INTEL DDP BEST PRACTICE IN CERTUSNET VBNG

Zhang Chixiang   Tech lead/Certusnet
Ma Jianwei   PAE/Intel
Sep. 7th 2019

Legal Disclaimer

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL'S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS' FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENT IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: http://www.intel.com/design/literature.htm%20 Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit Intel Performance Benchmark Limitations

All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.

, Intel, Intel logo, Intel Core, Intel Inside, Intel Inside logo, Intel Ethernet, Intel QuickAssist, Intel Flow Director,, Intel Solid State Drives, Intel Intelligent Storage Acceleration Librariy, Itanium,, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

64-bit computing on Intel architecture requires a computer system with a processor, chipset, BIOS, operating system, device drivers and applications enabled for Intel® 64 architecture. Performance will vary depending on your hardware and software configurations. Consult with your system vendor for more information.

No computer system can provide absolute security under all conditions. Intel® Trusted Execution Technology is a security technology under development by Intel and requires for operation a computer system with Intel® Virtualization Technology, an Intel Trusted Execution Technology-enabled processor, chipset, BIOS, Authenticated Code Modules, and an Intel or other compatible measured virtual machine monitor. In addition, Intel Trusted Execution Technology requires the system to contain a TPMv1.2 as defined by the Trusted Computing Group and specific software for some uses. See http://www.intel.com/technology/security/ for more information.

Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, virtual machine monitor (VMM) and, for some uses, certain platform software enabled for it. Functionality, performance or other benefits will vary depending on hardware and software configurations and may require a BIOS update. Software applications may not be compatible with all operating systems. Please check with your application vendor.

* Other names and brands may be claimed as the property of others.

Other vendors are listed by Intel as a convenience to Intel's general customer base, but Intel does not make any representations or warranties whatsoever regarding quality, reliability, functionality, or compatibility of these devices.   This list and/or these devices may be subject to change without notice.

Introduction to Dynamic Device Personalization

TRANSFORMATIONS IN PACKET CLASSIFICATION REQUIREMENTS
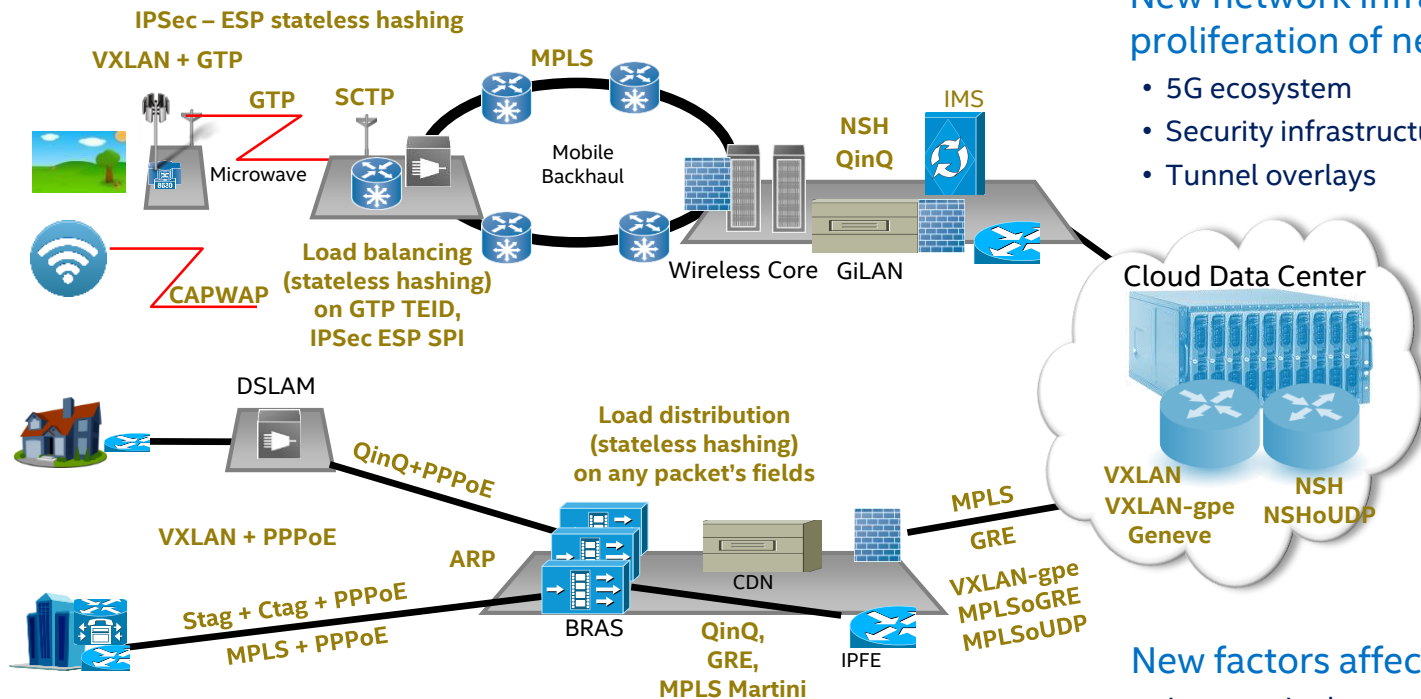
New network infrastructure is driving proliferation of new packet types

- 5G ecosystem
- Security infrastructure
- Tunnel overlays

IPSec – ESP stateless hashing

VXLAN + GTP

GTP

SCTP

MPLS

IMS

NSH
QinQ

Microwave

Mobile
Backhaul

CAPWAP

Load balancing
(stateless hashing)
on GTP TEID,
IPSec ESP SPI

Wireless Core   GiLAN

Cloud Data Center

DSLAM

Load distribution
(stateless hashing)
on any packet's fields

QinQ+PPPoE

VXLAN + PPPoE

ARP

MPLS
GRE

VXLAN
VXLAN-gpe
Geneve

NSH
NSHoUDP

Stag + Ctag + PPPoE

MPLS + PPPoE

BRAS

CDN

QinQ,
GRE,
MPLS Martini

VXLAN-gpe
MPLSoGRE
MPLSoUDP

IPFE

New factors affecting performance

- Increase in the amount of traffic
- Added protocol complexity
- Increased packet processing latency

WHAT IS DYNAMIC DEVICE PERSONALIZATION (DDP)?*

1. Loadable Packet Classifications (Parser) personalization for any well-defined traffic profiles for VNFs or appliances with a specific position in the network.

- **Extended support for protocols:**
  - GTP tunnel support, GTP-C and GTP-U as a separate flow types
  - IP protocols as new flow types, for example L2TPv3, ESP/AH for IPSec
  - Legacy protocols: PPPoE, PPPoL2TPv2
  - New protocols/standards: QUIC, Radio over Ethernet (RoE), MQTT-SNoUDP, SRv6
- **Modify existing protocols:**
  - VXLAN with IP payload, NSHoUDP
  - New PTYPEs on RX descriptor: for example, IPv4 TCP in one MPLS label
- **Make Smart NIC act no so Smart:**
  - Disable GRE classification
  - Disable IPv4 classification

2. Configurable tunnel filters for traffic steering (Classifier Switch)
  - Steering packets to a VM on QinQ
  - Steering packets to a VM on Tunnel ID (VXLAN, NVGRE, MPLS, GTP, NSH*)
  - User definable steering/mirroring rules, for example, VLAN or DSCP to a queue
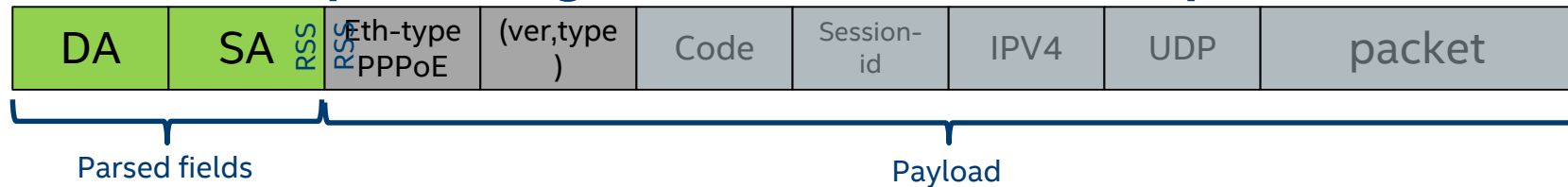
* Does not include Intel® Ethernet Connection X722

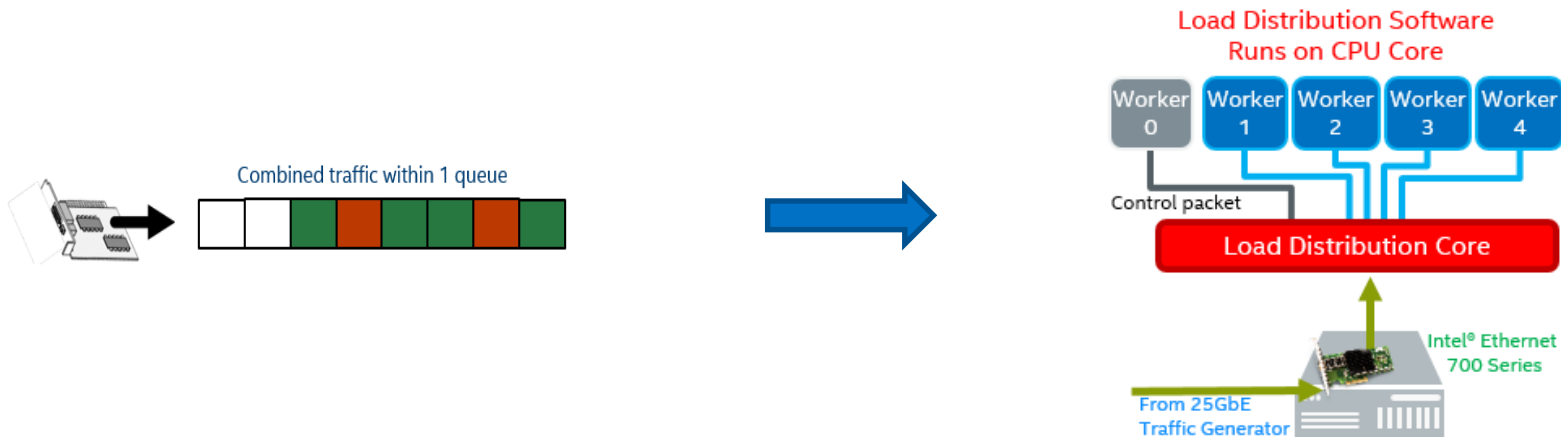# INTEL® ETHERNET 700 SERIES RX PROGRAMMABLE PIPELINE
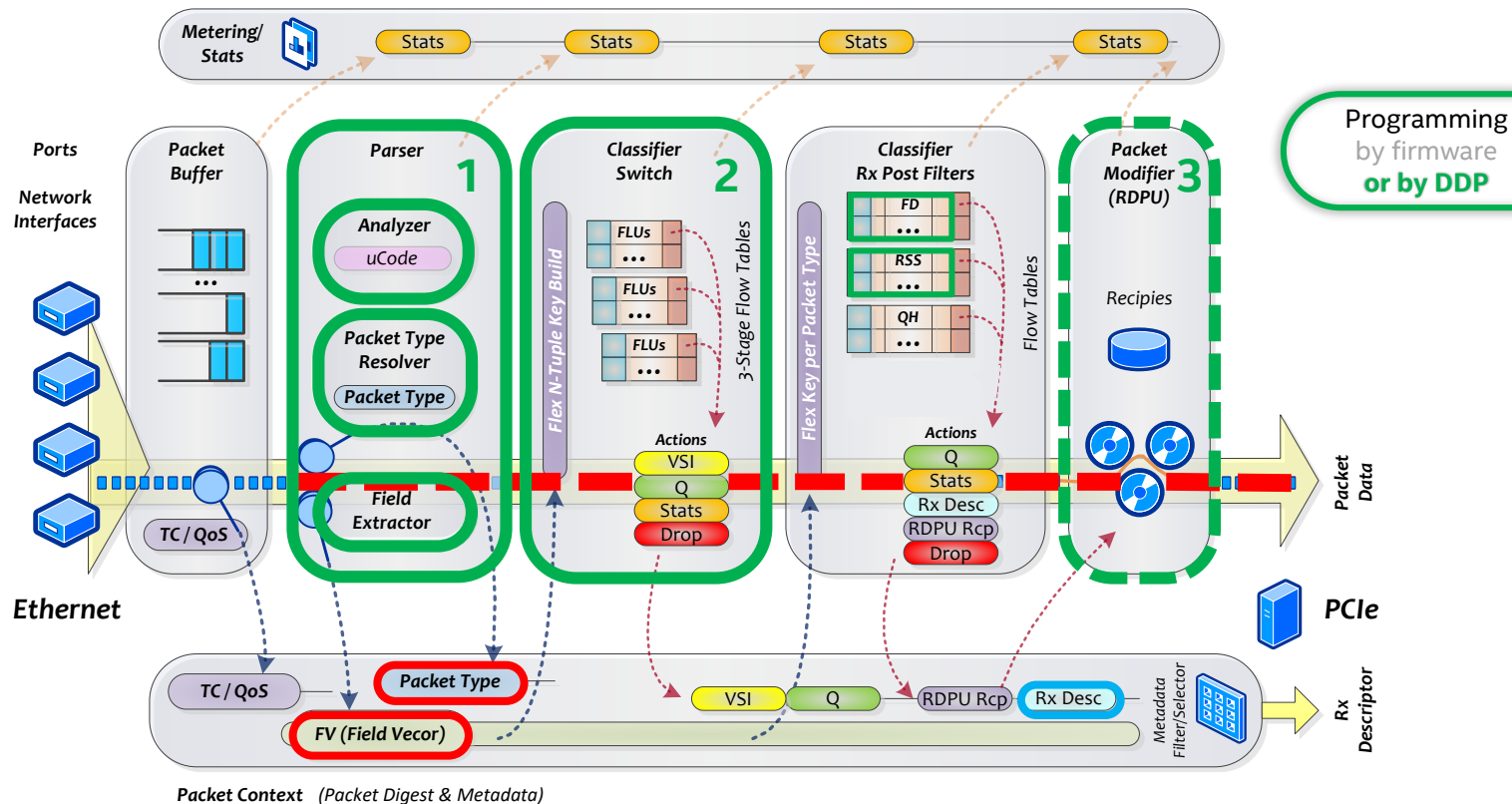
DYNAMIC DEVICE PERSONALIZATION PPPOE EXAMPLE

Default FVL6 processing (without Personalization profile):

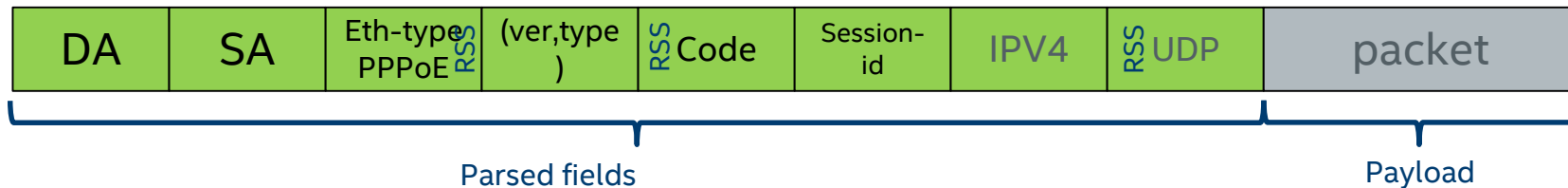| DA	SA	RSS	RSS	Eth-type PPPoE	(ver,type)	Code	Session-id	IPV4	UDP	packet

Parsed fields

Payload

PPPoE is unknown flow type, so no RSS, FDIR or other filters are possible on inner packet fields. FVL classifies PPPoE (Discovery and Session) packets as L2 payload.
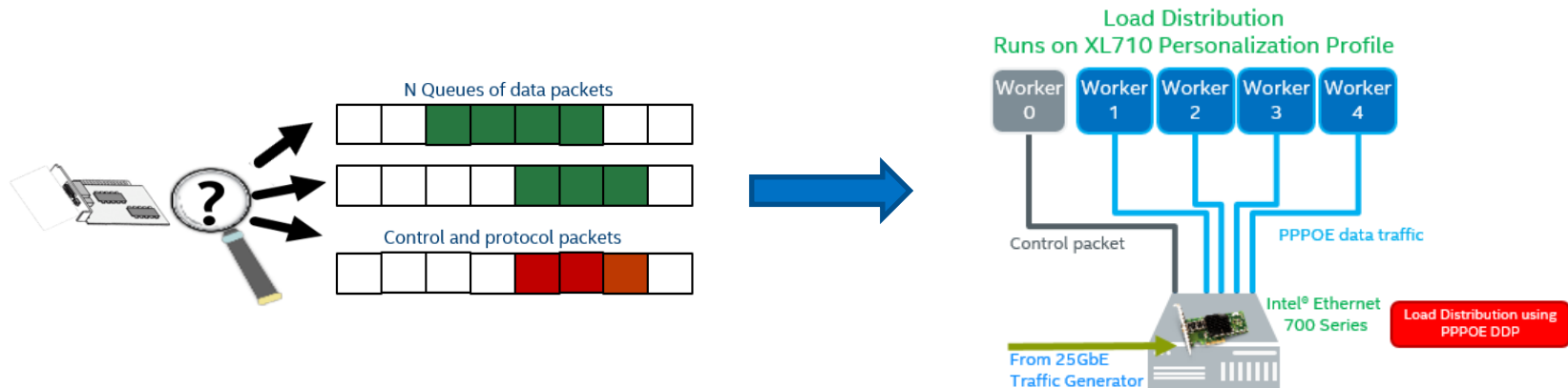

Load Distribution Software Runs on CPU Core

Worker 0 | Worker 1 | Worker 2 | Worker 3 | Worker 4

Control packet

Load Distribution Core

Intel® Ethernet 700 Series

From 25GbE Traffic Generator

Combined traffic within 1 queue

DYNAMIC DEVICE PERSONALIZATION PPPOE EXAMPLE

FVL6 with Personalization profile applied:

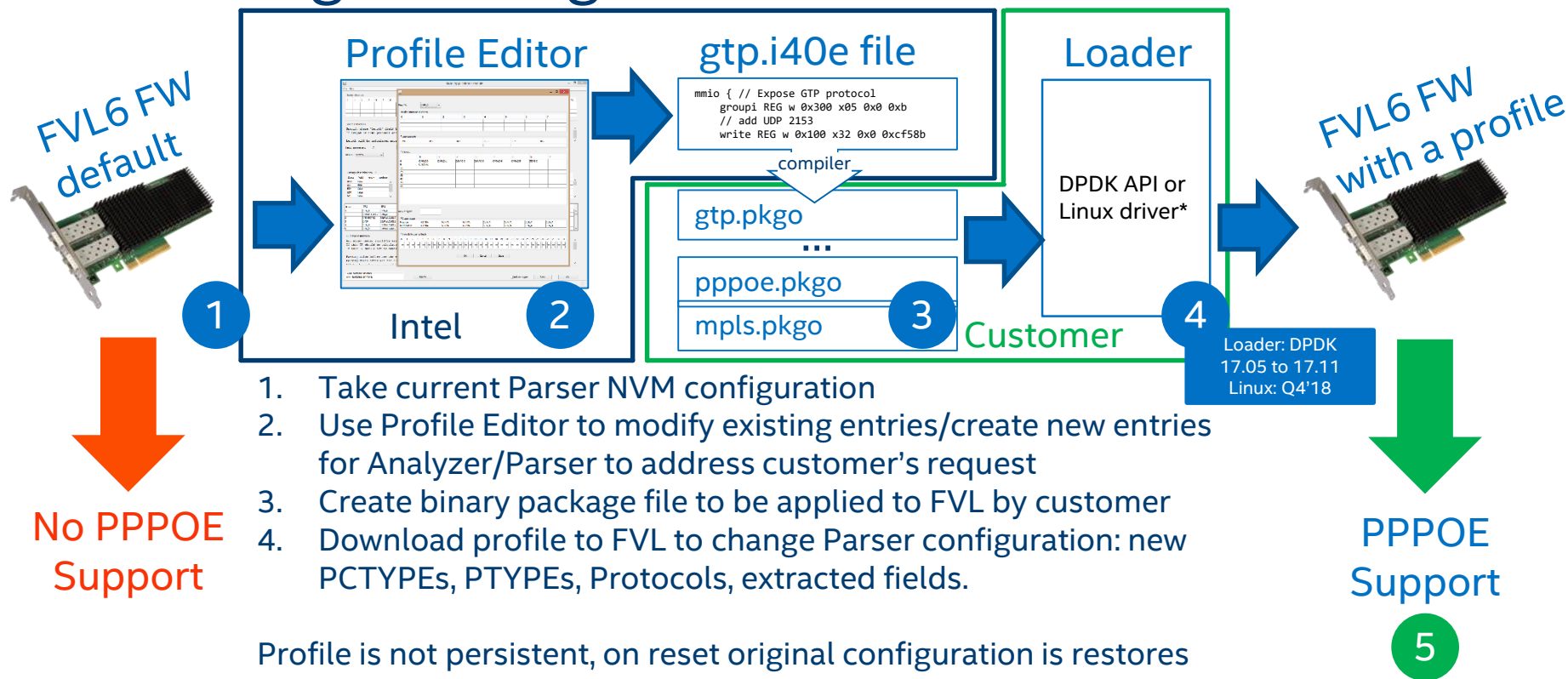| DA	SA	Eth-type PPPoE	RSS	(ver,type)	RSS Code	Session-id	IPV4	RSS UDP	packet

Parsed fields — Payload

With a DDP profile, PPPoE header can be parsed and encapsulated packet classified.
For example, for PPPoE session with UDP in IPv4, as shown above, RSS and FDIR filters can be used on UDP.
If required, RSS, FDIR and Tunnel filters can use PPPoE Session ID as well.

N Queues of data packets

Control and protocol packets

Load Distribution
Runs on XL710 Personalization Profile

Worker 0
Worker 1
Worker 2
Worker 3
Worker 4

Control packet

PPPOE data traffic

Intel® Ethernet
700 Series

Load Distribution using
PPPOE DDP

From 25GbE
Traffic Generator

DDP Programming Tool Chain

FVL6 FW default

Profile Editor

Intel

```
mmio { // Expose GTP protocol
    groupi REG w 0x300 x05 0x0 0xb
    // add UDP 2153
    write REG w 0x100 x32 0x0 0xcf58b
```

gtp.i40e file

compiler

Loader

DPDK API or Linux driver*

FVL6 FW with a profile

gtp.pkgo

...

pppoe.pkgo

mpls.pkgo

Customer

1 **2** **3** **4**

Loader: DPDK 17.05 to 17.11 Linux: Q4'18

No PPPOE Support

PPPOE Support

5

1. Take current Parser NVM configuration
2. Use Profile Editor to modify existing entries/create new entries for Analyzer/Parser to address customer's request
3. Create binary package file to be applied to FVL by customer
4. Download profile to FVL to change Parser configuration: new PCTYPEs, PTYPEs, Protocols, extracted fields.

Profile is not persistent, on reset original configuration is restores from the NVM. SW can restore original configuration without reset as profiles support "roll-back" transaction.

DDP can be applied in Transaction like mode
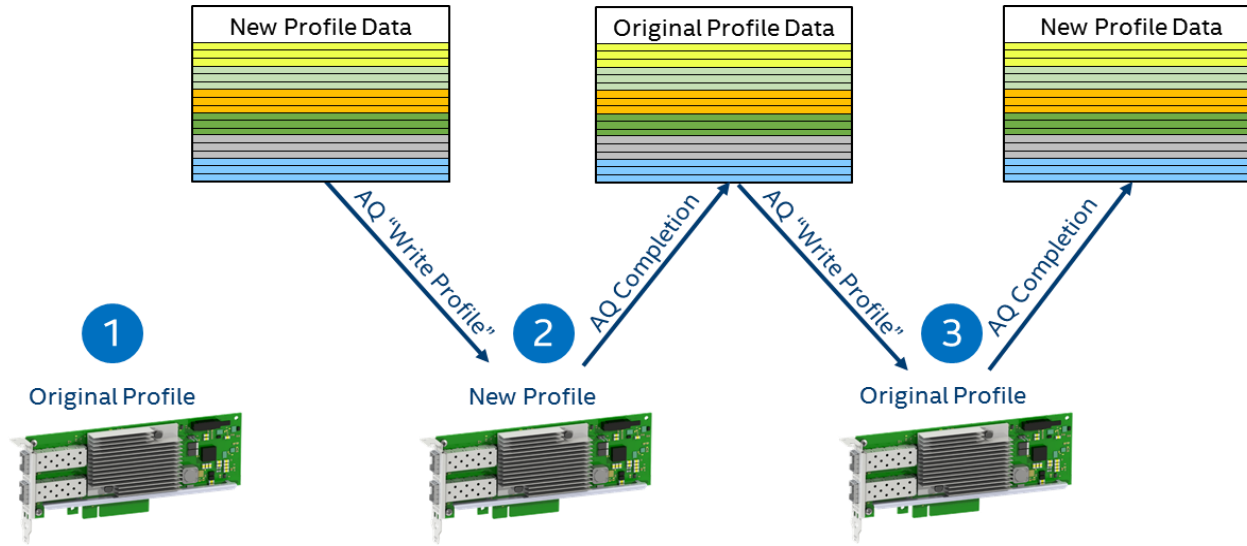
No need to reset NIC to restore original configuration:
"Write Profile" AQ returns original values ready to be applied.

Profile can be applied/removed with traffic running*.

* some packet loss is expected during profile download/rollback

赛特斯信息科技股份有限公司

软件定义通信

打造"软件定义一切"的产业生态体系

提供涉及云、管、端多个业务环节的解决方案与服务

VBNG产品功能特性

接口管理模块

报文处理模块

路由协议模块

DPI功能模块

配置管理模块

FW功能模块

BRAS接入控制模块

BASE Platform

| 业务与管理 | HQoS | EMS | CGN | ACL |
| | FW | ... | | |

| 认证与计费 | PPPoX | IPoX | AAA | Radius |
| | ... | | | |

| 管道协议 | MPLS | L3VPN | VPLS | L2TP |
| | VxLAN | IPsec | ... | |

| 基础协议 | OSPF | ISIS | BGP | ARP |
| | EVPN | ... | | |

intel
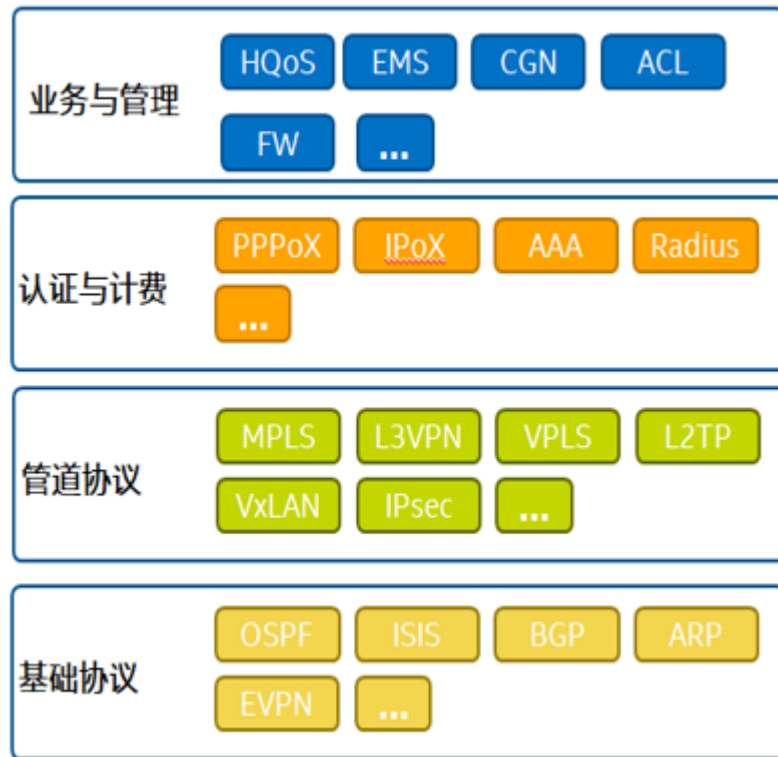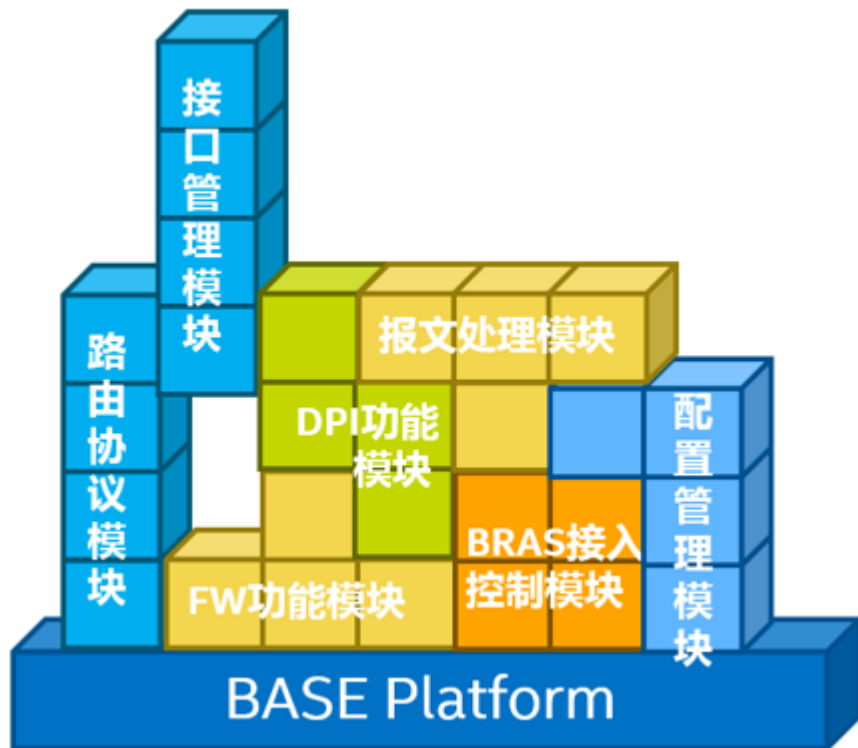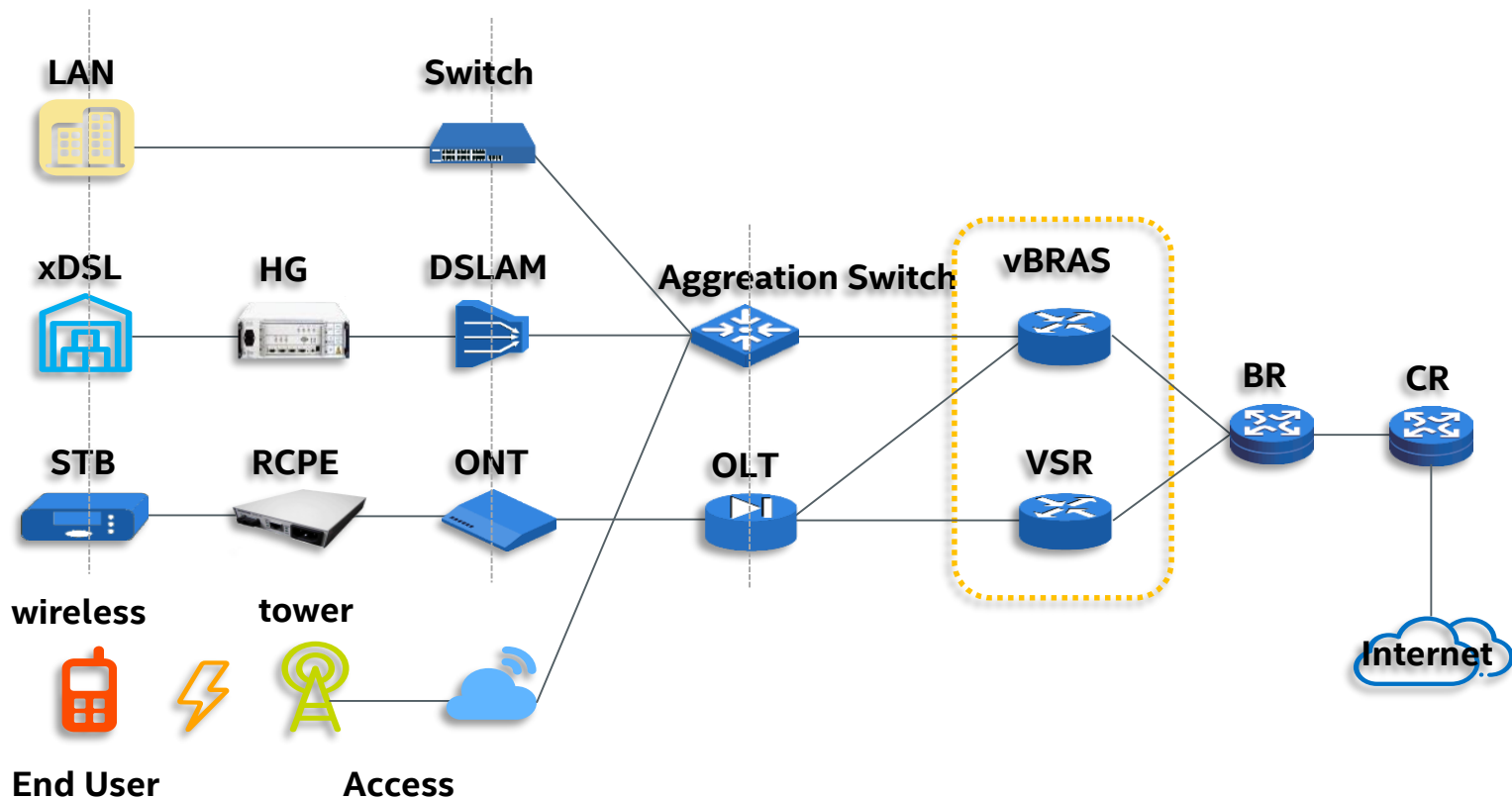
SD-WAN相关组件

赛特斯 *CertusNet*

NFV编排管理系统

白盒、虚拟化CPE

FlexEdge硬件盒子

MANO
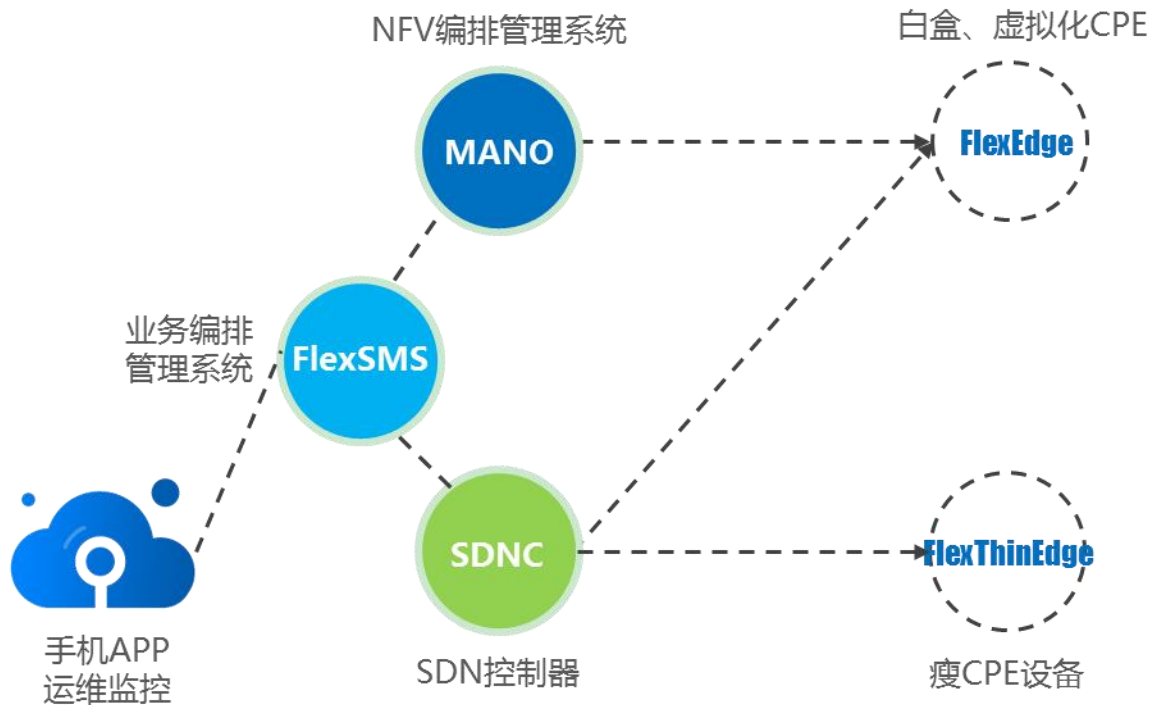
FlexEdge

VNF（FlexEdge）可安装在：FlexEdge硬件
盒子上、物理服务器上、云平台虚拟机上

业务编排
管理系统

FlexSMS

FlexThinEdge ARM硬件盒子

SDNC

FlexThinEdge

手机APP
运维监控

SDN控制器

瘦CPE设备

HC-EN100

HC-EN1000

intel

应用场景·多分支总部互连

某知名集团，

- 企业分支和总部数据中心互联，适应办公地点频繁迁移，打造企业专网体验。
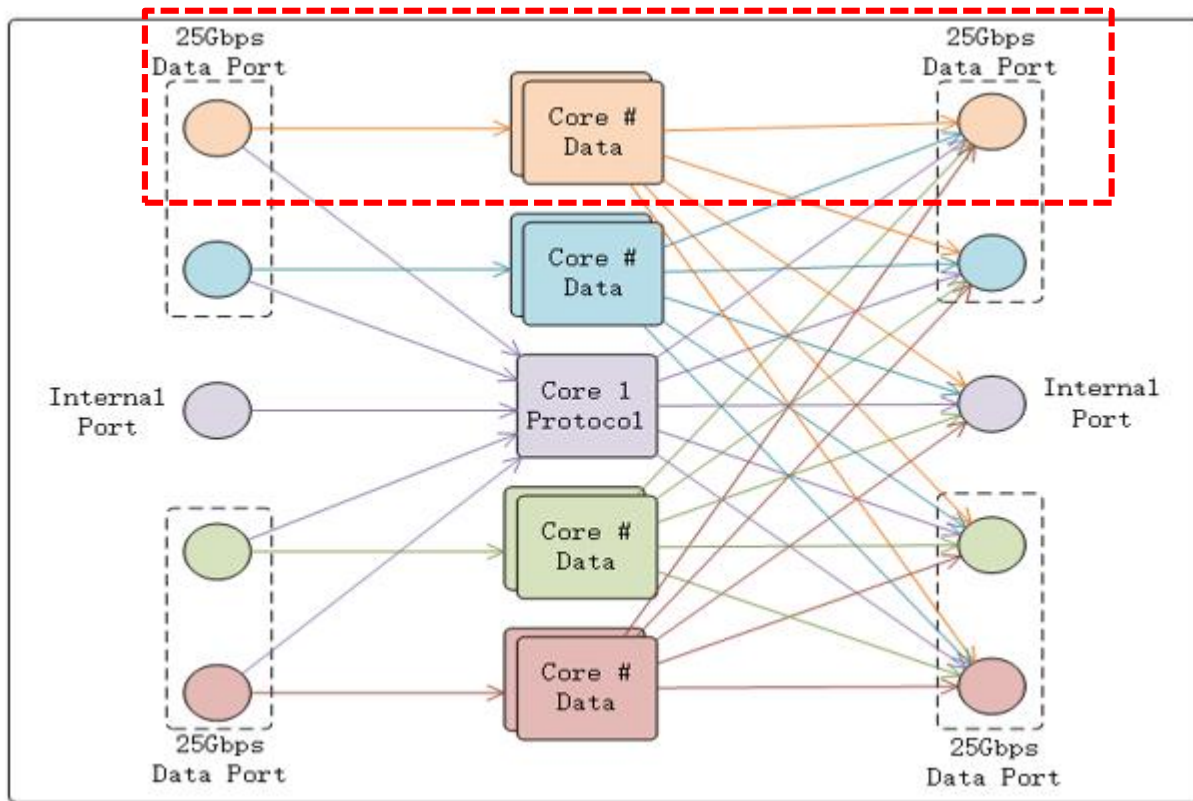- 节点数：50+。

某知名家居企业，

- 全国分支互联，取代原有SDH专有
- 线路，节约成本。
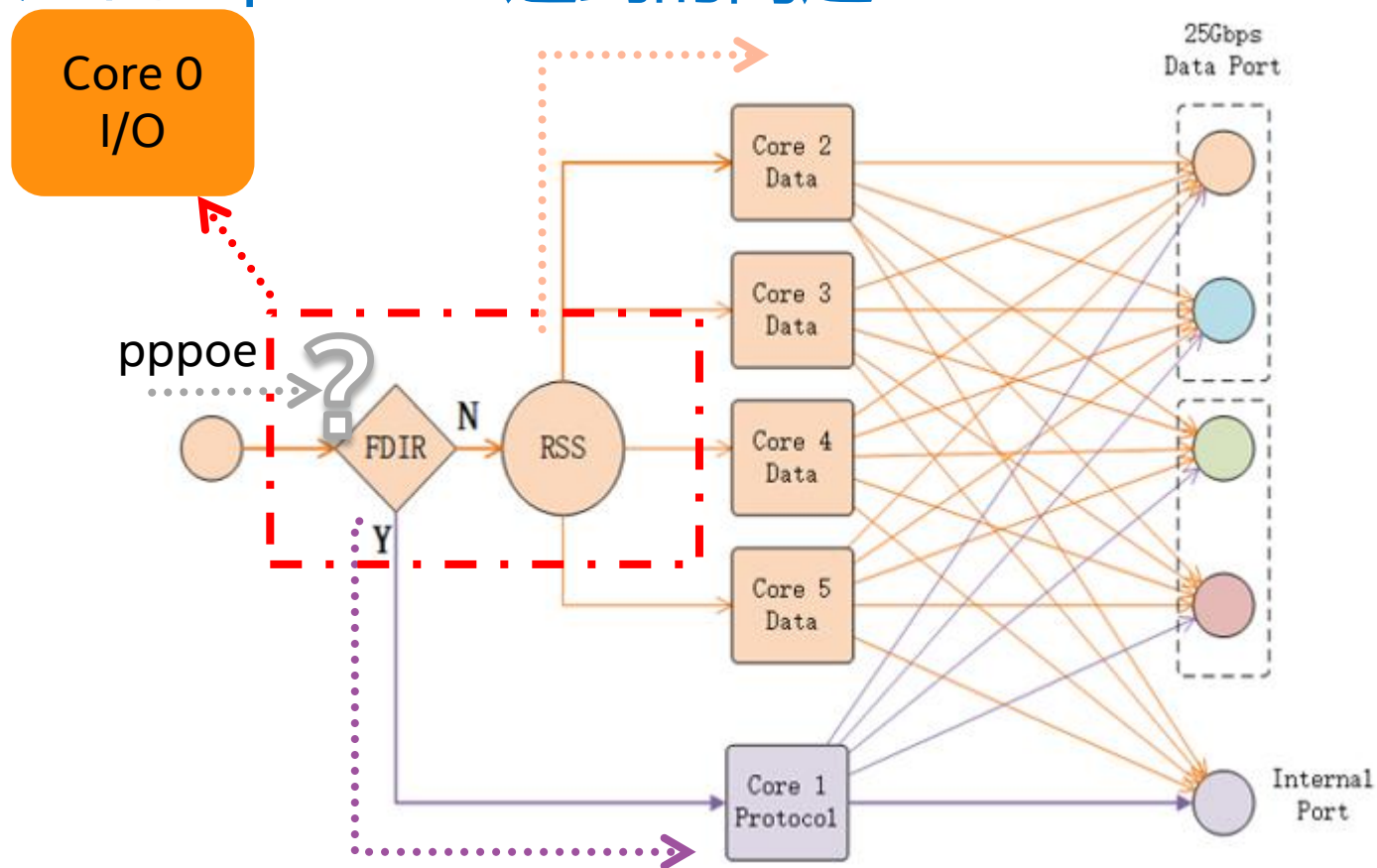- 节点数：600+。

intel

DDP的应用
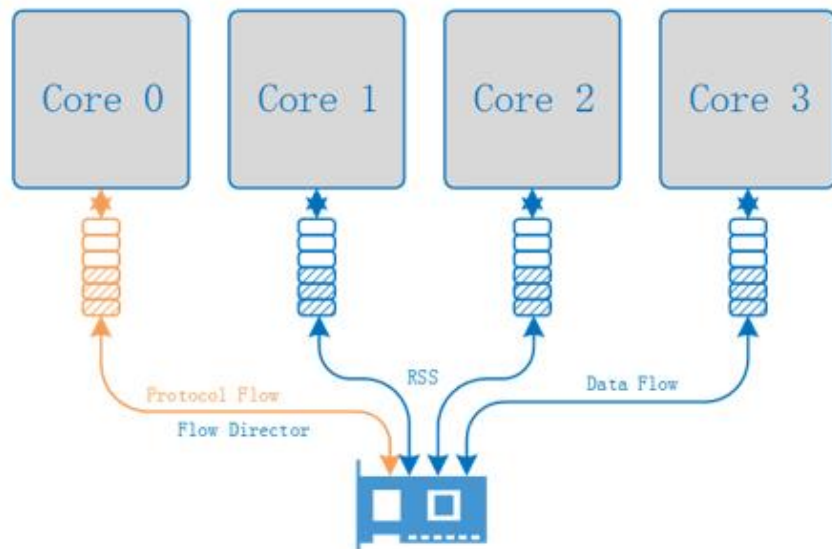
1.VBNG的Pipeline
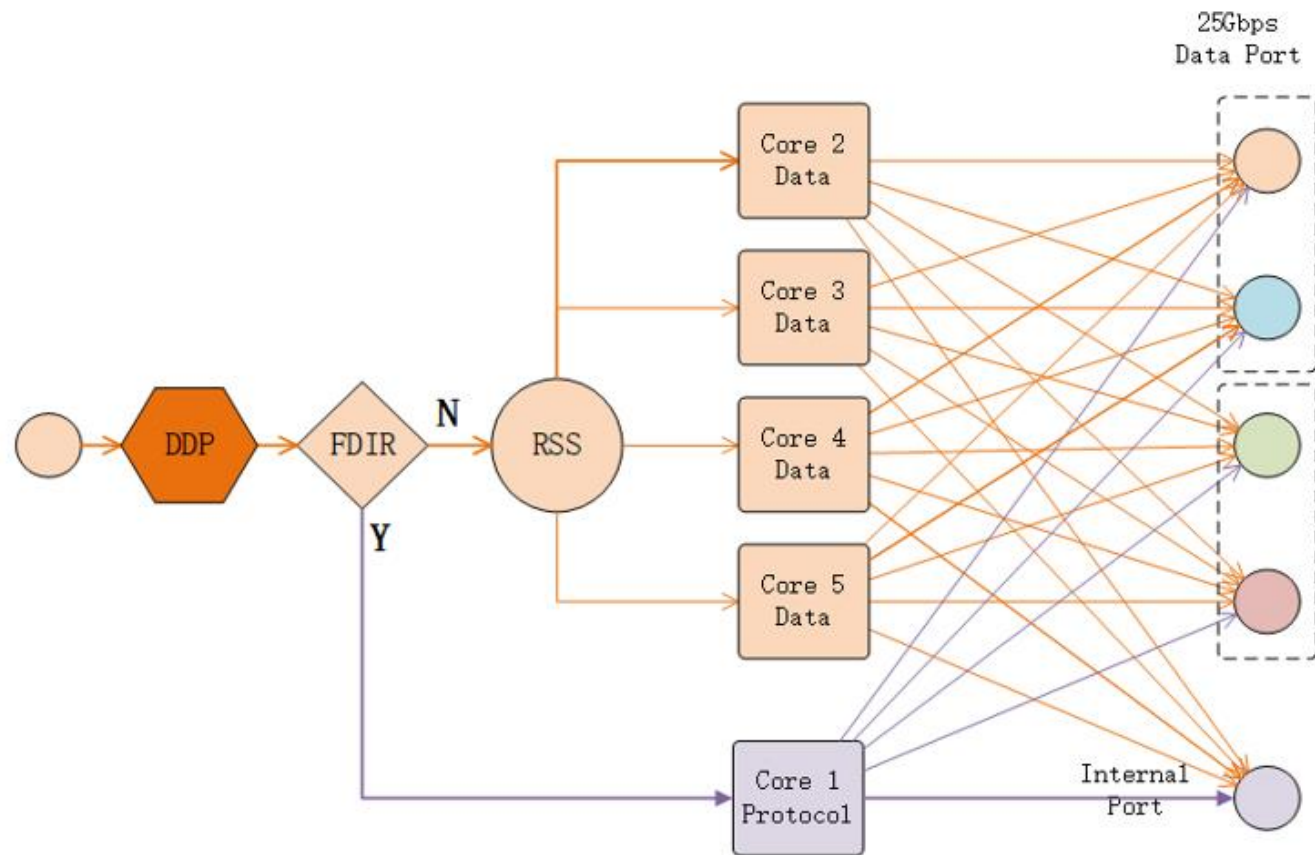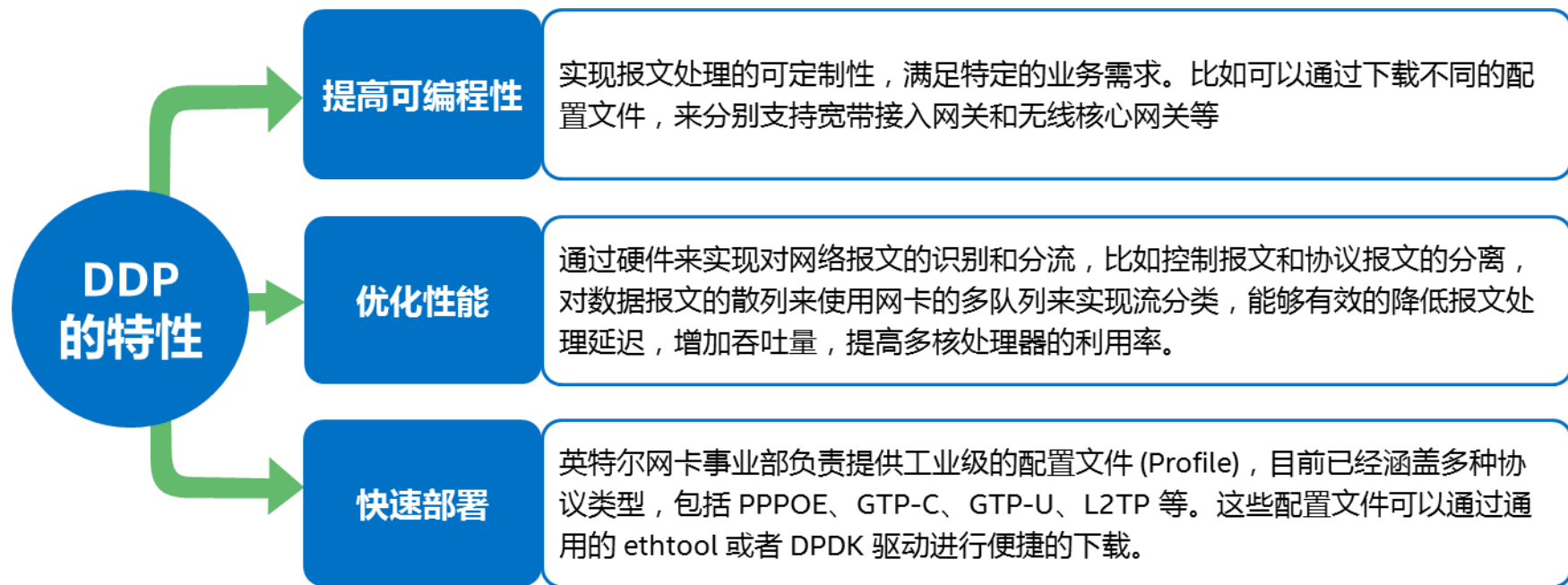
2.特殊报文带来的问题

3.DDP的使用

VBNG的Pipeline

单接口的Pipeline &遇到的问题

DDP 的使用

DDP(Dynamic Device Personalization)：
Intel设备定制化，可以对网卡添加新的协议，如PPPoE，使网卡能够识别并进行RSS和FDIR

使用DDP后的Pipeline

赛特斯 *CertusNet*

DDP
的特性

提高可编程性

实现报文处理的可定制性，满足特定的业务需求。比如可以通过下载不同的配置文件，来分别支持宽带接入网关和无线核心网关等

优化性能

通过硬件来实现对网络报文的识别和分流，比如控制报文和协议报文的分离，对数据报文的散列来使用网卡的多队列来实现流分类，能够有效的降低报文处理延迟，增加吞吐量，提高多核处理器的利用率。

快速部署

英特尔网卡事业部负责提供工业级的配置文件 (Profile)，目前已经涵盖多种协议类型，包括 PPPOE、GTP-C、GTP-U、L2TP 等。这些配置文件可以通过通用的 ethtool 或者 DPDK 驱动进行便捷的下载。

DDP带来的性能提升

1. 软硬件组网

2.业务流量模型

3.性能数据对比

硬件与软件配置

硬件配置	
Server	基于英特尔® 架构的通用服务器
CPU	英特尔® 至强® 铂金 8276L 处理器 @ 2.20GHz
Number of CPU	2
Number of Cores per CPU	28 (24 个 Core 用作数据面转发)
Memory	DDR4 @ 2400MHz, 112G (实际使用 64G)
NIC	英特尔® ETH NETWORK ADPTR XXV710 DA2
Intel Clear Regular	
软件配置	
FlexBNG	FlexBNG-v3.4
Guest OS	CentOS 7.5.1804
FlexCloud	FlexCloud-v5.1

业务流量模型

Item	Entries/Number	Comments
User Sessions	8K/UL Port	8 Ports used for UL 8 Ports used for DL
ACLs	4K/template	4 Template used
Routers	100K/socket	Global for all users
Qos(UL)	8K/UL port	One CAR entry per user
Qos(DL)	8K/UL port	One CAR entry per user