



**EE542**

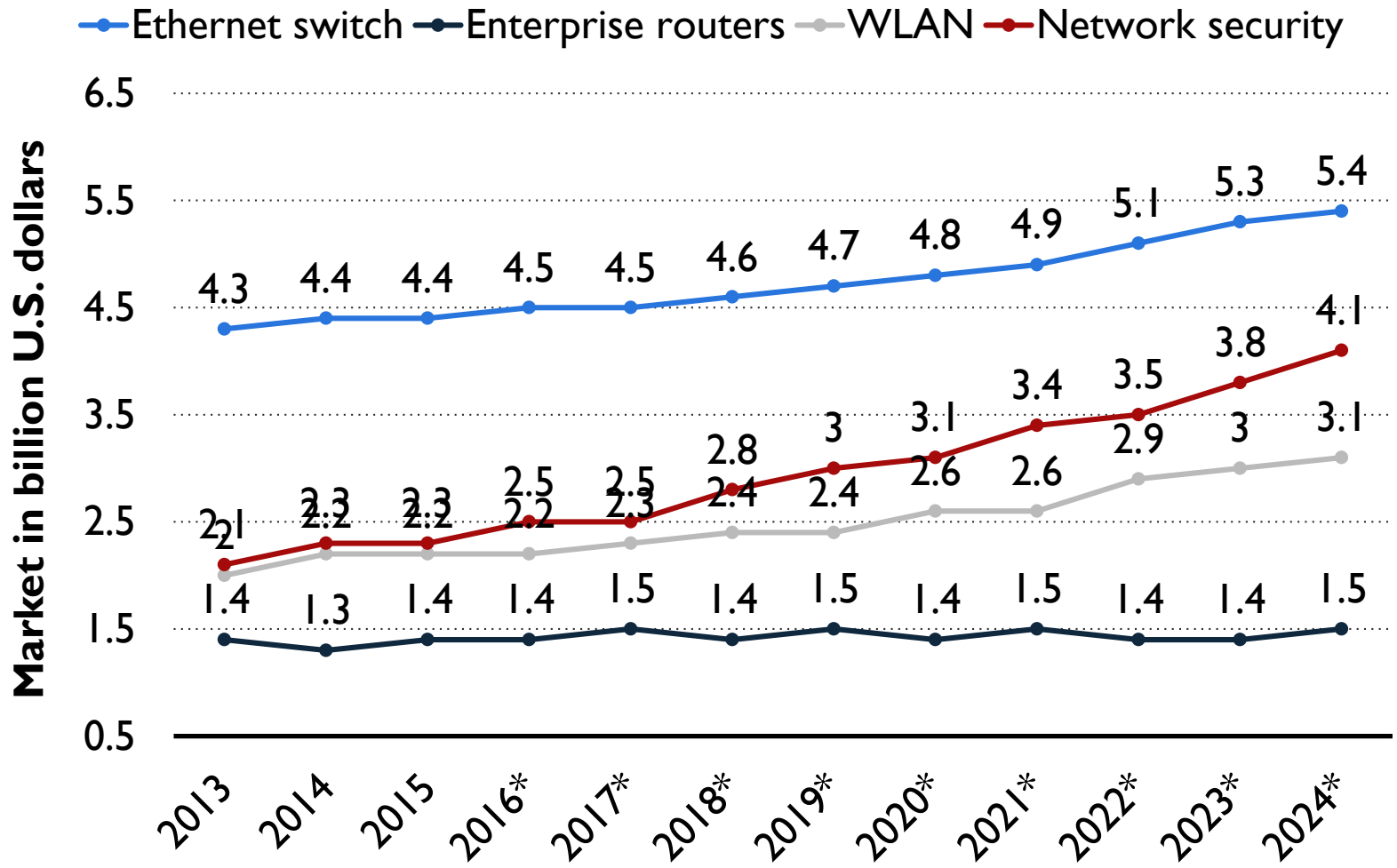
# **Lecture 7: Computer Networks for Cloud Internet and Cloud Computing**

Young Cho

Department of Electrical Engineering

University of Southern California

# Enterprise Network Market



**Note(s):** United States; 2017

Further information regarding this statistic can be found on [page 8](#).

**Source(s):** Statista estimates; Grand View Research; [ID 807581](#)

# Network Equipment Companies



a Hewlett Packard  
Enterprise company



**HUAWEI**



# Equivalent Command Line Interface

 <b>IOS XR</b>	 <b>JUNOS</b>	 <b>SROS</b>	 <b>HVRP</b>
--	---	--	--

## COMMAND LINE CHEAT SHEET

BASIC			
show	show	show	display
exit	exit / up	exit	quit
run	run	-	-
end	exit	exit all	return
include	match	match	include
... formal	display-set	-	-
reload	request system reboot	admin reboot now	reboot
GENERAL CONFIGURATION			
show running-config	show configuration	admin display-config	display current-configuration
show startup-config	-	-	display saved-configuration
configure terminal	configure / edit	configure	system view
hostname <i>hostname</i>	system host-name <i>hostname</i>	system name <i>systemname</i>	sysname <i>systemname</i>
show (after conf change)	show   compare	info (after conf change)	-
commit	commit	admin save	save

# Top 5 Certifications



# Components of Switch

- Control/Management Software
- Network OS
- Hardware Driver
- Mechanical Box
- Silicon



# White Box

- An Open Switch Architecture
- Separate switch software and hardware
- Openness, Flexibility and Programmability
- Transparent price
- Break free from vendor lock-in



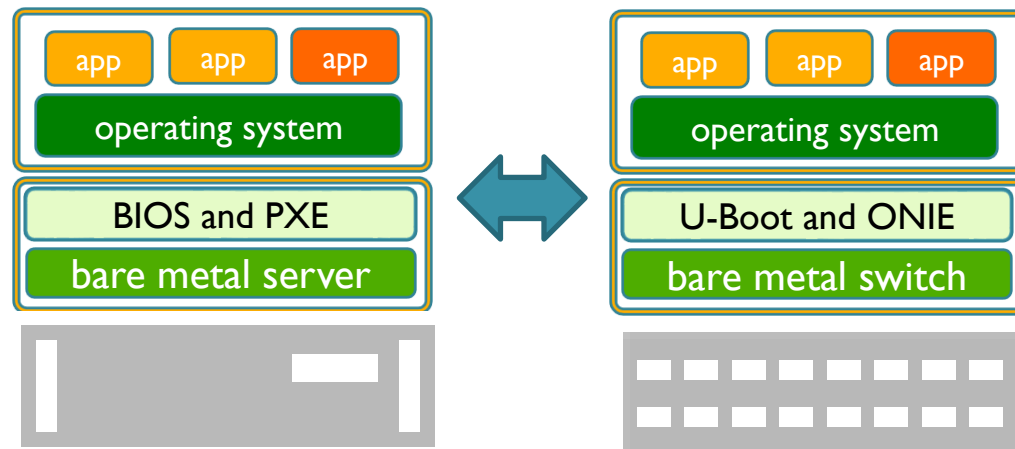
# White Box Market

Report Attribute	Details
Market size value in 2020	USD 8.8 billion
Revenue forecast in 2025	USD 25.5 billion
Growth Rate	22.0% from 2018 to 2025
Historical data	2015 - 2016
Forecast period	2018 - 2025

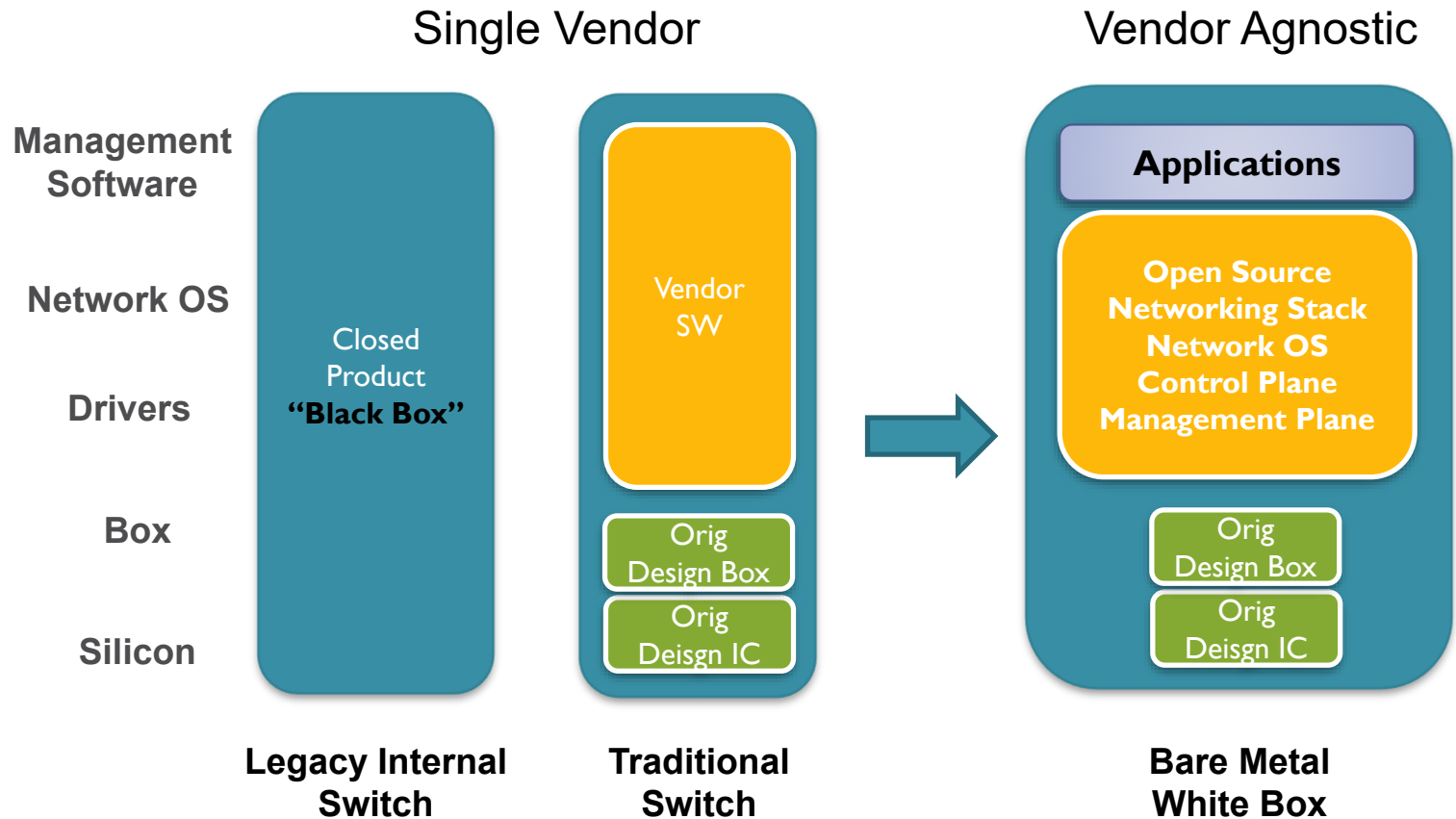


# Bare Metal Server vs. Switch

- BIOS + PXE → U-Boot + ONIE
- Open Network Install Environment
  - Supported hardware preloaded with ONIE
  - ONIE available on GitHub



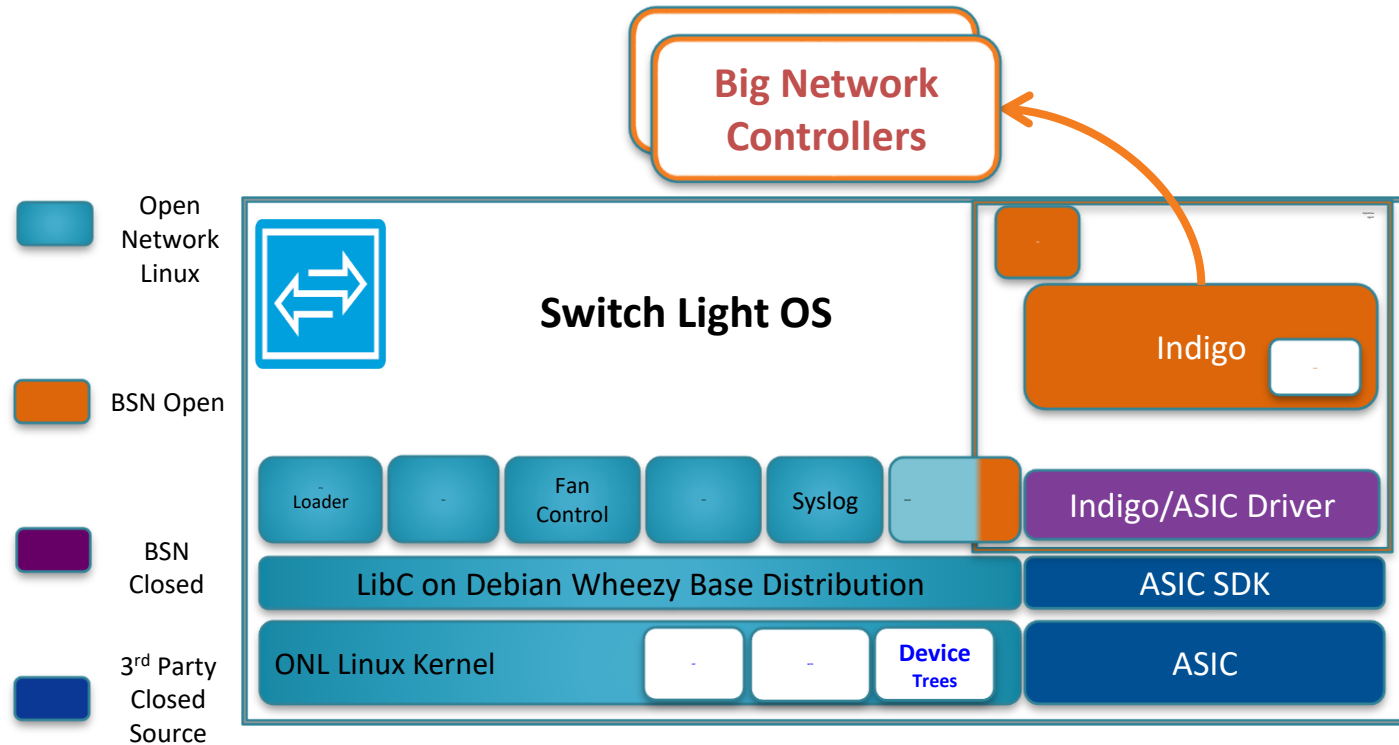
# Open Source Networking Switch



# Open Networking Software

- **Cumulus Network (Cumulus Linux)**
  - Linux NOS at web-scale
  - Configuration and Management Tools
- **Big Switch Networks (SwitchLight)**
  - Big Monitoring Fabric
  - Big Cloud Fabric via centralized controller
- **Pica8 (PicOS)**
  - Hybrid Networking: OpenFlow 1.3/1.4
  - Native L2 and L3 features
- **IP Infusion (OcNOS)**
  - Full function L2/L3 with MPLS/VPLS, Carrier Ethernet
  - OAM and telecom features
- **OpenSwitch**
  - Full function, open-source NOS
- **SnapRoute FlexSwitch**
  - L2 and L3 protocol suite
- **Open Network Linux**
  - OCP reference OS for switches
- **Open Network Foundation (CORD)**
  - OpenFlow Phy and Open VSwitches
  - ONOS OpenFlow controller with BGP
- **Open Compute Project**
  - ONIE: Universal NOS Loader
  - Switch Abstraction Interface (SAI)
- **Open Optical Monitoring API**
  - Open Networking in Cloud (SONiC)
  - Network OS with full L3 routing
  - Microsoft and co-contributors to OCP

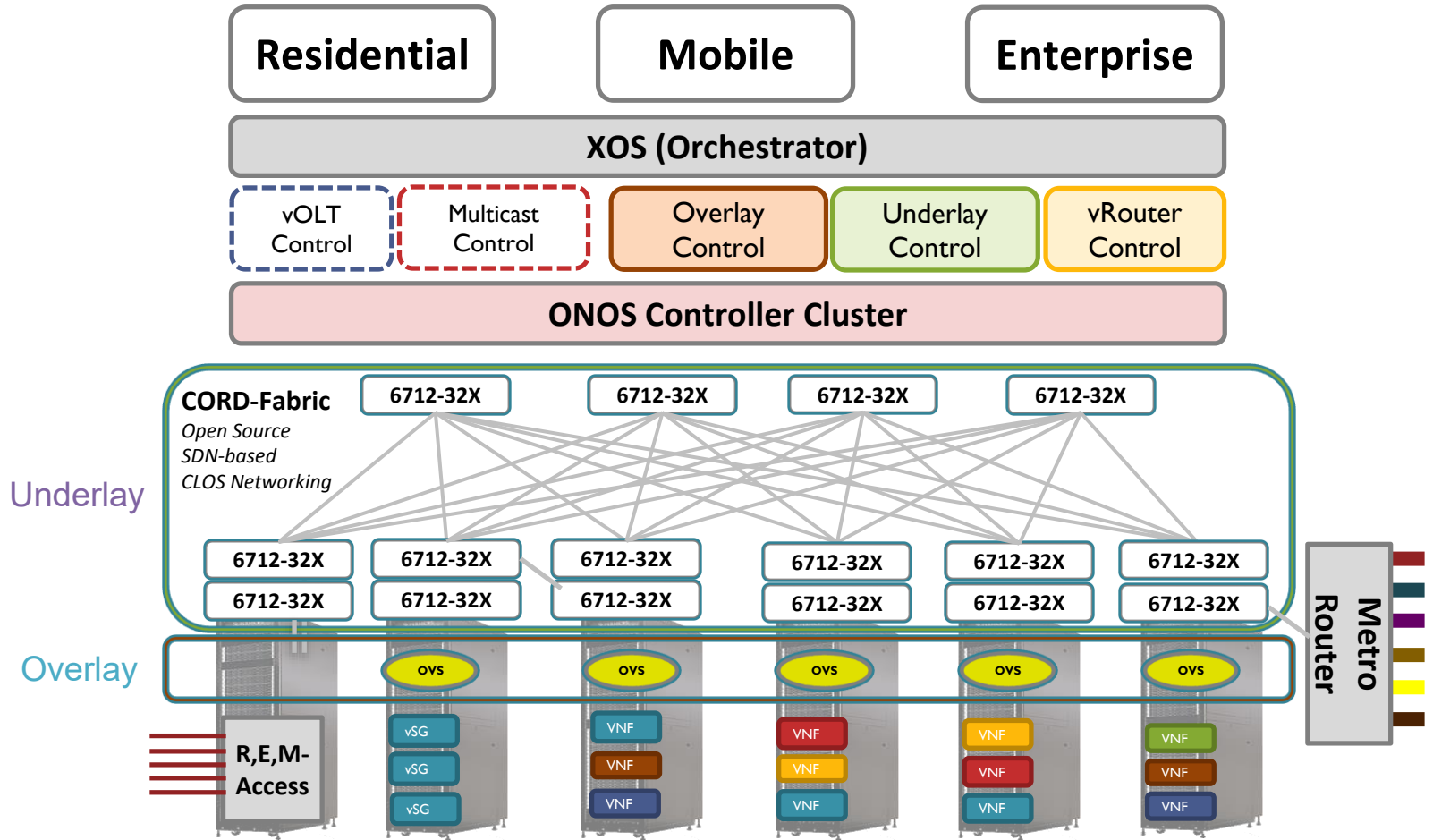
# Example: Switch Light Architecture



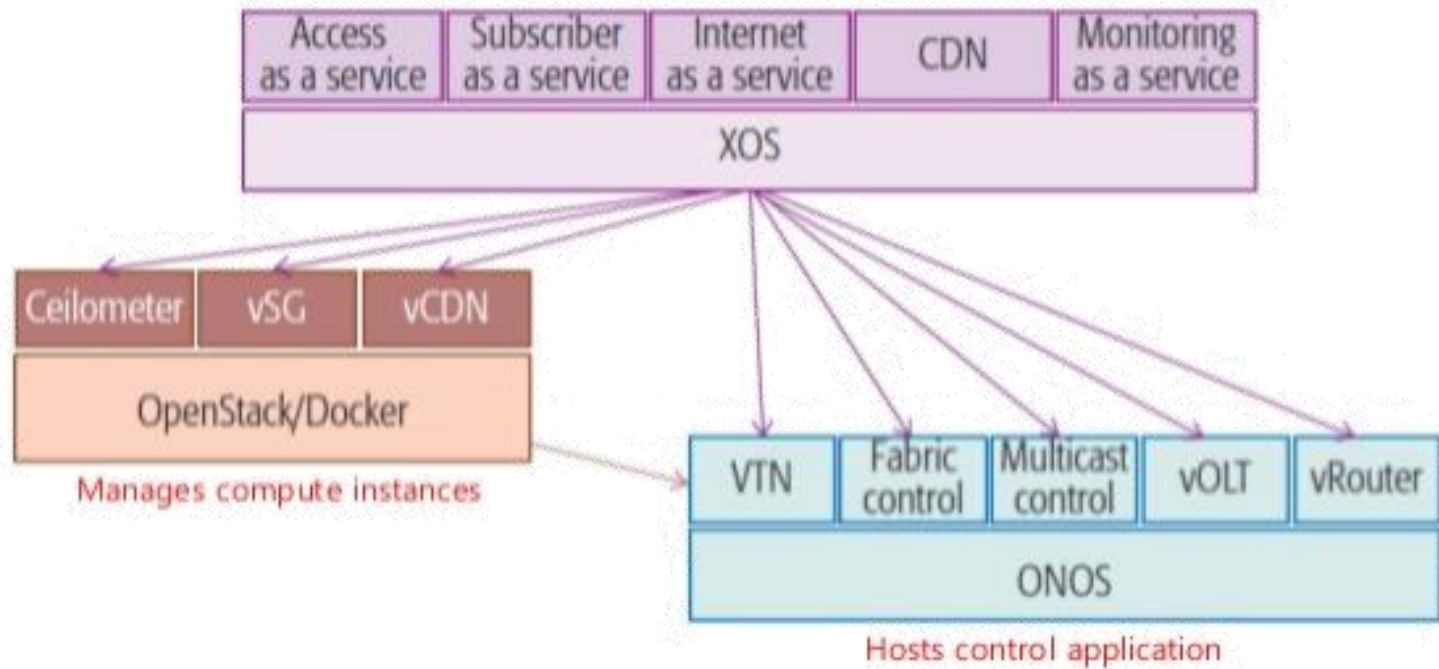
**Switch Light** is our **Indigo OpenFlow Agent** running on **Open Network Linux** on x86 or **ASIC-based** hardware.

# Example: CORD Architecture

Central Office Re-architected as a Datacenter (CORD)



# Components of CORD

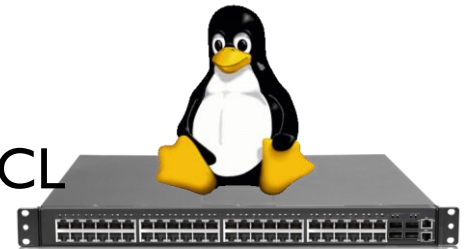


# Building Blocks of CORD

- OpenStack
  - Cluster Management Suite
  - Internet as a service (IaaS)
  - Create and Provision VMs and VNet
- Open Network OS (ONOS)
  - Collection of Control App
  - Manage SW and HW Switches
- Everything OS (XOS)
  - A service control plane on top of back-end service
  - VM-hosted VNFs, container-based services, and SDN control
  - Everything as a service (XaaS)
- Docker
  - Deploy and Interconnect Services
  - Deploy CORD

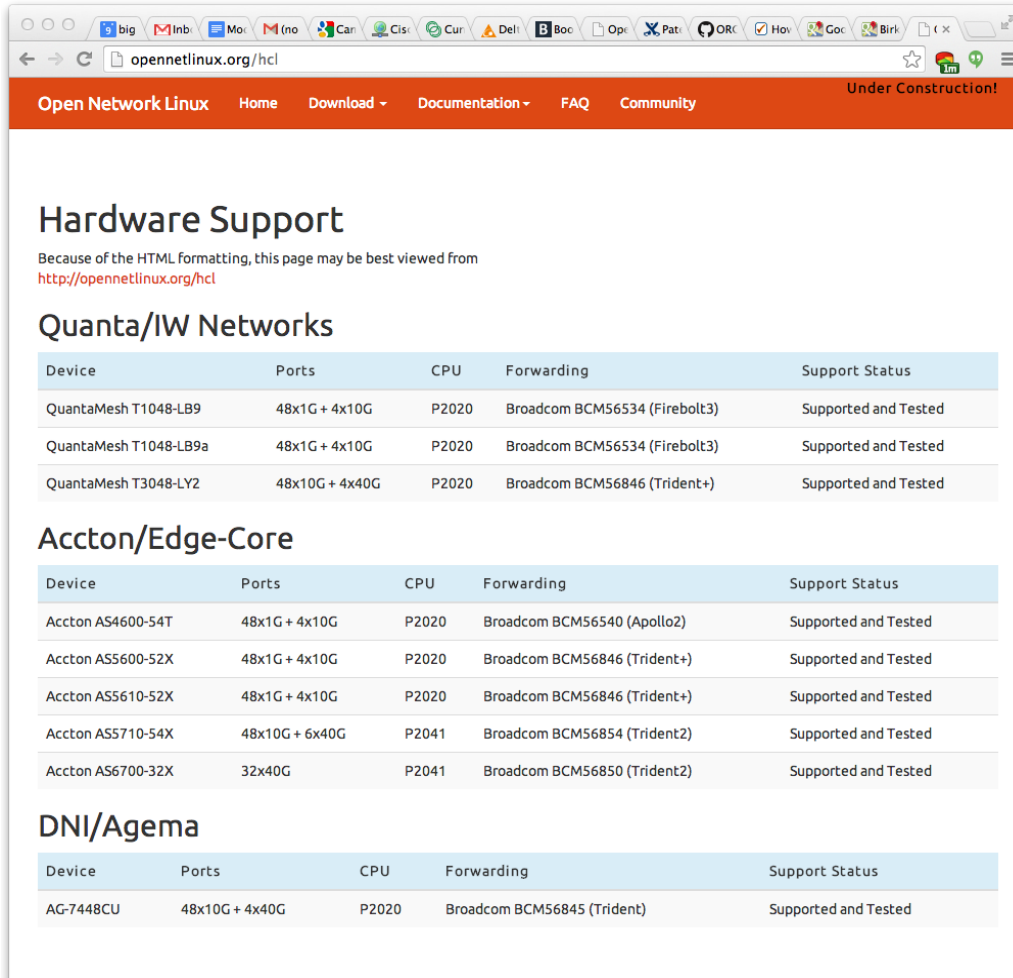
# Example: Open Network Linux

- A Linux distribution for bare metal switches
- Open-sourced build scripts, drivers from Switch Light
- Part of Open Compute Project (OCP)
- Project goals
  - Improve perception/de-risk bare metal
  - “Crowd source” larger Switch Light HCL
  - Released mid-January: Initial results
- Strong community response from ODMs, OCP
  - Accton self-supporting three new boxes, more ODMs interested
  - Other vendors investigating ONL for their own purposes
  - Info: <http://opennetlinux.org> – under construction





# ONL – Supported Hardware



The screenshot shows the Open Network Linux (ONL) website. The browser address bar displays `opennetlinux.org/hcl`. The website has an orange header with navigation links: Open Network Linux, Home, Download, Documentation, FAQ, and Community. A status message "Under Construction!" is visible in the top right. The main content area is titled "Hardware Support" and includes a note about HTML formatting. It lists supported hardware for three categories: Quanta/IW Networks, Accton/Edge-Core, and DNI/Agema. Each category contains a table with columns for Device, Ports, CPU, Forwarding, and Support Status.

## Hardware Support

Because of the HTML formatting, this page may be best viewed from <http://opennetlinux.org/hcl>

### Quanta/IW Networks

Device	Ports	CPU	Forwarding	Support Status
QuantaMesh T1048-LB9	48x1G + 4x10G	P2020	Broadcom BCM56534 (Firebolt3)	Supported and Tested
QuantaMesh T1048-LB9a	48x1G + 4x10G	P2020	Broadcom BCM56534 (Firebolt3)	Supported and Tested
QuantaMesh T3048-LY2	48x10G + 4x40G	P2020	Broadcom BCM56846 (Trident+)	Supported and Tested

### Accton/Edge-Core

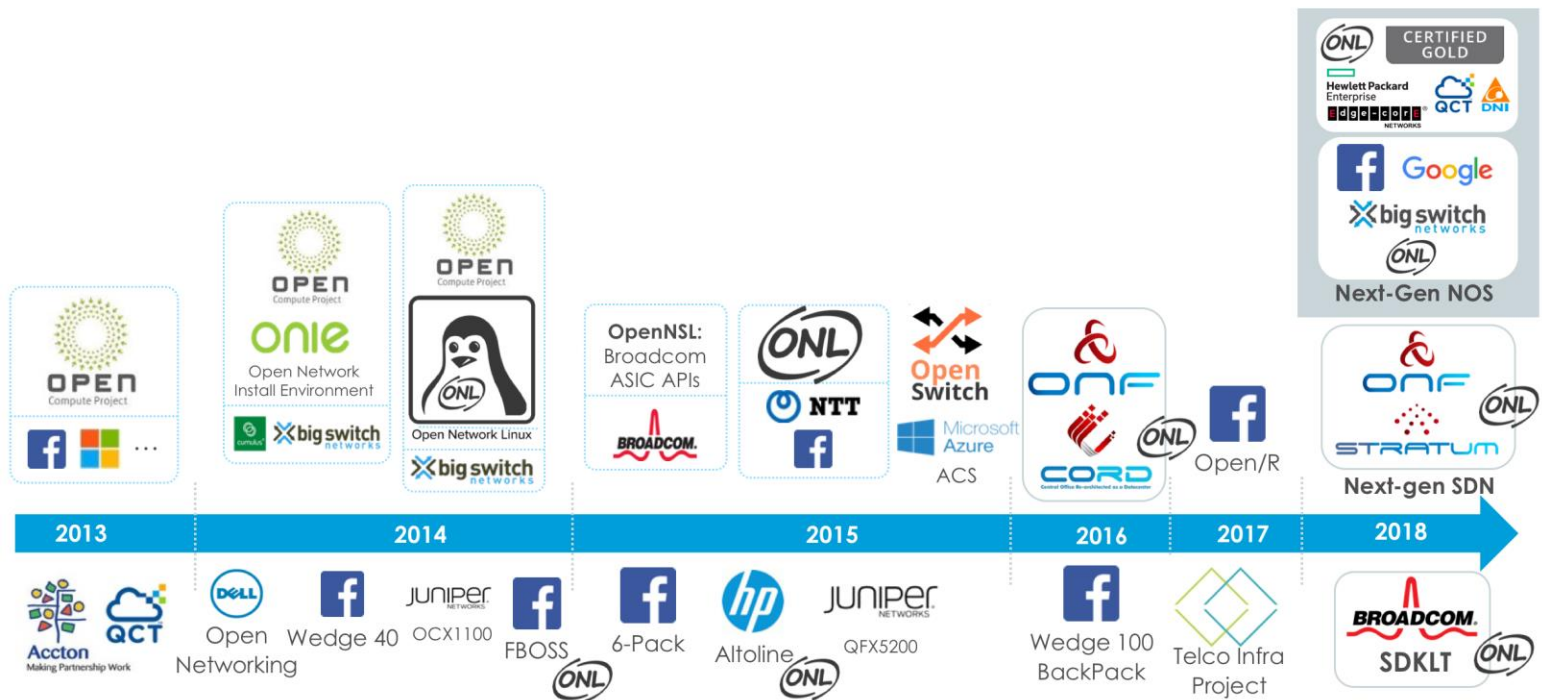
Device	Ports	CPU	Forwarding	Support Status
Accton AS4600-54T	48x1G + 4x10G	P2020	Broadcom BCM56540 (Apollo2)	Supported and Tested
Accton AS5600-52X	48x1G + 4x10G	P2020	Broadcom BCM56846 (Trident+)	Supported and Tested
Accton AS5610-52X	48x1G + 4x10G	P2020	Broadcom BCM56846 (Trident+)	Supported and Tested
Accton AS5710-54X	48x10G + 6x40G	P2041	Broadcom BCM56854 (Trident2)	Supported and Tested
Accton AS6700-32X	32x40G	P2041	Broadcom BCM56850 (Trident2)	Supported and Tested

### DNI/Agema

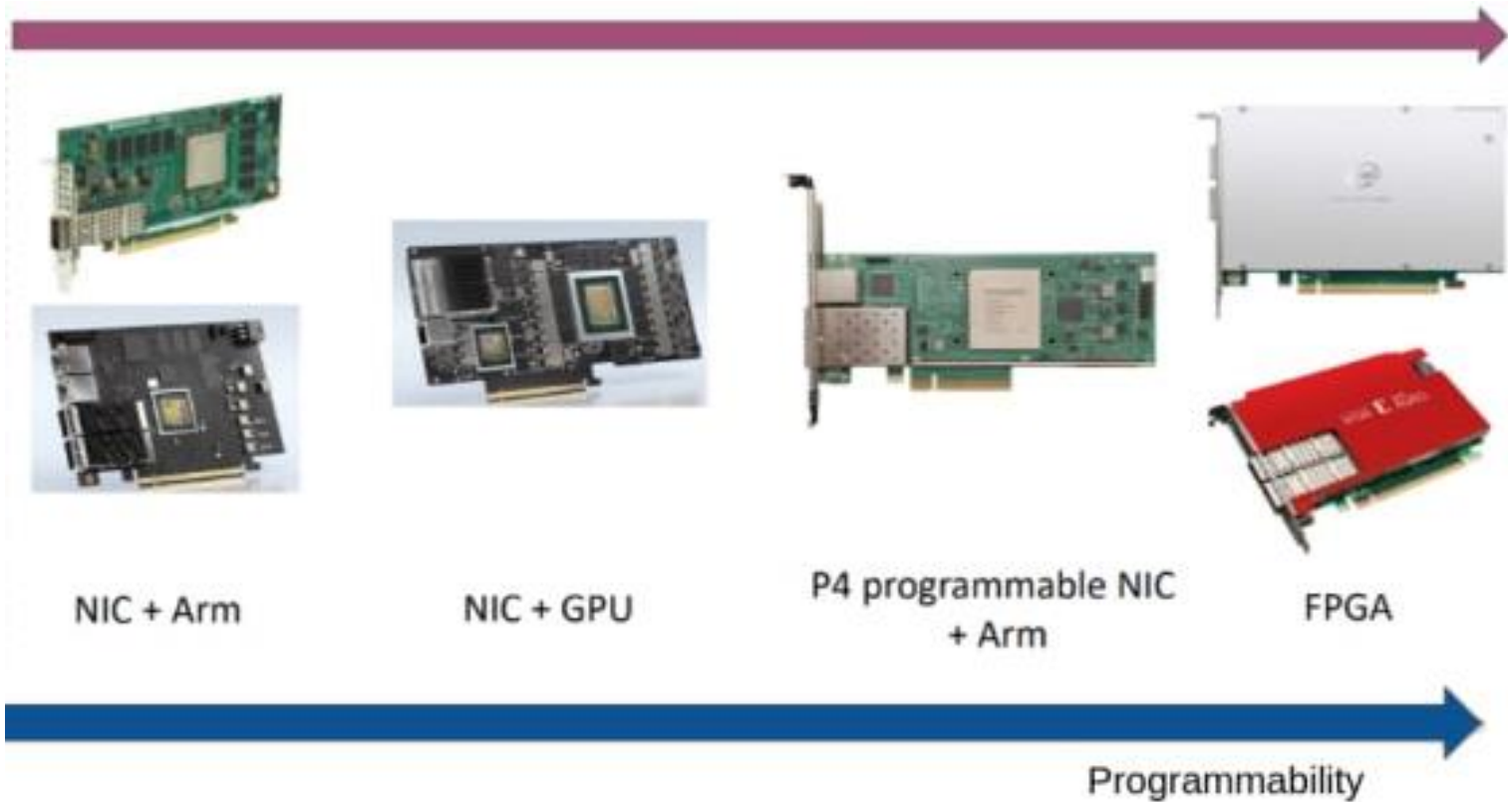
Device	Ports	CPU	Forwarding	Support Status
AG-7448CU	48x10G + 4x40G	P2020	Broadcom BCM56845 (Trident)	Supported and Tested

- ONL Support is Pre-cursor to Switch Light Support
- Working on OCP box from Interface Masters with x86 support

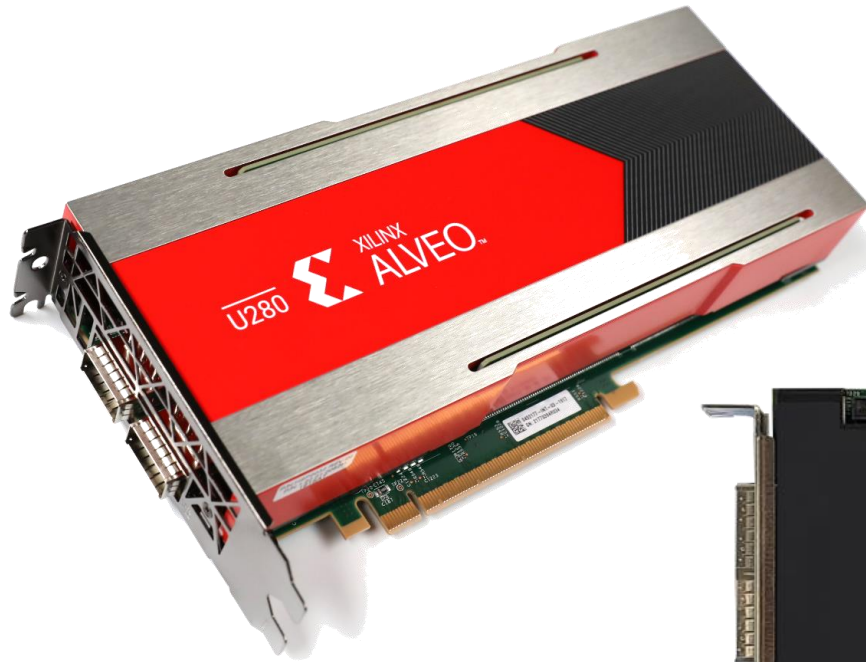
# Open Networking History



# SmartNICs

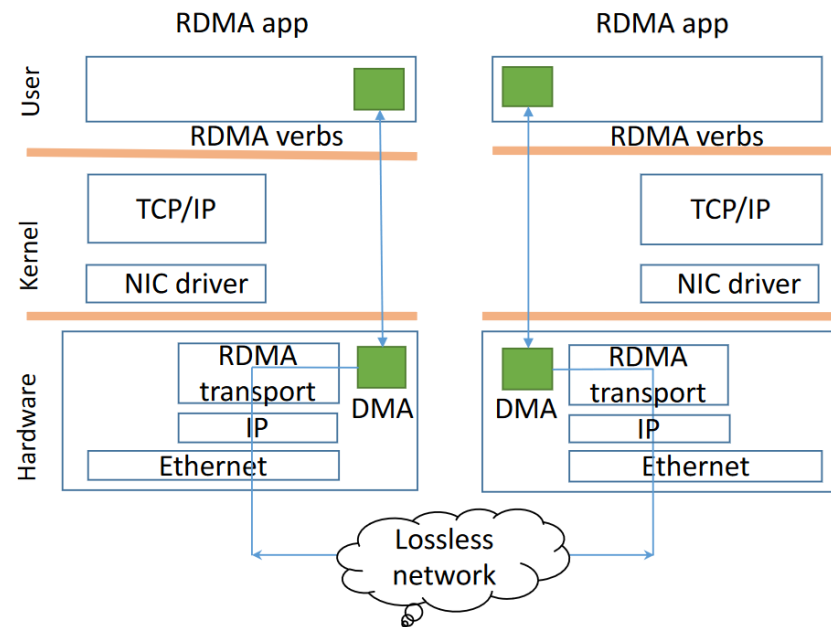


# Reconfig HW based SmartNICs



# Remote Direct Memory Access

- **RDMA enable Network**
  - Low and stable latency, e.g., **< 10us**; High throughput, e.g., **100Gbps**; **Low CPU overhead**
  - Widely **deployed** in companies such as Microsoft, Alibaba, ByteDance
- **Network stack in NIC**
  - Processing in dedicated NIC hardware
  - Bypass kernel
  - Zero copy



# Remote Direct Memory Access

## ❖ Remote

- data transfers between nodes in a network

## ❖ Direct

- no Operating System Kernel involvement in transfers
- everything about a transfer offloaded onto Interface Card

## ❖ Memory

- transfers between user space application virtual memory
- no extra copying or buffering

## ❖ Access

- send, receive, read, write, atomic operations

# RDMA Benefits

- ❖ High throughput
- ❖ Low latency
- ❖ High messaging rate
- ❖ Low CPU utilization
- ❖ Low memory bus contention
- ❖ Message boundaries preserved
- ❖ Asynchronous operation

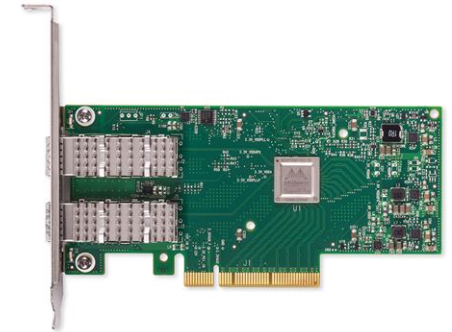
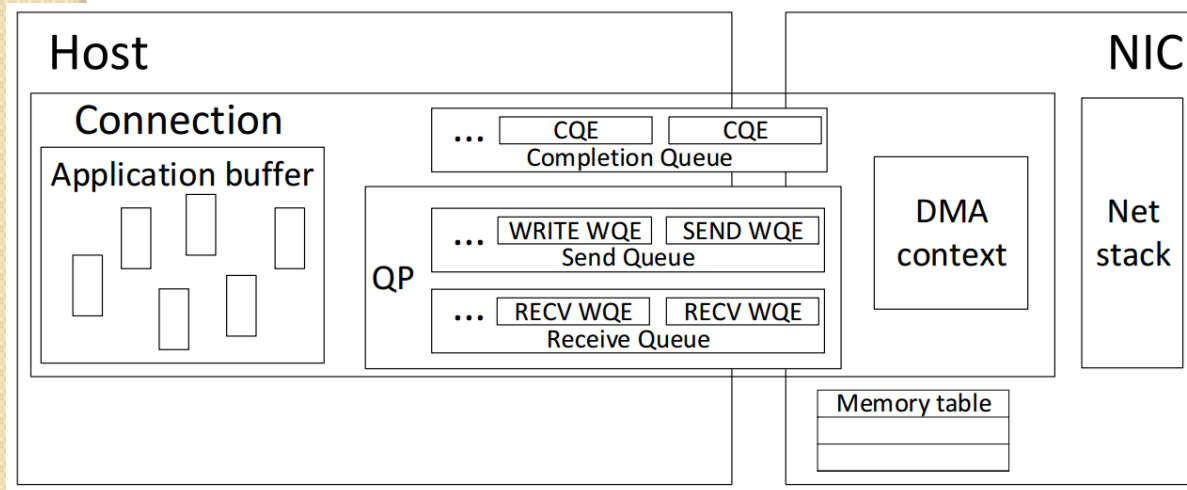
# RDMA Technologies

- ❖ InfiniBand – (41.8% of top 500 supercomputers)
  - SDR 4x – 8 Gbps
  - DDR 4x – 16 Gbps
  - QDR 4x – 32 Gbps
  - FDR 4x – 54 Gbps
- ❖ iWarp – internet Wide Area RDMA Protocol
  - 10 Gbps
- ❖ RoCE – RDMA over Converged Ethernet
  - 10 Gbps
  - 40 Gbps



# Uniqueness of RDMA stack

- Network stack on RDMA NIC (RNIC)
  - Maintain **massive connection-related states on RNIC**



e.g., MLX maintains **256B** states for each RDMA connection  
(`/include/linux/mlx4/qp.h`  
`is, ... mlx4_qp_context`)

Still growing...

# Lab 5 Preview

- Modifying Linux Kernel Module
- Read and understand the congestion control
- Fix/Break it to Not Back Off