Software Defined Networking Disruptive Technologies

Renato Recio
IBM Fellow & System Networking CTO
Agenda

- Software Defined Networking (SDN) Defined
- SDN Client Value
- SDN is a Discontinuous & Disruptive Technology
- SDN Adoption Curve Status
  - Current Products
  - Technology Investment Areas
- Summary
**SDN Platform – SDN Controller**

- Automates connectivity of network services (e.g. Firewall, IPS) used in multi-tier virtual systems, with multi-tenant network capability
- Optimizes traffic performance, availability and separation through fabric pathing services, with global network visibility and control
- Open APIs enable network applications

**Network Hypervisor – DOVE Network**

- Virtualizes the physical network thru a Network Hypervisor that enables a “wire once” physical network, analogous to Hypervisor for compute/IO

**Optimized Fabric – Ethernet & OpenFlow**

- Leverages OpenFlow to move network OS from physical switches to server based controller cluster, enabling: rapid protocol development time, workload aware network optimization, faster convergence times and global control
## SDN value to IT Constituents

SDN has value across Data Center constituencies; overall its disruptive potential is akin to server virtualization.

<table>
<thead>
<tr>
<th>Constituency</th>
<th>Today’s concerns</th>
<th>The benefit of SDN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CIO</strong></td>
<td>Computing models such as cloud and compute virtualization limited by human</td>
<td>Network becomes a “virtualized” asset with automated linkages to computing and</td>
</tr>
<tr>
<td></td>
<td>“middleware” needed to instantiate</td>
<td>applications.</td>
</tr>
<tr>
<td><strong>Line of Business Owner</strong></td>
<td>Exploiting analytics within enterprise requires IT agility. Rapid connectivity to</td>
<td>Network can be rapidly reconfigured and modified. Virtualized DMZ concept can</td>
</tr>
<tr>
<td></td>
<td>new sources of data and resources across-Departments, BU’s, Enterprises.</td>
<td>eliminate complicated security barriers and limitations.</td>
</tr>
<tr>
<td><strong>Application developer</strong></td>
<td>Multi-tier compose-able applications require complex interaction between</td>
<td>New connectivity service provides abstracted connectivity model, without tight</td>
</tr>
<tr>
<td></td>
<td>physically distributed systems and resources. Security, quality-of-service, etc</td>
<td>linkage to physical network configuration</td>
</tr>
<tr>
<td></td>
<td>need to be enforced.</td>
<td></td>
</tr>
<tr>
<td><strong>Systems Manager</strong></td>
<td>Provisioning, configuring, monitoring across server, storage, network is very</td>
<td>Simplified, virtual network model makes it much easier to integrate with data</td>
</tr>
<tr>
<td></td>
<td>complicated</td>
<td>center wide management systems</td>
</tr>
<tr>
<td><strong>Networking Manager</strong></td>
<td>Inability to evolve network rapidly enough to support changing workloads</td>
<td>Wire-once model limits need for physical network modification</td>
</tr>
<tr>
<td><strong>CTO architect/strategist</strong></td>
<td>Long cycle to deploy new, standard (IETF, IEEE) based networking functions</td>
<td>Rapid development cycle, leveraging OpenFlow’s control plane separation</td>
</tr>
</tbody>
</table>
SDN Value to Data Centers

Example deployment models

Integrated Network Software
- Provide dramatic improvement in business efficiency by reducing application deployment times

Automated Network Virtualization
- Provide business agility by making the infrastructure (network connectivity) completely dynamic

Optimized Fabric
- Provide finer control of network traffic flow, enabling higher fabric utilization

Integrated System (Pure)

Data Center

SDN Platform
Discontinuous Technologies

**Discontinuity** – a: the property of being not mathematically continuous; b: an instance of being not mathematically continuous; especially a value of an independent variable at which a function is not continuous
Current Discontinuous Technologies Examples

Discontinuous Technology → Impacted Technology

Flash & SSD → Magnetic Disk
CEE & FCoE → Fibre Channel
SDN* Overlays & OpenFlow → Traditional switching

* One can argue the InfiniBand Subnet Manager was an early SDN controller example.
Sustaining vs Disruptive Technologies

**Sustaining** – doesn’t affect existing markets, evolves existing one’s with better value.

See: Clayton M. Christensen, The Innovator's Dilemma
Sustaining vs Disruptive Technologies

**Disruptive** – innovation that creates a new market or displaces existing technologies in a market.
Current Disruptive Technologies Examples

PC & Microprocessor → Server markets
Handheld devices (e.g. iPad) → Notebook PCs
Overlays & OpenFlow → Traditional switching
CEE & FCoE → Fibre Channel
IBM SDN Deployments in Enterprise Client Test & Development

2 service providers

SDN Value:
• “Programmable hosting network”
• “Global visibility and control”

SDN (Overlay & OpenFlow) Client Examples

IBM SDN OpenFlow clients

- For Global Network Services network between data centers
  SDN Value: “Better network visibility & control”

- Selerity
  Provider of ultra-low latency real-time financial information
  SDN Value: “Policy driven content distribution & automated network configuration”

- Teravela
  Provider of distributed data fabric for global trading & risk analysis
  SDN Value: “Predictable network performance & rapid convergence”

Clients Evaluating World-wide Enterprise & Service Provider Data Centers
IBM Systems Networking SDN products and technology investments

Controller Platforms
- GA 2/2012 DVS 5000V Controller
- GA 10/2012 IBM Programmable Network Ctrl
- SDN Operating System
  - Network Apps
  - Open Source Based SDN Platform Investment
- DOVE
- OF

Network Virtualization
- GA 2/2012 standards-compliant layer-2 virtual switch
- DOVE Technology Investment

OpenFlow Physical Switches
- GA 11/2011 OpenFlow 1.0 10GB switch
- OpenFlow Technology Investment
Optimized OpenFlow Network

- Services run as Apps
  - Software Defined Network Stack
    - Mgt Plane: Telnet, SSH, SNMP, NTP, SYSLOG, HTTP, FTP/TFTP
    - Apps: Multipath, Security, FCF, ...

- Control Plane
  - Network topology, ACLs, Forwarding & Routing, QoS, Link Management

- Data Plane
  - Link, Switching, Forwarding, Routing

- OpenFlow Protocol
- Control plane is extracted from the network

- OS
  - Mgt Plane: Telnet, SSH, SNMP, NTP, SYSLOG, HTTP, FTP/TFTP
IBM RackSwitch G8264 OF Switch

1st OpenFlow single chip switch to pass the 1 Terabit per second barrier!

- OpenFlow-based flow handling in hardware at line rate (1.28Tbps)
- Support Layer 2 (MAC) forwarding table manipulated thru OF:
  - Layer 2 (destination MAC & destination MAC+_+VLAN) table: Max 128K flow entries
  - Layer 3 or anything other than layer-2 MAC/VLAN: Max 750 flow entries
- IBM and partner OF Controllers (PNC, BigSwitch Floodlight,...)

Specifications

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forwarding</td>
<td>• Delay less than 1us</td>
</tr>
<tr>
<td></td>
<td>• 1.28Tbps; 960Mpps</td>
</tr>
<tr>
<td>Number of ports</td>
<td>• 48 x 1 Gb/10 Gb SFP+ ports;</td>
</tr>
<tr>
<td></td>
<td>• 4 x 40 Gb QSFP+ ports</td>
</tr>
<tr>
<td></td>
<td>• Up to 64 x 1 Gb/10 Gb SFP+ ports with optional breakout cables</td>
</tr>
<tr>
<td>Model</td>
<td>• Airflow-type rear to front</td>
</tr>
<tr>
<td></td>
<td>• Airflow-type front to rear</td>
</tr>
<tr>
<td>Dimensions</td>
<td>• 17.3” wide; 19.0” deep; 1U high</td>
</tr>
<tr>
<td>Protocol version</td>
<td>• OpenFlow 1.0.0</td>
</tr>
<tr>
<td>Number of instances</td>
<td>• 1</td>
</tr>
<tr>
<td>Protocols</td>
<td>• No legacy protocols running in OpenFlow switch mode</td>
</tr>
<tr>
<td>Management</td>
<td>• Telnet, SSH, SNMP, sFlow</td>
</tr>
<tr>
<td>Redundancy</td>
<td>• Power/fan</td>
</tr>
</tbody>
</table>

4x 40 GE uplinks or 4x*10G with QSFP to SFP+ cable
IBM Programmable Network Controller

- High performance OpenFlow based controller, provides:
  - Highly reliable end-to-end fabric that works over any OpenFlow 1.0 compliant switches
    - Automatically discovers OpenFlow network topology
    - Intelligent and dynamic multipath routing based on business policy
  - Virtualizes an OpenFlow network using Virtual Tenant Network (VTN) Application, where a VTN provides a:
    - Customized layer-2 or layer-3 virtual network isolated from other virtual networks
    - Secure slice of the underlying physical network, with policy based networking
  - Automated, global end-to-end view & control of the network
    - Automated network topology discovery
    - Point and Click virtual network design
    - Network segments can be configured centrally
    - One touch point (IBM PNC) versus hundreds of touch points (Network Elements)
    - Policies are enforced throughout the network vs individual switch configuration
    - APIs to create, edit, and delete VTNs, as well as to add and remove policies
## Multi-path fabric with LAG to server

- **Pros:**
  - Scale well
  - Standard based multi-pathing protocols
- **Cons:**
  - Requires proprietary fabric to reduce managed switches
  - Long convergence times for Distance / Path vector algorithms (e.g. RIP / BGP), which iteratively send routing protocol packets to calculate routes.
  - Long time to add new functions to standard (e.g. OSPF, TRILL)

## OpenFlow based multi-pathing

- **Pros:**
  - Scales well with arbitrary topologies
  - Optimized multi-pathing with short convergence times*
  - New functions (e.g. disjoint multi-pathing can be added to SDN controller, without long-standard lead time)
  - Gated by SDN market adoption S-Curve
- **Cons:**
  - Established networking vendors will create FUD against this model

* Controller discovers switches & creates topology, when a device fails, neighbors report the event and controller loads new routing tables around the failure.
Automated connectivity of services (e.g. Firewall, IPS) used in virtual system patterns, with multi-tenant network capability

- Layer-3 DOVE switch decouples virtual networks from physical network

- Simple “configure once” physical network (vs configured per VM)
IBM contributions to Distributed Overlay Virtual Ethernet (DOVE) networking

- In early 2009, IBM Haifa research began work on overlay networking.
  - Virtual Application Networks (VANs) are a host side solution that allows hosts to construct a fully virtualized network service overlay on top of a standard IP physical network.
  - Filed patents on VAN related material

- At the Ethernet Summit Conference in February 2011, IBM introduced DOVE concept:

- At the September 2011 Data Center Converged And Virtual Ethernet Switching workshop IBM published a paper on DOVE networking:

- At the November 2011 IETF, IBM:
  - Proposed creation of an IETF workgroup to work on DOVE networking:
  - Provided a Problem Statement to begin the IETF standardization process:

- At the March 2012 IETF, IBM:

- At the July 2012 IETF approved IBM’s problem statement as an internet draft
DOVE Network Technology

Components

- **DOVE Management Console**
  - Provides GUI and APIs for management of DOVE network, groups and policies

- **DOVE Switches (DOVES)**
  - Provides layer-3 & layer-2 over UDP overlay (header format same as VXL AN)
  - Performs data and some control plane functions

- **DOVE Service Appliance**
  - Distributed Connectivity Service (DCS):
    - discovers & disseminates VM location (physical server)
    - maintains policy (e.g. allow, deny, insert service appliance) and works with DOVE switches to apply policy
  - Gateway Service: Connectivity to non-DOVE networks

- Virtual and Physical (network and storage) Appliances provide services for DOVE Network
Overlay Operation

**DOVE Header (VXLAN/OTV):**

<table>
<thead>
<tr>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>I</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>Reserved (24-bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DOVE Domain Group ID (24-bits)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reserved (8-bits)</td>
</tr>
</tbody>
</table>
Clients build DOVE Networks, which interconnect workload groups through a set of middle-boxes. A DOVE Network provides the network connectivity required to build virtualized multi-tier systems, that can be deployed multiple times with minimal manual configuration.

DOVE Networks provide an SDN Connectivity Service that enables Virtualized Systems

Instance 1 of a Virtualized System

Network connectivity for a Virtualized System

DOVE Network

Interconnected workload groups

• DOVE connects a set of workload groups thru middle-boxes and automates middle-box configuration

Workload Group

vNIC port set

• Logical grouping of workloads
• Workloads share network services

Workload

Virtual NIC port

• Layer-2 address (00:23:45:67:00:23)
• Layer-3 address (129.2.200.5)
• Port QoS attributes (e.g. # Gbps)

Services Middle-box

Interconnects Workloads

• Network service provider (e.g. Firewall, IPS, ADC)
• Virtual or physical
When combined with Virtual App Patterns:
the network connectivity service interconnects and configure the middle-boxes; and
the workload deployment manager (e.g. PSM) configures the pattern’s workloads.

Instance N of a Virtual App Pattern Instance

Instance 1 of a Virtual App Pattern Instance

Network connectivity for a Virtual App Pattern
Interconnected & configured workload groups

- DOVE connects & configures network
- Workload deployment manager configures workloads

Virtual NIC port
- Layer-2 address (00:23:45:67:00:23)
- Layer-3 address (129.2.200.5)
- Port QoS attributes (e.g. # Gbps)

Workload Group
- Logical grouping of workloads
- Workloads share network services

Services Middle-box
- Network service provider (e.g. Firewall, IPS, ADC)
- Virtual or physical

Interconnects Workloads
Layer-2 automated
+ Access switches are network aware, using IEEE 802.1Qbg standard.

Complex, limited & costly appliances
- Configuration complexity (requires per workload instance configuration of physical: NAT controllers, security appliances, workload balancers, …)
- Only within the boundary of security appliance are protected (migration across boundaries requires physical configuration changes)
- Expensive appliances; no open eco-system

Full Virtual Network Automation
+ Simple configuration based on DOVE network pattern
+ VMs are free to move around DC, VM’s network service attributes are not tied to physical location
+ Low cost appliances running on open SDN eco-system
- Gated by SDN market adoption S-Curve
Multi-tenant, Optimized Virtual Network

Today’s Network

Sub-optimal traffic flow, limited scale
- Traffic is not optimized across groups (cross-subnet VM-VM traffic must go North-South), even for virtual network services
- Limited scaling (lacks multi-tenancy)

Software Defined Networking

Optimized traffic flow, multi-tenant scale
+ Optimizes traffic within and across groups (cross-subnet VM-VM traffic stays in server)
+ Multi-tenant scaling for: cloud service providers; clients that consolidate infrastructure after a merger; …
Software Defined Networking Technology Value

- **Network Services value:**
  - Eco-system for network Apps vs today’s closed switch model
  - *Connectivity Service for DOVE automates layer-3 and above*
  - *Multi-pathing Service for OpenFlow lowers convergence time and time to new function*

- **DOVE Network value:**
  - *Automated network resource (layer-2 and above) provisioning*
  - De-couples virtual network from physical network
  - Simple “configure once” network (physical network doesn’t have to be configured per VM).
  - Cloud scale (e.g. multi-tenant)

- **OpenFlow value:**
  - *Global physical network control & visibility*
  - De-couples control plane from data plane
Thank You

Renato J Recio
IBM Fellow & Systems Networking CTO
11400 Burnett Road
Austin, TX 78758
512 973 2217
recio us ibm com
QUESTIONS