High performance Cloud with Hardware Acceleration

Cloud BU, Huawei
The world is changing - more devices, more conns, more data

- Billion
- Tens of Billions
- 100 ~ 1000 of Billions

Desktop Internet
- The PC brought the internet access to billions, during the 1990s

Mobile Internet
- The Mobile Revolution put computing and instant access to information in the hands of billions, in the early 2000s

Internet of Things
- Now!
Moore’s law is slowing

- Single-threaded perf: 1.1x per year
- Memory VS CPU: 1.5x per year

Keeping Moore’s law alive in Data Center

- Customized CPU
- Smart NIC
- Cloud SSD
- SDI (Service Driven Infrastructure)
- KunLun - Powerful X86 Server (768 cores, 32T)
- Heterogeneous computing platform
Host-based SDN:
separate a centralized control plane from a data plane on the host, and implement almost all virtual networking features, to connect VMs to physical network.
History of Huawei eVS

eVS – Elastic Virtual Switch

**eVS 1.0**
- Start from 2013
- Kernel-based OVS
- Enhanced features (CT/BUM)
- 8Gbps/900Kpps

**Limited by Kernel**

**eVS 2.0**
- Start from 2015
- DPDK
- General Packet Filtering Platform
- Optimized SIMD and Multi-threaded scheduling
- 20Gbps/5Mpps

**Limited by CPU frequency, memory bandwidth**
How to break the bottleneck of vSwitch?

Our way:
Combination of software and self-developed hardware

And additional benefits:
flexibility, high performance, low cost, high availability
eVS 3.0 - First Tens of Millions PPS virtual network switch

40Gbps 10Mpps

- Hardware Acceleration
- High Performance
- High Availability

Smart NIC Offloads
Virtio-Direct
Hot upgrade
Huawei Smart NIC: u-NIC

Self-developed smart chip: u-NIC

- RDMA
- TCP
- NIC
- OVS
- VxLAN
- PCIe, SR-IOV, Virt-IO
- Intelligent network protocol processor
- QoS
- Security
- Application
- Ethernet
- HD Memory
- p-Switch

Features

- Programmable high performance packet forwarding platform
- Network-specific optimization engine (PPE)
- Large flow table and security rules
- High precision hard QoS
Architecture: Integrated Flow Table and Offload

- eVS data plane
- GFP
- ovs datapath
- upcall
- ext action
- Offload Engine
- Flow Table (IFE)
- u-NIC
- virtio-net
- virtio-direct
- VM
Virtio-Direct: Virtual I/O

- **High performance virtio data path**
  - Offload virtio head
  - Zero Copy
  - IRQ Aggregation

- **High Availability, Smooth migration**
  - Standard virtio-net
  - Non-intrusive GuestOS
  - Support Live migration
Hot upgrade, both Hardware and Software

**No user impact**

Active/Standby Mode
Hot upgrade/downgrade
Independent hot upgrade
Joint hot upgrade
# Flexibility, Performance, Availability

<table>
<thead>
<tr>
<th></th>
<th>Virtio-Direct</th>
<th>Other Smart NIC SR-IOV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Virtual I/O Mode</strong></td>
<td>Software &amp; Hardware</td>
<td>Hardware</td>
</tr>
<tr>
<td><strong>Performance</strong></td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Non-intrusive GuestOS</strong></td>
<td>✔️</td>
<td>❌</td>
</tr>
<tr>
<td><strong>Live migration</strong></td>
<td>✔️</td>
<td>❌</td>
</tr>
<tr>
<td><strong>Hot upgrade</strong></td>
<td>✔️</td>
<td>❌</td>
</tr>
</tbody>
</table>
SDI
SDI

Release more CPUs to user
High perf on basic services:
storage and network
Based on self-developed chip
• Resource Competition
• 20+% Cores wasted
• Weak reliability

• Resource isolation
• 100% Cores for users
SDI Architecture

Cloud Scale Application

VM VM ... VM VM VM ... VM
Hypervisor A Hypervisor B Bare Metal APP

SmartEP Driver @ PCIe

Management Storage Network Accelerator

Euler OS

Storage Interface Network Interface PCIe Interface

Software Accelerator Hardware Accelerator

Extension interface External accelerator

18
Customized CPU

- Single Core Integer \(25\%\uparrow^*\)
- Single Core Linpack \(70\%\uparrow^*\)
- Memory Bandwidth \(62\%\uparrow^*\)

Application

Customized Server

Data Center

\(^*\): 6151 vs 2680v4
C3ne - with u-NIC, Virtio-Direct, SDI

**General Network Enhancement C3ne ECS**

Latest-generation Intel Xeon SkyLake CPUs and high-speed smart Hi1822 NICs offer powerful and stable computing performance, ultra-high network bandwidth, and high Maximum PPS.

Applications:
- High Maximum PPS scenarios, such as on-screen video comments and telecom business forwarding
- Enterprise-class applications with high network requirements
- Small-and medium-sized databases, cache, and search clusters
- Data analysis and computing

Specifications:
- CPU/Memory ratio: 1:2:1:4
- vCPUs: 2-60
- Fundamental frequency/Turbo Boost: 3.0/3.4 GHz
- Maximum PPS: 10,000,000
- Maximum intranet bandwidth: 40 Gbit/s

2.5x↑

2x↑
Simultaneous online players from 30k to 13M

- The first listed company that offers game-accelerating services in the world.
- lower ping, no packet loss, no login failure

Before:

- PPS: 3Mpps
- 30k Player

now:

- PPS: 13Mpps
- 13M players
Thanks