TATA CONSULTANCY SERVICES
Experience certainty.

Telecom

White Paper

Inter-SDN Controller Communication:

Using Border Gateway Protocol

About the Authors

Deepankar Gupta

Deepankar Gupta is an Associate Consultant and Lead Engineer working in the IP and Software Defined Networking domain with Tata Consultancy Services (TCS). Deepankar is responsible for ideation and identification of market trends. With over 14 years of experience in the IT industry, he has worked across a number of areas within telecom, such as wireless/wireline protocols and embedded technologies.

Rafat Jahan

Rafat Jahan is working as an R&D Analyst with TCS with a total experience of over 7 years in wireless technologies and cloud infrastructure. Rafat is involved in ideation and implementation of solutions based on next generation technologies, primarily on SDN. She is responsible for providing offerings as solution accelerators to help telecom customers with accelerated roll-outs of Software Defined Networking and Network Function Virtualization.

The widespread adoption of Software Defined Networking (SDN), -, has given rise to the need for more communication interfaces between SDN controllers. SDN is a new and increasingly popular network architecture that aims at decoupling the control and data planes. OpenFlow is a preferred protocol used for establishing such communication between the control and data planes.

Any SDN network abstraction comprises an application layer (management and service), a control plane, and a data plane. The SDN controller resides in the control plane, which is the middle and most important layer. It provides north-bound APIs to the user for applications and communicates with the data plane through the OpenFlow interface.

SDN east-west interface is one way to establish communication between multiple SDN controllers so as to share control plane parameters like Quality of Service (QoS), policy information, and so on. This white paper discusses the need for the Border Gateway Protocol (BGP) protocol and how it can be used for inter-SDN controller communication

Contents

List of Abbreviations

ALTO	Application Layer Traffic Optimization
AS	Autonomous Systems
BW	Bandwidth
BGP	Border Gateway Protocol
CDN	Content Delivery Networks
EGP	Exterior Gateway Protocol
IAAS	Infrastructure as a Service
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
ND	Network Domain
NLRI	Network Layer Reachability Information
NOS	Network Operating System
OFS	OpenFlow Switch
ONF	Open Networking Foundation
POP	Point of Presence
QoS	Quality of service
REST	Representational State Transfer
RFC	Request for Comments
RIB	Routing Information Base
SDN	Software Defined Networking
SIP	Session Initiation Protocol
SOHO	Small Office Home Office
SP	Service Provider
TCP	Transmission Control Protocol
VM	Virtual Machine
WAN	Wide Area Network

1. Introduction

SDN aims at providing networks an innovative approach in controlling data flow in the dataplane. SDN proposes a common, centralized control plane where the controller or NOS is responsible for maintaining an inventory of all network devices, packet switching, and allocating networking resources to applications running on top of the controller, as shown in Figure 1.

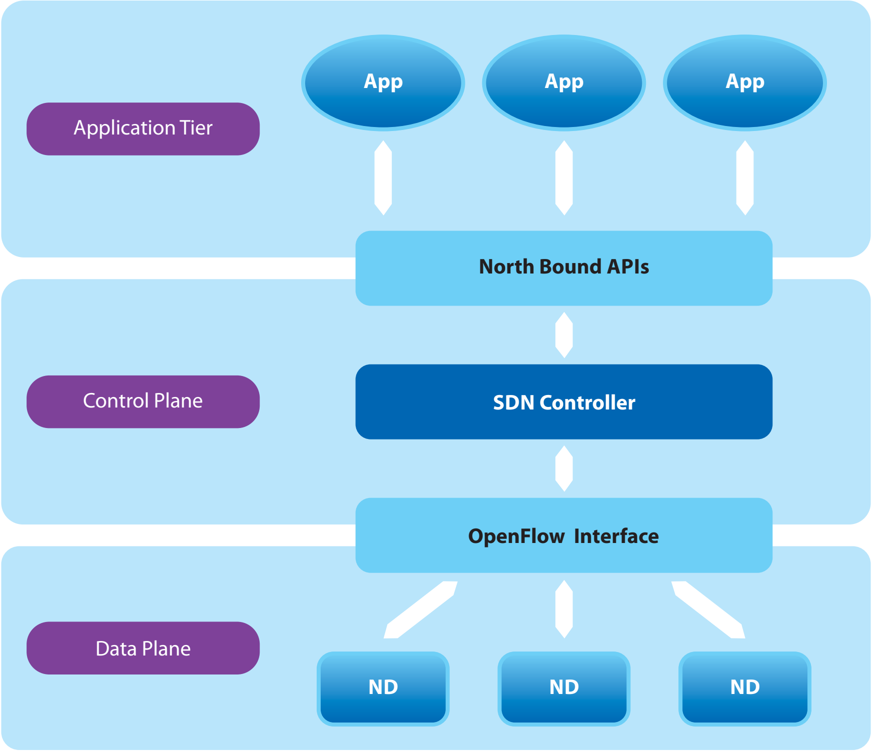

Figure 1: SDN (OpenFlow) Architecture

In SDN, the control plane is centralized. An SDN controller is able to control a small part of the whole network, termed as the SDN domain. Data centers, especially those in the cloud, are a logical example of an SDN domain. With the increasing use of SDN, exchange of information between multiple SDN domains will become an important need.

A network can have multiple SDN domains, each controlled by an individual SDN Controller. Interconnecting these controllers to share information and coordinate their decisions is important for routing information and providing quality of service. In the future, when SDN will be deployed in large-scale networks, operators of this large scale enterprise will want to divide the whole network into multiple connected SDN domains for better scalability and security.

In a legacy network, for ease of administration and control, a network is divided into several Autonomous Systems (AS), which have Interior Gateway Protocols (IGP). Networks communicate with the WAN using Exterior Gateway Protocols (EGP).

Today, networks are moving towards cloud architecture (data center) and SDN is becoming the key enabler of the same. To utilize data center resources efficiently, SDN controllers need to communicate. In a multi-SDN controller environment, each controller needs to be connected to the neighboring controller. This need is explored in more depth in the following section.

2. The Need for Inter-SDN Controller Communication

These use cases explain the need for inter- SDN controller communication.

2.1 Use Case: Bandwidth on Demand

In a network with a single SDN controller, a request for bandwidth change can be processed instantaneously. However, when the network resources are distributed among multiple SDN domains, controllers from each domain have to communicate with the SDN controller of the source domain to share information on parameters like QoS, bandwidth availability, and so on. This enables the SDN controller of the source domain to confirm and process the bandwidth requirement. Without communication with SDN controllers of various domains, the processing of requests like bandwidth-on-demand will not be possible.

2.2 Use Case: Content Delivery Networks

Service Providers (SPs) have to meet the content delivery requirements as per the committed QoS. SPs have several caches or replicas of the Content Delivery Network (CDN) server to meet these customer needs. In a scenario where the CDN server or cache nearest to the customer location is experiencing high loads and is unable to serve the customer, the request is sent to a CDN cache that is not loaded, but may be located in a different SDN domain in the network. Here, the source SDN controller will need to communicate with the other SDN controllers within the network to negotiate a path to the best possible CDN server that can meet the customer's QoS needs.

Here is an analogy to explain this: a football game broadcast in the USA is being viewed by a large number of fans. As more fans turn on their IP TVs or come online to view the game, the load on the CDN server increases, and the performance of the server decreases. In this situation, it is better to have customers get data from a CDN server located elsewhere, say in Europe, which has underused capacity.

Figure 2 depicts an SDN-based content delivery network[1].

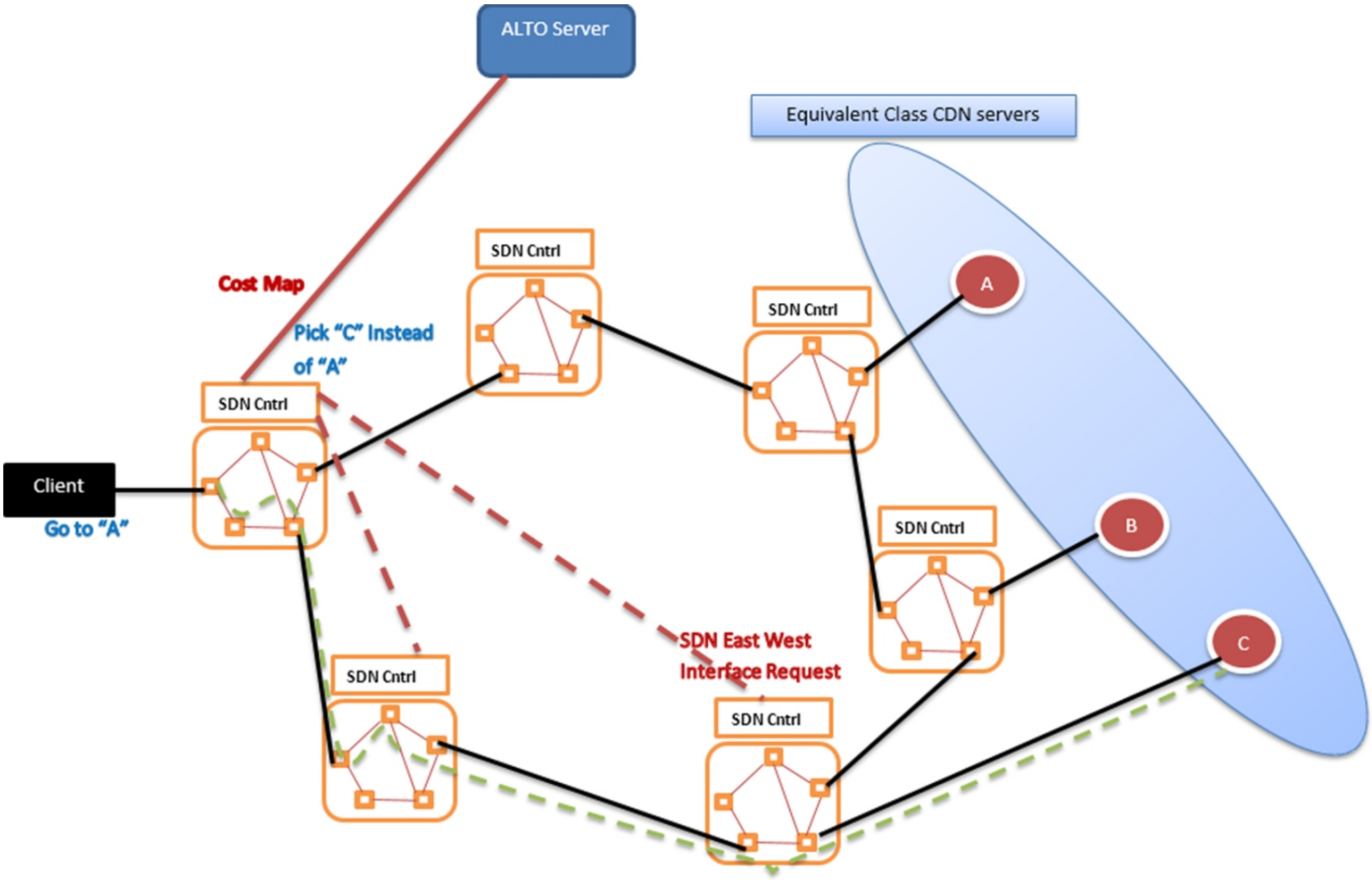

Figure 2: Content Delivery Network with SDN based Network

2.3 Use Case: Separate SDN Controllers for Various Network Functions

As the end-to-end wireless networks are not fully compliant with protocols like OpenFlow, it is difficult to have a single SDN controller for network components at access, edge, core, and data center. A common solution is to federate several SDN controllers, each controlling one set of individual network functions, as shown in . The long term goal is to have one single protocol handle both configuration and traffic engineering.

If the customer (retail, enterprise) demands a service that is provided by the data center, the SDN controllers of access, edge, core networks, and the data center need to communicate with each other. SDN controllers need to pass on QoS and other parameters like policy information from the access network to the core network to the data center. This is illustrated in Figure 3.

[1] Xie Haiyong, Tsou Tina, Lopez Diego, Yin Hongtao and Gurbani Vijay, "Use Cases for ALTO with Software Defined Networks", IETF Internet-Draft, Vancouver, Canada. July 2012, available at http://tools.ietf.org/id/draft-xie-alto-sdn-extension-use-cases-01.txt; Yin Hongtao, Xie Haiyong, Tsou Tina, Lopez Diego, Aranda P. and Sidi Ron, "SDNi: A Message Exchange Protocol for Software Defined/Driven Networks across Multiple Domains", IETF Internet-Draft, Vancouver, Canada. July 2012, available at http://tools.ietf.org/html/draft-yin-sdn-sdni-00; Xie Haiyong, Tsou Tina, Lopez Diego, Sidi Ron, Yin Hongtao, and Aranda Pedro Andres, "Software Defined Networking Debuted at IETF", the IETF Journal, October, 2012

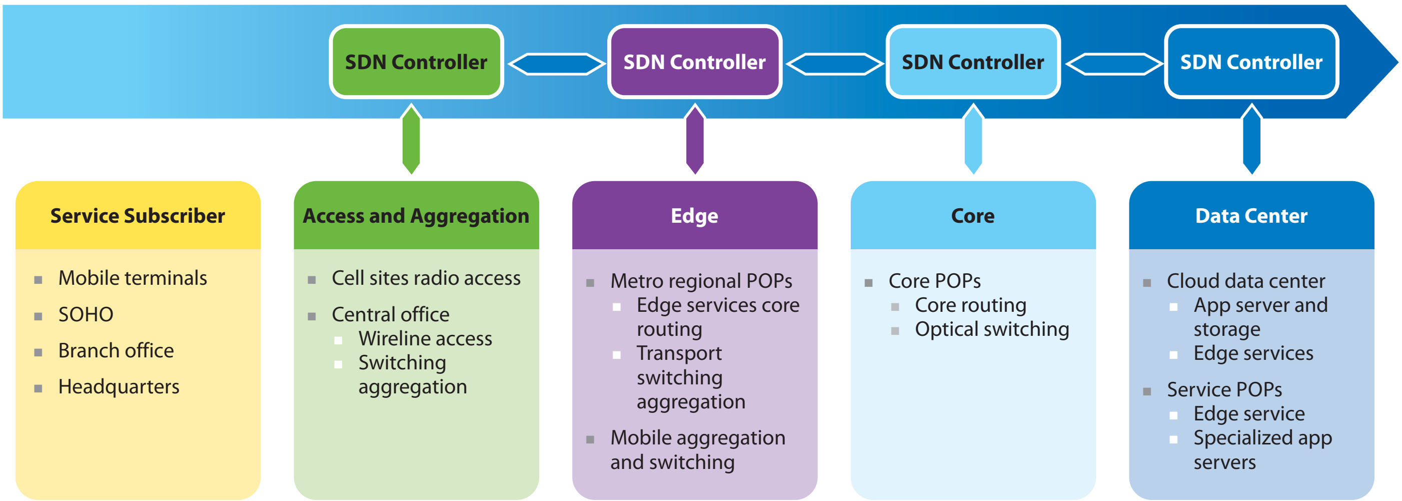

Figure 3: Federation of SDN Controllers

3. Approaches to Inter-SDN Controller Communication

Inter–SDN controller communication can be implemented using the vertical or the horizontal approach.

3.1 Vertical Approach

In the vertical approach, there is a master controller over the individual network controllers, as shown in . The master controller has a global view of the network across all connected SDN domains and can orchestrate the configuration in each domain.

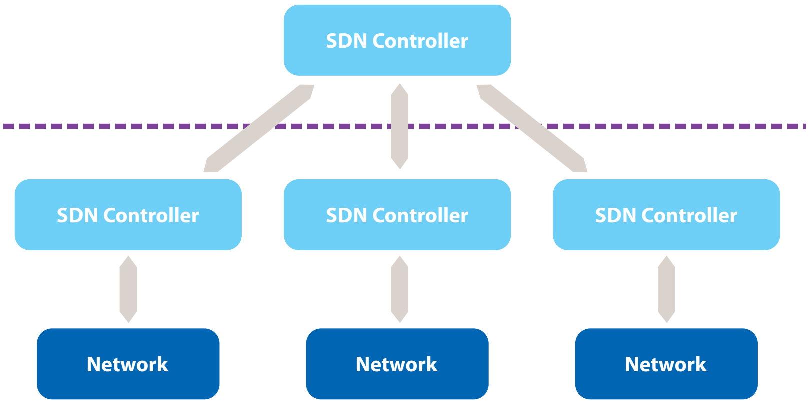

Figure 4: Vertical Approach to Inter-SDN Controller Communication

3.2 Horizontal Approach

In the horizontal approach, the SDN controllers establish peer-to-peer communication, as shown in Figure 5. Each controller can request for information or connections from its peers, that is, SDN controllers from other domains in the network. This is also called the SDN east-west interface.

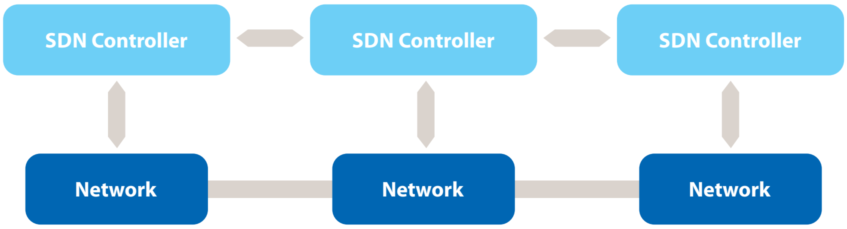

Figure 5: Horizontal Approach to Inter-SDN Controller Communication

3.3 Comparative Results

In the vertical approach, the controller of controllers, also known as the master controller, interfaces with the underlying controllers using RESTFul APIs. The master controller is responsible for sharing information across the underlying controllers based on individual SDN domain policies. Though this approach can work for data centers, problems may arise if this model is to be scaled up in an enterprise spread across different locations and geographies.

The horizontal approach seems more feasible for scaling across geographies, as controllers in this model can communicate with each other using standard protocol. This approach keeps the network SDN controller independent, with distinct policies and path setup to apply to network elements in its control, while maintaining a federation with the neighbouring networks.

We will discuss the horizontal approach between SDN controllers in greater detail.

3.4 A Closer Look at the Horizontal Approach

The basic SDN east-west interface functioning is shown in Figure 6.

In the east-west SDN interface, each controller is responsible for a specific domain in the network.

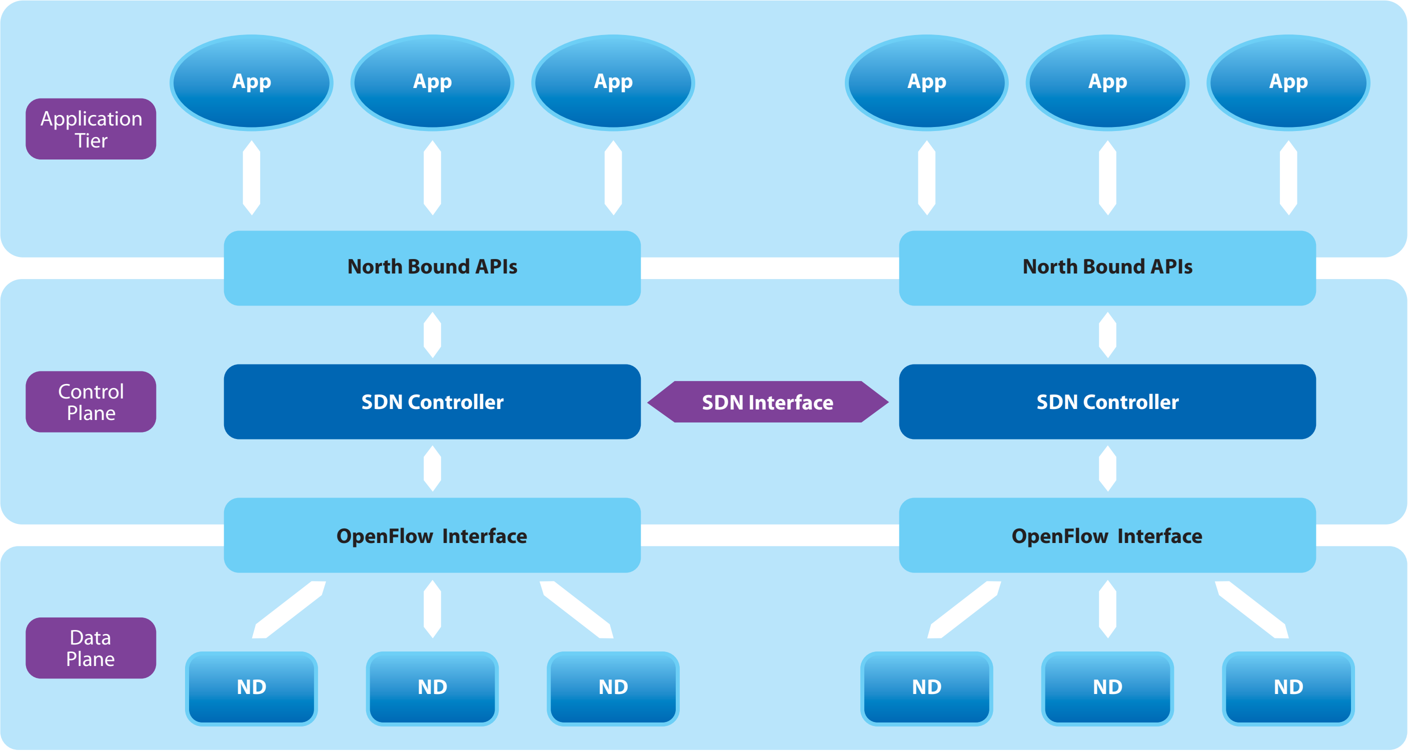

Figure 6: SDN East-West Interface Concept

3.5 Expectations from SDN East-West Interface

Path setup across SDN domains depends on the capability and resource availability (including factors such as Bandwidth, QoS, CDN server response time, and so on) within each domain. The SDN east-west interface should be able to coordinate the path setup requirement across multiple SDN domains, so it should be implemented in such a way as to incorporate the capabilities of the different controllers.

Controllers need to exchange information such as:

- **Reachability update:** Exchange of reachability information facilitates inter-SDN domain routing. This allows a single flow to traverse multiple SDNs and each controller can select the most appropriate path in the network.
- **Flow setup, tear-down, and update requests:** Controllers coordinate flow setup requests, which contain information such as path requirements, QoS, and so on, across multiple SDN domains.
- **Capability Update:** Controllers exchange information on network-related capabilities such as bandwidth, QoS and so on, as well as system and software capabilities available inside the domain.

4. Implementation Options

SDN east-west interface is used to exchange information between SDN domains that are under the control of single or multiple network operators. A session needs to be established between the two controllers by using either BGP[2] or Session Initiation Protocol (SIP) over Transmission Control Protocol (TCP) to exchange information.

SIP is a request-response protocol for initiating and managing communication. The protocol defines the messages that are sent between peers, which govern establishment, termination, and other essential elements of a call. SIP maintains a session but it is mostly used in establishing multimedia sessions.

BGP is an inter-autonomous system routing protocol. It maintains a session between routers and follows a state machine approach.

A session can be established by either of these application protocols, but BGP2 has significant advantages, as shown in Table 1.

4.1 Recommended Approach

BGP is a routing protocol required to share routing information between two autonomous systems. This makes BGP an ideal protocol as it can be easily adapted for inter-SDN controller communication, where various SDN domains will have access and control of QoS, policies, and other parameters through the SDN controllers.

BGP has the following features that are needed for SDN east-west interface:

- BGP messages can carry capability and reachability information as part of their message format.
- BGP is a standard and the most feasible protocol for any peer data to be exchanged.

	BGP	SIP
General Approach	Functional State Machine approach makes it easier to implement a BGP session.	It is based on request-response policy; SIP is an application-layer control protocol that can establish, modify, and terminate multimedia sessions.
	BGP can be used independently as an application layer protocol	SIP should be used in conjunction with other protocols, such as Session Description Protocol, Real Time Transport Protocol, etc., to provide complete services to users.
	The straightforward message format is in accordance with the OpenFlow messages.	Since messages are mostly text-based, retrieving data can be difficult when integrated with SDN Controller
Capability Information	The OPEN message has a capability information field, making it simpler to exchange QoS data. IETF RFC 3392 explains this approach in detail.	Capability exchange procedures to ensure that multimedia signals that are transmitted, can be received and processed appropriately by the receiving terminal. They do not serve the data that is expected .
Reachability Information	The UPDATE message in BGP maintains the reachability Information for a session.	Reachability information support is not available, though user availability information shares the willingness of the called party to engage in communications.
Flow set up and tear down	NLRI and Path Attribute field of UPDATE Message maintains the flow set up and tear down.	SIP is used for the call set up and tear down functionality.

Table 1: Comparison Between BGP and SIP

[2] *IETF, A Border Gateway Protocol 4 (BGP-4), January 2006, accessed January 2014, http://www6.ietf.org/rfc/rfc4271*

- BGP meets the conditions of existing legacy-based solutions such as standard eBGP.

Figure 7 depicts how reachability, capability, and flow setup information can be shared among multiple controllers.

4.2 Steps in Connection Establishment

Connection establishment typically takes place in this manner.

1. SDN Controllers need to be enabled as BGP speakers, each BGP speaker having state-machine logic. Once the controllers are up, the BGP functionality for each controller gets triggered by an event generated through the controller, which can be a BGP_START event.
2. For each SDN controller, information about the neighbor (another BGP speaker) will be configured manually by the network administrator. The SDN Controller tries to establish a TCP connection with its neighbour. If the

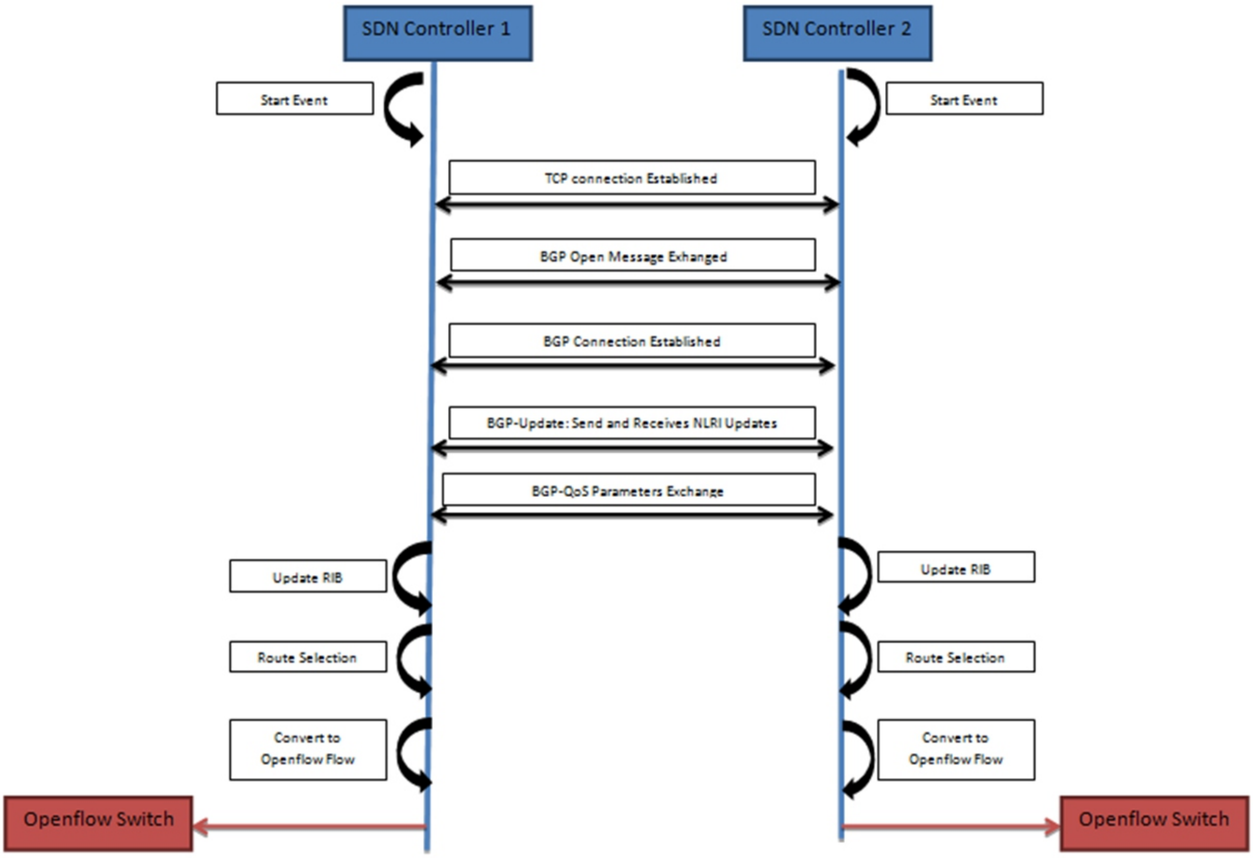

Figure 7: SDN East-West Connection Establishment, Route, and Flow Setup

connection is not established due to reasons such as TCP timeout or BGP message timer expiry, the BGP speaker has to establish connection with another neighbor, which has to be configured by the administrator.

3. BGP will use TCP as its Transport Layer Protocol. Once the TCP connection is established, OPEN message is sent by both the BGP speakers to each other and they move to OPEN state.
4. During the OPEN state, BGP speakers can negotiate capabilities of the session through OPEN messages (as per RFC 5492). This capability information about the SDN domain is obtained by the controller using OpenFlow protocol.
5. Once the BGP peers are in a session, they move to ESTABLISHED state. The BGP UPDATE messages are exchanged in this state. Reachability data is available in the NLRI parameter of this message. Reachability information helps in the selection of the most appropriate data path between SDN controllers. Information obtained through NLRI parameter is used to update routing information base. In the SDN controller, exchange of reachability information is required to facilitate inter-SDN routing.
6. Bandwidth information is propagated through BGP UPDATE messages between the SDN controllers. The EXTENDED COMMUNITY[3] attribute of this message helps in exchanging the bandwidth among the BGP speakers. The bandwidth of a link is carried in the Link Bandwidth Community of the attribute.
7. QoS Marking attribute is also encoded as a BGP EXTENDED COMMUNITY attribute, which can be carried in the BGP UPDATE message.
8. As the two controllers are BGP speakers, the reachability data is maintained in the RIB table of each controller. This information is converted into flows in OpenFlow4 switches.
9. Route selection is done when more than one path is available based on BGP process decision. Once the path is established, packets can traverse successfully between two SDN domains and data exchange can take place successfully through the BGP OPEN and UPDATE messages.

As depicted in figure 7, communication can be established between multiple controllers in the same manner.

5. Conclusion: Why SDN is the Future of Cloud Computing

Why SDN is the Future of Cloud Computing

As businesses rapidly adapt to cloud-based IAAS, there will be an increase in the number of switches, both real and virtual, as well as the number of Virtual Machines (VMs) or hosts that need to be supported. However, there is a limitation to the number of switches and hosts that an SDN controller can manage (a BigSwitch network controller can support 1,000 switches and 250,000 hosts).

[3] *Traina Paul, Chandrasekran Ravishanker, 'BGP Communities Attribute' available at as IETF RFC at http://tools.ietf.org/html/rfc1997; Th. Knoll, 'BGP Extended Community Attribute for QoS Marking' July 11, 2013 , IETF Draft available at http://tools.ietf.org/html/draft-knoll-idr-qos-attribute-12; Ramachandra Srihari, Tappan Daniel, Rekhter Yakov, 'BGP Extended Communities Attribute' , December 2001 IETF Draft http://tools.ietf.org/id/draft-ramachandra-bgp-ext-communities-09.txt*

[4] *Open Networking Foundation, Openflow Switch specification openflow-spec-v1.3.2, April 2013/, accessed January 2014 , https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.2.pdf*

For the cloud to scale further and be efficient, it will need to divide the network among separate SDN controllers. This will make inter-SDN controller communication much more critical. Besides interconnecting multiple controllers, an SDN east-west interface will also allow them to share control plane network information, resulting in a coordinated decision-making process.

Though SDN east-west interface is yet to be standardized, the need for the same seems imminent. The Open Networking Foundation (ONF) is working on standardizing southbound interface from controller to switch. We expect that the ONF will also look into standardization of the east-west interface. Using the BGP protocol and making the SDN controller its speaker will lead to innovation in application development.

There are various design considerations for the SDN east-west interface, such as connection between network topologies and control plane parameters that should be shared between SDN controllers. BGP over TCP seems to be the most open and flexible idea to implement this east-west interface. By implementing the horizontal approach, your organization gets the flexibility to incorporate any impending enhancements and scale up data exchange as required, making your network future-proof.

TATA CONSULTANCY SERVICES
Experience certainty.

About TCS Telecom Business Unit

TCS' Telecom Industry Service Unit is one of the largest verticals within TCS, contributing strongly to TCS' revenues. With a dedicated pool of professionals and an accumulated experience and ongoing associations with world-class telecom service providers and equipment manufacturers, TCS has acquired unique and holistic understanding of the telecom domain to offer services suited to every stage of the business life cycle of our customers.

TCS helps wireline, wireless, broadband, and cable service providers redefine their markets with innovative solutions that help them become more agile, reduce fixed operations costs, improve profit margins, and introduce next generation services. TCS sets customers apart from their competitors with instant access to industry solutions, best-in-breed technology, assets and frameworks.

Contact

For more information about TCS' consulting services, contact **global.telecom@tcs.com**

Subscribe to TCS White Papers

TCS.com RSS: http://www.tcs.com/rss_feeds/Pages/feed.aspx?f=w
Feedburner: http://feeds2.feedburner.com/tcswhitepapers

About Tata Consultancy Services (TCS)

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model^{TM}, recognized as the benchmark of excellence in software development. A part of the Tata Group, India's largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at **www.tcs.com**

IT Services
Business Solutions
Consulting