40Gbps Ethernet Accelerator Functional Unit (AFU) Design Example User Guide

For Intel® Programmable Acceleration Card with Intel® Arria® 10 GX FPGA

Updated for Intel® Acceleration Stack for Intel® Xeon® CPU with FPGAs: 1.1 Production
1. About this Document

This document provides an overview of the 40Gbps Ethernet Accelerator Functional Unit (AFU) design example included in the Intel® Acceleration Stack for Intel Xeon® CPU with FPGAs and instructions to quickly evaluate the network port capability of the Intel Programmable Acceleration Card with Intel Arria® 10 GX FPGA.

1.1. Intended Audience

This document is intended for AFU developers and systems engineers to use as a quick start guide for evaluating AFU design and system integration of the network port feature on the Intel PAC with Intel Arria 10 GX FPGA.

1.2. Conventions

Table 1. Document Conventions

<table>
<thead>
<tr>
<th>Convention</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>If this symbol precedes a command, enter the command as a root.</td>
</tr>
<tr>
<td>$</td>
<td>If this symbol precedes a command, enter the command as a user.</td>
</tr>
<tr>
<td>This font</td>
<td>Indicates file names, commands, and keywords. The font also indicates long command lines. For long command lines, press Enter only if the next line starts a new command, where the # or $ character denotes the start of the next command.</td>
</tr>
<tr>
<td>&lt;variable_name&gt;</td>
<td>Indicates placeholder text that you must replace with appropriate values. Do not include the angle brackets.</td>
</tr>
</tbody>
</table>

1.3. Acronym List

Table 2. Acronyms List

<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Expansion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFU</td>
<td>Accelerator Functional Unit</td>
<td>Hardware Accelerator implemented in FPGA logic, which offloads a computational operation for an application from the CPU to improve performance.</td>
</tr>
<tr>
<td>AF</td>
<td>Accelerator Function</td>
<td>Compiled Hardware Accelerator image implemented in FPGA logic that accelerates an application. An AFU and associated AFs are also referred as GBS (Green-Bits, Green BitStream) in the Acceleration Stack installation directory tree and in source code comments.</td>
</tr>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
<td>A set of subroutine definitions, protocols, and tools for building software applications.</td>
</tr>
<tr>
<td>ASE</td>
<td>AFU Simulation Environment</td>
<td>Co-simulation environment that allows you to use the same host application and AF in a simulation environment. ASE is part of the Intel Acceleration Stack for FPGAs.</td>
</tr>
</tbody>
</table>

*Other names and brands may be claimed as the property of others.
<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Expansion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCI-P</td>
<td>Core Cache Interface</td>
<td>CCI-P is the standard interface that AFUs use to communicate with the host.</td>
</tr>
<tr>
<td>FIU</td>
<td>FPGA Interface Unit</td>
<td>FIU is a platform interface layer that acts as a bridge between platform interfaces like PCIe*, UPI, and AFU-side interfaces such as CCI-P.</td>
</tr>
<tr>
<td>FIM</td>
<td>FPGA Interface Manager</td>
<td>The FPGA hardware containing the FPGA Interface Unit (FIU) and external interfaces such as interfaces for memory, and networking. The FIM is also referred as BBS (Blue-Bits, Blue BitStream) in the Acceleration Stack installation directory tree and in source code comments. The AF interfaces with the FIM at run time.</td>
</tr>
<tr>
<td>NLB</td>
<td>Native Loopback</td>
<td>The NLB performs reads and writes to the CCI-P link to test connectivity and throughput.</td>
</tr>
<tr>
<td>OPAE</td>
<td>Open Programmable Engine</td>
<td>The OPAE is a software framework for managing and accessing AFs.</td>
</tr>
<tr>
<td>HSSI</td>
<td>High Speed Serial Interface</td>
<td>This is a reference to the multi-gigabit serial transceiver I/O in the FIM and the corresponding interface to the AFU.</td>
</tr>
<tr>
<td>PR</td>
<td>Partial Reconfiguration</td>
<td>The ability to dynamically reconfigure a portion of an FPGA while the remaining FPGA design continues to function.</td>
</tr>
</tbody>
</table>

### 1.4. Acceleration Glossary

#### Table 3. Acceleration Stack for Intel Xeon CPU with FPGAs Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intel Acceleration Stack for Intel Xeon CPU with FPGAs</td>
<td>Acceleration Stack</td>
<td>A collection of software, firmware and tools that provides performance-optimized connectivity between an Intel FPGA and an Intel Xeon processor.</td>
</tr>
<tr>
<td>Intel Programmable Acceleration Card with Intel Arria 10 GX FPGA</td>
<td>Intel PAC with Intel Arria 10 GX FPGA</td>
<td>PCIe accelerator card with an Intel Arria 10 FPGA. Programmable Acceleration Card is abbreviated PAC. Contains an FPGA Interface Manager (FIM) that pairs with an Intel Xeon processor over PCIe bus.</td>
</tr>
<tr>
<td>Intel Xeon Scalable Platform with Integrated FPGA</td>
<td>Integrated FPGA Platform</td>
<td>Intel Xeon plus FPGA platform with the Intel Xeon and an FPGA in a single package and sharing a coherent view of memory via the Ultra Path Interconnect (UPI).</td>
</tr>
<tr>
<td>OPAE_PLATFORM_ROOT</td>
<td></td>
<td>A Linux shell environment variable set up during the process of installing the OPAE SDK delivered with the Acceleration Stack.</td>
</tr>
</tbody>
</table>
2. Overview

The 40Gbps Ethernet (40GbE) AFU design example in the Acceleration Stack installation allows you to evaluate the network port capabilities of the Intel PAC with Intel Arria 10 GX FPGA. The 40GbE AFU design example contains a 40GbE MAC instance with traffic generation and checking logic to send and receive ethernet packets on the QSFP+ network port. The Acceleration Stack installation includes OPAE tools, APIs and a sample host application to initialize and start packet transfers from the host, and subsequently retrieve port statistics.

This design example supports internal HSSI transceiver loopback, external QSFP+ port loopback, and Intel PAC-to-PAC modes of operation.

Figure 1. System Block Diagram
2.1. 40GbE Design Example AFU Hardware

The design example uses the Intel FPGA 40G Ethernet MAC and PHY IP core to send and receive 40GbE ethernet packets on the Intel PAC’s QSFP+ network port. The design example supports generating and checking all network traffic data on the Intel PAC only through the implemented traffic generation and checking module in the AFU.

- The MAC IP instance connects to the HSSI PHY using the HSSI device class interface defined by OPAE. For more information about HSSI interface and 40G Ethernet MAC and PHY IP core connection, refer to the *HSSI User Guide for Intel Programmable Acceleration Card (PAC) with Intel Arria 10 GX FPGA*.

- The HSSI PHY implemented in the FIM connects to the FPGA’s transceiver I/O.

The HSSI Controller in the FIM utilizes transceiver reconfiguration to set the desired mode of the HSSI PHY. The design example requires that the host set the HSSI PHY mode to 40GBASE-SR4 (PMA only).

The design example utilizes the PR Management Interface ports on the hssi interface for internal purpose. Intel recommends you to terminate these ports in your AFU designs, refer to the *HSSI User Guide for Intel Programmable Acceleration Card (PAC) with Intel Arria 10 GX FPGA* for more details.

*Note:* The Intel Acceleration Stack version 1.1 supports PHY modes of 4x10GBASE-SR (PCS/PMA) and 40GBASE-SR4 (PMA only).

Use OPAE tools and APIs from the host to initialize and control packet transfers, and collect port statistics.

**Related Information**
- Low Latency 40-Gbps Ethernet IP Core User Guide
- *HSSI User Guide for Intel Programmable Acceleration Card (PAC) with Intel Arria 10 GX FPGA*

2.2. 40GbE Design Example AFU Software

Use the OPAE driver, tools, APIs, and sample host application to configure the HSSI PHY mode, initialize and control packet transfers, and collect port statistics with the design example AFU. For more information, refer to the following README file in the OPAE SDK installation:

```
$OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/sw/README.md
```

For more information about managing the network port feature from the host using the OPAE driver, refer to the *HSSI User Guide for Intel Programmable Acceleration Card (PAC) with Intel Arria 10 GX FPGA*.

**Related Information**
- Open Programmable Acceleration Engine (OPAE) Tools Guide
- *HSSI User Guide for Intel Programmable Acceleration Card (PAC) with Intel Arria 10 GX FPGA*
3. Running the Design Example Tests

In the `$OPAE_PLATFORM_ROOT/hw/samples` directory, there are two reference AFUs containing packet generation—`eth_e2e_e10` (10G Ethernet), and `eth_e2e_e40` (40G Ethernet). These AFUs contain packet generators and can be exercised by the sample OPAE host application located in the `sw` subdirectory.

3.1. Setup Prerequisites

To install the Intel PAC and OPAE SDK on a supported platform, follow the *Intel Acceleration Stack Quick Start Guide for Intel Programmable Acceleration Card with Intel Arria 10 GX FPGA*. If you only want to evaluate network port operation using the pre-compiled AFs from the OPAE SDK installation, you do not need the Intel Quartus® Prime Pro Edition software.

The `$OPAE_PLATFORM_ROOT` environment variable points to the location where you installed the OPAE SDK, which is delivered as part of the Acceleration Stack for Intel PAC with Intel Arria 10 GX FPGA.

**Related Information**

*Intel Acceleration Stack Quick Start Guide for Intel Programmable Acceleration Card with Intel Arria 10 GX FPGA*

For more information about installation of the software and licensing requirements.

3.2. Running 40GbE Internal Loopback Test in Single Intel PAC System

1. Load the AF for the 40GbE AFU example.

   $ cd $OPAE_PLATFORM_ROOT
   $ sudo fpgaconf hw/samples/eth_e2e_e40/bin/eth_e2e_e40.gbs

2. cd `$OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/sw`

3. Run the following steps on your Intel PAC:

   a. Compile the library and application using the command:

      $ make

   b. To configure the transceiver channel into 40G mode, write `40` to the following `sysfs` entry:

      $ sudo sh -c "echo 40 > /sys/class/fpga/intel-fpga-dev.<instance_id>/intel-fpga-fme.<instance_id>/intel-pac-hssi.<instance_id>/auto/hssi_mgmt/config"
<instance_id> represents the consecutive numbering of device, fme, and hssi instances.

For example:

```sh
sudo sh -c "echo 40 > /sys/class/fpga/intel-fpga-dev.0\n/intel-fpga-fme.0/intel-pac-hssi.2.auto/hssi_mgmt/config"
```

**Note:** Run this command twice to ensure effective reset.

c. To allow non-root users to access the 40GbE AFU instance, you can provide read and write privileges to the port (/dev/intel-fpga-port.*) where * denotes the respective socket. For example, to provide read and write privileges on Port 0:

```sh
sudo chmod 666 /dev/intel-fpga-port.0
```

d. To resolve library dependency:

```sh
export LD_LIBRARY_PATH=`pwd`:$LD_LIBRARY_PATH
```

e. To enable the internal loopback on B:D:F - 00:0a:0b,

```sh
./pac_hssi_e40 -b 00 -d 0a -f 0b --action=loopback_enable
```

f. To clear PHY, transmit, and receive statistics:

```sh
./pac_hssi_e40 -b 00 -d 0a -f 0b --action=stat_clear
```

Sample output:

```
Cleared TX stats on channel 0
Cleared RX stats on channel 0
```

g. To transmit 0x1000 packets:

```sh
./pac_hssi_e40 -b 00 -d 0a -f 0b --action=pkt_send
```

Sample output:

```
Sent 0x10000 packets on channel 0
```

h. To get PHY, transmit and receive statistics:

```sh
./pac_hssi_e40 -b 00 -d 0a -f 0b --action=stat
```

For more details, refer to the README file located in the sw subdirectory to:

```
$OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/sw/README.md
```

### 3.3. Running 40GbE External Loopback Test in Single Intel PAC System

The setup and output from the commands in the external loopback test are similar to the internal loopback test. The only difference is that the traffic loopback is established after the Intel PAC’s QSFP+ network port.

1. Loopback the generated network traffic at the Intel PAC’s external QSFP+ network port. You can accomplish this loopback in several ways:
• installing a QSFP+ optical module loopback adapter, or
• installing a QSFP+ optical module with MPO connection and looping back through:
  — an inserted fiber loopback plug, or
  — external network equipment

2. Load the AF for the 40GbE AFU example (if the AF is not already loaded).
   $ cd $OPAE_PLATFORM_ROOT
   $ sudo fpgaconf hw/samples/eth_e2e_e40/bin/eth_e2e_e40.gbs

3. cd $OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/sw

4. Run the following steps on your Intel PAC:
   a. Compile the library and application using the command:
      $ make
   b. To configure the transceiver channel into 40G mode, write 40 to the following sysfs entry:
      $ sudo sh -c "echo 40 > /sys/class/fpga/intel-fpga-dev.<instance_id>/intel-fpga-fme.<instance_id>/intel-pac-hssi.<instance_id>\auto/hssi_mgmt/config"

      <instance_id> represents the consecutive numbering of device, fme, and hssi instances.
      For example:
      sudo sh -c "echo 40 > /sys/class/fpga/intel-fpga-dev.0/intel-fpga-fme.0/intel-pac-hssi.2.auto/hssi_mgmt/config"

      Note: Run this command twice to ensure effective reset.
   c. To allow non-root users to access the 40GbE AFU instance, you can provide read and write privileges to the port (/dev/intel-fpga-port.*) where * denotes the respective socket. For example, to provide read and write privileges on Port 0:
      $ sudo chmod 666 /dev/intel-fpga-port.0
   d. To resolve library dependency:
      export LD_LIBRARY_PATH=`pwd`:$LD_LIBRARY_PATH
   e. To disable the internal loopback on B:D:F – 00:0a:0b,
      $ ./pac_hssi_e40 -b 00 -d 0a -f 0b --action=loopback_disable
   f. To enable the required PMA settings:
      i. Download the PMA settings from Intel PAC webpage.
      ii. Extract the following scripts: hssitune.py and hssicon from:
          tar xvzf a10_gx_pac_ias_1_1_pv_pma_settings.tar.gz
      iii. Specify the CONFIG based on your mode and cable:
          python hssitune.py --help
          usage: hssitune.py [-h] -c CONFIG -b BDF
          optional arguments:
3. Running the Design Example Tests

3.4. Running 40GbE Intel PAC-to-PAC Test between two connected Intel PACs

In this procedure, you can install the Intel PACs in the same system or two separate systems with the Acceleration Stack. Unless shown otherwise, you can expect the commands to return similar outputs as the internal loopback test.

```
-g -b 00 -d 0a -f 0b --action=stat_clear
```

```
Sent 0x1000 packets on channel 0
```

```
$ ./pac_hssi_e40 -b 00 -d 0a -f 0b --action=stat
```

```
Note: • Once the transceiver channel is configured to 40G, you can run the PMA setting script anytime with internal loopback disabled.
• After every hot plug/unplug of the cables, you must run the PMA settings script as discussed above to adapt the values.

For more details, refer to the README file located in the sw subdirectory to:

```
$OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/sw/README.md
```
Figure 3.  System Setup with Intel PACs Installed on Same Host

Figure 4.  System Setup with Intel PACs Installed on Separate Hosts
1. Install a QSFP+ optical module in each Intel PAC and connect the QSFP+ ports with an optical cable.

2. Assuming the two Intel PACs are installed in the same system, find their PCI Bus:Device:Function mappings.

   $ lspci | grep 09c4

   Sample output:
   04:00.0 Processing accelerators: Intel Corporation Device 09c4
   06:00.0 Processing accelerators: Intel Corporation Device 09c4

3. Load the AF for the 40GbE AFU example on both Intel PACs.

   $ cd $OPAE_PLATFORM_ROOT
   $ sudo fpgaconf hw/samples/eth_e2e_e40/bin/eth_e2e_e40.gbs -b 0x04
   $ sudo fpgaconf hw/samples/eth_e2e_e40/bin/eth_e2e_e40.gbs -b 0x06

4. Run the following steps on your both Intel PAC:

   a. Compile the library and application using the command:

      $ make

   b. To configure the transceiver channel into 40G mode, write 40 to the following sysfs entry:

      $ sudo sh -c "echo 40 > /sys/class/fpga/intel-fpga-dev.<instance_id>/intel-fpga-fme.<instance_id>/intel-pac-hssi.<instance_id>.auto/hssi_mgmt/config"

      `<instance_id>` represents the consecutive numbering of device, fme, and hssi instances.

      For example:
      sudo sh -c "echo 40 > /sys/class/fpga/intel-fpga-dev.0/intel-fpga-fme.0/intel-pac-hssi.2.auto/hssi_mgmt/config"

      Note: Run this command twice to ensure effective reset.

   c. To allow non-root users to access the 40GbE AFU instance, you can provide read and write privileges to the port (/dev/intel-fpga-port.*) where * denotes the respective socket. For example, to provide read and write privileges on Port 0:

      $ sudo chmod 666 /dev/intel-fpga-port.0

   d. To resolve library dependency:

      export LD_LIBRARY_PATH=`pwd`:$LD_LIBRARY_PATH

   e. To disable the internal loopback on B:D:F - 00:0a:0b,

      $ ./pac_hssi_e40 -b 00 -d 0a -f 0b --action=loopback_disable

   f. To enable the required PMA settings:

      i. Download the PMA settings from Intel PAC webpage.

      ii. Extract the following scripts: hssitune.py and hssicon from:

         tar xzvf a10_gx_pac_ias_1_1_pv_pma_settings.tar.gz
iii. Specify the CONFIG based on your mode and cable:

```
python hssitune.py --help
usage: hssitune.py [-h] -c CONFIG -b BDF

optional arguments:
  -h, --help            show this help message and exit
  -c CONFIG, --config CONFIG
                          Legal [rx_10g_c1m, rx_10g_c3m, rx_40g_c1m, rx_40g_c3m]
  -b BDF, --bdf BDF     Bus(B), Device(D), Function(F) in the
                      format B:D.F
```

**Note:**
- Verify that some of the DFE tap values are non-zero. This ensures that the script run is successful.
- You can provide the above config arguments for an optical cable of any length.

Figure 5. **PMA Setting Sample Output**

```
$ ./pac_hssi_e40 -b 00 -d 0a -f 0b --action=stat_clear
```

Sample output:
```
Cleared TX stats on channel 0
Cleared RX stats on channel 0
```

h. To transmit 0x1000 packets:

```
$ ./pac_hssi_e40 -b 00 -d 0a -f 0b --action=pkt_send
```

Sample output:
```
Sent 0x10000 packets on channel 0
```

**Note:** `pac_hssi_e40` transmits broadcast traffic in the default configuration. The destination MAC Address is set to 0xFFFFFFFFFFFF. To send Unicast traffic, you must update the following file:

```
SOPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/sw/pac_hssi_e40.c
```

Line 299 with upper 32 bits of the destination MAC

Line 307 with lower 32 bits of the destination MAC

Then, recompile the design example using the `make` command.

i. To get PHY, transmit and receive statistics:

```
$ ./pac_hssi_e40 -b 00 -d 0a -f 0b --action=stat
```

**Note:**
- Once the transceiver channel is configured to 40G, you can run the PMA setting script anytime with internal loopback disabled.
- After every hot plug/unplug of the cables, you must run the PMA settings script as discussed above to adapt the values.
For more details, refer to the README file located in the `sw` subdirectory to:

```
$OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/sw/README.md
```
4. Using the Design Example as a Platform for Further Evaluation

Use the 40GbE AFU design example to perform further evaluation with the Intel FPGA MACPHY IP, third party IP, or your own MACPHY IP. The example design source is located in the same location as the sample AFUs included in the OPAE SDK installation:

```bash
$OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40
```

The RTL source for the example is at the following location:

```bash
$OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/hw/rtl
```

While recompiling the example AFU to regenerate an AF (.gbs), you require an installed version of the Intel Quartus Prime Pro Edition (version 17.1.1) software.

OPAE version 1.0.2 does not support the ASE flow for HSSI interfaces.

**Related Information**

- Intel Acceleration Stack Quick Start Guide for Intel Programmable Acceleration Card with Intel Arria 10 GX FPGA
  For more information about installation of the software and licensing requirements.

- Accelerator Functional Unit (AFU) Developer's Guide
  For more information about the OPAE SDK design flow for AFUs that target the Intel PAC with Intel Arria 10 GX FPGA.

4.1. Prerequisite while Evaluating with the Intel FPGA MACPHY IP

In addition to the Intel licensing requirements for Intel Quartus Prime Pro Edition and Intel FPGA IP specified in the Intel Acceleration Stack Quick Start Guide for Intel Programmable Acceleration Card with Intel Arria 10 GX FPGA, the regeneration of AFs for the 40GbE design example with the Intel FPGA MACPHY IP also requires the following license:

**IP-40GEUMACPHY Low Latency 40G MACPHY**

**Related Information**

Intel Acceleration Stack Quick Start Guide for Intel Programmable Acceleration Card with Intel Arria 10 GX FPGA
For more information about installation of the software and licensing requirements.
4.2. Evaluation with an Alternate MAC/PHY IP

You can use the 40GbE AFU example as a framework to evaluate MAC/PHY IP from third parties or your own IP. The design example instantiates the Intel FPGA MAC/PHY IP in the following AFU RTL source file:

```
OPAE_PLATFORM_ROOT/hw/samples/eth_e2e_e40/hw/rtl/eth_e2e_e40.v
```

Replace the `alt_eth_ultra_0` instance shown in this figure with your own MAC/PHY IP instance. You must provide any necessary wrapper shim logic to integrate your IP within the AFU design example framework.

Figure 6. 40GbE AFU Design Example Hierarchy
5. Document Revision History for 40Gbps Ethernet AFU Design Example User Guide

<table>
<thead>
<tr>
<th>Document Version</th>
<th>Intel Acceleration Stack Version</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018.08.06</td>
<td>1.1 Production (supported with Intel Quartus Prime Pro Edition 17.1.1)</td>
<td>Initial release.</td>
</tr>
</tbody>
</table>

Intel Corporation. All rights reserved. Intel, the Intel logo, Altera, Arria, Cyclone, Enpirion, MAX, Nios, Quartus and Stratix words and logos are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries. Intel warrants performance of its FPGA and semiconductor products to current specifications in accordance with Intel’s standard warranty, but reserves the right to make changes to any products and services at any time without notice. Intel assumes no responsibility or liability arising out of the application or use of any information, product, or service described herein except as expressly agreed to in writing by Intel. Intel customers are advised to obtain the latest version of device specifications before relying on any published information and before placing orders for products or services.

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