At a Glance:

- Meituan Open Services deployed Open vSwitch* with the Data Plane Development Kit (OvS-DPDK) to improve manageability and performance over vanilla Open vSwitch (OvS) requiring only one CPU core for the vSwitch.
- The company minimized capex and opex by running the software-based solution on existing hardware.
- It addressed Windows guest, vHost reconnection, vHost live migration, and huge page memory use. This technical collaboration covered Open vSwitch, QEMU and DPDK projects.

Case study

Leading cloud service provider optimizes data center resources to boost performance and cut costs with Intel and OvS-DPDK

China-based Meituan.com is one of the world’s leading online-to-offline (O2O) local life service platforms. It connects over 240 million consumers and 5 million local merchants through a range of e-commerce services, and it processes over 10 million orders and deliveries each day. It also operates a public cloud platform, Meituan Open Services (MOS), which offers efficient and reliable cloud solutions to enterprise customers.

As its customer base continues to grow rapidly, MOS needed to optimize the performance and efficiency of its data center resources while keeping costs down. To support the growing demand on networking traffic for its public cloud services in China, it deployed Open vSwitch* with the Data Plane Development Kit (OvS-DPDK) as software to enable accelerated, high performance virtual switching, powered by Intel® Xeon® E5-2650 processor v2 and Intel® Ethernet Controller X540-AT2.

Challenge

- With customer numbers and service diversity growing fast, Meituan Open Services (MOS) needs to ensure its data center resources are optimized to create and drive value for the business and its customers.
- Inefficiencies in Standard Kernel-based OvS switching between applications in the virtual network meant too much time, money and valuable CPU resources were being spent on non-billable activities.
- MOS needed a more efficient network switching solution that would also enable a smooth evolution to future technologies across the data center.
Solution

- MOS benchmarked performance for key workloads using both a kernel-based vSwitch based on Open Virtual Switch (OVS) and a vSwitch running in User Space based on OvS-DPDK
- The environment tested was powered by Intel Xeon E5-2650 processor v2 with 128 GB of memory, and an Intel Ethernet Controller X540-AT2 for network connectivity
- The Open vSwitch version was 2.4.90. The virtual infrastructure was based on a kernel-based virtual machine (KVM) with QEMU v2.6.0. The evaluation was based on 64-byte packets
- The vSwitch running on MOS platforms supported a wide range of online applications with various operating parameters. The objective was to have an efficient software based vSwitch that could support these applications with minimal overhead and maximum flexibility

Results

- OvS-DPDK offered greater manageability and improved performance over vanilla OVS and only required one CPU core for the vSwitch
- The software-based solution runs on existing hardware, helping to minimize capex and opex while delivering improved function and quality of service
- It enhanced vSwitch stability by fixing bugs across OvS, QEMU and DPDK
- vHost-user port reinitialization was reduced to just two seconds

Monetizing the CPU

Like all cloud service providers (CSPs) MOS aims to monetize as much of its available CPU resources as possible by using them to run customer services. This means that it must minimize the amount of processing power needed to run non-billable activities, such as managing its network functions.

The company used Open vSwitch (OvS) software to automate management of network switching for various workloads from a packet generator sending packet sizes of 64K bytes across its virtual environments. However, it was dissatisfied with the level of performance this approach delivered, as it limited the number of virtual machines (VMs) possible per core, in turn impacting total cost of ownership (TCO).

Benchmarking the options

MOS ran a benchmark test to compare the existing CPU-based model and that enhanced by implementing OVS with the Data Plane Development Kit (OVS-DPDK) – see figure 1.

Its environment includes dual 10Gbps ports for uplink bonding (Intel Ethernet Controller X540-AT2) and is powered by single-core Intel Xeon E5-2650 processors v2, running at 2.6G and 128GB memory. The kernel used was version 3.10.0, and OVS version 2.4.90. A QEMU open source-hosted hypervisor v2.6.0 was also used.

Figure 1: Benchmark comparison of OvS and OvS-DPDK
By adding OvS-DPDK to its existing environment, MOS was able to move its vSwitch functions from the operating system kernel to the user space. This freed MOS from the limitations inherent in using the kernel for network switching. For example, as more VMs are added to an environment, a traditional OvS model running in the kernel will need to use more and more cores to achieve the necessary performance. However, OvS-DPDK accelerates the network transitions, delivering a larger number of packets per second¹ (see figure 2). This means that more VMs can be added for the same throughput (or the same number of VMs for less throughput). Meituan found this option worked particularly well and was a more cost-effective option than transitioning to smart NICs across the network.

MOS also achieved a reduction in switch re-initialization based on its customization of OvS-DPDK. This enabled standard OvS-DPDK to run within one process, with multi threads handling packet I/O and processing, exception/slow path, openflow rule refresh, and housekeeping. Meanwhile, the main thread is able to take care of DPDK initialization, management, probing and configuration. By separating the single process into two in this way, the reinitialization was reduced to just seconds.

Driving value for the business
When comparing results of the benchmarks, MOS found that with OvS-DPDK, manageability and performance were both improved, as it could run the virtual switch on just one core for its existing workloads. By fixing bugs across OvS, QEMU and DPDK, it was also able to improve the stability of the vSwitch. Cost savings were also realized by being able to operate on existing hardware without additional investment. Pre-allocated CPU and memory resources could be set, allowing flexibility to occupy additional resources when network traffic is low. Meanwhile, as a software-based solution, OvS-DPDK could be easily adapted for MOS’s cloud platform while still allowing customized functionalities and improvements to be made easily, using common programming methods and with no specialized Register Transfer Level (RTL) skills or tools. Updates can be made quickly and simply using remote patches, rather than requiring regular site visits to maintain hardware. Functionality is improved by enhancing vHost recovery and live migration, while quality of service was also enhanced using the DPDK-BOND port and preserving the existing OvS-FILTER framework.

With live migration now possible, it was also quicker to upgrade from OvS to OvS-DPDK rather than moving to Smart NIC switches, helping to save time in operations as well. This time saving was important as re-initializing the system during an update used to take a few hours. With OvS-DPDK, the re-initialization process went down to just a few seconds, thereby minimizing the impact on the system.

“By working with Intel to test and optimize our existing network architecture, we were able to boost performance and achieve lower capital expenditure (capex) and operational expenditure (opex) than alternative options, which could require larger investments of funds and expertise to achieve,” says Jianyan Ye, director of the networking department at Meituan Cloud.

Advantages of Network Switching with OvS-DPDK

- Improved performance-per-core over OvS
- Creates additional CPU resource for more stable network and system performance
- Compatible with OvS on core
- Pure software framework allows performance upgrade on existing hardware

Technological advantages of OvS-DPDK

Operational Benefits of OvS-DPDK

- Stronger performance than vanilla OvS
- More efficient use of NIC capability based on performance demands of workloads
- Exclusive core resources for more stable performance
- Simple deployment for customized functionalities

Figure 2: Performance increase OvS-DPDK over vanilla OvS¹
Prepared for future growth
As MOS continues to grow its customer base and service portfolio, its data center and network management techniques will adapt and change over time. Intel® technology will remain a core element of its environment, with ongoing efforts underway to maintain and enhance the DPDK network software framework to further increase the reliability, manageability and scalability of running vSwitch on the core.

For larger workloads, with a throughput of over 25Gbps, Intel is also developing a Smart NIC roadmap to ensure a smooth transition to an integrated solution that supports virtual I/O and vHost.

Spotlight on Meituan Open Services
Meituan Open Services (MOS) is the public cloud platform offered by Meituan, China’s leading e-commerce platform for services. Meituan aims to transform China’s service industry by providing solutions to merchants, including targeted online marketing tools, cost-effective on-demand delivery infrastructure, cloud-based ERP systems, integrated payment systems, and supply chain and financing solutions. For more information visit www.mtyun.com

Lessons Learned
Maximizing your billable resources doesn’t mean investing in brand new technologies. Make the most of what you already have today by optimizing them with simple solutions such as OvS-DPDK.

Find the solution that is right for your organization. Contact your Intel representative or visit intel.com/CSP

Learn More
This solution brief is part of an entire solution kit of content that is full of key insights and learnings:

- Intel® Xeon Phi™ processor
- Intel® Xeon® Scalable processor
- Intel® Distribution for Caffe*
- Intel® Ethernet Converged Network Adapter X540-T2
- Open vSwitch* with DPDK Overview