

A method of creating experimental network with routing between virtual hosts

Seiichi YAMAMOTO¹, Eiji KAWAI²

1) yama@wide.ad.jp

- National Institute of Information and Communications Technology
- The University of Tokyo

2) eiji-ka@nict.go.jp

- National Institute of Information and Communications Technology

Summary

- A method of creating specific network route for some experiment with virtual hosts.
- Traffic engineering on IP routing structure with docker image.
- L2VPN on IP routing provides Layer 2 connection between the hosts.

Background

- Generally, a tunnel connection between hypervisors is used for network experiment between virtual hosts.
- The tunnel connection is the shortest path between hypervisors.
- So that it is difficult to make flexible path creation.
 - Ex) specific path for NFV verification, long-distance path for inter-planet network experiment.

Proposal

- Create router as traffic engineering point (like a “hinge”).
- Use IP routing, because of the ease of creating tunnel and selecting path.
- Practically, Use docker router image.
- “IP routing” is one of the choice to do a traffic engineering.
- However, “Segment routing” would be the another candidate to do the traffic engineering.

Implementation

- Use virtual machines or physical machines as docker nodes (i.e. Hypervisors).
 - Prepare docker swarm cluster to do unified control all docker containers.
 - Use docker-machine tool at docker manager node to control container all together.
- Prepare 2 type of containers, “experiment use” and “network use”.
- Create point-to-point network segment between containers with docker overlay network driver.
- Create IP routing network between network containers.
 - BGP routing is used with IP router container image (vyos) on network container.
 - Use vyos L2VPN, when Layer 2 connection between experiment containers is needed.
 - Detailed policy control is available with BGP routing.
- Provide experiment containers with IP routing network connection.

Reference

- Docker
 - <https://www.docker.com/>
- Mininet
 - <http://mininet.org/>
- Linux namespace
 - <http://man7.org/linux/man-pages/man7/namespaces.7.html>
- openstack neutron
 - <https://docs.openstack.org/security-guide/networking/architecture.html>
- segment routing
 - <http://www.segment-routing.net/>

Past presentation

- Seiichi YAMAMOTO, 活用研究会を通じたNICT総合テストベッド利用体験事例, スマートIoT推進フォーラム 技術戦略検討部会 テストベッド分科会, Sep. 2018. (Japanese)
 - <https://testbed.nict.go.jp/bunkakai/pdf/bunkakai-05-04.pdf>

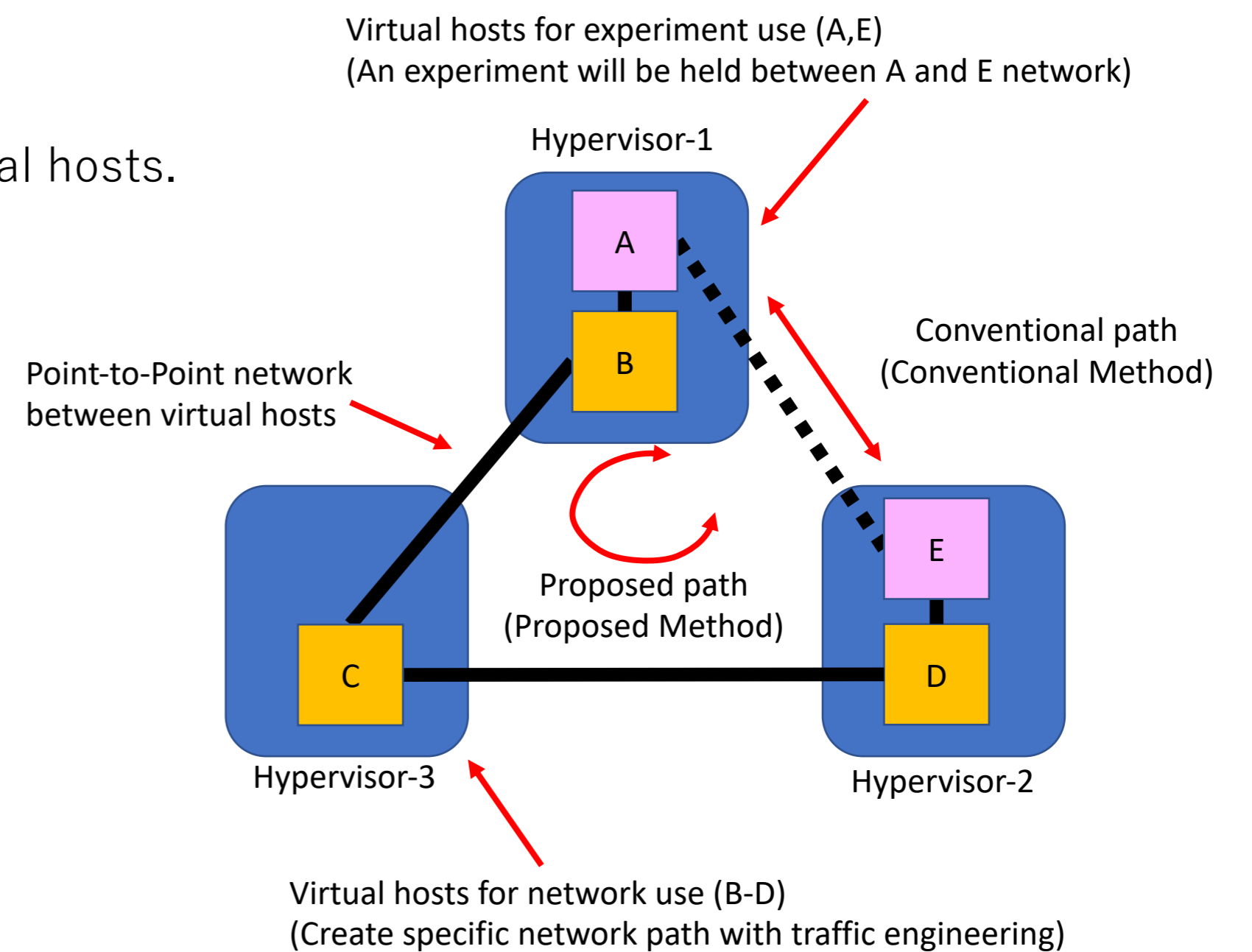


Fig 1. Proposed system

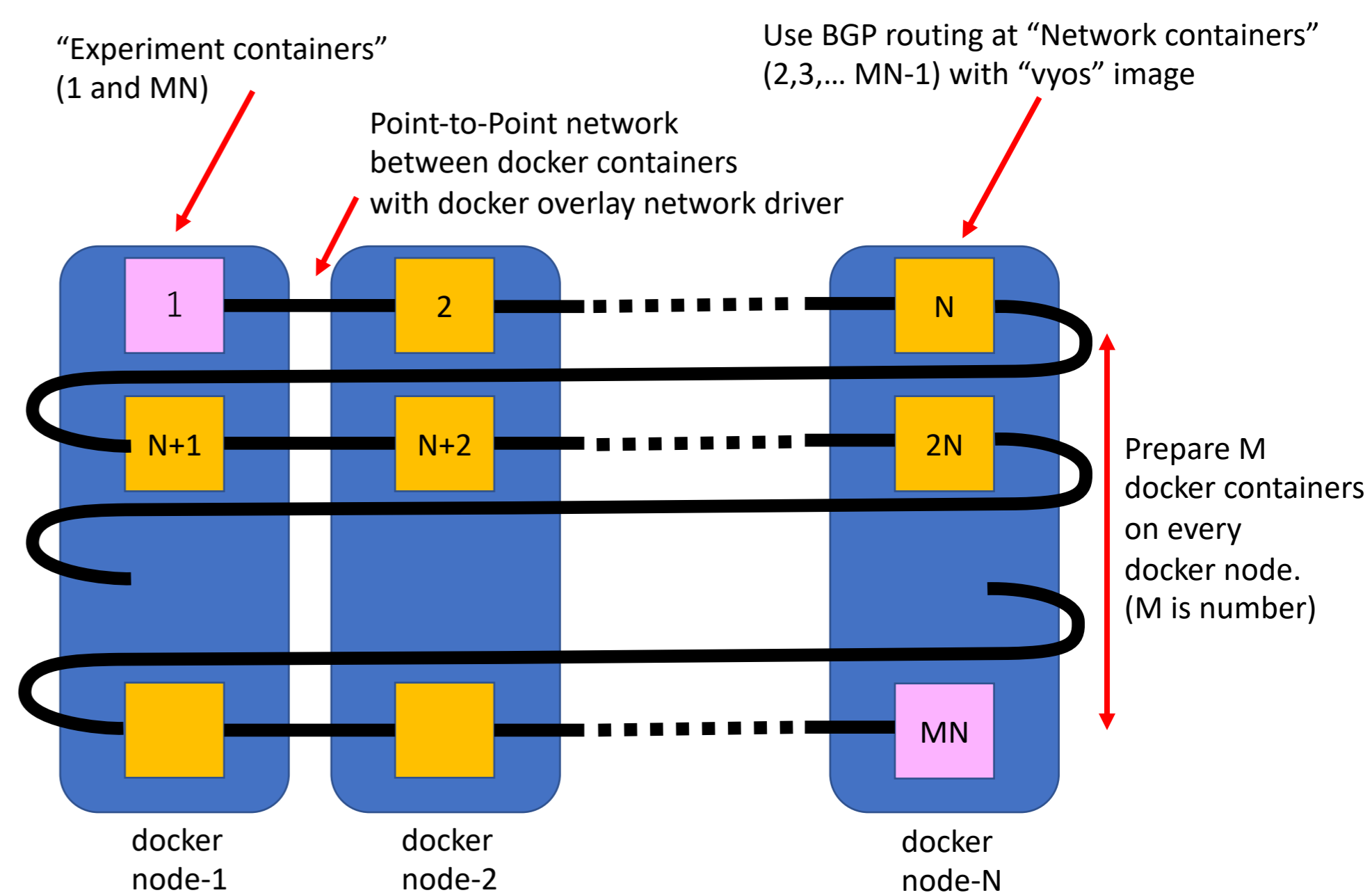


Fig 2. An example of network through multiple containers

```

yama@testbed1-yokosukai:~$ docker exec -it vyos-57 vtysh "-c sh ip bgp"
BGP table version is 0, local router ID is 0.0.0.57
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, B Stale, R Removed
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop         Metric LocPrf Weight Path
  * 10.199.34.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 i
  * 10.199.35.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 i
  * 10.199.36.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.37.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.38.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.39.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.40.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.41.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.42.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.43.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.44.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.45.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.46.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.47.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.48.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.49.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.50.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.51.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.52.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.53.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.54.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.55.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.56.0/24  10.199.56.2         0  56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 i
  * 10.199.57.0/24  0.0.0.0             1  32768 i
  * 10.199.57.0/24  0.0.0.0             1  32768 i

Total number of prefixes 24
yama@testbed1-yokosukai:~$
    
```

Fig 3. Routing information through 24 containers