

Low Latency GRE Processing Accelerator Evaluation

FPGA acceleration specialized for individual applications and network operations

About mixi Group

Following the corporate philosophy of “User Surprise First,” mixi Group is always working to exceed the imagination and expectations of users. Since its founding in 1997, mixi Group has created communication services for friends and family to enjoy together, including the social network mixi and the multiplayer mobile application—Monster Strike. The group will continue to work toward enriching lifestyles of the future under the mission of “For Communication” by developing new businesses and services via IT to inspire communication around the world.



Introduction

The Infrastructure Development Group of mixi Inc. utilizes the Data Plane Development Kit (DPDK) and other tools to process packets for its on-premises network environment, which runs the Evolved Packet Core (EPC) software and smartphone applications.

This white paper shows a workload using FPGAs to accelerate the decapsulations of Generic Routing Encapsulation (GRE), which is a Layer 3 tunneling protocol used for multi-cloud environments. mixi Inc. has built a seamless multi-cloud environment by connecting cloud servers and on-premises servers using the point-to-point GRE protocol to abstract cloud-side network constraints. To configure the point-to-point GRE tunnel, the endpoint IP addresses on both sides have to be defined to construct a network, when one of the end point IP address change, the list of IP addresses on the DNS server will require synchronization.

In the case of mixi Inc's network environment, the only required function in this process is to remove the GRE header in the on-premise environment, the additional functions of managing the GRE headers in the cloud and on the on-premise servers are no longer necessary. The goal of the test was to achieve low latency despite removing the GRE headers in the test environment.

This white paper describes the efficient and practical use of FPGA acceleration in the network design and the operation of a specific content provider while comparing to software-based processing such as using DPDK, by taking the GRE header removal which solved the constraints of multi-cloud connections.

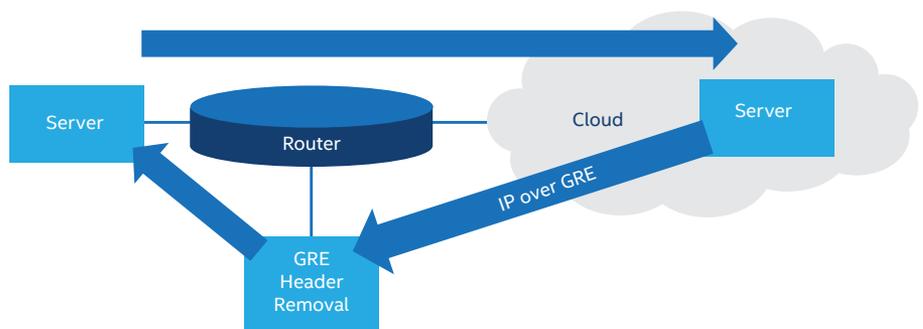


Figure 1. Example Communications Using GRE

Intel® Acceleration Stack for Intel Xeon® CPU with FPGAs

The Intel® Acceleration Stack for Intel Xeon® CPU with FPGAs [2] is a robust collection of software, firmware, and tools designed and distributed by Intel to make it easier to develop and deploy Intel FPGAs for workload optimization in the data center. The Intel Acceleration Stack for Intel Xeon CPU with FPGAs provides optimized and simplified hardware interfaces and software application programming interfaces (APIs), saving developers time so they can focus on the unique value-add of their solution.

The Intel FPGA Programmable Acceleration Card (Intel FPGA PAC) N3000 Prototype Board used in this white paper allows the implementation of user-specific packet processing functions in the User Programmable Logic (UPL) block between Ethernet MAC functions connected to QSFP 28/ QSFP+ transceivers and Ethernet MAC functions connected to the Intel Ethernet Controller XL710 (NIC), which enables users to focus on developing this core logic.

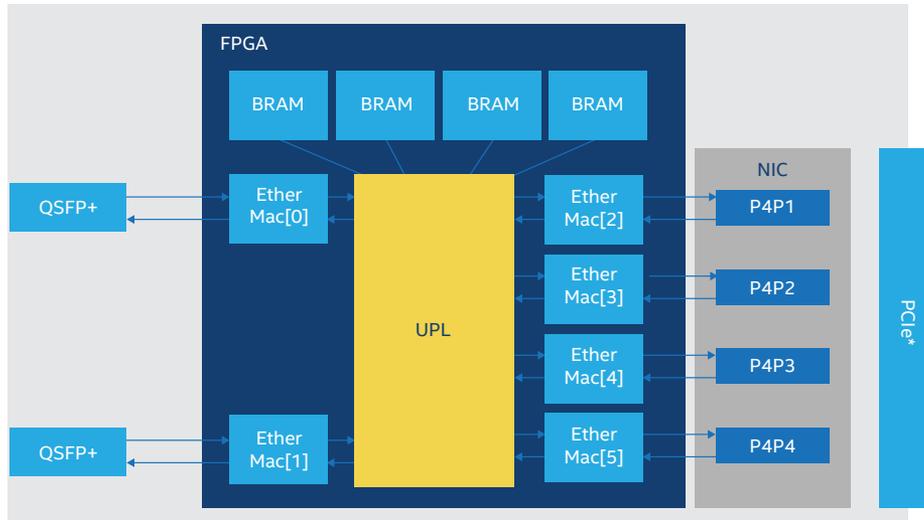


Figure 2. Intel FPGA Programmable Acceleration Card (Intel FPGA PAC) N3000 Prototype Board

Timing Chart - Decapsulation

The FPGA internal interface used in this evaluation can receive 256 bits (32 bytes) of packet data in one clock cycle. The first column shows the incoming packet data. As GRE packets tunnels the Layer 3 (L3), decapsulation processing could start once the second data was received.

First, the internal waveform was verified by adding a logic analyzer; after this verification, all data in a GRE packet of the minimum size (Ethernet frame (64B) + outer IPv4 header (20B) + GRE header (4B) = 88 bytes) was received in 5 to 7 clock cycles. In addition, using decapsulation to buffer the inbound packet data of two clock cycles allowed the implementation of pipeline processing in the delayed third clock cycle.

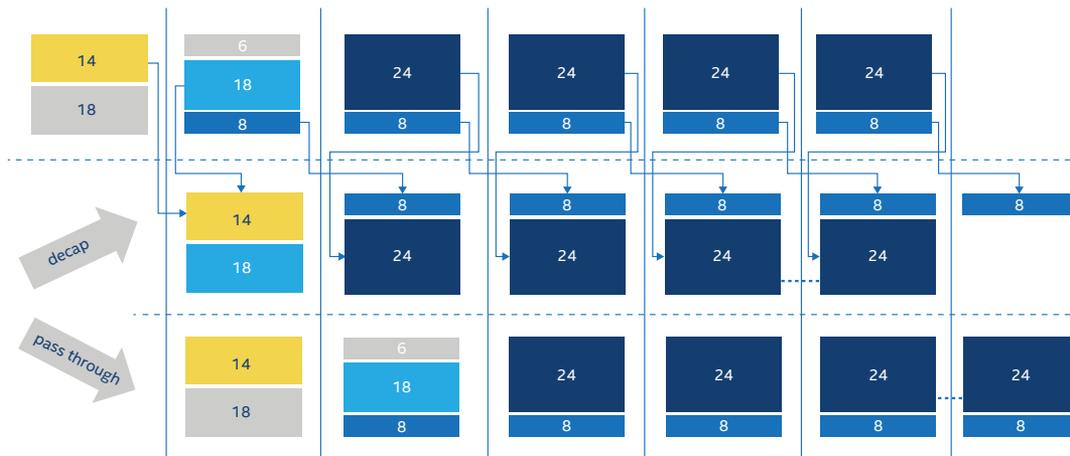


Figure 3. Timing Chart

State Machine

A state machine similar to the configuration shown in Figure 4 was implemented to perform the aforementioned decapsulation. From the initial state (INIT), it determines if the received packet data is a GRE packet and then depending on true/false of the valid signal, it transitions to the waiting state (WAIT) or payload processing state (PYLD). Finally, it transitions to the EOP state (EOP), which indicates the end of the packet before returning to the initial state (INIT).

```

XGRED_STATUS_INIT:begin
  if (in_eop==OFF&& in_sop==ON && in_val==ON) begin
    if (__in_dat[159:144] == 16'h0800 && __in_dat[71:64] == 8'd47) begin
      __internal_status_next = XGRED_STATUS_SOP;
    end else begin
      __internal_status_next = XGRED_STATUS_PASS;
    end
  end
  __in_dat_ss = __in_dat;
  __counter = 8'd0;
end else begin
  __internal_status_next = XGRED_STATUS_INIT;
  __counter = 8'd0;
end
end
    
```

Figure 4. Part of the State Machine Implementation

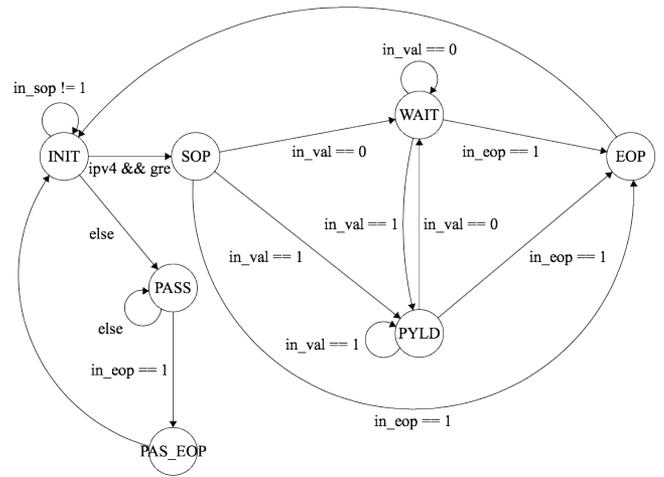


Figure 5. High-Level View of State Transition (partially omitted)

Platform Comparison

The latency and throughput performance of the GRE decapsulation process is compared in terms of FPGA acceleration, software processing using DPDK, and directly connected network equipment. Figure 6 shows the topology of the evaluation.

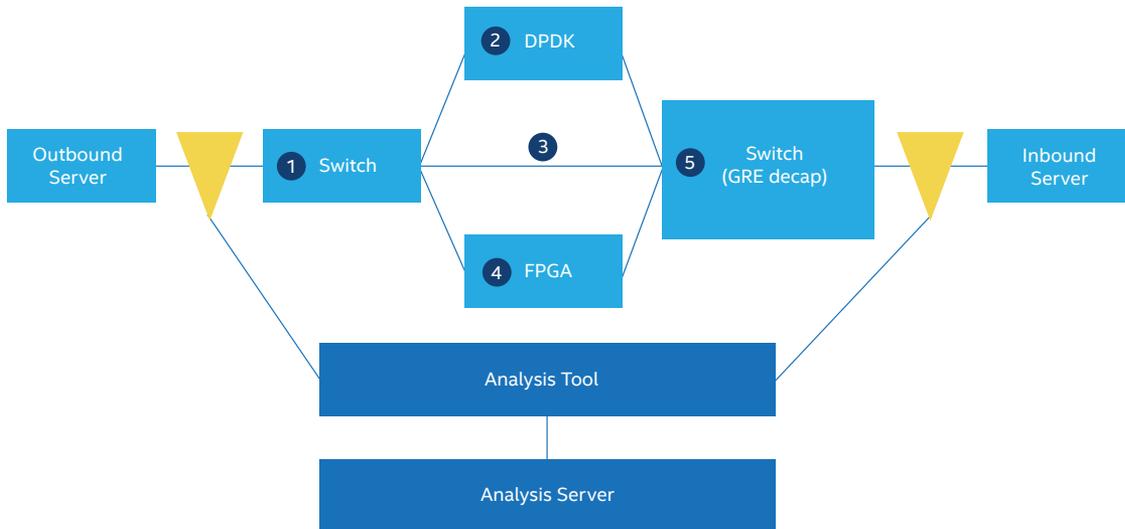


Figure 6. Platform Comparison

Platform Latency Comparison

Table 1 shows each latency value that is measured for one packet. DPDK is a platform for userland software processing, which delivers high-speed packet processing through CPU pole mode and batch processing. As such, the latency of one packet was, as expected, larger than the latency achieved with the FPGA hardware processing.

FPGA acceleration will result in less power consumption than DPDK, where the CPU is occupied with packet retrieval processing.

NO	PLATFORM	PATH	LATENCY
1	DPDK GRE	1 + 2 + 5	9.52 us
2	DPDK TCP		9.45 us
2	FPGA GRE	1 + 4 + 5	3.18 us
3	FPGA TCP		3.01 us
4	Switch GRE	1 + 3 + 5	2.29 us
5	Switch TCP		2.09 us

Table 1. Platform Latency Comparison

Measurement method: TCP segments used Ethernet frame length of 100 bytes, including frame check sequence (FCS). The GRE test used Ethernet frame length of 100 bytes, including FCS; following the removal of the IP/GRE header (24 bytes), the inbound server received a 76-byte Ethernet frame. One packet was sent five times and the average was calculated.

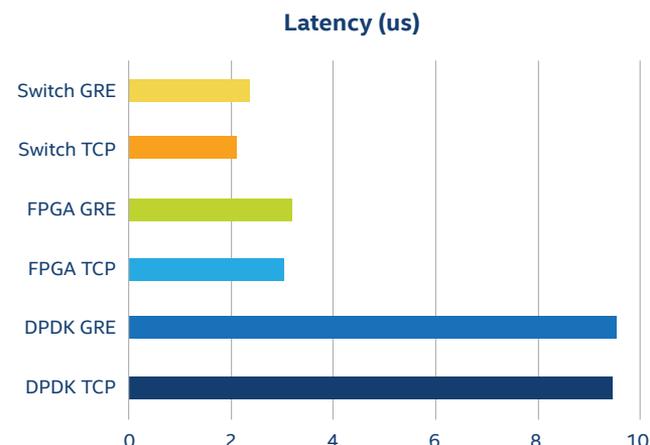


Figure 7. Latency Comparison

Platform Throughput Comparison

Table 2 lists the throughput values, which represents the sufficient level of performance that was achieved.

Measurement method: TCP segments used Ethernet frame length of 92 bytes, including FCS. The GRE test used Ethernet frame length of 92 bytes, including FCS; following the removal of the IP/GRE header (24 bytes), the inbound server received a 68-byte Ethernet frame. TCP segments and GRE-encapsulated packets were sent to the inbound server at a ratio of 1:1.

	LENGTH 92 BYTE (GRE TCP 1:1)			
	DPDK + QFX	QFX (DECAP)	FPGA	IDEAL VALUES
tx byte	55168165240	21304829160	12308110704	-
rx byte	47972317600	18525938400	10702704960	-
rx / tx byte (%)	0.8695652174	0.8695652174	0.8695652174	0.8695652174
tx pps	11.17 Mpps	11.17 Mpps	11.19 Mpps	-
rs pps	11.17 Mpps	11.17 Mpps	11.19 Mpps	-
loss packet	0	0	0	0

Table 2. Platform Throughput Comparison

Platform Comparison Consideration

As every solution has its own advantages and disadvantages, it is important to choose a solution that suits the situation or application in your company.

Table 3 lists the advantages and disadvantages for different applications.

APPLICATIONS	ADVANTAGES	DISADVANTAGES
Commercial network equipment	Operating stability and ease of assignment of human resources	Complicated configuration management
FPGA NIC hardware processing	Low latency, low power consumption, and efficient throughput	Difficult human resource assignment, development environments, and so on
DPDK-based software processing	Ease of assignment of human resources	Keeping development environments up to date requires high power consumption and latency

Table 3. Advantages and Disadvantages of FPGA Acceleration Solutions for Different Applications

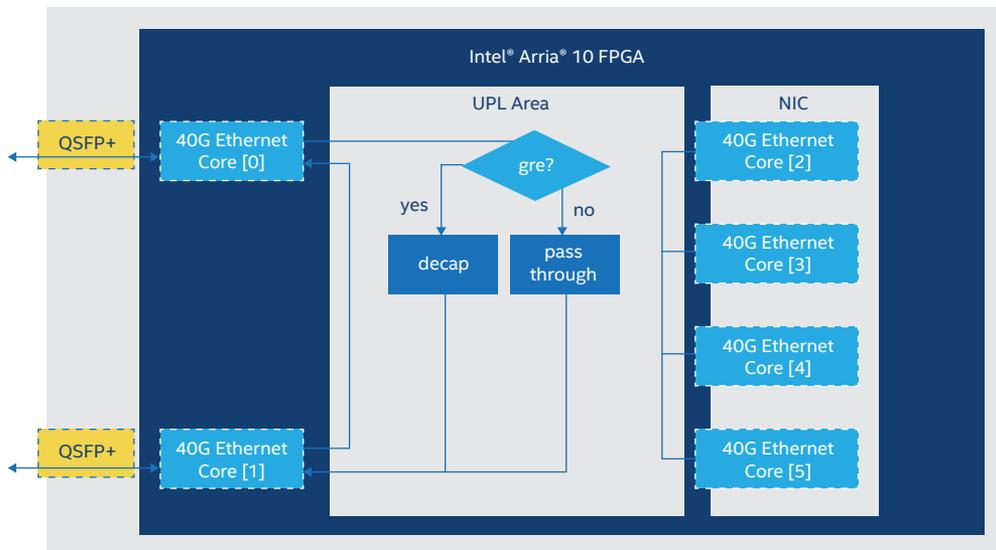


Figure 8. FPGA Module Logical Configuration

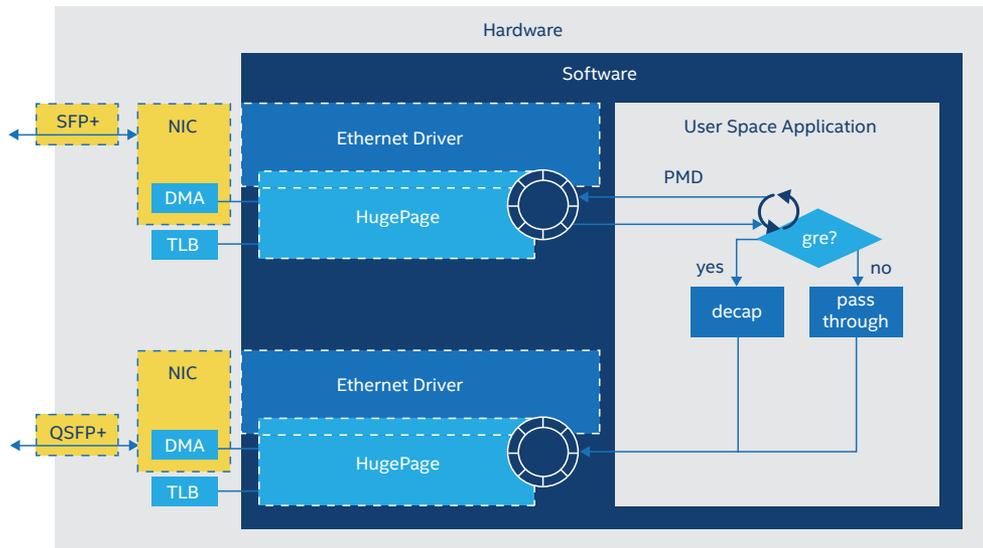


Figure 9. DPDK Module Logical Configuration

Summary

In response to the common refrain that network equipment scaling is not keeping pace with the recent cloudification and centralization of services, software-defined networking (SDN) and network functions virtualization (NFV) technologies that use server-based software processing to replace conventional equipment are often proposed as the solutions.

However, by comparing the commercial network equipment, FPGA NIC-based hardware processing, and DPDK-based software processing, this white paper shows that there are particular cases where SDN- or NFV-based software processing does not always provide the optimum solution.

FPGA-based hardware processing and DPDK-based software can accelerate some of the wide range of network functions that require dedicated network equipment. As this white paper demonstrates, depending on the network requirements of the company, it may be more efficient to replace some network functions with FPGAs specifically programmed for your situation. In such cases, hardware programmable FPGAs may be an option for more broadly defined SDNs and NFVs.

References

- DPDK website: www.dpdk.org/
- Intel FPGA Acceleration Hub site: www.intel.com/fpgaacceleration/



Intel does not control or audit third-party data. You should review this content, consult other sources, and confirm whether referenced data are accurate.

Tests measure performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to www.intel.com/benchmarks.

Performance results are based on testing as of August 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure.

System Configuration:

- DPDK: One Intel Xeon CPU E5-2670 v2 @ 2.50 GHz, Online Memory: 48 GB, SAS 146GB, Ubuntu 16.04
- FPGA: Intel FPGA PAC N3000 Prototype Board, Dell* PowerEdge R740 ((2U Server/GPU Install Kit Configuration), Intel Xeon Gold 6130 x2, 16 GB Memory x12, RAID Controller H730P, 300 GB SAS HDD x2 (No RAID setting), Management/iDRAC9 Enterprise (with OpenManage Essentials), NDC/ 1Gb QP, CentOS7.4)

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. Check with your system manufacturer or retailer or learn more at [\[intel.com\]](http://intel.com).

© Intel Corporation. Intel, the Intel logo, the Intel Inside mark and logo, Altera, Arria, Cyclone, Enpirion, Experience What's Inside, Intel Atom, Intel Core, Intel Xeon, MAX, Nios, Quartus and Stratix words and logos are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries. See Trademarks on intel.com for full list of Intel trademarks. *Other marks and brands may be claimed as the property of others.