

PBBN in the Data Center

Bob Sultan (bsultan@huawei.com)

Abstract—Provider Backbone Bridged Networking (PBBN) may be deployed in the data center to reduce the number of MAC addresses visible to core bridges. Extensions are proposed to allow PBBN to better meet the specific requirements of the data center.

Index Terms—Provider Backbone Bridged Network; data center.

I. INTRODUCTION

A recent article in the New York Times magazine¹ describes data centers operated by Microsoft, Google, Yahoo, Facebook, and Savvis (supporting financial markets). Microsoft's newest data center, under construction in the Chicago area, is described as occupying an area of 700,000 sq. ft. (65,000 sq. meters) and housing tens-of-thousands of servers. A Microsoft spokesperson suggests that a fully containerized data center design might further increase server density by an order of magnitude in the near future. In such a network, it is not hard to imagine that the number of MAC addresses visible to core bridges could exceed the capacity of even a large Filtering Database (FDB). One solution is the deployment of IEEE 802.1ah Provider Backbone Bridged Networking (PBBN) to reduce the number of MAC addresses visible in the network core. This paper describes extensions to PBBN that are useful when PBBN is deployed in the data center.

II. THE DATA CENTER NETWORK

A data center network supports connectivity among attached devices. The point of attachment of a device is a *NIC*. Each NIC is associated with an individual MAC address. The devices are generally interconnected by a hierarchy of *bridges*, as illustrated in Figure 1. Bridges deployed in the data center are known as *customer bridges* and are said to be *C-VLAN-aware*². Traffic exchanged between devices³ associated with different C-VLANs⁴ must

transit a *router*. For purposes of this discussion any device attached at the edge of the bridged network, and which is not a router, is called a *server*⁵. The hierarchical structure is often augmented with mesh interconnection in the central portion of the network to provide redundancy and load distribution. Load distribution may be achieved by deploying the Multiple Spanning Tree Protocol (MSTP) and mapping the C-VLANs to diverse trees. In the future, such distribution may be realized using equal-cost shortest paths provided by IEEE 802.1aq Shortest Path Bridging (SPB).

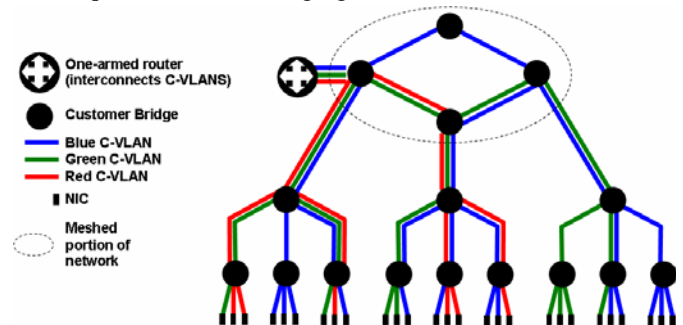


Figure 1: A data center network

The physical realization of the data center network can vary. In some cases, multiple levels of hierarchy are contained within a single physical chassis; for example, in a *blade server*. In this case, a multiplicity of *blades* can be inserted into slots of the blade server chassis. The blades are interconnected by a bridging device that is contained within the chassis. Further, each blade comprises multiple processors, each of which functions independently as a server. The multiple servers with the blade are interconnected by a bridging device integrated into that blade. Thus, the blade server contains two levels of the data center hierarchy within its chassis. Such a device may contain large numbers of individual servers and associated NICs.

In general, devices within, or closer to, the network core are stand-alone bridges⁶ with sufficient capacity to accommodate the increased load present in the central portion of the network.

Increasingly, data centers are deploying virtualization technology. In this environment, a single physical server contains multiple logical servers, each of which is known as a *virtual machine*. While such logical servers do not differ

¹ <http://www.nytimes.com/2009/06/14/magazine/14search-t.html?pagewanted=1&ref=todayspaper>

² Frames associated with a customer VLAN (C-VLAN) carry the TPID value 0x8100 in contrast to frames associated with a service VLAN (S-VLAN) that carry the TPID value 0x88A8. Some data centers may deploy both C-VLANs and S-VLANs in a hierarchical fashion. Some data centers may deploy S-VLANs *instead of* C-VLANs. For purposes of discussion, it is assumed that such data centers have essentially adopted the service provider model. In this case, the PBBN extensions described in this paper would not be applicable.

³ Or from one device to other devices in the case of broadcast/multicast.

⁴ The VLAN represents an IP Subnetwork from a Layer 3 perspective.

⁵ Such a server device may perform client functions with respect to other servers. In all cases, we refer to the device as a server. Servers may provide processing, storage, and other types of services.

⁶ In some cases with integrated routing function.

functionally from their physical counterparts, the trend further increases the number of MAC addressable NICs whose attachment must be supported. This increases the likelihood that the number of MAC addresses visible in the core will grow large enough to exceed the FDB capacity of the core bridges. Addressing this problem may require more expensive core bridges with greater FDB capacity or, if higher capacity bridges are not available, the result may be poor performance due to excessive flooding. The deployment of PBBN in the data center can address this problem by reducing the number of MAC addresses visible to bridges in the core of the network.

III. THE REQUIREMENT FOR A C-TAGGED SERVICE

The PBBN architecture⁷ was designed to support the interconnection of *Provider Bridged Networks (PBNs)* across a backbone network. A PBN contains *provider bridges*. Traffic carried within the PBN is associated with *service VLANs (S-VLANs)*. *Backbone Edge Bridges (BEBs)* are deployed at the boundary of the backbone. A BEB port supporting attachment of the PBN is called a *customer⁸ network port (CNP)*. S-VLAN traffic is carried across the backbone via a *Backbone Service Instance (BSI)*.

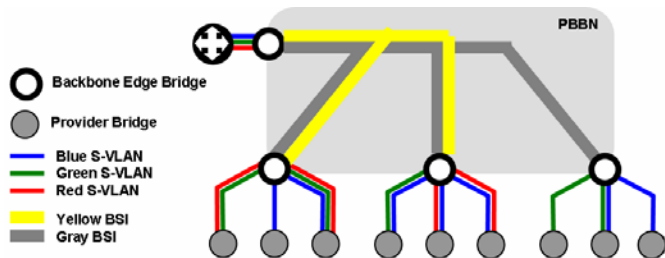


Figure 3: Mapping S-VLAN to BSI

The BSI can carry traffic associated with a single S-VLAN or multiple S-VLANs. In Figure 3, the green and blue S-VLANs are mapped to the gray BSI; the red S-VLAN is mapped to the yellow BSI.

The BEB supports two services for mapping PBN traffic to BSIs: a *port-based service* and an *S-tagged service*. The port-based service associates *all* traffic arriving on a CNP with a single BSI. The S-tagged service associates traffic carried on each S-VLAN with an associated BSI. The association between S-VLAN and BSI is established by provisioning.

When deployed in the data center, PBBN⁹ provides

⁷ The description of PBBN that appears here is *not* intended to be comprehensive. The intention is to provide just enough information to clearly explain the changes to PBBN needed to support the data center environment. Please notify the author of inaccuracies.

⁸ Calling this a *customer network port* is a bit confusing. Here, the term ‘customer’ is intended to denote that the PBN is a ‘customer’ of the PBBN. That is, the term ‘customer’ indicates that the PBN is a ‘client’ of the ‘service’ provided by the PBBN.

⁹ Strictly speaking, it is incorrect to use the term *PBBN* to describe a backbone that interconnects CBNs. It might be more accurate to change the term PBBN to Backbone Bridged Network (BBN), allowing it to be applied to

interconnection of *Customer Bridged Networks (CBNs)* across the data center core. A *customer bridge* is C-VLAN-aware, as contrasted with a provider bridge which is S-VLAN-aware. A customer bridge cannot attach to the PBBN via the S-VLAN service. It can attach *only* by means of the port-based service. *All* C-VLANs associated with a particular BEB CNP (ie., the port providing access to the PBBN) are mapped to the same BSI.

Thus, when PBBN is deployed in the data center, the C-VLANs associated with a CNP cannot be mapped to diverse BSIs. The result is that the data center administrator *cannot distribute load across the network at the granularity of the C-VLAN* as is currently the practice in the data center.

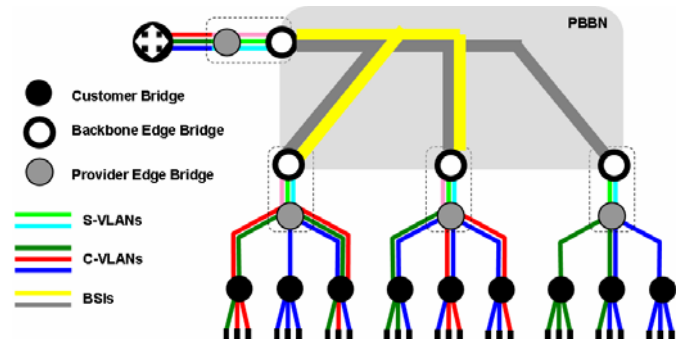


Figure 2: Inserting a Provider Edge Bridge

One way to overcome this limitation is to insert a *Provider Edge Bridge (PEB)* between the customer bridge and the BEB, as illustrated in Figure 2. Functionally, this places a PBN between the CBN and the PBBN. The PEB allows each C-VLAN to be mapped to an S-VLAN. The BEB can then map each S-VLAN to a BSI. This solution requires deployment of additional hardware (ie., the PEB) and increases complexity by requiring two mappings (C-VLAN \rightarrow S-VLAN \rightarrow BSI) where a single mapping (C-VLAN \rightarrow BSI) should suffice.

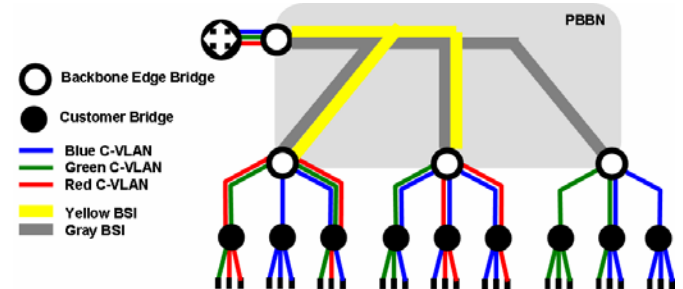


Figure 4: Mapping C-VLAN to BSI

A more efficient solution is extension of the PBBN architecture to support a *C-tagged service* as shown in Figure 4. The extension is straightforward as the C-tagged service is functionally identical¹⁰ to the existing S-tagged service except

the interconnection of PBNs or CBNs across a backbone. Since the term PBBN is now very well known, this change would bring more confusion than clarity, so we do not propose this.

¹⁰ Note that the proposed PBBN C-tagged Service differs from the existing PBN C-tagged Service described in IEEE 802.1ad. Specifically, the PBN C-tagged service does not allow direct bridging between C-tagged Service

that traffic carries a C-tag (Ethertype 0x8100) instead of an S-tag (Ethertype 0x88A8).

IV. REQUIREMENT FOR THE CUSTOMER BRIDGE AS BCB

At the PBBN edge, each BSI is mapped to a backbone VLAN (B-VLAN) for transit across the core. In Figure 5, the gray and yellow BSIs are mapped to the blue B-VLAN and the green BSI is mapped to the red B-VLAN. Source (B-SA) and destination (B-DA) backbone MAC addresses identify the endpoints of the BSI. Frames carrying this backbone encapsulated traffic are forwarded by *Backbone Core Bridges (BCBs)* in the interior of the backbone. A key feature of the PBBN architecture is that the BCB is identical to a provider bridge. This makes it unnecessary to purchase a new type of equipment for the backbone core. To support this feature, the *backbone tag (B-tag)* and the *service tag (S-tag)* are identified by the same Ethertype (0x88A8).

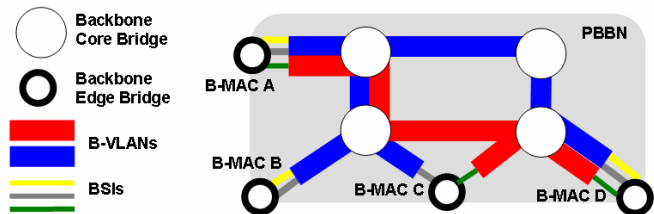


Figure 5: B-VLAN traffic transiting BCBs

The PBBN architecture specifies that a BCB is identical to a provider bridge. A data center bridge is, however, a customer bridge. Thus, deployment of a PBBN in the data center would require that data center core bridges change their mode of operation from ‘customer bridge’ to ‘provider bridge’. It is possible that the existing customer bridges do not support provider bridge operation, in which case, new equipment would be required. In the more likely case that the data center core bridges *are* capable of provider bridge operation, the administrator may prefer to allow these bridges to continue to operate as customer bridges. In this way, changes associated with the introduction of PBBN can be isolated to the BEBs.

The extension required to allow a BCB to be instantiated by a customer bridge is straightforward. The PBBN architecture is updated to specify that a BCB can be either a provider bridge or a customer bridge. The description of the interface between the BEB and the BCB is updated to describe a C-tagged interface between the BEB and the BCB when the BCB is instantiated as a customer bridge.

V. CONCLUSION

Two extensions to the PBBN architecture are proposed to allow effective deployment of PBBN in the data center. One

is support for a C-tagged service that would allow the mapping of an individual C-VLAN to a BSI. The second is support for the use of a customer bridge as the Backbone Core Bridge.

Acknowledgements—Thanks to Ben Mack-Crane for very useful editorial and technical comments. Other comments welcomed.

interfaces, while the PBBN C-tagged Service would permit this. In this respect, the behavior of the PBBN C-tagged Service is intended to be functionally the same as the PBBN S-tagged Service (except for the difference in associated Ethertype) and *not* the same as the existing PBN C-tagged service.