Benefits brought by the use of OpenFlow/SDN in the AmLight intercontinental research and education network

Julio Ibarra, Jeronimo Bezerra, Heidi Alvarez, Luis Fernandez Lopez
Florida International University (FIU)
Miami, Florida
{julio, jbezerra, heidi, llopez}@fiu.edu

Michael Stanton, Iara Machado, Eduardo Grizendi
Rede Nacional de Ensino e Pesquisa
Rio de Janeiro, Brazil
{michael, iara, eduardo.grizendi}@rnp.br

Donald A. Cox, III
Vanderbilt University
Nashville, Tennessee
chip.cox@vanderbilt.edu

Luis Fernandez Lopez
University of São Paulo
São Paulo, Brazil
lopez@ansp.br

***Abstract*—Operating unprotected network links for international collaboration between research and education communities, subject to a high-availability production service requirement, is challenging. Provisioning circuits, maintaining a loop-free network topology, and configuring multi-path redundancy to provide high availability are complex processes, which involve extensive coordination between, and manual configuration operations carried out by, multiple network operators, resulting in high operations costs. Moreover, network-oriented research applications increasingly require the capability to program network functions to satisfy particular requirements, such as high tolerance, low delay, end-to-end visibility, etc. We describe a solution, based on Software-Defined Networking (SDN), which significantly lowers the operations costs by automating most network operations and reducing coordination efforts between network operators. The design of the network, before and after SDN was deployed, is discussed. For each network function migrated to SDN, a comparative analysis is provided with metrics, first to represent real measurements before and after each SDN deployment scenario, and second, to describe findings of reduced operations costs.**

Keywords—international network links; research and education networks; automating network operations; software-defined networking; OpenFlow.

I. INTRODUCTION

Science Research and Education (R&E) communities communicate, cooperate and collaborate in an global context. Members of such communities access remote instruments, share data, and computational resources that are geographically distributed, in support of international research collaborations [1]. Networks designed to support these R&E community collaborations are interconnected internationally using inter-continental network links [1].

Americas Lightpaths (AmLight) is a project of the U.S. National Science Foundation International Research Network Connections (IRNC) program to facilitate science research and education between the U.S. and the nations of Latin America [2]. AmLight operates a number of international network links connecting U.S. R&E networks to similar networks in Latin America. The AmLight links are shared and operated collaboratively by Florida International University (FIU) [3], the Academic Network of São Paulo (ANSP) [4], and Rede Nacional de Ensino e Pesquisa (RNP) [5]. The AmLight network uses a double ring topology formed by four spatially-diverse unprotected (a.k.a. "linear") 10Gbps connections, providing redundancy in case of a fiber cut on one of the network links, by moving data in both clockwise and counterclockwise directions around both rings. From São Paulo, two links head east and north to Miami, Florida, with one stopping at RNP points of presence in Rio de Janeiro and Fortaleza. The remaining two diverse links head west and north to Miami, with one of them stopping in Santiago, Chile. The AmLight network topology is represented in Figure 1.

Figure 1 AmLight Network Topology

We describe two objectives of the AmLight project towards fulfilling its programmatic activities: (1) Improving operations efficiency; and (2) Providing the capability for applications to program network functions.

Improving Operations Efficiency: Operating inter-continental network links when the end-to-end path is not under the control of a single operator is challenging, especially when it involves multiple technologies, different equipment vendors and

The AmLight project is made possible through the funding support of the National Science Foundation (awards ACI-0963053, ACI-1341895, ACI-1140833), the Academic Network of Sao Paulo FAPESP grant# 2008/52885-8), Rede Nacional de Ensino e Pesquisa (RNP), the Association of Universities for Research in Astronomy (AURA).

management philosophies. Provisioning new services, maintaining a reliable network topology with multi-path redundancy to support a high-availability service requirement for both production and experimental (R&E) applications can be a complex process. In a network with multiple operators, these processes involve a high degree of coordination, and use of manual procedures between multiple network operators and, sometimes, even users. Operation of these processes has a high cost, and could lead to errors and unexpected downtime. As an example, consider a layer 2 circuit between two universities, one in Brazil, and one in Europe. Between these two universities, it is quite common for network traffic to transit five, six, or even seven separate R&E networks operating different technologies, from layer 1 to MPLS. So, deploying this new layer 2 circuit requires a high degree of coordination between all networks involved; e.g., VLAN ID selection, and bandwidth and Quality of Service requirements. A provisioning activity like this could take weeks. Moreover, troubleshooting these circuits is also a very complex activity.

Network Programmability: Network-oriented applications for science research increasingly depend on the capability to program network functions to achieve particular requirements, such as high tolerance, low delay, end-to-end visibility, multipath, etc. Big data [6], Science DMZ [7], HD video streaming [8] could benefit from network programmability to optimize their flows and react to network conditions.

In this paper, we describe our experiences using Software-Defined Networking (SDN) [9] and OpenFlow 1.0 [10] to improve operations efficiency and to support network programmability on the AmLight intercontinental network infrastructure. Provisioning and programmability are two use cases defined to measure operations efficiency.

Our hypothesis is that OpenFlow/SDN significantly simplifies provisioning and network management functions, resulting in a higher degree of operations efficiency by automating most network operations and reducing coordination efforts between network operators.

Network programmability is a new capability on AmLight as a result of the SDN deployment. Network programmability functions, along with potential applications will be described, for example Software-Defined Exchanges (SDX).

The rest of the paper is organized as follows: Section II describes the characteristics of the AmLight network before and after the deployment of OpenFlow/SDN. Measures for the provisioning process are provided. Section III discusses findings and lessons learned from the Openflow/SDN deployment and its impact on network management. Section IV discusses future work going forward using the SDN capability on the AmLight network. Section V summarizes our conclusions.

II. Network Management of a Multi-path Intercontinental Network

A. AmLight Network Description before SDN

When the current AmLight network was designed in 2012, the main focus of its configuration was aimed at increased resilience. To guarantee a resilient platform for innovation, AmLight links were configured creating two backbones, as illustrated in Figure 2: A) One Academic Layer 2 Ring; B) One Academic IP Ring.

To connect both backbones, two 10Gbps links were installed in São Paulo through the optical infrastructure provided by ANSP. These links are also used for IP and Layer 2 traffic exchange, and to enable redundancy between them. In the event of a double failure (fiber cut, devices outage, etc.) in one backbone, the other backbone can provide full connectivity. To increase the resilience, the AMPATH International Exchange Point[1] [11] in Miami, has two network devices – configured as a cluster - to terminate the international links. So, even in the event of downtime in one of the devices, the AmLight network remains operational. The Academic Layer 2 Ring is primarily used for academic traffic and experimentation, and its configuration will be the focus of this paper.

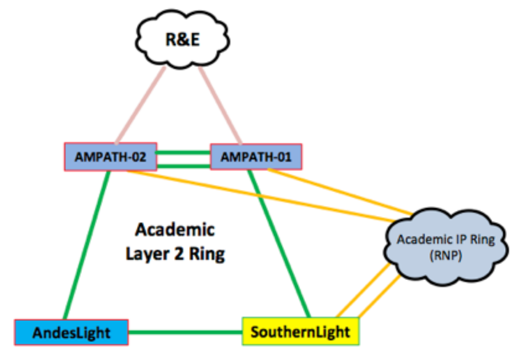

Figure 2 AmLight Topology, showing how the SouthernLight, AMPATH and AndesLight exchange points are connected

Before the migration to SDN, the AmLight Layer 2 Ring was based on Layer 2 technologies; specifically VLANs - *to encapsulate all traffic and to do the proper forwarding* -, and Brocade per-VLAN Rapid Spanning Tree [12] - *to guarantee a loop-free topology*. This infrastructure was used in many different demos and experiments, and almost all of these demos had a remote partner not directly connected to AMPATH, but connected through other R&E Networks (RENs) (for example, Internet2 [13] and ESnet [14]). This characteristic added complexity to all provisioning tasks as mentioned before, and it will be further described.

1) Provisioning before SDN:

Provisioning and troubleshooting scenarios involving multiple operators are the most complex and time-intensive activities, because the process normally involves a high degree of communication and coordination among the originating operator, the operators of the transit networks, and the operator of the destination network. For example, when a researcher requests a new layer 2 circuit for an experiment,

[1] AMPATH International Exchange Point is a high-performance Internet

he/she needs to contact his/her university's campus network team, which then needs to contact the last-mile service provider. Successively, each REN operator in the chain extending from the requesting researcher's campus to the destination site's last-mile provider will contact the next downstream REN. All these network operators have to discuss, agree on and deploy layer 2 circuits, MPLS pseudowires or dedicated layer 1 services in order to create a single international layer 2 domain. Working with different vendor products adds more complexity to the provisioning process. On average, all layer 2 circuits that crossed more than three network domains took, at least, one week to be fully provisioned and tested. Some circuits took almost eight weeks to be provisioned. Similarly, troubleshooting scenarios increased in complexity as the number of network domains in the end-to-end path increased. Table I describes coordination costs to the provisioning process as the number of network domains in the path increases.

TABLE I. COORDINATION COSTS FOR THE PROVISIONING PROCESS AS THE NUMBER OF NETWORK DOMAINS IN THE PATH INCREASES

Number of domains involved in the path	***Average number of days to provision a new circuit***	***Average number of e-mails exchanged***
Up to three	5	10
More than three	12	65
Domains between continents (America and Europe, etc.)	45	100

2) Managing an inter-continental multipath network

As described previously, the two rings that create the AmLight network are connected through two 10G links; so one backbone can provide resilience to the other. The AmLight IP Ring uses Juniper routers with MPLS; the AmLight Layer 2 Ring uses Brocade switches with VLANs and per-Vlan RSTP. The full configuration to provide the mutual redundancy was achieved using dedicated MPLS pseudowires, deployment of QinQ [15] and some dedicated 10G ports. This solution, while it meets functional requirements, has two drawbacks: It is complex, and it results in higher equipment cost (CAPEX). Case in point, at least three 10G ports are fully dedicated to handling occasional double failures. In 2013, only one such double failure occurred.

More significant than the high CAPEX for this solution was the human effort involved to achieve the resilience objective. A number of senior network engineers from RedClara [16], ANSP, FIU and RNP were assigned to discuss the solution and how it would affect each network using AmLight links. Due to the complexity of the preferred solution, the whole process took eight months and involved five network engineers. Most of the complexity was due to the different equipment vendors and technologies involved, and the resulting lack of interoperability between all different protocol implementations.

3) Programmability

Network programmability of the AmLight network was not part of its initial design. Its complexity made it impossible for researchers to program AmLight for experimentation. The only resource available to researchers was visibility: they could have access to the network devices through Simple Network Management Protocol (SNMP). With SNMP access, they could see which links were operational, as well as their utilization and interface errors, if any. This lack of programmability was one of the key drivers for SDN deployment in the AmLight network.

B. AmLight Network Description after SDN Deployment

The two main drivers for SDN deployment in the AmLight network were the optimization of provisioning activities, especially those involving multiple domains, and the provision of a programming capability. The SDN deployment consisted of two main phases:

- Phase 1: Modeling and reproducing the AmLight operations in a controlled environment using the same devices. The objective of this phase was to test the OpenFlow support of the AmLight switches to confirm their code and all SDN applications were ready to support the network functions in use;
- Phase 2: Migrating of the network functions in use to an SDN approach. The strategy and the migration plan were created alongside Phase 1. Phase 2 was deployed on August 31st, 2014.

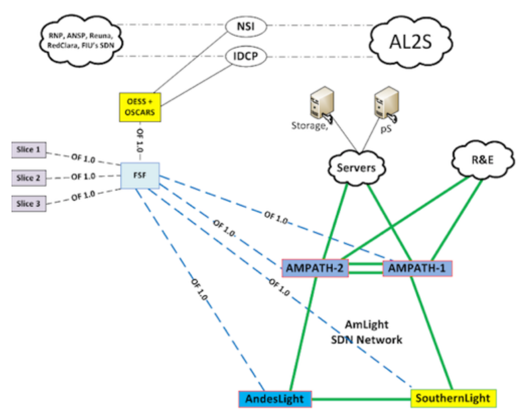

Figure 3 AmLight SDN Big Picture

Figure 3 provides a representation of the SDN implementations in the AmLight network. This Figure has three key pieces of information: (1) the control plane connections between switches to the FSF (FlowSpace Firewall[17], an OpenFlow firewall explained in Section B.2), using OpenFlow 1.0; (2) the FSF as a proxy between OpenFlow controllers and switches; and (3) the OESS+OSCARS server, responsible for network orchestration and inter-domain communication. Both OESS and OSCARS will be described in Section B.1.

The next section describes the effect on provisioning of network services as a result of the SDN deployment in the AmLight network.

1) Provisioning

The main idea of SDN is to move the control plane function from the network devices to a centralized network orchestrator. This network orchestrator has a full understanding of the network topology and, using this topology information, is able to send OpenFlow entries to all network devices, in order to configure their data plane actions.

Due to the academic and collaborative nature of AmLight, the network orchestrator adopted was the Internet2 Open Exchange Software Suite (OESS) [18]. It is the only orchestrator available with support for inter-domain communication, through the On-demand Secure Circuits and Advance Reservation System (OSCARS) [19]. OESS works through a Web User Interface, which makes it easy to manage. Due to its integration with OSCARS, OESS allows the provisioning of local and multi-domain circuits. For example, it is now possible to provision a circuit from SouthernLight [20], in São Paulo or AMPATH in Miami, to MANLAN [21], in New York City, using a secure web-based interface, with diverse Access Control Lists profiles. The Internet2 Advanced Layer 2 Services (AL2S) [22] network has been using OESS for almost two years, confirming it is a robust and stable platform for layer 2 service provisioning.

Having a single network orchestrator to manage AmLight, which includes the AMPATH, SouthernLight and Andes Light[2] exchange points allows a network engineer from ANSP, RNP or FIU to provision a layer 2 circuit without prior coordination with the other network teams, reducing to zero the number of emails exchanged. Moreover, with the multi-domain feature, network engineers will no longer need to contact the Internet2 NOC to request layer 2 circuits within the Internet2 nor the ESnet national backbone networks, since they both support OSCARS. In the future, with the Network Service Interface protocol [23] (NSI), more academic networks will be reachable through OESS.

TABLE II. COORDINATION COSTS TO THE PROVISIONING PROCESS WITH SDN DEPLOYED

Domains involved in the path	***Average time to provision a new circuit***	***Average number of e-mails exchanged***
RNP, ANSP, RedCLARA, AmLight, Internet2, ESnet	< 2 minutes	0
With other domains using OSCARS or NSI support	< 2 minutes	0
With domains not using OSCARS or NSI support, with up to three networks in the path	5 days	10
With domains not using OSCARS or NSI support, with more than three networks in the path	12 days	65
With domains in other continents not using OSCARS or NSI support	45 days	100

[2] AndesLight is not yet an exchange point. It is currently operating as a Network Access Point in Chile, supporting interconnectivity for AmLight.

The results in Table II, compared with those in Table I, show that by using SDN, the provisioning activity was measurably both less complex and less time-consuming. The complexity of the provisioning in the past was caused both by the coordination required between the network operators, and by the complexity of the network configuration due to the multiple protocols involved. A single orchestrator can now handle all devices at once, because OpenFlow provides a common interface.

Figure 4 below shows a Layer 2 circuit, created using the OESS user interface. It is now possible through this single interface to manage this Layer2 circuit, see its utilization, and to confirm if link protection is working properly.

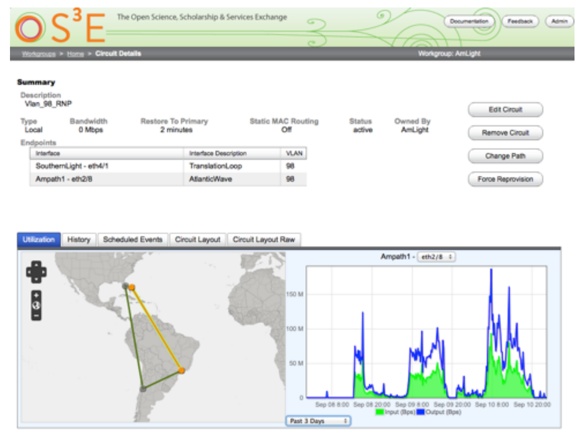

Figure 4 Layer 2 circuit provisioned using OESS

2) Network Programmability

The introduction of a network programmability capability is the biggest achievement of this new network. In this new environment, researchers can now deploy their network-oriented applications and use AmLight as a real platform for innovation. Being network-aware means that these applications will be able to provision their circuits, including capacity on demand, and to react to network conditions, such as increasing delay and packet loss.

Network programmability was deployed in the AmLight network using Internet2's FlowSpace Firewall (FSF) [17] - an OpenFlow firewall that controls what OpenFlow controllers can do to the OpenFlow switches. FSF makes it possible to create a new service called a network "slice" - a dedicated virtual network where a user can perform experimentation - with specific ports and VLAN ranges [24][25]. A network slice allows multiple tenants to share the same physical infrastructure. A tenant can be a customer, requiring his own isolated network slice; a sub-organization that needs its own slice; or an experimenter who wants to control and manage some specific traffic from a subset of endpoints. With slices, a controller in one slice cannot interfere with other slices; for example, it cannot remove flow entries or overlap them. Within its VLAN range, an OpenFlow controller can create flow entries using any field from layer 2 and/or layer 3 headers, giving the researcher the possibility of highly customizing his application, or even of creating his own network protocols.

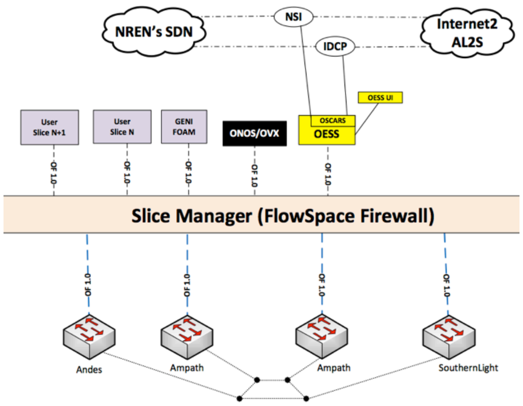

Figure 5 FlowSpace Firewall at AmLight

Figure 5 shows how applications and switches interact, giving a better overview of all layers involved. It is possible to see that FlowSpace Firewall is the software that virtualizes all switches and links. Furthermore, not only academic researchers could benefit from the network programmability capability in the AmLight network. Network engineers could also provision a slice for tests and learning; for example, to test network applications before putting them into production. Operators could use a slice to test a new controller or new orchestrator, or even develop their own; vendors could use a slice to test new features through a secure approach in a production network.

III. Findings

For approximately four months, the AmLight Engineering team discussed, designed and tested the orchestrator software, all switches, and the FlowSpace Firewall, before they were sufficiently convinced they could develop a plan to safely deploy SDN in the AmLight network. Although non-academic IP VLANs represent only 3% of all VLANs in use, these VLANs represent 60-70% of all traffic. It was decided to initially only move the academic VLANs to SDN. Tests were designed to learn how to keep both networks operating in parallel without impacting each other.

In spite of the fact that OpenFlow 1.0 was released more than four years ago, and equipment vendors deployed it more than two year ago, it is still considered "new" compared with more traditional network protocols. So, the AmLight Engineering Team was prepared for errors and restrictions to appear in all the code involved, and every effort was taken to limit the impact of these errors. Even a Disaster Recovery Plan was created to avoid extended downtime.

The following key lessons were learned: (1) Network devices still face important limitations when addressing interoperability issues configuring both legacy protocols and OpenFlow in the same switch. (2) Features such as link aggregation, sFlow and some Layer 2 control protocols might not be supported on OpenFlow ports. (3) The number of VLANs per port in hybrid ports (ports that support OpenFlow and IP traffic at the same time); the number and kinds of statistics per flow and per line card; and the control plane communication are all challenges to be overcome when deploying SDN in the current production devices.

Some of the features mentioned, for example, link aggregation, are a real issue when deploying OpenFlow 1.0, since its specification does not make it clear enough how vendors should support it. At AMPATH, this became a problem, since all devices are connected through aggregated links. This limitation was overcome with new connections established only for OpenFlow traffic. Instead of reducing capital expenditures and freeing 10G ports, more 10G ports were necessary to overcome the link aggregation issue. Another key issue was the fact some switches don't support control plane messages over Openflow entries. Due to its ring characteristics with all links configured as OpenFlow ports, control plane communications between the FlowSpace Firewall and all switches had to be built over one of the AmLight's member networks, acting as an out-of-band access. As AmLight is a collaborative project, counting on one of its members for out-of-band access wasn't a problem, but it might be for other networks.

As shown in Figure 5 and based on the findings, it is possible to understand why the network community is so interested in OpenFlow and Software-Defined Networking: having a centralized controller with a standard southbound interface makes most of the network activities simpler and more efficient.

Our hypothesis was confirmed after a short time of operation in the new network, when provisioning became almost completely automated, lowering the coordination time from weeks to minutes.

IV. Future Work

The possibilities of the current AmLight network are still restricted by the limited number of features of OpenFlow 1.0. Speeding up convergence, supporting QinQ and metering comprise the next focus for AmLight. These features are only available, or are better supported, by OpenFlow 1.3 [26], which is on the roadmap for the switches currently deployed at AmLight.

Since AmLight provides connectivity between AMPATH and SouthernLight, it plays the role of a distributed Internet peering fabric. AMPATH is also part of another distributed Internet peering fabric, called AtlanticWave [27]. AtlanticWave connects MANLAN, AMPATH, Southern Crossroads (SoX) [28] and Mid-Atlantic Crossroads (MAX) [29], using the Internet2 AL2S network. So, future work will explore the integration of AmLight and AtlanticWave, to create a single distributed Internet peering fabric, with full support for OpenFlow. By means of this unified intercontinental distributed peering fabric, we envision the need to interconnect our users' SDN networks, to extend atscale experimentation for our researchers. To achieve this level of connectivity, AmLight is proposing a new project, called Software-Defined Exchanges, or SDX.

An SDX could provide a capability to prototype an OpenFlow network where Autonomous Systems members of each Internet peering fabric could exchange traffic based in layer 2, layer 3 or even layer 4 fields of the frames [30]. The key idea is to have a new entity, called an SDX controller, responsible for creating OpenFlow entries in all switches of AmLight and AtlanticWave, through the FlowSpace Firewall installation of each domain. Figure 6 represents the idea, where each AS would have a different view of the IXP, using sandboxes created by the SDX controller.

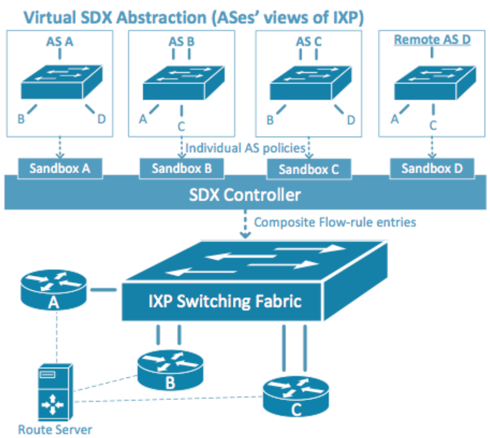

Figure 6 SDX diagram showing the SDX abstractions of each AS' view of the peering fabric.

SDX would introduce an important capability for AmLight, since our users are starting to deploy SDN in their own networks and need to interconnect to other SDN domains. Resources such as NSI or OSCARS are focused on resource allocation in a circuit-oriented approach only. With SDX, we expect to give our users a wide range of possibilities to manage how their flows will be forwarded on AmLight and AtlanticWave.

V. Conclusions

The deployment of SDN and OpenFlow on the AmLight network has improved operations efficiency and introduced programmability as a capability for the research and education community. The time spent in provisioning end-to-end circuits across multiple network domains was reduced by many orders of magnitude. OpenFlow 1.0, which was recently shipped by many vendors, offers some risks, and can create incompatibilities with legacy protocols. The extra usage of 10G ports to overcome the link-aggregation restriction increased the total cost of the solution and some existing monitoring components were lost due to some legacy line card limitations. But, to conclude, the solution worked properly, confirming the hypothesis that SDN and OpenFlow, even with all their risks at this moment, created a worthy solution, especially for academic environments.

Acknowledgment

The authors would like to thank Florida International University, Florida LambdaRail and Internet2 for their support of the AmLight project.

References

[1] International Research Network Connections (IRNC), National Science Foundation Program Solicitation 14-554, http://www.nsf.gov/pubs/2014/nsf14554/nsf14554.htm

[2] AmLight - America's Lightpaths, http://www.amlight.net/

[3] Florida International University, http://www.fiu.edu/

[4] ANSP - Academic Network of Sao Paulo, http://www.ansp.br/

[5] RNP - Rede Nacional de Ensino e Pesquisa, http://www.rnp.br/

[6] Paul Zikopoulos, Chris Eaton, Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data. ISBN:0071790535 McGraw-Hill Osborne Media, 2011

[7] Science DMZ, https://fasterdata.es.net/science-dmz/

[8] Rao, A; Legout, A; Lim, Y; Towsley, D; Barakat, C., Dabbous, Walid; Network characteristics of video streaming traffic. ACM Digital Library, 2011

[9] Software-Defined Networking, https://www.opennetworking.org/sdn-resources/sdn-definition

[10] OpenFlow Switch Specification 1.0, https://www.opennetworking.org/sdn-resources/onf-specifications/openflow

[11] AMPATH - America's Pathways, http://www.ampath.net

[12] Brocade Per-VLAN Rapid Spanning Tree, http://www.brocade.com/documentation

[13] Internet2, http://www.internet2.edu

[14] ESNET, http://www.es.net

[15] IEEE Standard, 802.1ad-2005, ISBN 0-7381-4874-1,2006

[16] RedClara, http://www.redclara.net

[17] FlowSpace Firewall: http://globalnoc.iu.edu/sdn/fsfw.html

[18] OESS, http://globalnoc.iu.edu/sdn/oess.html

[19] OSCARS, http://www.es.net/services/oscars/

[20] SouthernLight, http://wiki.glif.is/index.php/SouthernLight

[21] MANLAN, http://wiki.glif.is/index.php/MAN_LAN

[22] Internet2 AL2S, http://www.internet2.edu/products-services

[23] Guy Roberts, Tomohiro Kudoh, Inder Monga, Jerry Sobieski, John MacAuley, Chin Guok, NSI Connection Service v2.0, OpenGridForum GFD.212

[24] Network slice with Flowvisor, http://onlab.us/flowvisor.html

[25] Ronald van der Pol, D1.2 OpenFlow. Availeble on http://www.surf.nl/binaries/content/assets/surf/en/2013/RoN-2011-D1.2.pdf

[26] OpenFlow Switch Specification 1.3, https://www.opennetworking.org/sdn-resources/onf-specifications/openflow

[27] AtlanticWave, http://www.atlanticwave.net

[28] SOX - Southern Crossroads, http://www.sox.net

[29] MAX GigaPoP, http://www.maxgigapop.net

[30] A. Gupta, L.Vanbever, M. Shahbaz, S. Donovan, B. Schlinker, N. Feamster, J. Rexford, S. Shenker, R. Clark, E. Katz-Bassett. "SDX: A Software Defined Internet Exchange.", (2013).