DPDK
DATA PLANE DEVELOPMENT KIT

THE LINUX FOUNDATION

DPDK SUMMIT CHINA 2017

主办方： intel

参与方： 腾讯云 ZTE 美团云 Panabit 太一星晨 Balance Your Networks UnitedStack 有云 云杉网络 Yunshan Networks

协办方： SDNLAB 专注网络创新技术 视频支持方： IT大咖说

DPDK
DATA PLANE DEVELOPMENT KIT

THE LINUX FOUNDATION

Practice of Network Monitoring and Security Technologies in Cloud Data Center

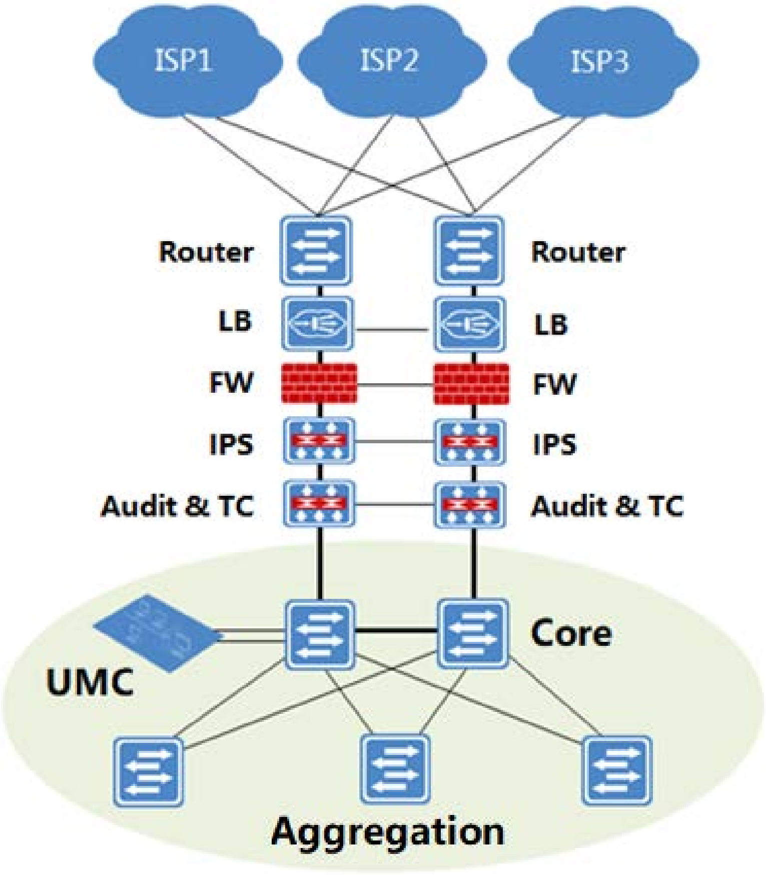
Kai, Wang
YunShan Networks

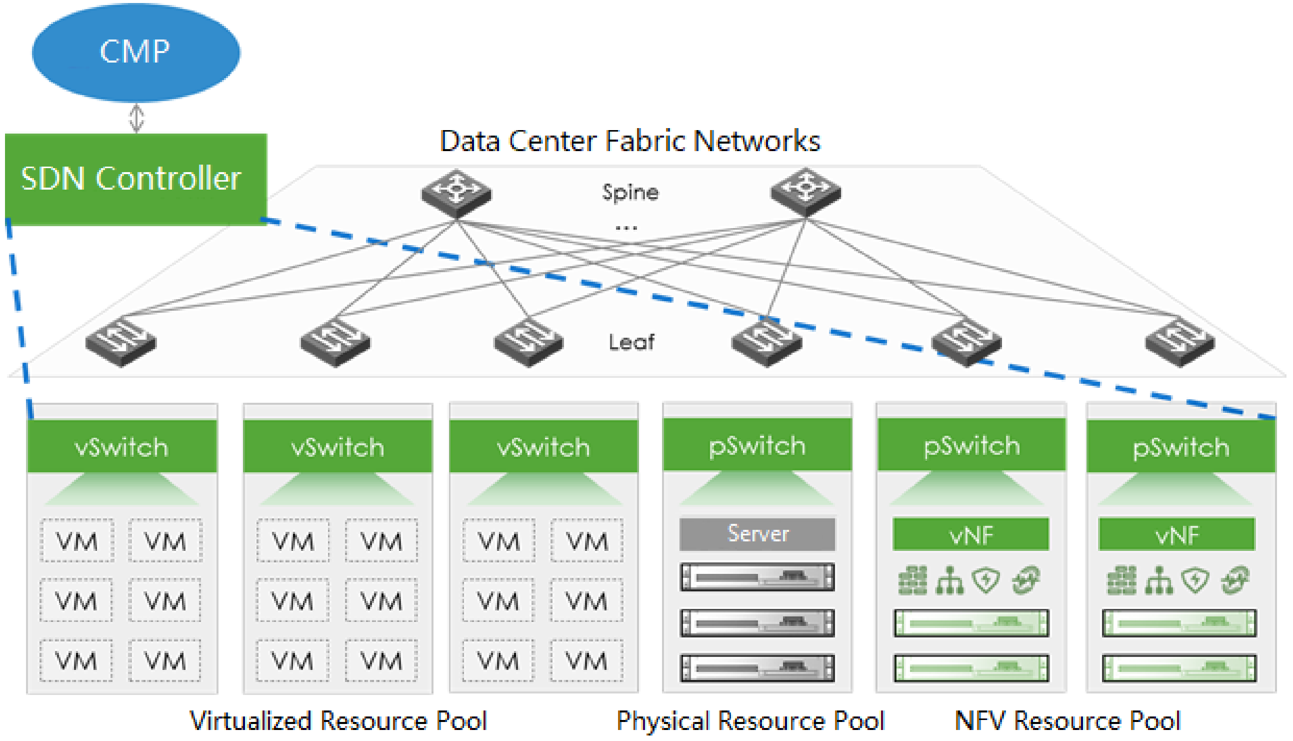
主办方： intel

参与方： 腾讯云 ZTE 美团云 Panabit 太一星晨 云杉网络

协办方： SDNLAB　　视频支持方： IT大咖说

Data center is evolving to be cloud based and software defined

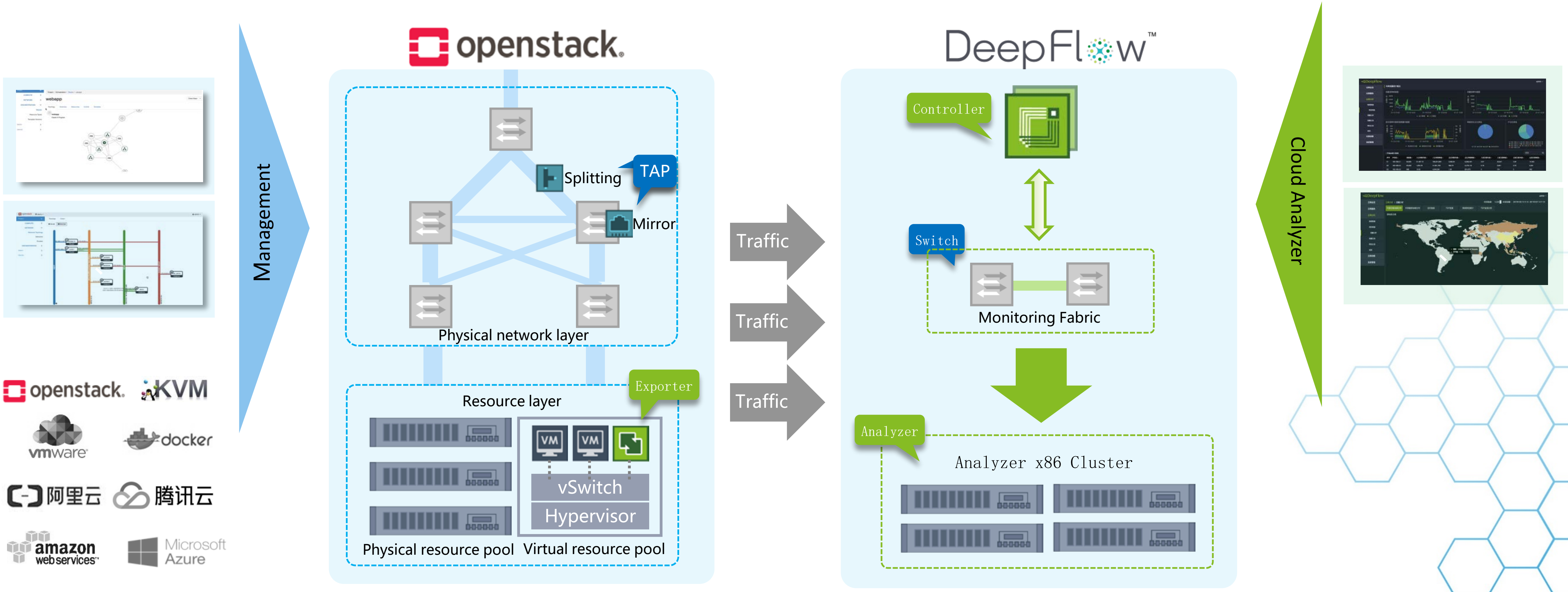




The monitoring and security problems in SD-CDC

- **The logical topologies become more and more complex**
  - Difficult to quickly find and locate the network problems in the tenant business
- **The collection of network data is inefficient**
  - Netflow/sFlow/IPFIX: Sampling, per-packet interrupt & netlink upcall
  - Limited variety of supported fields for collected flows
- **The analysis of overlay traffic is insufficient**
  - Unable to do flexible & find-grain traffic collection on demand
  - Unable to distinguish duplicated traffic from multiple tenants
  - Unable to effectively aggregate the overlay packets in tunnel capsulation and IP fragments
- **The physical boundaries of network security disappear**
  - Zero trust for the nodes in internal network

The monitoring solution



Management

openstack

Splitting

TAP

Mirror

Physical network layer

Exporter

Resource layer

VM VM

vSwitch

Hypervisor

Physical resource pool

Virtual resource pool

Traffic

Traffic

Traffic

DeepFlow™

Controller

Switch

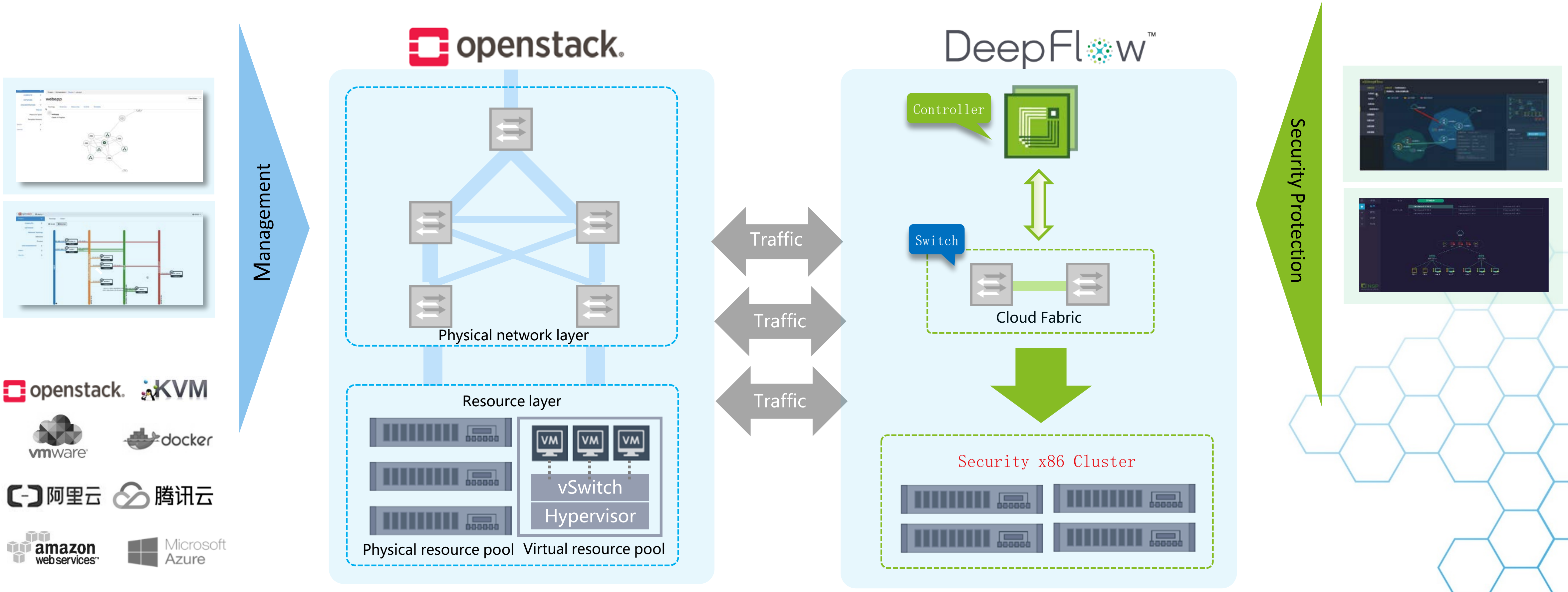
Monitoring Fabric

Analyzer

Analyzer x86 Cluster

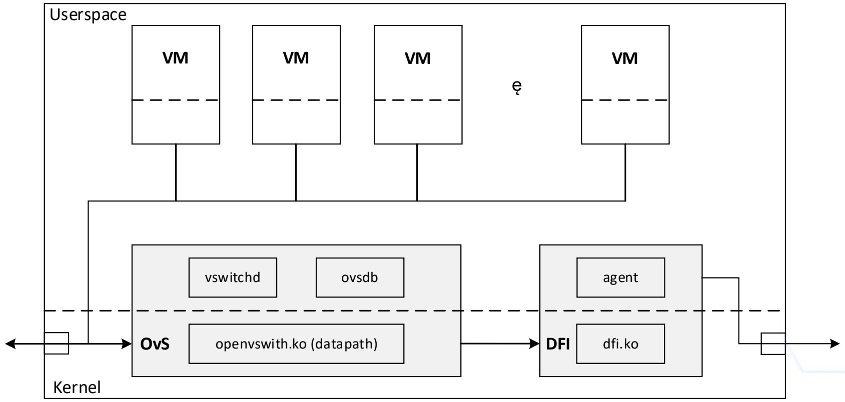
Cloud Analyzer

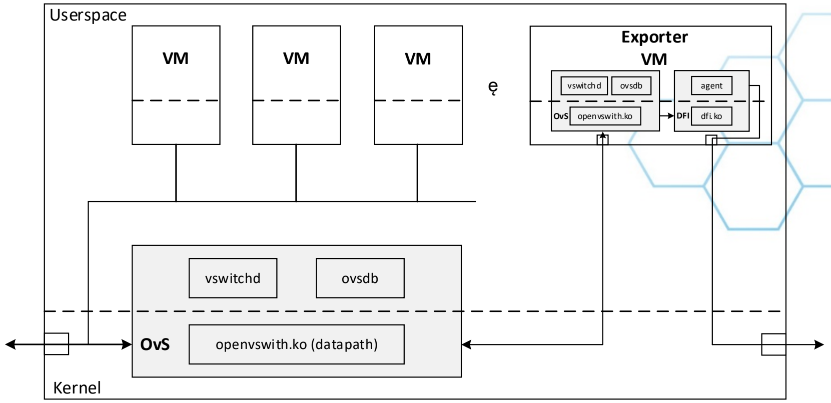
The security solution



Technology evolution for virtualized networks monitoring

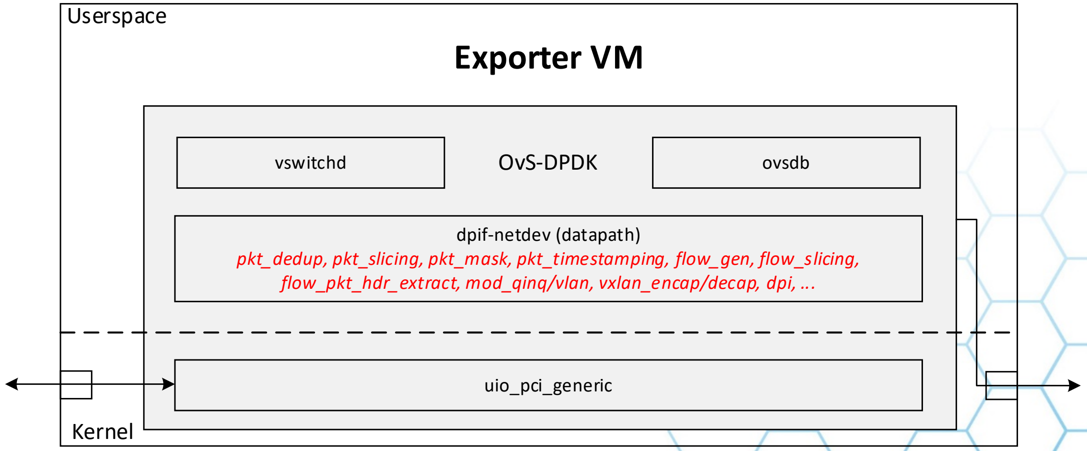
- Our solution: hypervisor based DFI (Deep Flow Inspection)
  - Probe utilizing OvS in Hypervisor
  - Overlay traffic collection
  - Kernel module + Userspace agent + OvS action
  - Cons: invasive deployment
    - Stability Problems: crash, soft lockup
    - Influence to tenant business
- Our solution: VM based DFI
  - Deployed in VM
  - Mirror overlay traffic to VM
  - Performance bottleneck





Technology evolution for virtualized networks monitoring

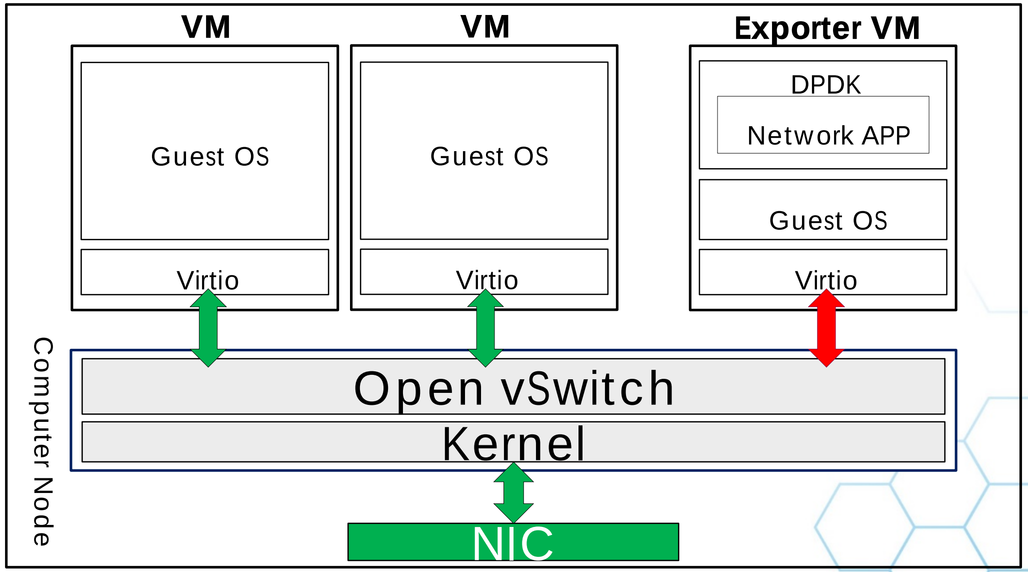
- Our current solution: DPDK based
  - Utilizing OvS-DPDK
    - Fully exploit the compute resource of VM
  - Extend functions based on OvS-DPDK conntrack
    - ACL
    - Flow generation
    - Packet header extraction and compression
    - DPI
    - NPB
  - SDN
  - More efficient, flexible, benefit for debug
  - Used for physical networks monitoring as well



Userspace
Exporter VM
vswitchd
OvS-DPDK
ovsdb
dpif-netdev (datapath)
pkt_dedup, pkt_slicing, pkt_mask, pkt_timestamping, flow_gen, flow_slicing, flow_pkt_hdr_extract, mod_qinq/vlan, vxlan_encap/decap, dpi, ...
uio_pci_generic
Kernel

Further optimization for exporter

- NIC Multi-queue & Symmetric RSS
  - VM template
- Parallelize conntrack processing
  - Make it scalable
- Optimize the datapath classifier (dpcls) algorithm Tuple Space Search (TSS)
  - HyperSplit algorithm
- Intel vTune Amplifier
  - Lock, Polling & Interrupt

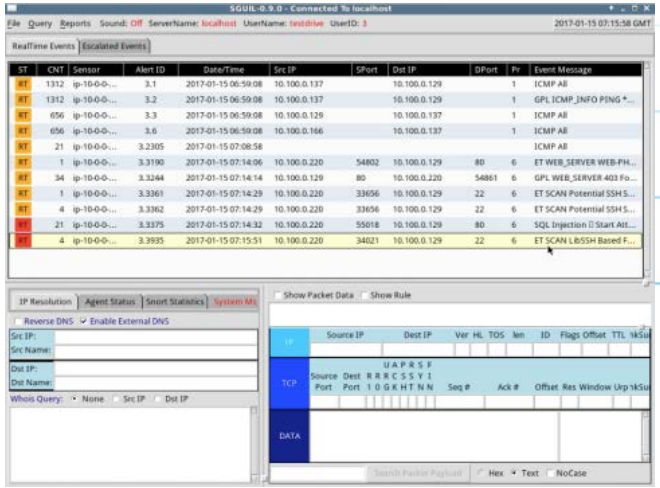


Analysis & Visualization

- Cluster-based analyzer
  - Use Storm to do real-time analysis
    - DDoS/Port Scan
    - Abnormal connections/transactions, Abnormal login
    - ARP/MAC/IP Spoof
    - Loop detection
  - Use Spark to do off-line analysis
    - Security analysis model
  - Use ElasiticSearch/Kibana to do search and visualization
    - Customized statistics in different dimensions
    - Trace back of historical events
- Third-party analysis tool
  - E.g. SQUIL, SQL injection detection

177.109.115.253 > 122.115.40.66
protocol : TCP
BPS : 428302
PPS : 2580
tcp flags : SYN,RST,ACK
server ports : 21-8080
client count : 2130

122.115.40.65 > 7.7.7.0
protocol : IP
min ip : 7.7.7.0
max ip : 7.7.7.255
time start : 2015-05-18 09:52:49
time end : 2015-05-18 09:59:28
server count : 37
flowscount : 37



SGUIL-0.9.0 - Connected To localhost

File Query Reports Sound: Off ServerName: localhost UserName: testdrive UserID: 3 2017-01-15 07:15:58 GMT

RealTime Events | Escalated Events

ST	CNT	Sensor	Alert ID	Date/Time	Src IP	SPort	Dst IP	DPort	Pr	Event Message
RT	1312	ip-10-0-0-...	3.1	2017-01-15 06:59:08	10.100.0.137		10.100.0.129		1	ICMP All
RT	1312	ip-10-0-0-...	3.2	2017-01-15 06:59:08	10.100.0.137		10.100.0.129		1	GPL ICMP_INFO PING *...
RT	656	ip-10-0-0-...	3.3	2017-01-15 06:59:08	10.100.0.129		10.100.0.137		1	ICMP All
RT	656	ip-10-0-0-...	3.6	2017-01-15 06:59:08	10.100.0.166		10.100.0.137		1	ICMP All
RT	21	ip-10-0-0-...	3.2305	2017-01-15 07:08:58						ICMP All
RT	1	ip-10-0-0-...	3.3190	2017-01-15 07:14:06	10.100.0.220	54802	10.100.0.129	80	6	ET WEB_SERVER WEB-PH...
RT	34	ip-10-0-0-...	3.3244	2017-01-15 07:14:14	10.100.0.129	80	10.100.0.220	54861	6	GPL WEB_SERVER 403 Fo...
RT	1	ip-10-0-0-...	3.3361	2017-01-15 07:14:29	10.100.0.220	33656	10.100.0.129	22	6	ET SCAN Potential SSH S...
RT	4	ip-10-0-0-...	3.3362	2017-01-15 07:14:29	10.100.0.220	33656	10.100.0.129	22	6	ET SCAN Potential SSH S...
RT	21	ip-10-0-0-...	3.3375	2017-01-15 07:14:32	10.100.0.220	55018	10.100.0.129	80	6	SQL Injection Start Att...
RT	4	ip-10-0-0-...	3.3935	2017-01-15 07:15:51	10.100.0.220	34021	10.100.0.129	22	6	ET SCAN LibSSH Based F...

IP Resolution | Agent Status | Snort Statistics | System Ms

□ Reverse DNS ☑ Enable External DNS

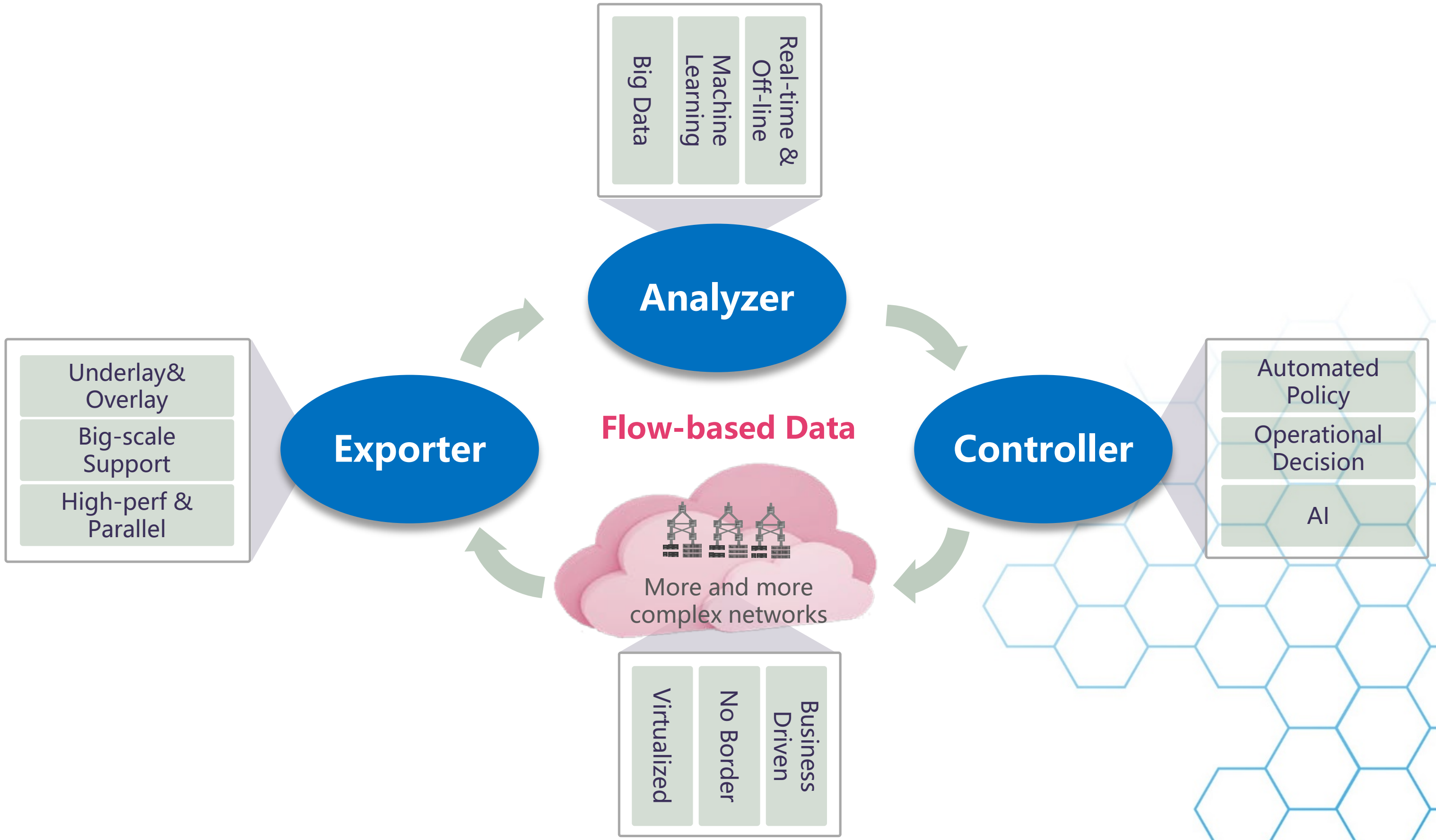
Src IP:
Src Name:
Dst IP:
Dst Name:

Whois Query: ● None ○ Src IP ○ Dst IP

□ Show Packet Data □ Show Rule

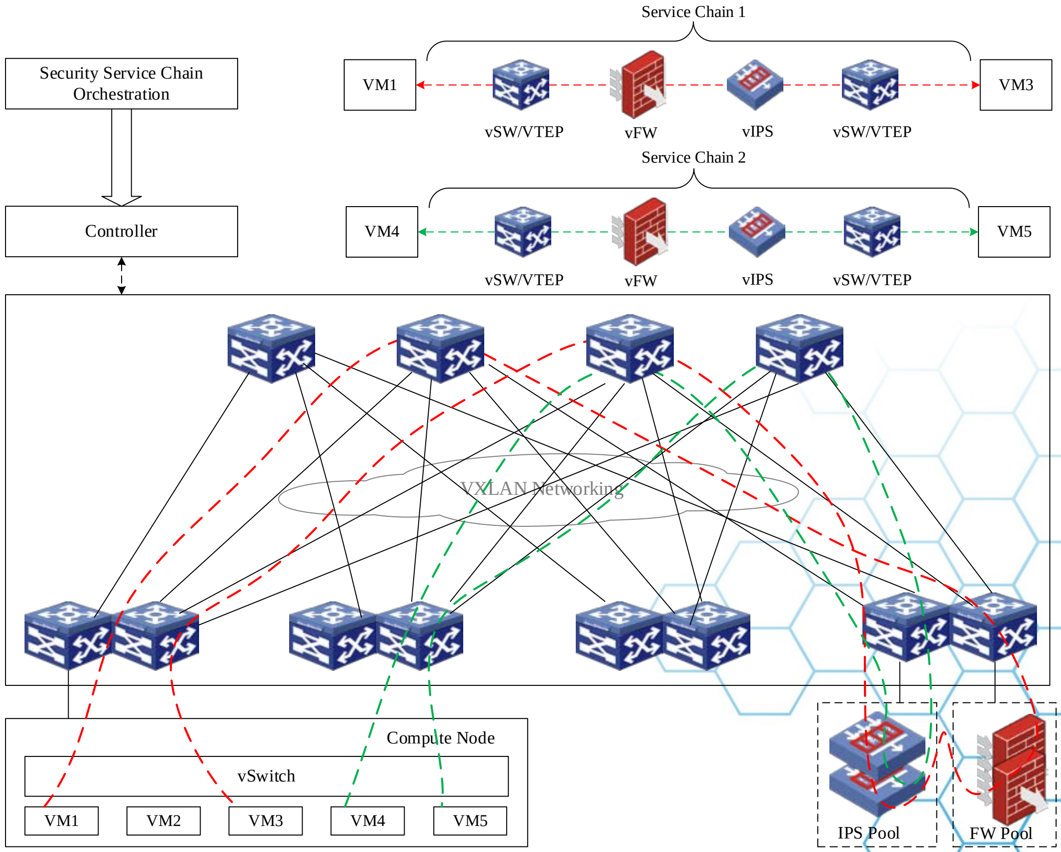
From monitoring to security control

- Use the monitoring results to generate security policies
  - Exporter
    - Overview the security problems & risks in cloud networks
  - Analyzer
    - Locate the problematic nodes or areas
  - Controller
    - Prevent/Protect these nodes or areas via SDN



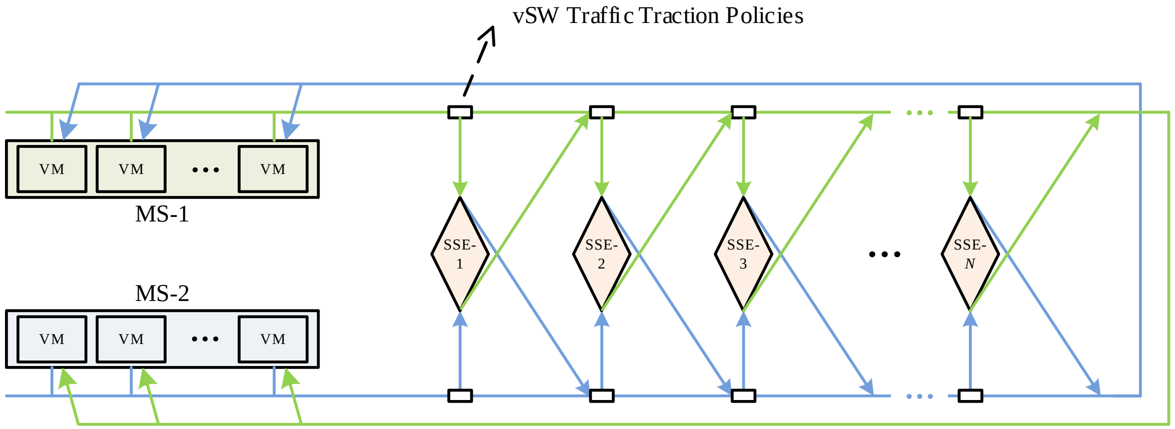
Security service chain and problems

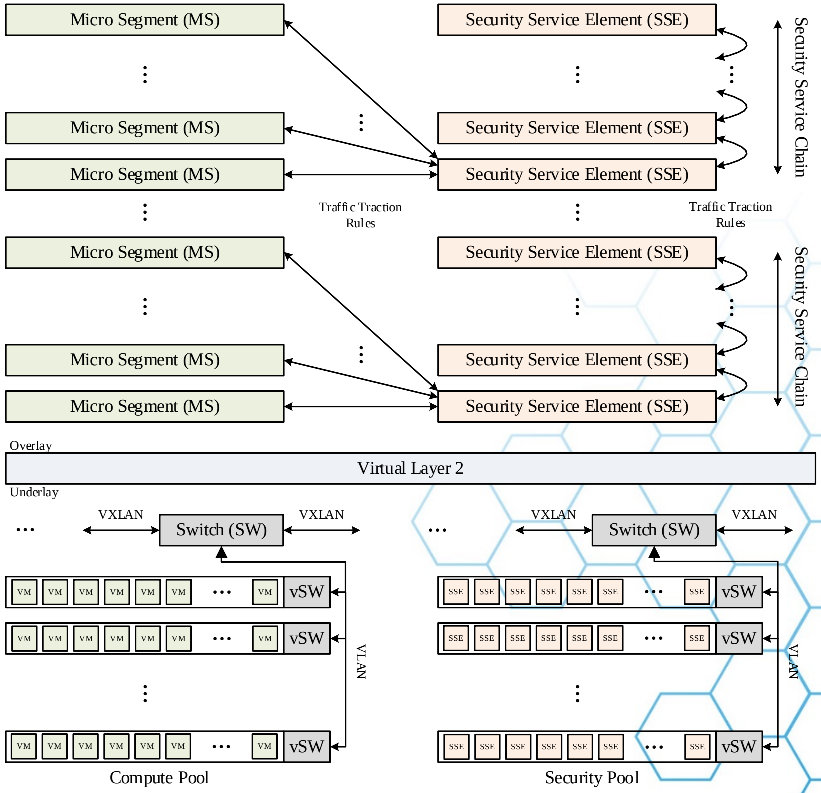
- Use VNF to do security detection/prevention
  - Based on VXLAN
- Pros
  - Elastic and flexible
- Cons
  - Inefficient and low-performance, hard to cover the large-scale east-west traffic
    - VXLAN encap/decap load
    - Poor scalability of security service chain
    - vSwitch and VNF performance bottlenecks



Performance optimization

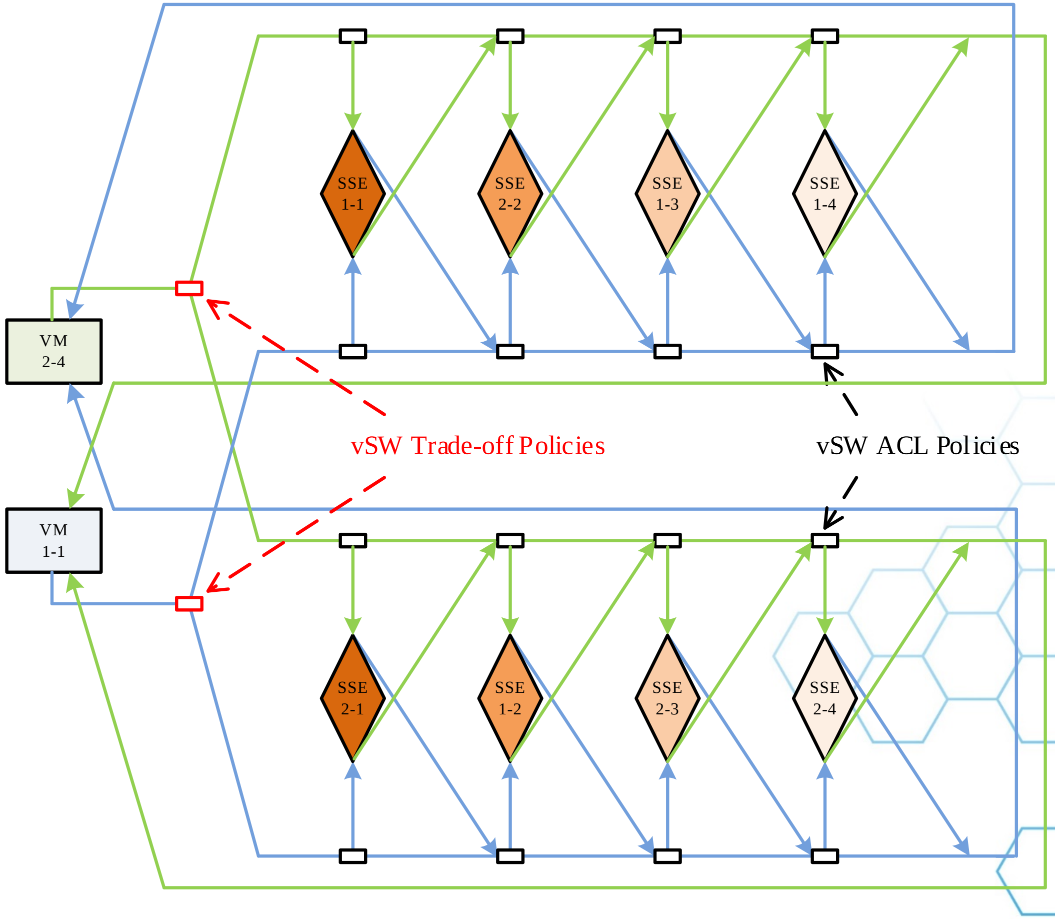
- Use VLAN instead of VXLAN to introduce traffic to assigned security nodes
  - Offload VXLAN encap/decap to ToR switch, save more CPU for SSE processing
    - table=0,priority=202,dl_vlan=2000,ip,actions=output:20
    - table=0,priority=102,in_port=10,dl_vlan=0xffff,ip,actions=mod_vlan_vid:2000,resubmit(,0)





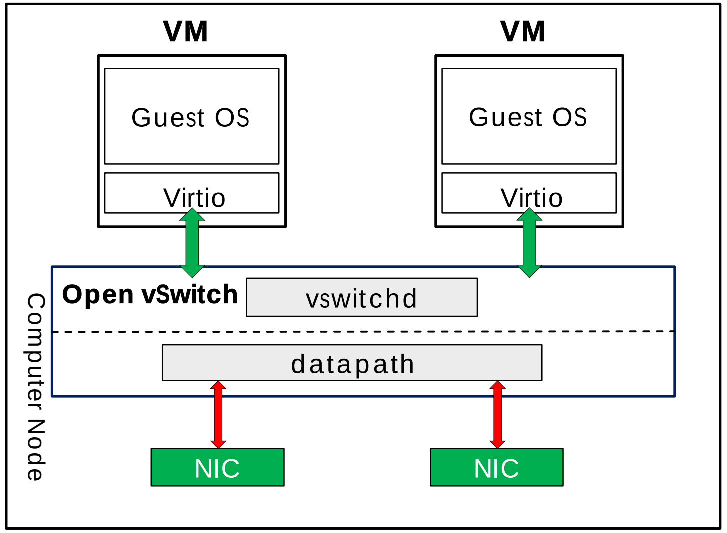
Performance optimization

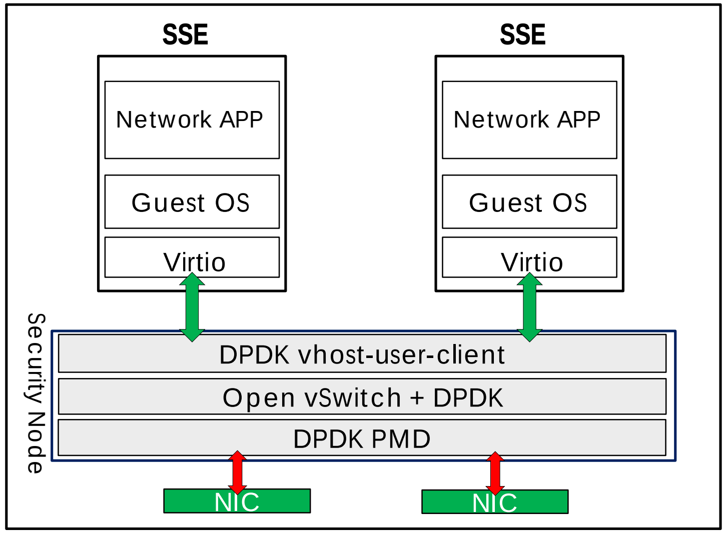
- Single VNF/SSC has limited performance
- Use SDN policies based trade-off to dispatch traffic to multiple chains
  - Based on pseudo node
  - Linearly increase the performance
- E.g.
  - priority=401,table=0,dl_vlan=1000,ip,tcp, tp_src=0/0x0001,tp_dst=0/0x0001,actions =mod_vlan_vid:2000,resubmit(,0)
  - priority=401,table=0,dl_vlan=1000,ip,tcp, tp_src=1/0x0001,tp_dst=1/0x0001,actions =mod_vlan_vid:2000,resubmit(,0)
  - priority=401,table=0,dl_vlan=1000,ip,tcp, tp_src=0/0x0001,tp_dst=1/0x0001,actions =mod_vlan_vid:3000,resubmit(,0)
  - priority=401,table=0,dl_vlan=1000,ip,tcp, tp_src=1/0x0001,tp_dst=0/0x0001,actions =mod_vlan_vid:3000,resubmit(,0)

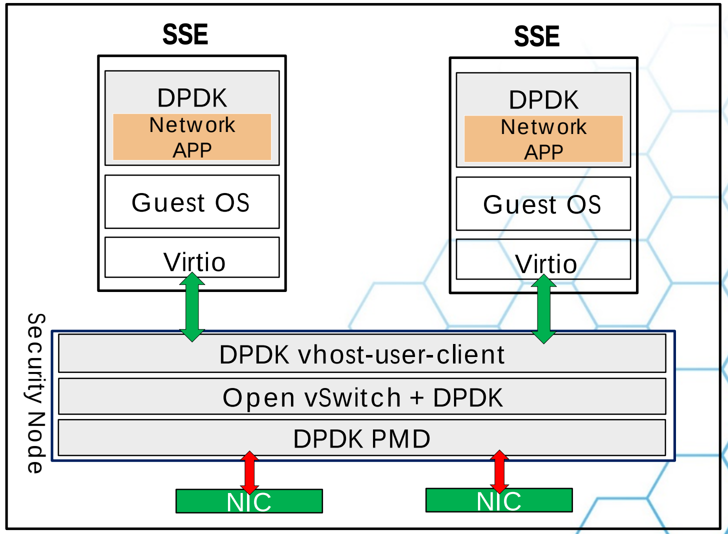


Performance optimization

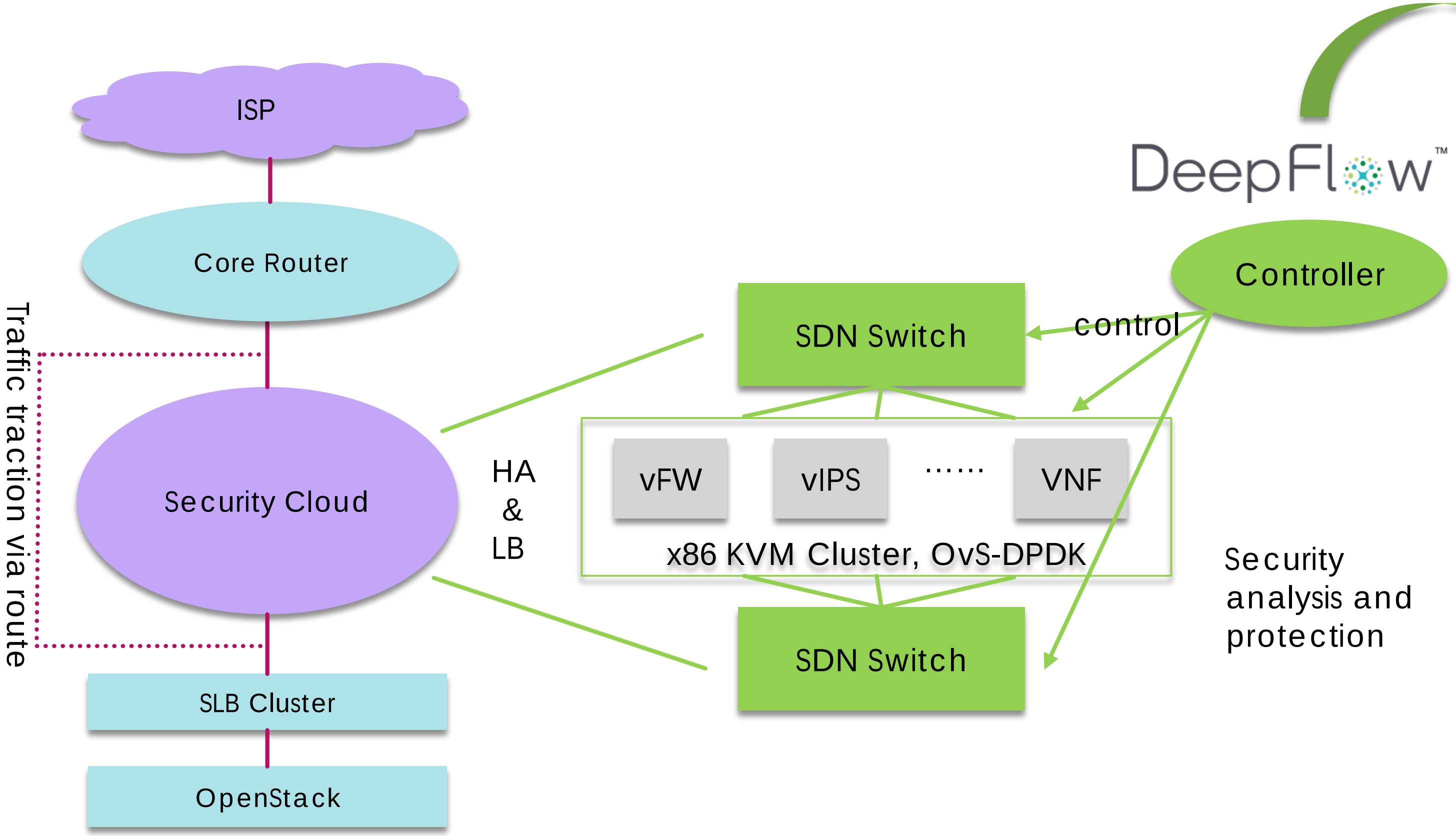
- Use OvS-DPDK to accelerate the networking in security resource pool
- Use DPDK to accelerate SSE
  - TOPSEC

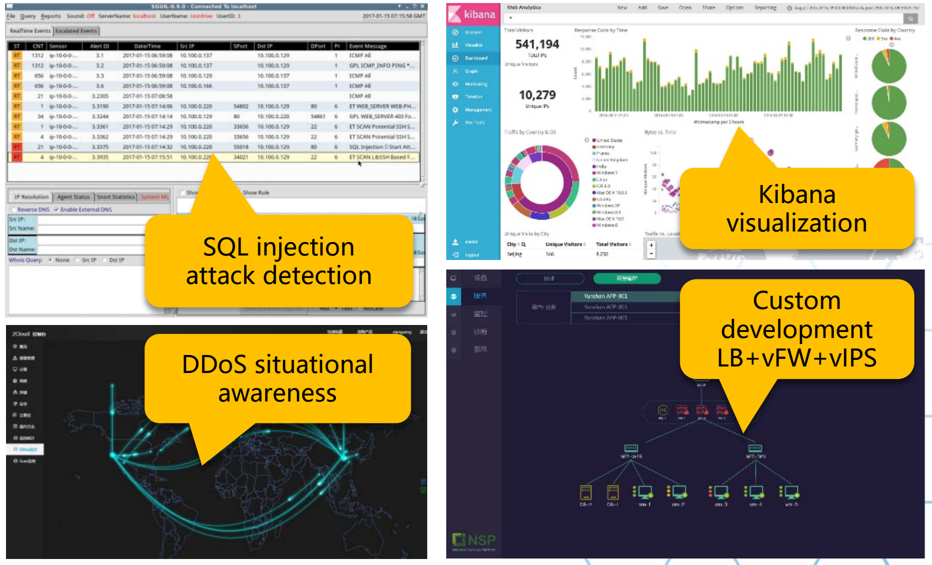






Security cloud




SQL injection attack detection
Kibana visualization
DDoS situational awareness
Custom development LB+vFW+vIPS
kibana
541,194
10,279
NSP

Thanks!!

欢迎关注**DPDK开源社区**