White Paper

Intel® FPGAs
Open FPGA Stack

SAP Prototypes a Containerized Compression Workload Leveraging the Open FPGA Stack (OFS) Infrastructure

Computationally intensive workloads can benefit from FPGA-based acceleration. To expedite this development, SAP employs the OFS hardware and software infrastructure.

Interested in Developing a Containerized Application with OFS?

Read this white paper to learn how Intel partner, SAP, used an Intel® FPGA-based acceleration platform and Open FPGA Stack to expedite development of a proof-of-concept (PoC) deployed in a Docker container.

If you're interested in learning more about OFS, a proven, pre-developed FPGA based infrastructure, visit the open source OFS Repository on GitHub.

Authors

Dr. Christian Färber
TS - Field Application Engineer
Intel Programmable Solutions Group

Tamara Lin
Product Marketing Specialist
Intel Programmable Solutions Group

Vincent Riesop
Linux Developer
SAP SE

Executive Summary

Developers at SAP SE are creating a proof-of-concept (PoC) of cloud-based Compression as a Service (CaaS). They chose to leverage FPGAs to accelerate this computationally intensive workload, and plan to run their workload in Docker containers on their High Performance Analytics Appliance (HANA) cloud using their distribution of the Garden Linux operating system (OS). Using Open FPGA Stack, or OFS, they leverage a scalable, source-accessible hardware and software infrastructure designed to ease custom board and workload development. The flexibility provided by OFS enables them to develop a PoC in their preferred configuration, and the modular architecture expedites their development.

The OFS release for Intel® Agilex™ FPGAs includes a Docker image file that can be used to jump-start development with containers. You can access the Docker image files, including a user guide, in the open source OFS Repository.

Background

SAP SE is a German multinational software corporation based in Walldorf, Baden Württemberg, that develops enterprise software to manage business operations and customer relations.

SAP HANA is a relational database management system developed and marketed by SAP SE. Columnar data in SAP HANA is encoded with dictionaries. String dictionaries may contain vast amounts of textual data that needs to be compressed to minimize memory requirements. There are many different compression algorithms available (LZ77, LZR, LZSS, LZMA, ZStandard, etc.), but these algorithms are typically employed to compress large quantities of information into a block. If a dictionary is compressed this way, the entire file would have to be decompressed to access a single entry, which would be extremely inefficient in terms of time, computation, and power consumed. Alternatively, using these algorithms to compress each dictionary entry individually would present its own inefficiencies because they are not optimized to compress small amounts of data.

Re-Pair is a compression algorithm that is well suited for applications such as string dictionaries that require random accesses to compressed data. Unfortunately, Re-Pair is a computationally intensive and expensive algorithm that has not enjoyed widespread use in the data management community due to its prohibitively high compression and decompression times when implemented on central processing

units (CPUs). However, the programmable fabric in field-programmable gate arrays (FPGAs) can be configured to perform algorithmic processing in a massively parallel fashion. This means that algorithms like Re-Pair can be executed quickly on FPGAs while consuming relatively little power.

Garden Linux, the operating system leveraged by SAP, is a Debian GNU/Linux derivate crafted to provide small, auditable Linux images for use by cloud service providers (CSPs) and bare metal deployments. Garden Linux runs the current LTS Linux Kernel with additional driver support for cutting-edge enterprise server hardware. Providing a system to build images with secure and tested building blocks, it is highly versatile and customizable to various workload requirements.

The challenge was to prototype the Re-Pair compression workload deployment using Docker containers at SAP's HANA Cloud employing its Garden Linux OS with a high-performance PCI Express (PCIe)-based FPGA acceleration card. Using containers facilitates initial deployment, and moving applications from one server to another becomes easier as everything required to run those applications already in their containers.

SAP Deployment

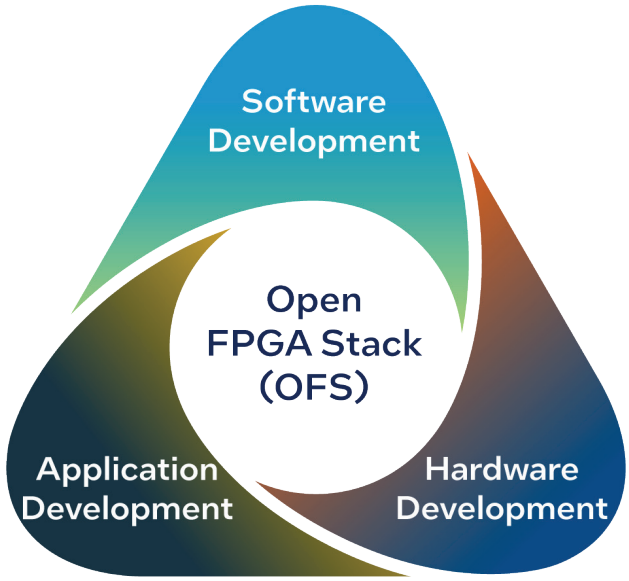
First, SAP took their stand-alone Intel Acceleration Stack 2.0 (DCP2.0) OpenCL™ compression workload using a shared virtual memory (SVM) board support package (BSP) and ported this to OFS. This was easy, as only the new universal shared memory (USM) BSP had to be used with the OFS stack to recompile the FPGA programming file. The host code did not have to be modified, only re-compiled with the OFS libraries.

Next, a Dockerfile was written using CentOS 7 as the base image and the following steps were taken in the Dockerfile:

1. Ensure that dependencies are installed or configured.
2. Define a set of environmental variables.
3. Clone the OFS git repository.
4. Build OPAE.
5. Install the Intel® FPGA SDK for OpenCL™ v20.4.
6. Set additional environmental variables.
7. Add the compression host code and FPGA programming file *.aocx.
8. Build the host executable.
9. Write the run scripts.

A server was then set up with an Intel® FPGA Programmable Acceleration Card (Intel® FPGA PAC) D5005, which required OFS firmware and Garden Linux, which already included the OFS Device Feature List (DFL) drivers. OFS DFL drivers have been included in the Garden Linux kernel since the 5.15 version, meaning SAP did not need to separately install the DFL drivers. The Docker container was then deployed on the server and started.

Why did SAP use Intel OFS?



OFS is supported by and validated on Intel® FPGA-based reference platforms. Customers and partners can leverage these platforms for application evaluation and development before porting to their custom or third-party OFS-based production board.

In this case, SAP developers wanted to deploy on an Intel FPGA PAC D5005, which is a high-performance PCI Express (PCIe)-based FPGA acceleration card for use in data centers. Because the OFS reference platform for Intel® Stratix® 10 FPGAs is the Intel FPGA PAC D5005, SAP was able to deploy their containerized workload with very little modification to the FPGA Interface Manager (FIM) hardware or software.

The OFS hardware consists of a FIM and an Accelerator Functional Unit (AFU) region. Using OFS, board (FIM) developers can leverage the provided infrastructure (base FIM) to quickly create a tailored, customized FIM for their board based on the interface and protocol requirements for a targeted workload. Meanwhile, AFU (workload) developers can leverage the provided FIM and AFU region with standardized Arm AMBA 4 AXI interfaces (Advanced eXtensible Interface) to focus on creating a custom workload/AFU without having to worry about the rest of the FPGA communication design.

Similarly, using the OFS provided software framework, developers can leverage up-streamed and open-sourced kernel drivers to accelerate integration into common application frameworks.

The SAP developers leveraged the provided infrastructure and documentation to quickly port their Re-Pair compression workload to their Intel® FPGA PAC D5005. Being able to quickly and easily generate this PoC will help SAP plan their deployment of FPGA-based solutions in production environments.

OFS Functions and Features Leveraged by SAP

OFS enables software, hardware, and application developers to use standard interfaces and application programming interfaces (APIs) to accelerate workload development and enable code reuse.

OFS also provides flexibility for different OS distributions and deployment models, including bare metal, virtualized, and containerized. In this case, SAP developers were able to deploy their workload using Docker containers. All of this was facilitated by the fact that OFS DFL FPGA drivers have been included in the Garden Linux kernel since version 5.15.

SAP developers leveraged FPGA-based re-programmability without requiring tremendous levels of FPGA expertise by leveraging the FIM provided by OFS. Modifications to the FIM can be made quickly and easily thanks to a modular, easily customizable architecture.

Additionally, OFS allows you to take full advantage of FPGA re-programmability by offering two configuration options: flat designs or designs that support partial reconfiguration (PR). PR provides a high level of flexibility by allowing portions of the FPGA to be reconfigured while the device is running—keeping the overall infrastructure intact and operating—thereby allowing changes to be made with no interruption to the system as a whole. OFS enabled SAP developers to leverage PR for their PoC inside their Docker containers.

By following the OFS deployment flow, SAP developers were also able to leverage the high-level design (HLD) shim. The HLD shim is a component integrated into the AFU Region that enables OpenCL and oneAPI kernels to communicate with runtime software by constructing correct packet protocol. The initial PoC of SAP's CaaS currently runs using OpenCL, with plans to adopt oneAPI in future iterations. Similarly, although the initial PoC was developed on the Intel Stratix 10 FPGA-based PAC D5005, SAP has plans to migrate to an Intel® Agilex™ FPGA-based card in the future.

Summary

OFS addresses the challenges associated with designing FPGA-based acceleration platform solutions for use in Intel® Xeon® processor-based servers. Employing OFS makes it easier for hardware, software, and application developers to create custom acceleration platforms and solutions.

OFS provides standard interfaces and APIs to enable greater code reuse, accelerate development, and speed deployment. For example, by means of OFS, developers can deploy containerized workloads, such as Docker containers with ease.

The result is that months of development time can be expedited by leveraging the infrastructure, source code, and documentation provided by OFS coupled with an FPGA-based reference platform.

Additional Resources (Intel)

- Access OFS source code, documentation, and the Docker container image for the OFS Agilex FPGA release by visiting the OFS Repository on GitHub.
- Learn more at: www.intel.com/OFS

Additional Resources (SAP)

- Learn more about the SAP HANA Cloud
- Learn more about the SAP Garden Linux Distribution on GitHub

WP-01320-1.1