A Core/Edge Separation Based Bridging Virtualization Approach Oriented to Convergent Network

Tao Yu
CS Department
Tsinghua University,
Beijing, P. R. China
yutao@csnet1.cs.tsinghua.edu.cn

Jun Bi
Network Research Center
Tsinghua University,
Beijing, P. R. China
junbi@tsinghua.edu.cn

Jianping Wu
CS Department
Tsinghua University,
Beijing, P. R. China
jianping@cernet.edu.cn

Abstract—**Internet has been developing for more than three decades ago. Considering its popularity, ubiquity and scale, it causes immense difficulty for other networks with different architectures to be deployed or even get adequately evaluated. This paper introduces a core/edge separation based architecture model, which allows core networks with various architectures to coexist, enables end users to choose whichever core network at will, and simplifies the design, deployment, operation and management of edge networks. In this model, Carrier Grade Ethernet, which is still on-going evolving, is selected as transport technology in the edge networks. A radically new approach called Bridging Virtualization is used to form a convergent access platform to these different core networks for end users. Besides the increasing flexibility for the edge transport carriers, this model also improves end users' independence from the access provider. Moreover, it makes it possible the execution of network convergence step-by-step and can be used to realize it in the edge first. Finally, we apply this architecture model to design the campus networks of Tsinghua University, analyze its pros and cons and make some conclusion.**

Keywords-Virtualization; network convergence; Carrier Grade Ethernet; transport; Future Internet

I. INTRODUCTION

Due to the research and development effort revolved around the concept of New Generation Networks (NGN), Carriers' networks have been improved significantly. The access and aggregation networks especially, even the core networks, are undergoing architectural evolution. Although there are different approaches to NGN, convergence, ubiquity, mobility and security are the common goals for all of them.

The concept of network convergence has twofold meanings, i.e., first and foremost it means that there is a uniform packet transport platform; second it means various kinds of applications can be built on a common network architecture such as CCN (Content Centric Network) or some other virtualization based future network. Here we distinguish them as two functions and locate them in different layers of the network system which is divided into edge part and core part. For the second meaning, it is mainly implemented in the core part; for the first meaning, it is implemented in both the core and the edge, and the architecture and technology of the implementations in both of them may be different depending on their individual characteristics and requirements. For us, network convergence not only facilitates the deployment and application of new network services, but also encourages the emergence, setup and evaluation of innovative network architectures.

Currently, there are two main alternatives to achieve network convergence: IP/MPLS and Carrier Ethernet. In this paper we focus on the Carrier Ethernet proposal for the edge networks, especially the Carrier Grade Ethernet (CGE) solution aiming at a transport technology for carriers' networks, since the former one is closely related to IP technology, whereas the latter does not care about whatever the upper layer architecture and protocol will be and fits in with future Internet initiatives much more likely.

About Carrier Grade Ethernet, there are several technological components such as Provider Bridge (PB) [12], Provider Backbone Bridge (PBB) [13], Provider Backbone Bridge – Traffic Engineering (PBB-TE) [14] and Shortest Path Bridging (SPB) [15]. To date, all of them are IEEE standards except the SPB, of which PB and PBB address scalability and management issues, PBB-TE adds traffic engineering and SPB introduces a link-state protocol into Ethernet. In addition, there are other standards pertaining to Ethernet OAM. All of the above existing standards together with the future ones will finally make Ethernet an adoptable packet transport technology for carriers.

In order to join different transport carriers' edge network together and increase the ubiquity of NGN, virtualization technique is utilized to achieve a virtual connectivity platform, where different Carrier Grade Ethernet technologies coexist and cooperate closely to provide access to various core networks for end users. Each edge transport Carrier has its own geographical footprint, and it provides network connectivity for all the potential users without the discrimination of which core network clients they are (It can charge the core network providers for their customers' usage of its transportation infrastructure). This way ensures the customer independence from the individual edge transport carriers, which not only brings flexibility for the edge transport networks but also increases the mobility for the core networks. Different edge transport carriers' networks are interconnected through virtual bridging approach to form a

 ISBN: 978-1-61208-148-9

complete platform to provide access to each kind of core networks, such as the current Internet, the next generation Internet based on IPv6 and some other totally different networks as the Future Internet (FI) candidates.

By this way, we separate the whole network system into two parts, the edge part and the core part. The infrastructure in the edge just provides the convergent access platform for end users to different core networks, while at the core various networks with different architectures compete with each other to meet end users' increasing service requirements. Users can switch to different protocol stack or utilize different virtual machines in their terminals to choose among services provided by these core networks, the final judgment and decision about what architecture the future network will be depend on the most people who approve its services, then other core networks will diminish gradually and the transition to the Future Internet can be realized smoothly.

In this paper, we first present Carrier Grade Ethernet as a solid alternative to achieve convergent edge transport infrastructure in carrier's domain. Second bridging virtualization approach is proposed to provide isolate transport platform for each core network, and then Click tool is used to apply this design to Tsinghua University campus network (THUNET). Finally, we analyze, evaluate this model and complete this article with some conclusions.

II. NET CONVERGENCE AND RELATED WORK

What we mean network convergence is that people can use various network services through just one type of network. Although it is universally agreed the future network will be a convergent network, to what degree that the convergence will reach still remains uncertain. In this paper, we adopt the strictest definition of it: convergence of various services and convergence of connectivity. Meanwhile we need to point out that this definition of network convergence will not affect its realization step by step.

In the scenario of network convergence, the whole network will consist of three different parts; those are the users' networks, networks of connectivity providers and networks of service providers. Just as we say in the introduction, this paper concentrates on the convergence of connectivity providers' networks in the edge, and carrier grade Ethernet is chosen as the basis for this kind of convergence.

There are three main options for transporting Ethernet frame: Ethernet over SONET/SDH, Ethernet over MPLS and Carrier Grade Ethernet. For the first solution, its drawback lies in that the SONET/SDH equipment is costly and it has severe problems for multipoint solutions. The second solution does allow both point-to-point and multipoint connections, but again it is complex and rather expensive because it needs a routable network and IP/MPLS solutions at the core. For Carrier Grade Ethernet, it is the most suitable technology to provide Ethernet services since almost all the Internet traffic originates and ends in Ethernet. What's more, organizations for standardization including IEEE, IETF, ITU and MEF are still striving toward extending the scale and scope of Ethernet, especially aiming it at a suitable transport technology for carrier's networks. In the following we briefly describe and analyze the evolving process of Carrier Grade Ethernet.

Some requirements must be met for Ethernet if it can be considered as a transport alternative for carriers. First of all, it must provide standardized services without the need to change customer equipments. Besides this, it must support different types of quality of service in terms of bandwidth, packet loss, delay or jitter. Also, it must be scalable to be able to provide services for millions of customers accessing simultaneously to voice, video and data services. To guarantee reliability, resilience is an essential and a recovery time inferior of 50 ms (of SDH technology) is mandatory. Finally, standardized mechanisms for network monitoring, diagnosis and management are indispensable.

To fulfill the above requirements, first traffic from different customers must be segregated. IEEE802.1q is available for this purpose, but because the VLAN tag in 802.1q frame has only twelve bit length, at most 4094 VLANs can be used, which is not enough for carriers. Moreover, many customers need to assign VLAN ID for their internal networks by themselves and want to maintain this VLAN designation across the carrier's network, so there need much cooperation between carrier and customer which imposes great difficulty for carrier's operation if it is not impossible.

IEEE 802.1ad is developed to solve this problem, which uses a separate VLAN tag called S-tag for carrier's network in the 802.1ad frame. Similarly, the length of S-tag is twelve bits too, which has the same scalability problem as 802.1q. Besides, because the carrier's network and the customer's network share the same MAC address space in 802.1ad, which make them seem to be in a same bigger network, any changes in customer's network will have an impact on carrier's one, which is not desired by the carrier. From the customer's point of view, if their internal MAC addresses are exposed to the carrier's network, it may cause security concern. The most important of all, the BPDUs from carrier's network and customer's network should not interact with each other, which otherwise may cause unpredicted serious consequence. But currently in 802.1ad, there is no efficient way to distinguish them.

For all of the above reasons, IEEE802.1ah is proposed, whose frame format is as Figure 1 shows. 802.1ah uses a new carrier's MAC header to encapsulate (decapsulate) incoming (outgoing) frame. In this header, instead of a 12-bit S tag, a 24-bit field called I-SID (I-tag) is used to differentiate the service instances, which drastically resolve scalability problem. The forwarding is based on the new header's fields (B-DA, B-SA and B-VID), totally isolated from customer's addressing scheme. But 802.1ah still maintains flooding and STP mechanisms.

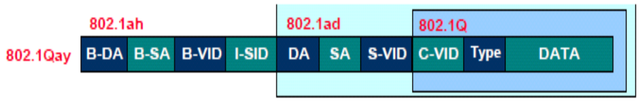

Figure 1. Carrier Grade Ethernet frame format

For carriers, traffic engineering capability of transport network is critical, but STP does not have such one and flooding wastes so much bandwidth that the carrier cannot bear. IEEE 802.1Qay aims to solve this problem. It is based on 802.1ah encapsulation, by disabling some well known mechanisms of Ethernet like STP, flooding, broadcasting or MAC learning, it eliminates the problems faced by the 802.1ah but needs additional management plane/control plane to populate its forwarding table. The forwarding decision is based on the combination of destination MAC address and VLAN ID (60 bits), providing enough capacity to traffic engineering.

IEEE 802.1aq [15] (Shortest Path Bridging) is another development that proposes an alternative to STP dependence. 802.1aq is a draft standard that uses 802.1ah data plane combined with the well-known link state protocol IS-IS [6]. This enhancement adds carrier-grade any-to-any infrastructure capabilities to the 802.1Qay point-to-point model. This is done by changing from a management system to IS-IS associated states and protocols to rule forwarding behavior. 802.1aq technology enables Ethernet to use the shortest path from any source to any destination, thereby allowing full use of the entire mesh connectivity and eliminating the need for complex Multiple Spanning Tree Protocols.

OAM (operation, administration and maintenance) is critical for Carrier-Class Ethernet. There are several standards regarding these fields: IEEE 802.1ag [7], it provides a mechanism for service fail proactive signaling; IEEE 802.3ah, it defines OAM capabilities for the first mile; IEEE 802.1AB, it allows topology discovery; ITU-T G.8031, it adds Ethernet protection mechanisms; and ITU-T Y.1713 [8], it gives additional management capacities to 802.1ag.

With all of the above developments in Carrier Grade Ethernet, carrier's network will converge to common transport platform based on native Ethernet and customers will be able to access the specified core networks via it.

III. System Architecture and Prototype Implementation

A. System Architecture

Today telephone network, Internet and cable TV network are still separate networks through which people access voice service, data service and video service accordingly, and these networks each have their own architecture and underlying technology. Even if there are some telecommunication operators or cable TV operators who claim that they can provide all these three services by themselves, users still hope that the communication media, terminal equipment and even the connectivity technique for these services could be unified too. This object seems simple and clear at a sudden, but the technical factor supporting this uniformity are extremely complex and difficult, lots of work is still in process of revision, especially for deployment in large scale.

As we mentioned before, the network convergence we think should include the service convergence and the connectivity convergence, and we separate the whole network into core and edge. In this scenario, there will be three different business roles that do not match those of today exactly. They are users, transport carriers and service providers.

Because there are already Ethernets deployed in users' premises, and Ethernet has the advantage of low cost, ease to use, maintain, upgrade, etc., we continue to use Ethernet as a convergent communication platform in user's domain.

In the domain of edge transport carrier lying between the user's network and the core network, based on the analysis in the second section, we choose Carrier Grade Ethernet (CGE) to construct a convergent packet transport platform, in which various technical components of CGE are utilized depending on actual needs to aggregate users' networks into the POPs of core networks. There users' networks are connected to various types of core networks through different gateways. These gateways may be IPv4 router, IPv6 router, CCN router, and so on.

At the interface between user's network and the edge transport network, a layer two device is used instead of a layer three router. The system architecture will look like as Figure 2 shows.

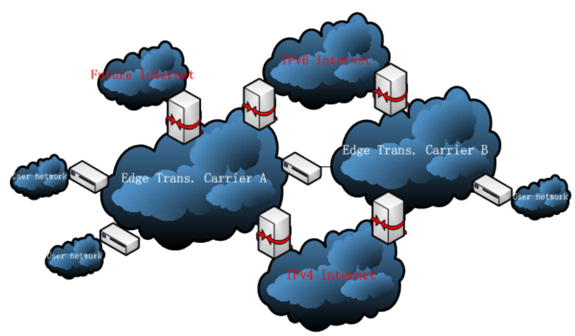

Figure 2. System architecture

Due to the large number of users' networks, heavy routing and processing burden will be imposed on the core network's gateway, e.g., they are demanding for huge port number, link bandwidth and processing power, etc. About the problem of port number, port virtualization can be used to subdivide one physical port to lots of sub interfaces or virtual interfaces, and then the technique of "router on a stick" may be used to provide routing among huge number of networks in one device. The problems of link bandwidth and processing power of a single router will be resolved or improved by the distributed router technology. Furthermore, along with the advancement of router virtualization, some device which merges the IPv4 router, IPv6 router and CCN

 ISBN: 978-1-61208-148-9

router will finally appear and get ready for the utmost realization of core network convergence.

In the edge transport network between user's network and core network, in order for end users to connect to different types of core network, network virtualization is utilized to partition the end users' traffic for different core networks into different network instances. Because the transport network embeds the users' Ethernet frame into a new Ethernet frame when it reaches and de-encapsulates it when it leaves, which separates the transport carrier's network from the user's network hierarchically and processes the user's traffic recursively, and all these actions happen in the layer two, we call it connectivity virtual bridging or bridging virtualization. Any Layer 3 protocol from the core networks can use this connectivity transparently. In addition, due to the limited footprint of every edge transport carrier's network, similar virtual bridging technique among carriers can be used for expansion. Hence there are two layer bridging virtualization in our approach. The first layer virtualization is within each carrier's domain, where there is an individual virtual transport platform per core network. The second layer virtualization is across carriers, which further virtualizes the virtual networks in the first layer in each carrier which are for the same core network.

The second layer virtualization has two purposes. First it can be used to extend the geographical footprint of the Future Internet candidates with innovative architectures. For these networks, its initial deployment is in small scale and provided by few carriers, by this way, more carriers can provide access to it and lots of end users are able to try out it at its beginning. Second for extensively deployed IPv4 and IPv6 Internet, this type of virtualization could provide reliability and performance guarantees for end users' networks by multi-homing.

The schematic diagram of bridging virtualization is demonstrated in Figure 3.

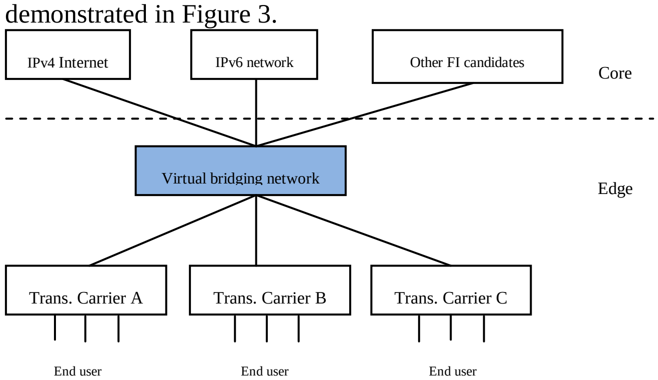

Figure 3. Schematic diagram of bridging virtualization

The advantage of this architecture model is that new network instance can be added easily and any solutions in control plane other than IEEE proposal can be used. This means that any new network architecture can be incorporated into a common transport platform by adding a new instance whose forwarding behavior can be customized for it.

B. Prototype Implementation

We use Click tool [9] to prove this network prototype for Tsinghua University Campus Network (THUNET). This tool was developed at MIT initially and it allows simulation networks to interact with real network nodes. It needs to be pointed out that because there is only one operator for THUNET, so we don't consider virtual bridging across multiple carriers in our initial prototype implementation. If Tsinghua University or CERNET wants to expand the footprint of IPv6 Internet or the future FI, it can permit other telecommunication operators (such as China Mobile or China Unicom) to provide access to these networks for ambient social users through this virtual bridging platform of Tsinghua University, meanwhile keep them separate from THUNET.

The Click design for the component at the interface between user network and edge transport network is shown in Figure 4.

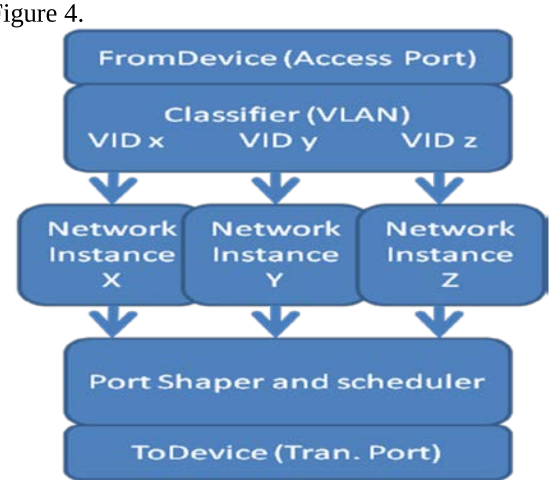

Figure 4. Click design for virtual bridging

When a frame enters the bridge node, first it must go through a classification process to determine which network instance it belongs to, then it will be processed and encapsulated with new header, after that some specific component of CCE will determine and forward this frame to proper output port. Since all the network instances share each access port of the bridge node, there must be a scheduler for the frame forwarding among these network instances.

In addition, in the Click design, the edge bridge and the core bridge are different. The edge bridge need encapsulate, process and forward frames, while the core bridge need only forward frame according to control protocol or management system. The edge bridge sets value of fields in the new header according to the type of packet, the destination gateway, the operator's policies, and so on. Once the frame finishes encapsulation process and enters some specific network instance, many protocols such as STP, IS-IS, proprietary protocols and (or) management system can be applied in this network instance to direct its forwarding action.

IV. Discussion

This type of network architecture model has the following advantages:

First, it eliminates the use of low-end branch routers completely and decreases the campus IP network's burden of distributed network configuring, managing and troubleshooting.

Second, by using core routers to complete the function of routing, security enforcement, user authentication, protocol and address translation, etc., it will improve the IP network's reliability, stability, security, manageability and controllability, and increase the utilization rate of core router's backplane, which makes campus network construction much more cost efficient.

Third, since low-end branch routers or layer 3 switches do not support IPv6 and FI related protocols generally, elimination of them will make easy or possible for the deployment or expansion of IPv6 Internet and Future Internet, and the application based on them will spread, which allow these types of internets to get adequate usage and efficient evaluation.

Fourth, in the edge transport network, it is easy to implement traffic engineering to maximize its usage and improve the average Internet bandwidth per user.

Fifth, there is no need to make any change of the current IP routers and end hosts for this network architecture to work.

Sixth, this type of network architecture is more likely to support multiple architecturally different networks including Future Internet, facilitating their deployment and coexistence.

However, this architecture model has its drawbacks too:

First, for big organizations with multiple VLANs, because the traffic across inner VLANs need to be transported to the POP for routing, their delay will increase, though this will not affect the routine applications with no strict requirement noticeably. Besides, this model will cause security concern for those who has some security requirement on the traffic across their inner VLANs. But there are solutions for all of the above problems, for example, they can be solved by the scalable enterprise Ethernet architecture, for more details, please refer to SEATTLE [18].

Second, to aggregate users' networks and transport them to the operator's PoP for routing, the number of VLANs accumulated there will be very high, so the VLAN ID should be carefully designed and the map of routers' ports to VLAN IDs must be planned in detail to avoid confliction, which will be more demanding for the capability of network planning and management than the current IP solution.

In summary, this network architecture does not intend to replace IP protocol of present Internet; instead, it still needs protocols including IP to interconnect with other networks. But it can free end users from tyranny of local ISP, simultaneously support internetworks with different IP versions, and much likely support Future Internet and its alternatives deployed in the core.

The purport of this architecture is trying to expand the scope of edge networks, minimize the impact of IP protocol, which is mainly located or applied in the core in our model, and pave the way for the appearance and transition to post-IP Internet paradigm conceptually and practically.

Our contribution lies in the following three aspects:

First, our model changes the architecture of edge networks. It corresponds to Cloud Computing trend in network field, uses virtualization-ready core routers with rich function, high performance and high reliability to route, enforce security, translate between protocols and addresses, etc.

Second, this model design is oriented to the convergent new generation networks, it separates whole network into core and edge, and makes possible for the realization of network convergence step by step. Moreover, it realizes convergence in the edge firstly and preliminarily, which is helpful and promotive to convergence in the core.

Third, the two-layer virtualization approach makes possible for coexistence of multiple core networks, lowering the barrier to entry for network innovation, and fulfills the function of multi-homing for user's network in layer 2.

V. Conclusion

Until now, there is still confusion about which way to take for the future Internet design: incremental approach or clean-slate approach [17]. Unsettlement of this problem impedes the rapid development for Future Internet. The approach introduced in this paper can bring an end to this argument and allow two kinds of design to coexist and develop in parallel, while the final decision about what the future Internet will look like depends on the choice of most users.

In this paper, based on the concept of network convergence, we propose to partition the whole network into two parts: the core and the edge. The different networks with various architectures, such as IPv4 Internet, the next generation Internet based on IPv6 and other future Internet candidates, lie in the core. While at the edge, a two-layer bridging virtualization approach is designed to form a uniform access platform for all the core networks, enabling end users choose among multiple core networks freely. This model lowers construction cost of campus networks with better security and manageability, makes them scalable for new network architectures, and improves user networks' reliability and performance.

With its advantages and advancement of Carrier Grade Ethernet, we are optimal the deployment and execution of this architecture model will become a reality.

VI. Future Work

In our architecture model, we define business roles as transport carriers and service providers, who cooperate to provide the network connectivity and network services. Considering how to setup the proper business relation between them to spur the development of this network

 ISBN: 978-1-61208-148-9

architecture is valuable. Moreover, when network faults or service interruption happen, how to settle responsibility among them or improve the accountability of this type of network is meaningful.

The technique for scalable Ethernet architecture is very helpful to our architecture model, which not only facilitates its execution but also can be explored for its application in the edge transport networks. So the next step for us is to adapt this technique and apply it to CGE, proposing new Carrier Grade Ethernet draft and promoting it to be standardized.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation of China under grant 61073172, Specialized Research Fund for the Doctoral Program of Higher Education of China under Grant 200800030034, and Program for New Century Excellent Talents in University. Many thanks to Kilnam Chong and Yu Zhu for their useful feedback.

REFERENCES

[1] Y. Zhu, R. Zhang-Shen, S. Rangarajan, and J. Rexford, “Cabernet: connectivity architecture for better network services,” Proceedings of the 2008 ACM CoNEXT Conference, pp. 1-6, December 09-12, 2008, Madrid, Spain

[2] A. Kasim, Delivering Carrier Ethernet: Extending Ethernet Beyond the LAN, McGraw-Hill. 2008

[3] A. Meddeb, “Why Ethernet WAN Transport?” IEEE Communications Magazine, Nov 2005, pp. 136-141

[4] D. Allan, N. Bragg, A. Mcguire, and A. Reid, “Ethernet as Carrier Transport Infrastructure” IEEE Communications Magazine, Feb 2006, pp. 95-101

[5] D. Fedyk and D. Allan, “Ethernet Data Plane Evolution for Provider Networks” IEEE Communications Magazine, March 2008, pp. 84-89

[6] D. Allan, P. Ashwood-Smith, N. Bragg, and D. Fedyk, “Provider Link State Bridging” IEEE Communications Magazine, Sept 2008, pp. 110-117

[7] IEEE Std. 802.1ag, Virtual Bridged Local Area Networks – Connectivity Fault Management, 2007

[8] ITU-T Y.1731, OAM functions and mechanisms for Ethernet based networks, 2006

[9] Click Project, http://read.cs.ucla.edu/click/ [last access: June 28, 2011]

[10] MEF TS 6.1, “Ethernet Services Definitions – Phase 2”, Technical Specification, April 2008.

[11] IEEE Std. 802.1Q, Virtual Bridged Local Area Networks, 2005

[12] IEEE Std. 802.1ad, Virtual Bridged Local Area Networks: Provider Bridges, 2006

[13] IEEE Std. 802.1ah, Virtual Bridged Local Area Networks: Provider Backbone Bridges, 2008

[14] IEEE 802.1Qay, Virtual Bridged Local Area Networks: Provider Backbone Bridge Traffic Engineering, 2009

[15] IEEE 802.1aq/D0.4, Draft Standard, Shortest Path Bridging, 2008

[16] J. Matias, E. Jacob, M. Aguado, and J. Astorga, “Enhancing NGN’s Versatility for Multi-Service Support: the Bridging Virtualization Approach”, The 10th International Conf. on Telecommunications, ConTEL 2009, pp. 367-374.

[17] Feldmann, “Internet clean-slate design: what and why?”, SIGCOMM Computer Communication Review 37 (3) (2007), pp. 59–64.

[18] C. Kim, M. Caesar, and J. Rexford, “Floodless in seattle: a scalable ethernet architecture for large enterprises”, Proceedings of the ACM SIGCOMM 2008 conference on Data communication, pp. 3-14, August 2008, Seattle, WA, USA

 ISBN: 978-1-61208-148-9