Accolade Technology

Flow Classification

Hardware Based Classification of Up to 32 Million IP Flows

What is an IP Flow?

To limit any potential confusion, let's define what we mean by an Internet Protocol (IP) flow. We define an IP flow as a sequence of packets sent from a source (e.g. a laptop) to a destination (e.g. a web server). Both the source and destination have a unique IPv4 or IPv6 address assigned to them. For the rest of this paper we will often drop the term "IP" and just refer to these sequence of packets as a "flow" with the implicit understanding that we are ONLY talking about Internet Protocol related flows: sometimes also called network flows, packet flows or traffic flows.

One flow is distinguished from another by its header contents. Specifically, we use the following 3 or 5 parameters, referred to as a 3-tuple or 5-tuple respectively, in combination to uniquely identify a flow. The difference between a 3-tuple or 5-tuple is simply whether or not the IP packet in question is a TCP/UDP packet.

1. **Source IP address** (IPv4 or IPv6)
2. **Destination IP address** (IPv4 or IPv6)
3. **IP protocol** (e.g. TCP, UDP or some other such as ICMP, GRE, MPLS, PIM, OSPF, etc.)
4. **Source TCP/UDP port** (if IP protocol is TCP or UDP)
5. **Destination TCP/UDP port** (if IP protocol is TCP or UDP)

For further clarification, the same source and destination pair can have multiple "conversations" going on between them and these will be treated as separate and unique flows if they are communicating with different protocols or TCP/UDP port numbers. For example, two computers may have an IP phone conversation going on between them as well as an email exchange and both transactions will be treated as separate flows and tracked by an ANIC adapter appropriately.

Accolade Flow Classification

Each current generation ANIC packet capture adapter has built in hardware-based flow classification that can be enabled or disabled at any time. This unique functionality is included in the cost of the adapter. The following are a few high-level flow classification characteristics offered with ANIC adapters:

- Can track up to 32 million unique IP flows per adapter
- Actions such as forward, drop or re-direct can be requested on a per flow basis
- Both directions of a flow are tracked and recorded
- Information such as total packet count, byte count and the last time a packet was seen is maintained for every flow

Flow Classification is performed inside the FPGA on an ANIC adapter. Figure 1 shows a visual representation of an ANIC adapter and each green box depicts a unique feature or function that is provided in the onboard FPGA (also referred to as the Advanced Packet Processor).

Implementing these flow classification functions in an FPGA (as opposed to software) enables the highest performance, lowest latency, scalability and precision required in sophisticated, mission critical network monitoring and security applications.

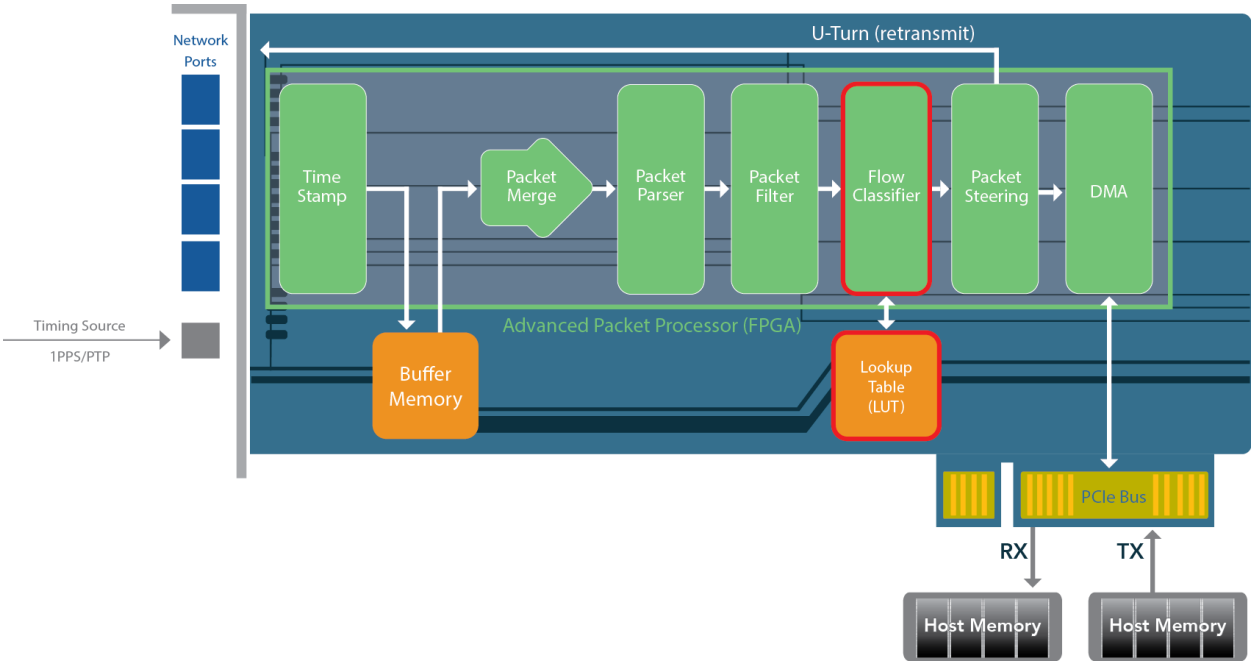

Figure 1: Flow Classification Functionality in ANIC Adapter

The blocks outlined in red in Figure 1 depict the entities inside an ANIC adapter that perform flow classification. Specifically, there is flow classifier logic which sees every packet that comes through an ANIC adapter. This logic inspects each packet and makes flow classification decisions. These decisions result in a flow classification entry being created or updated in a lookup table (LUT); which is physically a bank of memory (DRAM). It is in this LUT or DRAM that flow entries for up to 32 million unique flows are stored. The structure of each flow entry is shown in Figure 2.

For a detailed description of each ANIC feature please visit:
https://accoladetechnology.com/features/

Flow Entry Format

As noted, the flow classification table is stored in ANIC onboard DRAM. The table can accommodate up to 32 million unique IP flows which will correspond to 32 million flow entries. Figure 2 shows a graphical representation of how a **single flow entry** is stored in the flow table.

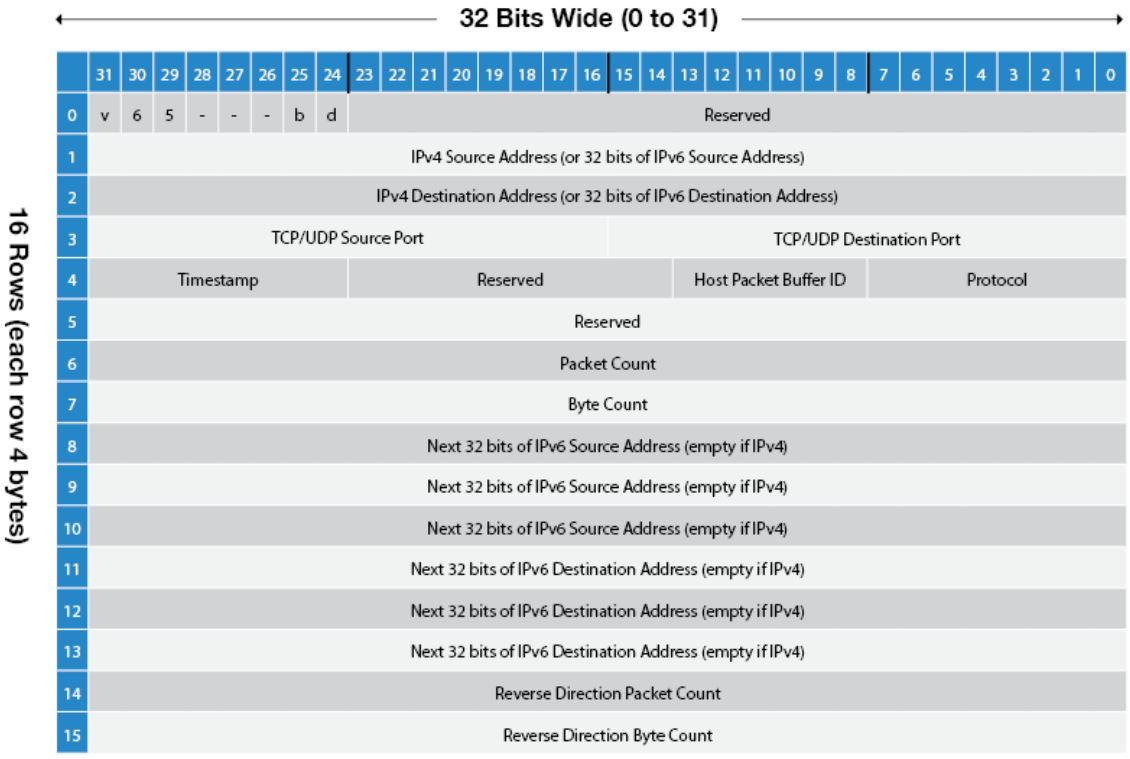

Figure 2: Flow Entry Format

Each flow entry is 64 bytes in size (16 rows of 4 bytes each) and a total of 2GB of DRAM is required in the LUT to store 32 million unique flows. Each field of a flow entry is detailed below, starting in the upper left corner (row 0, bit 31):

"V" bit – Indicates if a flow entry is present at this location. Set to 1 if a flow entry is present and 0 if empty.

"6" bit – Indicates IPv4 or IPv6. Set to 1 if IPv6 and 0 if IPv4.

"5" bit – Indicates whether record contains a 3-tuple or 5-tuple flow. Set to 1 if 5-tuple and 0 if 3-tuple

"B" bit – Indicates if flow should be redirected out a network port on the ANIC adapter. Set to 1 to invoke redirect action and 0 otherwise.

"D" bit – Indicates if flow should be dropped. Set to 1 to invoke drop action and 0 otherwise.

NOTE: If both "B" and "D" bits are 0, then the default action of forwarding flow information to host memory is observed. Also "B" and "D" bits cannot both be 1 as this is an illegal condition.

Row 1 – Contains the 32-bit IP address of the source for this flow. If this is an IPv6 flow, the field contains the first 32 bits (out of 128 bits) of the source IPv6 address. The remaining 96 bits (3 bytes) are contained in rows 8, 9 and 10.

Row 2: Contains the 32-bit IP address of the destination for this flow. If this is an IPv6 flow, the field contains the first 32 bits (out of 128 bits) of the destination IPv6 address. The remaining 96 bits (3 bytes) are contained in rows 11, 12 and 13.

Row 3: Contains the Source and Destination (2 bytes each) TCP or UDP ports. If this is a 3-tuple flow these fields are empty.

Row 4:

> **Timestamp:** A timestamp of the last packet seen in this flow. Note, this is a low fidelity value only used for internal purposes to age out old flows.
>
> **Host Packet Buffer ID:** An identifier for which host packet buffer (HPB) the flow entry should be forwarded to. For more information on HPB please visit: https://accoladetechnology.com/portfolio-item/host-packet-buffer/
>
> **IP Protocol:** Indicates which IP protocol this flow is carrying. TCP and UDP are most common, but others such as ICMP, GRE, MPLS, PIM, OSPF etc. may also be present

Row 5: Reserved for future use

Row 6: Running count of number of IP packets that have been received for this specific flow

Row 7: Running count of the number of bytes of data received for this specific flow

Rows 8, 9, 10: If this is an IPv6 flow, contains 96 bits (of 128 total) of the source IPv6 address. The other 32 bits are in row 1. If this is an IPv4 flow these rows are empty.

Row 11, 12, 13: If this is an IPV6 flow, contains 96 bits (of 128 total) of the destination IPv6 address. The other 32 bits are in row 2. If this is an IPv4 flow these rows are empty.

Row 14: Running count of the number of packets received in the opposite direction (from destination IP to source IP)

Row 15: Running count of the number of bytes of data received in the opposite direction (from destination IP to source IP)

How does it work?

Below is a high-level description of the various mechanisms involved when a single packet arrives in to the flow classifier. Keep in mind that an ANIC adapter must perform these procedures for each and every packet that arrives without dropping or altering even a single packet. All of this must be performed flawlessly at up to 100Gbps speeds.

1. For each packet that arrives in the flow classifier the relevant 3-tuple or 5-tuple (at start up, the host application selects which to evaluate) header fields are extracted and run through a proprietary hash algorithm to produce a 24-bit **flow-id**. The hash algorithm is implemented such that the two directions of a bi-directional flow will have the same flow-id. In other words, if a packet arrives with source IP address A and destination IP address B the flow-id generated will be the same as for a packet with source address B and destination address A; assuming the protocol field and TCP/UDP port numbers (in case of 5-tuple) are the same.
2. The flow-id is used as an index value into the flow table to locate the matching flow entry. If this is the first packet received for a flow (i.e. new flow), the flow-id should point to an empty location ("V" bit set to 0) in the flow table; this location will be populated with a new flow entry using the header information from this first packet.
3. If the flow-id points to an existing flow entry, the header information in the arriving packet is compared to that in the flow entry to make sure there is a match. In the event of a match, the flow entry is updated with information from this newly arrived packet (e.g. increment packet and byte count).
4. If the headers don't match, a so called "**collision chain**" is followed. Simply put, a collision chain is a set of alternate locations for a flow entry. Each alternate location is reviewed to find a match or an empty record. If an empty record is found it is used to establish a new flow. In the unlikely event that neither a matching nor empty flow entry is found, the arriving packet is marked as "unclassifiable".
5. Once a matching or new flow entry is found, several fields are updated. The packet count is incremented by one to account for the packet that just arrived. The byte count is incremented by the number of bytes of data present in the packet. The timestamp is updated with the current time, to indicate that a new packet just arrived for this flow. The timestamp is used to age out old flow entries. The host application can set a flow **idle threshold** (via the API) of between 1 and 60 seconds; this is basically a time out value. A background process continually walks through all established flow entries and compares the timestamp field with the current time and all records that have been idle for a threshold number of seconds are removed from the table to make room for new flows.

6. If a packet arrives that is in the reverse direction (source and destination IPs are swapped) the reverse direction packet and byte count fields are updated along with the timestamp.

Application Programming Interface (API)

We've just articulated how flow classification functions, but how does a host application interface with this powerful feature?

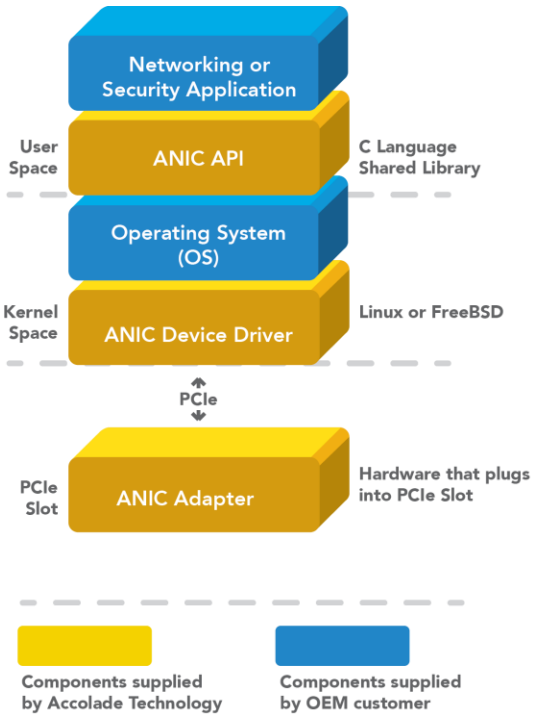

Figure 3: ANIC Stack

As with all Accolade products, the host application interacts with the adapter via an application programming interface (API); a lightweight, C language shared library which is linked to the host networking or security application. In the case of flow classification, a variety of standard API calls are provided to perform functions such as clearing the flow classification table, retrieving a specific flow entry, setting filtering actions (e.g. drop or re-direct specific flows) or setting which flow entries (e.g. only new ones) to forward to the host application. Based on specific customer requirements additional API calls can also be added.

Figure 3 shows a complete depiction of the various components involved to integrate an ANIC adapter into a host appliance.

Use Cases

Inline Deep Packet Inspection (Inline DPI)

Flow classification can be used as a mechanism to selectively drop unwanted flows in live network traffic. The flows could be dropped for many reasons for instance if they are deemed malicious or if they violate some terms of service.

A representative implementation is as follows:

1. A new flow arrives and a table entry is created. This new table entry triggers a **"new flow notification**" - an event to which the host can subscribe. The host application only subscribes to notifications for new flows (not existing flows) to reduce the amount of traffic it must analyze.

Note: In this scenario, all new flows are by default set to redirect ("B" bit set to 1) out an ANIC network port.

2. The host application receives the payload data associated with the new flow and performs deep packet inspection. Based upon the inspection a decision is made to either let the flow continue to be redirected (the default scenario, so no further action required) or decide to drop or block the flow.
3. To drop the flow, the host application would simply make an API call to the ANIC adapter instructing it to drop all subsequent packets that arrive for this flow. The ANIC adapter implements this request by setting the "D" bit for the corresponding flow entry.

Blacklist Matching

Flow classification can be used to drop or block known bad IP addresses (IPv4 or IPv6). In this scenario, the host application provides the ANIC adapter (via the API) a list of IP addresses and if the source IP address of a flow matches one of the IP addresses in the blacklist, the flow is immediately dropped or blocked and the payload data is sent to the host application for analysis.

Company Profile

Accolade is the technology leader in advanced, lossless packet capture and acceleration adapters and OEM acceleration platforms. Accolade's 1-100GE ANIC FPGA-based adapters and ATLAS series of acceleration platforms help accelerate network/cyber security and monitoring applications developed by the world's leading networking companies. ANIC adapters are fully PCIe compliant and seamlessly integrate into standard servers offered by companies such as Cisco, Dell, HP, Super Micro and others. Accolade's OEM customers offer products for network security and monitoring, flow classification, deep packet inspection, network test and measurement, video stream monitoring, high frequency trading (HFT), and more. ID: 170903

Corporate Headquarters:
124 Grove Street, Suite 315
Franklin, MA 02038
T: 877.653.1261
F: 208.275.4679
www.accoladetechnology.com

Silicon Valley:
980 Mission Court,
Fremont, CA 94539
T: 877.793.3251

South East U.S. Regional:
2997 Cobb Parkway,
Atlanta, GA 30339
T: 877.897.4833