White Paper

IT services, telcos
Enterprise, Service providers

PANTHEON.tech*, Delivers Fast Data and Control Planes

Data plane performance is important to containerized networking applications. PANTHEON.tech Co.'s StoneWork*, running on Intel® architecture processors, supports high-speed, containerized packet processing and high bandwidth.

Table of Contents

Network functions virtualization (NFV) pioneered the virtualization of telecom network functions. NFV was followed by cloud native functions (CNFs), and they both have become a mainstream technology deployed in production networks. According to the Cloud Native Computing Foundation*, 96% of all organizations surveyed in 2021[1] are using or evaluating a cloud native operating system called Kubernetes*.

A high-performance / high-scale Kubernetes implementation requires high-speed / low-latency virtual networks within a server to interconnect containers / pods running on the server to each other and to the outside world.

Routing or switching packets in a virtual network requires a data plane and a control plane. The data plane defines the part of the process that decides what to do with packets arriving on an inbound interface. Most commonly, the router looks up the destination address of an incoming packet in its forwarding table and retrieves the information necessary to determine the path from the receiving interface through the internal forwarding fabric to the proper outgoing interface.

The control plane configures packet forwarding policies, handles forwarding exceptions, and further optimizes data plane processing. The separation of data plane from control plane in a pure software-based virtual network has proven useful in packet switching.

Traditionally a virtual network that connects application containers/pods has utilized networking functionality built into the Linux operating system. All forwarded packets must be processed by the Linux kernel to determine their destination. But the kernel is subject to interrupts and intracore locking, which can negatively impact throughput.

Technologies have been introduced to move the packet processing from the kernel to the user space which enables high-speed, low-latency packet processing virtual appliances that can be packaged in containers/pods or VMs. One of the first efforts was the Intel-developed Data Plane Development Kit (DPDK), a set of data plane libraries and network interface controller polling-mode drivers for offloading TCP packet processing from the operating system kernel to processes running in user space. DPDK is now an open source project managed by the Linux Foundation*.

After DPDK came vector packet processing (VPP) which provides layer 2 / layer 3 switching and improves performance and lowers latency by processing multiple packets at the same time, with low latency.

PANTHEON.tech, an Intel® Network Builders ecosystem participant, has combined these technologies and more into StoneWork, a high-performance data plane/control plane for next-generation container networking.

StoneWork: Data Plane, Control Plane Framework Solution

PANTHEON.tech has more than two decades of experience in network software development. Recognizing the importance of cloud native performance, PANTHEON.tech created StoneWork, a high-throughput data plane and modular control plane, vital elements in moving information across a virtualized network.

StoneWork offers a complete and fast-virtualized data path solution that integrates a VPP data plane built on the open-source Fast Data.io (FD.io) project with a management agent built using the open source Ligato project. VPP moves packet processing from kernel space to user space, thus delivering higher performance. The first packet in the vector 'warms up' the instruction cache, so remaining packets are processed extremely fast - sharply reducing the processing time for each subsequent packet in the vector.

FD.io is flexible and can be used to support a wide range of network functions. Today, several major communication network providers and equipment manufacturers rely on it to deliver world-class network functions. FD.io VPP has a growing list of network traffic management and security features to support gateway use cases, such as a broadband network gateway.

FD.io VPP also supports entry-level hardware options from a number of hardware vendors for building low-cost customer premise equipment (CPE) devices that can support all the WAN and LAN needs of a branch office, including cloud load balancers and intrusion prevention systems. FD.io VPP has four different access control list technologies; ranging from simple IP-address whitelisting to sophisticated FD.io VPP classifiers. FD.io is being used in discrete appliances, virtual network functions (VNFs), and for cloud native functions.

StoneWork Enterprise is based on a foundation of DPDK and uses the poll mode drivers to enable receiving and sending packets with a minimum number of CPU cycles. With DPDK, some packet processing functions have been benchmarked up to hundreds of millions frames per second, using 64-byte packets with a PCIe NIC[2].

Imperative Control Plane Based on Ligato*

Ligato is the control plane software framework that is used in the StoneWork solution. As seen in Figure 1, it provides a VPP agent and a Linux agent with a full set of high-level networking services and programmable interfaces. It also provides a KVScheduler* that configures data-plane network functions using a dependency resolution feature to keep the data plane in sync with the intended configuration even if multiple configuration requests are coming from different sources.

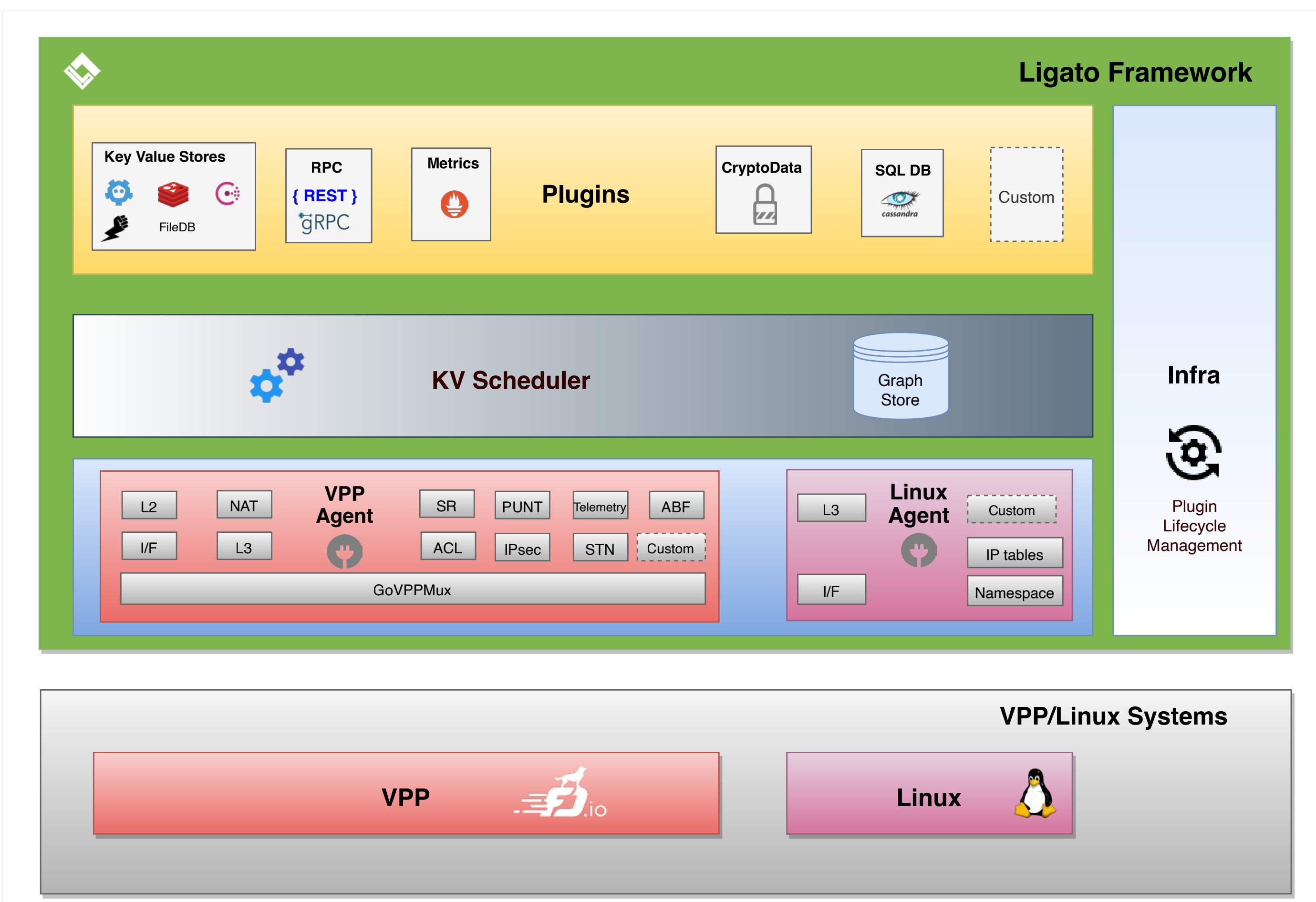

Figure 1. Ligato block diagram.

Ligato supports either declarative or imperative programming; StoneWork provides a declarative interface for users of the system and translates the desired intent by calling Linux and VPP imperative programmable interfaces for data plane configuration.

With the declarative approach, a YAML file containing the desired configuration will be read and applied towards the declarative statement. A KVScheduler then translates the desired configuration into a set of steps that are executed via VPP binary APIs to apply the intended configuration.

With the imperative approach, developers describe how to do something. They explain to VPP in detail how to configure a network service. Imperative management can deliver a faster introduction to managing deployments, with more control over each of the steps in the process, but can be more challenging in case of dependencies between individual programmable interfaces. An example is a virtual interface. In order to use a virtual interface (e.g., loopback, VLAN, tunnel, etc.) in combination with another network function (e.g., VRF, bridge, cross connect, etc.) the developer must first create the interface, before it can be used somewhere else in the system.

In addition, StoneWork's management features include a programmable, protobufs-modeled northbound application programming interface (API) that makes configuration and state data management lightweight, transport-agnostic, and language-neutral. The system can be accessed via a variety of technologies: transport protocols, such as gRPC, REST, or HTTP, reading/writing values to/from key-value stores such as ETCD, Redis, or Bolt, using Kubernetes CRDs and via message brokers, such as Kafka or RabbitMQ.

Business Results

StoneWork is a high-performance data plane and a modular control plane. The open source solution provides a framework for building, deploying, and managing CNF applications.

The solution features optimized packet interfaces supporting a multitude of use cases:

- An integrated vhost user backend for high speed VM-to-VM connectivity
- An integrated memif container backend for high speed container-to-container connectivity
- An integrated vhost-based interface to punt packets to the Linux Kernel

The same optimized code-paths execute on the host and inside VMs and Linux containers. As a result, developers leverage best-of-breed open source drivers. Its linear core scaling has been tested with millions of flows and MAC addresses.

These features have been designed to take full advantage of common CPU optimization techniques, such as reducing cache and TLS misses by processing packets in vectors, realizing IPC gains with vector instructions such as SSE, AVX and NEON, eliminating mode switching, context switches and blocking, and supporting cache-lined aligned buffers for cache and memory efficiency.

StoneWork Enterprise, a solution offered by PANTHEON.tech, integrates StoneWork and network functions from PANTHEON.tech's CDNF.io portfolio. The StoneWork Enterprise version contains additional integrated features like IPsec internet key exchange (IKE), BGP, OSPF, IS-IS, DHCP, IDS/IPS. That makes StoneWork Enterprise a universal solution for deployments on the edge, in the cloud, or on-premises.

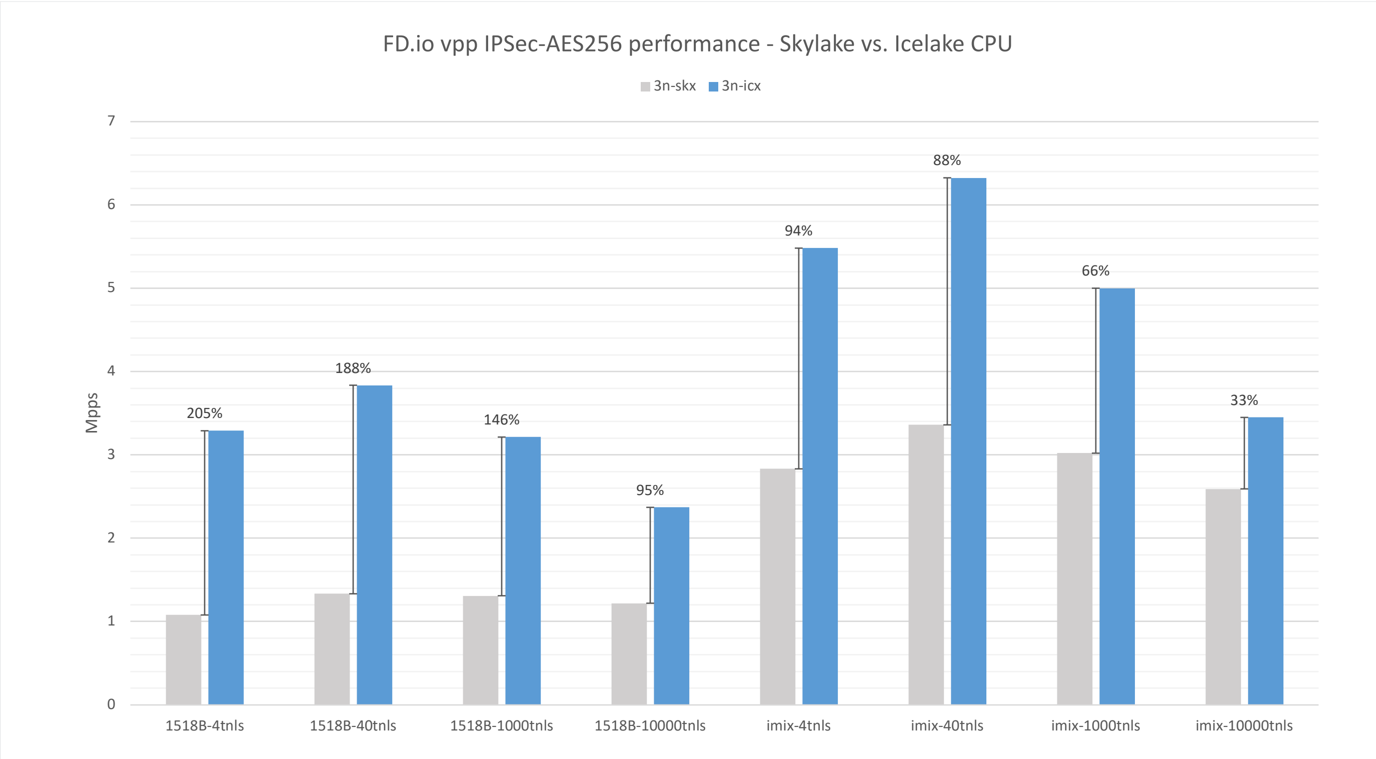

Figure 2. FD.io CSIT - IPsec packet forwarding performance.

Test 1: 3rd generation Intel® Xeon® Scalable processor vs. previous generation

To show the data plane performance improvement of the 3rd generation Intel® Xeon® Scalable processors versus the previous generation CPU, a test was set up to measure FD.io VPP packet throughput on two servers connected to a traffic generator.

The testing methodology uses multiple test methods to obtain repeatable test results. Multiple loss ratio search (MLRsearch), based on Request for Comment 2544 (RFC2544) an IETF standard for measuring overall throughput of the system with zero packet loss. Maximum receive rate (MRR), MLRsearch complementary tests providing maximum throughput of a device under heavy load. Probabilistic loss ratio search (PLRsearch), extends RFC2544 to provide a suitable benchmark test for software-based network functions. Detailed descriptions of test methodology and hardware, used for performance test, is available as a part of FD.io CSIT documentation[3].

IPsec encryption/decryption performance was also tested (see Fig. 2) using both 1518-byte packets and a random mix of data packet sizes typically found in internet traffic (IMIX). Tunnel counts of 4, 40, 1000, 10,000 were also used. And finally, the tests measured the strength of the crypto algorithms including AES256GCM, AES128CBC, and HMAC512SHA to evaluate impact of changing these parameters. Most significant improvement can be seen on IPSec-AES256 tests. The 3rd Generation Intel Xeon Scalable-based server provides almost 200% improvement, which means the forwarding is three times faster compared to the previous generation CPU. For IMIX packets with four configured tunnels there is more than 100% improvement in the number of forwarded packets, which is double the throughput of the tests run on the previous generation CPU.

Test 2: StoneWork Enterprise vs. FD.io VPP performance

To show the impact on performance when using the StoneWork Enterprise advanced control-plane features the same test that was performed on Intel Edge and Networking Testbed 3-node topology. Detailed description of the testbed setup can be found in Appendix 1.

The goal of the tests was to compare the performance of VPP 21.06 vs. StoneWork Enterprise 21.06 on a 3-node topology. IPsec was used to compare the impact of externalized StoneWork Enterprise control-plane compared to the VPP data-plane only. Packets were sent from a traffic generator (TG) to a device under test 1 (DUT1) via an encrypted IPsec tunnel to a device under test 2 (DUT2). DUT2 decrypted the packets and forwarded them back to the TG. Both DUTs performed encryption and decryption at the same time as the traffic, sent by the TG in a 3-node circular topology, was bi-directional.

The main difference between StoneWork Enterprise and VPP tests was the IPsec tunnels between DUTs, which were statically created by using VPP binary APIs, where in the case of StoneWork Enterprise tests, the tunnels were dynamically created by the enhanced IPsec control-plane plugin via Internet Key Exchange version 2 (IKEv2). The same test methodology was used as in previous test (MRR) to determine maximum throughput of the environment. Observed performance results can be found in Table 1.

TEST CASE	STONEWORK ENTERPRISE (VPP 21.06)	CSIT (VPP 21.06)
1518b-1c-ethip4ipsec4tnlsw-ip4base-int-aes256gcm	3.11 Mpps	3.00 Mpps
imix-1c-ethip4ipsec4tnlsw-ip4base-int-aes256gcm	4.85 Mpps	4.85 Mpps

Table 1. IPsec performance comparison between StoneWork Enterprise and VPP.

The results validated that the StoneWork Enterprise enhanced control-plane does not affect VPP data-plane performance. That result is expected, as the StoneWork Enterprise data-plane uses the open-source VPP version, and with proper CPU pinning, the control-plane containers (processes) do not interfere with the data-plane worker threads.

Conclusion

Telecommunications service providers' and enterprises' reliance on their networks has been increasing. They need a simple-to-deploy, high performance platform. PANTHEON.tech's StoneWork platform leverages the Intel microprocessor technology and delivers the high-bandwidth, low-latency virtual networking required in today's cloud environments.

Learn More

PANTHEON.tech StoneWork

FD.io

Ligato

Data Plane Development Kit

Intel® Xeon® Processors

Intel® Smart Edge

Intel® Network Builders

Appendix 1. Intel® Edge and Networking Testbed description

System Description & Configuration

	TG	DUT1/DUT2
PLATFORM	Cascade Lake	Ice Lake
SOCKETS	2	2
CPU	Intel® Xeon® Platinum 8260M CPU @ 2.40GHz	Intel® Xeon® Gold 6338N CPU @ 2.20GHz
CORES/SOCKET	Core(s) per socket: 24	Core(s) per socket: 32
INTEL® HYPER-THREADING TECHNOLOGY	Enabled	Enabled
INTEL® TURBO BOOST TECHNOLOGY	Enabled	Enabled
BIOS VERSION	SE5C620.86B.02.01.0012	05.01.01
SYSTEM DDR MEM CONFIG: SLOTS / CAP / SPEED	24/24/ Freq: 2933 MHz	32/32/ Freq. 2667 MHz
TOTAL MEMORY	12*32GB DDR4 - 2933 MHz	16*32GB DDR4 - 3200 MHz
NICS	1* Ethernet Controller E810-C for QSFP 1* Ethernet Connection X722 for 10GBASE-T	2* Ethernet Controller E810-C for QSFP 1* Ethernet Controller 10G X550T

Software Configuration

	TG	DUT1/DUT2
OS	Ubuntu 20.04.4 LTS	Ubuntu 20.04.4 LTS
KERNEL	5.4.0-107-generic	5.4.0-107-generic
WORKLOAD & VERSION	CSIT, Trex 2.88	StoneWork 21.06, VPP 21.06 StrongSwan swanctl 5.9.0dr2

Network topology

Description: Figure 3 illustrates the testing topology and data flows. Tests were performed on a circular, 3-node topology. Traffic generator (TG) is connected to 2 Device Under Test devices (DUTs) with 2 different interfaces. Both interfaces are used as Rx and Tx at the same time to simulate symmetric, bi-directional traffic.

Traffic flows:

1. TG - DUT1 - DUT2 - TG
2. TG - DUT2 - DUT1 - TG

IPsec tunnels are configured between DUT1 and DUT2

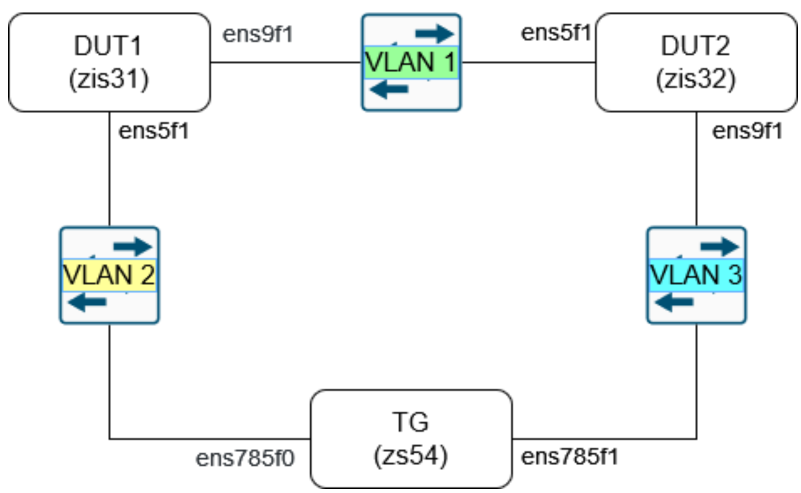

Figure 3.

Notices & Disclaimers

[1] https://www.cncf.io/announcements/2022/02/10/cncf-sees-record-kubernetes-and-container-adoption-in-2021-cloud-native-survey/
[2] https://www.dpdk.org
[3] https://docs.fd.io/csit/rls2202/docs/

Performance varies by use, configuration and other factors. Learn more on the Performance Index site.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. *Other names and brands may be claimed as the property of others.

0922/TM/HO9/PDF ♻ Please Recycle 352385-001US