Alveo Host CPU Offload Features Overview

An Overview of Features and Functions Available with Alveo Accelerators/SmartNICs

Accolade Host CPU Features Support for Xilinx Alveo Accelerators

Xilinx's Alveo line of FPGA-based accelerators/SmartNICs help accelerate security and networking applications developed by the world's leading security and networking vendors. The Alveo accelerators are fully PCIe compliant and seamlessly integrate into standard servers and network appliances. A common API across all Alveo products (C language shared library) is used by the host application to configure and control each accelerator.

Alveo accelerators/SmartNICs are available in a variety of port configurations with speeds ranging from 1 to 100 Gbps.

Feature Categories

- 100% Packet Capture
  - Buffer memory
- Precise Time Stamping
  - Nanosecond precision
  - 1PPS, PTP, Host OS
  - Gigamon, Arista timestamp
- Packet Merging
  - Multi-port
  - Multi-accelerator
- Packet Slicing
- Packet Parsing
  - L2/L3/L4 header parsing
  - Tunneling and encapsulation protocols (VLAN, VXLAN, MPLS, GTP, GRE)
  - IP fragment handling
- Packet Filtering
- Deduplication
- Flow Classification
  - Flow Based Filtering
- Flow Shunting
- Host Packet Buffer (HPB)
- Packet Steering
  - Steering methods
  - U-Turn (Retransmit)
- DMA (Direct Memory Access)
  - Multi-core DMA
  - PF_RING
- Statistics
  - RMON1 (RFC 2819)
  - Alveo Onboard Sensors

Accolade
Technology
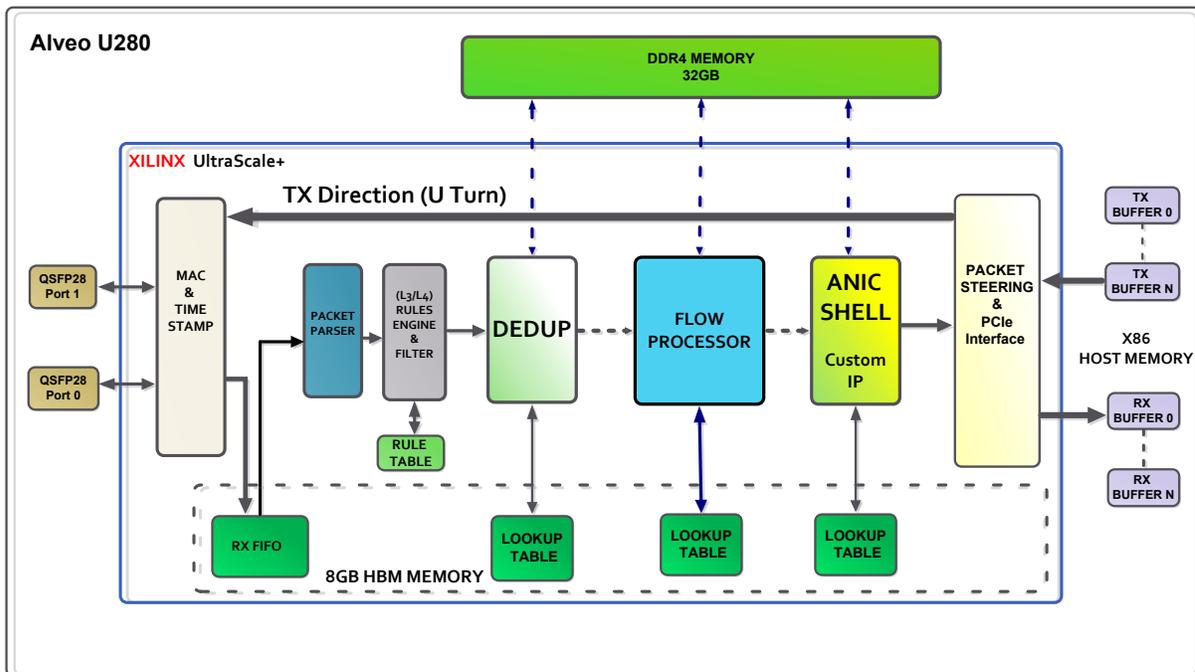
100% Packet Capture

Each Alveo accelerator port captures 100% of the incoming traffic, irrespective of packet size (i.e. 64 byte vs. jumbo frames), without dropping a single packet. Furthermore, line rate packet capture is supported irrespective of which packet processing functions (e.g. packet filtering) are enabled. Abundant onboard buffer memory is available to absorb abnormally large bursts of traffic.

Precise Time Stamping

Each packet that enters an Alveo accelerator is tagged with a timestamp with up to 4nS (nanosecond) time precision. The Alveo accelerator must be disciplined from a timing source and there are four options as follows listed in order of popularity:

1. Host OS - The host operating system acts as the time source and can in turn be disciplined by any other source such as PTP, 1PPS, NTP or the like.

2. GPS/CDMA - Alveo accelerator can be directly attached (via a front port) to a 1PPS (One Pulse Per Second) time source such as GPS or CDMA.

3. Another Alveo accelerator - One Alveo accelerator can be the time source for another by attaching them via the onboard "card-to-bus."

4. Free Running - All timing is handled by the Alveo accelerator onboard clock. This is the least precise mechanism but easiest to utilize

Accolade
Technology

An Alveo accelerator can also parse out a Gigamon or Arista generated timestamp and propagate it forward to the host application.

Packet Merging

Multi-port

Data coming into the Alveo on different physical ports (e.g. 2 separate 100G ports) can be optionally merged together into a single (in timestamp order) combined stream of data packets. This is often required when an application needs to analyze both the receive and transmit directions of the same connection, but the data from the respective directions comes in on different Alveo ports. Once data is merged together, the host application can still determine which physical port a given packet came in on based upon a packet descriptor (metadata) that is provided for each received packet.

Multi-accelerator

Packets can also be merged into a single stream from ports on different Alveo accelerators. This merge is accomplished either via an external cable or the "card-to-card" bus available on each accelerator.

Packet Parsing

Each Alveo accelerator has a very powerful and flexible L2/L3/L4 packet parser. The header information from each packet that enters the system is extracted and processed to inform the host application about relevant packet details and also as input for packet filtering.

Tunneling and Encapsulation

The parser is able to recognize various tunneling and encapsulation protocols such as VLAN, VXLAN, MPLS, GTP and GRE. The accelerator is then able to intelligently strip away the tunnel encapsulations and analyze the relevant packet information contained inside the tunnel.

Packet Slicing

If a host application does not require analysis of all packet data in its entirety, packet slicing is a perfect solution to limit the amount of data that an Alveo accelerator presents. Examples where this functionality might be relevant is for traffic engineering, billing or protocol analysis type applications. Packets can be optionally sliced to include no packet data at all (only packet headers) or any arbitrary number of bytes, for example just the first 128 bytes of packet data. Different byte slice lengths can be assigned per port or a single value can be applied to all incoming packets.

IP Fragment Handling

Fragmented IP packets are identified and processed as if they were a single IP packet. For example, the fragments will be classified as the same flow and will be steered to the same host packet buffer.

Packet Filtering

Each packet can be optionally compared against a number of pre-defined filtering rules. A rule can be defined to trigger on most any L2, L3 or L4 header field(s). The most common fields for filtering rules are IP address (source and/or destination), TCP/UDP port (or a range), protocol,

Accolade
Technology

VLAN id and MPLS tag. Rules can be defined so that they trigger on both sides of a bi-directional UDP or TCP session.

Once a packet filter matches, actions such as; drop, U-turn (local retransmit) or steer traffic to a specific host packet buffer can be applied to guide a packet along a desired path.

Deduplication

Most network monitoring appliances receive a significant number of duplicate packets; sometimes as high as 50% of all traffic. This is often because a SPAN port is configured to copy ingress and egress data from every switch port, which leads to duplicate packets for every packet that goes into and then out of a network switch. Alveo accelerators can discard all duplicate packets in hardware before they ever reach the host application, thus saving a tremendous number of wasted processing cycles.

Flow Classification

An Alveo accelerator can optionally perform flow classification on each incoming packet. If this feature is enabled, information on up to 16 million unique flows is maintained in onboard memory as part of a lookup or flow table. For each new flow a unique "flow id" is generated such that the same id is calculated for both directions of a bi-directional flow. Information such as total packet count, total byte count and the last time a packet was seen is maintained for each and every one of the up to 16 million identified flows.

A flow is identified by either a 5-tuple (source IP address, source TCP/UDP port, destination IP address, destination TCP/UDP port and IP protocol) or a 3-tuple (source IP address, destination IP address, IP protocol).

Flow Based Filtering

If flow classification is enabled, an Alveo accelerator can be configured to filter out (i.e. drop) or filter in (i.e. only forward matching flows) the desired traffic flows. This capability enables security features such as blacklist matching. In addition, flows can be marked for steering to a specific host packet buffer(s).
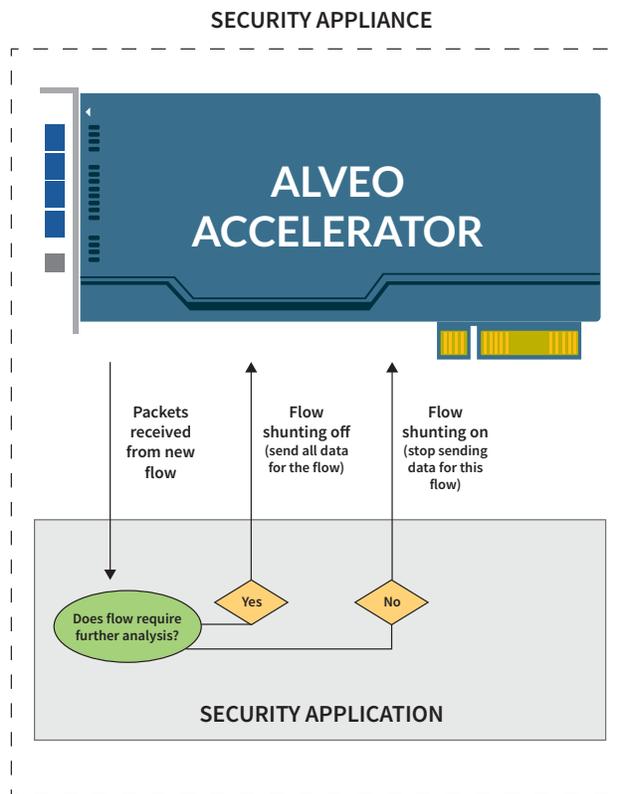
Flow Shunting

At a high level, flow shunting allows an application to programmatically turn packet transmission on or off—for a given flow (based on 5-tuple). In other words, the application can decide from which flow(s) it does and does not want to receive data traffic. By intelligently "toggling" the flow shunting switch, an application can greatly reduce the amount of data it has to analyze, thereby freeing up CPU resources for more pressing tasks.

The diagram to the right clearly illustrates the flow shunting process. Inside a security appliance there is an Alveo accelerator (in PCIe slot) and the security application. The security application communicates with the accelerator via a well-defined API (natively integrated into Suricata, PF_RING etc.) which it uses to configure and control the accelerator. The accelerator classifies each flow and subsequently sends the entire packet (header + data payload) to the application. The application in turn examines the packet (or more likely many packets in a row) and decides whether this particular flow requires further analysis or not. If the flow is not of interest, then the application tells the accelerator to turn flow shunting on or in other words to stop sending any packets from that flow. If for some reason the situation changes, flow shunting

can always be turned off for this flow, in which case packets will resume being forwarded to the application. There may be instances when "toggling" flow shunting on and off is necessary.

There are many reasons why an application may not want to continue receiving traffic for a given flow. For example, if the application cannot process encrypted traffic there is no point in receiving encrypted flows. Or an application may not want to examine video traffic (e.g. NetFlix) because it doesn't pose a threat or wastes too much disk space, so all video traffic could be shunted away. Or perhaps the application has an IP blacklist (or whitelist) on which it operates and therefore any flows which don't match the list should be shunted aside. The value of flow shunting is clearly that it puts control in to the hands of the application, so that dynamic decisions such as which traffic flows should be analyzed can be made based on programming logic.

SECURITY APPLIANCE

ALVEO ACCELERATOR

Packets received from new flow

Flow shunting off (send all data for the flow)

Flow shunting on (stop sending data for this flow)

Does flow require further analysis?

Yes

No

SECURITY APPLICATION

Host Packet Buffer (HPB)

In order to support multi-core CPUs and multithreaded host applications, ALVEO accelerators utilize a flexible host packet buffer (HPB) technique. Host memory is segmented into a number of fixed size blocks. The block size is configurable but is typically 2MB or 4MB each. A collection of these host memory blocks is then dynamically pooled together to form a host packet buffer. A specific application thread (often tied to a CPU core) is then explicitly assigned or linked to a given HPB and will only process data that is transferred into its own HPB. Up to 64 independent HPBs can be created (per Alvro accelerator) and in turn assigned to up to 64 host application threads.

The memory blocks (typically 2MB or 4MB in size) assigned to a host packet buffer (HPB) do not have to be contiguous. In other words, each HPB is composed of blocks of host memory that are randomly spread out in various areas of physical memory. In addition, memory blocks are temporarily assigned to a given HPB by the Alveo accelerator and once an application thread has finished processing all the data from a given memory block, that block can be assigned to a different HPB.

Packet Steering

Steering Methods

An Alveo accelerator is configured to intelligently steer packets in to specific host packet buffers (HPB). The benefit of packet steering is that each thread in a multithreaded application (often utilizing multiple CPU cores) can process packets from its own HPB. In this way a security or

Accolade Technology

networking application can take advantage of parallel processing of data thus achieving higher levels of speed and efficiency.

There are three different ways to steer packets into a HPB:

1. Alveo accelerator is configured to use its own internal algorithms to evenly and efficiently distribute or load balance packets across a specified number (from 1 to 64) of HPBs. This is done to ensure that no processing thread is overwhelmed with data while others are starved.

2. Based upon the results of packet filtering, packets can be steered to specific HPBs. For example, packets that match a specific packet filter rule might all be steered to the same HPB for processing.

3. Based upon flow classification, packets are steered to specific HPBs. In other words, specific flows are identified and explicitly steered to a specific HPB for processing.

U-Turn (Retransmit)

Packet traffic is typically transferred across the PCIe bus (DMA) for consumption by the host application. However there may be circumstances under which select traffic must be locally redirected or retransmitted out of one of the Alveo network ports. Packet filtering or flow classification can be used to identify which specific packets or flows must be redirected out a given port.

DMA (Direct Memory Access)

After all packet processing concludes, an Alveo accelerator efficiently transfers all relevant packets and associated packet descriptors (metadata) across the PCIe bus directly in to host memory for consumption by the host application. The main advantage of DMA is the host CPU is not burdened with memory transfer and hence is available to perform other more important tasks.

Multi-core DMA

Multi-core DMA is a technique that makes processing by multiple host CPU cores more efficient. For example, assume the host Intel CPU has 4 cores (up to 64 cores are supported) with each operating independently of the other 3 cores. The Alveo accelerator is programmed to write data in to 4 independent host packet buffers (HPB) and each CPU core (and related application thread) is in turn programmed to process only its corresponding HPB. In this way a network security or monitoring application can take advantage of parallel processing of data thus achieving higher levels of speed and efficiency.

PF_RING

PF_RING is an open source, Linux kernel module for packet capture that is supported by most open source IPS/IDS solutions such as Snort, Suricata, and Bro as well as other applications such as Wireshark and Argus. PF_RING has native support for Alveo which makes integration with any application that uses PF_RING seamless.

Statistics (RMON1)

Alveo accelerators provide all RMON1 (RFC 2819) related per-port statistics such as packets received, packets in error, dropped packets and broadcast/multicast packet count. The

Accolade
Technology

implementation also goes beyond the RFC to provide a packet count for various packet sizes ranging from the minimum 64 bytes up to and including jumbo frames (greater than 1518 bytes).

ALVEO Onboard Sensors

A variety of information is provided with respect to the condition of the Alveo hardware itself. Some of the information is available via LEDs on the front panel of the accelerator and other information via the software API. The following is a sample of information provided:

- Optical interface (e.g. SFP, CFP) temperature
- FPGA temperature
- PCB temperature (multiple readings from different locations)
- Optical power (e.g. SFP, CFP)
- Ethernet link status
- Time synchronization status

Company Profile

Accolade is the technology leader in FPGA-based Host CPU Offload and 100% Packet Capture PCIe ANIC and Xilinx Alveo SmartNICs/Accelerators as well as the ATLAS Service Nodes. Accolade and Xilinx's Alveo line of 1-100GE products enable 100% packet capture, flow classification, deduplication, packet filtering and more. Our customers are both governmental entities and commercial entities which are global leaders in network monitoring & cybersecurity applications as well as network test/measurement, telecom and the data center space.

Corporate Headquarters:	Silicon Valley:	South East U.S. Regional:
124 Grove Street, Suite 315	980 Mission Court,	2997 Cobb Parkway,
Franklin, MA 02038	Fremont, CA 94539	Atlanta, GA 30339
T: 877.653.1261	T: 877.793.3251	T: 877.897.4833
F: 208.275.4679		

www.accoladetechnology.com

Accolade
Technology