# 갈라파고스의 이구아나

인프라에 오픈소스를 올리다.
그래서 보이는 오픈스택


장재민

# 스마일서브

- IDC 운영 지식공유
  - http://idchowto.com
- 화사대표 도메인
  - http://cloudv.kr
- 게임엔진서비스
  - http://igamev.com
- 2002.06
- 69 : 60

# 목차

- 베어메탈 서버 배포
- 서버환경
  - 전원관리
  - 콘솔
- 배포할 운영체제
- 운영체제 배포
- 네트워크
- 네트워크 보안
- ARM Server 그리고 POE


for your Safety Life !
ICAM
인텔 82599칩 사용 범용 10G 듀얼 랜카드
10G 듀얼 광 랜카드
intel
10Gbps SFP+
10Gbps
Windows Base
Linux
PCI Express


SMILESERV

# 전원 관리

- Power on
  - IPMI 2.0 / wol
  - apt-get install ipmitool wakeonlan
- Power off
  - Ipmitool / shutdown(?)
- Power reset
  - ipmitool

콘솔

```
e_mangle iptable_nat nf_conntrack_ipv4 nf_conntrack_netlink nf_nat nf_conntrack nfnetlink 8021q iptable_filter ip_tables x_table
s deflate zlib_deflate twofish twofish_common camellia serpent blowfish des_generic cbc ecb geode_aes blkcipher aes_generic aes_
i586 xcbc sha256_generic sha1_generic crypto_null af_key lp loop ipv6 parport_pc parport evdev serio_raw psmouse pcspkr containe
r button ac i2c_piix4 i2c_core intel_agp agpgart shpchp pci_hotplug ext3 jbd mbcache sg sd_mod floppy pata_acpi mptspi mptscsih
mptbase scsi_transport_spi pcnet32 mii ata_piix ata_generic libata scsi_mod thermal processor fan fuse vesafb fbcon tileblit fon
t bitblit softcursor
[  236.317101]
[  236.317696] Pid: 5741, comm: conntrackd Not tainted (2.6.24-18-server #1)
[  236.318326] EI
[  236.319076] EI
[  236.319761] EA
[  236.320347] ES
[  236.321004]  D
[  236.321741] Pr
[  236.326187] St
[  236.328297]
[  236.329068]
[  236.329797] Ca
[  236.331121]  [
[  236.332014]  [
[  236.333260]  [
[  236.335210]  [
[  236.336055]  [
[  236.336800]  [
[  236.337573]  [
[  236.338278]  [
[  236.339071]  [
[  236.339885]  [
[  236.340542]  [
[  236.341164]  [
[  236.341854]  [
[  236.342637]  [
[  236.344100]  [
[  236.345344]  [
[  236.346016]  [
[  236.346730]  [
[  236.347455]  [
[  236.348460]  [
[  236.349476]  [
[  236.356625]  [
[  236.358221]  [
[  236.359039]  [<c0106f3a>] sysenter_past_esp+0x6b/0xa1
[  236.359917]  =======================
[  236.360627] Code: 48 8b 6c 24 08 89 4c 24 04 83 c7 0e 83 c5 0c eb 0e 8d 74 26 00 85 db 89 de 0f 84 ba 00 00 00 8b 1e 0f 18 03
 90 8b 56 18 8d 4a 0c <0f> b6 41 26 3a 84 24 96 00 00 00 75 dd 8b 01 3b 44 24 70 75 d5
[  236.362476] EIP: [<c8c64b66>] nf_nat_setup_info+0x446/0x6d0 [nf_nat] SS:ESP 0068:c7075af4
[  236.364291] Kernel panic - not syncing: Fatal exception in interrupt
```

```
[<ffffffff80424477>] acpi_sc
[<ffffffff80424188>] acpi_event_ini
[<ffffffff80407a5c>] init+0x1f9/0x2
[<ffffffff8005efb1>] child_rip+0xa/
[<ffffffff801824cc>] acpi_ds_init_
[<ffffffff80407863>] init+0x0/0x2f
[<ffffffff8005efa7>] child_rip+0x0/

Code: 48 8b 55 00 48 39 da 74 1b 48
RIP  [<ffffffff80154df4>] __list_ad
 RSP <ffff810107f37820>
CR2: 0000000000000000
 <0>Kernel panic - not syncing: Fat
```


MegaRAC SP-X


Video Keyboard Mouse Options Media Keyboard Layou Video Recorc Power Active Users Help
Zoom Size : 100%
50
100
150
Welcome to CentOS 6.4!
Install or upgrade an existing system
Install system with basic video driver
Rescue installed system
Boot from local drive
Memory test
Press [Tab] to edit options
Automatic boot in 57 seconds...
CentOS 6
Community ENTerprise Operating System
LALT LCTRL RALT RCTRL Num Caps Scroll

## 더미 서버와 스마트 관리 서버의 차이점

**더미서버**

- IDC방문 또는 작업의뢰 신청
- 담당자 접수 확인
- IDC내 담당자 서버실 출입
- 해당 서버 콘솔 연결
- 작업시작

VS

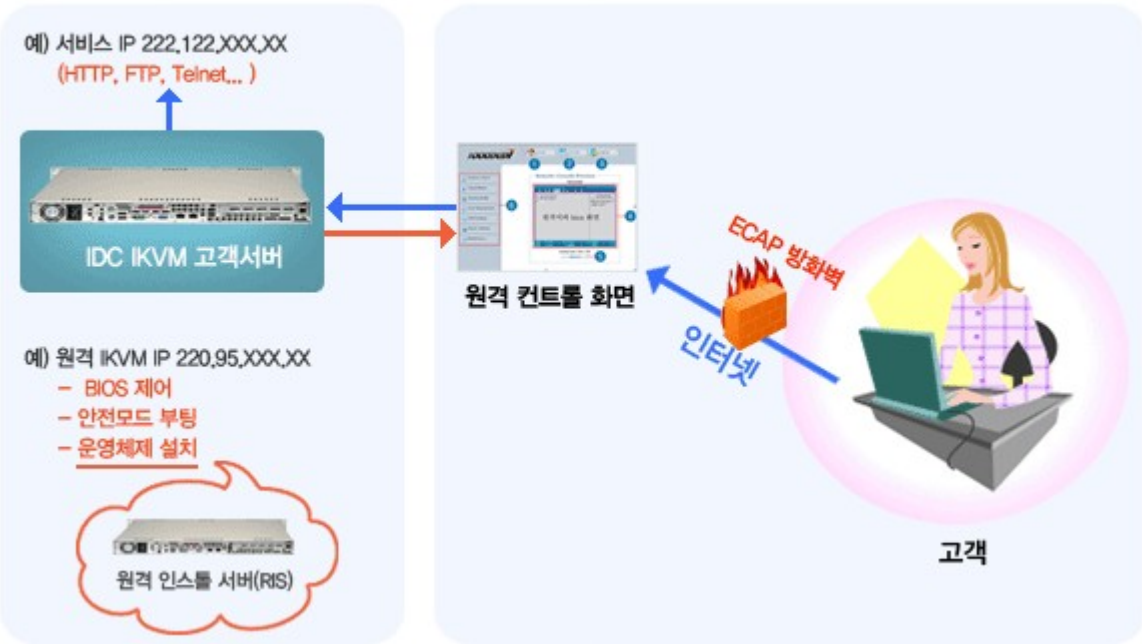
**스마트 관리 서버**

- 원격서버 관리페이지 접속
- 5분
- 작업시작

시간: 최소 30분(작업의뢰시)
비용: 최소 5천원(콘솔연결 30분)

시간 : 5분
비용 : 없음

**[IKVM 서비스 매커니즘]**



Key Features

1. Intel® Xeon® X3400 / L3400 series,

Core™ i3 & Pentium® processors

with LGA 1156 socket2. Intel® 3420 Chipset3. Up to 32GB DDR3 1333/1066/800MHz

ECC Registered DIMM / 16GB

Unbuffered DIMM4. Dual Intel® 82574L Gigabit Ethernet

Controllers5. 6x SATA (3 Gbps) Ports

RAID 0, 1, 5, 106. 2 (x8) PCI-Express 2.0,

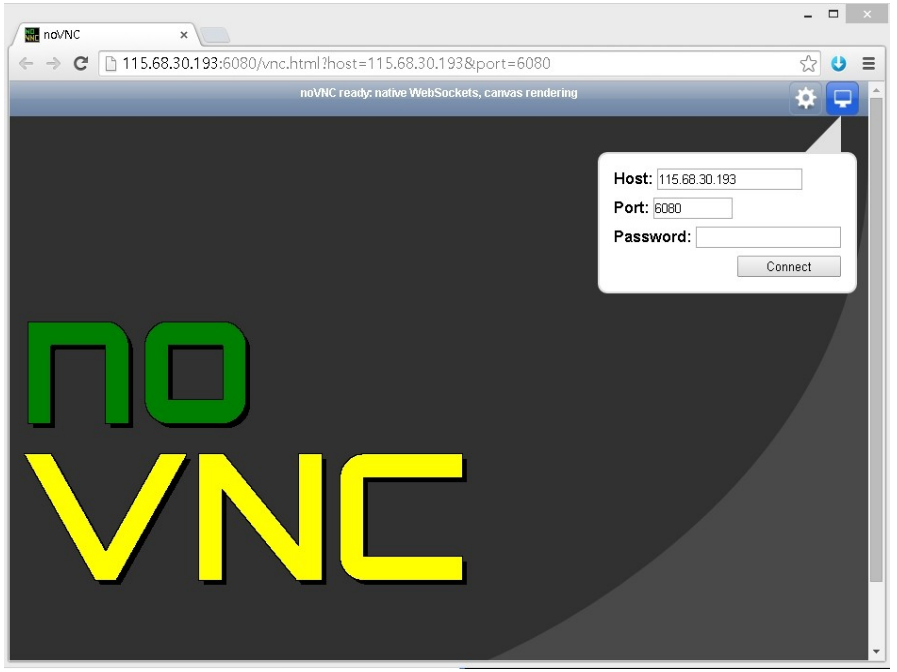
1 (x4) PCI-Express (using x8 slot),

1 32-bit PCI slot7. **Integrated IPMI 2.0 with KVM** and

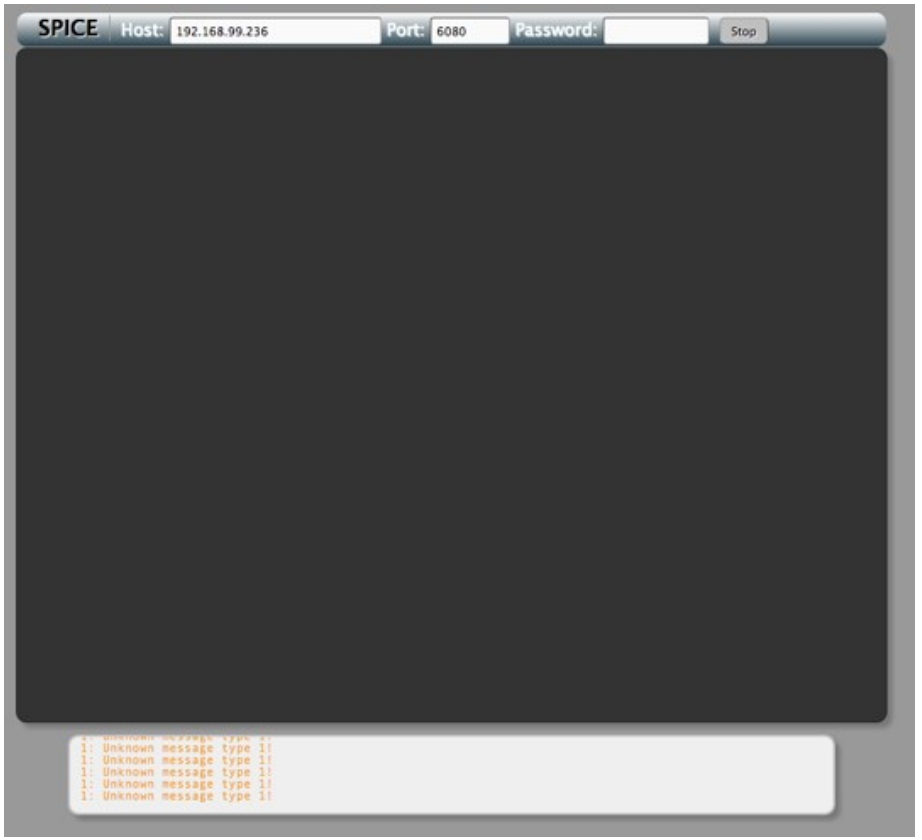
Dedicated LAN8. 7x USB (2 rear, 1 on-board, 2 headers)

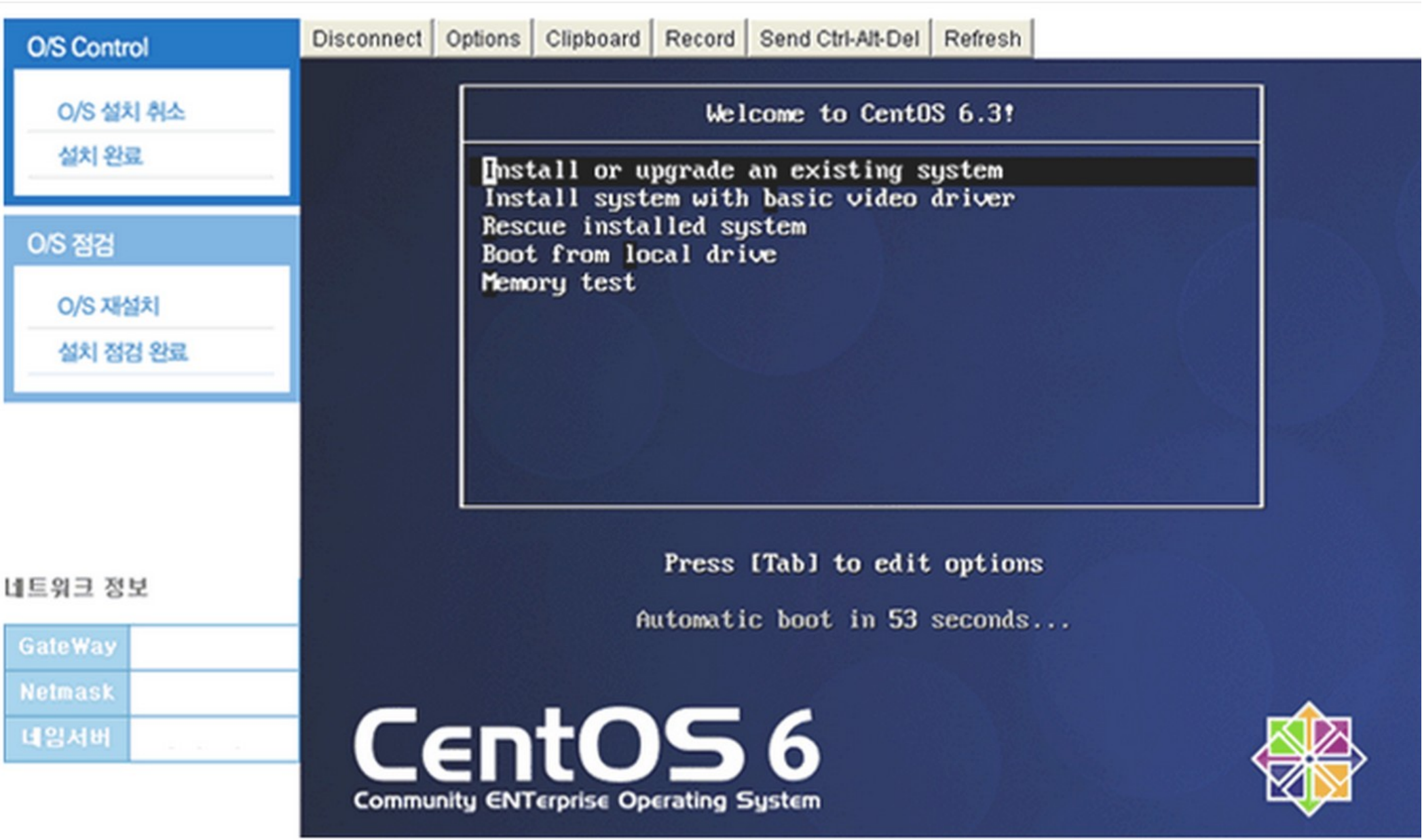
# VNC vs SPICE

- NoVNC
  - vncproxy




- Spice : HTML5
  - Spice + websockfy




O/S Control
O/S 설치 취소
설치 완료
O/S 점검
O/S 재설치
설치 점검 완료
네트워크 정보
GateWay
Netmask
네임서버
Disconnect
Options
Clipboard
Record
Send Ctrl-Alt-Del
Refresh
Welcome to CentOS 6.3!
Install or upgrade an existing system
Install system with basic video driver
Rescue installed system
Boot from local drive
Memory test
Press [Tab] to edit options
Automatic boot in 53 seconds...
CentOS 6
Community ENTerprise Operating System

| 구분 | 버전 | |
|---|---|---|
| CentOS | 32bit | 5.X, 6.X, 7.X (Final version) |
| | 64bit | 5.X, 6.X, 7.X (Final version) |
| Fedora | 32bit | 15, 16, 17, 18, 19, 20 |
| | 64bit | 15, 16, 17, 18, 19, 20 |
| Redhat | 32bit | 7, 8, 9 |
| | 64bit | 제공되지 않음. |
| SULinux | 32bit | 1.0, 1.5, 2.0 |
| | 64bit | 2.0 |
| OpenSuse | 32bit | 10.3, 11.4, 12.1 |
| | 64bit | 10.3, 11.4, 12.1 |
| Ubuntu | 32bit | 10.X, 11.X, 12.X, 13.X |
| | 64bit | 10.X, 11.X, 12.X, 13.X, 14.X |
| Debian | 32bit | 5.X, 6.X |
| | 64bit | 5.X, 6.X |
| FreeBSD | 32bit | 6.X, 7.X, 8.X, 9.X |
| | 64bit | 6.X, 7.X, 8.X, 9.X |
| Gentoo | 32bit | 2011, 2012, 2013 |
| | 64bit | 2011, 2012, 2013 |
| 안녕Linux | 32bit | 제공되지 않음 |
| | 64bit | 1.2, 1.3, 1.3-R2, 1.3-R3, 1.3-R4, 1.3-R5 |
| ArchLinux | 32bit | 2012.10.06, 2013.11.01 |
| | 64bit | 2012.10.06, 2013.11.01 |
| AsteriskNOW | 32bit / 64bit | 3.0.0 |

# 배포할 운영체제 이미지

- Kickstart Installation
- Preinstalled OS Image
  - Sysprep
    - debootstrap
    - yumbootstrap
      - tool for installing Yum-based distributions (Red Hat, CentOS, Fedora) in a chroot directory
      - https://github.com/dozzie/yumbootstrap
- http://ftp.cloudv.kr
  - http://ftp.cloudv.kr/openstack-image

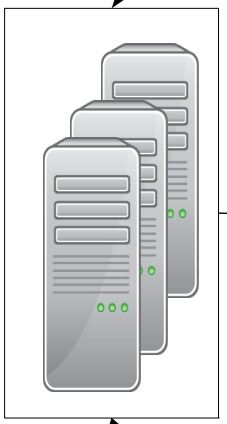
- Ubuntu Cloud Image
  - Ubuntu Cloud Images are pre-installed disk images that have been customized by Ubuntu engineering to run on cloud-platforms
  - https://cloud-images.ubuntu.com/

- Redhat Cloud Image
  - Downloading Pre-Built Images for OpenStack
    - https://openstack.redhat.com/Image_resources
  - Image builder
    - Diskimage-builder
    - Image factory
      - http://imgfac.org/

- Openstack
  - Get Images
    - http://docs.openstack.org/image-guide/content/ch_obtaining_images.html
  - Linux Image Requirements
    - Disk partitions and resize root partition on boot (cloud-init)
      - Cloud-intiramfs-growroot
    - No hard-coded MAC address information
    - Disable Firewall and ssh server running

# 운영체제 배포

Fixed IP
NIC or MB : iscsi support



Baremetal DB

1) request to deploy

2) Select Barebone

4) Power on

Node-controller

3) enable pxe

12) disable pxe

5) DHCP request

6) DHCP reply with bootparams

7) Download boot image

8) connect iscsi

dhcpd

tftpd

10) iscsi boot
11) copy image to local
12) fix the ip address

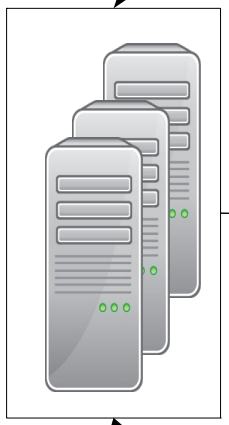
11) copied and reboot

Iscsi target

Image DB

Preinstalled OS Image DB

Fixed IP
NIC or MB : iscsi support



Baremetal DB

1) request to deploy

2) Select Barebone

3) enable pxe

4) Power on

Node-controller

12) disable pxe

5) DHCP request

6) DHCP reply with bootparams

7) Download boot image

8) connect iscsi

dhcpd

tftpd

10) iscsi boot

11) deployed

Iscsi target

Image DB

Preinstalled OS Image DB

# Deployment (PXE)

Bare Metal Nodes

IPMI + PXE

glance

quantum

nova-scheduler

6) fetch images

5) plug VIFs

2) apply filters &
find available node

1) nova boot ...

nova-api

nova-computer

3) ComputeManager calls
driver.spwan()

message
queues

4) get info & claim mode

9 ) write image
10) reboot

12) update instance state

nova-conductor

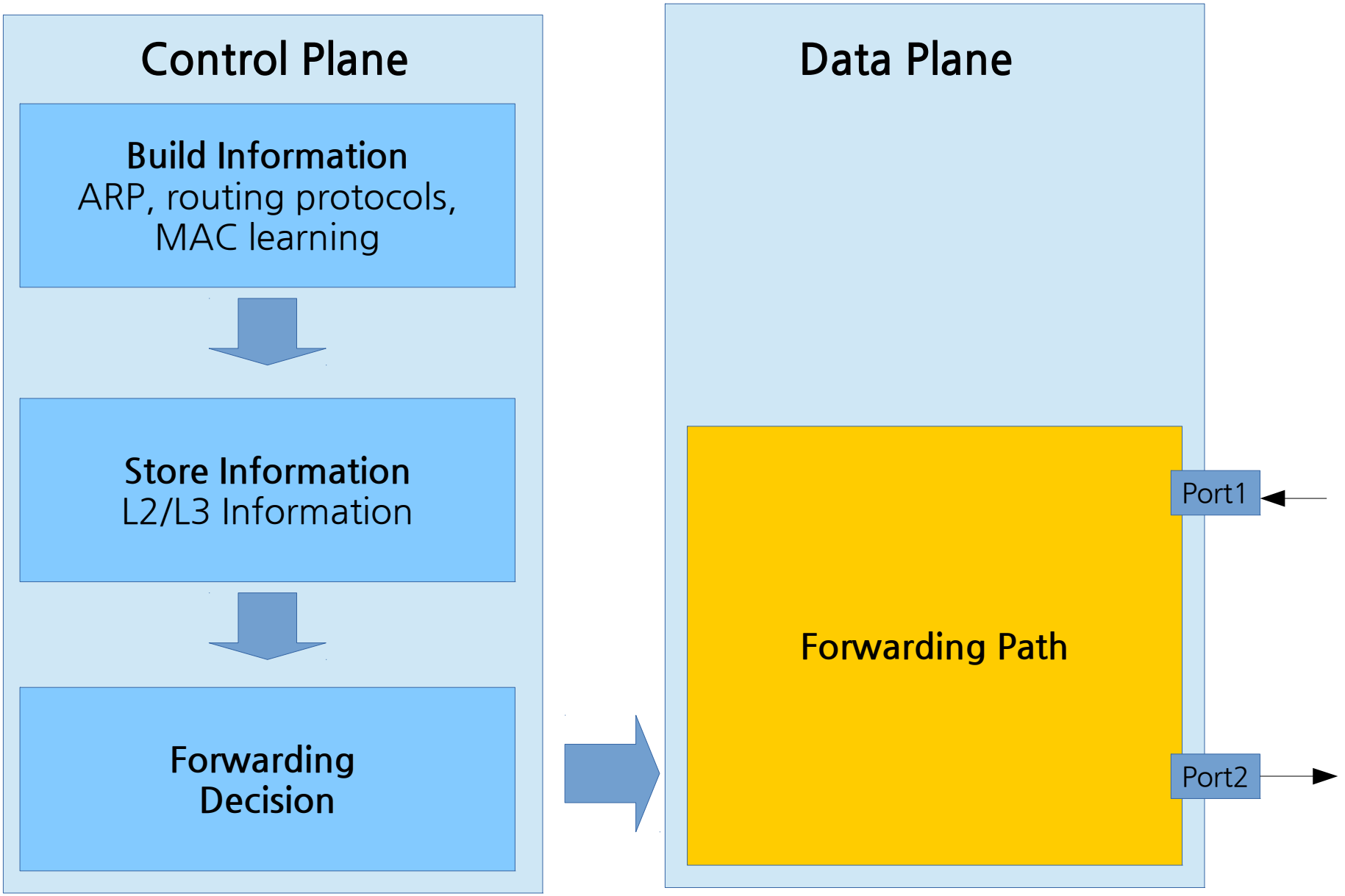
Nova-baremetal-deploy-helper

11) update status of node

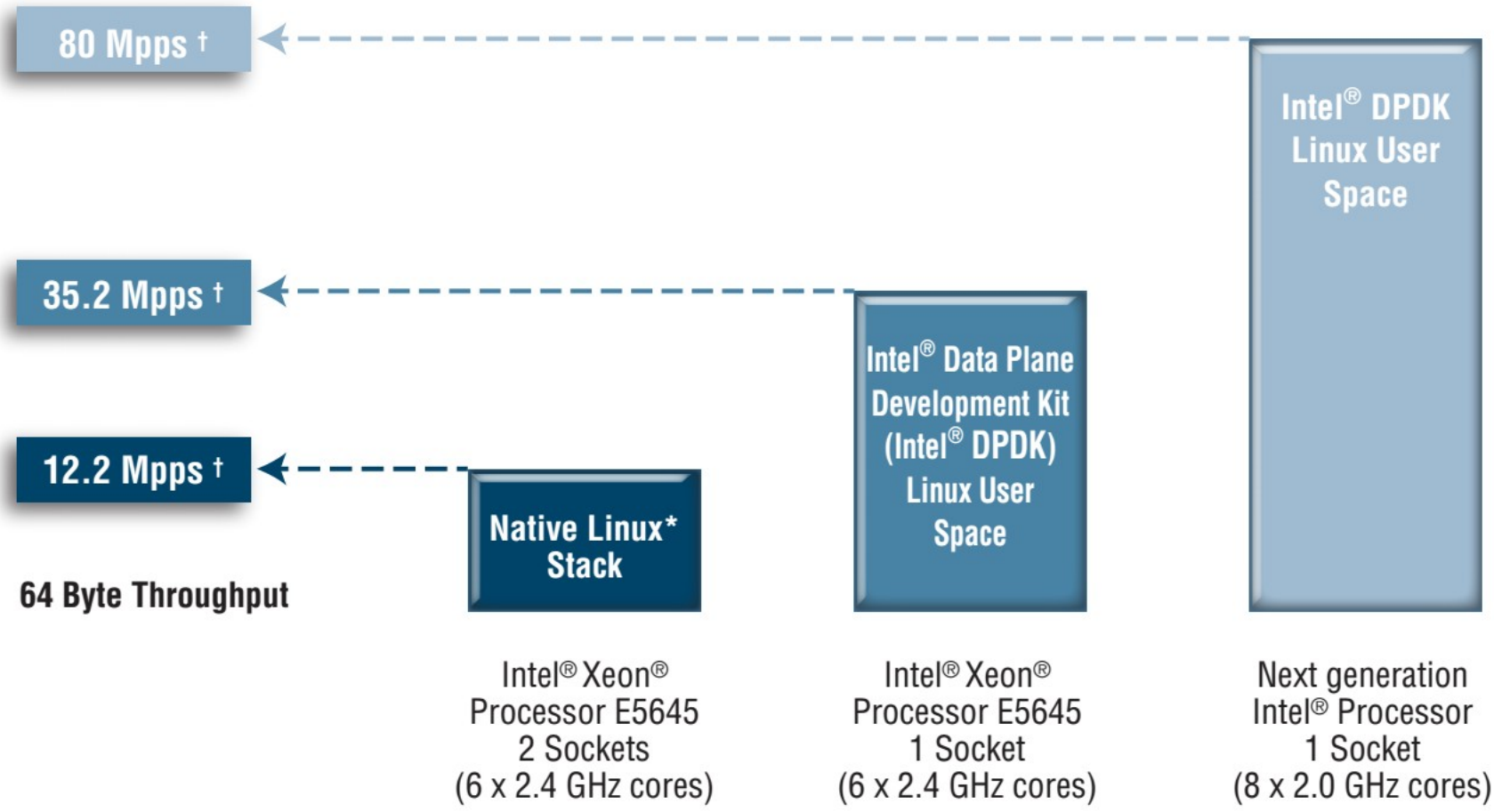
Baremetal
Database

Nova
Database

# Network Switch



# Packet Fowarding




† Measured performance. IO Performance may vary with platform configuration

†† Estimates are based on internal analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

**Figure 2.** Breakthrough Data Plane Performance with Intel® Data Plane Development Kit (Intel® DPDK) L3 Packet Forwarding

- Ports
  - 10GbE SFP+  48 Ports. 10GbE or 1GbE
  - 40GbE QSFP  4 Ports. 40GbE or 4x10GbE
- Performance
  - Forwarding  960 Mpps
  - Throughput  1.28 Tbps line rate L2 and L3 switching
- Software
  - Installer Loaded with Open Network Install Environment (ONIE)
  - SwitchOS  Compatible with Cumulus Linux r1.5.1 and later

# **O**pen **N**etwork **I**nstall **E**nvironment

- Combines a boot loader with a modern Linux kernel and BusyBox
- Disruptive, liberating users from a captive, pre-installed network OS
- Manage your switches like you manage your Linux servers
- Provides an environment for installing any network OS
- Helps automate large scale data center switch provisioning

# Pica8 PicOS

- Network operation system using user space standard Debian Linux environment
- Leverage vast array of standard Linux tools as a common management and operations framework
- Zero Touch Provisioning (ZTP) functionality coupled with ONIE delivers a true bare metal to application environment.
- Rich Layer-2 protocol stack with MLAG, seamlessly integrating into existing architectures
- Full Layer-2 & Layer-3 ACL support
- IPv4 & IPv6 Static Routing
- Open vSwitch (OVS) Mode provides Industry-leading OpenFlow 1.3 support and integration with open networking software like OpenStack

  https://github.com/Pica8/Openstack-Neutron/wiki

  ml2_conf.ini mechanism_drivers = openvswitch,pica8

Netfilter

Stateful firewall

Conntrack

table full, dropping packet

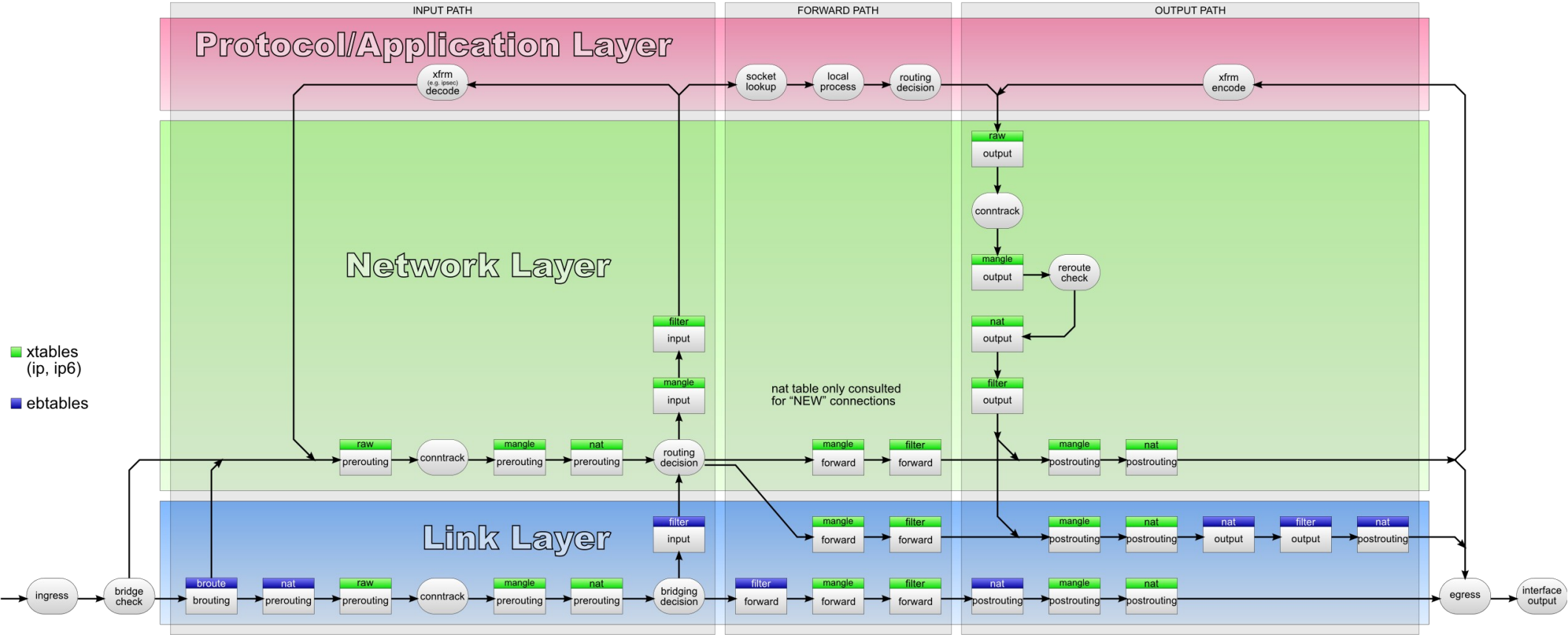
# Elcap firewall service

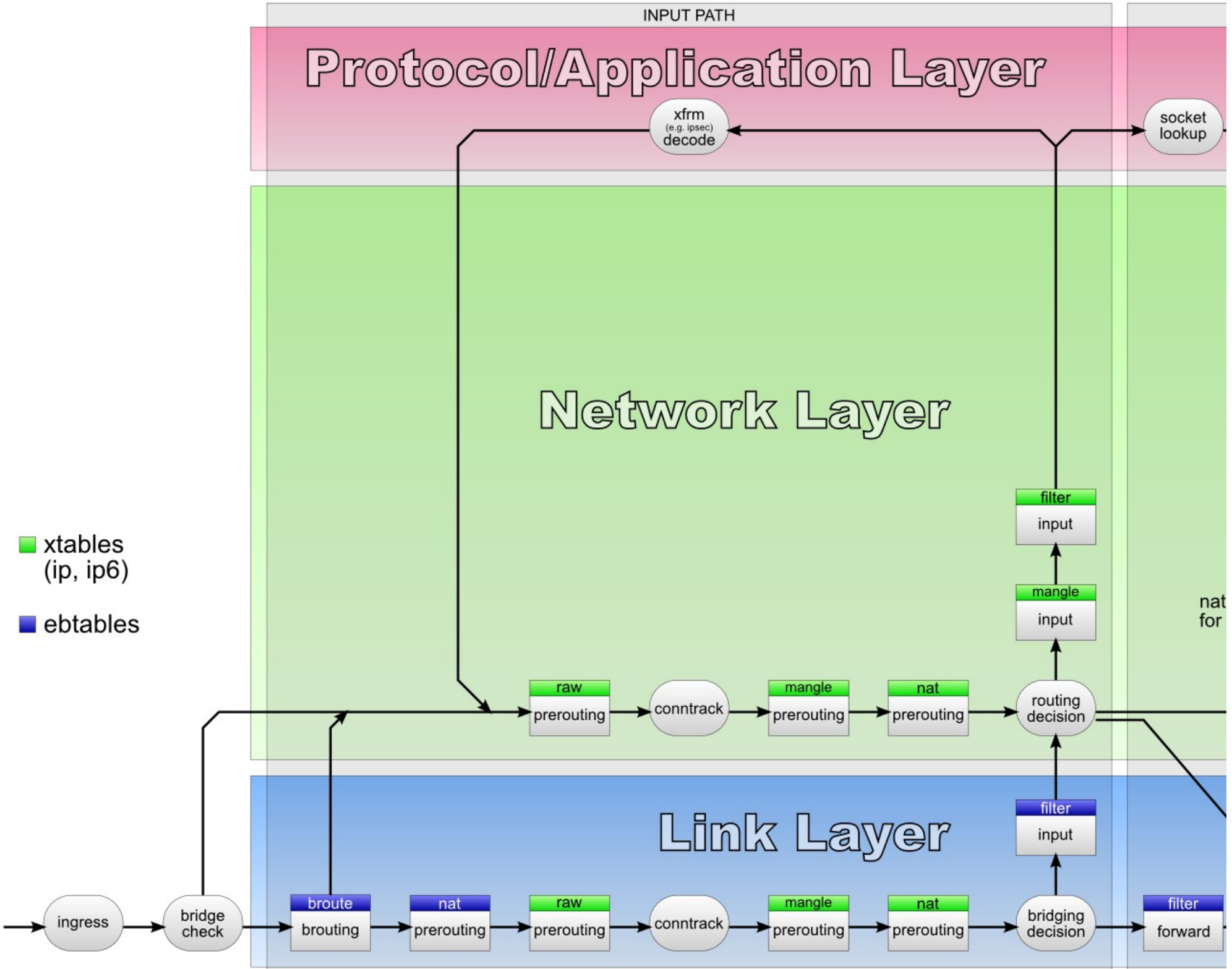
- Netfilter control hub
  - Geoip, ipset
  - Stateless firewall
  - Global ruleset
  - Network traffic meter
    - IDC내 트래픽 무료
    - 클라우드 솔루션 / 클러스터 솔루션 설치
  - lkvm gateway

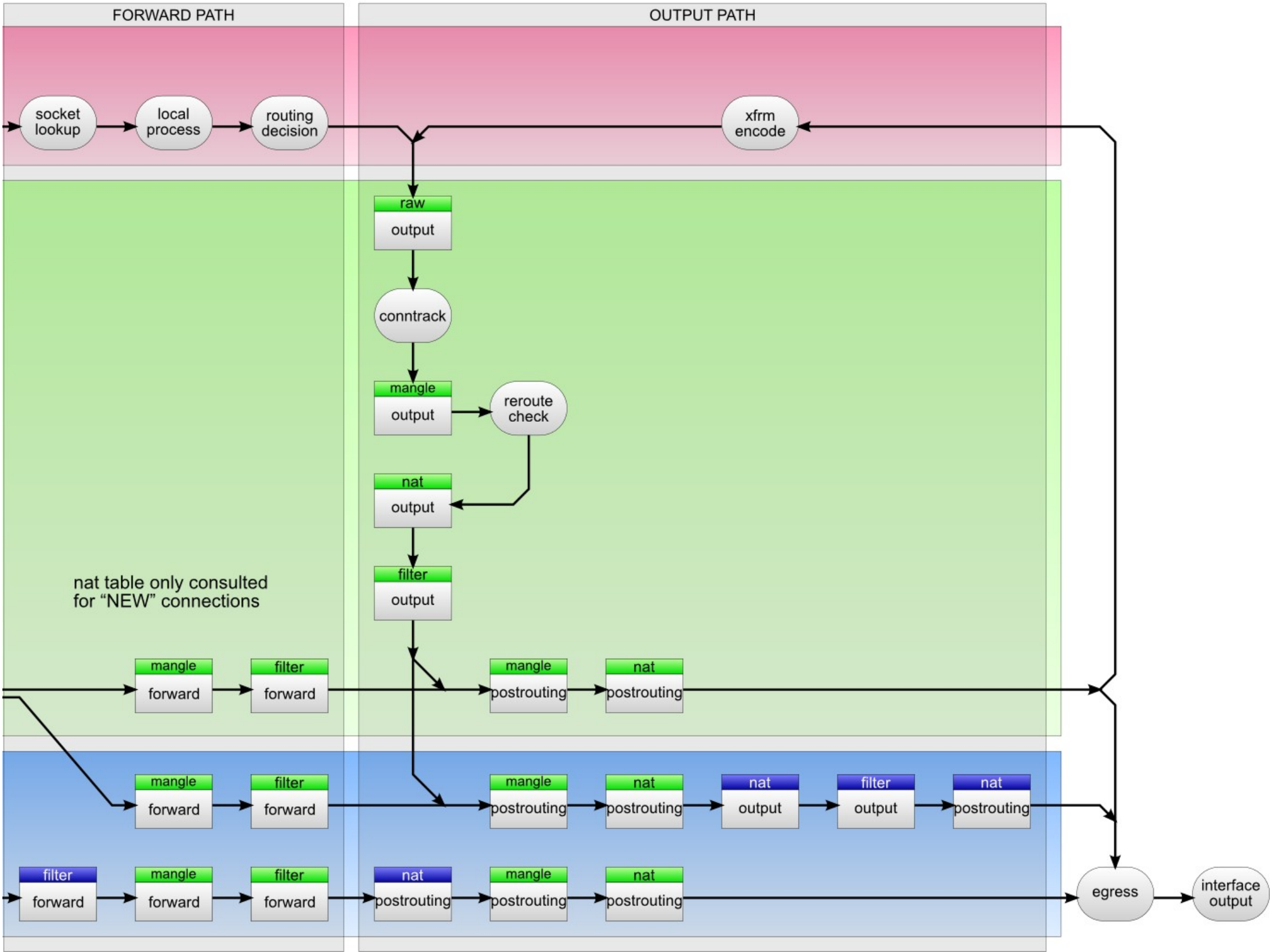
| Speed | Bits/sec | Bytes/second | Max PPS |
|---|---|---|---|
| 10Mbps | 10,000,000 | 1,250,000 | 14,881 |
| 100Mbps | 100,000,000 | 12,500,000 | 148,810 |
| 1Gbps | 1,000,000,000 | 125,000,000 | 1,488,095 |
| 10Gbps | 10,000,000,000 | 1,250,000,000 | 14,880,952 |

*by Jan Engelhardt, last updated 2009-11-27*
*based in part on Josh Triplett's graph*

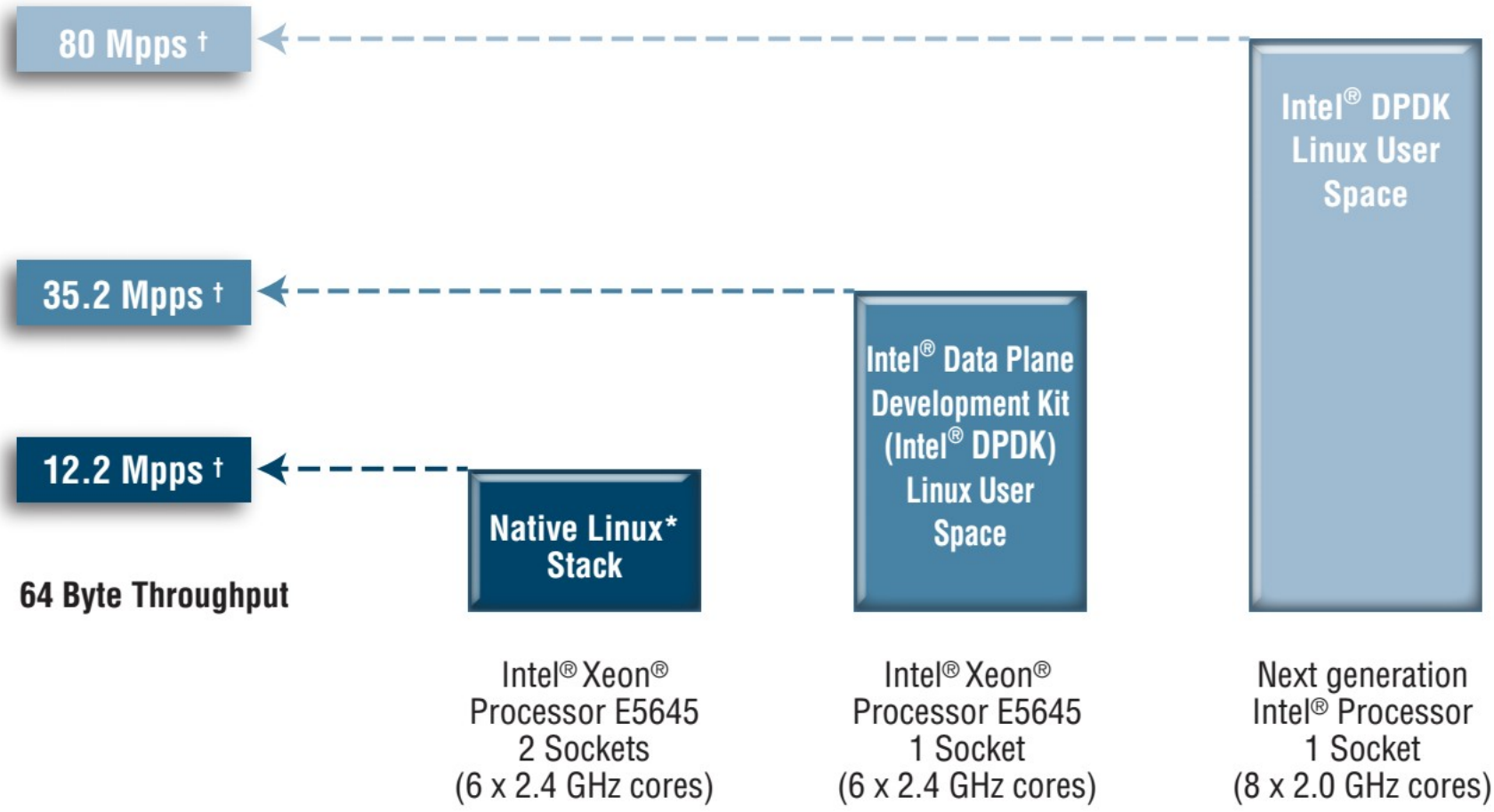
# *Netfilter packet flow; hook/table ordering*




INPUT PATH
Protocol/Application Layer
xfrm
(e.g. ipsec)
decode
socket
lookup
Network Layer
xtables
(ip, ip6)
ebtables
filter
input
mangle
input
nat
for
raw
prerouting
conntrack
mangle
prerouting
nat
prerouting
routing
decision
Link Layer
filter
input
ingress
bridge
check
broute
brouting
nat
prerouting
raw
prerouting
conntrack
mangle
prerouting
nat
prerouting
bridging
decision
filter
forward


FORWARD PATH
OUTPUT PATH
socket lookup
local process
routing decision
xfrm encode
raw
output
conntrack
mangle
output
reroute check
nat
output
filter
output
nat table only consulted for “NEW” connections
mangle
forward
filter
forward
mangle
postrouting
nat
postrouting
mangle
forward
filter
forward
mangle
postrouting
nat
postrouting
nat
output
filter
output
nat
postrouting
filter
forward
mangle
forward
filter
forward
nat
postrouting
mangle
postrouting
nat
postrouting
egress
interface output

# Packet Fowarding




† Measured performance. IO Performance may vary with platform configuration

†† Estimates are based on internal analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

**Figure 2.** Breakthrough Data Plane Performance with Intel® Data Plane Development Kit (Intel® DPDK) L3 Packet Forwarding

# Intel DPDK

- Netfilter ?
  - DPDK is a user-mode tcp/ip stack and replaces the entirety of the kernel stack including netfilter and iptables.
- Intel DPDK + OVS
  - Requirements
    - Intel Xeon Processor E5 Family  / Intel Atom Porcessor C2000 Family
  - Intel DPDK 1.7.1
  - https://github.com/01org/dpdk-ovs
  - https://wiki.linaro.org/LNG/Engineering/OVSDPDKOnUbuntu
- Netvm

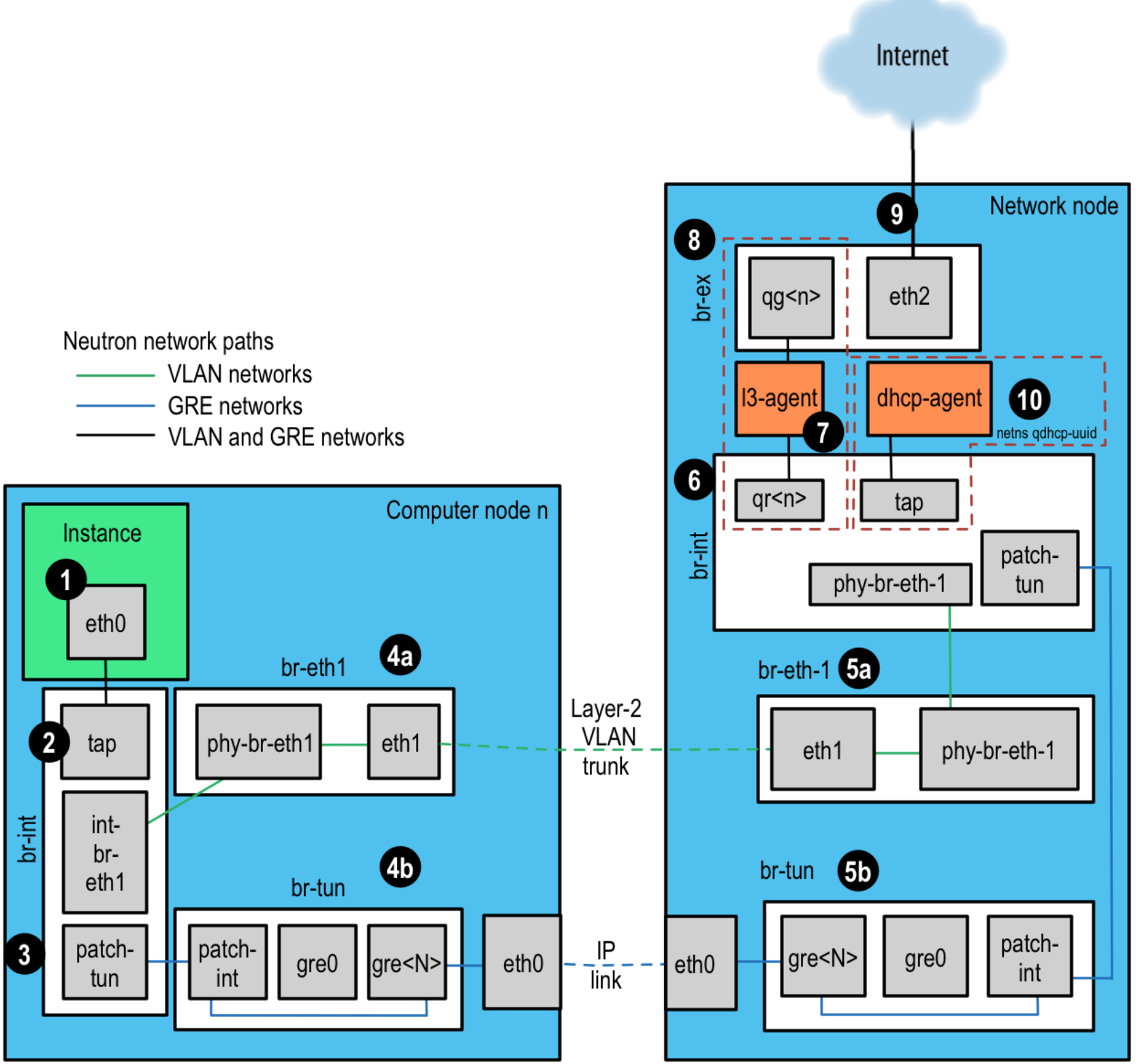
https://www.usenix.org/conference/nsdi14/technical-sessions/presentation/hwang

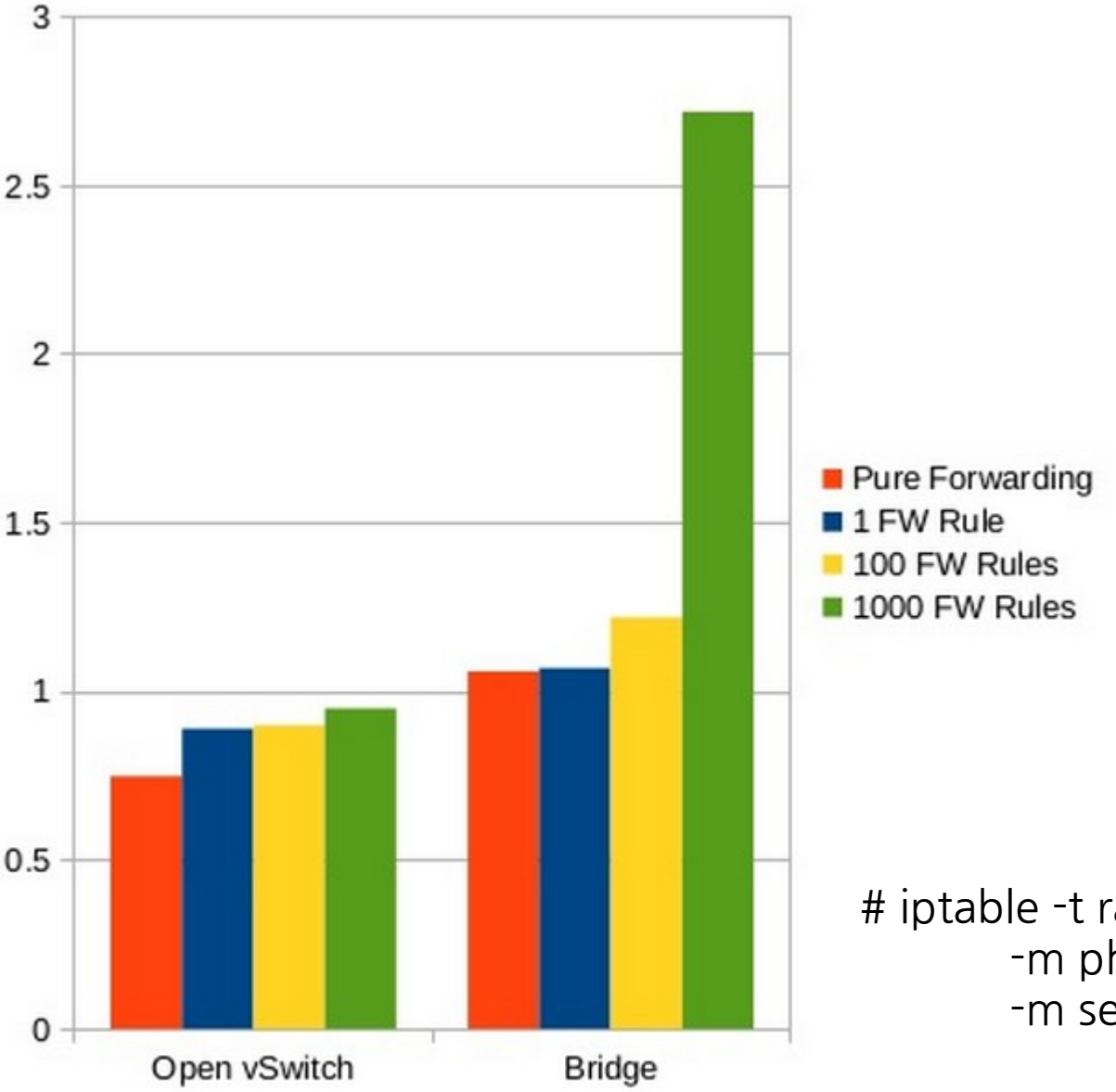
# netflow vs sflow

- Fprobe-ng
- Packet sampling => pseudo sflow
  - # -A  -m  statistic - - mode nth 10 -j ULOG
  - # apt-get install fprobe-ng

# Openstack, Linux, Netfilter

# DDOS Attack


Internet
Network node
9
8
br-ex
qg<n>
eth2
l3-agent
dhcp-agent
10
netns qdhcp-uuid
7
6
qr<n>
tap
br-int
patch-tun
phy-br-eth-1
br-eth-1
5a
eth1
phy-br-eth-1
br-tun
5b
eth0
gre<N>
gre0
patch-int
IP link
Layer-2 VLAN trunk
Neutron network paths
VLAN networks
GRE networks
VLAN and GRE networks
Computer node n
Instance
1
eth0
br-eth1
4a
2
tap
phy-br-eth1
eth1
br-int
int-br-eth1
4b
br-tun
3
patch-tun
patch-int
gre0
gre<N>
eth0




Number of gigacycles per second required to saturate a 10Gbps link with netperf TCP_STREAM. (Lower is better)

OVS + Conntrack

VS

Bridge + iptables

```
# iptable -t raw -I PREROUTING
          -m phydev -physdev-in eth1
          -m set  --set BLACK_HOLE src -j DROP
```

http://www.slideshare.net/ThomasGraf5/ovs-41973875
https://github.com/justinpettit/ovs/tree/conntrack

# Openstack and arp posioning

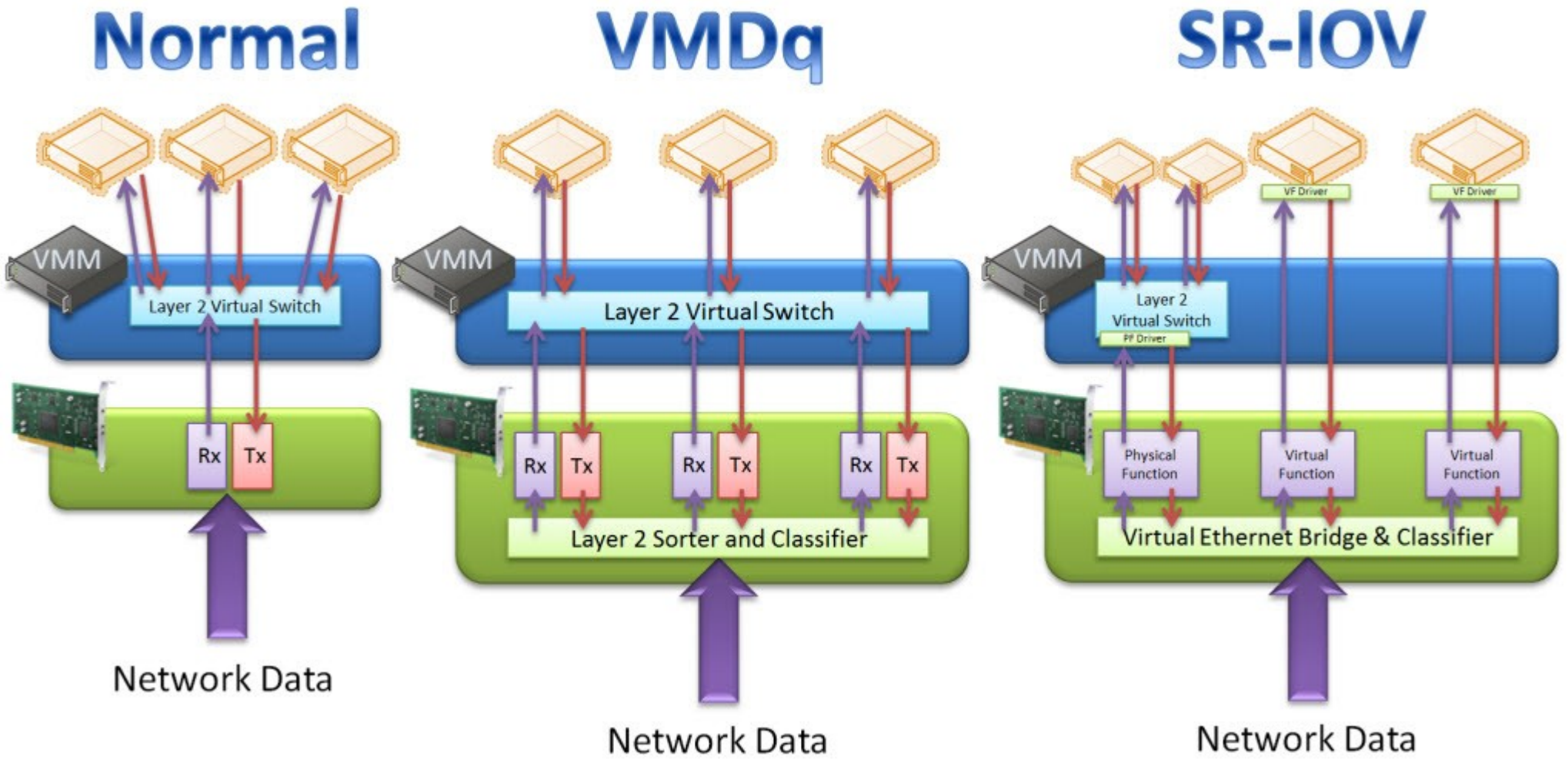
. ebtables manager

Netfilter

ebtables -A FORWARD -p IPv4 --ip-src 172.16.1.4 -s ! 00:11:22:33:44:55 -j DROP

iptables -A FORWARD -s 172.16.1.4 -m mac ! --mac-source 00:11:22:33:44:55 -j DROP

https://github.com/openstack/neutron-specs/blob/master/specs/kilo/arp-spoof-filtering-ebtables.rst
https://launchpadlibrarian.net/164709417/iptables_firewall_basic_spoofing.patch




Intel VT-d
AMD-Vi

CPU Freescale i.MX6 1 GHz Quad Core
Memory 1 Gybtes
Flash  4 Gbytes
LAN : 10/100/1000 Mbps
Power : POE 전용 / 802.3AF
OS : Ubuntu linaro 11.x / 12.04
Kernel : 3.0.35.custom

# Nework namespace

- CLONE_NEWNET
- CONFIG_NET_NS since LINUX 2.6.29
- Separate network stack
  - network addresses
  - nftables/netfilter rules
  - loopback interface for name space
- veth intreface(CONFIG_VETH), ip netns


Install Guides
Operations And Administration Guides
Cloud Administrator Guide
High Availability Guide
Operations Guide
Security Guide
Virtual Machine Image Guide
Architecture Design Guide
Configuration Guides
Architecture Design Guide
Configuration Reference
Cloud Administrator Guide
High Availability Guide
Operations Guide
Security Guide
Virtual Machine Image Guide
User Guides
API Quick Start
End User Guide (includes Python SDK)
Admin User Guide
Command-Line Interface Reference
Open source software for application development

SMILESERV

KOREA NUMBER ONE SERVICE

www.smileserv.com | www.1000dedi.net | www.bighdd.com


# 감사합니다.