T-CAP (Converged Appliance Platform)

2016. 6
Converged N/W Appliance (T-CAP)

Data Center Networking is changing.
→ New Architecture for Virtualization, Big Storage, Overlay N/W, ...

- Computing & Storage & Networking are in x86
- Performance Issue (Network Storage connection, Virtual Switch, ...)
- Traffic Flow Issue (in-out Traffic congestion, ...)
- High Speed Connection Needs Increase
- High Speed/Availability/Deployment Needs From Hyper-scale DC
- N/W Application Adoption in L2/L3 node (Linux OS in Switch)
- Performance Issue (Low CPU, Low Bandwidth between CP & DP)
- Optimized for Service

Hyper Converged Infrastructure,
Service-Optimized Whitebox Switch
High Performance N/W Appliance
Converged N/W Appliance (T-CAP)

Open Hardware Architecture for Software defined & Service optimized Infra

**Hyper-Converged System**
- Nutanix
- Vmware with Supermicro
- Cisco
- Dell / EMC

**Service-Optimized Whitebox Switch**
- Facebook
- Google
- Linkedin

**High Performance N/W Appliance**
- Znyx, Pluribus, ...

Trends
Converged N/W Appliance (T-CAP)

Converged N/W Appliance, with High Performance Server & Data Center Switch, as a Flexible & Cost-effective NFV Solution & Service-optimized Hybrid N/W Solution

- Provides L2/L3 Functions & N/W Application Services
  - L4, Security, Network Analytics, etc.
  - extensible based on 3rd Party Applications

- Supports multi-OS & Virtualization Environment
  - Linux KVM, Vmware ESXi for x86 Server Environment
Converged N/W Appliance (T-CAP)

Use Cases

Integrated Network Service Appliance

- TOR Switch
- Load Balancer
- Firewall, VPN
- Data Backup...
- Server Farm

Multiple N/W Service appliances integrated into single ToR N/W Appliance

Network Analysis

Performs Inline Analytics for the traffic through ToR switch
Converged N/W Appliance (T-CAP)

Appliance H/W Architecture Requirements

- High Speed Xeon CPUs
- Multi Host Option
- PCIe Slots for Options (Flash Accelerator, HBA for Storage, etc)
- Internal Storage Option

- High Bandwidth Between Compute node & Switch node
- Data Center grade Switching Capacity
- Dedicated CPP

- 10G/40G Interface
- 25/50G Standard supported
- Modular Architecture
Converged N/W Appliance (T-CAP)

Dual Xeon E5 Compute Node + Intel RRC based Switch Node

H/W Architecture

Compute Node

Switch Node

Application Processor

Dual Intel Xeon E5-2600 v3

Intel Atom CPP

Slot #1

Slot #2

Slot #3
Converged N/W Appliance (T-CAP)

H/W Architecture

T-CAP (RRC based)
- Intel Xeon E5-2658 v3 (Application)
- Intel RRC 10840
- Atom (CPP)
- PCIe v2
- 4X PCIe v3 (200Gbps)
- Slot #1: 10G 12port
- Slot #2: 10G 12port
- Slot #3: 10G 12port

40G Switch (Trident2 based)
- Intel Atom (CPP)
- PCIe v3 x2 (16Gbps)
- BCM 56850
- 40G Switch
- 40G 16port

N/W Appliance (Trident+ based)
- Intel Xeon E5 (Application + CPP)
- PCIe Switch
- BCM 56842
- 82599
- 82599
- 82599
- 20Gbps * 3
- 10G 16port +
H/W Specifications

- 2U / 19 inch Rack-mount type
- Dual Intel Xeon E5-2600 v3 CPU (Haswell-EP) Compute Node + Intel RRC based Switch Node
- Front Loading Switch Port Module (10G 12 Port / 10G 4 Port + 40G 2 Port)
- 4x 2.5 inch SATA HDD/SSD Front Hot-swap Bay
- 2x PCIe Gen3 Slots as rear side (Flash Accelerator, HBA, RAID card and more, supported)
- 1+1 Redundant PSU (Power Supply Unit)
- Front & Rear Panel Air hole, Internal Flow Guide for Cooling Optimization
Converged N/W Appliance (T-CAP)

H/W Components

- Dual Xeon E5 v3 Server Board
- CPP Board (Switch Control)
- Switch Board
- Switch Port Module

- PCIe v3.0 Slot (2EA)
- PSU(1+1)
- Cooling FAN
- 2.5 inch HDD/SSD (4EA)
Converged N/W Appliance (T-CAP)

Server Mainboard
- Intel Xeon E5-2600 v3 Processor (Dual Socket)
- Max 256GB Memory DDR4 RAM
- PCI Express Slots for Flash Accelerator, HBA Card, ...
- PCIe v3 interface (4EA) for connecting Switch Board

Switch Board
- Intel RRC N/W Controller based
- PCIe v3 interface (4EA) for connecting Server Board
- Modular type Ethernet Interface Slot (3 EA)
  - 10G & 40G supported
  - 1G/2.5G/5G/25G/50G/100G supportable
CPP (Control Plane Processor) Board for Switch

- Intel Rangeley/Avoton CPU based
- SATA-DOM for OS Boot
- 2x USB2.0, 1x External Console(RJ45 Type), 1x GbE for Management

Ethernet Module

- Front Loading type Ethernet Daughter card
  - 12 port 10G Daughter card
  - 4 port 10G & 2 port 40G Daughter card
Open Architecture based Network OS – Commercial & Open Source NOS

Converged N/W Appliance (T-CAP)

Open NOS Integration
- OpenSwitch and other
- Integration roadmap with SDN Platform

Partnership with Commercial NOS
- IPI OcNOS / Launch in July `15
- Quick deployment with advanced Features

Network OS with ONIE
SK Telecom to Integrate IP Infusion’s OcNOS™ Network Operating System Into Their Converged Network Solutions for Mobile and Customers Networks

**OcNOS solution allows for new, innovative services to be launched faster and more cost effectively**

June 07, 2016 11:00 AM Eastern Daylight Time

SANTA CLARA, Calif.--(BUSINESS WIRE)--IP Infusion, Inc., a leading provider of intelligent network software for telecom and data communications services, today announced that SK Telecom, Korea's largest telecommunications company, will use the OcNOS network operating system as part of their open networking strategy to deliver enterprise and carrier-grade solutions, which will allow them to reduce network costs, increase flexibility, and to deploy new features and services quickly. In addition to integrating OcNOS as the network operating system for their own networks, SK Telecom will provide its networking solution combined with OcNOS to networking customers.

“**We are proud to work with SK Telecom to deliver solutions that require less time and lower capital and operating expenditures.**”

SK Telecom is creating a converged network appliance (i.e. T-CAP) based on open networking architecture combining a high performance server and a data center scale switch together, as part of their strategy to build modular data center solutions for services at the edge to serve VNF and mobile computing needs. With this approach, the carrier can move service intelligence distributed and closer to the edge of the
<table>
<thead>
<tr>
<th>Requirement</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 2</td>
<td>Virtual LANs with Port-based VLANs</td>
</tr>
<tr>
<td></td>
<td>IP Subnet-based VLAN classifiers</td>
</tr>
<tr>
<td></td>
<td>Broadcast/Multicast/Unknown Unicast Storm Recovery</td>
</tr>
<tr>
<td></td>
<td>Jumbo Ethernet Frames</td>
</tr>
<tr>
<td></td>
<td>Static MAC Filtering</td>
</tr>
<tr>
<td></td>
<td>MAC based VLANs</td>
</tr>
<tr>
<td>Layer 3</td>
<td>Ethernet ARP</td>
</tr>
<tr>
<td></td>
<td>Transmission of IP Datagrams over Ethernet</td>
</tr>
<tr>
<td></td>
<td>IP Broadcast</td>
</tr>
<tr>
<td></td>
<td>IP Broadcast in the Presence of Subnets</td>
</tr>
<tr>
<td></td>
<td>IP Subnetting</td>
</tr>
<tr>
<td></td>
<td>ICMP Router Discovery Messages</td>
</tr>
<tr>
<td></td>
<td>Classless Inter-Domain Routing (CIDR); an address assignment and aggregation strategy</td>
</tr>
<tr>
<td></td>
<td>Classless Inter-Domain Routing (CIDR); an address assignment and aggregation strategy</td>
</tr>
<tr>
<td></td>
<td>Requirements for IP Version 4 Routers</td>
</tr>
<tr>
<td></td>
<td>Route Redistribution across OSPF and BGP</td>
</tr>
<tr>
<td></td>
<td>VLAN Routing</td>
</tr>
<tr>
<td></td>
<td>Dual IPv4 / IPv6 TCP/IP Stack</td>
</tr>
<tr>
<td></td>
<td>Border Gateway Protocol, Version 4</td>
</tr>
<tr>
<td></td>
<td>Capabilities Negotiation with BGP-4</td>
</tr>
<tr>
<td></td>
<td>BGP Support for Four-Octet AS Number Space</td>
</tr>
<tr>
<td></td>
<td>Open Shortest Path First Version 2</td>
</tr>
<tr>
<td></td>
<td>OSPF Enhancement for NSSA/VAS-External and Reducing CPU Loads</td>
</tr>
<tr>
<td></td>
<td>Passive Interface Support for OSPFv2</td>
</tr>
<tr>
<td></td>
<td>OSPF Not-So-Stubby-Area (NSSA) Option</td>
</tr>
<tr>
<td></td>
<td>Passive Interface Support in OSPFv3</td>
</tr>
<tr>
<td>QoS</td>
<td>DiffServ Field in IPv4/IPv6 Headers</td>
</tr>
<tr>
<td></td>
<td>Architecture for Differentiated Services</td>
</tr>
<tr>
<td></td>
<td>QoS Access Control Lists (ACLs) Permit/Deny</td>
</tr>
<tr>
<td></td>
<td>for Inbound/Outbound IP traffic based on:</td>
</tr>
<tr>
<td></td>
<td>a) Type of Service (TOS) or DiffServ (DS) DSCP field</td>
</tr>
<tr>
<td></td>
<td>b) Source IP address</td>
</tr>
<tr>
<td></td>
<td>c) Destination IP address</td>
</tr>
<tr>
<td></td>
<td>d) IP Protocol Number</td>
</tr>
<tr>
<td></td>
<td>QoS Class of Service (CoS) direct user configuration of the following:</td>
</tr>
<tr>
<td></td>
<td>a) IP DSCP to Traffic Class Mapping</td>
</tr>
<tr>
<td></td>
<td>b) Interface Traffic Shaping Rate</td>
</tr>
<tr>
<td></td>
<td>c) Minimum and Maximum Bandwidth Per Queue</td>
</tr>
<tr>
<td></td>
<td>d) WRR/WFQ/SP Scheduling Per Queue</td>
</tr>
<tr>
<td>VXLAN</td>
<td>VXLAN tunneling for IPv4</td>
</tr>
<tr>
<td>xSTP</td>
<td>Spanning Tree (STP)</td>
</tr>
<tr>
<td></td>
<td>Multiple Spanning Tree Protocol (MSTP)</td>
</tr>
<tr>
<td></td>
<td>Rapid Spanning Tree (RSTP)</td>
</tr>
<tr>
<td></td>
<td>Rapid Spanning Tree (RSTP) optimization for rings</td>
</tr>
<tr>
<td>VRRP</td>
<td>Virtual Router Redundancy Protocol (VRRP) for IPv4</td>
</tr>
<tr>
<td>OpenFlow</td>
<td>Multiple Flow table Support</td>
</tr>
<tr>
<td></td>
<td>Connecting to multiple controllers for HA</td>
</tr>
<tr>
<td></td>
<td>Hybrid Mode Support (May need to understand the level of hybrid support in the Intel)</td>
</tr>
<tr>
<td>Hardware Diagnostic</td>
<td>Over temperature protection</td>
</tr>
<tr>
<td></td>
<td>Fan control</td>
</tr>
<tr>
<td></td>
<td>Power: Current and Voltage monitor</td>
</tr>
<tr>
<td></td>
<td>CPU and Memory usage</td>
</tr>
<tr>
<td></td>
<td>LED settings</td>
</tr>
<tr>
<td>System Information</td>
<td>Board Serial Number/Device Identifier, MAC Address</td>
</tr>
<tr>
<td></td>
<td>Part numbers for CPU, FAN, PSU …</td>
</tr>
<tr>
<td></td>
<td>H/W controller information</td>
</tr>
<tr>
<td></td>
<td>BIOS info</td>
</tr>
<tr>
<td></td>
<td>OS Information</td>
</tr>
<tr>
<td>Management</td>
<td>SNMP</td>
</tr>
<tr>
<td></td>
<td>sFlow</td>
</tr>
</tbody>
</table>
Converged N/W Appliance (T-CAP)

H/W Independent

Management daemons
- CLI, Rest, Chef, ...

Monitoring daemons
- pmaclt, hsflowd, broadview...

System daemons
- sysd, intfd, vland, ...

L2/3 protocol daemons
- bgpd, lldpd, lacpd, ...

HW Support Daemons
- tempd, fand, powerd, pmd, ...

Virtual L2/3 interfaces

I²C and other drivers

Kernel

x86 Rangeley CPU
Intel RRC N/W Controller

Credit: www.openswitch.net
Converged N/W Appliance (T-CAP)

On-Demand N/W Service Platform

Management Cluster

Service Cluster

Service Network
Management Network

Dell R720 Server #3 (for VMs)
VM
UI Controller
VM
File
VM
...

Dell FX2S
Controller Server #1 (Active)
Service Controller
SDN Controller
Database

Controller Server #2 (Standby)
Service Controller
SDN Controller
Database

CNA#1 (Active)
C-Node (172.19.196.22/24)
LB
FW
VID
IDS
KVM
C-Node Agent
N-Node (172.19.196.21/24)
Inter-Connection (40G x 2)

CNA#2 (Standby)
C-Node (172.19.196.24/24)
LB
FW
VID
IDS
KVM
C-Node Agent
N-Node (172.19.196.23/24)

Nutanix #1
Nutanix #n

Switch #13-1 (1G)
Switch #13-2 (1G)
Switch #14 (1G)

BB #12-1
BB #12-2
BB #20-1
BB #20-2
WAN
Converged N/W Appliance (T-CAP)

**SDV System for advanced monitoring**

- NPB: Packet Filtering/Distribution/Copy, VLAN stripping, Time stamping
- Collector: SNMP, sFlow, NetFlow
- Flow Analyzer: s/n/pFlowGen
- NPM: TCP Performance, Link/Flow Statistics
- Forwarder: Packet Forwarding, External I/F, API

**AF-Ceph Storage Mgmt. node**

- 20Gb (10Gb x 2) Ethernet
- 4 OSD/SSD 40EA
- NVRAM for Journal
- Flash Optimized
- Usable 12.8TB

Cinder Storage Backend

Service N/W

CEPH Mgmt. node

SKT All-flash CEPH

OSD node

Storage Replication
Converged N/W Appliance (T-CAP)

- NOS option
- BMC controller support
- More core for Xeon Processor E5-2600 v3 series, 6 to 18 cores
- 100G network interface options
- NFV Applications supported
- 48V power option available
- ...
End

sohn.minho@sk.com