napa:tech;
RECONFIGURABLE COMPUTING

SOLUTION DESCRIPTION

Scaling Suricata Performance to 100 Gbps With Napatech SmartNICs

Using stateful flow processing technology

napa:tech;
RECONFIGURABLE COMPUTING

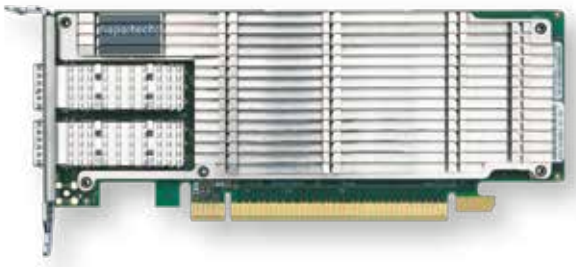
SOLUTIONS
Cybersecurity
Monitoring
Infrastructure
Cloud and Edge
Mobile
Financial

PLATFORM
Link™ Capture Software Napatech
Link™ Capture Software Intel®
Link™ Inline Software
Link™ Virtualization Software
Link™ SmartNICs
Link™ SmartCards
Link™ Programmable
FPGA Cloud Crypto

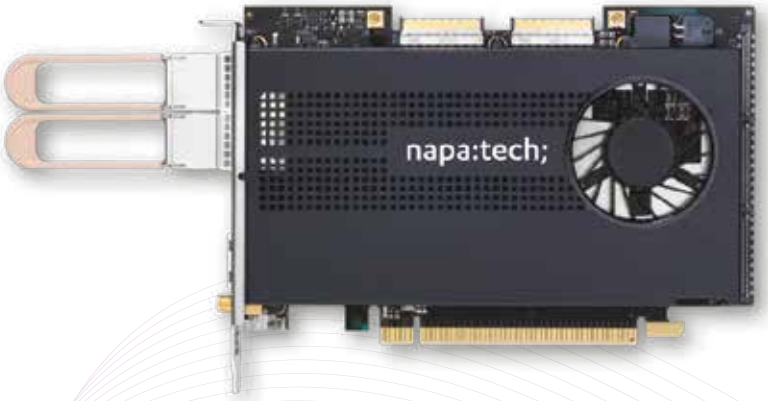
APPLICATION PERFORMANCE
Suricata
Zeek
Snort
n2disk
Wireshark
TRex
+ More

SERVICES
Professional Services
Custom Development

Link-25, 2-port 25 gigabit Ethernet SmartNIC



Link-100, 2-port 100 gigabit Ethernet SmartNIC



Accelerating Suricata with Napatech stateful flow processing technology

Suricata is a free open source, mature, fast and robust network threat detection engine. The Suricata engine is capable of real time intrusion detection (IDS), inline intrusion prevention (IPS), network security monitoring (NSM) and offline PCAP processing. Suricata inspects network traffic using a powerful and extensive rules and signature language and has powerful Lua scripting support for detection of complex threats. If a threat or anomalous behavior is detected, Suricata will send an alert to the administrator and optionally attempt to block or stop it.

The growth in network traffic and increasing sophistication of attack vectors are placing enormous strain on the CPUs of IDS/IPS devices. One of the ways these systems are adapting is by offloading more of the packet processing workload from the CPU to an FPGA-based SmartNIC.

Napatech SmartNICs have been shown to dramatically increase the lossless throughput of Suricata implementations by providing zero-copy DMA (Direct Memory Access) of packets directly from the NIC to the application, freeing CPU cycles otherwise required for that task.

The latest Suricata version can be accelerated to even greater performance by directing a Napatech SmartNIC to handle a significant portion of the packet processing in the FPGA on a per-flow basis without needing to process the traffic on the host CPU.

Suricata Acceleration

Evaluating every packet in real time against thousands of attack signatures for hundreds or thousands of concurrent flows would require enormous computing power. To make the problem tractable, since release 3.2 Suricata has allowed administrators to specify that traffic matching certain criteria should be excluded from further analysis once a determination has been made regarding the desired handling of that traffic. This is an optimization that allows packets to bypass the full packet analysis pipeline once a pass/drop verdict on the traffic has been made.

Common situations where this “bypass” feature is applied include:

- Encrypted traffic: The Suricata application layer parser has the ability to stop requiring further processing of encrypted traffic (TLS, IPsec, Kerberos) after the initial handshake.
- Streaming services: Where large flows are expected from known providers (e.g. music, movies, etc. from Netflix, YouTube, Spotify, Apple Music), traffic may be passed or blocked according to policy without the need to decode every packet.
- Special cases for certain types of traffic: Other expected large flows (e.g. a backup stream), or where certain hosts or applications may be trusted.

Improving Suricata performance with Napatech SmartNIC
Networking and security applications like Suricata are extremely CPU intensive and flow aware, but not all packets necessarily need to be processed by the application at all times. Only the first few packets in a flow may be interesting, or specific protocols can be ignored, or only packets to/from specific individuals may be worthy of inspection. The Napatech solution uses a flow-based action processing to bypass x86 altogether (FWD, block, redirect) using a workload-specific FPGA for flow action processing. This requires tight integration between x86 applications and SmartNICs via APIs to allow applications to change actions on a per-flow basis.

The basic idea of a bypass is simply to stop processing packets that don't need further inspection as early in the packet processing pipeline as possible. This feature is enabled in Suricata software by using the keyword 'bypass' in a rule signature, signifying that once a verdict is reached, all other traffic in the same flow can safely be excluded from full decode and evaluation.

For example, here's a signature to bypass a flow matching on IP address:

```
pass ip 1.2.3.4 any <> any any (msg:"pass all
traffic from/to 1.2.3.4"; bypass; sid:1;)
```

Napatech extends this idea by enabling the SmartNIC hardware to handle bypassed traffic on the FPGA without further involvement from Suricata or the host CPU. Without a Napatech FPGA-based SmartNIC, this bypass is done in software. When a Napatech SmartNIC is used, the bypass action is executed on the programmable SmartNIC, further increasing performance and reducing load on the general-purpose processors.

Support for this feature was recently merged into the Suricata git repository maintained by the Open Information Security Foundation (OISF) and will be included in all future Suricata releases.

The implementation is based on the Napatech Flow Management API which enables an application tracking network traffic on a per-flow basis to offload some or all of that processing to the SmartNIC FPGA: application-level analysis remains on the host CPUs while per-packet actions are executed in hardware on the network interface card.

The Flow Management API is capable of managing network flows through the full life cycle:

- Maintaining a table of active flows
- Adding new flows
- Executing per-packet actions
- Synchronizing packet/byte counters
- Updating flow records on termination or timeout

The Suricata source code offers a working example of the Napatech API for developers of other applications.

With Napatech Flow Manager functionality integrated to the application software, no programming is required of Suricata users: hardware-accelerated flow management is accessible via the familiar Suricata rules interface. Enabling a flow bypass action in hardware is simply a matter of compiling Suricata with Napatech support and including the keyword 'bypass' in the rule signature.

Once Suricata determines that a flow can be bypassed, all packets from that flow will be handled in the SmartNIC hardware for the duration of the flow. This frees the host CPUs from packet processing, allowing the application to achieve higher throughput.

Hardware bypass demonstration
To illustrate the value of hardware bypass, this demonstration compares the throughput of Suricata on a standard NIC using software bypass with an otherwise identical configuration on a Napatech SmartNIC with zero-copy packet DMA and hardware bypass.

Hardware Configuration
The test configuration was based on a dual-socket HP DL380 G8 with Intel® Xeon® E5-2690 v2 3.0 GHz CPUs (40 HT cores), running CentOS 7.5. Traffic was generated by PCAP replay of an actual network traffic capture comprising a mix of TCP (88%) and UDP (12%) packets in over a million unique flows. The traffic is typical of what might be seen on an enterprise network with hundreds of users running dozens of interactive client-server applications and internet access.

Traffic was sourced by a Napatech PCAP replay application on an NT200 2x40GE SmartNIC and carried over 40 GE links to the device under test (DUT) configured with either a Napatech NT200 SmartNIC (2x40GE) or an Intel XL710 (1x40GE) network interface card.

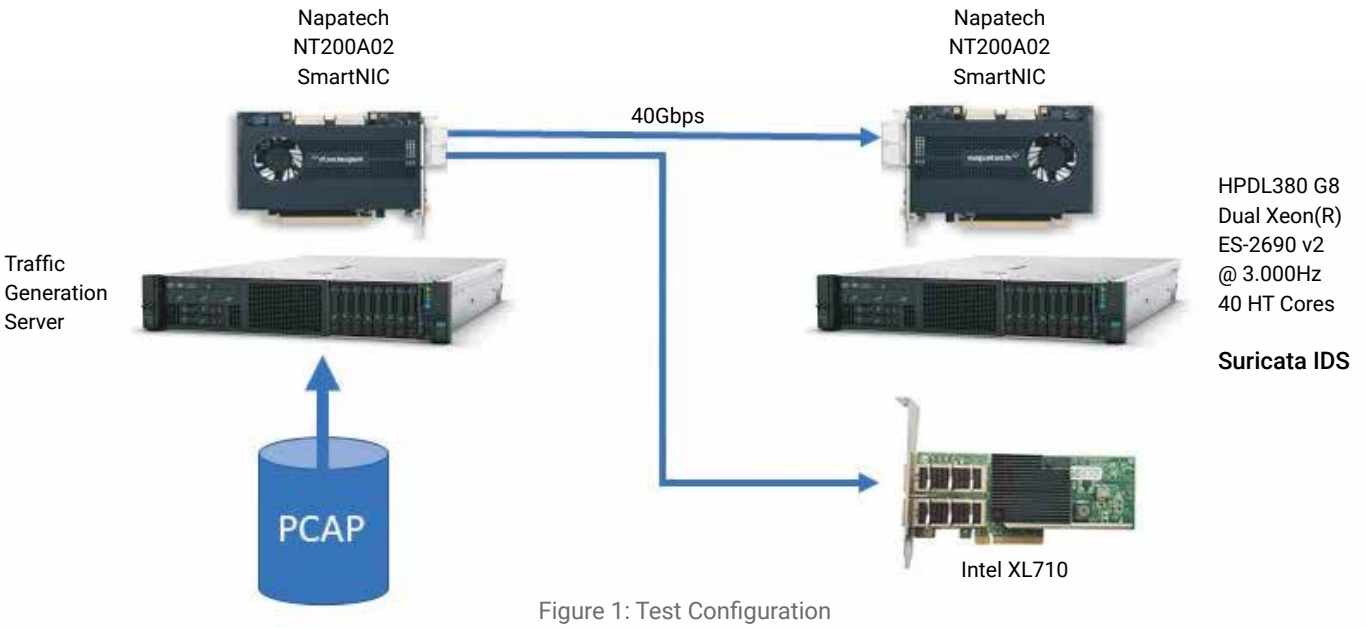


Figure 1: Test Configuration

Suricata Configuration
Organizations nearly always need to calibrate their IDS/IPS to recognize what normal traffic on the network looks like as compared to potentially malicious activity. These adjustments involve fine-tuning the rule set and enabling bypass as appropriate.

As noted above, the user enables bypass by using the 'bypass' keyword in a Suricata rule signature. In this test, a single bypass rule was added to the base Suricata rule set. The syntax of the additional rule is this:

```
drop tcp any [!80,!20,!21,!23,!25,!53,!67,!
68] <> any any (msg: "SURICATA Test rule";
priority:1;classtype:protocol-command-decode;
bypass; sid:1000001; rev:2;)
```

In words, this rule indicates that any TCP traffic should be bypassed except traffic to/from the specified set of port numbers. The same type of rule could easily provide a similar type of bypass for UDP traffic, traffic to/from a set of addresses, specific applications protocols, etc.

With the traffic pattern used in the test, the effect is that approximately 70% of the traffic was bypassed and the remaining 30% of traffic was handled entirely by Suricata software on the host CPUs. The SmartNIC still provides zero-copy DMA to the application in this case, but every flow must be decoded, reassembled and fully processed in software.

In a practical application the amount of bypass traffic could range from zero to 100%, and the more traffic that is offloaded to the SmartNIC, the higher the potential capacity of the system.

Test Results

For an IDS/IPS, lossless performance is an absolute requirement for maximum security. Testing has shown that just 3% packet loss can lead to 10% missed alerts; a mere 0.4% packet loss can cause 10% file extraction failure rate.

In this test, lossless throughput with hardware bypass on the Napatech SmartNIC was 2.5 times higher than the same test with a standard NIC and bypass in software.

Measured lossless Suricata throughput:

- IPS using the Emerging Threats Open Ruleset with no bypass: 15 Gbps
- IPS using the Emerging Threats Open Ruleset with software bypass: 30 Gbps
- IPS using the Emerging Threats Open Ruleset with SmartNIC hardware bypass: 76 Gbps

CPU utilization

On the standard NIC, the Suricata workload was very unevenly distributed, with roughly half the cores running at 100% utilization while others were at less than 50% utilization. The reason for this is that many other solutions exhibit asymmetric traffic distribution or exact a performance penalty for non-local memory access.

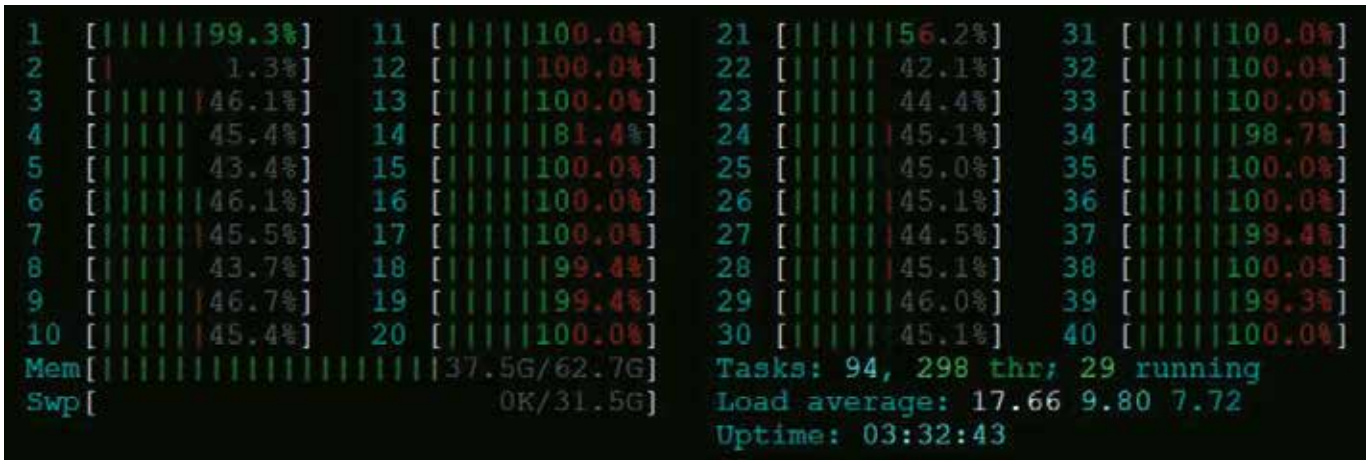


Figure 2: CPU utilization (standard NIC)

On the Napatech SmartNIC, the workload was much more evenly distributed across all cores, with essentially all cores loaded at less than 80%.

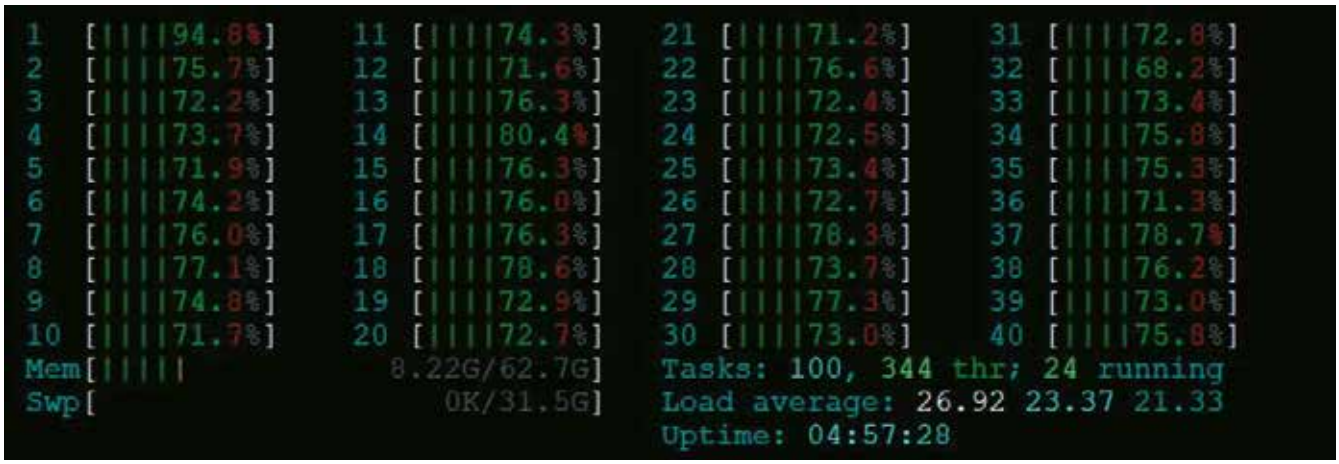
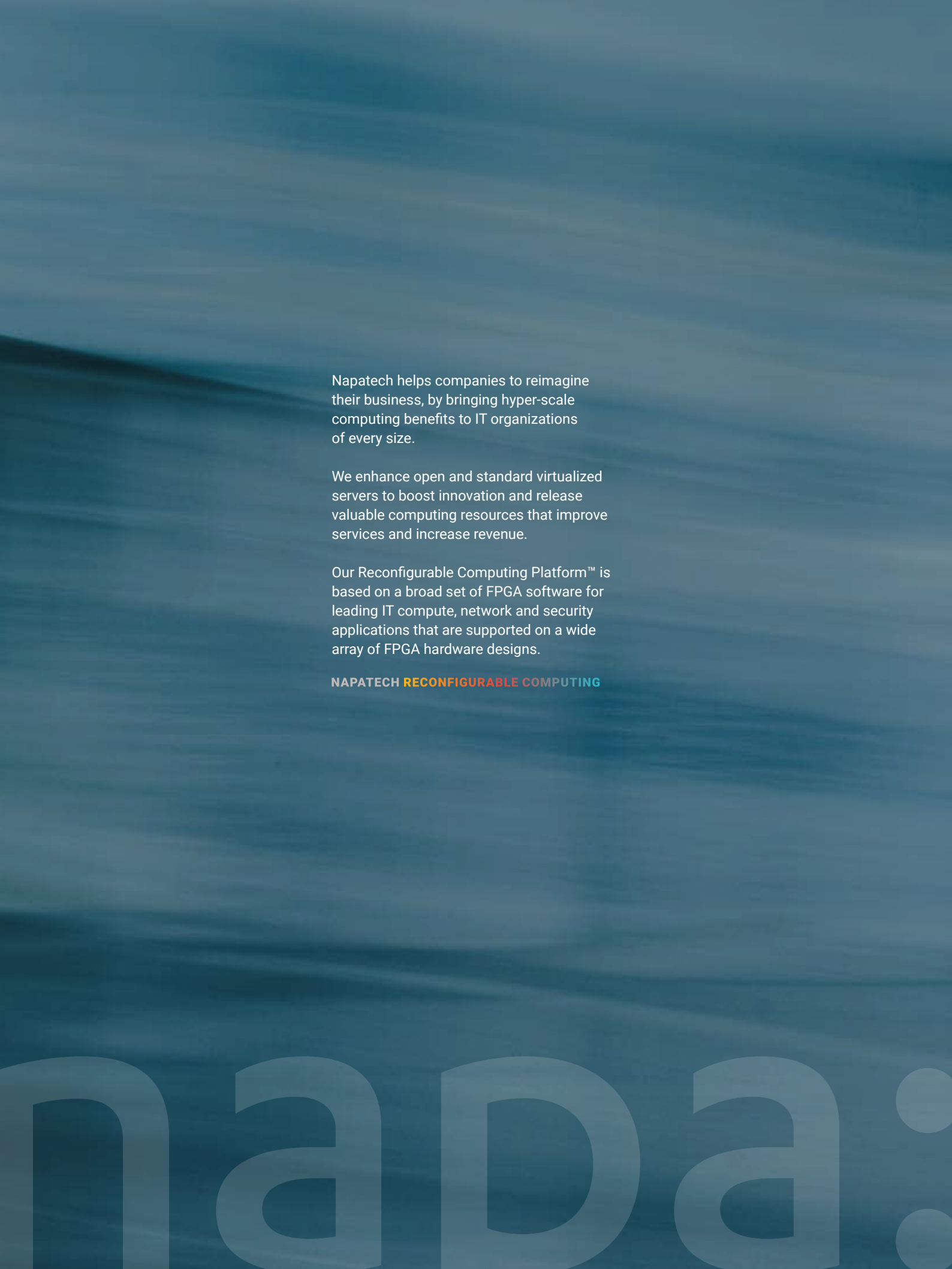


Figure 3: CPU Utilization (Napatech SmartNIC)

Better balanced workflow means higher system efficiency because a few CPU cores won't be overloaded and causing packet loss without the system being at full utilization.



Napatech helps companies to reimagine their business, by bringing hyper-scale computing benefits to IT organizations of every size.

We enhance open and standard virtualized servers to boost innovation and release valuable computing resources that improve services and increase revenue.

Our Reconfigurable Computing Platform™ is based on a broad set of FPGA software for leading IT compute, network and security applications that are supported on a wide array of FPGA hardware designs.

NAPATECH RECONFIGURABLE COMPUTING

Summary

This test demonstrates Suricata processing 76 Gbps without packet loss by handling bypassed traffic on the Napatech FPGA-based SmartNIC. In comparison, the same system doing bypass in software achieves 30 Gbps throughput without packet loss.

Additionally, the Napatech NIC demonstrated a Suricata workload balanced across all available CPU cores in the system, minimizing processor starvation or overloading with buffer overflow.

Suricata hardware bypass feature and Napatech Flow Manager API are supported with Napatech's Link™ Capture and Link™ Inline Software on the NT200A02 and NT50B01 FPGA-based SmartNIC hardware.

Technical specifications

SmartNIC Hardware Feature Highlights

General Hardware Specs	NT200A02	NT50B01
Dimensions	½-length, full-height PCIe	½-length, ½-height PCIe
Network Ports	2×QSFP28	2xSFP28
Data rate	8×10Gbps,2×10/25Gbps, 4×10/25Gbps,2×40Gbps, 2×100Gbps	2×10/25 Gbps
Onboard memory	12GB	10 GB
PCIe configuration	16-lane 8 GT/s PCIe Gen3	16-lane 8 GT/s PCIe Gen3

Shared Hardware Feature Highlights

- Zero packet loss for all frame sizes from 64 bytes to 10,000 bytes
  - keep or discard errored frames
- Addressing up to 1 TB application buffer memory
- Packet or segment delivery to application
- Hardware-acceleration for:
  - Multi-port packet merge
  - Load distribution across up to 128 CPU cores
  - Filtering based on e.g. L3/L4 criteria
  - Stateful flow management
  - GTP, IP-in-IP, GRE and NVGRE tunnel support
  - IP fragment handling
  - Slicing at fixed or dynamic offset
  - De-duplication in hardware
  - RMON1 counters, including jumbo frames
- IEEE 1588-2008 PTP and PPS time synchronization
- Hardware Time Stamp: Resolution: 1 ns
  - Stratum 3 compliant TCXO
  - OS time synchronization

Napatech Software Suite Feature Highlights

- NTAPI: Common API for all Napatech SmartNICs
- Simple programming interface allowing fast integration
- Zero copy transfer of data to host
- Zero copy transfer of data between SmartNICs
- Packet-based interface providing abstraction of hardware details
- Stateful Flow-aware match and action processing
  - Matching – flows are completely configurable based on flexible key matcher
- Actions
  - Drop (Block) - Drop unwanted traffic, reducing system resource reduction and increasing performance
  - Forward (cut-through/bypass) - Local, on card, port to port forwarding, reduces CPU demands and PCIe bottlenecks, lowers latency to increase application performance
- Flow-based Load Balancing (RSS) - efficient and distributed CPU utilization
- Per flow statistics (packets / bytes) with directional state per packet
- Flow Aging – Monitor TCP flow state, configurable timers for UDP
- Flow Metadata - information sharing between host application and SmartNIC offload engine
- Advanced features including:
  - Data merging of port data from multiple SmartNICs into a single data stream
  - Data sharing of captured data between multiple customer applications without the need for replication
- Operating systems: Linux, Windows
- ibpcap, WinPcap and DPDK
- IEEE 1588-2008 PTP stack
- SDK tools included in source code for debugging and prototyping and as application examples

napa:tech;
RECONFIGURABLE COMPUTING

Napatech helps companies to reimagine their business, by bringing hyper-scale computing benefits to IT organizations of every size.

We enhance open and standard virtualized servers to boost innovation and release valuable computing resources that improve services and increase revenue.

Our Reconfigurable Computing Platform™ is based on a broad set of FPGA software for leading IT compute, network and security applications that are supported on a wide array of FPGA hardware designs.

NAPATECH RECONFIGURABLE COMPUTING

CONTACT
Tel. +45 45 96 15 00
Tel. +1 888 318 8288
info@napatech.com
www.napatech.com

DN-1260 Rev. 1