### Introduction

In the cloud computing era, distributed architecture is used to handle operations of mass data, such as the storage, mining, querying, and searching for data in a data center. Distributed architecture leads to a huge amount of collaboration between servers, as well as a high volume of east-to-west traffic. In addition, virtualization technologies are widely used in data centers, greatly increasing the computational workload of each physical server.

These data center service requirements drive the evolution of data center network architecture and lead to the need for a large Layer 2 network architecture. To build large Layer 2 networks, the Transparent Interconnection of Lots of Links (TRILL) is introduced.

The following analyzes the network architecture requirements of data centers in the cloud computing era and introduces the Huawei TRILL-based solution, which can help build data center networks that meet cloud computing service requirements.

Network Architecture Requirements of Data Centers in the Cloud Computing Era
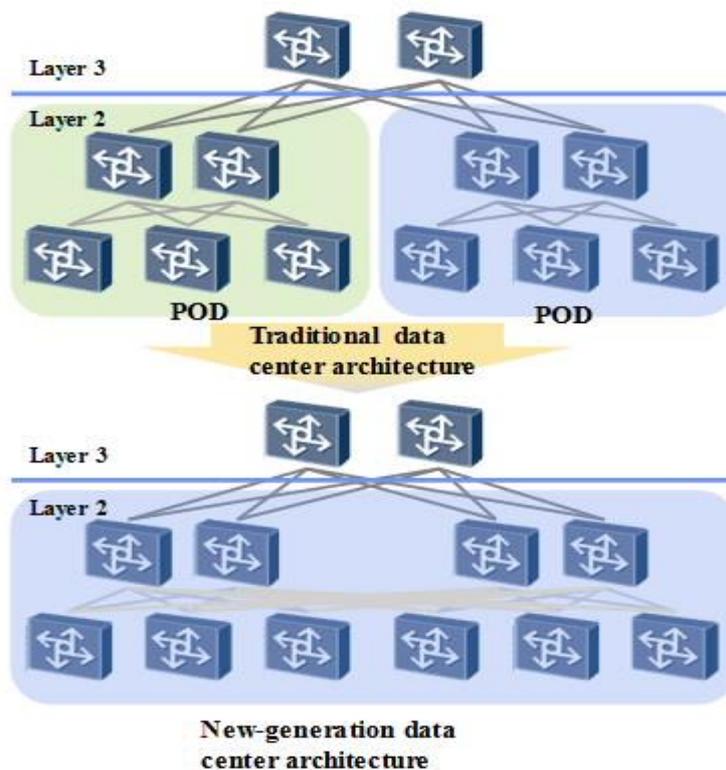
Smooth Virtual Machines (VM) migration

Server virtualization is a commonly used core cloud computing technology. To maximize service reliability, reduce IT and O&M costs, and increase service deployment flexibility in a data center, VMs must be able to dynamically migrate within the data center, instead of being restricted to an aggregation or access switch.

A traditional data center often uses Layer 2+Layer 3 network architecture, in which Layer 2 networking is used within a power of delivery (POD) and Layer 3 networking is used between PODs. In such network architecture, VMs can only be migrated within a POD. If VMs are migrated across a Layer 2 area, their IP addresses have to be reassigned. Applications will be interrupted if the load balancer configuration remains unchanged under this situation..

Non-blocking, low-latency data forwarding

The data center traffic model in the cloud computing era is different from the traditional carrier traffic model. For example, most traffic in a data center is east-to-west between servers, and the data center network acts as the bus between servers.

In traditional Layer 2 networking, xSTP is required to help eliminate loops and prevent Layer 2 broadcast storms, allowing only one of the N links to be used to forward data. This method reduces bandwidth efficiency and fails to meet data center network requirements in the cloud computing era. To ensure that services are running properly, non-blocking and low-latency data forwarding is required in the cloud computing era. Flat and fat-tree topology must be supported to make optimal use of multiple data forwarding paths between switches.

Multi-tenancy

In the cloud computing era, a physical data center is shared by multiple tenants instead of being exclusively used by one tenant. In this situation, each tenant has its own virtual data center so that they have exclusive servers, storage resources, and network resources. To ensure security, data traffic between tenants needs to be isolated. Currently, in traditional Layer 2 networking, the number of supported tenants is limited by the number of VLANs – a maximum of 4K. As cloud computing technology develops, the number of tenants supported by the future data center network architecture must be able to exceed 4K.

Large-scale network

In the cloud computing era, a large data center must support potentially millions of servers. To implement non-blocking data forwarding, several hundreds or thousands of switches are required for a large data center. In such large-scale networking, networking protocols must be able to prevent loops. Fast network convergence is necessary to enable services to recover rapidly when a fault occurs on a network node or link.

TRILL Characteristics

TRILL is defined by the IETF to implement Layer 2 routing by extending IS-IS. TRILL has the following characteristics:

High-efficiency forwarding

On a TRILL network, each device acts as the source node to calculate the shortest path to all other nodes through the shortest path tree (SPT) algorithm. If multiple equal-cost links are available, load balancing can be implemented when unicast routing entries are generated. In data center fat-tree networking, forwarding data along multiple paths can maximize network bandwidth efficiency. A traditional Layer 2 network using xSTP to prevent loops is like a single-lane highway, whereas a TRILL network is analogous to a multi-lane highway.

On a TRILL network, data packets can be forwarded using equal cost multipath (ECMP) or along the shortest path. Therefore, using TRILL networking can greatly improve the data forwarding efficiency of data centers and increase the throughput of a data center network.

Loop prevention

TRILL can automatically select the root of a distribution tree, allowing each router bridge (RB) node to use the root as the source node to calculate the shortest paths to all other RBs. In this manner, the multicast distribution tree to be shared by the entire network can be automatically built. This distribution tree connects all nodes on the network to each other and prevents loops when any Layer 2 unknown unicast, multicast, or broadcast data packets are transmitted on the network.

The route convergence time of nodes may be different when there is a change to the network topology. Users can use reverse path forwarding (RPF) checks to discard the data packets received from an incorrect interface so as to prevent loops. In addition, the Hop-Count field, carried on a TRILL header, can reduce the impact of temporary loops and effectively prevent broadcast storms.

Fast convergence

On a traditional Layer 2 network, the TTL field is not carried by an Ethernet header and the xSTP convergence mechanism is also inadequate. As a result, network convergence is slow and may last tens of seconds when network topology changes, failing to provide high service reliability for data centers. TRILL uses routing protocols to generate forwarding entries, and a TRILL header carries the Hop-Count field to allow temporary loops. These characteristics provide fast network convergence when there is a fault on a network node or link.

Convenient deployment

The TRILL protocol is easy to configure because many configuration parameters, such as the nickname and system ID, can be automatically generated and can retain default settings. On a TRILL network, the TRILL protocol manages unicast and multicast services in a unified manner. This is different from a Layer 3 network, where multiple sets of protocols manage unicast and multicast services separately. Moreover, a TRILL network is still a Layer 2 network, which has the same characteristics as a traditional Layer 2 network: plug-and-play, and ease of use.

Easy support for multi-tenancy

Currently, TRILL uses the VLAN ID as the tenant ID and isolates tenant traffic using VLANs. During the initial phase of cloud computing technology development and large Layer 2 network operation, a maximum of 4K VLAN IDs will not cause a bottleneck. But, as the cloud computing industry begins to mature, TRILL must break through the 4K tenant IDs limitation. To address this issue, TRILL will use FineLable to identify tenants. A FineLable is 24 bits long and can support a maximum of 16M tenants, which can meet tenant scalability requirements.

Smooth migration

Networks using the traditional Layer 2 Ethernet technology MSTP can seamlessly connect to a TRILL network. Servers connected to the MSTP networks can communicate with those connected to a TRILL network at Layer 2, and VMs can be migrated within a TRILL network.

These characteristics have enabled TRILL-based network architecture to better meet data center service requirements in the cloud computing era.

Huawei TRILL-based Large Layer 2 Network Solution
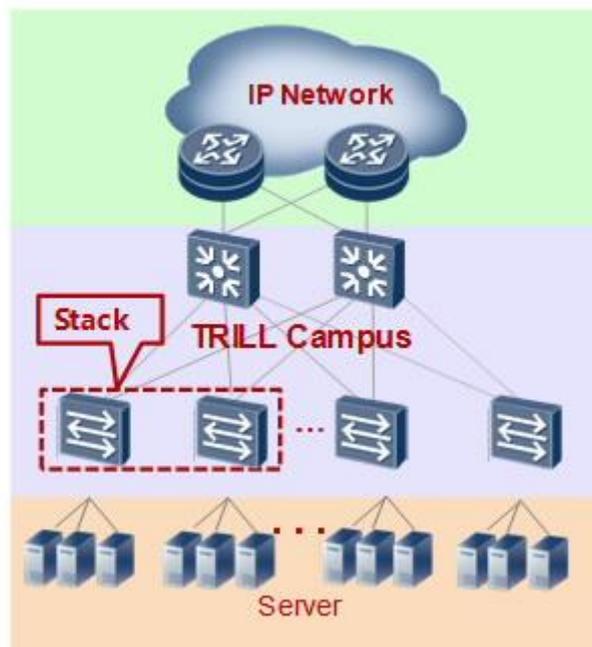
The preceding figure shows a typical Layer 2 fat-tree networking, in which all access switches and core switches run the TRILL protocol. Access switches are either top-of-rack (TOR) or end-of-row (EOR) switches that support stacking. Gateways can be deployed independently or together with router bridges (RBs).

The TRILL network functions as the bus of a data center. To ensure that services are running properly in a data center, the TRILL network must efficiently forward traffic between servers, and between servers and Internet users.

Huawei TRILL Solution Characteristics and Advantages

TRILL is the highway of a data center network and has the following characteristics:

Flexible networking that supports TOR networking and EOR networking

Data center switches all support TRILL protocol, TOR networking, and EOR networking, while all boards support TRILL forwarding. This feature ensures networking flexibility. To increase reliability, access switches also support stacking.

The TRILL protocol can be deployed from core switches to edge access switches of a data center network. In a data center, the TRILL network functions as a bridging fabric to provide stable core network architecture. As cloud computing technology continues to develop, users can easily add physical servers, IPv4/IPv6 gateways, firewalls, and load balancing devices to data center networks.

Flexible gateway deployment

TRILL supports two gateway deployment modes:

- Independent deployment of Layer 3 gateways: Layer 3 gateways and core RBs are directly connected. In a large data center, multiple gateways can be deployed to implement VLAN-based load balancing.
- Deployment of Layer 3 gateways on core RBs: Virtual switch (VS) technology virtualizes a core RB into two VSs, with one VS running the Layer 3 gateway function, while the other runs the TRILL protocol.

Gateways need to forward data traffic between Internet users and virtual servers in a data center and inter-segment traffic within a data center. The gateway deployment mode directly affects the forwarding performance of south-to-north traffic and inter-segment east-to-west traffic of a data center.

Deploying Layer 3 gateways on core RBs applies to small- and medium-scale data centers, while independent deployment of Layer 3 gateways applies to large-scale data centers. Users can select the gateway deployment mode according to service requirements.

Supports various operation, maintenance, and management methods
TRILL supports a variety of operation, maintenance, and management methods. CLI, SNMP, and Netconf administrators can log in to the VLANIF interfaces of RBs through a TRILL network. The management network can share the same physical network with a TRILL service network. To detect connectivity within a TRILL network, the TRILL ping can be used for fault location.

More multicast distribution trees

Link load balancing can be performed on TRILL unicast traffic using the ECMP algorithm. The ingress RB of multicast traffic can select different multicast distribution trees for VLANs to implement link load balancing for multicast traffic. A TRILL network supports a maximum of four multicast distribution trees, facilitating the multicast traffic processing of different tree roots and allowing for fine-grained load balancing on multicast traffic of the entire TRILL network.

Large network scale and optimal performance

A TRILL network supports more than 500 nodes, provides less than 500 ms network convergence in the case of a node or link fault, and supports load balancing among a maximum of 16 links. These advantages can meet the network scale and performance requirements of large data centers.

Two data center evolution modes

TRILL supports two data center evolution modes:

- Mode 1: If network expansion devices supporting TRILL are added to a data center using the traditional Layer 2 Ethernet technology MSTP, the traditional MSTP network can seamlessly connect to a TRILL network to form a large Layer 2 network. This mode maximizes the return on investment.

- Mode 2: The type of network that the C-VLAN servers connect to can be specified to implement smooth data center evolution. In new data centers, devices support TRILL. TRILL can be initially deployed only for specific services, while MSTP is deployed on other services. This allows users to test the operation of TRILL, and to gain experience before migrating all services to a TRILL network.

Summary

A TRILL network functions as the high-speed bus in a data center and has many advantages including smooth virtual machines migration, non-blocking and low-latency data forwarding, multi-tenancy, and large network scale. Therefore, a TRILL network can meet data center service requirements in the cloud computing era. Huawei works together with its partners to promote the application, development, and evolution of TRILL technology in data centers.