed lambda-IX with L2-based connection

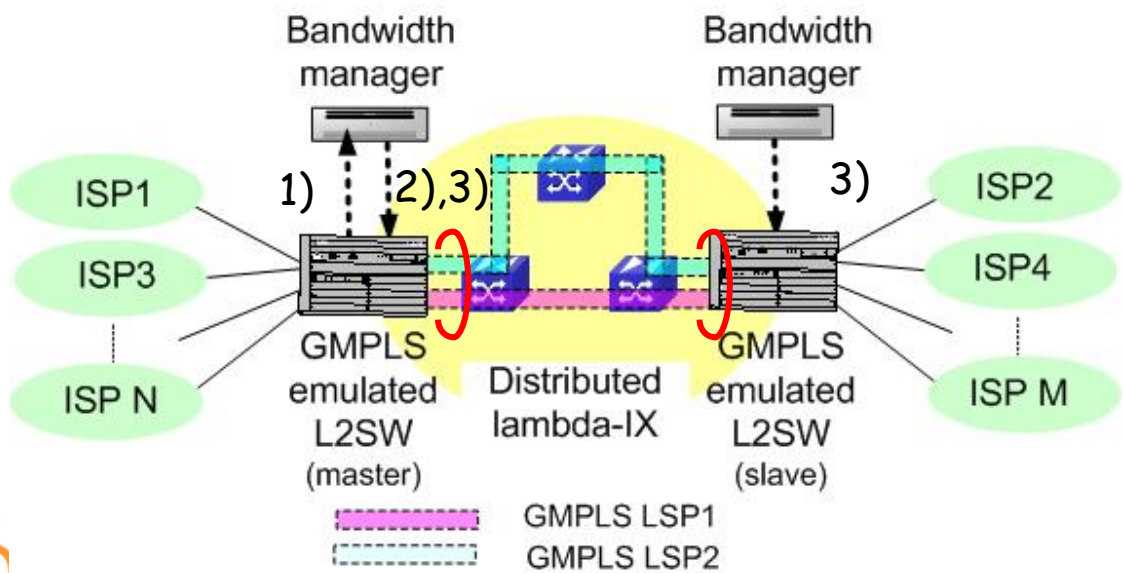
- Requirements for distributed lambda-IX with L2-based connection
  - Traffic management functions as one of the enhancements of L2-IX
    - To effectively utilize network resource (lambda) in lambda-IX
    - To meet unexpected traffic increase due to P2P applications, video transport application and etc..

GMPLS-controlled link bandwidth modification



GMPLS-controlled link bandwidth modification (1)

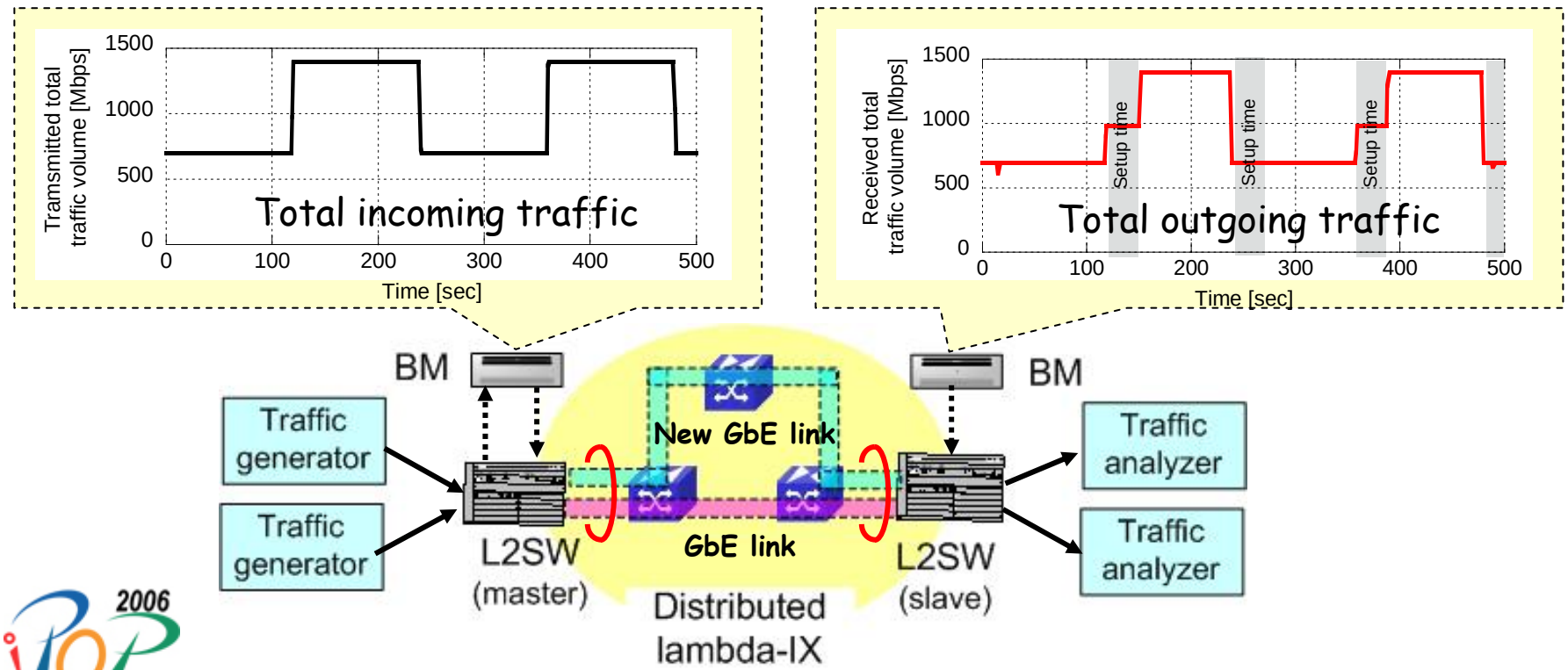
- Procedure of link bandwidth modification
  0) Master L2SW monitors the incoming and outgoing traffic volume exchanged between ISPs over that WAN interfaces
  1) Bandwidth manager system collects the traffic information from the L2SW and compares the information with the targeted threshold.
  2) Once the traffic exceeds the threshold, the L2SW automatically creates a new GMPLS-LSP and then aggregates the added link into original links



1) Traffic monitoring using SNMP
2) Request for additional LSP creation using CLI
3) Request for Link aggregation of two links using CLI

GMPLS-controlled link bandwidth modification (2)

- Evaluation results
  - Utilization ratio of link bandwidth with 1Gbps
    - Upper threshold : 90% (900Mbps per link)
    - Lower threshold : 50% (500Mbps per link)



Conclusion

- We presented distributed lambda-IX as one of next generation-IX models and GMPLS services
- Our proposed lambda-IX has been demonstrated
  - With direct lambda interconnection between ISPs
    - Using eBGP over GMPLS loosely routed LSP
    - Using GMPLS fast restoration mechanism
  - With L2-based interconnection between ISPs
    - Using link bandwidth modification
- We believe the distributed lambda-IX can be introduced in the future Internet world thanks to GMPLS control plane and PXCs
- Future discussion
  - Operation and management technique
  - Accounting system
  - Multi-homing technique

Acknowledgement

- This work was supported by Ministry of Internal Affairs and Communications (MIC), JAPAN
  - Asia Broadband Program (http://www.asia-bb.net/en/index.html)