Bit Index Explicit Replication (BIER)
Multicasting in Transport Networks

A. Giorgetti(1), A. Sgambelluri(1), F. Paolucci(1), N. Sambo(1), P. Castoldi(1), F. Cugini(2)
(1) Scuola Superiore Sant’Anna, Pisa, Italy
(2) CNIT, Pisa, Italy,

ONDM 2017
Outline

- Segment Routing (SR)
  - technology
  - Path encoding
  - TE performance (ECMP vs strict routing)
  - Use cases

- Bit Index Explicit Replication (BIER)
  - technology
  - Experimental validation
Segment Routing

- **Segment Routing (SR)** is a traffic engineering (TE) technique compatible with traditional **MPLS data plane**.

- Using SR, a **signaling protocol is not required** and path state is not maintained in intermediate nodes → simplified control plane operation.

- Each packet is forwarded according to an header composed of **segment identifiers** (SIDs), e.g., representing a specific network node.

- SIDs are advertised by properly extended IGP (e.g., OSPF-TE).

- Intermediate nodes forward the received packet along the **shortest path** toward the node indicated in the top SID.
Segment Routing (SR) basic behavior (1/2)

Target path: 1, 2, 4, 5  \rightarrow  segment list: 5
Segment Routing (SR) basic behavior (2/2)

Target path: 1,2,3,4,5  \rightarrow  segment list: 3, 5
Equal Cost Multi Paths (ECMP)

Default behavior → load balancing on ECMPs
Target paths: 1,2,3,4,5 AND 1,2,6,4,5 → segment list: 8
Strict route selection avoiding ECMPs

Target path: 1,2,3,4,5,7,8  \rightarrow  segment list: 3, 8
Path encoding: Segment List with minimum SLD

- **Problem:** given a strict route, identify the segment list having the minimum Segment List Depth (SLD).
- Routers may **not** support large SLD
Path encoding: Segment List with minimum overhead

In ECMP, Deep Packet Inspection (DPI) is needed to perform load balancing on a per-flow basis, avoiding packet misordering → additional HW requirements, not needed in case of strict routes.

One route may have excessive latency, or reliability issues.

Larger segment lists
SR Traffic engineering performance

• Network usage for different topologies

• Default SR behavior exploiting ECMP typically leads to inefficient network usage

• Segment Routing with large Segment List Depth (e.g., 8) guarantees optimal TE solutions

• Surprisingly, Segment Routing with just a Segment List Depth of 3 is able to guarantee optimal TE solutions

<table>
<thead>
<tr>
<th></th>
<th>ECMP</th>
<th>SHP</th>
<th>$\text{SegMR } \kappa = 8$</th>
<th>$\text{SegMR } \kappa = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid $2 \times 2$</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Grid $3 \times 3$</td>
<td>7.00</td>
<td>6.00</td>
<td>6.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Grid $4 \times 4$</td>
<td>18.63</td>
<td>16.00</td>
<td>16.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Grid $5 \times 5$</td>
<td>36.75</td>
<td>30.00</td>
<td>30.00</td>
<td>30.00</td>
</tr>
<tr>
<td>Eurocore</td>
<td>4.96</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td>NFSNET</td>
<td>15.33</td>
<td>13.00</td>
<td>13.00</td>
<td>13.00</td>
</tr>
<tr>
<td>EON</td>
<td>25.00</td>
<td>18.00</td>
<td>18.00</td>
<td>18.00</td>
</tr>
<tr>
<td>UKNET</td>
<td>27.63</td>
<td>21.00</td>
<td>19.00</td>
<td>19.00</td>
</tr>
<tr>
<td>ITALNET</td>
<td>48.38</td>
<td>33.00</td>
<td>28.00</td>
<td>28.00</td>
</tr>
<tr>
<td>Arpanet</td>
<td>35.88</td>
<td>33.00</td>
<td>33.00</td>
<td>33.00</td>
</tr>
<tr>
<td>Eurolarge</td>
<td>131.60</td>
<td>88.04</td>
<td>66.00</td>
<td>66.00</td>
</tr>
</tbody>
</table>

Number of flows in the most utilized link
• ECMP: default SR behavior ($k=1$)
• SHP: shortest path, no ECMP
• SEGMR: SR with $k=8$ and $k=3$

$k$: segment list depth

Uniform traffic matrix is assumed. Results are confirmed for non-uniform matrices.

Use cases for SR: Optical bypass

• Dynamic selection of pre-established optical bypass, with no signaling

A Sgambelluri, F Paolucci, A Giorgetti, F Cugini, P Castoldi
“Experimental demonstration of segment routing”, JLT 2016
Use cases for SR: multi-domain

- No e2e signaling (critical in multi-domain/vendor scenarios)
- Compressed Segment Lists to limit label stacking
- Confidentiality

Other use cases for SR

- Recovery:
  - SR-FAILOVER: rerouting to the destination
  - SR-DETOUR: rerouting to the next(-next) hop

- OAM

- Service chaining
Bit Index Explicit Replication (BIER)

- BIER has been recently proposed for P2MP
- As in SR,
  - No signaling protocol
  - No forwarding state at intermediated nodes
  - the ingress router applies a specifically designed label (here called BitString) which defines the forwarding actions

- In the BIER BitString, each bit represents exactly one egress router in the domain.
- Forwarding is then performed by each intermediate node by just processing and updating the BitString

- In large networks, a hierarchical structure of the BIER header is used.
## BIER
### Basic behavior

**Bit Indexed Forwarding Tables**

<table>
<thead>
<tr>
<th>BitMask</th>
<th>Neighbour</th>
<th>Port</th>
<th>BitMask</th>
<th>Neighbour</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000001</td>
<td>B1</td>
<td>1</td>
<td>00000100</td>
<td>-</td>
<td>local</td>
</tr>
<tr>
<td>00000100</td>
<td>-</td>
<td>local</td>
<td>11011011</td>
<td>B2</td>
<td>1</td>
</tr>
<tr>
<td>00110000</td>
<td>B4</td>
<td>2</td>
<td>00100000</td>
<td>B6</td>
<td>2</td>
</tr>
<tr>
<td>00100100</td>
<td>B3</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11000000</td>
<td>B8</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

BITF table scales with the number of outgoing links and not with the number of flows traversing the node.
BIER Implementation with OF controller

OpenFlow

SDN controller

BIER agent

Open vSwitch
BIER implementation

- SDN network controller implemented in Ryu, with OpenFlow 1.3
- Packet nodes: Open vSwitch + a BIER agent in each node implemented on a local instance of the Ryu controller
- The BIER agent stores the list of nodes reachable with a shortest path using each of the outgoing links (BIFT table - provided and updated by the network SDN controller.
- When a multicast request arrives, the SDN controller only has to enforce the proper BIER header at the ingress (no signaling protocol).
- A group is created with the id equal to the numerical value of the BitString including all recipients joining the multicast address.
- If packets need to be replicated at the specific node, the group will include a number of buckets.

<table>
<thead>
<tr>
<th>Node 2: Flow Table</th>
<th>Node 2: Group Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow match</td>
<td>Action</td>
</tr>
<tr>
<td>MPLS label 10100</td>
<td>group 10100</td>
</tr>
<tr>
<td>MPLS label 10010</td>
<td>group 10010</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Capture

(a) - message exchange

(b) GROUP_MOD
Node C

(c) FLOW_MOD
Node C
Conclusions

- Overview of Segment Routing technology and use cases
  - Path encoding
  - Traffic engineering (limited SLD are typically adequate)
  - Multi-domain (no e2e signaling)
  - Recovery (straightforward)
  - OAM (easy mechanism to probe the network)
  - Service chaining (services can be described with Segment ID)
- BIER
  - Experimental demonstration
thank you!

email:
filippo.cugini@cnit.it