

Interconnecting Millions Of Endpoints with Segment Routing

draft-filsfils-spring-large-scale-interconnect-01

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

J. Tantsura, Ericsson

IETF 94 SPRING, November 2015, Yokohama

What's new since version 00?

- Very simple, no much change
- Add new co-author
 - Jeff Tantsura, Ericsson

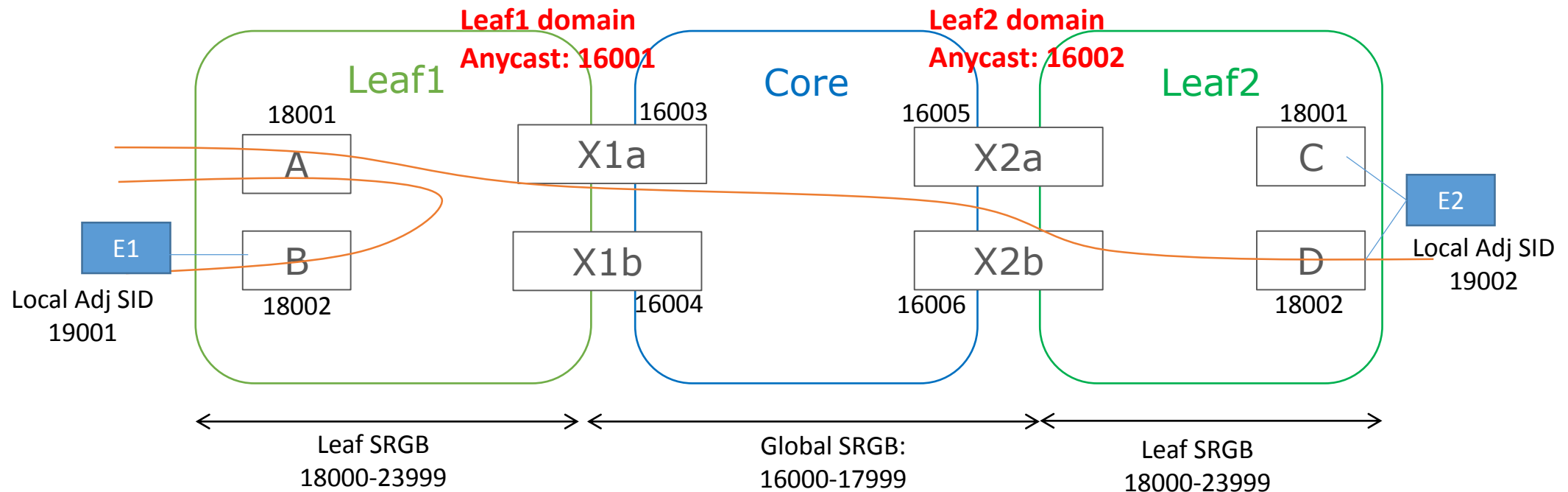
Problem Statement Re-cap

- Not like IP, there is no such concept of the label summarization and default label
- For the MPLS network, each node need specific label for the forwarding
- Millions of nodes/endpoints means millions of RIB/FIB, which is not desirable for the low cost DC switches or SP metro access nodes

The Principle and Reference Design

Using hierarchical label stack to solve large scale MPLS network

- Network is divided into 2 or 3 layers: core, leaf and sub-leaf (or local endpoint) optionally
- Each leaf domain is reachable via domain label (thinking zip code). Domain label is anycast SID which is advertised by the leaf border routers
- Endpoint use local adj SID which is behind one or multiple leaf nodes

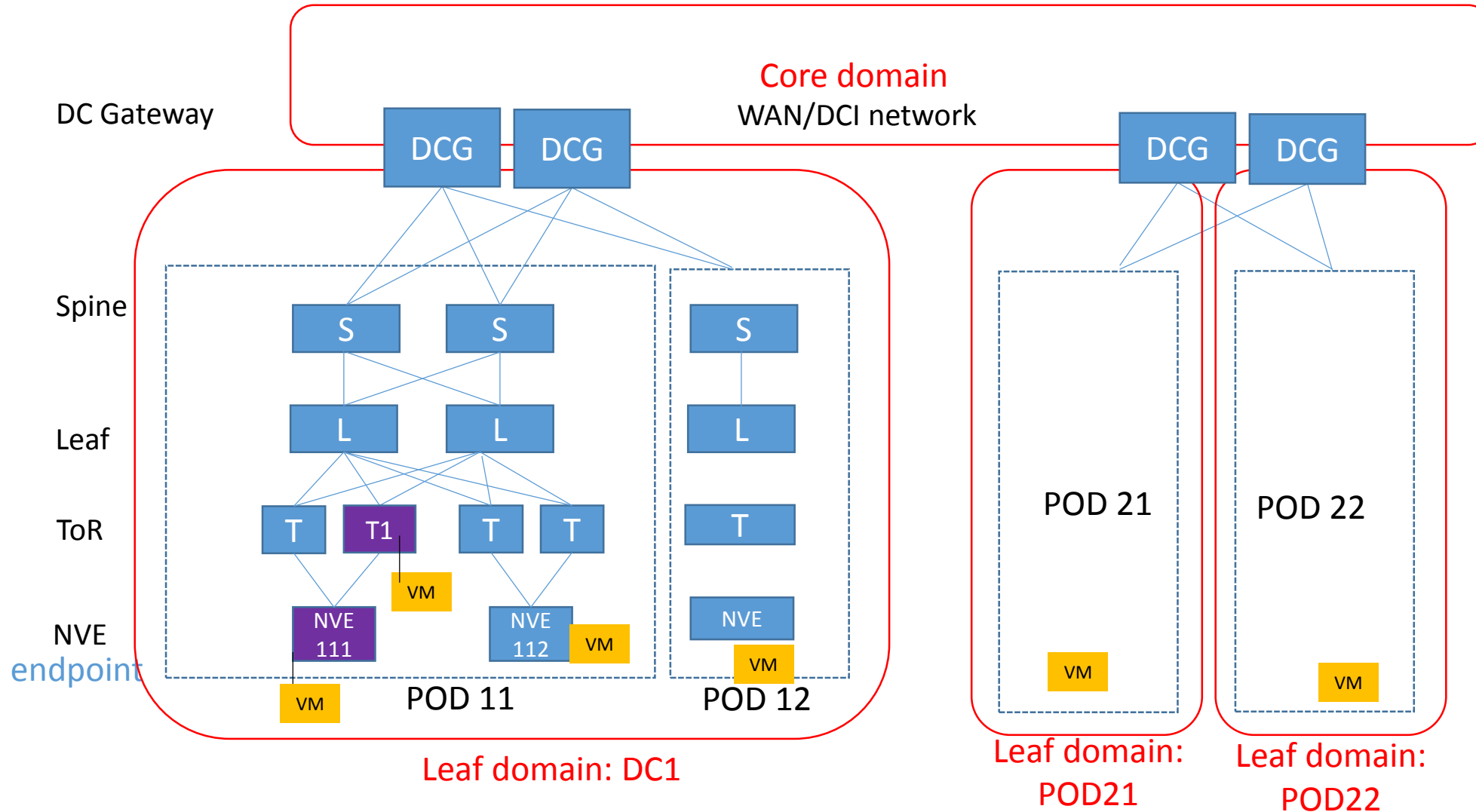


Intra-leaf (A → B): shortest-path {18002}

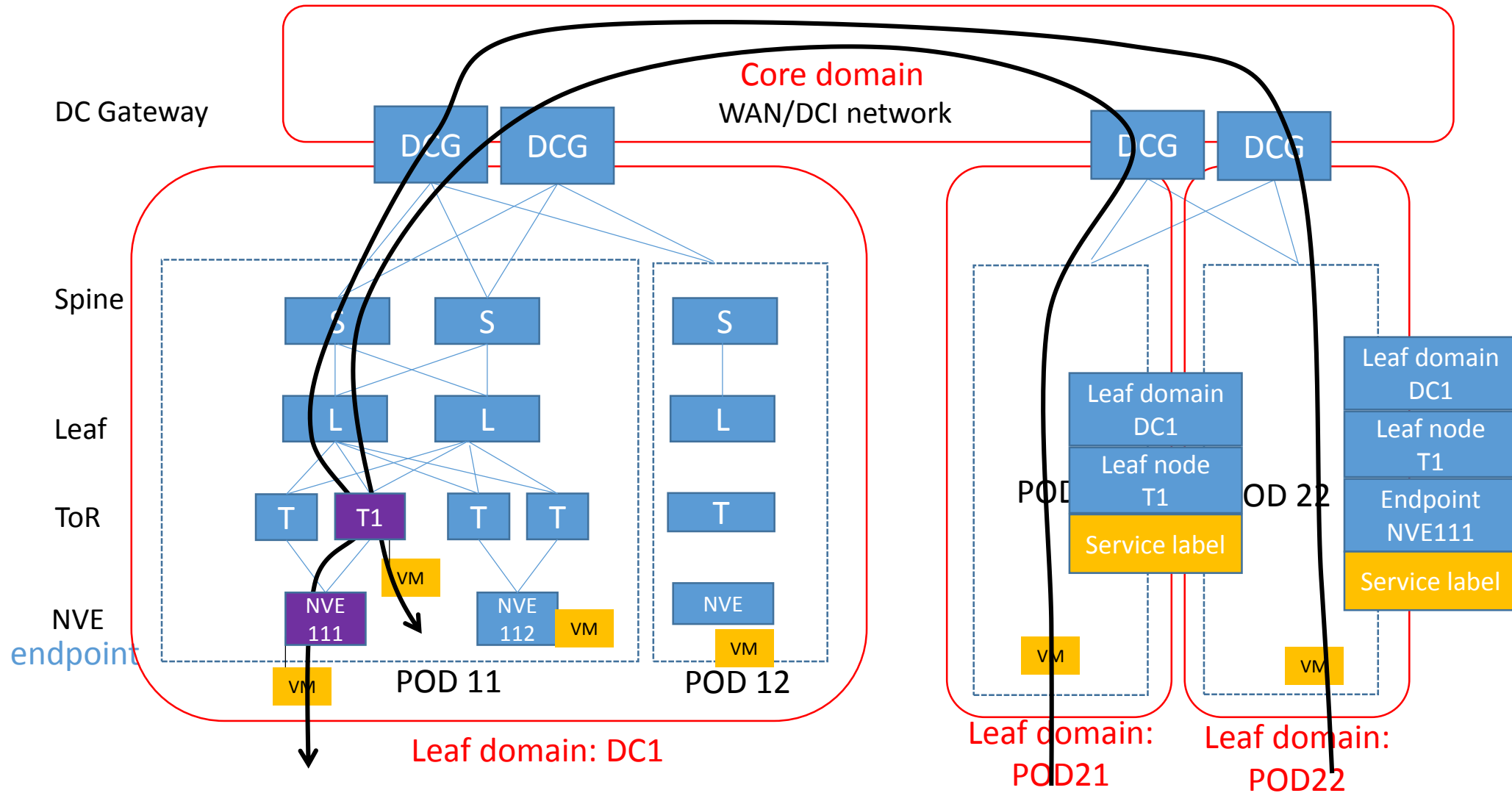
Inter-leaf (A → D): shortest-path via X {16002, 18002}

Inter-leaf (E1 → E2): shortest-path via X {16002, 18002, 19002}

Large-scale DC Network Use Case

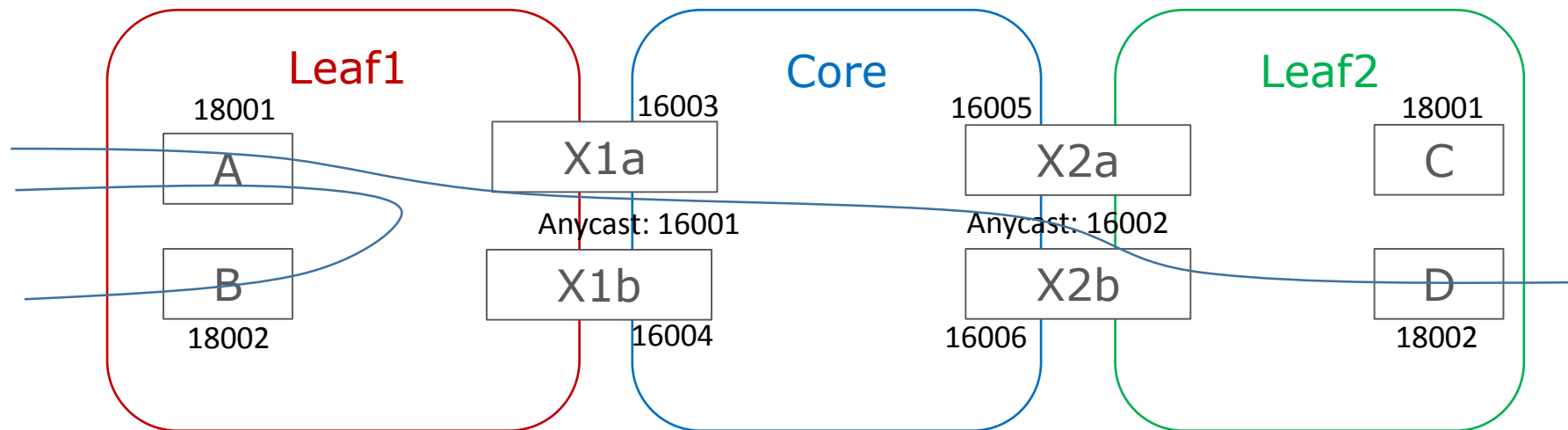


Large-scale DC Network Use Case



Benefit

- A simple way to scale MPLS network using existing SR, no protocol changes
- Fully leverage the distributed SR in each domain: ECMPs, TI-FRR and SR-TE
- Simple “X” node (border node) redundancy using anycast SID
- Fully inter-operate with existing network protocols and design: LDP, seamless MPLS



Questions/Comments?