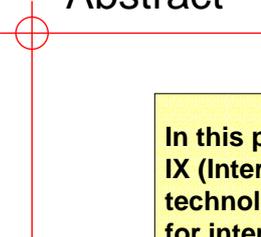


## ***MPLS-IX***

An implementation of new IX architecture  
using MPLS technology

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

## Abstract



**In this presentation, we propose a next generation IX (Internet eXchange) architecture using MPLS technology. That is, we are trying to apply MPLS for interconnection (or peering) between ISPs.**

**We have been tested the MPLS based IX (MPLS-IX) architecture and functionality since 1999. We also started an experimental testbed in Oct, 2001. We report these MPLS-IX activities, too**

# Agenda

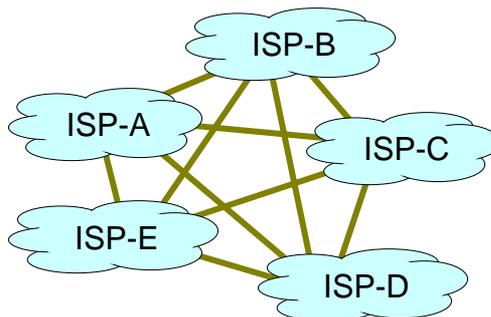
- Background
  - Existing IXes and issues
  - LAN based IX and ATM based IX
- Proposing a new IX architecture
  - Based on MPLS technology
  - Independent on data-link medium
  - Widely distributed IX environment
  - Hierarchical architecture and Scalability
- Next Generation IX Consortium
  - Introduction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

Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

# Interconnection model (1)

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

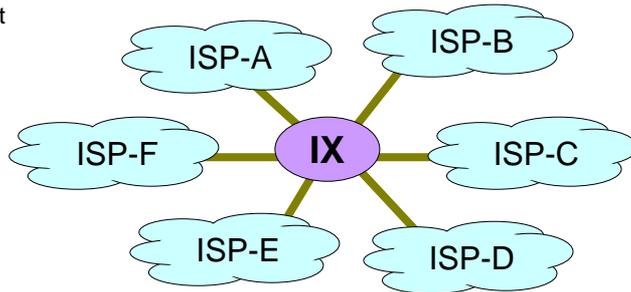


Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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

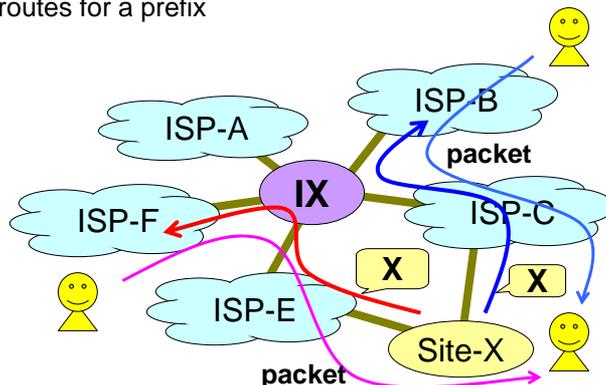


Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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

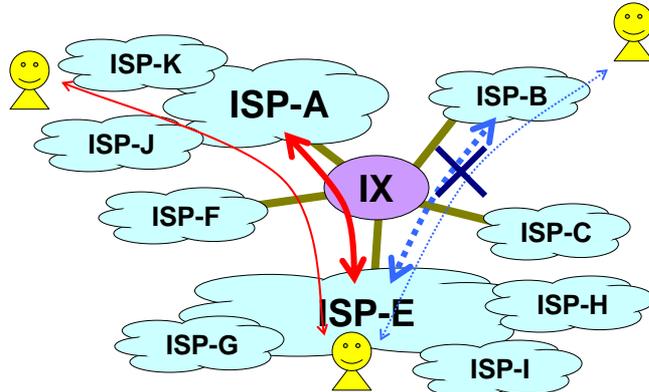


Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## IX policy model (2)

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



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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 / Cheap
    - 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

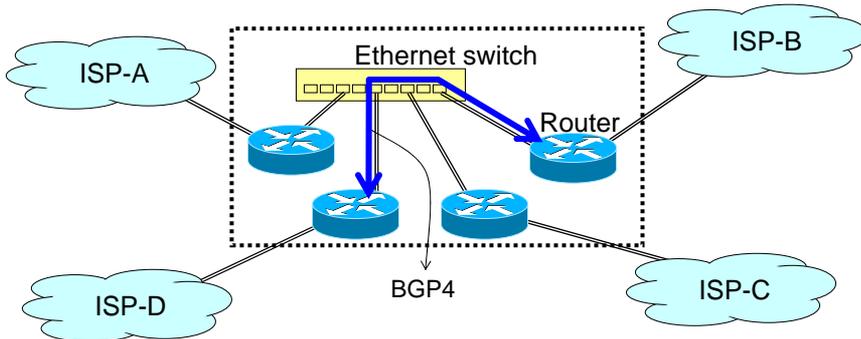
Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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



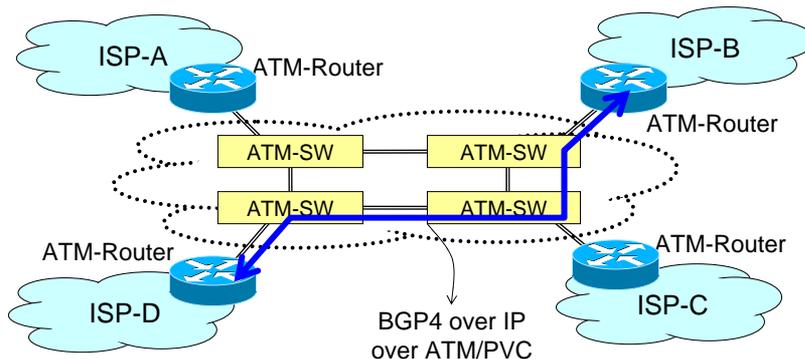
Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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, (ex.) OC-192 or OC768, ...
  - No architectural limitation of interface speed
- Widely distributing architecture
  - Hierarchical and scalable architecture
  - IX points in widely distributed areas
  - No more additional routers in IX points
- And of course, “*bilateral*” policy control is required

**Using MPLS technology**

Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## MPLS

- What is MPLS?

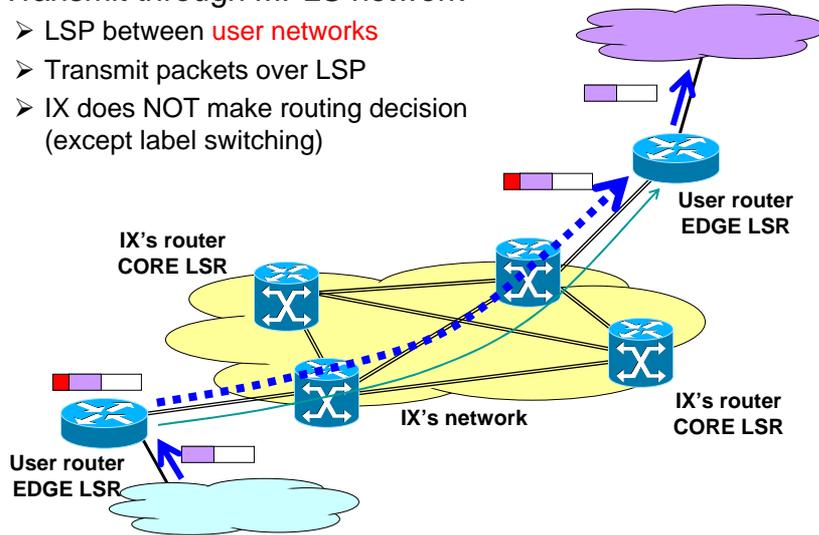
Snip, today.

Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

# MPLS-IX model

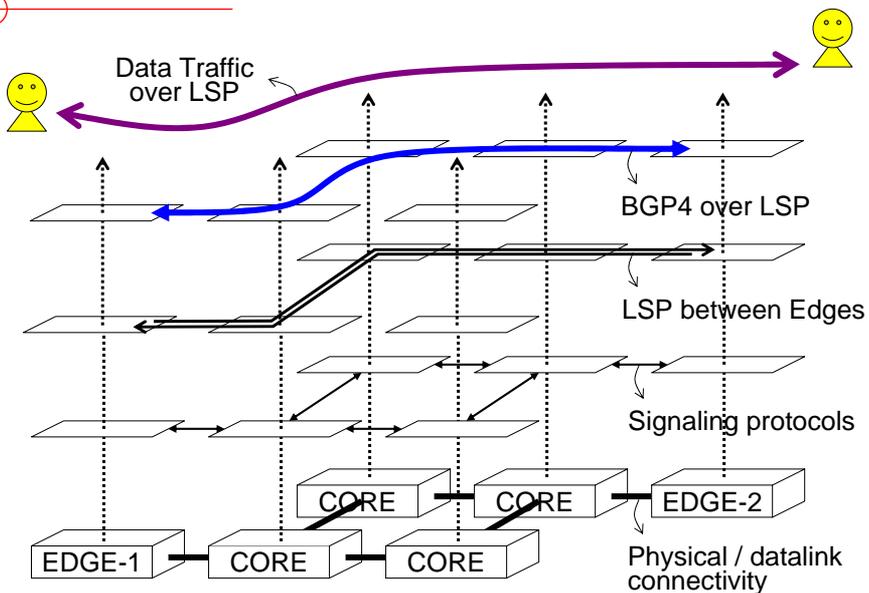
- Transmit through MPLS network
  - LSP between **user networks**
  - Transmit packets over LSP
  - IX does NOT make routing decision (except label switching)



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

# MPLS-IX architecture

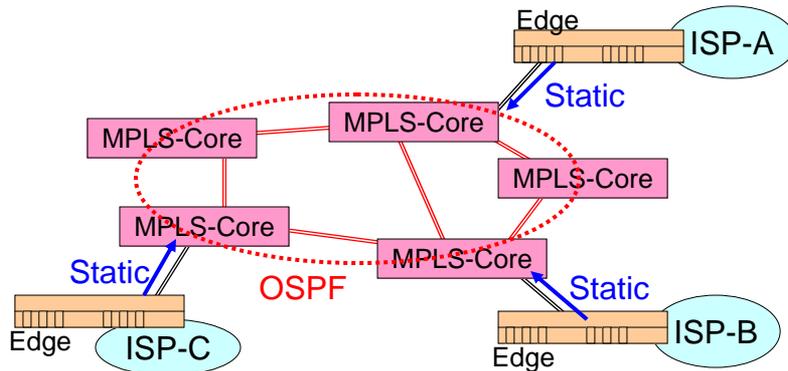


Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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 distributing model
  - Widely distributing 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Key features of MPLS-IX (Cont'd)

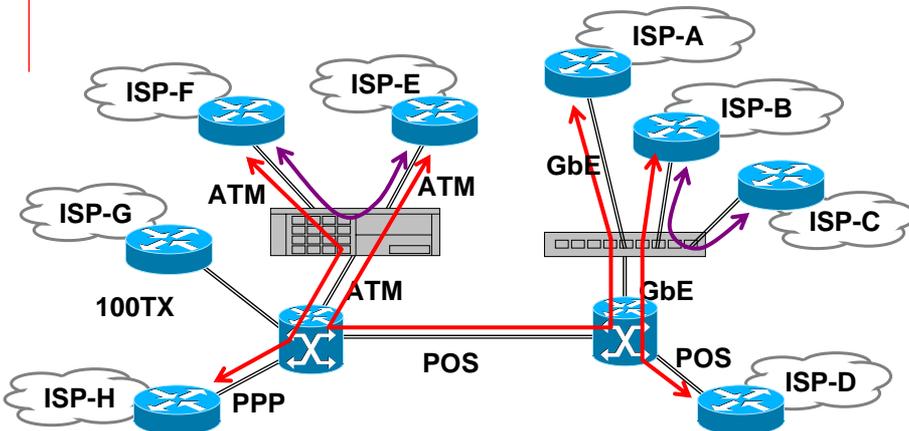
- Flexibility
  - MPLS backbone is a kind of IP network
  - Hierarchical architecture
  - Flexible to extend MPLS network  
(No need to renumber address/mask)
- Operational benefit
  - Core routers have only topology information
  - Only host routes and BGP4 configuration required inside participants'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

Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Interconnection with existing IXes

- Interconnect multiple L2 IXes
- Hierarchical IX architecture



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Possibilities of MPLS-IX

- Many useful MPLS features
  - Traffic Engineering
  - QoS/CoS - DiffServ features
  - Interconnection between MPLS Domains
  - IPv6 interconnection over MPLS network
  - Source address routing by RSVP
  - Solution to multi-home sites
  - Extension to optical switch IX
  - etc...

Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Next Generation IX Consortium

- Consortium for experimental research project
  - Core members:
    - Chair: Hiroshi Esaki
    - Board: Ikuo Nakagawa, Kenichi Nagami, Yutaka Kikuchi
    - Advisory: Shinji Shimojo
  - Core research team
    - Based on TAO's research project
  - 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Next Generation IX Consortium (cont'd)

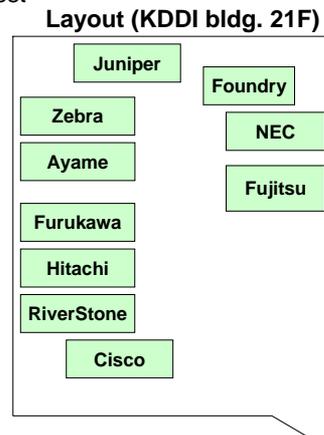
- Working groups:
  - router working group
    - Chair/Co-chair: Yutaka Kikuchi, Kenichi Nagami
    - summarize requirements for MPLS router implementations
    - interoperability test between numerous implementations
  - IX provider working group
    - Chair:Co-chair: Hiroshi Esaki, Ikuo Nakagawa
    - summarize requirements for MPLX-IX provider
    - develop operational/management technology for MPLS-IX
    - interconnect 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Interoperability test

- An activity of router working group
  - Interoperability test for MPLS-IX architecture
    - on 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, multihop
  - etc, ...
- The result is here:
  - <http://www.distix.net/router-wg/>
- Next test will be on 1/28-2/1



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Interoperability test (cont'd)



**Core routers (Juniper M20),  
And Router testers**

**Developers / engineers from  
Router vendors joined the test**



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

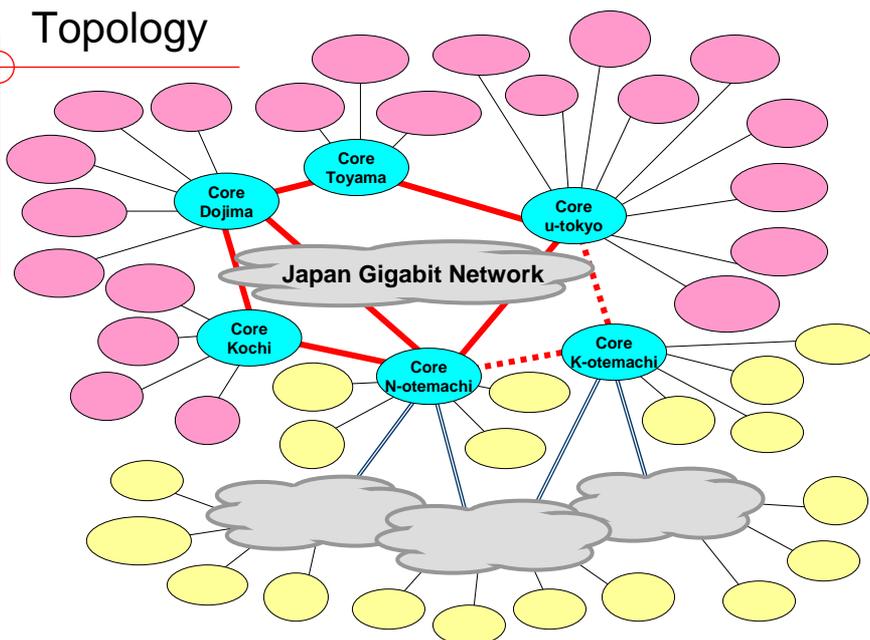
# Implementation of experimental MPLS-IX

- 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

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Topology

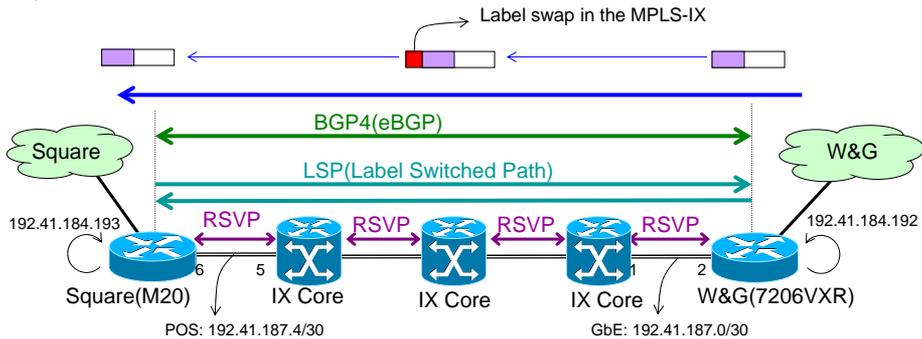


Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration

- Square - INTEC W&G
  - Enabling MPLS/RSVP configuration
  - Establishing LSP between Square and W&G
  - Establishing BGP4 neighbors via LSP

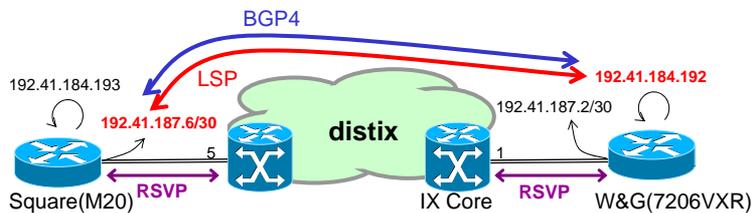


Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

- Physical configuration
  - Square: POS/OC-12 with Juniper M20
  - W&G: GigabitEthernet with Cisco 7206VXR
  - (Between Core LSRs: ATM)
- LSP&BGP4 configuration (control plain)
  - Square: 192.41.187.6 (interface address)
  - W&G: 192.41.184.192 (loopback address)



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

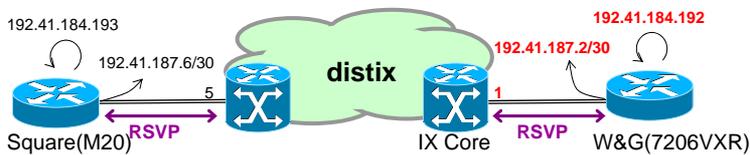
### ➤ Basic configuration for W&G (7206VXR)

```
interface loopback0
 ip address 192.41.184.192 255.255.255.255

interface GigabitEthernet2/0
 ip address 192.41.187.2 255.255.255.252

ip route 192.41.184.0 255.255.252.0 192.41.187.1
```

address space for "distix" = 192.41.184.0/22



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

### ➤ MPLS/RSVP configuration

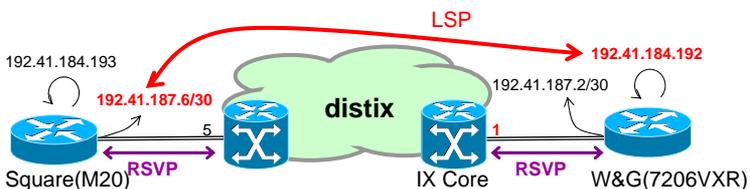
- enable MPLS
- enable RSVP tunnel
- configure explicit-path (try verbatim option!)

```
mpls traffic-eng tunnels
mpls traffic-eng signalling advertise implicit-null
no tag-switching ip propagate-ttl
tag-switching tdp router-id Loopback0
```

```
interface Tunnel0
 ip unnumbered Loopback0
 tunnel destination 192.41.187.6
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng path-option 10 explicit ¥
 name to-square verbatim
```

```
ip explicit-path name to-htcn enable
 next-address 192.41.187.1
```

```
ip route 192.41.187.6 255.255.255.255 Tunnel0
```



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

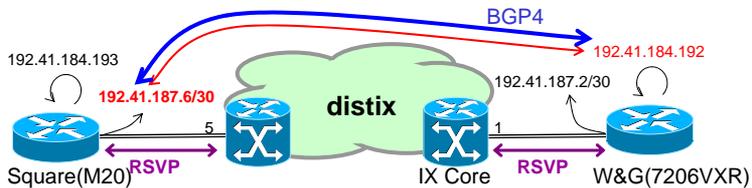
- BGP4 configuration
  - configure neighbor information
  - use "ebgp-multihop 2"

```

router bgp 9612
neighbor 192.41.187.6 remote-as 17685
neighbor 192.41.187.6 ebgp-multihop 2
neighbor 192.41.187.6 update-source Loopback0
neighbor 192.41.187.6 send-community
neighbor 192.41.187.6 filter-list 1 out
    
```

```

ip as-path access-list 1 permit ^$
    
```



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

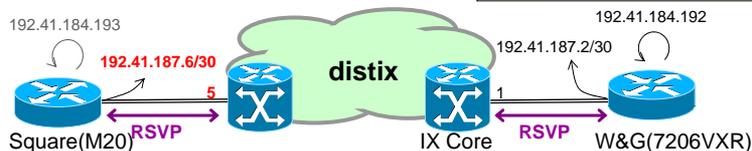
- Basic configuration for Square (M20)
  - configure interface address (and enabling mpls)
  - configure static routing entry

```

interfaces {
  so-0/0/0 {
    unit 0 {
      family inet {
        address 192.41.187.6/30;
      }
      family mpls;
    }
  }
}

routing-options {
  static {
    route 192.41.184.0/22 {
      next-hop 192.41.187.5;
      preference 30;
    }
  }
}
    
```

address space for "distix" = 192.41.184.0/22



Nov, 2001

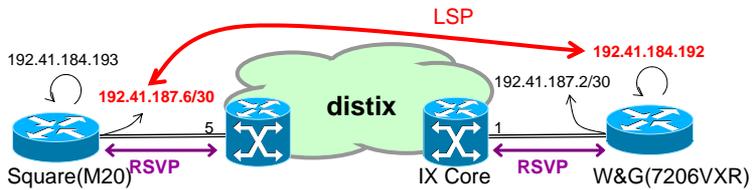
Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

- MPLS/RSVP configuration
  - enable MPLS
  - enable RSVP
  - configure label-switched-path with no-cspf option

```

protocols {
  rsvp {
    interface so-0/0/0.0;
  }
  mpls {
    traffic-engineering bgp-igp;
    no-propagate-ttl;
    label-switched-path from-square-to-wag {
      from 192.41.187.6;
      to 192.41.184.192;
      no-cspf;
    }
    interface so-0/0/0.0;
  }
}
    
```



Nov, 2001

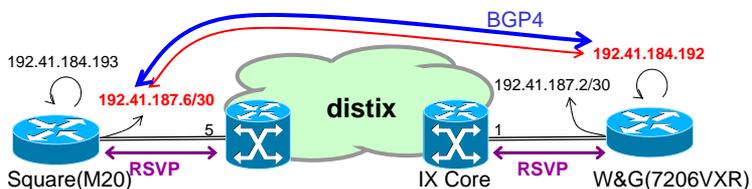
Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

- BGP4 configuration
  - configure neighbor information
  - use "multihop ttl 1"

```

protocol {
  bgp {
    local-as 17685;
    group wag {
      multihop ttl 1;
      local-address 192.41.187.6;
      peer-as 9612;
      neighbor 192.41.184.192;
    }
  }
}
    
```



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Example configuration (cont'd)

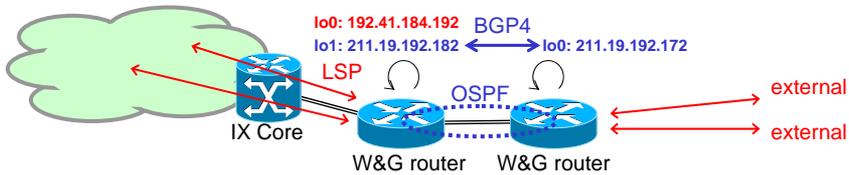
- Configuring internal OSPF with yet another loopback

```
tag-switching tdp router-id Loopback0

interface Loopback1
 ip address 211.19.192.182 255.255.255.255

router ospf 1
 network 211.19.192.182 0.0.0.0 area 0
 network ....

router bgp 9612
 neighbor 211.19.192.172 remote-as 9612
 neighbor 211.19.192.172 update-source Loopback1
```



Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

## Applying optical switch to the MPLS-IX

- Extending MPLS-IX toward optical-IX

Snip, today.

Nov, 2001

Copyright (c) 2001, INTEC W&G, All rights reserved.

# Conclusion

- MPLS-IX is a new IX architecture
  - Data-link media independent IX
    - No more limitation in data-link types or 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
  - It’s open to join - **Please join us!**