Software Defined Networking: the need for performance-optimized networking (under virtualization)

Speaker: Saverio Niccolini, NEC Laboratories Europe, niccolini@neclab.eu

Thanks to Felipe Huici, Mohamed Ahmed (NEC) and Luigi Rizzo (UniPI) for backing this talk up with more than slideware

Thanks to Hans-Joerg Kolbe (NEC) and Bob Briscoe (BT) for the architectural and inspiring discussions
Networking Paradigms*: the SDN Vision

Control Plane:
Software Defined Networking (SDN)

- Creates abstractions and orchestrates
- Reduces OPEX and CAPEX, Space and Power Consumption
- Challenge: Scalable, Feature-Rich Control Plane Software

Data Plane

Value added network functions (i.e., middleboxes)
Commodity network functions (e.g., L2 SW)

*inspired from BT’s vision
OpenFlow is a technology enabler for SDN
Networking Paradigms*: the SDN Vision

Control Plane: Software Defined Networking (SDN)

Challenge: Scalable, Feature-Rich Control Plane Software

Reduces OPEX and CAPEX, Space and Power Consumption

SDN is not only control plane: ever thought of the data plane?

*inspired from BT's vision
Industry Trends

"Replace HW-based NW appliances with SW on virtualized commodity PCs"¹
- BT’s proof-of-concept for deploying SW-based network functions and services at edge
  - E.g., software-BRAS co-located with caches²
- Targets
  - Total cost of ownership: [-33%; -50%]
  - Energy: [up to -50%]
- Challenge
  - “performance of virtualized network functions on high-end commodity servers”

“Move access network functions to the Data Center”
- DT recently launched the “Terastream” project
- Simplify access networks, easier to virtualize and move

“Reduce OPEX by simplifying home networks”
- FT’s proposal for virtual home gateway to BBF
- Increase flexibility in service deployment at the edge
- Supported by Telefonica, Portugal Telecom, China Telecom

Layer 123 SDN & OpenFlow World Congress: announcement of an ETSI ISG on “Network Functions Virtualization”
- AT&T, BT, DT, KDDI, NTT, FT/Orange, TI, Telefonica, Verizon

¹http://www.btplc.com/Innovation/News/NetworkVirtualization.htm
Networking Paradigms*: a more complete SDN Vision

- Control Plane: Software Defined Networking (SDN)
  - Creates abstractions and orchestrates

- Reduces OPEX and CAPEX, Space and Power Consumption (even more)

- Challenge: Fast, Feature-Rich Virtualized Data Plane on Commodity Hardware

- Network Functions Virtualization
  - Value added network functions (i.e., middleboxes)
  - Commodity network functions (e.g., L2 SW)

*inspired from BT’s vision
Network Functions Virtualization

Virtualize value added network functions on commodity PCs (that are normally implemented on specialized hardware)

| SW | | | |
| --- | --- | --- |
| Application program | Application program | Application program |
| Server mngt. software | Server mngt. software | Server mngt. software |
| NW management software | NW management software | NW management software |

Equipment / Server + software

| HW | | | |
| --- | --- | --- |
| OpenFlow controller | Router switch | Other architecture NW equipment |
| OpenFlow switch | OpenFlow switch | Legacy architecture |
| Other architecture | Other architecture | Other architecture |

High speed network functions on virtualized commodity hardware: optimized data plane under virtualization
Free your middleboxes down to the data plane with tiny, fast network VMs
Traditional Virtualization

Virtual Machines (VMs)

- Fat
- Big overhead
- Poor performance
- Few per PC

Thus: not suited for network functions
Virtual Machines (VMs)

- Small
- Small overhead
- Good performance
- Many per PC
- Natively supports network functions (i.e., middleboxes)
Requirements for network functions virtualization

- Flexible processing (for network functions)
- Good performance (for millions of packets)
- Isolation (to protect from faults)
- Flow migration / per-flow state (to support mobility)
- Support for potentially many “domains”
  - So ideally would be like “small” domains (i.e., not full OSes)

Diagram:

- Minimalistic solutions: L4, Exokernel, MINIX
- Minimalistic OSes
- Container-based: KVM, Google native client
- Full virtualization: VMware, VirtualBox
- Full-fledged solutions

© NEC Corporation 2012
ClickOS*: tiny, fast network VM

- Runs under XEN
- It is a small (and secure/isolated) domain
  - Small domain: based on MiniOS (minimalistic OS)
  - Secure: MiniOS originally intended to run drivers
- Provides natively network functions
  - MiniOS + Click = ClickOS
  - Click provides the necessary network functions (firewall, NAT, shaper, vpn, pppoe, etc.)

* http://conferences.sigcomm.org/sigcomm/2012/paper/sigcomm/p293.pdf
ClickOS: Features

- Inherits features provided by XEN
  - Domain **isolation**
  - Driver support
  - Performance (optimization needed, work in progress)

- ...but is still **small and scalable** (~1.4MB)
  - Can run many (>3000) on a single commodity PC
  - Boots and migrates **quickly** (~70 msec)

- Inherits the **processing flexibility** from Click
  - i.e., it already provides **most of the network functions needed to operate a network**
Data plane performance: where are we at today?

Average delay overheads →

Throughput

- **Before**: steady rate of \(~8,000\) pps for 1500-byte packets (96Mb/s, < 1 % of line rate)
- **Now**: 950Kp/s for max-sized pkts (> 10gb line rate, \~810Kp/s)
- **Smaller packets? Work in progress, stay tuned...**
Freeing middlebox functions: a vision

Traditional Networking | SDN with OpenFlow | SDN with ClickOS

Management / Orchestration Applications

API₁
API₂
... APIₙ

SW + Commodity HW

Equipment Software

HW

Control Plane

Data Plane

Control Plane

Data Plane

Control Plane

Data Plane

(Relative) Scalable Cloud

Distributed Scalable Cloud

Virtualiz.

Virtualiz.

Routers, Switches, Firewalls, DPI nodes, etc.

© NEC Corporation 2012
In a nutshell: why should you care?

- Unprecedented latency and throughput of virtualized network functions
- Flexibility and speed in service deployment
- Network simplification and consolidation
  - Reduction of OPEX, CAPEX, space and power consumption
(Example)
Fields of application
Use cases

Application areas where ClickOS technology can be helpful

- **Broadband Remote Access Server** (BRAS, B-RAS or BBRAS)
  - can also be referred to as a Broadband Network Gateway (BNG)
  - terminates PPPoE, routes and shapes traffic to/from Digital Subscriber Line Access Multiplexer (DSLAM) on an Internet service provider's (ISP)

- **Residential Gateway**

- Deep Packet Inspection
- Caching solutions
- Enterprise firewalling
- Mobile phones
- Etc.

All of them software-based, virtualized, easily migratable... and high-speed 😊
Use case: Software BRAS

COTS device

Tiny VMs with optimized virtualization

NASS Module (PPP)
Indiv. FW & rate limit

NASS Module (PPP)
Indiv. FW & rate limit

NASS Module (PPP)
Indiv. FW & rate limit

NASS Module (PPP)
Indiv. FW & rate limit

VPN Routing

Global Firewall

Rate Limit per Access Node

PPPoE

FW

Traffic Shaper

SW-defined NW Functions in ClickOS
Example: hierarchical shaping

Other example taken from http://www.juniper.net/techpubs/en_US/junos11.4/topics/concept/cos-subscriber-access-adjustments.html
Residential Gateway Today

Interfaces | Security | Network | MNG
---|---|---|---
Eth WAN Port | Anti Virus | IGMP | DLNA
Eth LAN Ports | Firewall | QoS | Statistics
DECT | Remote Access | PPPoX | Diagnostics
WiFi | Parental | ALG | GUI
FXO | | DHCP | TR-69
FXS | | NAT | Port map
USB

Network

TV/STB
VoIP
PC
Laptop
Virtual Residential Gateway with ClickOS

- **Interfaces**
  - Eth WAN Port
  - Eth LAN Ports
  - DECT
  - WiFi
  - FXO
  - FXS
  - USB

- **MNG**
  - TR-69
  - QoS

- **Network**
  - IGMP
  - QoS
  - Remote Access
  - Parental
  - Port map

- **Security**
  - Anti Virus
  - Firewall
  - Parental

- **MNG**
  - DLNA
  - Statistics
  - Diagnostics
  - GUI

- **Network**
  - NAT
  - Port map

- **Optimization of virtualization**
  - COTS device

- **SW-defined NW Functions in ClickOS**
  - PPPoE
  - FW
  - Traffic Shaper

ONU + Home Bridge

- TV/STB
- VoIP
- PC
- Laptop
ClickOS a technology for virtualizing network functions (thus freeing them)

- With unprecedented speed on data plane thanks to the cooperation of other technologies (e.g., netmap, VALE) (and still counting…)

We have more than slides 😊

- Running software for real use cases (e.g., BRAS)

- But still a lot of work ahead…
Acknowledgement

ClickOS, Netmap, VALE are technologies developed inside the CHANGE project.

This work has been partially funded under the contract number INFSO-ICT-257422.

Coming up: Trilogy2 (from 01/2013)

Architectural discussions on fields of applications have been conducted jointly with the OFELIA project (contract number: 258365).

Use cases testing to be conducted in the OFELIA islands.