Enabling autonomic provisioning in SDN cloud networks with NFV service chaining

Robert Cannistra [1], Benjamin Carle[1], Matt Johnson[1], Junaid Kapadia[1], Zach Meath[1], Mary Miller[1], Devin Young[1], Casimer DeCusatis[2], Todd Bundy[3] , Gil Zussman[4], Keren Bergman[4], Aparicio Carranza[5], Carolyn Sher-DeCusatis[5], Andrew Pletch[6], Raymond Ransom[6]

1- Marist College, Poughkeepsie, NY (robert.cannistra@marist.edu)
2- IBM Corporation, Poughkeepsie, NY (decusat@us.ibm.com)
3- Adva Optical Networks, New Fairfield, CT (tbundy@advaoptical.net)
4-Columbia University, New York, NY (bergmen@ee.columbia.edu)
5-New York City College of Technology, City University of New York, Brooklyn, NY (csher-decusatis@citytecy.cuny.edu)
6-State University of New York, New Paltz, NY (ransomer@newpaltz.edu)

Abstract: Experimental results are presented from an SDN/NFV testbed with automated, dynamically provisioned, 125 km optical WAN. Live VM migration for NFV video serving is demonstrated, along with Layer 0-3 orchestration using Open Daylight, OpenFlow, and DOVE.

1. Introduction

For many years, servers have taken advantage of abstractions which facilitate programmability, such as operating systems, programming languages, file systems, and virtual memory. As a result, modern servers have become highly virtualized, capable of supporting tens to hundreds of virtual machines (VMs) per physical server, each of which can be dynamically created, moved to a different host, modified, or deleted in a few minutes. The many benefits of server virtualization are well established, including significant reductions in capital and operating expense, reduction in the physical footprint of devices, lower energy consumption, faster provisioning times, and higher utilization [1]. Further, this approach has enabled new applications and revenue streams which would not have been technically possible or economically feasible otherwise.

Unlike servers, however, data networking (as used in both conventional enterprise data centers as well as cloud computing or service provider applications) has not yet evolved its own set of fundamental abstractions. Until recently, these abstractions were not required since many workloads were static and based on single application servers. Conventional data networks can require 5 days or more to provision the necessary service chains within a data center, and weeks or longer to re-provision service between data centers [2]   Current industry trends, including dynamic workloads, mobile computing, multi-tenant cloud computing, warehouse scale data centers, and big data analytics, have led to a need for much richer network functionality [3]. The industry has recently recognized that creating abstractions for data networking enables benefits similar to those derived from server virtualization. This concept is known as software-defined networking (SDN), and represents a significant disruption to existing network architectures [4-6]. While there are many interpretations of SDN, generally it includes separation of the data and management/control planes (with the management and control functions moving to software running on a server cluster, known as a network controller); a shift towards centralized management (preferably through cloud middleware such as the OpenStack Quantum interface); and virtualizing the Layer 2-3 network attributes through an overlay network with waypoints of network function virtualization (NFV). Making network resources programmable implies that they can be automated and eventually optimized, resulting in a workload or application aware network. SDN solutions are preferably based on industry standards or supported open source software, to facilitate interoperability among multiple networking platforms and promote sustainable innovation [4].

2. Software-Defined Network Test Bed

In this paper, we present experimental results from an SDN/NFV test bed constructed as part of the New York State Center for Cloud Computing and Analytics SDN Innovation Lab. Established in 2013, this center is a consortium based at Marist College (a member of Internet2), which currently includes several academic partners (Columbia University, City University of New York, and State University of New York) as well as corporate partners (IBM, ADVA, NEC, and BigSwitch). The goals of this test bed include demonstrating practical use cases for SDN/NFV network abstractions, promoting standards-based, open source development communities, and developing new academic curricula for networking professionals. As shown in figure 1, the test bed currently consists of three data centers connected by a 125 km ring of single-mode optical fiber. Each data centers is

connected to the optical ring using ADVA FSP3000 wavelength division multiplexing (WDM) equipment, running at 10 Gbps/wavelength. Within each data center, a combination of IBM PureSystems, System Z enterprise servers, and various IBM storage devices are connected by a 10 Gbps Ethernet network of IBM G8264 OpenFlow switches. All of the data center switches and WDM equipment (Layer 0-3) is currently orchestrated through an open source Floodlight controller running OpenFlow 1.0.

Figure 1 – SDN/NFV Dynamic Infrastructure Test Bed using 125 km Fiber Network
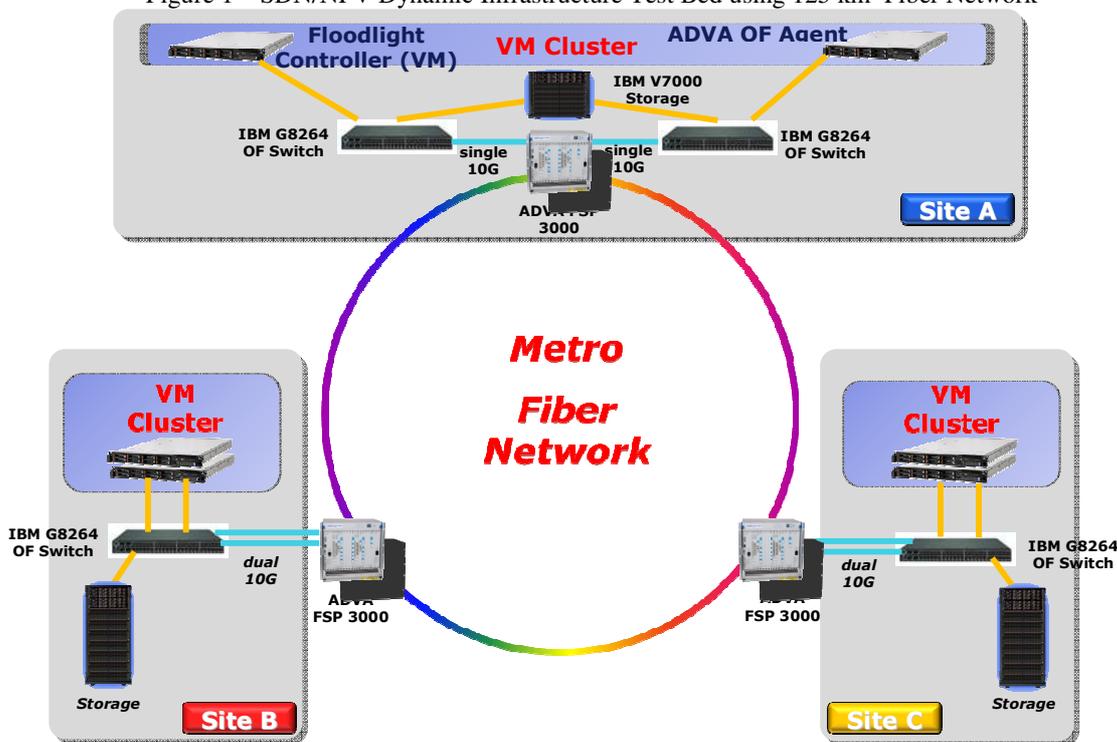
We have written several applications for this environment, including an open source management graphical user interface called Avior, which controls the OpenFlow devices from a mobile device such as a tablet or smart phone. Avior allows us to monitor network statistics, configure traffic flows, and administer firewall properties, without requiring the user to code scripts. It incorporates a static flow pusher which can deploy pre-configured network profiles or schedule network configuration events for touch-free provisioning. A second application called Advalanche is called by Avior to orchestrate the WDM equipment. This is a pre-standard implementation, since OpenFlow does not yet support physical layer orchestration (we intend to propose this for a future OpenFlow release). We use the open source application Ganglia to monitor events such as server utilization or available memory; when a preset threshold is exceeded, Ganglia executes an action such as provisioning a network profile, migrating a VM, or cloning a VM.

We demonstrate automatically triggering live VM migration (uninterrupted operation of functions in the VM) when server utilization exceeds 75%. When over-utilization of the server occurs (possibly resulting from many NFV applications running in a highly virtualized environment), Ganglia triggers a VM migration. The SDN controller automatically provisions all the switches in the source and target data center, as well as wavelengths on the optical network (subject to available physical resources and workload priority levels). End-to-end network provisioning can be accomplished in about a minute. VMWare VSphere 5.0 is then used to live migrate the VM from a server in the source data center to another server in the target data center. The VM is running a video streaming application, which continues uninterrupted service during the migration; in principle, any NFV application could be used. Once migration is complete, the SDN controller reverts all network devices to their original states.

Fixed wavelengths in most metropolitan area optical networks are currently under-utilized and high cost, since they must be provisioned for peak network capacity. This approach is not well suited to dynamic workloads such as cloud

bursting (an Infrastructure-as-a-Service (IaaS) solution which transfers excess workload from private enterprise data centers into the public cloud). During the initial VM migration of a cloud burst, network bandwidth of 10 Gbit/s or more will be required; however, for most of the IaaS lifecycle, bandwidth requirements are much lower, often less than 1 Gbit/s. Other workloads which benefit from more agile network provisioning include backup and disaster recovery, multi-site load balancing, or bandwidth calendaring. Colorless, directionless, agile core optical networks have been enabled by component technologies such as wavelength tunable optics, gridless reconfigurable optical add/drop multiplexers (ROADMs), and hybrid optical amplifiers [4]. However, the management information required to intelligently reconfigure the wide area network (WAN) has not been readily accessible, and there has been a lack of industry standard programming interfaces (APIs) for switches and optical equipment. One of the goals of this test bed was to overcome this limitation as a first step towards enabling dynamic network infrastructure, capable of re-provisioning on the same timescale as virtual server provisioning (minutes to tens of minutes).

The VM migration time is a function of the VM memory size, M, page dirty rate per second, W, and network bandwidth in pages/second, R. Applications such as VMWare perform live migration (uninterrupted operation of functions in the VM) using a variant of stop and copy technique with a pre-migration phase (periodically suspending the VM for a stop time, S,, and iteratively transferring its active memory page contents, as well as its execution state, and architectural configuration to another physical host where it is re-installed). The migration time, T, is given by

$$T = M / (R - W(T-S)/T) \qquad (1)$$

We present simulation results of migration time across the optical network test bed under various conditions, which show that migration time can vary nonlinearly for certain workloads. We have begun working on transfer of larger data sets in collaboration with CERN research labs, and migration of other NFV applications [5].

In practice, reconfiguration of optical WANs requires mitigating the impacts of the impairments caused by adding and dropping wavelengths. In [7] we present network-wide optimization algorithms that deal with these impairments by leveraging OPM measurements to dynamically control the wavelengths' power levels, thereby allowing dynamic optical networking. The algorithms proposed in [7] were evaluated in a small-scale test bed and via simulation. We plan to evaluate the performance of the algorithms in SDN/NFV test bed and to obtain insights that would allow extending the algorithms and to include various control parameters. We also discuss a mobile management application for Android smart phones using messaging queuing middleware based on AMQP protocols. Context appropriate content of SDN logs is delivered to a mobile app using RabbitMQ 3.0. Finally, we review efforts to develop an SDN education curriculum, including service education for the International Society of Service Industry Professionals (ISSIP), OpenFlow functions from the GENI network, and hands-on student labs using OpenFlow and the IETF standard Distributed Overlay Virtual Ethernet or DOVE) with NetFPGA adapters.

3. References

[1] C. Isci, J. Liu, B. Abal, J.O. Kephard, and J. Kouloheris, "Improving server utilization using fast virtual machine migration", IBM Journal of Research and Development vol 55 no 6 paper 4 (November/December 2011)

[2] J. Manville, "The power of a programmable cloud", OFC 2012 annual meeting, Anaheim, CA, paper OM2D.2 (March 18-22, 2013)

[3] C.J.S. DeCusatis, A. Carranza, and C. DeCusatis, "Communication within clouds: open standards and proprietary protocols for data center networking", IEEE Communications Magazine, October 2012

[4] C. DeCusatis, I. Marty, R. Cannistra, T. Bundy, C.J. Sher-DeCusatis, "Software defined networking test bed for dynamic telco environments", Proc. SDN & OpenFlow World Congress, Frankfurt, Germany (October 22-24, 2013).

[5] C. Dixon, D. Olshefski, V. Jain, C. DeCusatis, W. Felter, J. Carter, M. Banikazemi, V. Mann, J. Tracey, R. Recio, J. Kidambi, A. Biswas, and U. Shankar Nagaraj, "Software defined networking to support the software defined environment", IBM Journal of Research & Development, to be published (January 2014)

[6] M. Haley, C. DeCusatis, T. Bundy, R. Cannistra, R. Wallner, J. Parraga, R. Flaherty, "Dynamic, software defined service provider network infrastructure and cloud drivers for SDN adoption", Proc. IEEE International Conference on Communications (ICC), June 9-13, Budapest, Hungary (2013)

[7]] B. Birand, H. Wang, K. Bergman, D. Kilper, T. Nandagopal, G. Zussman, "Real-time power control for dynamic optical networks - Algorithms and experimentation," Proc. 21st IEEE International Conference on Network Protocols (ICNP'13) (Oct 2013)