Multicast only FRR

An IP solution to achieve fast (<50ms) multicast convergence

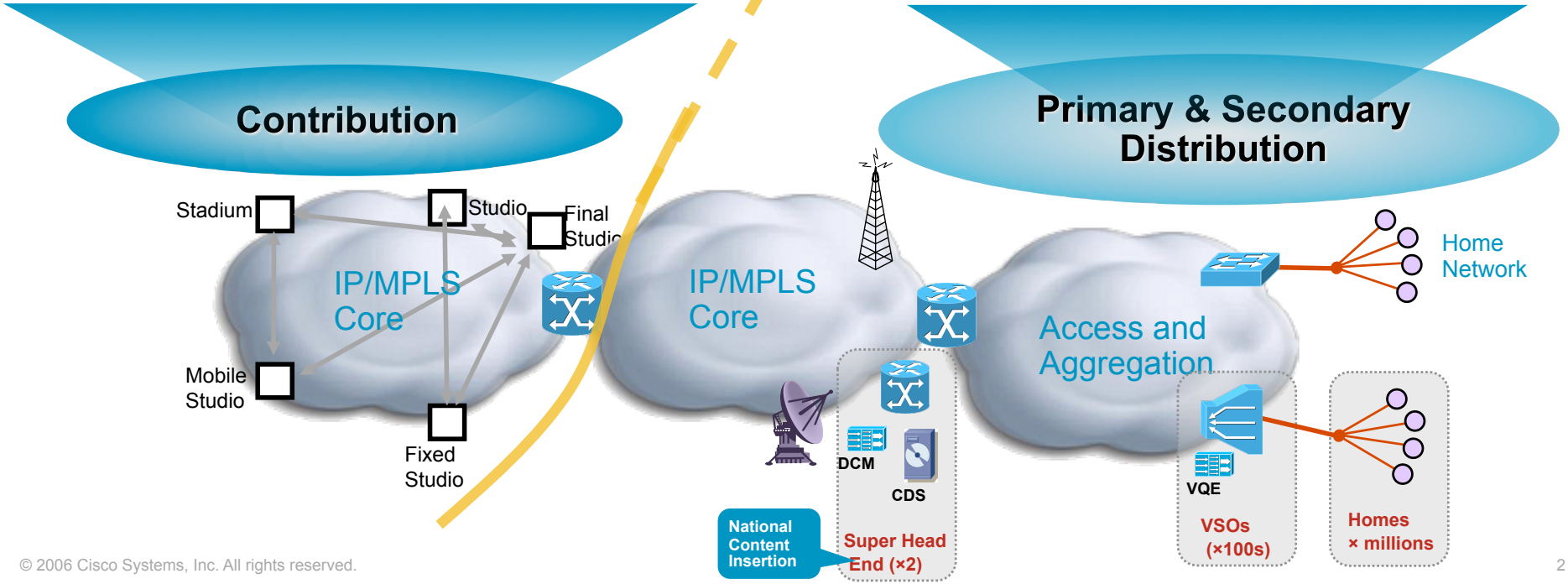
Santanu Dasgupta

© 2006 Cisco Systems, Inc. All rights reserved.

IP/TV & Video Service Provider Requirements

- Very high bandwidth, less end-points
- P2P and P2MP transport
- Mostly uncompressed video format is used (270Mbps – 3Gbps/stream)
- Extremely high degree of reliability is required, *hitless switchover (0ms)* desired by broadcasters nowadays – 50ms is claimed to be insufficient!

- Low bandwidth/stream, thousands to millions of end-points
- P2MP transport
- Compressed video (MPEG) is used
- 50ms requirement is a myth, 50ms switchover may cause artifact lasting 300ms. However, many people feel they would be better of with 50ms!



© 2006 Cisco Systems, Inc. All rights reserved.

Multicast Fast Convergence tools

- IGP / PIM Fast Convergence
- P2MP TE / LSM
- Multicast VPLS with TE-FRR
- Multicast-only FRR (MoFRR)
  - http://tools.ietf.org/html/draft-karan-mofrr-00.txt
- Tunnel based Multicast FRR (TMFRR)
  - http://tools.ietf.org/html/draft-lwei-pim-tmfrr-00

© 2006 Cisco Systems, Inc. All rights reserved.

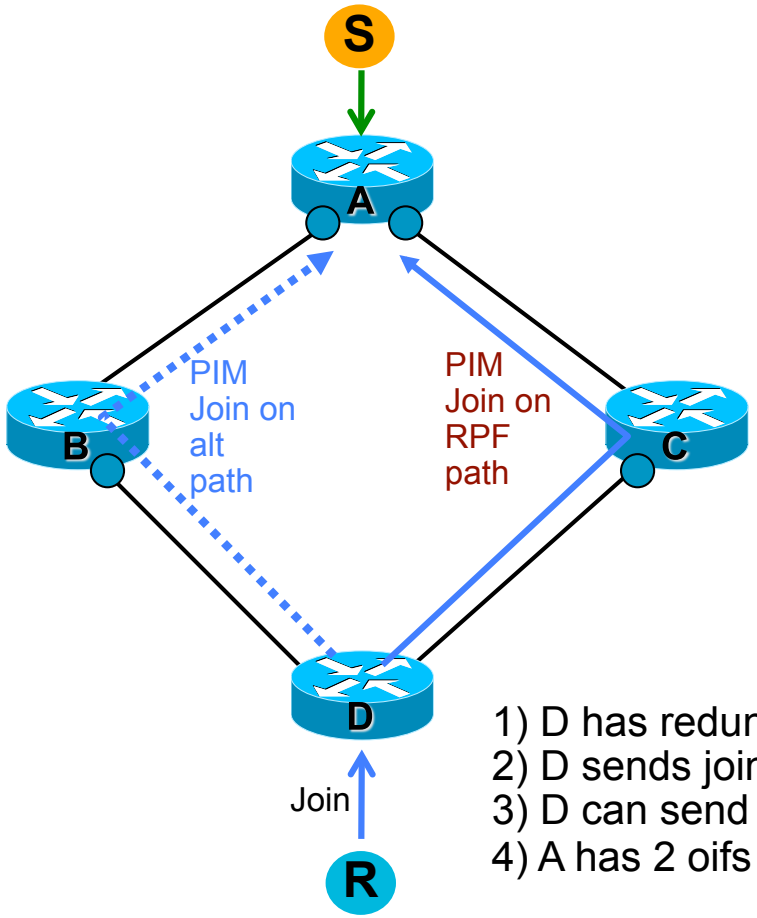
Multicast only FRR (MoFRR)

- A simple IP based solution
  - works with simple enhancement on PIM of sending redundant PIM join over an alternate path
- Can provide sub-50ms multicast convergence
  - the PE locally switch to the backup path upon failure on primary
- Purely local behavior on the last hop PIM routers
  - e.g. P routers don’t need to be aware of MoFRR
- The concept as such is capable of hitless delivery
  - PE uses the two branches to repair losses and present lossless data (with RTP sequence numbers) to its IGMP neighbors

© 2006 Cisco Systems, Inc. All rights reserved.

MoFRR Working Principle – 1

Receiver PE sends Redundant PIM Joins over Alternate Paths



1) D has redundant paths {BA, CA} to S
2) D sends join on RPF path through C
3) D can send alternate-join on BA path
4) A has 2 oifs leading to a single receiver

Data path

RPF path (RPF join)

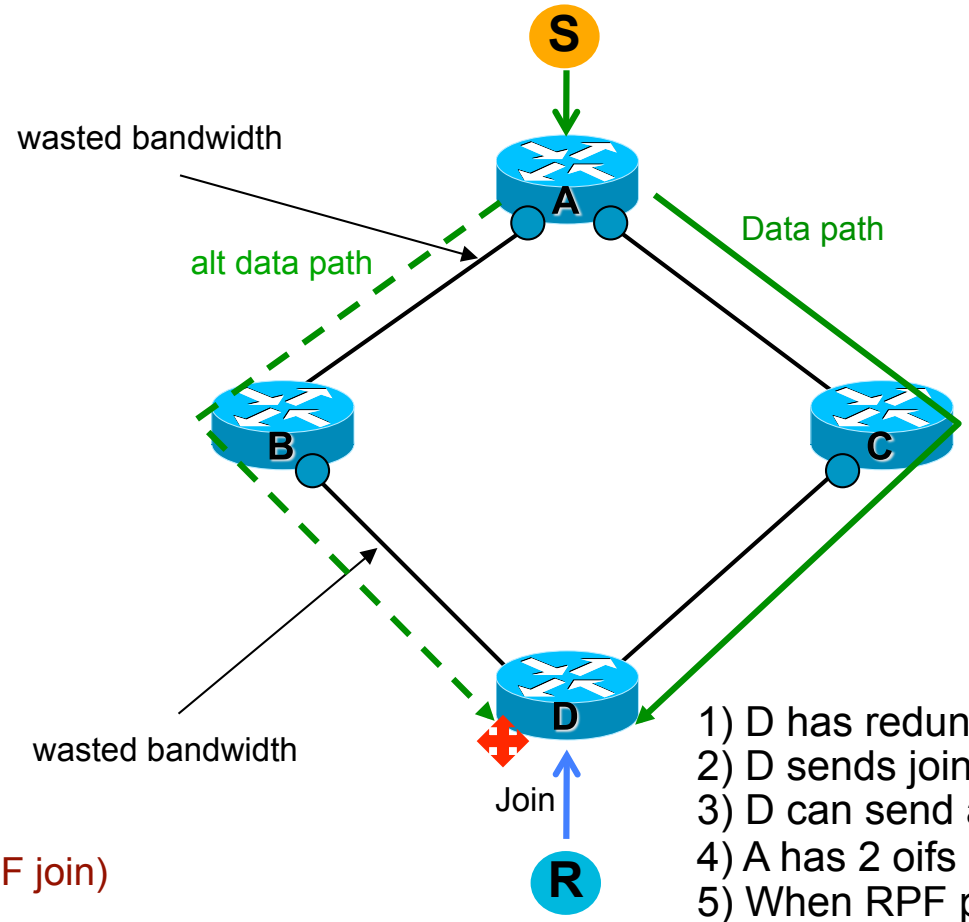
Alt join (sent on non-RPF)

interface in oif-list

© 2006 Cisco Systems, Inc. All rights reserved.

MoFRR Working Principle – 2

Redundant data path is set, redundant path RPF failed at receiver PE



1) D has redundant paths {BA, CA} to S
2) D sends join on RPF path through C
3) D can send alternate-join on BA path
4) A has 2 oifs leading to a single receiver
5) When RPF path is up, duplicates come to D
6) But D RPF fails on packets from B

Data path

RPF path (RPF join)

Alt join (sent on non-RPF)

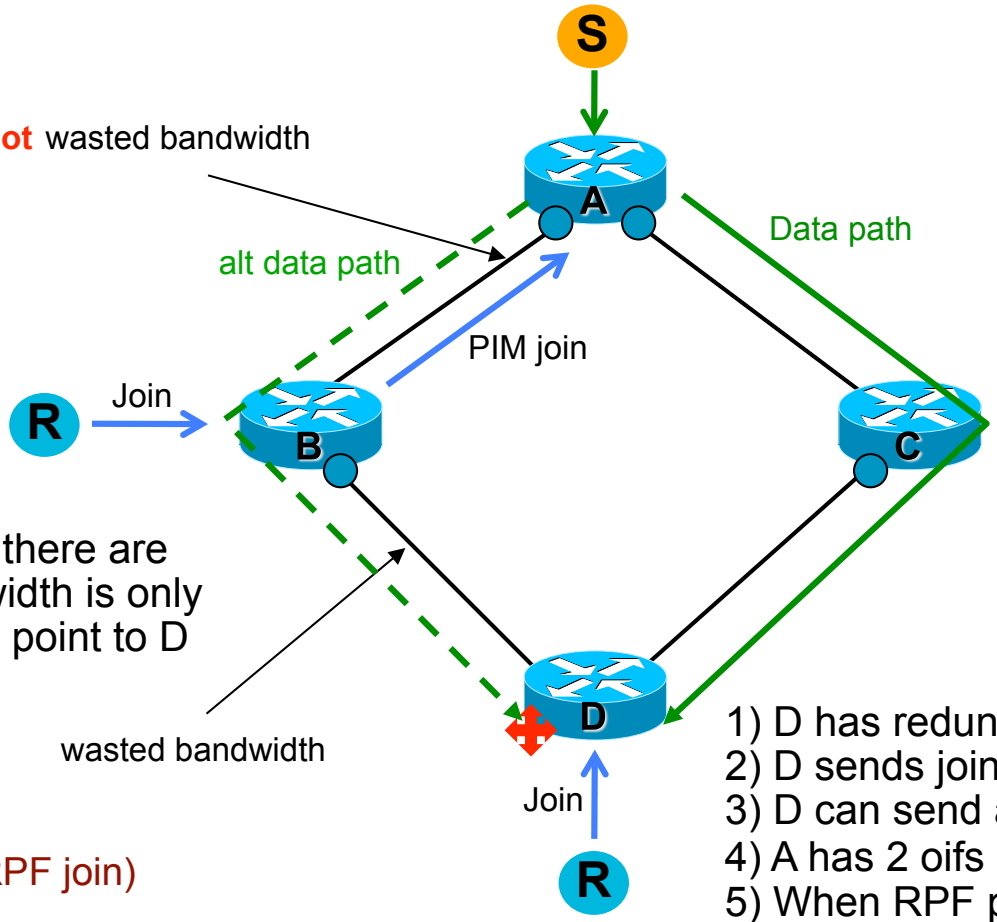
interface in oif-list

link down or RPF-failed packet drop

© 2006 Cisco Systems, Inc. All rights reserved.

MoFRR Working Principle – 2 (Contd.)

No bandwidth is wasted if intermediate nodes has receivers



7) If upstream of D there are receivers, bandwidth is only wasted from that point to D

1) D has redundant paths {BA, CA} to S
2) D sends join on RPF path through C
3) D can send alternate-join on BA path
4) A has 2 oifs leading to a single receiver
5) When RPF path is up, duplicates come to D
6) But D RPF fails on packets from B

Data path

RPF path (RPF join)

Alt join (sent on non-RPF)

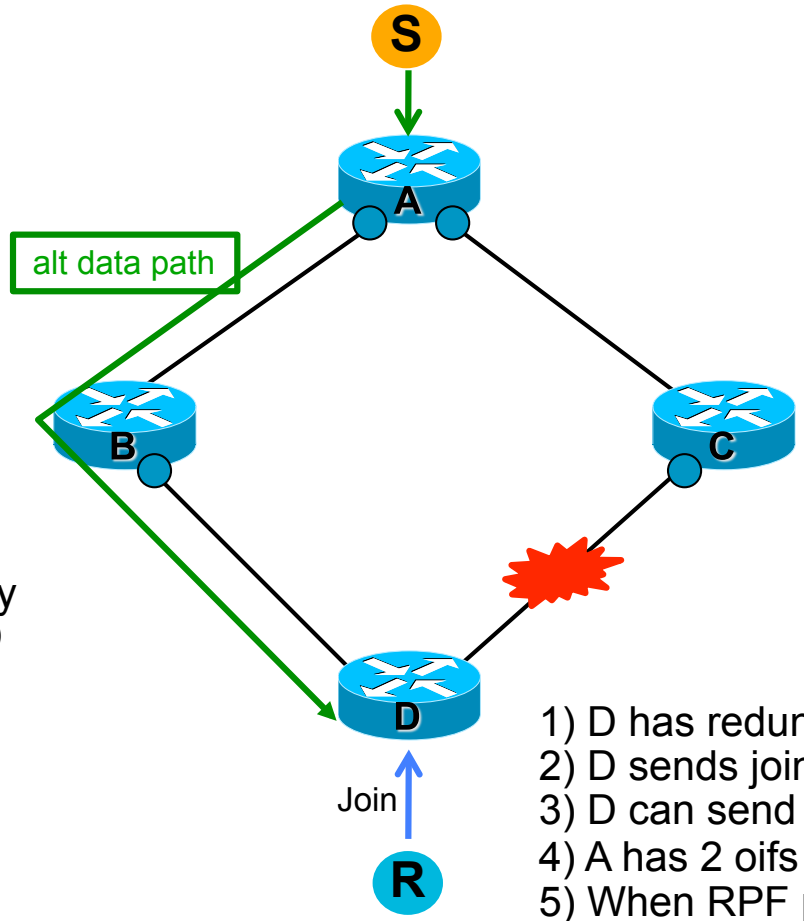
interface in oif-list

link down or RPF-failed packet drop

© 2006 Cisco Systems, Inc. All rights reserved.

MoFRR Working Principle – 3

Upon failure on the primary path, receiver PE makes local decision to accept packets from backup path without waiting for the IGP to converge



7) If upstream of D there are receivers, bandwidth is only wasted from that point to D

1) D has redundant paths {BA, CA} to S
2) D sends join on RPF path through C
3) D can send alternate-join on BA path
4) A has 2 oifs leading to a single receiver
5) When RPF path is up, duplicates come to D
6) But D RPF fails on packets from B

Data path

RPF path (RPF join)

Alt join (sent on non-RPF)

interface in oif-list

link down or RPF-failed packet drop

8) When C fails or DC link fails, D makes local decision to accept packets from B

9) Eventually unicast routing says B is new RPF path

© 2006 Cisco Systems, Inc. All rights reserved.

Failure trigger for MoFRR

- Two mechanisms: IGP based and Flow based.
- **IGP based: (protocol based)**
  - Switchover triggered by IGP primary route detection
  - On average, switchover time is in hundreds of ms.
- **Flow based: (hardware based)**
  - Switchover triggered by traffic loss detection in hardware on the primary path
  - Flow based hardware counter (such as vidmon) can be used
  - Switchover time is under 50 ms.

© 2006 Cisco Systems, Inc. All rights reserved.

Summary

- MoFRR is a simple IP based approach for multicast fast convergence for IPTV providers
- Make-before-break solution for sub-50ms
- No new protocols, no interoperability requirements, no forwarding/hardware changes
- Only receiver PE’s need to be aware of MoFRR
- Concept is capable of doing hitless switchover (0ms)!

© 2006 Cisco Systems, Inc. All rights reserved.

© 2006 Cisco Systems, Inc. All rights reserved.