THE CAPACITY ERA

Storage Days Tallinn

31 August – 1 September 2026

TALLINN, ESTONIA

Why head depopulation matters now

A failing head is not a failed drive — recovering capacity across the recertified fleet today and the HAMR generation next

David Gerstein CTO · Leil

leil

Agenda

01 Why a failing head is not a failed drive — the 5% problem

02 Two sides of the coin — today’s recertified fleet and tomorrow’s HAMR

03 Doing it in software — LeilFS, LeilOS, and the SMR roadmap

The cost of treating a drive as binary

~5%

of a drive’s capacity is all that a single failing head actually costs.

~40%

of real-world drive rejections are caused by just one failing head.

100%

of the drive is rebuilt across the cluster when HDD is marked dead.

One mechanism, two payoffs

TODAY'S FLEET

Used & recertified drives

Short warranty, higher single-head wear. Detect the degrading head early, evacuate its data cleanly, depop it, and return the drive at ~95%. Reuse capacity you already paid for — 2–3× more usable life.

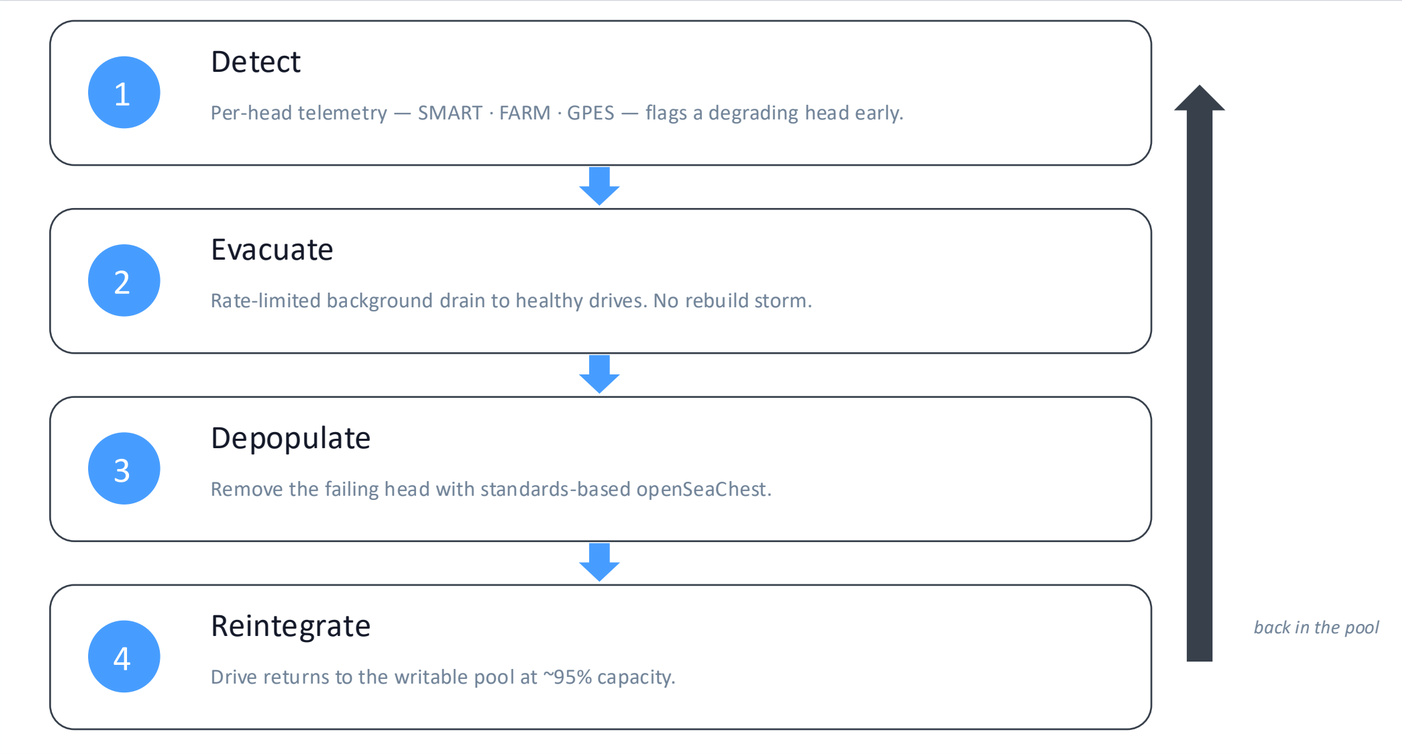
TOMORROW'S DRIVES

HAMR and beyond

HAMR write elements wear faster, and drives keep gaining platters and heads on the way to 100 TB by 2029. As head count climbs, one failing head strands ever more capacity — so data-preserving depop matters most on the newest drives.

The predictive depop loop

LeilOS · DETECT → EVACUATE → DEPOPULATE → REINTEGRATE



LeilOS acts on early degradation, while the head still reads. Data drains cleanly ahead of failure, the head is depopped, and the drive returns to the pool at reduced capacity.

Emergency in LeilFS, predictive in LeilOS

CMR head depop ships in two modes. The open core handles the emergency; LeilOS adds the proactive loop.

- LeilFS (open source): when a CMR head fails, recover the drive in place at ~95% — no whole-drive rebuild, no RMA.
- LeilOS (enterprise): detect the degrading head early, evacuate ahead of failure — 0% reconstruction, tier-aware policy.
- SMR data-preserving depop is on the roadmap — gated on the kernel support Christoph showed. We adopt it when it lands.

Where Leil fits: the last mile

WHAT'S ARRIVING

Commands and kernel, converging

Uwe covered WD's depop commands. Christoph showed data-preserving depop in XFS: after a depop, REPORT ZONES marks the head's zones Offline — no reformat. Kernel support is landing (7.3–7.4), and openSeaChest already drives modify-zones.

WHAT LEIL ADDS

The predictive loop, at fleet scale

LeilOS runs the proactive loop on CMR — detect, evacuate, depop, reintegrate — with warranty- and tier-aware policy; open-source LeilFS covers the emergency path. SMR data-preserving depop follows once the kernel support lands. One ask to WD: advertise the standard modify-zones feature on the Ultrastar SMR SKUs.

Takeaways

01 A failing head is a ~5% capacity event, not a dead drive — treating it as binary throws away the other 95%.

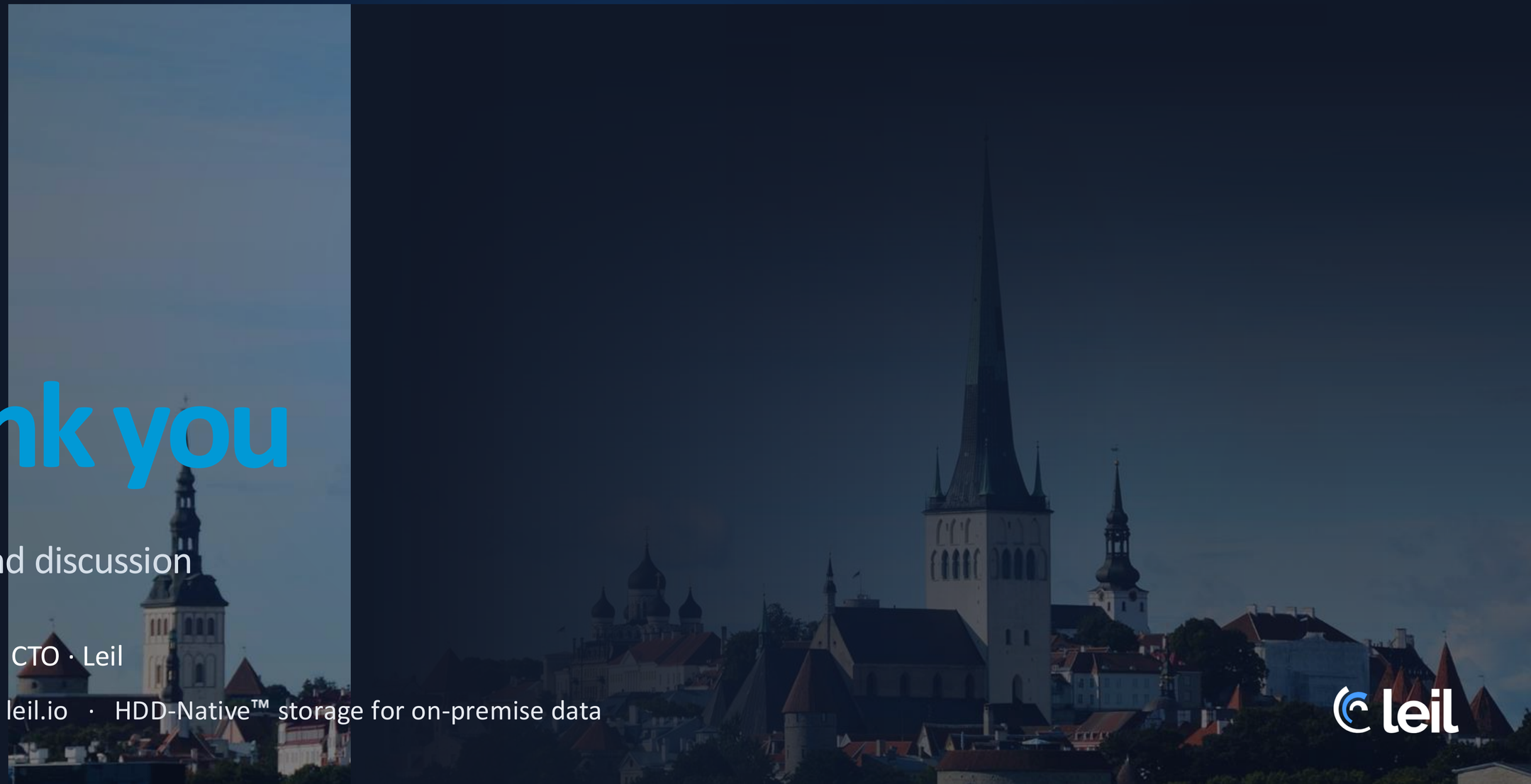
02 The same predict → evacuate → depop → reintegrate loop pays off twice: it rescues today's recertified fleet and protects tomorrow's 100 TB HAMR capacity.

03 CMR depop ships now — emergency in open-source LeilFS, predictive in LeilOS. SMR data-preserving depop follows the kernel work Christoph showed. That is the last mile Leil is building.

Q&A

Implementation, SMR specifics, and the asks for WD.

Storage Days Tallinn



Thank you

Questions and discussion

David Gerstein CTO · Leil

david@leil.io · leil.io · HDD-Native™ storage for on-premise data

leil

Appendix — the openSeaChest depop flow

The manual baseline that LeilOS automates for CMR, using Seagate's standards-based tooling on WD drives. Detect, evacuate, remove the element, reformat, reintegrate.

- Detect — --showPhysicalElementStatus reports per-head Health (1 = in limit · 254 = in progress · 255 = depopulated).
- Depop — --removePhysicalElement <N> (truncate) or, on a zoned drive, --removeElementModifyZones (data-preserving); evacuate first, then poll status.
- Recover — reformat and remount at reduced capacity; --repopulateElements (REAR) restores full size.