SPDK v1.0.6

High Performance API for Network Monitoring

Over

Intel® DPDK

White Paper, May 2014

Efficient and scalable packet processing infrastructure for Monitoring and Network Visibility implementations, at 10GbE wire speed and beyond, without loss and with a breathing commodity Intel® CPU is now, and for quite some time actually, within reach.

Packet Processing

Quick Historic Overview

It said that Intel® co-founder, Mr. Gordon E. Moore, once coined a law stating that approximately every two years, the number of transistors is doubled on an integrated circuit. Roughly this means that a standard grade commercial processor's power would double as well. Starting from 2007, Intel® exhibit more or less a consistent Tick-Tock approach with its microarchitecture, where Tick means die shrink, and Tock means new microarchitecture.
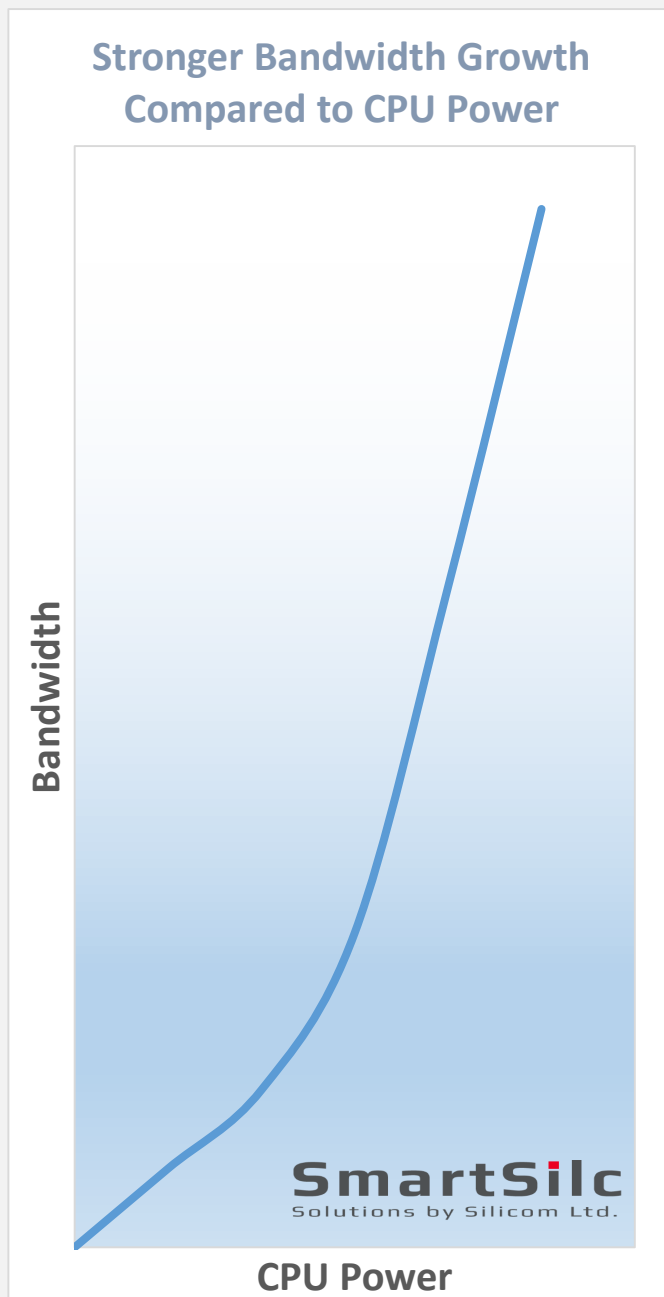
Hand in hand with the processing power evolution, the Ethernet network bandwidth has evolved as well, from the days of 10Mbps to the 10GbE and 40GbE standard and to 100GbE around the corner.

Monitoring

However, CPU power on one hand, and bandwidth demands have grown separately in different proportions over time, while the bandwidth demands and technology were growing faster.

The widening spread of the networking use, brought along the need for network monitoring and network visibility tools, to be used for network security, and network stability and integrity requirements.

 Those tools became more and more complex with logic, while the number one requirement from such tools was to uphold ever growing bandwidth.

Stronger Bandwidth Growth Compared to CPU Power

One thing that has remained pretty much unchanged over time was the basic software infrastructure for packet processing in commodity general purpose CPU, which had and still has roughly three layers. One is the network controller driver. The second internetworking protocol layers, and above it, user programming interface. For Network processing.

Few years back, when 1GbE controller was the high end standard for network controller, it was already apparent the handling wire speed packet processing with commodity general purpose CPU, is gradually becoming a challenge. When traffic was flowing in at wire speed, up to 1.48 million frames per second, then the CPU had to serve up to that many interrupts each second. While one single 1GbE interface was something that could be coped with, 4 such instances was more or less the limit. Commercial vendors that were striving to exhibit screaming performance opted mostly to dedicated network processor, as a standalone silicon, in order to at least relieve the main CPU to its applications tasks, if not to withstand ever growing bandwidth.

The Game Changes for Network Monitoring

As part of Silicom's 10GbE network interface cards (NIC) offering, and for several years now Silicom is offering commercial grade software solutions for 10GbE wire speed packet processing.

Intel® DPDK, went open source, and immediately became the platform of choice as an infrastructure for wire speed packet processing for Silicom's solutions, mainly because of the following important reasons:

- Technically, Intel® DPDK is superior in resource handling and scalability, as will be shown herein;
- Intel® DPDK was tailored bottom up, looking forward with virtualization in mind;
- Intel® DPDK large eco system which is backed by Intel®.

The tests results brought herein support the initial motivation to employ Intel® DPDK as an infrastructure, while more is revealed once the implementation is tested and fine-tuned:

- Buffers management is much more robust and stable
- Wire speed processing is enabled, provisioned and powered with full scalability and stability at the bare minimal costs and nominal efficiency and power consumption

Content of SPDKv1.0.6 Package – A Monitoring API

The current SPDKv1.0.6 features the following:

- Monitoring and network visibility API
- Capture & Replay application API
- True 10GbE processing power
- Efficient Wireshark and Tcpdump interface via libpcap library
- Smart and easy, buffer management
- Improvement over DPDK in terms of CPU utilization, and memory management
- Sample applications utilizing the delivered API
- Background running `SPDK` daemon and `SPDK.conf` configuration file

The package is delivered in two flavors. One flavor is optimized to Intel® Sandy bridge microarchitecture, and another flavor optimized to Intel® Ivy bridge microarchitecture.

Testing the New Package

The Numbers are Unveiled

Two types of tests were performed on Silicom's lab with the new SPDKv1.0.6 software release utilities. One is the suite of industry accepted tests, under accepted standards, such as RFC2544. The second suite of tests was done to profile out how SPDKv1.0.6 scales with resources and optimizes the use of available system resources.

The results not only met the initial expectation, but in several important aspects surpassed them, as brought herein.

Overall Performance

Performance testing conducted with SPDKv1.0.6 revealed immediately the strength of Intel® DPDK, and more specifically, how SPDKv1.0.6 improvements have an effect.

Testing streams of variable length of frames, injected simultaneously from two ports, immediately revealed the SPDKv.1.0.6 performed very close to maximal capacity, right from the shortest length of 64 bytes frames.
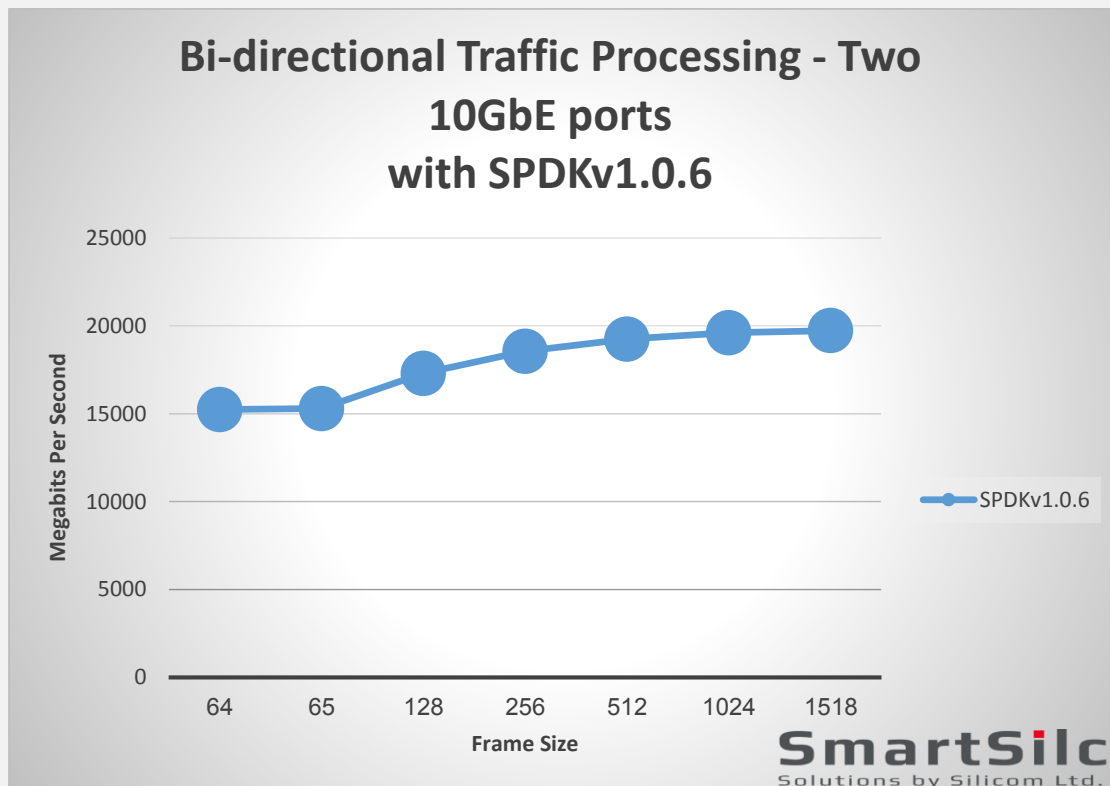
Notorious 65 Bytes Long Frames

Even better, when tested with 65 bytes long frames – perhaps the most problematic frame length, that often cause poor buffer utilization within a packet processing application – even then, SPDK with its **advanced buffer management**, managed to uphold the traffic without degradation.

Best Approach

Testing SPDKv1.0.6 bi-directional processing rate with various frame length on a single processing core, reveals an impressive behavior right from the start. With 64 bytes long

frames, SPDKv1.0.6 kick starts with close to 75% of full wire speed, and **immediately approaches theoretical maximal rate** soon as frame length increases.
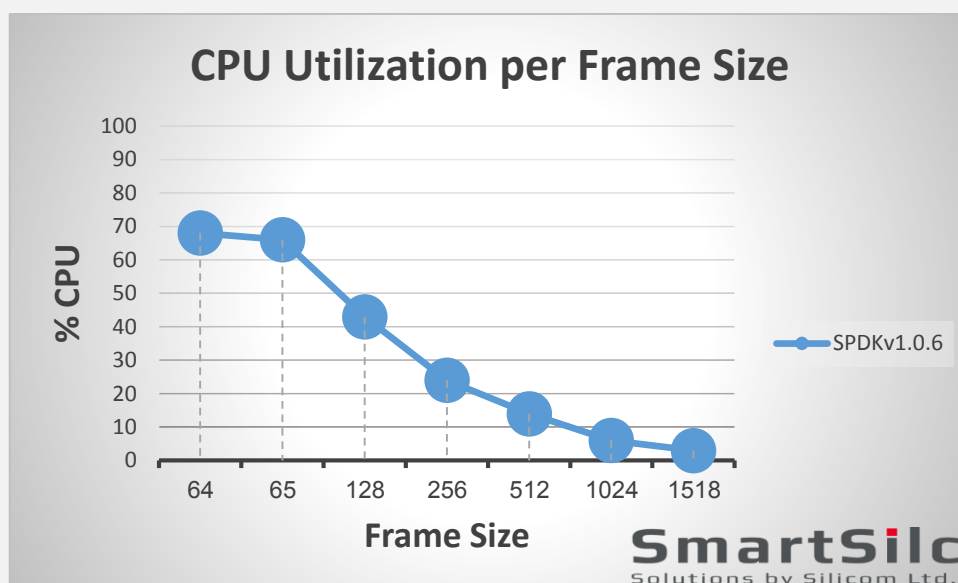

With 128 bytes long frames, SPDKv1.0.6 reaches close to the theoretical maximal bandwidth, it practically there with 256 bytes long frames. Taking into account that the rule-of-thumb average length of a data frame in the internet is ~400 long frame, we conclude that at a real life scenario, the SPDKv1.0..6 operation on a single commodity core, enabled true wire speed packet processing.

CPU Utilization

Not only that SPDKv1.0.6 performs better in bandwidth, it does it with significantly less CPU power, thus maximizing the CPU utilization, and getting really close to the theoretical CPU-cycles-per-frame numbers.

CPU cycles

Measurements of CPU utilization percentage with SPDKv1.0.6 reveals an important fact. There is room for more business logic of packet processing for an application, with SPDK. SKDP consumes just the right amount of CPU power it requires, while leaving the rest of the resources free and available for an application. Even in the most demanding cases where frame are short, the CPU breaths with quite large percentages, thus enabling further processing on the same core. Longer frame consumes practically insignificant amount of resources.
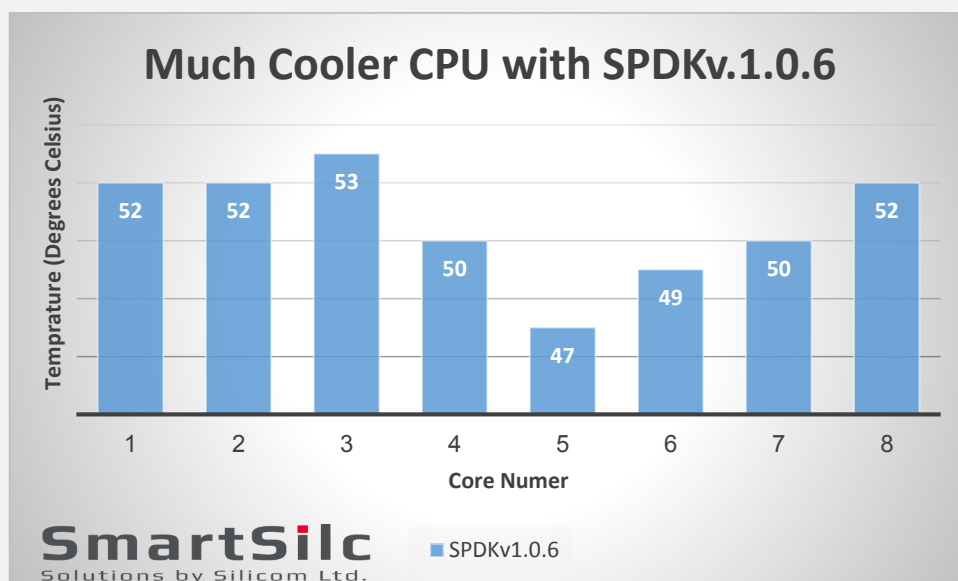
CPU power consumption emerges as yet another strength of SPDKv1.0.6, where all over fame sizes, CPU would never saturate.

This is a SPDK specific improvement, and it is present even over open source DPDK in that respect.

More Power with Less Fuel – The Cool Factor

Further research into performance with SPDKv.1.0.6 reveals another important fact. Using less CPU per bandwidth, brings along savings with electricity power consumption, and less hit dissipation.

The long term implication for that fact stems right away. More longevity to the hardware as a whole, less cooling burdens, less overall power consumption, with best wire speed performance.

RFC2544 - Throughput and Latency

Silicom conducts regularly standard throughput and latency tests, whether unidirectional, or bidirectional, with its SPDK releases. Tests are conducted with strict compliancy both in procedures and in measurements and results, to RFC 2544.

Throughput and latency figures, with all frames length, always fall within the required industry standard definition. For more information, please contact Silicom representative.
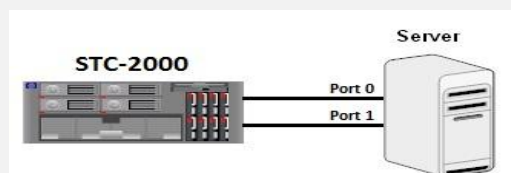
Summary

New SPDKv1.0.6 release brings about new improved package of ten of gigabits of packet processing at wire speed, focusing on:

- True 10GbE processing power
- Capture & Replay application API as well as Monitoring and network visibility API
- Smart buffer management
- Best CPU utilization
- Based on Intel® DPDK with a long road map ahead

About the Testing Methodology

Testing were conducted with persistent installation, with commercial grade test equipment, and by using accepted industry standards as well as RFC standards for testing. The test bed was comprised of:

- Server: Intel Xeon, CPU - E5690 2900 MHz, memory - 32 Gb - L1 Cache 256kB - L2 Cache 2048 kB - L3 Cache 20 MB
- Network adapters: PE 210G2SPI9 - SR, slot PCI-E x8 v. 3.0 RSS=0,0
- OS: CentOS 6.5 kernel 2.6.32-358.el6.x86_64
- Software under test: SPDK-1-0-6 with shp_cap -d sio0 -c 0; shp_cap -d sio1 -c 2; Other existing zero copy tools
- Traffic Generator: STC-2002HS, module - CV-10G-S8
- Utilities: i7z_64bit --socket0 0 --socket1 1

About SmarSilc

SmartSilc is Silicom's new suite of advanced networking solutions designed to streamline processes for cloud application vendors, network application vendors and data centers evolving towards next-generation virtualization. With a focus on acceleration and offload, SmartSilc products leverage all of the field-proven fundamentals of Silicom's network devices, NICS and accelerators, including the use of top-vendor, best-of-breed silicon and the provision of full support for production-grade software suites.

About Silicom

For more than 25 years, Silicom has provided innovative, state-of-the-art connectivity and networking solutions to the continuously evolving networking industry. Silicom's shares trade on the NASDAQ Global Select Market (NasdaqGM) and the Tel Aviv Stock Exchange (TASE) under the symbol SILC.