

# **LHCONE NETWORK SERVICES: GETTING SDN TO DEV-OPS IN ATLAS**

Shawn McKee/Univ. of Michigan

LHCONE/LHCOPN Meeting, Taipei, Taiwan

March 14th, 2016



# Context for this Presentation

- Within LHCONE we have had a point-to-point service effort for quite a while.
  - It has been challenging to make progress beyond a few limited demonstrations
- Within the LHC experiments there has been interest in what might be possible with networking and especially with how a future production quality software defined networking capability would fit with the way the experiments manage, operate and orchestrate their globally distributed resources
  - Network device support for SDN has not really been “production quality”...hard to interest the experiments in even testing because of problems getting anything enabled between sites of interests
- How best to make some progress?



# Challenges Getting SDN into LHC Production Systems

- While we have dabbled as a community for years with various SDN capabilities we have never managed to effectively bridge the gap into the core LHC experiment middleware and workflow systems. *Why?*
  - The experiments have their own “stove-pipes” of effort and there hasn’t been much interaction with networking
  - The experiments focused on what they perceive as bigger problems they must face
    - We have helped ensure the network has been the most reliable and capable component of their distributed infrastructure
  - Our test implementations are typically one-offs designed to demonstrate features and capabilities but not then easily translated into use with existing production systems.
  - SDN itself (both software and hardware) has not been near “production quality” to-date.
    - This is improving rapidly, new hardware/chipsets are much more capable, problems with software usability improving.



# Getting SDN to the Ends using Dev-Ops

- To make progress with SDN capabilities for LHC we need to start focusing on enabling new SDN features in production instances, blending production and development.
  - Software and technology development has called doing this “dev-ops” (Development and Operations)
- One shortcoming in the P2P effort to-date has been the significant challenge in getting all the way to the ends: to the servers that source and sink our data.
  - We have been able to create WAN circuits but it then gets “messy” for how those are actually used to carry the right traffic for production activities
- We now have an interesting option to help us: **Open vSwitch** ([openvswitch.org](http://openvswitch.org)).
  - This is well tested, supported software to create virtual switches on Linux (and other OSs) with traffic control and shaping and **OpenFlow** and **OVSDB** support.



# Details on Deploying Open vSwitch (OVS)

- There is a web page on the Wiki below documenting both the creation of RPMS for Redhat/CentOS/SL 6.x and their deployment onto existing hosts:
  - <https://www.aglt2.org/wiki/bin/view/AGLT2/InstallOpenvSwitch>
  - This web site will soon provide some detailed and tested configuration examples for implementing OVS on hosts with various types of network configuration (bonded, VLANs, multiple interfaces, etc)
- **The idea is to move your systems IP addresses off from their existing physical (or virtual OS) NICs and onto the OVS bridge you will bring up.**
- ***OVS can be installed and turned on without any impact to the running system (install RPM, activate service)***
  - It is actually moving the IP that is potentially disruptive and must be done with some care. The URL above has details.

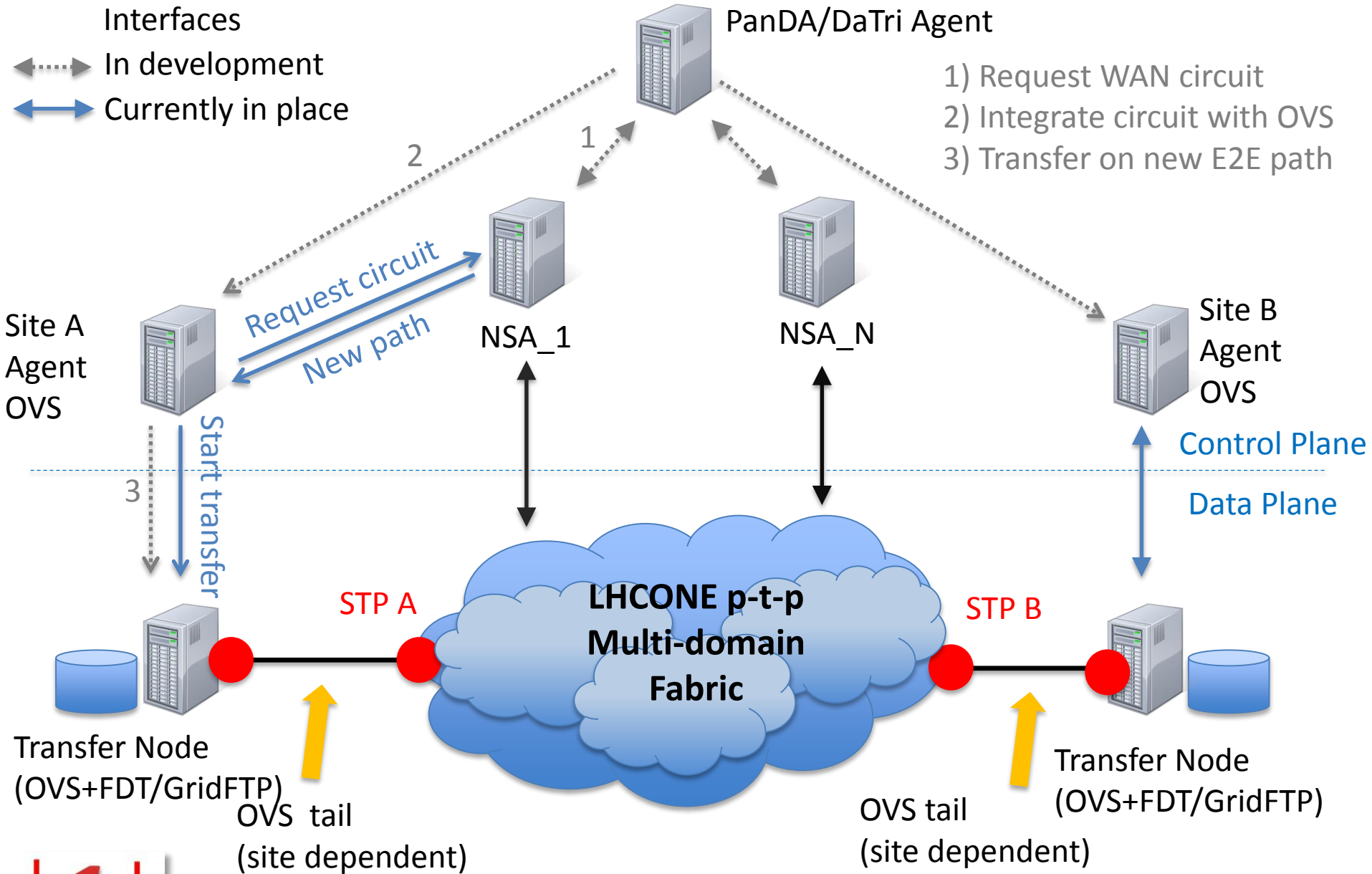


# Advantages of OVS on Production Instances

- By getting OVS in place on LHC production storage systems we immediately gain visibility and control all the way to the sources and sinks of data-flows for LHC
- We have verified that OVS has almost no measureable impact when shaping traffic on 10G NICs (See Ramiro's presentation at the last LHONE meeting:  
[https://indico.cern.ch/event/401680/contribution/16/attachments/178611/1705261/LHCONE-AM\\_SDN\\_OVS\\_rv1.pdf](https://indico.cern.ch/event/401680/contribution/16/attachments/178611/1705261/LHCONE-AM_SDN_OVS_rv1.pdf))
- Having OVS running on production systems with the IPs moved to the OVS bridge allows us to continue to operate all production services identically to how they were operated prior to installation and configuration
  - The big win is that we can start to do simple tests incorporating specific flows or sets of servers into end-to-end circuits.
  - Gradually, we can verify the impact of using such capabilities with LHC production systems and, if positive, it makes a strong argument for other sites to begin joining the effort.



# Diagram of Possible Future SDN Dev-Ops Testbed



# Challenges

- While having OVS “at the ends” will be a huge step forward for our Point-to-Point work, there remain a number of challenges
- The primary challenge is integrating existing circuit creation systems with OVS as a participant
  - How can we incorporate the OVS-enabled end-systems seamlessly into the end-to-end circuit?
- How best to use the many OVS features to improve the overall performance of the circuit?
- The main “meta-question”: ***How can SDN capabilities improve the LHC experiments ability to manage, utilize and optimize their global infrastructure?***
  - There is a lot of work to do to investigate this: Getting SDN “in-line” with production LHC work is our first step!



# Next Steps

- Finalize testing of OVS Configuration to support various network configurations
- **AGLT2** (Michigan, Michigan State) and **MWT2** (Illinois, Indiana and University of Chicago) have agreed to deploy OVS onto their ATLAS dCache storage systems
  - Total of 8.7 Petabytes of storage between the two
  - Most system dual 10G connected; sites 80 Gbits to WAN
  - This will provide an example to experiment with SDN end-to-end using real ATLAS production traffic
- We want to expand as soon as is feasible. Interest from
  - DE-KIT
  - Possible Canadian participation
  - Seeking additional sites with real use cases (at least one more in North America and in Europe)
- **Timescale** April-May 2016 for initial tests (assumes documentation and initial OVS configurations documented and tested by end of March)
- **Email Shawn McKee if your site is interested in participating**

# QUESTIONS & COMMENTS

Shawn McKee [smckee@umich.edu](mailto:smckee@umich.edu)

