



Open source SDN platform: Features, scalability and Deployment

Juniper Networks
Tatsuya Naganawa





Agenda

- Tungsten Fabric Introduction
- Key feature for scalability
- Recommendation for scalable deployments



Self introduction

- Currently working on Contrail product in Juniper Networks
- Support customers for IaaS, NFVI, for 5G deployment



Tungsten Fabric Introduction

- Originally implemented as a software vRouter, based on VPNV4 MP-BGP and MPLS over IP
- L3 VXLAN is also supported from v5.0.2 (L2 VXLAN is supported from v2.0)
- VXLAN becomes one of favorable deployment scenario, for integration with many more boxes than originally intended
 - no need to support MPLS



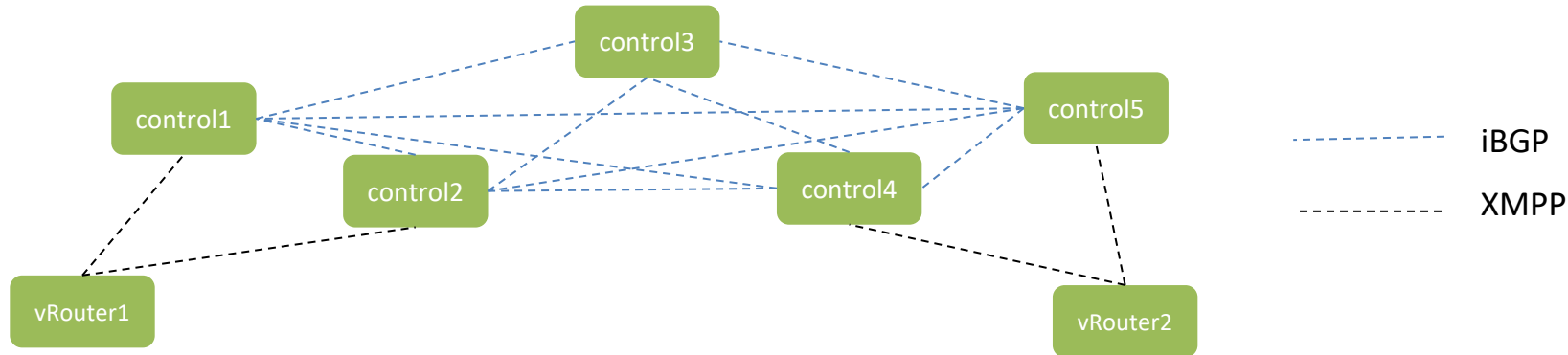
Key feature for scalability

1. Two XMPP connections and iBGP between controllers
2. Route Target Filtering
3. ERM-VPN



Two XMPP connections and iBGP between controllers

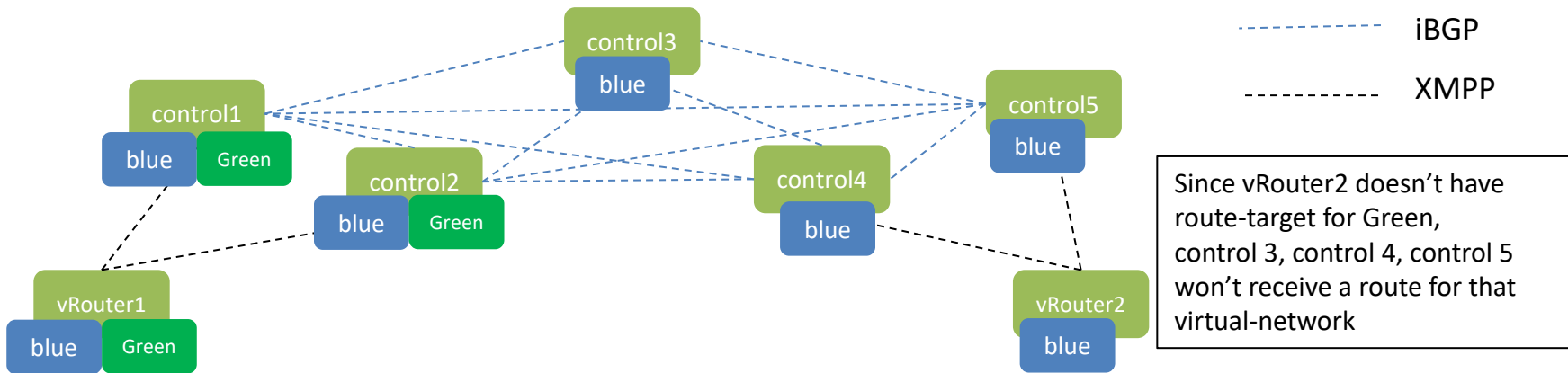
- vRouter will choose only two controllers to connect, although much more controllers can be used to add more capacity
- It can be used for horizontal scale out for its control plane
- All the routes between them can be exchanged by iBGP between them





Route Target Filtering

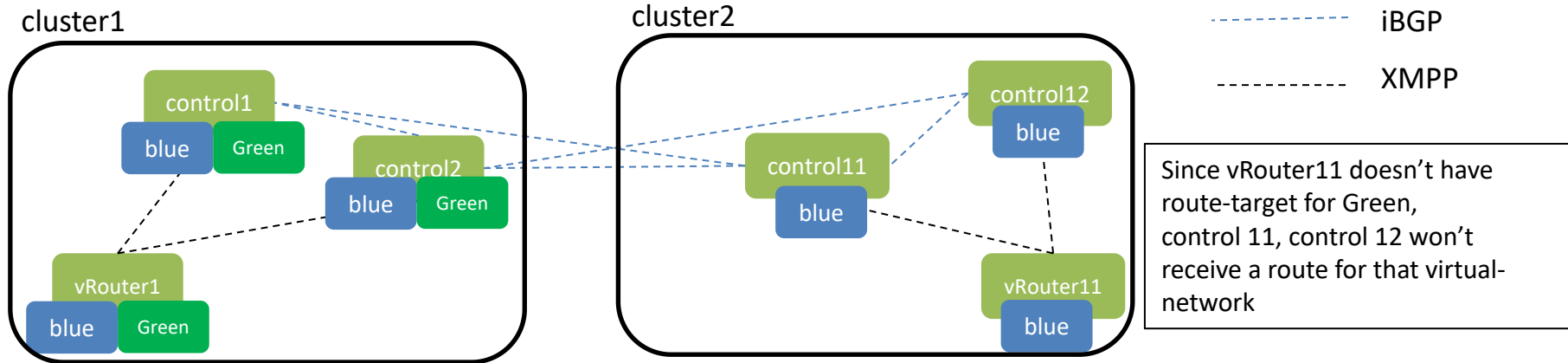
- When a controller receives routes from other controllers, **it won't receive that route, if none of vRouters which are directly connected needs that route**
- 'Route target' community is used to identify if each VPNV4 route is marked by specific virtual-network, and that prefix needs to be imported by control process and vRouters





Route Target Filtering

- Route Target Filtering also applies when each control and vRouter belongs different TF clusters
- So, when there are several clusters and they mainly serves different tenant, even if they have bgp connection between them, they don't need to import all the prefixes





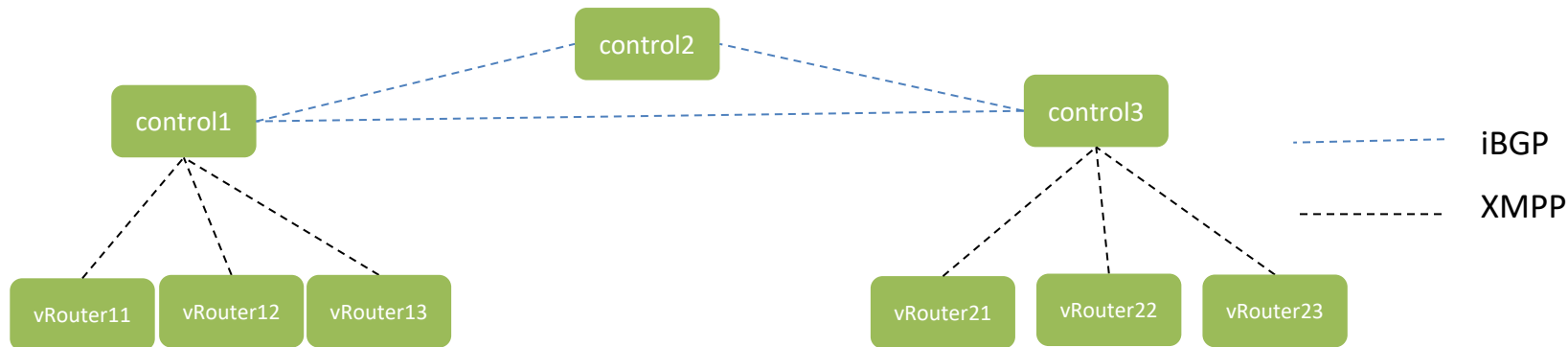
ERM-VPN

- To replicate BUM traffic from vRouter, **when number of vRouters are large, they replicate packets a lot of times** (even 1,000 times), and potentially could cause traffic spike
- To overcome this, control process firstly calculates one specific BUM trees based on the data of connected vRouters, and exchange that data between controllers, with ERM-VPN bgp family
- As a result, **one large set of BUM tree will be formed, which includes all vRouters**
- Because of that, each vRouter only needs to send up to other 5 vRouters, so even if large number of vRouters have the same virtual-network, it won't cause much traffic spike because of BUM.

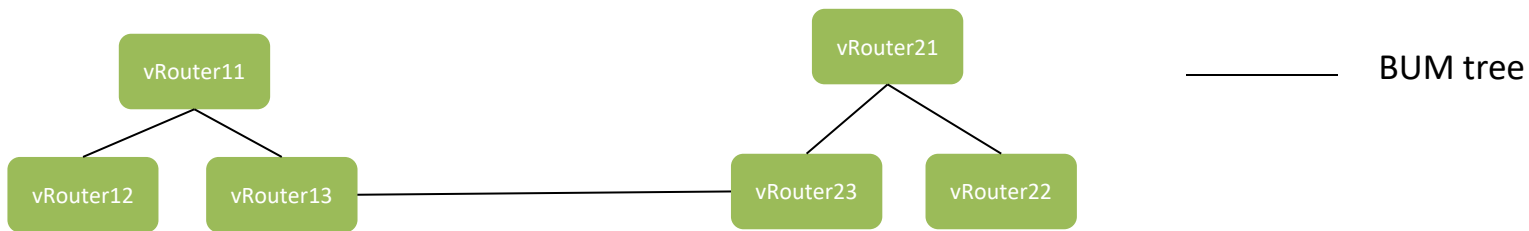


ERM-VPN

XMPP / iBGP setup



BUM tree formed





Recommendation for scalable deployment

When large cluster (1k vRouters with 15 control nodes, or 2k vRouters with 5 control nodes ...) needs to be implemented, those point need to be carefully considered before implementation.

1. control process can be horizontally scaled out
2. config process cannot be scaled out, scale up is the way to go
3. XMPP hold timer also is an important to balance faster failure detection and CPU power consumption, so this also needs to be carefully planned
 - XMPP_KEEPLIVE_SECONDS parameter can be used to tune this



Communication channel

- slack

<https://app.slack.com/client/T0DQ21YM9/C0DQ4JPCZ>



Reference

- Control plane Scalability

<https://github.com/tnaganawa/tungstenfabric-docs/blob/master/TungstenFabricPrimer.md#ii-scalability>

<https://github.com/tnaganawa/tungstenfabric-docs/blob/master/TungstenFabricPrimer.md#sizing-the-cluster>

<https://github.com/tnaganawa/tungstenfabric-docs/blob/master/TungstenFabricKnowledgeBase.md#vrouterscale-test-procedure>

- ERM-VPN

<https://github.com/tnaganawa/tungstenfabric-docs/blob/master/TungstenFabricKnowledgeBase.md#erm-vpn>

A large, glowing sphere composed of many small dots and lines, creating a sense of depth and light. The colors transition from purple on the left to pink on the right, with a bright white highlight at the top. The sphere is centered behind the main text.

4th THANKS FNDC