SAMSUNG

CMM-B

CXL Memory Module - Box

White Paper

Table of Contents

Introduction 3
CMM-B Overview 3
A Use Case: SAP HANA In-Memory Database 4
Summary 5

SAMSUNG

Introduction

The CMM-B (CXL Memory Module - Box) is an advanced memory pooling solution designed for rack scale computing environments, utilizing Compute Express Link (CXL) technology to facilitate disaggregated, extensible, and composable memory architectures with enhanced software compatibility. The CMM-B enables flexible allocation of memory resources by supporting the connection of up to 22x E3.S CXL memory devices (CMM-D) to host systems. It is compatible with CXL 1.1 and CXL 2.0 protocols, incorporating CXL SoC switch chip from Xconn. The CMM-B, a 4U rack-mountable system, is capable of significantly scaling memory capacity and bandwidth. Its performance has been evaluated with industry leading SAP HANA IMDB application, demonstrating 32% improvement in TPC-DS performance through software interleaving with two Up Stream Ports (USPs), which connect seven Samsung CMM-D devices in CMM-B, compared to a single USP configuration that connects one CMM-D device in CMM-B in addition to capacity scalability.

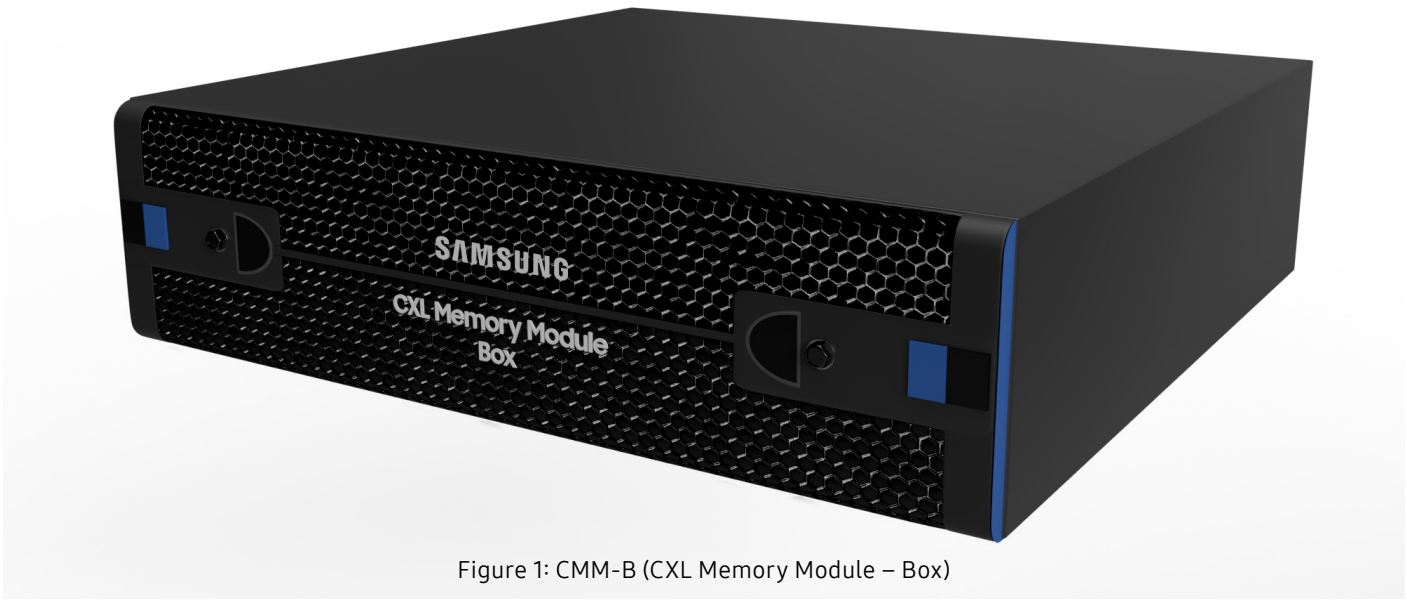

Figure 1: CMM-B (CXL Memory Module – Box)

CMM-B Overview

The CMM-B (CXL Memory Module - Box) facilitates both static and dynamic configurations through a Fabric Manager (FM): Samsung Cognos Management Console system manages operational states via Fabric Manager Application Programming Interface (FMAPI) aligned with the CXL specification. Currently, the CMM-B supports Single Logical Device (SLD) memory pooling as a part of CXL 2.0 specification. It treats the CXL switch as a singular virtual device that is recognized by the host through the Device-Specific Configuration Space (DVSEC) in the connected upstream port. This DVSEC details the size of devices bound in the downstream port.

The validation platform for the CMM-B utilizes the Intel Emerald Rapid servers. The validation involved CMM-D devices, each with a configuration of PCIe Gen5 x8 and a capacity of 256GB. Performance assessments were conducted using the Intel Memory Latency Checker (MLC) to evaluate the CXL switch.

Evaluations were performed on PCIe Gen5 x16 and Gen5 x8 upstream ports (USPs), with downstream ports (DSPs) set to PCIe Gen5 x8. The maximum throughput achieved was 35 GB/s with a 1:1 read-write ratio when 1-USP and 1-DSP/CMM-D were used, utilizing the PCIe Gen5 x8 interface on Intel MLC benchmark. Throughput variability was noted, influenced by the connection of the USP to the host socket and the execution location of the benchmark program. Specifically, throughput could achieve up to 28GB/s on 100% read workload if the USP and all threads were on Socket 0, but it could decrease to 15GB/s if threads were executed on a remote node, such as Socket 1. Latency also varied with a minimum of 390ns and an average of 520ns. Without CPU affinity settings, throughput would average between 25GB/s and 23GB/s for 100% read workload, as thread execution would be split between local and remote sockets.

For multiple USPs use case performance evaluation, software (SW) interleaving was facilitated by 'numactl'. This assessment utilized three USPs with 7x CMM-Ds and SW interleaving among triple USPs and seven virtual nodes configurations, on a single Emerald Rapid server. In the case of software interleaving with three USPs, the host system recognizes three HDM ranges, resulting in the operating system creating three NUMA nodes for the three USP endpoints. This configuration achieved a throughput of up to 60 GB/s and a latency of 596 ns for 100% read workload. For the software interleaving setup on seven virtual nodes, a single virtual NUMA node was allocated for each CMM-D device. This required modifications to the Linux kernel to create a NUMA node for each specified memory range associated with the CMM-D devices. The throughput reached 92 GB/s for 100% read workload with a maximum achievable throughput of 123 GB/s at a 1:1 read-write ratio, when 1 USP of PCIe Gen5 x16 and two USPs of x8 lanes were used. These performance metrics are measured by loaded_latency and idle_latency options.

Regarding capacity, the CMM-B supports Single Logical Device (SLD) memory pooling with up to 1:22 device binding capacity, allowing for the connection of 22 CMM-D devices to a single USP and enabling up to 5.6TB connection per USP. Despite the lack of hardware interleaving support, this configuration facilitates capacity scaling on a singular port. Evaluations were conducted across various scenarios involving up to three USPs and twenty-two DSPs on two Emerald Rapid servers and demonstrating the system's capability to enhance memory capacity effectively.

A Demonstration Use Case: SAP HANA In-Memory Database

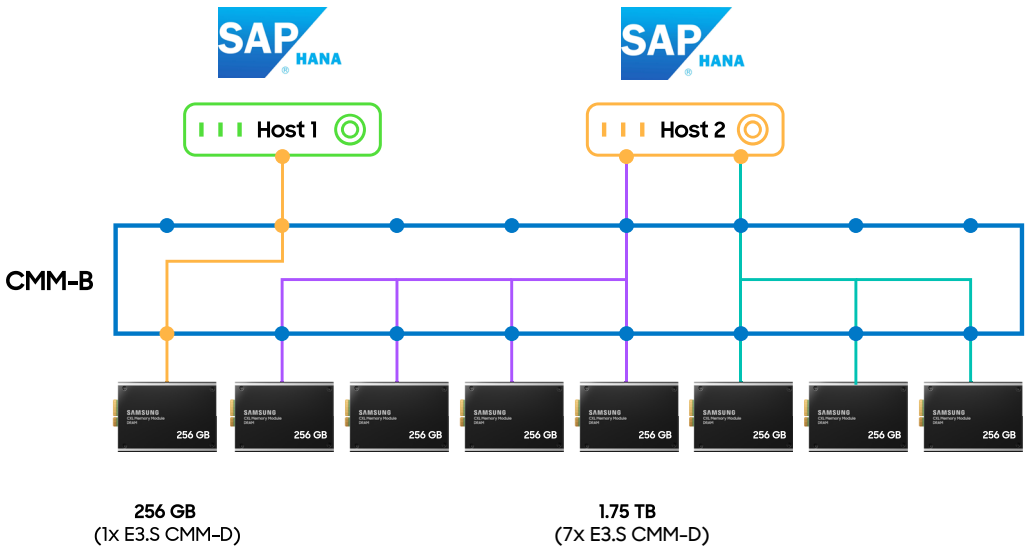

Figure 2: System Configuration

SAMSUNG

In our demo scenario, the integration and validation of the CMM-B were successfully completed with SAP HANA database. Despite SAP HANA's lack of native support for CXL memory, it can utilize DAX-enabled devices for its main storage. As we discussed in the previous section, the memory capacity scales as more CMM-D devices are attached to a host. To evaluate performance across multiple CXL devices, configurations like NUMA interleaving and Device Mapper striping were applied. SAP HANA throughput improvements were measured using the TPC-DS benchmark with 16 client threads. The findings revealed that employing NUMA interleaving with two USPs led to 32% performance boost over a single USP setup. A single USP connects to a single CMM-D device (256GB) and two-UPS configuration connects to seven CMM-D devices (1.75TB).

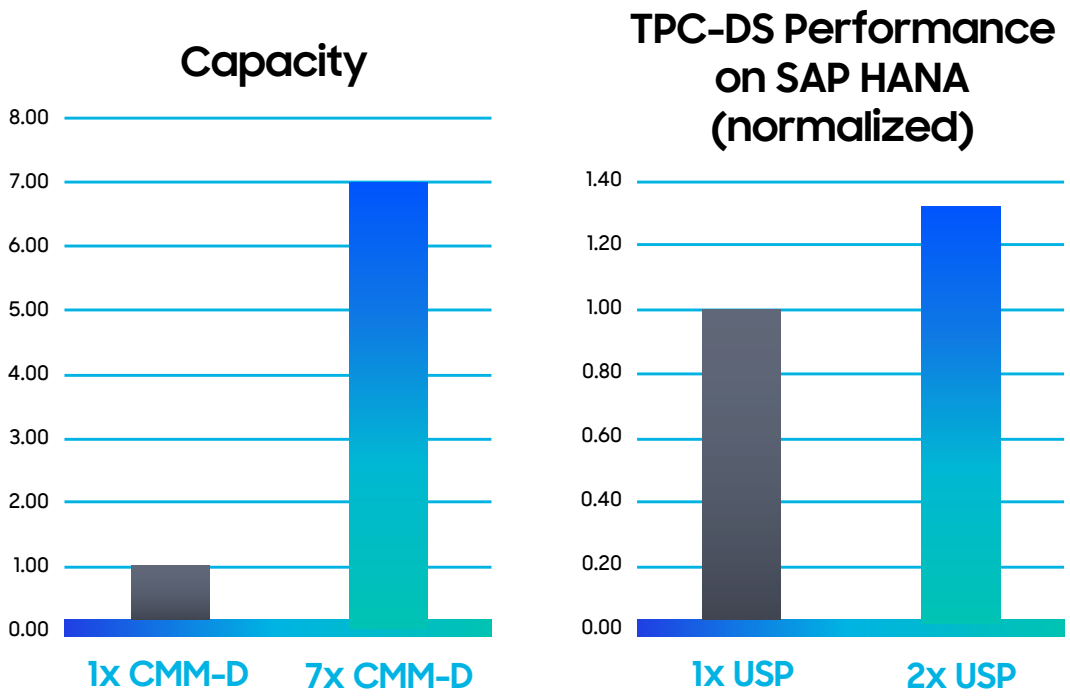

Figure 3: Memory Capacity Scalability and End-to-End Application Performance Enhancement

Summary

CMM-B provides support for CXL memory pooling functionality. It currently enables the integration of CMM-D devices with capacities up to 5.6TB with 22x 256GB E3.S CMM-D devices and supports up to three hosts with the ability to scale their capacity and bandwidth. We evaluated memory capacity and bandwidth scalability on multiple scenarios and also demonstrated that memory intensive application such as SAP HANA could achieve big performance benefit as more CXL memory bandwidth is attached.

For more information

For more information about the Samsung Semiconductor products, visit semiconductor.samsung.com.

About Samsung Electronics Co., Ltd.

Samsung Electronics Co. Ltd inspires the world and shapes the future with transformative ideas and technologies. The company is redefining the worlds of TVs, smartphones, wearable devices, tablets, digital appliances, network systems, memory, system LSI and LED solution. For the latest news, please visit the Samsung Newsroom at news.samsung.com.

Copyright © 2024 Samsung Electronics Co., Ltd. All rights reserved. Samsung is a registered trademark of Samsung Electronics Co., Ltd. Specifications and designs are subject to change without notice. Nonmetric weights and measurements are approximate. All data were deemed correct at time of creation. Samsung is not liable for errors or omissions. All brand, product, service names and logos are trademarks and/or registered trademarks of their respective owners and are hereby recognized and acknowledged.

Fio is a registered trademark of Fio Corporation. Intel is a trademark of Intel Corporation in the U.S. and/or other countries. Linux is a registered trademark of Linus Torvalds. PCI Express and PCIe are registered trademarks of PCI-SIG. Toggle is a registered trademark of Toggle, Inc.

Samsung Electronics Co., Ltd.
129 Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677, Korea **www.samsung.com** 1995-2021

SAMSUNG