

Virtual NVDIMM Status in XEN Design Discussion

Discussion I

Obstacles to expose NVDIMM to Dom0

George :Added a Xen specific label:

Todo: Init/active/map/unmap nvdimm pmem address space.

We will build the frame table when active/promote the pmem to DomU.

Trust that dom0 has unmapped before active.

George: Where to put the frame table?

Option a: A super block with single namespace?

Option b: Inside each namespace?

Discussion II

Xen Dom0 doesn't have huge page support.

Yu : Size of NVDIMM is tremendous, paging structures mapping to the NVDIMM and even to the struct page will occupy huge RAM size, access latency of NVDIMM is apparently higher than DRAM, these paging structures should be put in the DRAM which is much smaller than NVDIMM. Driver team solved this, by using huge page in DRAM to map NVDIMM and to map management information like struct page. But in Xen, this cannot be achieved without drastic changes in the PV MMU logic, which has no PSE exposed to dom0.

George: We could try to use PVH Dom0 instead.

Discussion III

fsdax vs devdax:

Yi: A file in the fs_dax should be easy management, but the dev_dax will have a better performance than that. also we have a fs rearrangement issue while using fs_dax file as backend.

Andrew: for fs rearrangement issue, we can try the llvm on the dax.