



Contribution ID: 33

Type: **not specified**

Device and Specific Purpose Memory MC

CFP ends on September 30th (CLOSED)

The Device and Specific Purpose Memory Microconference is proposed as a space to discuss topics that cross MM, Virtualization, and Memory device-driver boundaries. Beyond CXL this includes software methods for device-coherent memory via ZONE_DEVICE, physical memory pooling / sharing, and specific purpose memory application ABIs like device-dax, hugetlbfs, and guest_memfd. Some suggested topic areas include, but not limited to:

NUMA vs Specific Purpose Memory challenges

Core-MM services vs page allocator isolation

CXL use case challenges

Hotness Tracking and Migration Offloads

ZONE_DEVICE future for Accelerator Memory

ZONE_DEVICE future for CXL Memory Expansion

PMEM, NVDIMM, and DAX ``legacy'' challenges

Memory hotplug vs Device Memory

Memory RAS and repair gaps and challenges

Dynamic Capacity Device ABI (sparse memfd?)

Confidential Memory challenges

DMABUF beyond DRM use cases

virtiomem and virtiofs vs DAX and CXL challenges

Peer-to-peer DMA challenges

CXL Memory Pool Management

Device Memory testing

Why not the MM uConf for these topics? One of the observations from MM track at LSF/MM/BPF is that there is consistently an overflow of Device Memory topics that are of key interest to Memory device-driver developers, but lower priority to core MM developers.

Key Attendees:

Rajneesh Bhardwaj

Terry Bowman

Davidlohr Bueso

John Groves

Jason Gunthorpe

David Hildenbrand

John Hubbard

Alistair Popple

Gregory Price

Jonathan Cameron

Dave Jiang

David Rientjes

Ira Weiny

Balbir Singh

Konstantin Sinyuk

Viacheslav Dubeyko

Shivank Garg

Ben Cheatham
Yu Zhao
Changyuan Lyu
Liang Yan
Neeraj Kumar

Progress made on topics discussed at 2024 Plumbers:

Merged: CXL EDAC support for Memory Repair: <http://lore.kernel.org/20250521124749.817-1-shiju.jose@huawei.com>
Launched: CXL Management Library: <https://github.com/computexpresslink/libcxlmi>
Patches Available: FAMFS over FUSE: <http://lore.kernel.org/20250703185032.46568-1-john@groves.net>
Patches Available: Dynamic Capacity: <http://lore.kernel.org/20250413-dcd-type2-upstream-v9-0-1d4911a0b365@intel.com>
Patches Available: Type-2 CXL Accelerators: <http://lore.kernel.org/20250624141355.269056-1-alejandro.lucero-palau@amd.com>

“Device Memory” Background:

“Device Memory” is a catch-all term for the collection of platform technologies that add memory to a system outside of the typical “System RAM” default pool. Compute Express Link (CXL), a coherent interconnect that allows memory and caching-agent expansion over PCIe electricals, is one such technology. GPU/AI accelerators with hardware coherent memory, or software coherent memory (ZONE_DEVICE::DEVICE_PRIVATE), are another example technology.

In the Memory Management track of the 2025 LSF/MM/BPF Summit it became clear that CXL is one of a class of technologies putting pressure on traditional NUMA memory policy. While solutions like memory-interleave-sysfs and device-dax mitigate some of the issues there are still lingering concerns about memory of a certain performance class leaking into allocations that assume “default memory pool” performance.

The problem is how to keep Device / Specific Purpose memory contained to its specific consumers while also offering typical core-mm services. Solutions to that problem potentially intersect mechanisms like numactl, hugetlbfs, memfd, and guest_memfd. For example, guest_memfd is a kind of specific-purpose memory allocator.

Primary author: WILLIAMS, Dan (Intel)

Co-author: MANZANARES, Adam (Samsung Electronics)

Presenters: MANZANARES, Adam (Samsung Electronics); WILLIAMS, Dan (Intel)