IGNITION
DESIGN LABS

Large-Scale Networking

Technical White Paper V1.1

Disclaimer

This documentation and all information contained herein (“material”) is provided for general information purposes only. IDL and its licensors make no warranty of any kind, express or implied, with regard to the material, including, but not limited to, the implied warranties of merchantability, non-infringement and fitness for a particular purpose, or that the material is error-free, accurate or reliable. IDL reserves the right to make changes or updates to the material at any time.

Limitation of Liability

In no event shall IDL be liable for any direct, indirect, incidental, special or consequential damages, or damages for loss of profits, revenue, data or use, incurred by you or any third party, whether in an action in contract or tort, arising from your access to, or use of, the material.

Website:

https://ignitiondl.com/

Table of Contents

Chapter 1 Overview

In recent years, with the popularity of mobile applications, the wireless Internet becomes more significant and the WLAN network scale becomes larger. As a result, more and more APs need to be managed.

In the following scenarios, an AC will manage massive APs:

- Thousands to tens of thousands of APs in the campus Wi-Fi network
- Tens of thousands to hundreds of thousands of APs in the wireless city network
- Tens of thousands to hundreds of thousands of APs in IoT
- Hundreds of thousands to millions of APs in carrier-level network

An ordinary AC can manage 4,000 APs at most. In large-scale networking environments, it might need tens of to thousands of ACs to manage APs. This puts lots of stress on the configuration and management of devices.

However, our ACs feature large capacity and high performance, facilitating the management and configuration of devices in large-scale networking environments.

Chapter 2 ACs with Large Capacity and High Performance

The following describes the highest performance of an AC with our PP-8 board card (version 2) in large-scale networking environment:

- Managing 80,000 APs connected with the AC via encrypted CAPWAP DTLS
- Managing 900,000 STAs and user tables
- Managing 40,000,000 Session tables
- The performance of Data Path (DP) is as follows:
  Dealing with up to 8,000,000 data packets per second
  Supporting Max. 80 Gbps forwarding bandwidth per second
- DHCP performance: Assigning up to 3,200 IP addresses per second
- Online rate of users: up to 1,500 users getting online per second

Therefore, just one AC can manage tens of thousands of APs in large-scale networking environment. Even though in the MAN scenarios with hundreds of thousands of APs, several ACs are sufficient.

Chapter 3 Key Technologies

3.1 Main Technologies Involved in DP

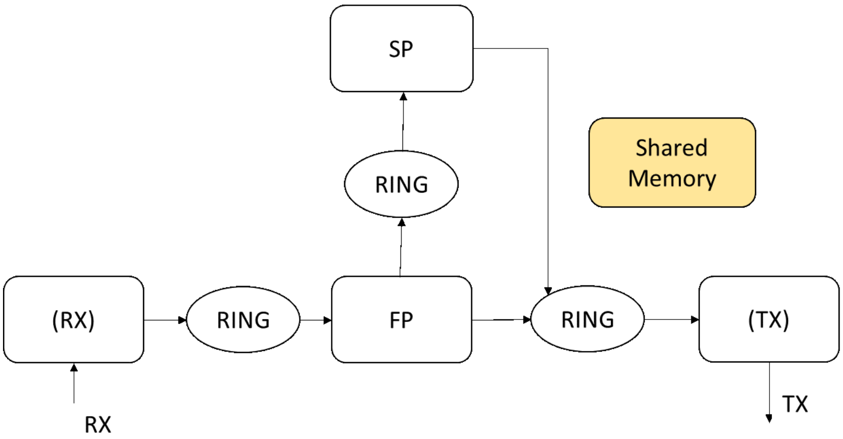
3.1.1 Fast Path (FP)

Our DP achieves the fast forwarding of user data by using DP.

The AC modules can be classified into CP (Control Path), SP (Slow Path), and FP based on the functions and each module running on different CPUs.

Specifically,

- The FP engages in the fast forwarding of data packets and specializes in forwarding services without disturbing the streamline of CPU. Therefore, it satisfies the DP performance requirements in large-scale AP networking scenarios.
- The SP takes charge of the creation of DP tables such as Session table, Bridge table, and ARP table. The SP synchronizes all the DP tables to FP by sharing memories with each other, so it possesses high performance.
- The CP is responsible for interaction with external complex protocols, such as the creation of the user table, Portal authentication, and the assignment of DHCP addresses.



3.2 Main Technologies Involved in the CP

Since ACs manage massive APs, large capacity and high performance are required for the processes of CPs. Here are the key technologies used in the CP:

3.2.1 Multithreaded Technologies and Lock-Free Technologies

The CPUs used on ACs have several cores and CPU threads. For instance, our company's PP-8 board card has two Intel CPUs which contain 16 cores and 32 CPU threads.

Generally, the single-threaded technology of CP's applications cannot meet the requirements of performances in large-scale networking. Therefore, each process will use the multi-threaded technology to utilize the CPU, thus improving the performance.

The multi-threaded technology will raise issues like resource competition and concurrency. So the mutex is required in case that multiple threads need to access the shared variables. However, the mutex will dramatically degrade the performance of the multi-threaded applications.

Instead of using the mutex, our products employ following technologies to handle different sharing data, thus achieving the multi-threaded, high-performance, and high-concurrency data process.

Global HASH Table Lock-Free Mechanism

Generally, there are frequent access to the internal global HASH table of an AC process, involving the user table, AP table, BSS table, STA table and Lease table If accessing the global HASH table requires a mutex, the global HASH table will become a bottleneck.

As a result, we adopt the Lock-Free global HASH table access technology to assign the global HASH table into each thread based on the IP address or MAC address. In this way, each thread is allowed to access its own HASH buckets rather than accessing the HASH buckets based on the management of thread, thereby achieving the Lock-Free mechanism. To implement the traversal of the HASH table, a dispatch thread is required to distribute the information to each processing thread and then collect the traversed data from each processing thread.

For example, the following processes adopt the global HASH table Lock-Free mechanism:

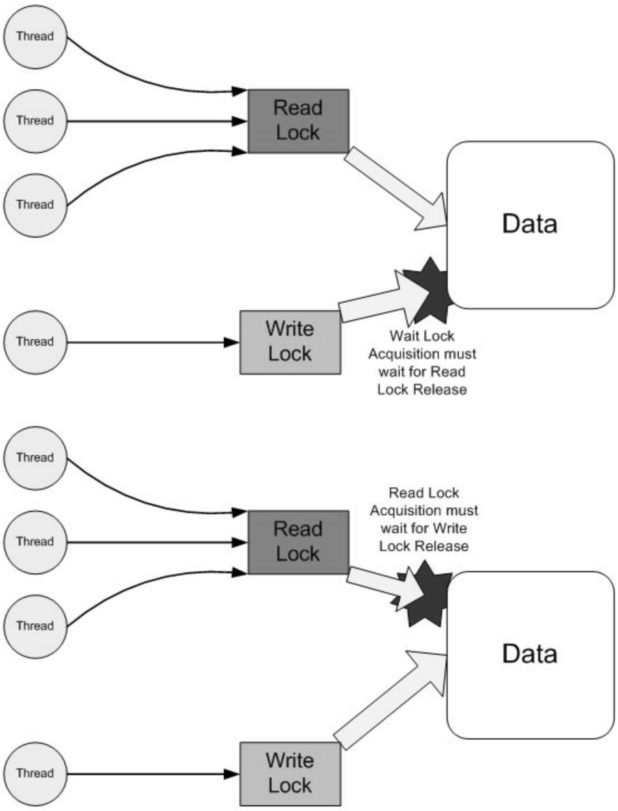
- The Hostapd process implements the HASH operation and assigns processing threads on the basis of MAC address.
- The DHCP process implements the HASH and assigns processing threads on the basis of IP address.

Small-grained Read-Write Lock Mechanism

There might be several global HASH tables in some processes. For example, the CAPWAP process has the AP table, BSSID table, and STA table. In this situation, the complete lock-free multi-threaded mechanism cannot be achieved. If the AP table uses the lock-free mechanism by the HASH, the BSS table and STA table have to add locks.

The working principle of the read-write lock is shown as follows and suits for the conditions with more reads and fewer writes.

Figure 3-1 Working principles of the read-write lock


Thread
Thread
Thread
Read Lock
Data
Thread
Write Lock
Wait Lock Acquisition must wait for Read Lock Release
Thread
Thread
Thread
Read Lock
Read Lock Acquisition must wait for Write Lock Release
Data
Thread
Write Lock

The small-grained read-write lock mechanism is adopted to ensure the high performance while the lock is created.

- The read-write lock mechanism guarantees the shared read (query or traversal) among several threads. Only editing the HASH table (adding and deleting tables) cannot be shared. Thus, parts of the HASH table can still be shared while the lock is used.
- The granularity of the marginal resources protected by the read-write lock should be as small as possible to reduce resource conflicts. For example, a read-write lock should protect a HASH bucket of one HASH table rather than a global HASH table.
- Take the CAPWAP process for example. An STA table contains 50,000 buckets which need to be protected by 50,000 locks accordingly. The CAPWAP process has 5 processing threads. If each thread handles the STA packets randomly (not complete random actually), the ratio that two threads simultaneously access the same HASH bucket is 0.0001. In this way, the read-write lock's performance improves greatly.

Atomic Operations of the Counter

For those statistical counters like the statistics of the packets' number, it will affect the performance heavily if each thread involves the increments or decrements and requires the lock protection.

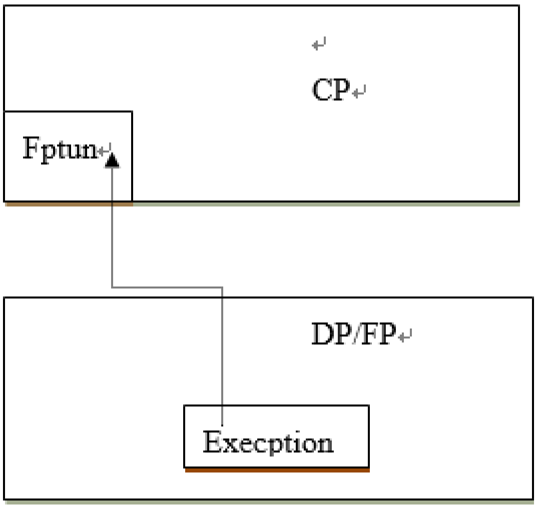
The AC defines a set of atomic operations based on Int32 and Int64 via assembly languages to achieve the autoincrements and autodecrements of the counter, thus avoiding performance penalty caused by the creating of counter's lock.

3.3 Firewall-CP for Protecting the CP

In large-scale networking environments, lots of packets from APs and users are handled, accompanying excessive packets asking for higher handling ability of the CP or DoS attacks from malicious users. Thus, the AC uses the Firewall-CP to guarantee the normal services of CP, thereby serving for most users. Specifically,

When the DP of an AC receives protocol packets (known as exception, involving DHCP, CAPWAP, OSPF, RADIUS, VRRP, SSH, and Telnet) that it cannot handle, it directly sends the packets to the CP over the FPTUN interface. However, a malicious user might send massive protocol packets to exhaust all CPUs or memory resources of the CP.

Figure 3-2 Schematic diagram of DP Exception



ACs use the firewall to prevent CPs from malicious traffic in terms of the following two parts.

3.3.2 ACL Protection for Firewall CP

ACL protection for firewall CP is implemented by using the ACL to control users' access permissions. The processes are as follows:

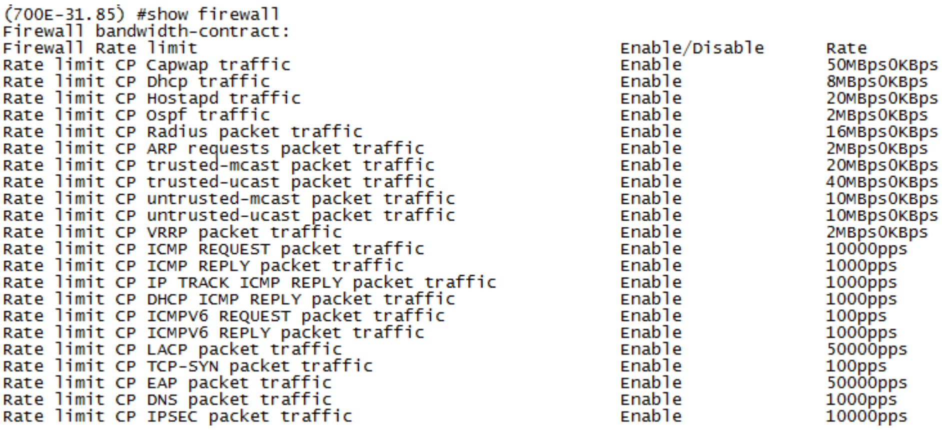
- Users are not allowed to access the Telnet service of the CP.
- Only users from 192.168.0.0/24 are allowed to access the SSH service.

3.3.3 ACL Permit Service

To prevent malicious attacks, the rate-limiting function for preventing DoS attacks is developed, and only packets that can be processed by the CP per second are sent to the CP.

The following figure illustrates the default value of a certain AC’s firewall:

Figure 3-3 Default value of a certain AC’s firewall



```
(700E-31.85) #show firewall
Firewall bandwidth-contract:
Firewall Rate limit                                         Enable/Disable      Rate
Rate limit CP Capwap traffic                                Enable              50MBps0KBps
Rate limit CP Dhcp traffic                                  Enable              8MBps0KBps
Rate limit CP Hostapd traffic                               Enable              20MBps0KBps
Rate limit CP Ospf traffic                                  Enable              2MBps0KBps
Rate limit CP Radius packet traffic                         Enable              16MBps0KBps
Rate limit CP ARP requests packet traffic                   Enable              2MBps0KBps
Rate limit CP trusted-mcast packet traffic                  Enable              20MBps0KBps
Rate limit CP trusted-ucast packet traffic                  Enable              40MBps0KBps
Rate limit CP untrusted-mcast packet traffic                Enable              10MBps0KBps
Rate limit CP untrusted-ucast packet traffic                Enable              10MBps0KBps
Rate limit CP VRRP packet traffic                           Enable              2MBps0KBps
Rate limit CP ICMP REQUEST packet traffic                   Enable              10000pps
Rate limit CP ICMP REPLY packet traffic                     Enable              1000pps
Rate limit CP IP TRACK ICMP REPLY packet traffic            Enable              1000pps
Rate limit CP DHCP ICMP REPLY packet traffic                Enable              1000pps
Rate limit CP ICMPV6 REQUEST packet traffic                 Enable              100pps
Rate limit CP ICMPV6 REPLY packet traffic                   Enable              1000pps
Rate limit CP LACP packet traffic                           Enable              50000pps
Rate limit CP TCP-SYN packet traffic                        Enable              100pps
Rate limit CP EAP packet traffic                            Enable              50000pps
Rate limit CP DNS packet traffic                            Enable              1000pps
Rate limit CP IPSEC packet traffic                          Enable              10000pps
```

NOTE

CP processing capabilities and the default value of firewall of different ACs vary.

The Firewall uses Token Bucket to achieve the rate limiting. Token Bucket is commonly used in Traffic Shaping and Rate Limiting. Typically, Token Bucket controls the quantity of the data sent to the network and allows the Burst data to be sent.

The working process consists of three phrases:

- Generate the Tokens.
- Consume the Tokens.
- Decide whether the packets are allowed to pass.

Two parameters are involved:

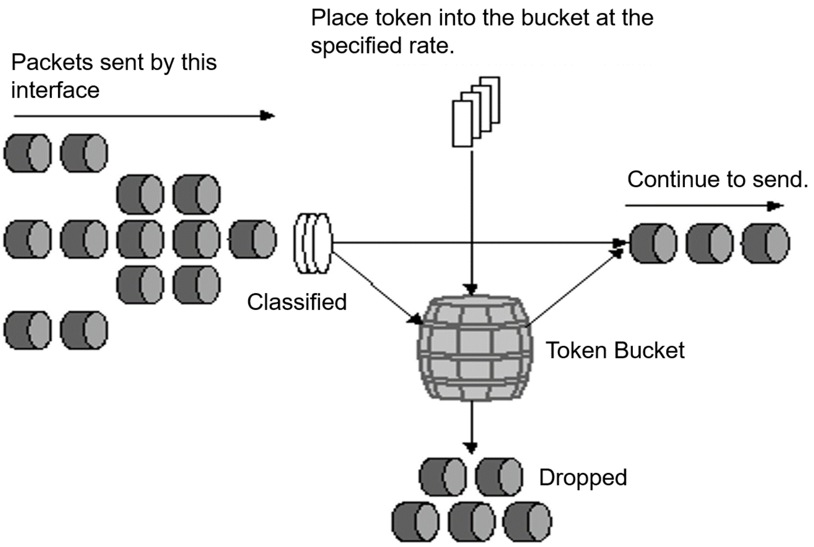
- Rate of generating the Tokens: CIR (Committed Information Rate) / EIR (Excess Information Rate)
- Size of Token Bucket: CBS (Committed Burst Size) / EBS (Excess Burst Size).

Step 1 Generate the Tokens: to periodically increase the Tokens at the rate of CIR/EIR. If the number of the Tokens exceeds the value of CBS/EBS, extra Tokens will be dropped.

Step 2 Consume the Tokens: the input packets will consume the Tokens in the Bucket. During network transmission, the size of packets is different. Larger packets consume more Tokens.

Step 3 Decide whether the packets are allowed to pass: after the input packets pass through Token Bucket, the packets can be output packets and dropped packets. When the number of the Tokens in the Bucket meets the requirement of packets, the packets will be put out, otherwise they will be dropped.

Figure 3-4 Schematic diagram of Token Bucket


Place token into the bucket at the specified rate.
Packets sent by this interface
Continue to send.
Classified
Token Bucket
Dropped

3.4 Level-2 AC Networking

Although the AC uses various technologies to optimize the performance of CP and DP, one AC has limited capabilities to process data. Therefore, in some scenarios, a level-2 AC network with separated CP and DP is used.

For details, please refer to the *Distributed Second-level AC Networking Technical White Paper*.

Chapter 4 Conclusions

One AC can manage industry’s highest number of APs and STAs, thereby deploying fewer ACs while providing the same user experience in large-scale networking.

Fewer ACs can bring customers the following benefits:

- Lower cost of deployment
- Lower cost of configuration and management
- Free from complex scenarios of AC roaming

A Abbreviations

Abbreviations	Full Name
AC	Access Controller
AP	Access Point
CAPWAP	Control And Provisioning of Wireless Access Points Protocol Specification
CIR	Committed Information Rate
CP	Control Plane
DHCP	Dynamic Host Configuration Protocol
DoS	Disk Operating System
DP	Data Plane
DTLS	Datagram Transport Layer Security
OSPF	Open Shortest Path First
RADIUS	Remote Authentication Dial In User Service
SSH	Secure Shell
STA	Station
VRRP	Virtual Router Redundancy Protocol
WLAN	Wireless Local Area Network