

CAMEL's CXL Solution

The World-first CXL-based Direct Accessible, High-performance Memory Disaggregation Framework

As the big data era arrives, resource disaggregation has attracted significant attention thanks to its excellent scale-out capability, cost efficiency, and transparent elasticity. Disaggregating processors and storage devices well does break the physical boundaries of data centers and high-performance computing into separate physical entities. In contrast to the other resources, it is non-trivial to achieve a memory disaggregation technique that supports high performance and scalability with low cost. Many industry prototypes and academic simulation/emulation-based studies explore a wide spectrum of approaches to realize such memory disaggregation technology and put significant efforts into making memory disaggregation practical. However, the concept of memory disaggregation has not been successfully realized by far due to several fundamental challenges.

CAMEL provides a large memory system with the world-first CXL solution framework that can achieve outstanding performance in big data applications, such as machine learning, in-memory database, and real-world graph analytics. CAMEL's CXL solution opens up the new direction for memory disaggregation, and it ensuring a direct accessible and high-performance capabilities.

CAMEL's CXL Solution Enables High Performance Memory Disaggregation

The Challenges: High-Cost, Limited Scaling, Heavy Data Copies, and Host Dependency

The basic idea of memory disaggregation is to connect a host with one or memory nodes, such that it does not restrict a given task execution due to limited local memory (DRAM) space. Most existing technologies for memory disaggregation employ remote direct memory access (RDMA) in moving data from the remote memory to the host's local memory. However, all the techniques are limited to scale out and significantly increase the system building and maintenance costs. There are two root causes. First, DRAM and its memory interface (e.g., DDR) are designed towards entirely passive device modules, which cannot operate without the assistance of a host-side CPU and the memory controller therein. As more memory nodes are added to the systems, the number of other resources such as the computing process to hold the remote memory grows, exponentially increasing the cost. Second, RDMA introduces redundant memory copies and software fabric intervention, which in turn makes the latency of disaggregated memory multiple order of magnitude longer than that of local DRAM accesses.

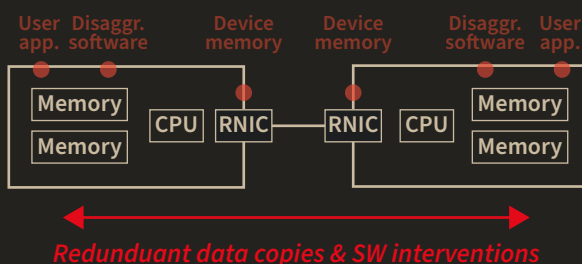
The Solution: Directly Access Remote Memory Resources with CXL

CAMEL has prototyped the world-first CXL solution (POC) that directly connects a host processor complex and remote memory resources over CXL protocol. CAMEL's CXL solution framework includes a set of computing express link (CXL) hardware and software IPs, including CXL switch, processor complex IP, and CXL memory controller. The solution framework can completely decouple memory resources from computing resources and enable high-performance, fully scale-out memory disaggregation architecture. The current prototype of CAMEL's CXL solution consists of:

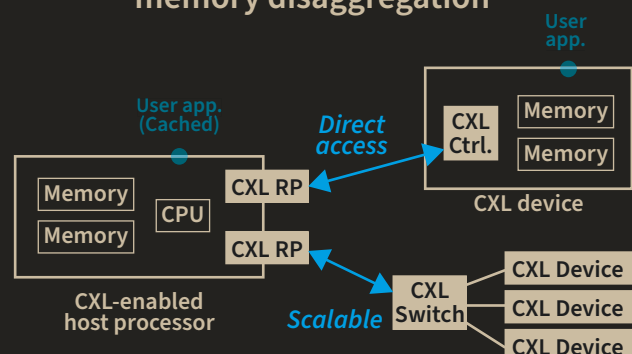


1. CXL device, which is a pure passive module that can have many DRAM DIMMs with its own hardware controller.
2. CXL-enabled host processor, which contains one or more CXL root ports (RPs).
3. CXL network switch, which allows connecting more than five hundred memory resources to simply expand (e.g., scale up) the memory space.

Conventional RDMA-based memory disaggregation





CAMEL's CXL-based memory disaggregation



CAMEL's CXL Solution Performance

As a case study, we compared the performance of our CXL solution framework to that of the conventional memory expansion method (e.g., RDMA-based memory disaggregation).

Experiment Setup

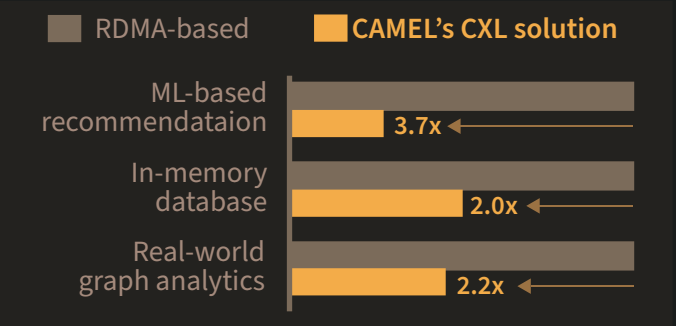
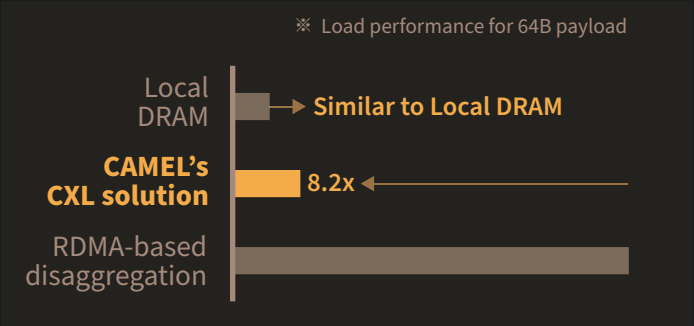
(Conventional) RDMA-based method	· RDMA NIC: Mellanox ConnectX-3 VPI InfiniBand RNIC (56Gbps) · RDMA driver: Mellanox OpenFabric Enterprise Distribution (OFED) v4.9		
Workload (Real)	ML-based recommendataion 	In-memory database 	Graph analytics 

RDMA vs. CXL

CAMEL's CXL solution shows 8.2x faster performance than RDMA-based solution (and achieves DRAM-like performance). The main cause of this gain is that CAMEL's CXL solution directly connects compute/memory nodes using PCIe, while RDMA requires protocol/interface changes between InfiniBand and PCIe. It also translates load/store requests from LLC into the CXL message, while RDMA uses DMA to read/write data from/to memory.

Real Workloads

CAMEL's CXL solution exhibits 2.3x better performance than RDMA-based memory disaggregation for three real workloads. This is because the RDMA-based method does not understand the characteristics of applications, and it swap-outs least-recently-used pages to the memory nodes. This makes frequent RDMA requests for page exchange and shows the worst performance.



Conclusion

CAMEL's CXL solution framework is protected by one or more patents. CAMEL's solution is a high-performance and scalable memory disaggregation method that allows multiple hosts to access remote memory resources through load/store instructions. It can exhibit DRAM-like performance when the workload can enjoy the host processor's cache. Finally, it can achieve a large memory capacity essential in handling emerging big data applications such as machine learning, in-memory database, and real-world graph analytics.

For more information and the latest news, please visit the CAMEL at camelab.org.