Packet-Switching

Packet Forwarding Devices: Types and Implementations

Outline

- Introduction
  - What is a packet forwarding device?
  - Which are its functions?
- Elements of a packet forwarding device
  - Buffers, Fabric, Ports
  - Management, scheduling
- Types of packet forwarding devices

Packet Switching

- Information is split into packets that are sent to the network
- A number of network elements process the packets until they eventually reach their destination
- The processing can include:
  - Performing checks on the packets
  - Marking the packets
  - Dropping the packets
  - Forwarding the packets

Packet Forwarding Devices

- An element that receives packets and processes them with the aim of getting them to the destination
- A wide range of devices:
  - The Internet access router at home
  - An Ethernet switch on a campus or office
  - A core router of the Internet that handles the traffic of millions of users
  - A data center switch
  - With additional functionality: firewall, load balancer, bandwidth manager, NAT, …

Packet Forwarding Devices

- **Depending on the layer at which the forwarding is done the devices have been classified as:**
  - **Switches (or bridges) when they use L2 information, for example Ethernet**
  - **Routers when they use L3 information, for example IPv4 or IPv6**
- **Software Defined Networking (SDN) will make the processing more general and flexible**
- **In fact, most devices already support multiple functionality (L2/L3 forwarding, ACLs, MPLS,...)**

Block diagram of a PFD

Elements of a PFD

- **Ports**
  - **Packets are sent & received over them**
  - **Can have different speeds from a few Mb/s to Tb/s**
  - **Different transmission media, cable, fiber, wireless**
  - **Different physical layer implementation**
- **Decision logic**
  - **Processes the packets, typically the header fields and decides what to do with them**
  - **In many cases processing is done in a sequence of rule tables lookup operations known as <u>pipeline</u>**
  - **Need to work at <u>wire speed</u>**

Elements of a PFD

- **Buffers**
  - **Store the packets until a decision is made at the input**
  - **Store packets until they are sent at the output**
  - **Shared and per port**
- **Fabric**
  - **Moves the packets from the input to the output port**
- **Scheduling**
  - **Decides when to send the packet to enforce QoS**

Elements of a PFD

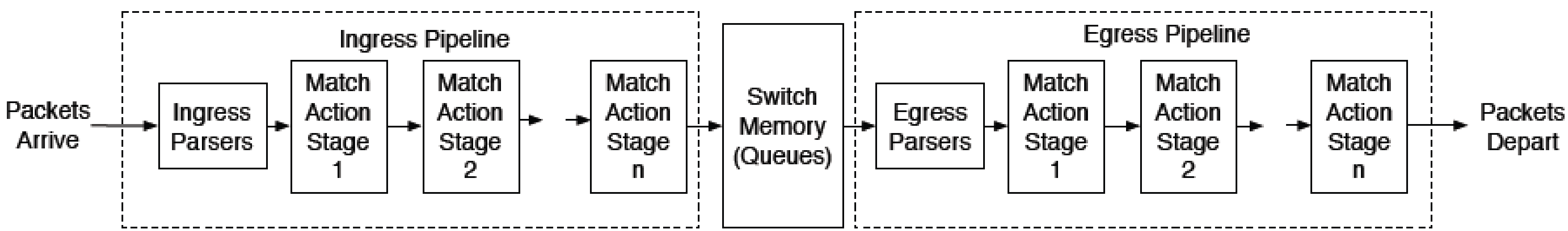
- **Control and Management**
  - Configures the decision logic, buffers, etc
  - Implements the control plane functionality for example routing protocols or SDN agents
  - Processes packets that cannot be handled by the data plane or that have the PFD as source/destination
- **Telemetry**
  - Gathers data of packets and traffic
  - Increasingly important for network monitoring

Decision Logic

- Ethernet switching
- LPM (Longest Prefix Match) “routing”
- Access Control List or Firewalls
- A PFD would in general have more than one of the above in its decision logic
- Another alternative is to have a generic table-based processing as proposed in SDN (Openflow, P4)

Decision Logic

- Parsing packets to extract the information in the headers
- Check tables to obtain actions to be applied to the packet
- Commonly done on ingress and egress



Fabric and buffers

- The interconnection fabric can be as simple as a shared memory(ies) or quite complex
- Buffers are needed per queue. Sharing can be beneficial but can also lead to undesired effects
- Scheduling is commonly linked to buffers and done at the egress port to apply the QoS configured
- Buffering increases delay so it should be used with caution

Control/Management

- Typically implemented in software
- Enables configuring the PFD (tables, buffers, etc.)
- Handles control plane protocols like BGP
- Keeps statistics on performance (evolving to Telemetry that enables detailed logging of events)
- In SDN most of its functionality is offloaded to a central controller and the PFD has only limited functionality

Virtual Routers

- **Implemented in Software and used in Virtual Environments**
- **A number of libraries and tools are used to achieve good performance in terms of latency and packet per second (pps)**
  - **Data Plane Development Kit (DPDK)**
  - **Open Data Plane (ODP)**
  - **Vector Packet Processing (VPP)**

Packet Forwarding Devices

- **Each type of packet forwarding device has different requirements:**
  - **Number of ports and speeds**
  - **Processing complexity and flexibility**
  - **Modularity**
  - **Availability**
  - **Cost**

Number of ports and speeds

- Determines the amount of traffic that we need to process
- Combined with the minimum packet size gives the worst-case number of packets per second (pps) that will be handled by the device
- The number of ports ranges from just a few to hundreds
- The speeds also vary from a few Mb/s to hundreds of Gb/s, soon Tb/s

Performance of routers

- The performance of a router is stated in terms of throughput expressed in b/s.
- However, as routers forward packets, it is more important to know how many packets they are capable of forwarding in a second: packets per second (pps).
- $T = P \times R$
  - P = # of ports
  - R = line rate of each port
  - T = throughput
- $R = S \times Ps$
  - S = packet size
  - Ps = packets per second
- ⇨ T = PxSxPs
- S = ? ⇨ Use smallest size (40 bytes : TCP ACK)
  - ✓ Usually: 64Bytes

Processing complexity

- **What do we need to do with the packets?**
  - It can be as simple as Ethernet switching with only a few MAC addresses
  - It can be complex as
    - ✓ Looking up in an Internet scale routing table with more than 800K entries (https://bgp.potaroo.net/as6447/)
    - ✓ Performing an Access Control List (ACL) check against thousands of rules
    - ✓ Tracking existing connections as in NAT
    - ✓ Performing classification on many packet fields as in SDN
  - It can be worse when several of the above are combined

Processing flexibility

- **What do we need to do with the packets?**
  - **Fixed pipeline that applies the same processing to all packets**
  - **Configurable pipeline that allows a set of predefined functions to be composed to process packets**
  - **Software defined processing that implements a generic set of packet processing primitives that can be used to define the processing**
- **Flexibility vs speed**
  - **In general hardware is fast, software is slow**
  - **Software is flexible, hardware is not**

Modularity

- Has an impact on cost
- Modularity helps with
  - Reliability and availability (redundant modules, module swap)
  - Capacity growth (adding ports)
  - Adding new features
- Commonly used in large devices

Availability

- **What are the effects of the device failure?**
  - **Service degradation/failure**
  - **Number of users affected**
  - **Services affected**
- **Is critical for some network devices/functions**
  - **Use redundant modules or even devices**
  - **Handle failures at network level (re-routing)**
  - **Adding new features**
- **Cost/Benefit analysis**
  - **Improving availability has a cost**
  - **Failures have also a cost**

Cost

- It is always important
- Cost is driven by
  - Research and Development
  - Bill of Materials (BoM)
  - Marketing and sales
- Integration to lower costs
  - System on Chip (SoC) if there is enough volume
  - Reduces PCB size
  - Reduces power consumption and associated costs
  - Initiatives to define common platforms for packet processing devices (https://www.opencompute.org/projects/networking)

Summary of router types

	Performance	Thr O()	Cost	Modularity	Reliability	Sample products	Interfaces	FIB size	Specifics
Residential	low-medium	10Gb/s	very low	none	low	Cisco 1100	Wifi, GPON, Ethernet	1E+02	NAT
SOHO and branch office	low-medium	10Gb/s	very low	low	low	Cisco SOHO 90, 1900, 2900, 3900, 4000, 800	Wifi, GPON, Ethernet	1E+02	NAT, firewall, VPN w/hardware encryption, voice
Enterprise	medium	100Gb/s	low	medium	high	Cisco 3660, 7200, Teldat Atlas	100M,1GE,10GE, WAN	1E+03	VLAN, IP telephony, QoS
Access Concentrators	medium-high	1 Tb/s	medium	high	high	Cisco 6400, 7600	POS: STM1-16; 10-40GEth	1E+04	service provisioning, management
MAN/regional	high	10Tb/s	high	high	very high	Cisco 10700	10Gb/s,100Gb/s	1E+05	integration w/ WDM
Core	very high	10Tb/s	very high	very high	very high	Cisco CRS, NCS6000, Juniper T1600, Nokia 7670RSP	10Gb/s,100Gb/s	1E+06	performance, big route table, ROADM
Data Center / HPC	high	6 Tb/s	high	high	high	Cisco Nexus 5000, Catalyst 4900M, 4948	10G, 40GEth, FCoE	1E+05	cut-through, low latency, FC, programmability
Virtual Routers	medium	10Gb/s	low	very high	high	Cloud Services Router, XRv 9000	100M,1GE,10GE	any	software running on hypervisor

Table only for informative purposes

References

- "Network Routing: Algorithms, Protocols and Architectures. D Medhi and K. Ramasamy. Morgan-Kauffman 2007. "
- "Inside Cisco IOS Software Architecture", R. White, V. Bollapragada and C. Murphy. Cisco Press 2000 (Available in the library).
- Chapters 2 and 3 of "Switching Theory", A. Pattavina. Wiley 1998 (Available in the library).
- "Next Generation Routers" H. Jonathan Chao in Proceedings of the IEEE Septermber 2002. (www.ieee.org)
- "Internet Backbone Routers and Evolving Internet Design" C. Semeria, Juniper Networks 2000. (www.juniper.net)
- "Implementing a Flexible Hardware-based router for the new IP infrastructure" C. Semeria, Juniper Networks 2000. (www.juniper.net)
- "Requirements for IPv4 Routers" RFC 1812 IETF (www.ietf.org)