「MPLSを用いた広域分散IXの実現」

中川郁夫	江崎浩	永見健一	菊池豊
INTEC W&G	東京大学	東芝	高知工科大学

概要

本発表では、IX(Internet eXchange) に MPLS(Multi-Protocol Label Switching) を適用した次世代 IX のアーキテクチャ MPLS-IX の実現手段を提案する。
IX は自律的に運用されているネットワーク同士を相互接続する技術のひとつで、インターネットにおいては、ISP(Internet Service Provider; プロバイダ) 間のトラフィック交換を実現する手段のひとつとして IX は重要な役割を担っている。IX では ISP のボーダルータ (境界ルータ) を接続し、他の ISP との間で経路制御を行うことによりピアリング (ISP 間の相互接続) を行う。これまで IX は単純なデータリンク層、すなわち OSI7 階層モデルにおけるレイヤ 2 の技術によって実現されてきた。既存の IX のほとんどは ATM 技術を用いた IX、あるいはイーサネットの技術を用いた IX のいずれかに分類される。一方で、インターネットにおいて ISP 間で交換されるトラフィック量は急激に増加し、IX においても、必然的に処理できるトラフィック量の増加、あるいは信頼性、拡張性などの面が重要な要素になってきている。
本発表では、IX に MPLS 技術を応用した次世代 IX のアーキテクチャMPLS-IX を提案する。MPLS-IX はデータリンク層の技術に依存せず、広域分散環境での相互接続が可能であるなどの特徴を持つ。著者らは通信放送機構の委託研究を受けて、同技術の研究・実証実験を進めているが、今回、同研究を中心に次世代 IX 研究会を設立し、MPLS-IX による実証実験を広くオープンに展開している。本発表では、同実証実験についても報告する。

Abstract

In this presentation, we propose a next generation IX architecture which is based on MPLS(Multi-Protocol Label Switching) technology.
IX is a technology by which numerous autonomous networks interconnect each other. In the current Internet, many ISPs(Internet Service Providers) exchange a large volume of traffic over IXes. An ISP connects its border router to an IX and establishes ‘peering’ by exchanging routing information between other ISPs. Historically, most IXes are based on simple data-link media, i.e., Layer 2 technology in OSI 7 layer model. Some of existing IXes use ATM technology, and others use Ethernet or similar LAN technologies. On the other hand, total volume of traffic exchanged between ISPs grows rapidly, and many requirements such as more higher performance in traffic exchange, security, and scalability are very important in the next generation IX technologies.
In this presentation, we propose a new IX architecture using MPLS technology. This IX architecture has a benefit of independency on data-link media for interconnection, and allows ISPs to interconnect each other in widely distributed areas. We also involve the experimental project of MPLS-IX to research and test several features for the MPLS-IX architecture. We established the ‘Next Generation IX Consortium’ based on the research project, and started the experimental IX testbed. We report the current status of this experimental IX project, too.

MPLS-IX –
An implementation of new IX architecture
using MPLS technology

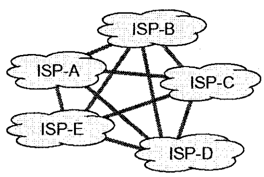
Ikuo Nakagawa, INTEC W&G
<ikuo@intec.co.jp>

Abstract

- Background
  - Existing IXes and issues
  - LAN based IX and ATM based IX
- Proposing a new IX architecture
  - Based on MPLS technology
  - Independent on data-link medium
  - Widely distributed IX environment
  - Hierarchical architecture and Scalability
- Next Generation IX Consortium
  - Introduction of experimental research project
  - Working groups and current status

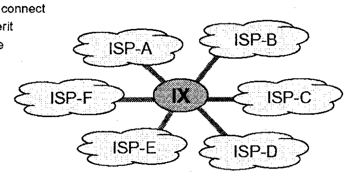
Interconnection model (1)

- Private peering
  - Individually configured peering between ISPs
    - Physical circuits on a peer basis
    - Fully meshed topology requires O(NxN) circuits
  - Scalability issues
    - How large is N (# of ISPs)?
    - Link costs?
    - Operational cost?
    - Negotiation?
    - etc.



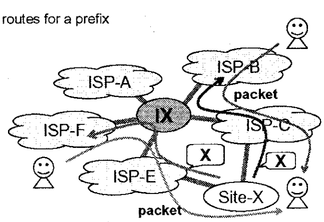
Interconnection model (2)

- IX (Internet eXchange)
  - IX is a field for interconnection between ISPs
    - Same functionality with private peerings
  - Only 1 physical link for an ISP
    - Full meshed connection requires O(N) circuits
  - Efficient!!
    - Easy to connect
    - Cost merit
    - Scalable



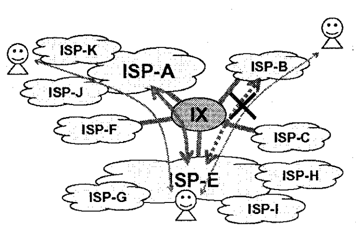
IX policy model (1)

- Routing policy is "bilateral"
  - IX doesn't care routing policy between ISPs
  - ISPs negotiate routing policy, each other
  - (ex) multiple routes for a prefix



IX policy model (2)

- 'bilateral' provides flexibility about peering
  - Peering is a business relationship
  - Large ISPs may disagree to peer with some ISPs



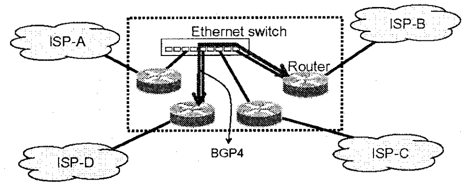
Current IX technologies (1)

- LAN-IX or ATM-IX
  - LAN-IX (IX based on LAN technologies)
    - FDDI or Ethernet switch based IX
    - Shared subnet - "exchange subnet"
    - Easy / Cheep
    - Concentrated model
  - ATM-IX (IX with ATM PVC technology)
    - Based on ATM switch or ATM network
    - Establish PVC(Permanent Virtual Circuit) between ISPs
    - Numerous PVCs over a single physical circuit
    - Distributed model

Nov, 2001

Current IX technologies (2)

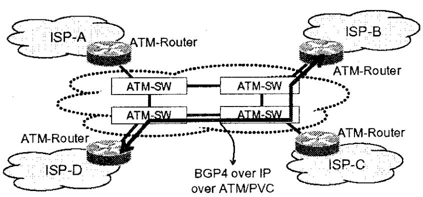
- LAN-IX
  - Switching speed (up to 1Gbps) is not enough
  - Less scalability especially from operational point of view
  - Third party next-hop problem (security issues)
  - Additional routers in the IX's housing space



Nov, 2001

Current IX Technologies (3)

- ATM-IX
  - Switching speed is a critical problem
  - Cell tax, which is why most IP engineers hate ATM
  - operational cost is also a critical problem



Nov, 2001

Design of new IX architecture

- Goal
  - Independent on data-link medium
    - No more "only Ethernet" or "only ATM"
    - Of course, we want to use POS, such as OC-192 or OC768, ...
    - No architectural limitation of interface speed
  - Widely distributed IX architecture
    - Hierarchical and scalable architecture
    - Widely distributed IX points
    - No more additional routers in IX points
  - And of course, "*bilateral*" policy control is required

Using MPLS technology

Nov, 2001

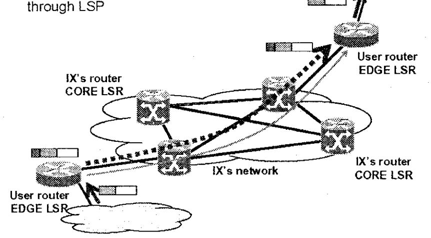
MPLS

- Multi-Protocol Label Switching
  - One of label switching technology (layer 2.5)
  - Many standards in IETF
  - Transmit packets with label information
  - Flexible routing with LSP(Label Switched Path)
  - Independent on lower layer (L2) medium
  - Independent on upper layer (L3) protocols
  - Traffic engineering
  - Supported by many routers

Nov, 2001

MPLS-IX model

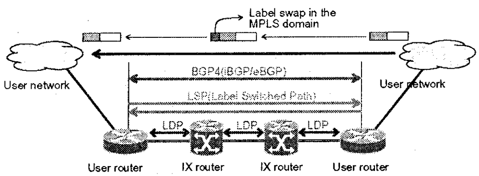
- Transmit through MPLS network
  - LSP between user networks
  - Transmit packets with label information through LSP



Nov, 2001

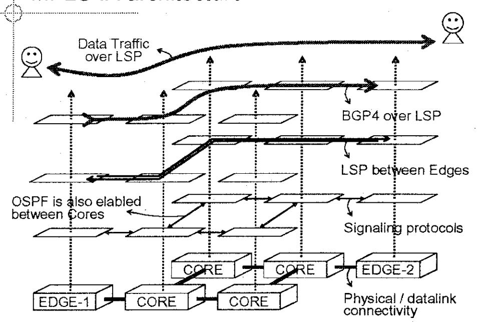
MPLS-IX model (cont'd)

- Virtual path (LSP) between participants' routers
  - Participant's router also uses MPLS signaling
  - LDP: Label Distribution Protocol
  - LSP: Label Switched Path
  - BGP4: Border Gateway Protocol version 4



Nov, 2001

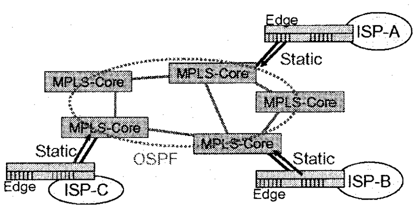
MPLS-IX architecture



Nov, 2001

Routing inside MPLS-IX

- Edge routers require only static routes
  - Static routes for peering (Edge) routers
  - OSPF or ISIS between Core routers
  - Avoid routing trouble depends on participating routers



Nov, 2001

Key features of MPLS-IX

- Independent on data-link medium
  - POS, ATM, GbE, FDDI, PPP, and any media
    - OC-192 or OC-768 POS also works fine
    - No need to negotiate about interface with peer
- Widely distributed IX
  - Widely distributed interconnection network
    - POS would be great for Core-Core interface
  - Participants can connect with any interface or circuit
    - No more additional routers or space required

Nov, 2001

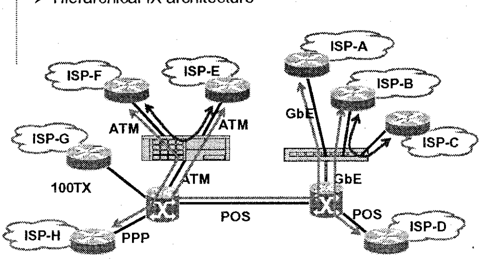
Key features of MPLS-IX (Cont'd)

- Flexibility
  - MPLS backbone is a kind of IP network
  - Hierarchical architecture
  - Flexible to extend MPLS network
    (No need to renumber address/mask)
- Operational benefit
  - Core routers have only topology information
  - Only host routes and BGP4 configuration required inside participants' (Edge) routers

Nov, 2001

Interconnection with existing IXes

- Interconnect multiple L2 IXes
- Hierarchical IX architecture



Nov, 2001

Possibilities of MPLS-IX

- Many useful MPLS features
  - Traffic Engineering
  - QoS/CoS - DiffServ features
  - Interconnection between MPLS Domains
  - Source address routing by RSVP
  - Solution to multi-home sites
  - etc...

Nov, 2001

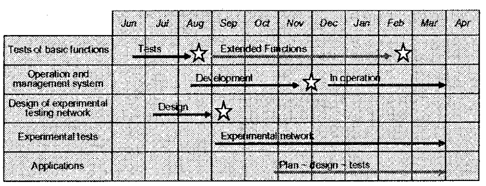
Next Generation IX Consortium

- Consortium for experimental research project
  - Core members:
    - Chair: Hiroshi Esaki
    - Board: Ikuo Nakagawa, Kenichi Nagami, Yutaka Kikuchi
    - Advisory: Shinji Shimojo
  - Meeting:
    - Open to join; one in a every two months
  - Engineering and experiment are also open to join
    - Both in testbed and in full connected network
    - At your own risk ☺
  - Disclose result or outcome of the research
    - Write papers
    - Results of private research, is up to you

Nov, 2001

Schedule

- Technical tests for basic functionalities
- Experimental test in fully connected network
  - In widely distributed areas in Japan



Nov, 2001

Next Generation IX Consortium (cont'd)

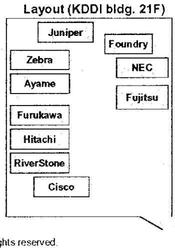
- Working groups:
  - router working group
    - Chair/Co-chair: Yutaka Kikuchi, Kenichi Nagami
    - summarize requirements for MPLS router implementations
    - interoperability test between numerous implementations
  - IX provider working group
    - Chair:Co-chair: Hiroshi Esaki, Ikuo Nakagawa
    - summarize requirements for MPLX-IX provider
    - development operational/management technology for MPLS-IX
    - interconnection between multiple MPLS-IX providers
  - IX user working group
    - Chair/Co-chair: Ikuo Nakagawa, Kenichi Nagami
    - summarize requirements for MPLS-IX users (participants)
    - summarize configuration and tips for participating MPLS-IX
    - implement and operate experimental MPLS-IX

Nov, 2001

Interoperability test

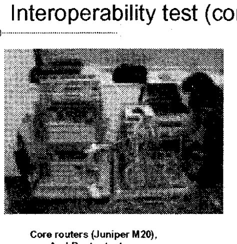
- An activity of router working group
  - Interoperability test for MPLS-IX architecture (10/15-19)
  - In total, 10 vendors participated to the test
    - Ayame, Cisco, Foundry, Fujitsu, Furukawa, Hitachi, Juniper, NEC, RiverStone, Zebra
  - Many test items
    - Signaling protocols: RSVP, LDP
    - Addressing: Loopback, Interface
    - TTL option: no-propagate-ttl
    - BGP4: addressing, ebgp-multihop
    - etc, ...
  - The result is here:
    - http://www.distix.net/router-wg/
  - Next test will be on 1/28-2/1

Layout (KDDI bldg. 21F)



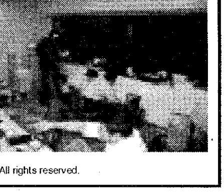
Nov, 2001

Interoperability test (cont'd)



Core routers (Juniper M20), And Router testers

Developers / engineers from Router vendors joined the test



Nov, 2001

Requirements for MPLS Implementations

- Treat an LSP as a tunnel
  - Using a LSP for both;
    - actual data traffic transfer (end-to-end communication)
    - BGP4 packet transmission (LSR-to-LSR communication)
  - no ttl decrement option or 1-hop LSP mode is required
  - traffic and packet counter MIB are required
- Addressing behavior
  - Need to handle both loopback and interface address for;
    - End addresses for RSVP / Transport addresses for LDP
    - End address for LSP (FEC)
    - BGP4 source/destination addresses
- Routing information
  - Run OSPF inside Core routers
  - Configure static routes end Edge routers

Nov, 2001

Requirements for MPLS implementations

- Signaling protocols
  - Support both LDP and RSVP
  - LDP;
    - DOD or DU / Ordered or Independent / Liberal
    - Filtering FEC for inbound and outbound
    - LSP keepalive or fail detection
    - LDP over RSVP is required
  - RSVP;
    - Configurable without OSPF/IS-IS (e.g., with static routes)
    - Filtering signaling packet from unauthorized routers
    - LSP keepalive or similar mechanism

Nov, 2001

Requirements for MPLS implementations

- QoS Support
  - Minimum bandwidth guarantee
    - New carrier's service such as virtual leased line
    - extra traffic will be treated
  - Traffic rate limiting
    - policing and shaping mechanism
- RFC3107 Support
  - Inter-Carriers label exchange mechanism
  - Using BGP4 between carriers
- RFC2547 extension
  - Carriers' carrier model
  - VRF which supports LSP/LDP (seen to customers)

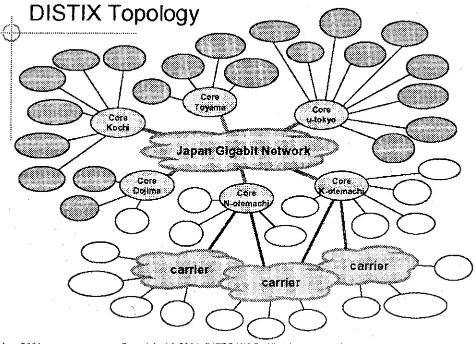
Nov, 2001

Implement/operate experimental MPLS-IX

- Running by IX-user working group
  - "distix" - experimental MPLS-IX operated by the consortium
  - 6 Core routers for experimental MPLS-IX
  - Using JGN as the MPLS-IX backbone
  - Many (over 40) networks will join/connect to the distix
- Participants can connect to the "distix" via
  - JGN – paths between participants and Core routers
  - Direct connect – to the one of Core routers
- Participants have to prepare
  - MPLS available router (with RSVP or LDP or both)
  - Circuits to the JGN or Core routers

Nov, 2001

DISTIX Topology



Nov, 2001

Conclusion

- MPLS-IX is a new IX architecture
  - Data-link media independent IX
    - No more limitation of data-link bandwidth
  - Widely distributing IX
    - Flexible and scalable IX architecture
  - Of course, it provides "bilateral" policy control
- distix – Next Generation IX Consortium
  - Three working groups such as:
    - Router interoperability working group
    - IX provider working group
    - IX user working group
  - Open and free to join - Please join us!

Nov, 2001