FAR: A Fault-avoidance Routing Method for Data Center Networks with Regular Topology

http://datatracker.ietf.org/doc/draft-sl-rtgwg-far-dcn/

Please send comments to rtgwg@ietf.org

- Drawback and future work
- Advantages
- Use case
- Principle and framework
- Differences
- What is FAR
- Background

Outline

Background

- With rapid development of cloud computing tech nologies, scale of a data center is growing up qu ickly.

Background

- Traditional tree-like architectures and routing protocols are not suitable for building large-scale networks.
- Some new network architectures, such as Fat-tree, BCube, are applied to data center networks.
- To maximize benefits of new architectures, some new routing methods are proposed according to the features of Fat-tree, BCube’s topologies.

What is FAR

- FAR is a generic routing method and framework for large-scale data center networks.
- FAR protocol is well designed to fully leverage th e regularity in the topology of networks.
- FAR is a high-performance routing method whic h computes routing tables in a simplistic manner.

Differences between FAR and other routing methods

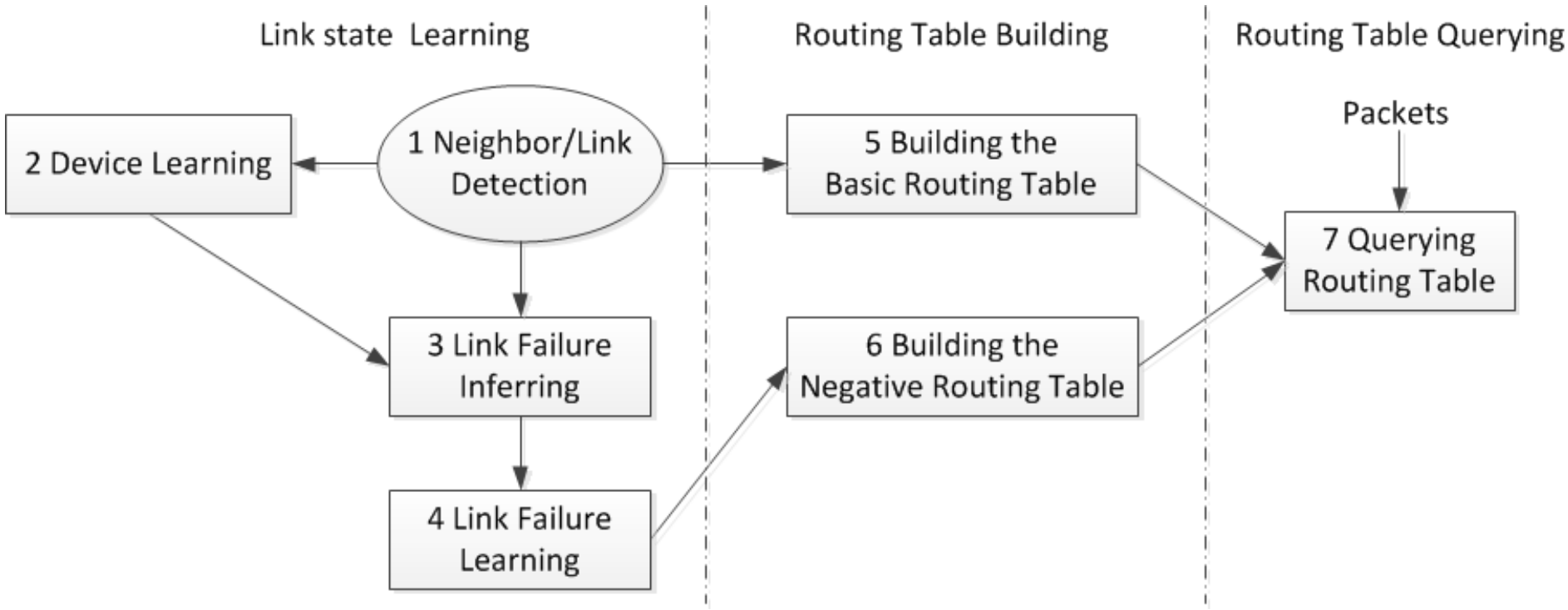
- OSPF, IS-IS or RIP works in an arbitrary network, but FAR is d esigned for regular topologies.
  - A regular topology means the distribution of nodes, addressing and conn ections are well designed, so a node knows the whole topology without l earning in a network.

- Other than some routing methods for specific networks such a s Fat-tree and BCube, FAR is a generic routing method suitab le for any network with a regular topology.

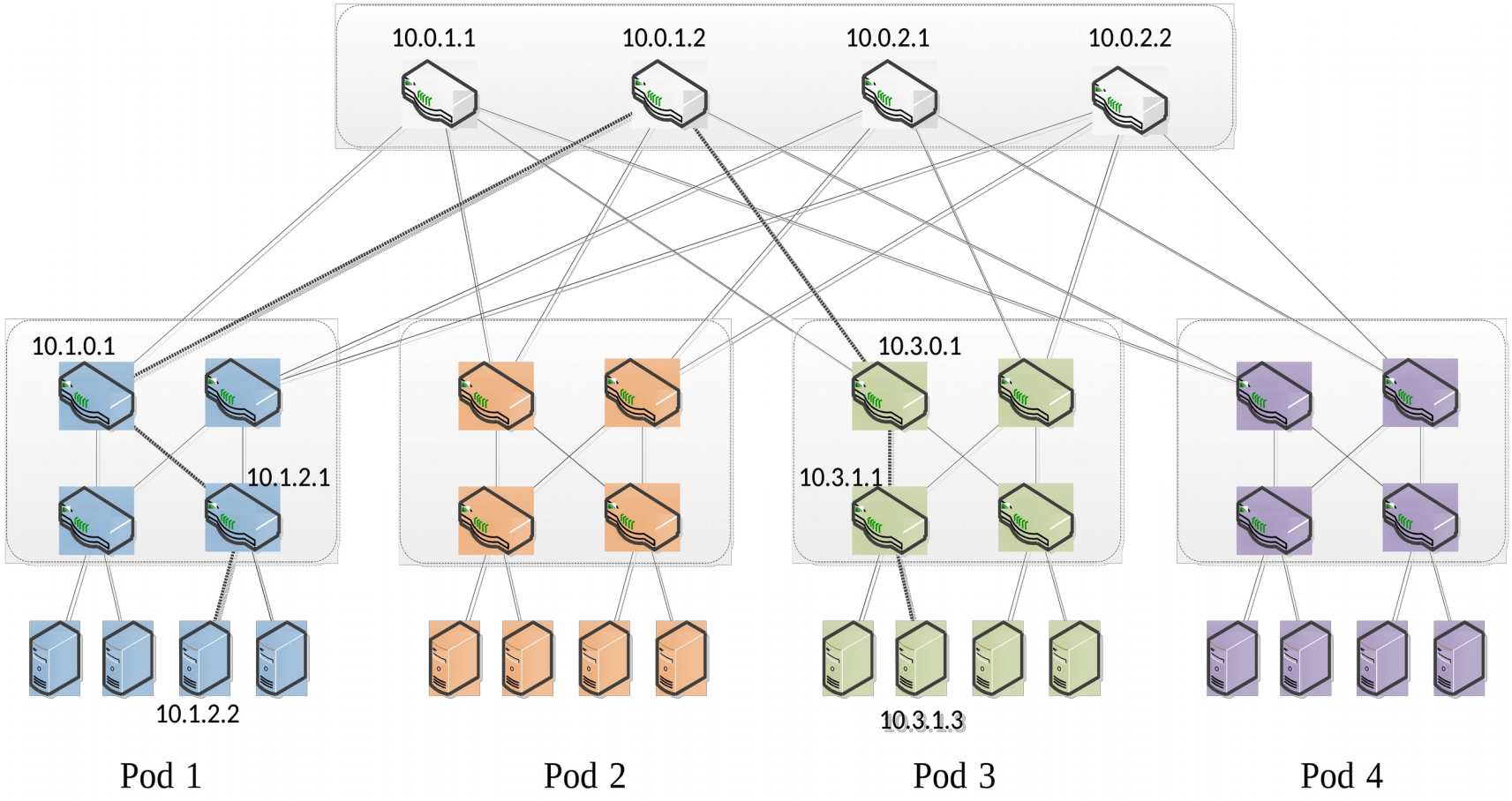
The Principle of FAR

- Network devices, including routers, switches, and servers, are a ssigned IP addresses according to their location in the network.
- A basic routing table (BRT) is built based on local topology.
- A negative routing table (NRT) is built based on link and device f ailures in the entire network.
- Look up both BRT and NRT to determine the final route in a rou ting procedure.
- Final routes = matched routes in BRT - matched routes in NRT.

The Routing Framework of FAR



Use Case (Fat-tree Network)



The BRT of aggregation switch 10.1.0.1

- It is easy to build a BRT for a router according to its local topol ogy
- We take 10.1.0.1 as an example. Its BRT is:

Destination/Mask	Next Hop
10.1.1.0/255.255.255.0	10.1.1.1
10.1.2.0/255.255.255.0	10.1.2.1
10.0.0.0/255.0.0.0	10.0.1.1
10.0.0.0/255.0.0.0	10.0.1.2

The NRT of aggregation switch 10.1.0.1

- A router's NRT is determined by locations of link or device fail ures in the network.
- Suppose the link between 10.0.1.2 and 10.3.0.1 fails, The NR T of 10.1.0.1 is:

Destination/Mask	Next Hop
10.3.0.0/255.255.0.0	10.0.1.2

Node 10.1.0.1 forward a packet to node 10.3.1.3

- 1) Calculate candidate hops. 10.1.0.1 looks up its BRT and ob tains the following matched entries:

Destination/Mask	Next Hop
10.3.0.0/255.255.0.0	10.0.1.1
10.3.0.0/255.255.0.0	10.0.1.2

So the candidate next hops = {10.0.1.1; 10.0.1.2}.

- 2) Calculate avoiding hops. 10.1.0.1 looks up its NRT and obt ains the following matched entries:

Destination/Mask	Next Hop
10.3.0.0/255.255.0.0	10.0.1.2

So the avoiding hops = {10.0.1.2}

- 3) Calculate applicable hops.

applicable next hops = {10.0.1.1; 10.0.1.2} – {10.0.1.2} = {10.0.1.1}

- 4) Finally, forward the packet to the next hop 10.0.1.1.

Advantages of FAR

- FAR is a generic routing method suitable for most data center s with regular topologies.
- FAR is a high-performance routing method which supports ver y large-scale networks.
- A FAR switch is simple and cheap, so it can lower the constru cting and operating cost of a data center.

Drawback of FAR and future work

- FAR in this proposal doesn’t give an universal method to calc ulate routing tables for various of network topologies. We sho uld design different method for each type of topology.

- Now we are solving the problem above. We have invented a T DL (topology definition language) to describe a regular topolo gy, and based on TDL, we can design an universal method to calculate routing tables for FAR switches.

Requested actions from the WG

- Routing methods based on regular topology have great a dvantages in large-scale next-generation data centers.
- In the past, no draft has discussed routing problem in reg ular network topology in Data Centers.
- All we need to do now is to propose the problems in the I ETF.
- Requesting IETF Rtg WG to consider adoption of this dr aft and then standardize the solutions.

Thanks and Q&A!