

Interconnecting Millions Of Endpoints with Segment Routing

[draft-filsfils-spring-large-scale-interconnect-00](#)

C. Filsfils, D. Cai , S. Previdi, Cisco

W. Henderickx, Alcatel-Lucent

R. Shakir, BT

D. Cooper, F. Ferguson, Level3

T. LaBerge, S. Lin, Microsoft

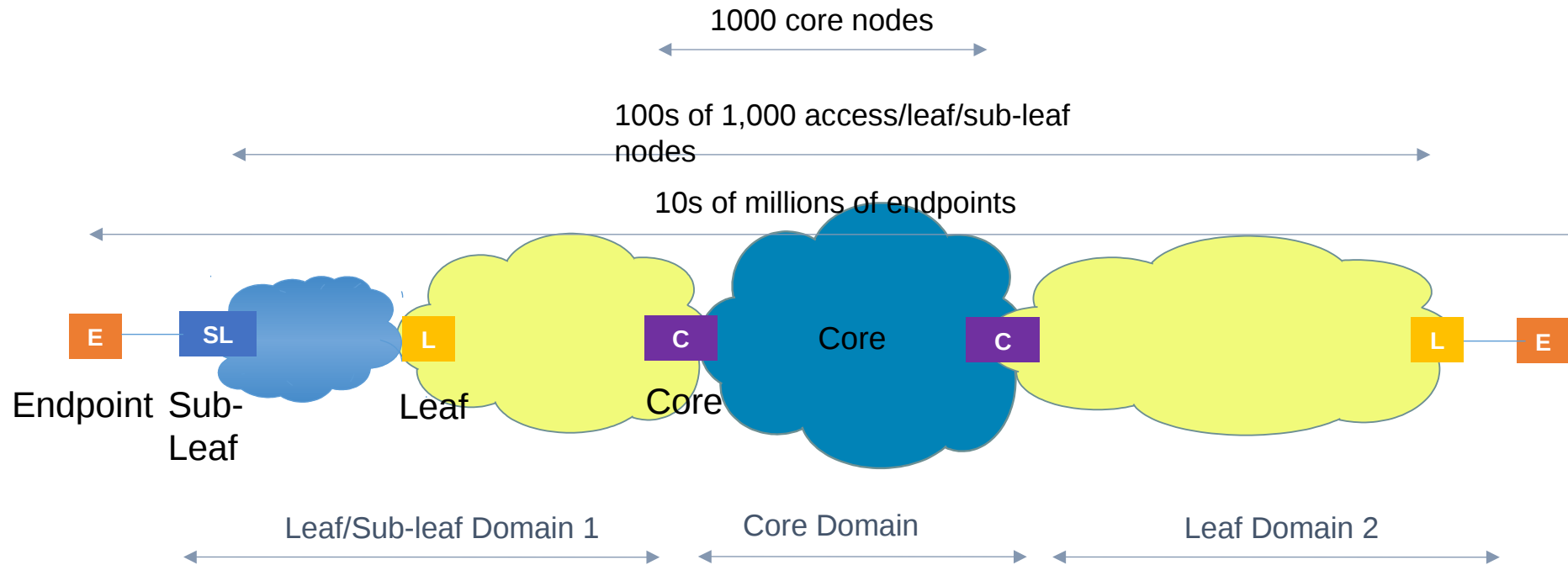
B. Decraene, Orange

L. Jalil, Verizon

IETF 93 SPRING, July 2015, Prague

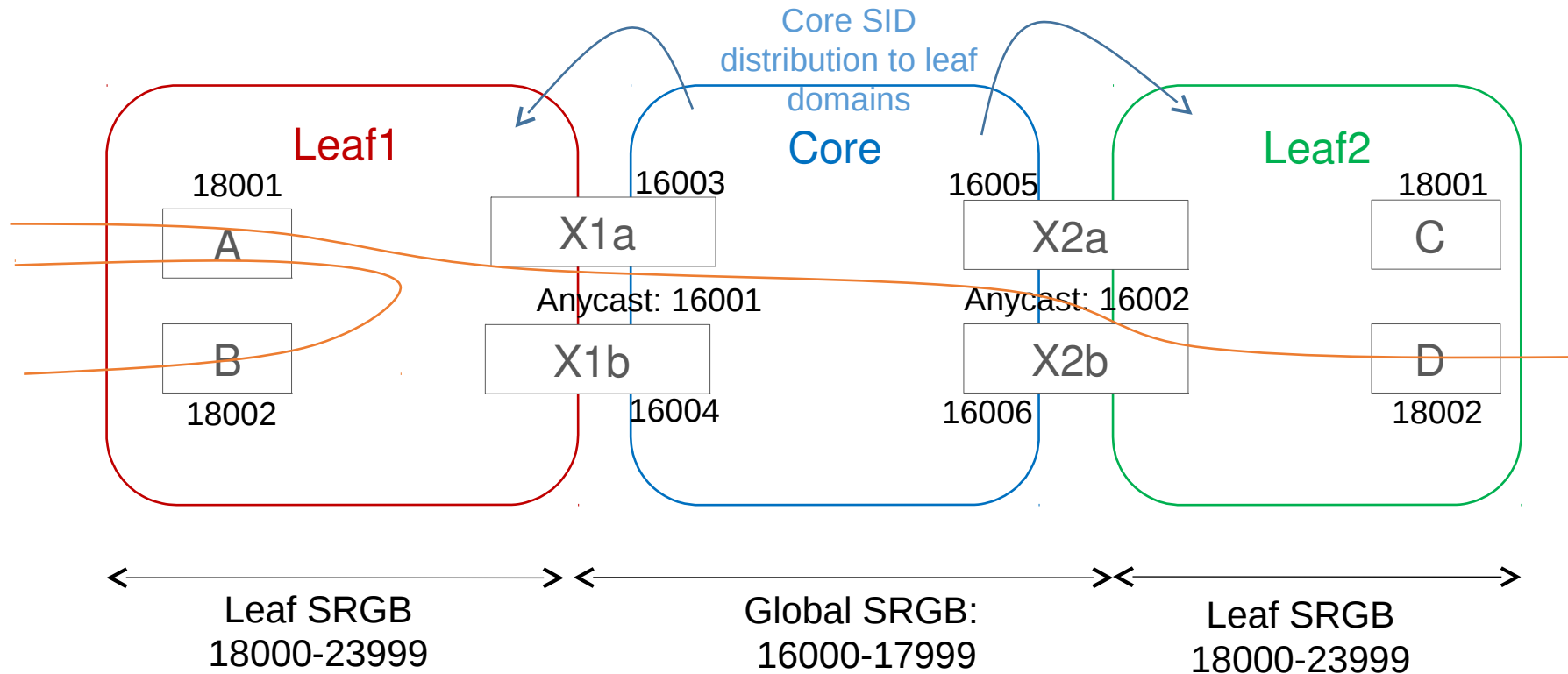
Problem Statement: Use SR for Large-scale Network Design

100s of 1000s of network nodes, 10s of millions of endpoints



Reference Design

Intra-leaf (A → B): shortest-path {18002}
Inter-leaf (A → D): shortest-path via X
{16002, 18002}



Scale Example

1 core domain + 100 leaf domains
 Core domain: 200 nodes, 300 SIDs
 (node+anycast)

Leaf domain: 6,000 nodes, 6,000 node SIDs
 each leaf: 500 endpoints, 500 adj SIDs

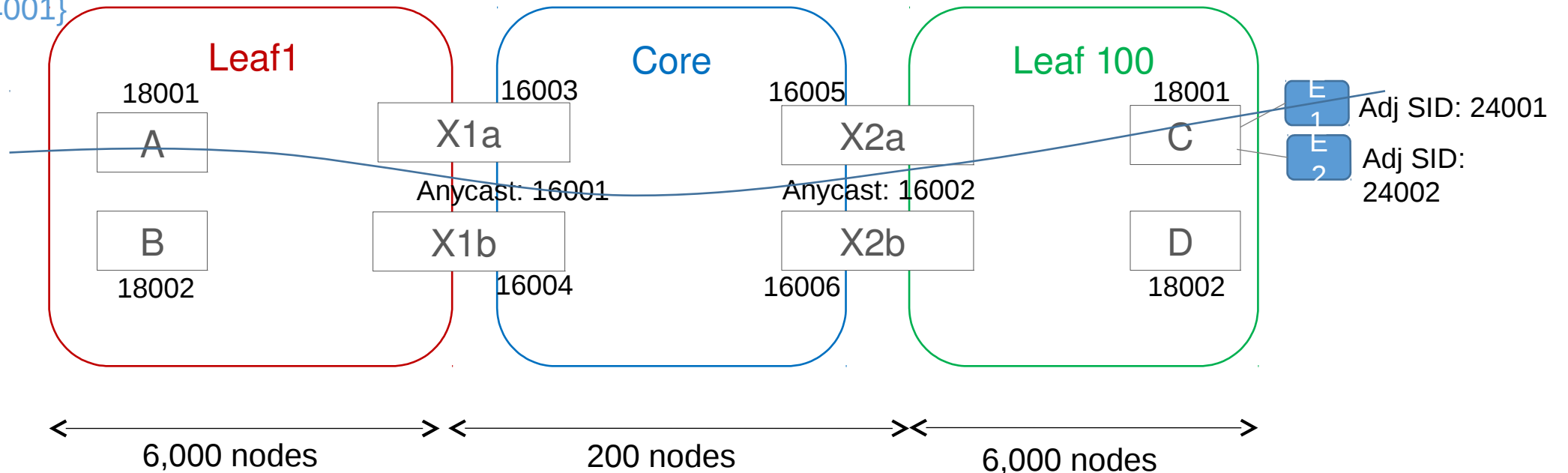
Network-wide scale*

- Node scale: $100 \times 6,000 = 600,000$
- Endpoint scale: $100 \times 6,000 \times 500 = 3M$

Per-node scale*

- Leaf node = $300(\text{core}) + 6,000(\text{leaf}) + 500(\text{adj}) = 6,800$
- Core node = $300(\text{core}) + 6,000(\text{leaf}) = 6,300$

Inter-leaf (A → E1): shortest-path via X {16002, 18001, 24001}



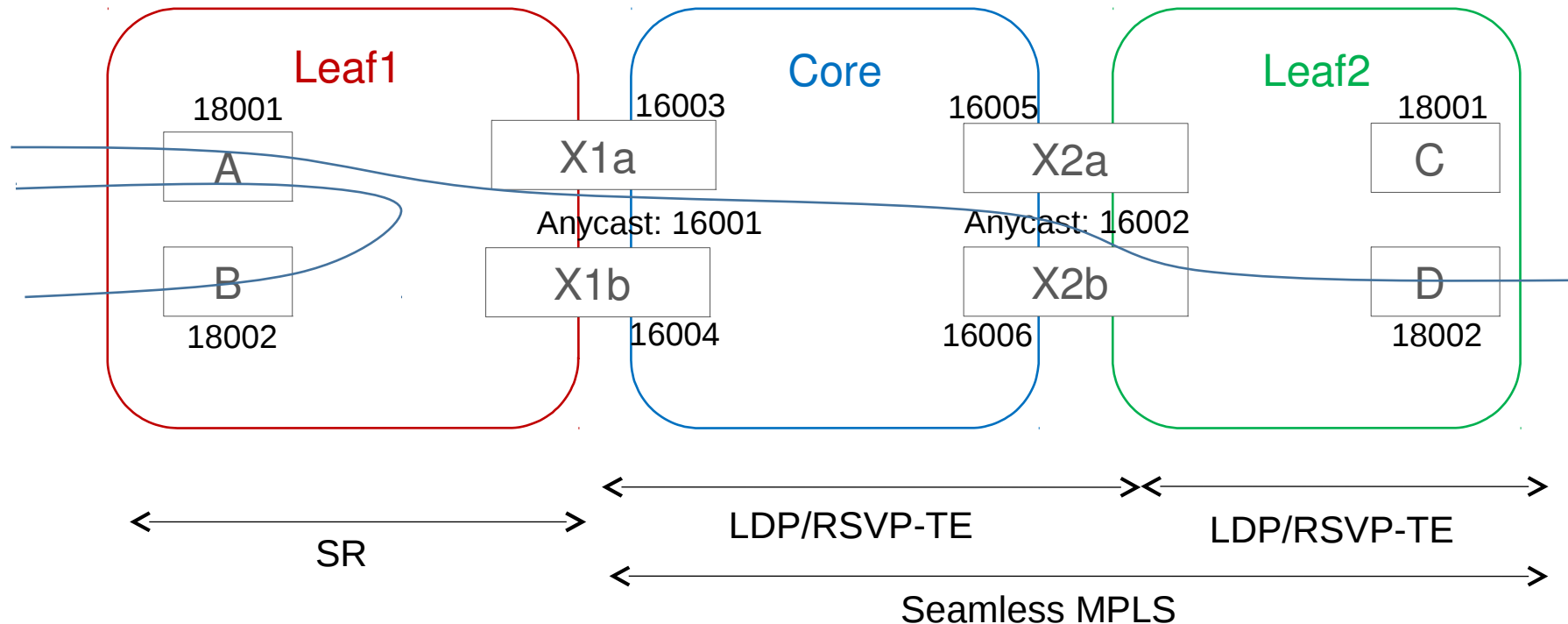
* Link adj SID is not included

Optional Designs

- Core node SID distribution to leaf domains
 - With distribution: two SIDs required for inter-leaf SR path
 - Without distribution: three SIDs required for inter-leaf SR path
- Sub-leaf domain
 - Sub-leaf domain can be created for low end leaf node or for design flexibility
- Traffic engineering
 - TE can be used for both intra-leaf and inter-leaf paths
 - Biding SID can be used to “compress” the SID list on the leaf node

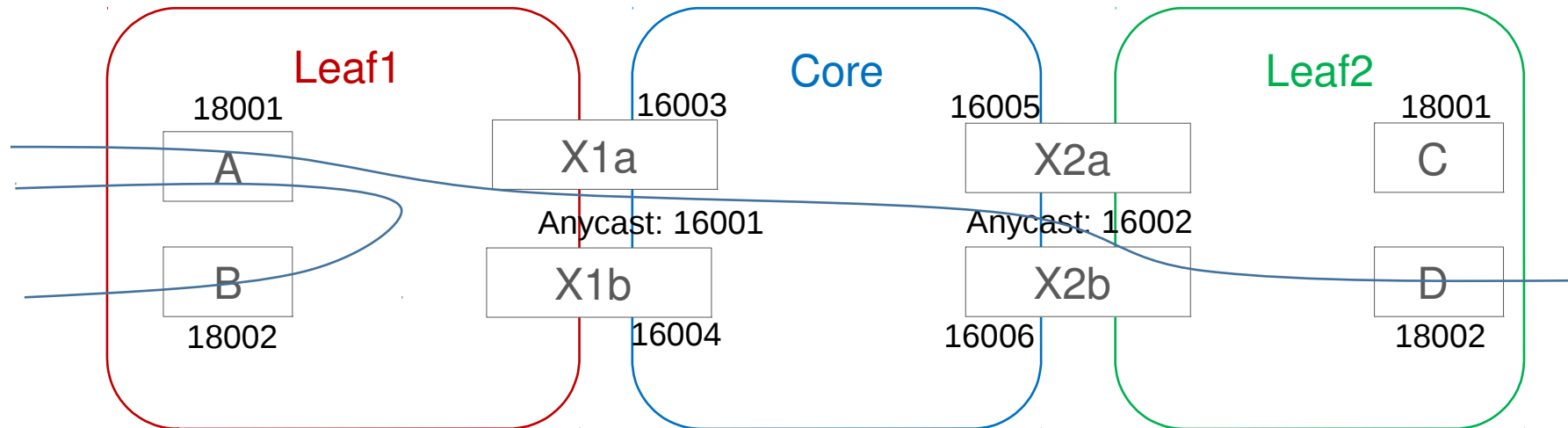
Deployment Models

- Green field: end-to-end SR
- Brown field: Inter-operate with LDP/RSVP-TE and seamless MPLS network



Benefit

- ECMPs
- TI-FRR
- Simple “X” node redundancy



Summary

- Use SR to scale the network to support hundreds of thousands of network nodes, and tens of millions of underlay endpoints
- Flexible network design, e.g. flexible size of the leaf domain, sub-leaf domains
- Inter-operability with existing network design, e.g. LDP/RSVP-TE, seamless MPLS design
- Optimized forwarding and HA: ECMPs, TI-FRR, simple node redundancy