SRP Stream Path Selection

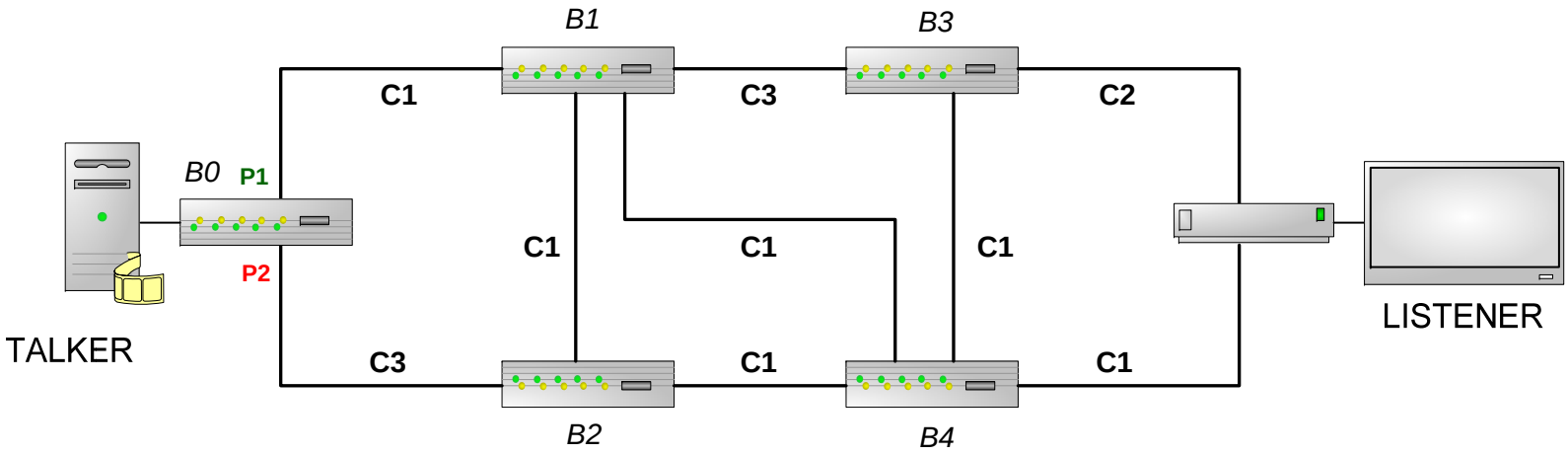
IEEE 802.1 AVB WG
Singapore - Mar 2011
Santa Fe – May 2011

Presented by: Philippe Klein, Broadcom
philippe@broadcom.com

SRP Stream Path Selection

- The goal:
  - Taking advantage of all the paths available in given network topology to
    - Increase the overall bandwidth
    - Increase resilience (failover or redundancy)
- The current limitations:
  - Standard 802.1D Rapid Spanning Tree Protocols:
    - Block redundant links (and therefore covers only a partial topology)
    - uses one static link metric only (LineSpeed) to select the “best” path between adjacent nodes
  - Multiple STP (MSTP)
    - complexity to configure the network
- Link State Routing protocols
    - TRILL
    - 802.1aq Shortest Path Bridging
- … could be potential solutions but not simple protocols…

SRP Talker Advertise

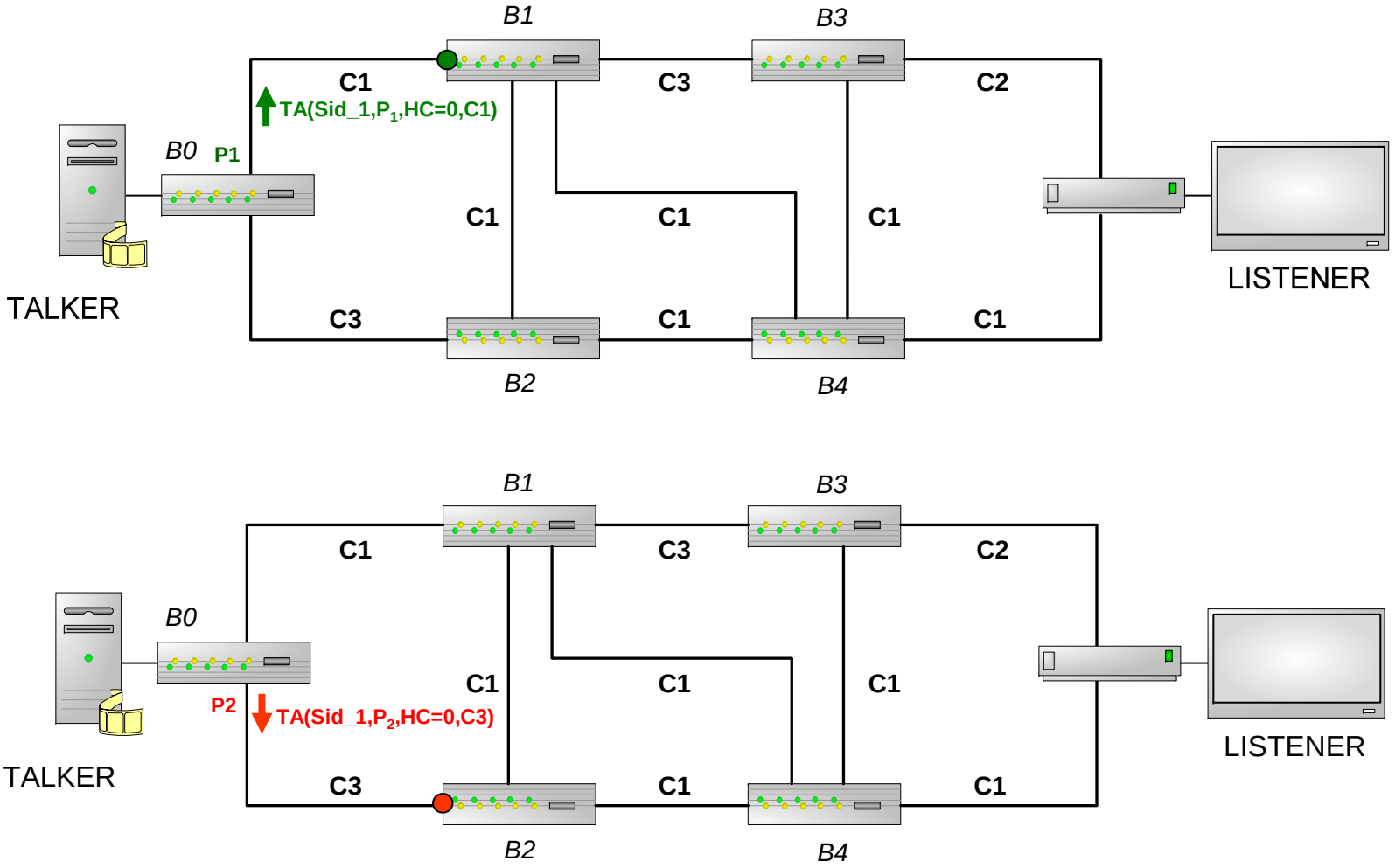


TA is propagated over the 2 ports of B0

●● = Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

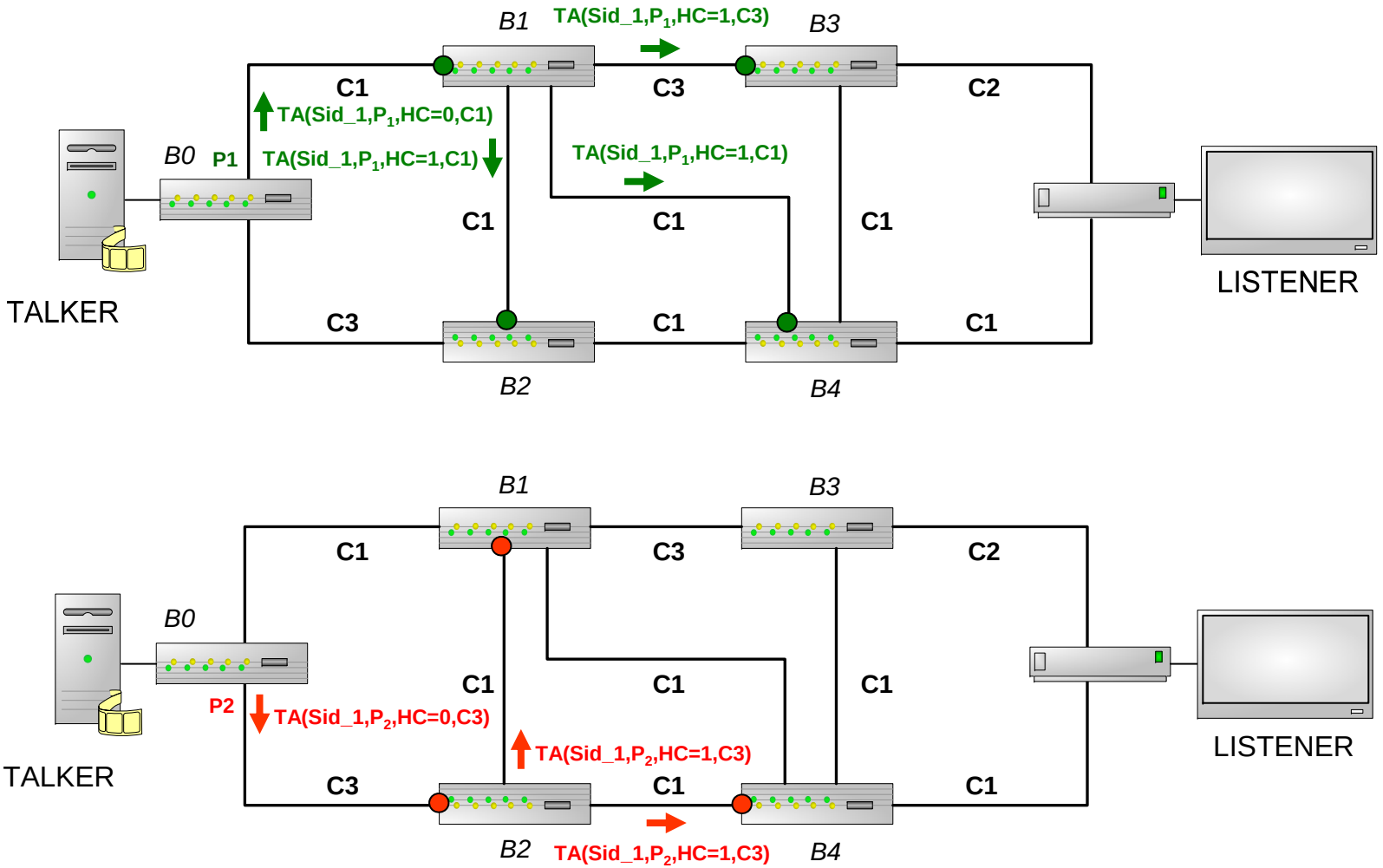
SRP Talker Advertise – Step 1



= Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

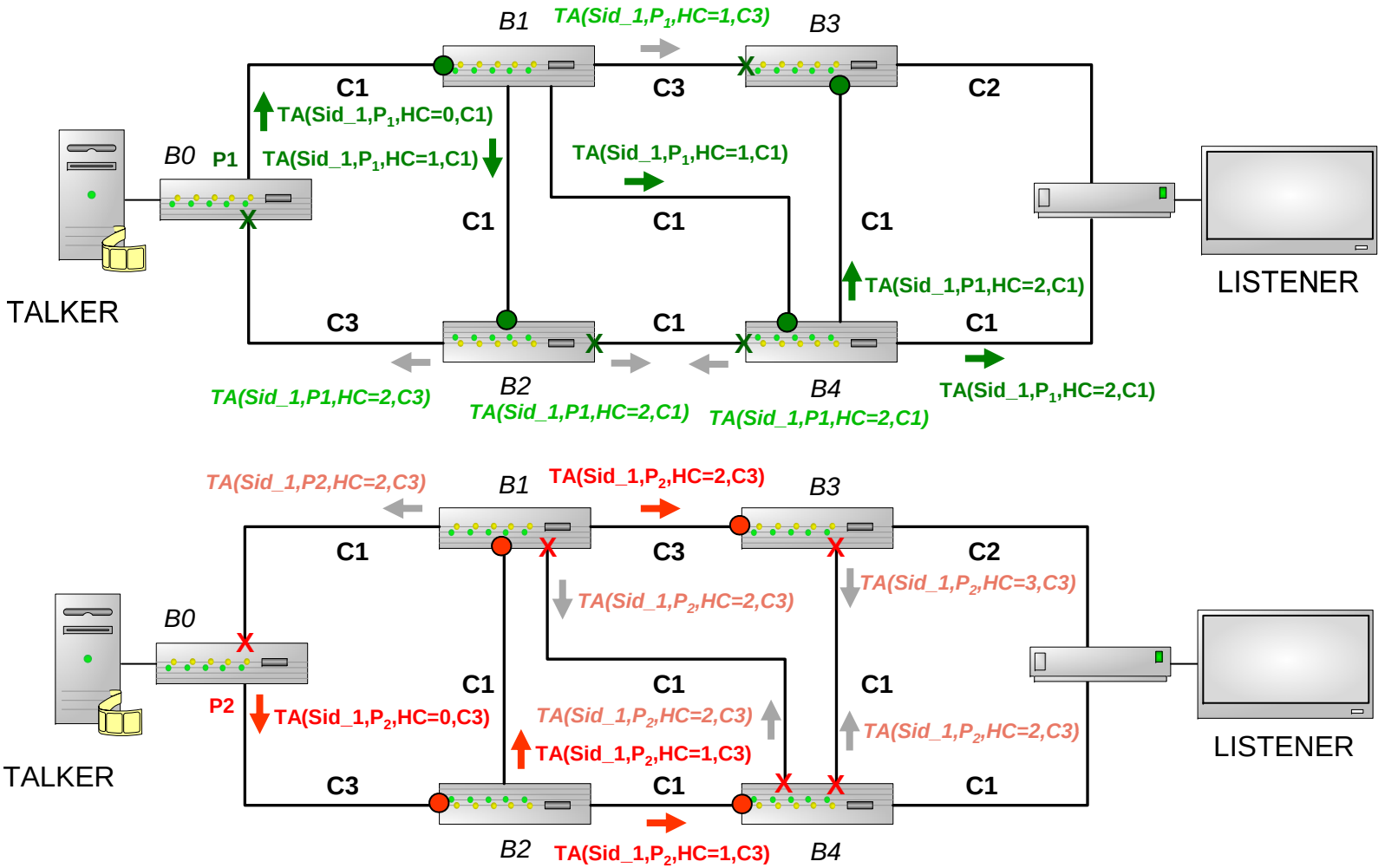
SRP Talker Advertise – Step 2



●● = Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

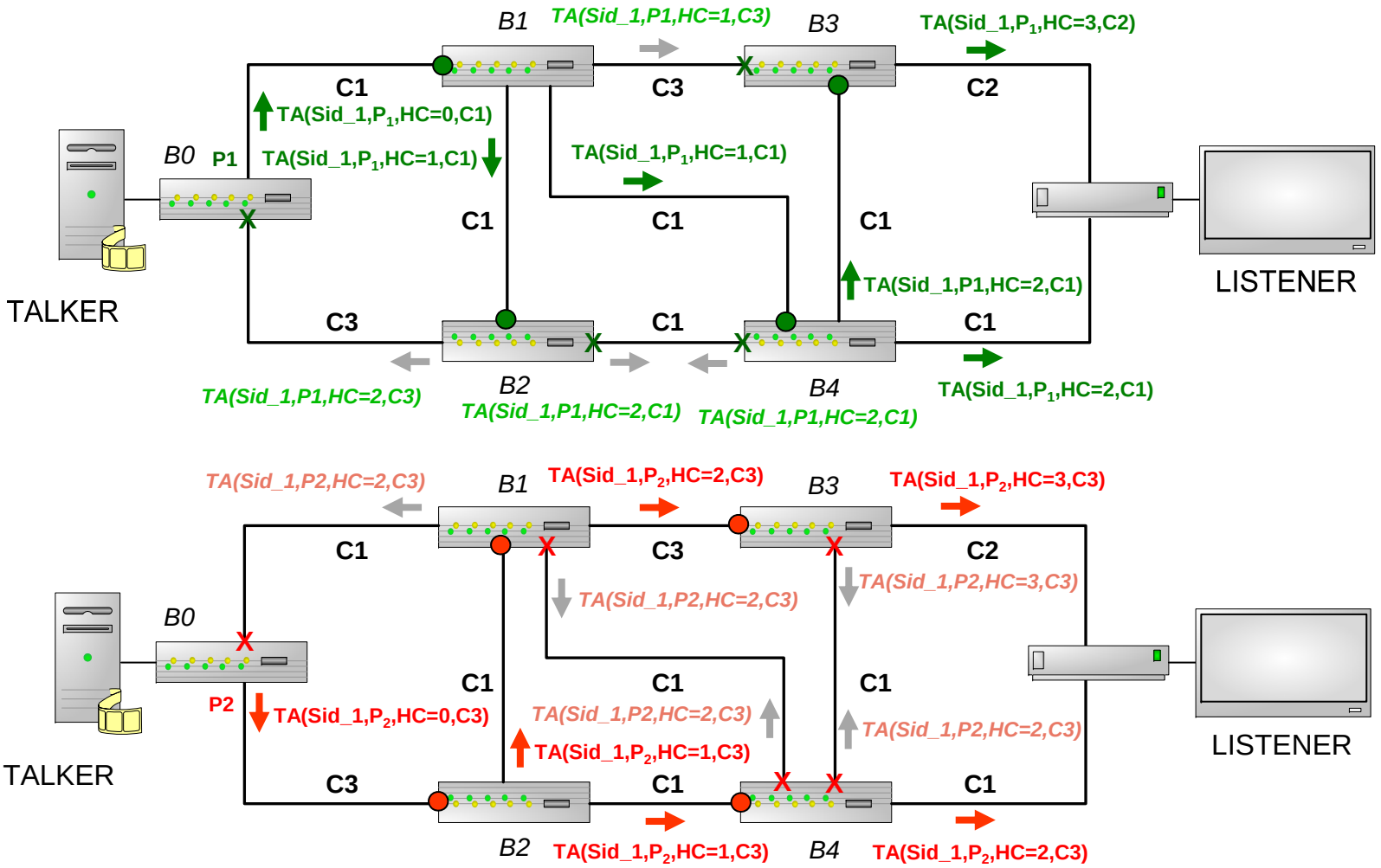
SRP Talker Advertise – Step 3



= Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

SRP Talker Advertise – Step 4

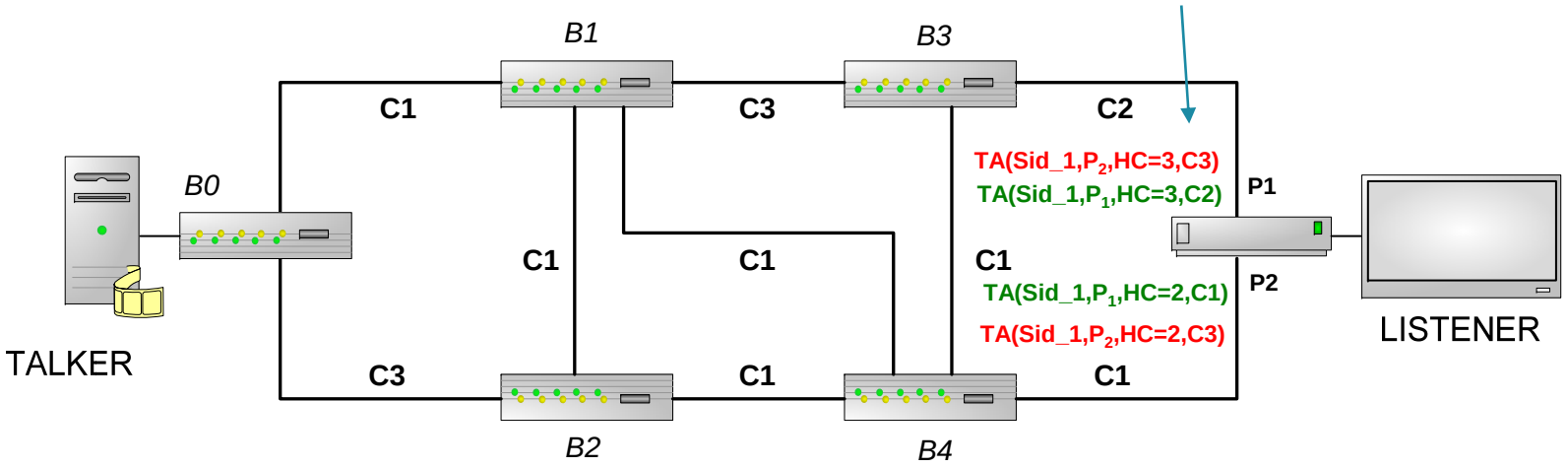


= Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

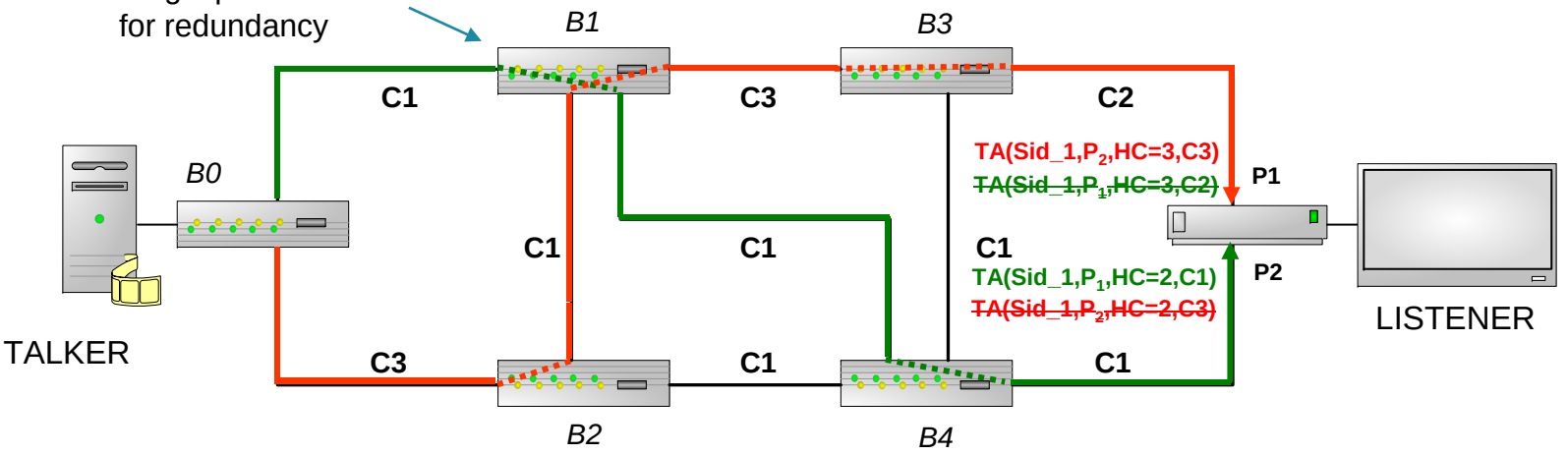
SID = Stream ID
Pi = Port ID
HC = Hop Count

Path Selection

Listener selects **TA(Sid_1,P_1,HC=2,C1)** on P2 as the best path, then could optionally selects **TA(Sid_1,P_2)** on P1 for redundancy



Single point of failure for redundancy



●● = Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

MSRP to select a Stream Path

- Bridge could be configured to select “duplicated” TA on
  - HopCount (delay)
  - or Available BW

- Avantages:
  - Effective: Best available path at the time it is needed with no latency; “on-demand” selection.
  - Autonomous: provide all the current information (link loads and current topology) to make a selection
  - Simple: Zero configuration protocol
  - Compatible: Coexists with STP protocols

Discussion

- What should be the link metrics added to TA ?
  - Available bandwidth
  - Delay
  - PER
  - Link Type
  - Link Speed
  - Max BW ?
- How long should a TA port be blocked ?
  - LeaveAll period ?
- How and how fast does an upstream bridge detect that the downstream bridge is “stalled” (link is up but stream is not forwarded anymore)?
  - if it is detected by not receiving LeaveAll , it is probably not fast enough…

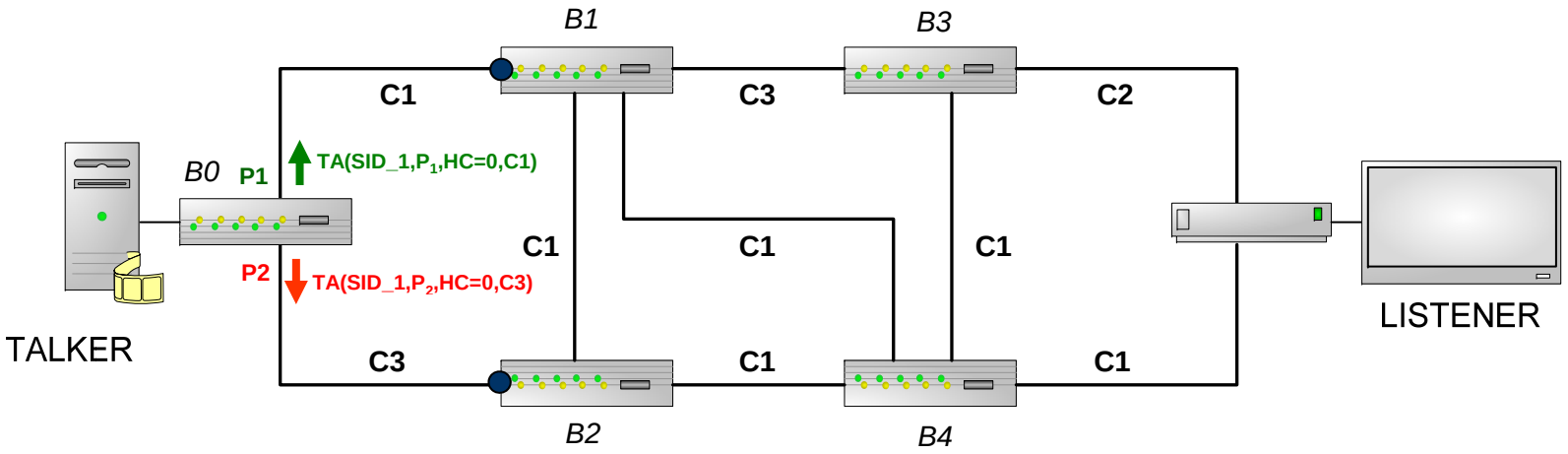
Discussion

- What is the worse case / avg recovery latency ?
  - between retransmitting a TA and the stream resuming at the Listener port….
- How does SRP handle locally configured unicast addresses ?

Proposed Additions To Support Redundancy with /without Single Point of Failure

- New SRP Talker Advertise Attributes:
  - Source Port ID (could either be the Talker ports or any subsequent "splitting" bridge's ports)
  - Link Cost (computed from multiple metrics) or Link Metrics
  - Single Egress (for Redundancy without Single Point of Failure)
- TA are flooded over every bridge's egress port
  - Loop detection by blocking duplicated TAs
- Bridge select between same Stream TA based on:
  - Source Port & Link Cost (configurable computation) / Metrics (configurable precedence)
  - Link Cost AND Single Egress (if no_Single_Point_of_Failure mode is selected)

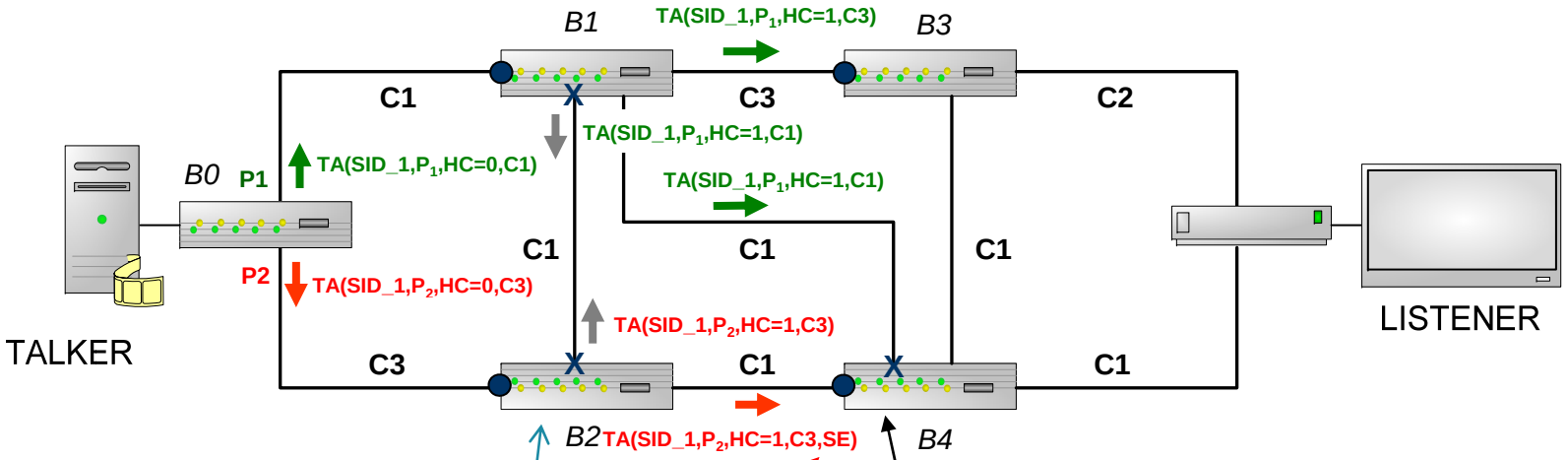
SRP Stream Redundant Path Selection w/o Single Point of Failure – Step 1



● = Talker Port
X = Blocked TA Port
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

SRP Stream Redundant Path Selection w/o Single Point of Failure – Step 2



TA (P_1) has a better cost that the TA (P_2) but since TA (P_2) is a single Egress ,TA (P_2) takes precedence

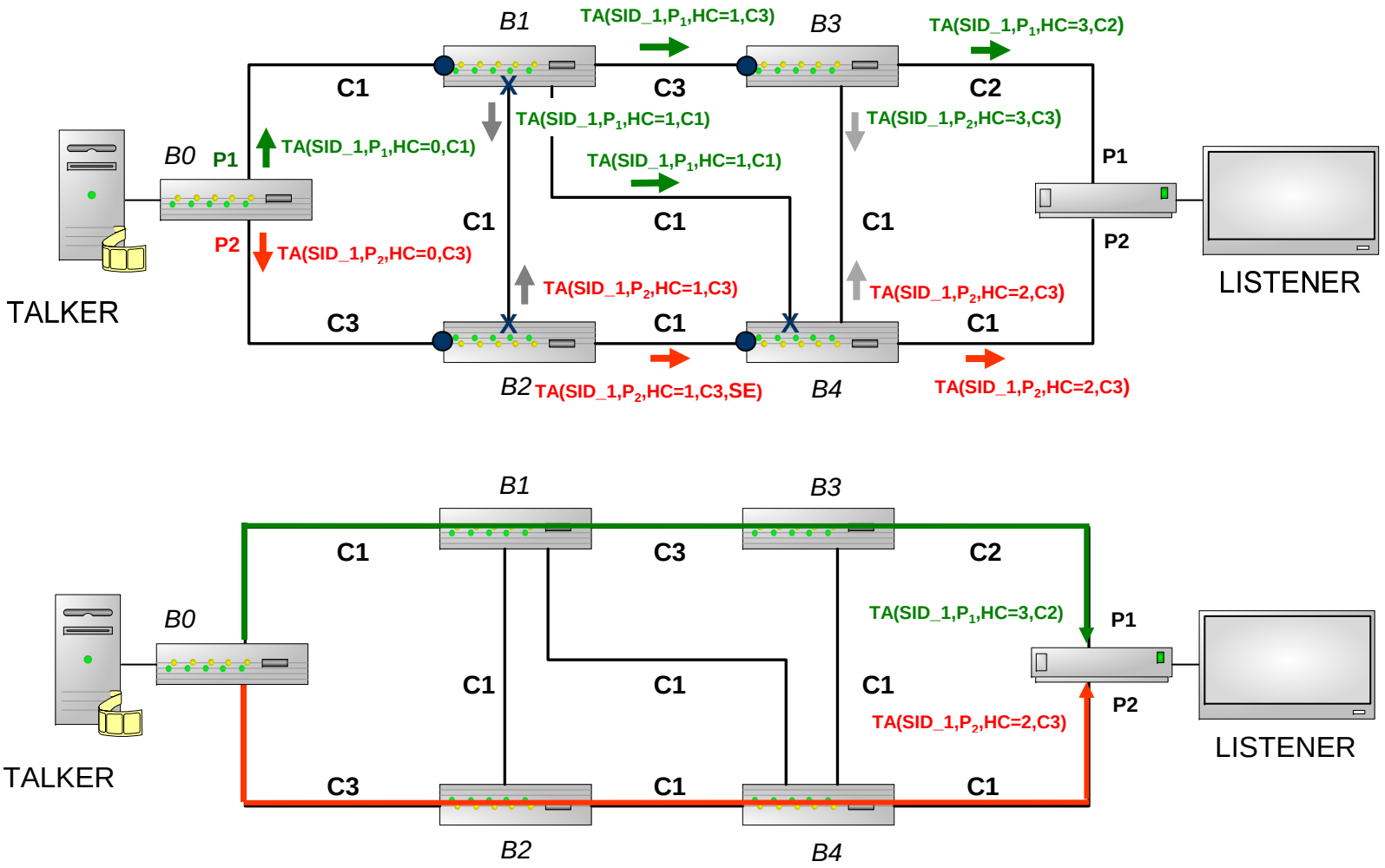
If TA (P_1) was selected as the TA port before PA (P_2) was received, TA (P_1) will be de-registered

Issue: B2 should know that TA(B2-B1) is blocked, in order to detect that TA(B2-B4) is the only egress unblocked port

● = Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

SRP Stream Redundant Path Selection w/o Single Point of Failure – Step 3



●● = Talker Port (SID)
X = Blocked TA Port (SID)
Ci = Link Cost (C1 < C2 < C3)

SID = Stream ID
Pi = Port ID
HC = Hop Count

Notes:

- If a bridge receives multiple TAs with a different port ID and both are Single Egress, the bridge should propagate the TAs but should flag them to indicate that the TA paths are not totally disjointed and therefore cannot provide redundancy without a single point of failure
  - Either the bridge or the listener could signal the management as its indicates a “faulty” (for the redundancy stand point) engineering of the network. How to fix it is out of scope here….
- This scheme cannot guaranty that at least 2 disjointed paths could reach a listener but it is up to the listener to signal it to the management if redundancy without a single point of failure is mandatory
  - Here too its indicates a “faulty” (for the redundancy stand point) engineering of the network. How to fix it is out of scope here….

Thank You