Clovis Object

a single top-level “Object” in Mero

Clovis block (IO) size is LCM of unit sizes in layout

FID — 8mb

FIDs are 120b globally unique

Composite Layout

layered set of individual layouts

any layer may itself be complex, thus layouts are hierarchical

FID_0 — extent1, extent2

FID_1 — extent1, extent2

…

Writes go to top layer in the composite layout. Reads fall though. Or layers can be read/written directly using FID_{sub}

PDClust Layout

{layout type, layout function f(), parameters, mapped read & write extents} stored as object MD

8+2+2 parity, unit size 4K-32M

FID_{GOB} — 1mb

parity

spares

Unit

$f(FID_{GOB}, params)$

layout function maps a global object onto a set of component objects

offset-in-file → cob-id, offset-in-cob

FID_{COB} — 1mb

Component Objects on each *device* in *pool*

Each GOB is (eventually) spread over the entire pool.

{size, checksum, attributes, AD} stored in COB MD

Allocation Data

Storage Object (STOB) — 4k

AD maps offset-in-COB to actual data location on device, e.g. LBAs on block device.
Blocks are not allocated until written; COBs are sparse.
There are multiple COBs on each device, 1 per GOB.