Presentation Title: Software-Defined Infrastructure at RNP

Presenter: Marcos Schwarz (RNP)

Authors names and affiliations: Michael Pietro Hernandez (RNP), Samuel Kopp

(TVoD/BrainyIT), Marcos Schwarz (RNP), José Rezende (RNP)

Presentation description: In this talk, we will describe the software-defined infrastructure and the corresponding orchestrator that is being deployed at RNP, the Brazilian NREN.

This infrastructure is composed by a completely virtualizable SDN Overlay Network and a two-tier private cloud. The orchestrator is based on ONOS and it is responsible for the lifecycle management of virtual resources such as SDN/L2 software switches, NSI/VXLANs/VLANs circuits, VMs/containers, and storage blocks. The presentation will contain an overview of the architecture and its main features, along with the description of two use cases: dynamic provisioned CDN on SDN and containers; and Softwarization of Science DMZ Infrastructure.

Abstract: The SDI project (Software-Defined Infrastructure) at RNP aims at providing an orchestration solution for computing and communication resources that will allow the offering of the next generation services to RNP clients, demonstrating the potentialities of integrating SDN, NFV and cloud computing paradigms for NRENs. This solution will enable the creation, configuration and monitoring of virtual instances of network, processing and storage resources in an agile and fully automated way. For this to be possible, it was necessary to deploy a fully virtualizable and programmable physical infrastructure. In this talk, we will detail the architecture used in this infrastructure, hereinafter referred to as SDI@RNP (Software Defined Infrastructure of RNP), as well as the software components used to integrate its various components.

SDI@RNP is composed of high performance physical servers dedicated and interconnected by a fully programmable and sliceable network, as well as computing resources of the RNP's private cloud. In this 2-tiers cloud architecture, level 1 nodes represent RNP's private cloud capabilities, which provided services are scalable and highly reliable, but do not have the granularity of localization required by latency sensitive and high-performance applications such as CDN and Science DMZ. The second level, i.e. the distributed local cloud, consists of nodes that are connected at the edge of the network, that is, in the PoPs (Point of Presence) of the RNP backbone, providing granularity at the level of large urban centers in Brazil. As previously stated, this second level of the system is connected directly to an overlay network composed of whitebox switches, which are servers in the x386 architecture with a multi-port Ethernet card supporting DPDK and 10Gbps transceivers. These whiteboxes are also installed in the PoPs and virtually interconnected through the RNP backbone, thus forming an overlay SDN network, referred to hereafter as SDNOverlay. This overlay network enables multiple concurrent instances of SDN applications with different characteristics. The physical servers that compose the Level 2 of the cloud architecture are physically connected to the whitebox switches, thus forming a distributed edge cloud interconnected by the SDNOverlay network.

In the current development stage of SDNOverlay, ten SDN whiteboxes were installed in different PoPs and connected to the core routers of the RNP backbone. In parallel, the SOLO orchestration system was developed, based on the ONOS controller, which allows the dynamic instantiation of isolated and programmable virtual networks, as well as computational resources in the form of dockers.

The current phase of the project consists in the expansion of the SDN network with three new whiteboxes, and the addition of thirteen high performance physical servers with storage capacity, 10G network connections. These new acquisitions extends the SDNOverlay network to thirteen PoPs and build a distributed edge cloud. The SOLO system was extended to allow the instantiation of containers base services through Kubernetes, which is an open software system for the orchestration of containers. This infrastructure in conjunction with the orchestration solution will enable the deployment of advanced networking services, such as SDN based CDN and Science DMZ on demand.

The first use-case allows a CDN service that can dynamically provision content reflectors on nearest RNP’s Point of Presence based on the users prefixes. Using a micro-service oriented CDN system integrated to a SDN topology, the CDN service can discover the closest Point of Presence to the user based on BGP announcements, and dynamically provision containerized content reflectors on the it. The content reflector control plane uses SDN to automate the network configuration and provision virtual circuits with guaranteed bandwidth. This approach provide dynamic allocation of production environment infrastructure in a programatically maner defined by users demand. And enables the creation of controlled staging environments for load tests, using the instantiation of parallel virtualized SDN networks, content reflectors and consumers.

The second use-case is composed by the encapsulation and isolation of physical resources required for the deterministic behaviour of each component of a Science DMZ and bundle it within containers. These containers are composed of applications for each component (Open vSwitch for SDN switch, GridFTP for data transfer and perfSONAR for network measurement) and resource requirements (Compute: CPU, RAM; Storage: high-density/low-speed and low-density/high-speed volumes; Network: Bandwidth). To provide resource allocation we are leveraging the Kubernetes and for high performance networking in containers, Intel Multus network plugin, SR-IOV and DPDK.

Another advantage of this approach is the ability to deploy a single instance of DTN using the whole resources of the bare-metal server or whenever necessary deploy multiple instances of lower capacity DTNs in the same infrastructure, providing multi-tenancy, facilitating resource sharing of expensive physical resources and reducing idle time.