Stop buying the same

How to reuse multi-vendor enterprise storage, decide when a new Ceph cluster is actually justified, and spend the remaining 2026 budget without buying tomorrow's technical debt.

+70%
NAND CONTRACT PRICE · Q2 2026 QOQ · FORECAST [4]

750TB
FASTEC 6+2 USABLE PER 1 PB RAW · THEORETICAL (VS 333 TB REPLICATED)

20/20
QUESTIONS YOUR ARCHITECTURE SHOULD ANSWER

AUTHOR
SIXE

EDITION
v1.0 · 2026.07

FOR
Architects · CIOs · Storage teams

Contents.

00 EXECUTIVE SUMMARY

Six points, one committee.

For the reader whose steering committee starts at 11:00.

01 · DEFAULT

No silo by default.

Don't buy array N+1 out of fear of Ceph. Ask whether the existing estate can serve part of the requirement through Cinder first.

02 · SHARED

Shared Ceph as candidate for new pooled capacity.

A well-designed external Ceph service can serve several OpenStack environments separated by pools, CephX identities, quotas, CRUSH rules.

03 · DEDICATED

Dedicated clusters only for explicit boundaries.

Regulation, sovereignty, geography, ownership, incompatible SLA, blast radius, extreme scale. "New OpenStack cloud" is not a boundary.

04 · BASELINE

Ceph Tentacle 20.2.2 = production baseline.

FastEC, improved RBD migration, NVMe-oF, integrated SMB and management gateway materially improve the consolidation case. [1]

05 · ANCHOR

OpenStack Gazpacho = H2 anchor.

Current maintained SLURP. Hibiscus estimated Sept 30, 2026: laboratory now, critical path later. [3]

06 · BUDGET

Spend year-end on optionality, not media.

Architecture, migration, network, backup, spares, automation, observability, training — more long-term value than another rushed pallet.

> The goal is not to put Ceph everywhere. The goal is to **stop buying the same usable terabyte twice.**

ONE WARNING

Architectures are solutions to specific problems. They are not religions. Anyone selling you "always Ceph," "always the array," or "always hyperconverged" before asking a single question is selling you their answer, not solving your problem. This paper gives you the questions. Judge them, not the vendor.

DISCLOSURE

SIXE sells Ceph consulting on community Ceph **and** IBM Storage Ceph. We also sell and support IBM enterprise storage (FlashSystem, Storage Virtualize) — we make money when you keep your arrays, which is why the five tests in section 04 exist. Distro choice for Ceph is not the point of this paper. Architecture is.

PART I / VI

Where

The first-half price shock made storage-media economics change faster than an annual budget. Silo-by-default now charges four taxes — every year.

The shortage didn't end. The panic changed shape.

The silo-by-default problem & its four taxes.

01 MARKET CONTEXT

The panic changed shape.

H2 2026 is not "after the storage shortage." It is **after the first price shock**, while scarcity, allocation pressure and long lead times continue to shape architecture.

55–60%

Q1 2026 NAND QOQ · FORECAST [2]

TrendForce projection; enterprise SSD projected to rise 53–58%.

70–75%

Q2 2026 NAND QOQ · FORECAST [4]

Forecast raised from Q1; enterprise-grade allocation pressure continues.

H2 26

DIRECTION, NOT RESET

Q3 moderation expected but not a price reversal; nearline HDD lead times give no comfortable escape.

Figures above are TrendForce forecasts (single-source). Realized prices may deviate; IDC / DRAMeXchange / Objective Analysis publish their own methodologies. Direction is more robust than any single percentage.

> Which capabilities do we already **own**, which should be **shared**, and where does a real boundary justify another platform?
>
> — THE QUESTION THAT REPLACES "WHAT SHOULD WE BUY?"

TAKE INTO NEXT PROCUREMENT MEETING

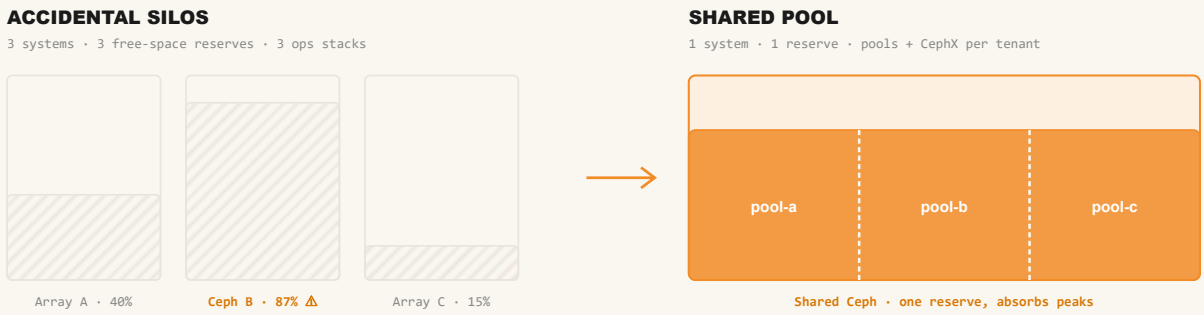
- How many independently managed free-space reserves are we funding — could we say the number without checking?
- If our fastest-growing environment needed 200 TB in 90 days, which pool would supply it, at what confirmed lead time?
- Which pending quotations have committed delivery dates versus "subject to allocation"?

02 THE SILOS

Silo-by-default charges four taxes.

A silo isn't a vendor logo. It's a boundary that capacity, policy, lifecycle and data mobility cannot cross economically. Useful when intentional. Expensive when accidental.

FIG 01 · THE SILO PROBLEM, VISUALISED



TAX #1

Capacity

Every platform funds its own free space, failure reserves, expansion headroom. Pooling doesn't eliminate reserves — it reduces the number separately stranded.

TAX #2

Operations

Each cluster brings its own mons, mgrs, OSDs, certs, alerts, upgrade rehearsals, DR docs. Small platforms carry most of the surface with fewer options.

TAX #3

Consistency

Separate systems drift into different volume-types, snapshot policies, retention, encryption. Five "Gold" tiers, none of which mean the same thing.

TAX #4

Migration

Data placed in isolation eventually needs to leave. Cinder gives you drivers, not a plan.[5] Cross-backend retype can require detach and generic copy depending on the driver pair; snapshots and QoS may not travel. Design the exit before the first volume.

ARCHITECTURE REVIEW QUESTIONS

- For each platform: who asked for the boundary, and what problem was it solving? Written down anywhere?
- What would it cost — in people-days, not euros — to move our largest tenant from platform A to platform B this quarter?
- How many of our "Gold" tiers would survive a side-by-side comparison of what they actually guarantee?

PART II / VI

Fear

Buying another array to avoid Ceph doesn't neutralise risk — it converts it. Four taxes, invoiced annually, across N+1 platforms. Fear as a strategy pays compound interest.

03 FEAR CHECK

"Ceph will eat our business."

Turn each fear into a question with an evidence-based answer.

"Ceph loses data."

Q: can we name the failure we fear, and has anyone reproduced it in a lab? Durability is design + rehearsed recovery — on Ceph **and** on arrays. If it has a name, it can be tested and priced. If not, it's a mood.

"There's no one to sue."

Q: what does our current vendor contract entitle us to when data is lost — and have we read it? Multi-vendor foundation. Enterprise distros with 24/7. Specialised L3 partners with written SLAs.

"Our team can't operate it."

Q: permanent fact about our people, or a training budget never approved? Most legitimate fear on the list — and the one with the most predictable remediation: structured training with real labs + support partner.

"Nobody gets fired for buying [vendor]."

Q: what happens to that career logic when the post-incident review shows we paid a fear premium for capacity we already owned? The defensible position isn't a logo — it's a decision gate, a benchmark and a rehearsed procedure.

Coexistence model that survives contact with reality

- **Arrays keep what they're demonstrably best at** — the regulated DB, the latency tier, the replication your DR depends on. They stay.
- **Ceph becomes default for new pooled capacity** — general-purpose block, images, object, dev/test, backup. Per-tenant pools, quotas and rehearsed isolation from day one.
- **"Default" = burden of proof, not outcome.** New array must justify why shared won't do. New Ceph move must show measurable value. Symmetric scepticism.
- **Fear gets a de-risking path, not a veto.** Ceph enters via small, reversible workloads first, then promotes to critical tiers only after it earns it on your gear, with your team.

THE FEAR CHECK

- Can we name the specific Ceph failure we fear — and spent one lab-week testing it, versus everything spent avoiding it?
- What is the five-year cost of the systems we bought defensively, including their four taxes?
- Which of our fears came from a measured incident, and which came from a vendor slide?

Two-week de-risking sprint with a team that has broken and fixed production clusters — the named failure reproduced on a correctly designed cluster. Not a full architectural PoC (that lives in the October slot of section 12), but enough to convert the fear into either confidence or a documented legitimate boundary. Rounding error of one defensive array. Both outcomes are wins.

PART III / VI

Your existing storage

OpenStack being new is not a reason to throw away an IBM / Dell / NetApp / Pure / Hitachi / HPE estate that already earns its place. Five tests decide whether an array stays.

04 FIVE TESTS

The five tests. Five yeses or nothing.

Cinder is designed for multiple back ends. But not every array automatically earns a place. Five questions with evidence — or the array doesn't stay.

01 **Is the exact integration supportable?** Not the manufacturer name — the exact array family, code level, Cinder driver version, distro, host OS, transport, multipath, snapshot/clone workflow, replication mode, failure behaviour.

02 **Does it have genuine headroom?** Nominal free capacity ≠ controller performance, front-end ports, cache, replication bandwidth or expansion runway. Plan throughput, IOPS, latency percentiles and recovery load — not just TB.

03 **Do its characteristics match a real workload class?** Sensible home for regulated DBs, latency-sensitive transactional, established sync/async replication. Poor economic home for disposable dev, bulk object, or rapidly growing general-purpose cloud capacity.

04 **Does its lifecycle fit the cloud roadmap?** An array approaching EOS shouldn't become the unexamined foundation of a five-year cloud. Transitional tier ≠ permanent default.

05 **Is there an exit path?** Before onboarding: document how the workload moves to another back end, Ceph RBD, another region, or an app-level replication target. The service isn't complete until its exit is understood.

> Five yeses: the array **stays.** Anything else: you found exactly where the validation work has to happen.

05 SERVICE CATALOGUE

Cinder should expose intent — not purchasing history.

Tenants should not choose between `flashsystem-01`, `powerstore-prod`, `old-san-fast`. That catalogue describes history.

PUBLIC VOLUME TYPE	INTENDED SERVICE
`block-critical-ha`	Low-latency persistent block with strongest availability policy
`block-standard`	General-purpose persistent storage
`block-capacity`	Capacity-efficient for validated workloads
`block-encrypted-dr`	Encrypted with approved remote recovery workflow
`block-high-iops`	Performance tier with explicit latency & throughput acceptance criteria

The back end behind a volume type may be an array today and Ceph tomorrow. Tenants request service, not brand. Cinder standardises workflows; it does **not** make every system semantically identical — snapshots, replication, encryption, QoS, migration still depend on driver capabilities. Represent those in policy and tested — not hidden behind optimistic naming.

CATALOGUE OWNER QUESTIONS

- Could a tenant guess which vendors we bought from by reading our volume-type list? If yes, catalogue describes history.
- Which exact operations — snapshot, retype, failover, migration — have we tested end-to-end per tier? Where's the test record?
- If we swapped the back end behind `block-standard`, would any tenant have to change anything?

PART IV / VI

The architecture

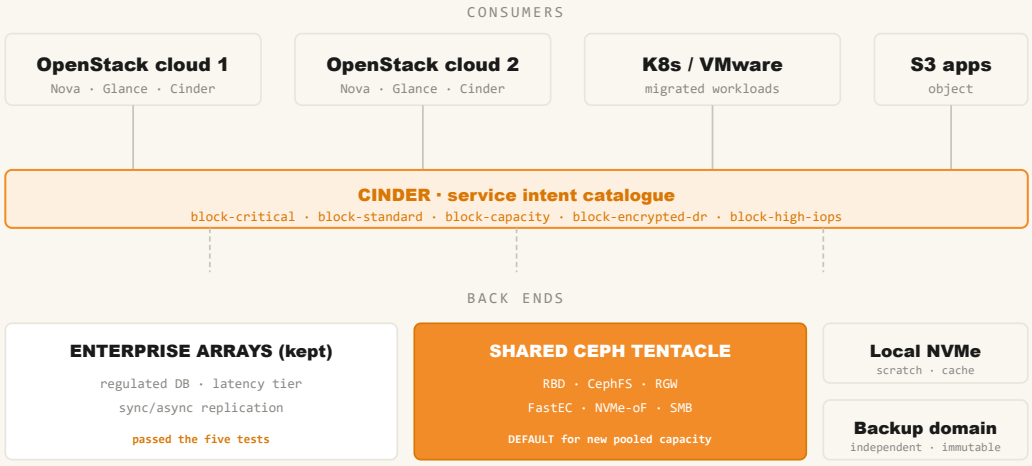
Five patterns. Some need much more justification than others. Dedicated clusters are for explicit boundaries — not new OpenStack control planes.

06 FIVE PATTERNS

The five architecture patterns.

PATTERN	WHEN	ADVANTAGE	RISK
Existing arrays via Cinder	Supported systems w/ headroom & data services	Preserves investment & ops maturity	Cost, lifecycle, driver limits
Shared external Ceph	Several clouds need pooled block/file/object	Pooling, open interfaces, unified lifecycle	Shared blast radius, coordinated upgrades
Dedicated Ceph per cloud	Hard legal, geographic, ownership boundary	Autonomy & containment	Duplicated capacity, ops, spares
Hyperconverged	Small / edge sites w/ predictable ratios	Efficient footprint, simple deploy	Contention, coupled scaling
Specialised silo	Measured need shared can't meet	Optimised for one protocol/workload	Lock-in, low utilisation, costly exit

FIG 02 · THE HYBRID PATTERN THAT USUALLY SURVIVES SCRUTINY



Shared ≠ uncontrolled. A shared Ceph design must define per-cloud pools + CephX identities, quotas, CRUSH rules + device classes, network SLOs, noisy-neighbour controls, maintenance ownership, client compatibility, upgrade coordination, and a documented emergency isolation procedure.

HONEST ABOUT WHAT SHARED SHARES

Per-tenant pools and CephX **do not** eliminate every failure mode. MON quorum loss, mgr crash, RADOS-wide config error, storage-network partition and CRUSH map errors affect every consumer of the shared platform simultaneously. Per-tenant boundary makes each tenant's **data** path explicit; it does not privatise the control plane.

This is why dedicated clusters remain valid for workloads whose recovery objective demands independence from another cloud's control-plane events (see section 07), and why **sub-500µs consistent-latency tiers stay on dedicated arrays with dedicated controllers**. Ceph on NVMe with replication reaches p99 in the low millisecond range; below that is array territory.

07 DEDICATED CEPH

When a dedicated cluster makes sense.

Not "always separate." Not "one giant cluster at any cost." Justified when the **boundary itself** delivers measurable value.

Regulatory / contractual

Customer, country, security domain, regulated workload requires independent admin, keys, audit scope, physical placement.

Geographic / network

Edge / disconnected env. Don't stretch a shared failure domain across an unsuitable WAN just to avoid deploying a cluster.

Independent ownership

Neither BU can accept the other's upgrade calendar or incident authority. The org boundary is already real — acknowledge it.

Blast radius / scale

Explicit RTO/RPO requires independence, or scale is large enough to split admin domains.

FIVE MANDATORY QUESTIONS FOR EVERY NEW-SILO REQUEST

1. What precise boundary are we creating?
2. Which risk does it reduce?
3. How will that reduction be measured?
4. What duplicated cost does the boundary introduce?
5. Why can pools, credentials, CRUSH rules, device classes, volume types or org controls not satisfy the requirement?

No evidence-based answers? The organisation doesn't need another cluster — **it has a preference**. Preferences shouldn't be disguised as architecture.

PART V / VI

The tech

Ceph Tentacle 20.2.2 is the production baseline. OpenStack Gazpacho is the H2 anchor. Umbrella and Hibiscus deserve laboratory time — not 2026 promises.

08 CEPH TENTACLE

Tentacle 20.2.2: what materially changes.

Strategic capabilities arrived with 20.2.0. Two point releases of production hardening later, 20.2.2 (June 16, 2026) is the current baseline.[1]

FASTEC

Capacity conversation changes

New EC I/O implementation; opt-in via `allow_ec_optimizations`; default plug-in for new clusters is ISA-L (not Jerasure).

RBD MIGRATION

Supports de-siloing

Import from another Ceph cluster in native format, or from external sources via NBD. RBD mirroring gains namespace-remapping.

PROTOCOLS

Consolidation options

NVMe/TCP gateway groups + secure listeners + DHCHAP; integrated SMB via Samba+CTDB; RGW bucket resharding; IAM Accounts; management gateway w/ OAuth2/OIDC.

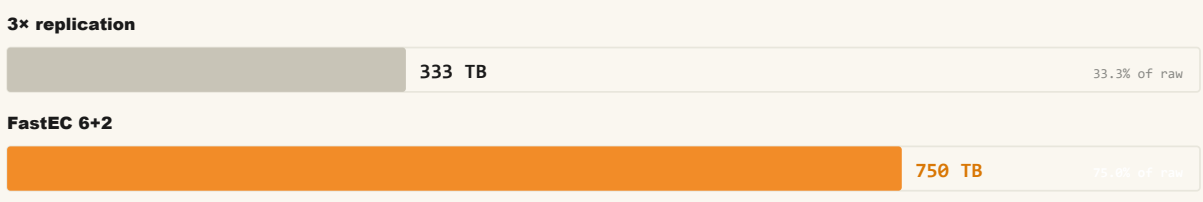
Multi-tenancy note: at scale, prefer RBD/RGW namespaces over one-pool-per-tenant. Pool count and PGs-per-OSD have practical ceilings.

QUIET WINS

BlueStore + OMAP

Improved compression, faster WAL, faster OMAP iteration (RGW bucket listing, scrubs). Foundations, not headlines.

FIG 03 · FASTEC · THEORETICAL USABLE CAPACITY PER 1 PB RAW



6+2 chosen as a mid-scale-common reference. Conservative 4+2 gives ~667 TB/PB (66.7%), wide 8+3 gives ~727 TB/PB (72.7%). The economic direction holds across profiles. Figures exclude operational reserves, BlueStore overhead, uneven utilisation, metadata, failure-domain constraints — not sizing results. FastEC ≠ replication: recovery, tail latency, object size, write ratio, queue depth still matter. Upstream has not yet published headline % improvement figures for FastEC; **benchmark locally against your profiles before committing capacity plans.** Latency-critical tiers keep replicas.

> Which workloads did we **reject** for erasure coding in 2023–2025, and have we rebenchmarked any since the economic boundary moved?

TECH-PREVIEW REMINDER

Crimson, SeaStore, pool-level Data Availability Score: tech preview.[7] Deserve lab time. **Roadmaps are useful. They are not purchase orders.**

09-11 WHAT COMES NEXT

Umbrella & Hibiscus enter the laboratory.

Design production around delivered features. Prepare readiness for upcoming ones. Don't justify a 2026 purchase with an unreleased feature.

CEPH · UMBRELLA (v21)

Prepare, don't depend.

Umbrella is identified upstream as the release after Tentacle. Anticipated: transparent pool migration at RADOS layer[8] — extremely valuable for moving between replicated/EC. Monitor-store backup may appear;[9] upstream is explicit it complements, not replaces, recovery. No official release date yet — treat every date you hear as tentative.

Readiness: keep clients/kernels on supportable paths · inventory old Ceph clients · test recovery procedures · reserve lab capacity.

OPENSTACK · GAZPACHO

The H2 anchor.

SLURP release. Rolling upgrade from Epoxy skipping Flamingo.[3] Nova parallel live migration, IOThread default, async volume-attachment API (2.101), tunable `qemu-img` address space for Ceph.[10]

Read like an engineer: Gazpacho improves machinery. Architecture still determines the result.

OPENSTACK · HIBISCUS

Estimated Sept 30, 2026.

Nova-conductor / nova-compute expected to use native threading by default. Upstream recommends separating the concurrency-mode change from the version upgrade.[12]

Not a SLURP: shorter maintenance horizon. Choose Hibiscus for features knowingly, not for the calendar.

RULE

Separate lifecycle decisions.

OpenStack and Ceph interact — they don't need to be upgraded as one indivisible event. Never combine major OpenStack upgrade + major Ceph upgrade + replicated→EC + storage-network redesign + workload migration in one weekend.

A platform may need all five changes. It doesn't need all five surprises on the same weekend.

> OpenStack Gazpacho does not require **"Ceph Gazpacho Edition"** — because no such dependency exists.

PART VI / VI

Making the

A calendar with deadlines. A budget that buys optionality. Twenty questions that decide the architecture. And a map from your gaps to the next phase.

12 BUDGET CALENDAR

What to do with the rest of 2026.

The worst year-end outcome isn't an unspent budget. It's a fully spent one that creates five years of avoidable operating cost.

1

BY AUG 31

Establish the truth

Estate + workload assessment. Capacity, utilisation, latency, EOL, drivers, growth.

2

BY SEP 30

Defend the architecture

Score 3 options: array reuse · hybrid · shared Ceph. Reserve long-lead media in parallel.

3

BY OCT 31

Test something

PoC w/ real workloads: p95/p99, degraded, failure, recovery, migration, upgrade steps.

4

BY NOV 30

Migration-ready

Waves, rollback, runbooks. Team trained before cutover. A runbook written during an incident is a transcript.

5

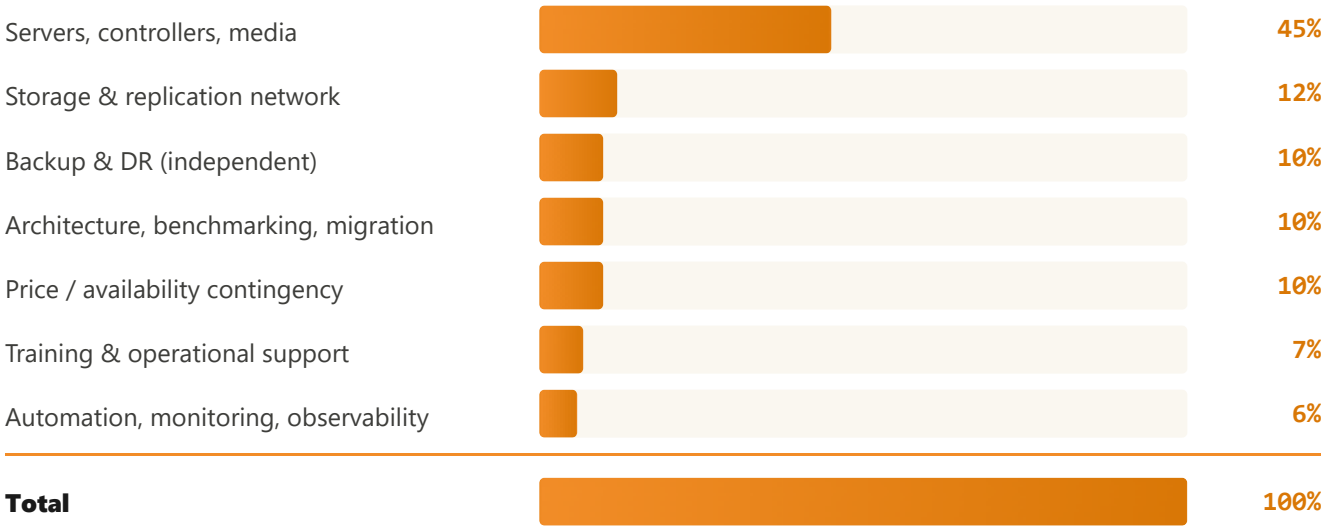
BY DEC 31

Buy architecture, not just HW

Network fabric, spares, backup, monitoring, automation, docs, training, support.

Illustrative 2026 budget model

Not a quotation. A starting point for the balance conversation.



A rack full of NAND without tested recovery, qualified networking or trained operators isn't a storage service. **It's inventory.**

13 SELF-ASSESSMENT

Twenty questions that decide the architecture.

Print. Take to next infrastructure meeting. Where you answer with evidence — you're fine. Where you answer with adjectives — you found your next project.

FOR BUSINESS & SEMI-TECHNICAL READERS

01 Can we state total usable vs consumed capacity across all platforms — on one page?

02 Do we know which platforms exit support during the life of the cloud we're building?

03 If our top-priority environment needed significant capacity in 90 days, do we know where it comes from and at what lead time?

04 Does every storage boundary have a written owner and a written reason?

05 Are our storage tiers defined by service commitments (latency, availability, recovery) or by the vendor we bought them from?

06 What would leaving each platform cost — and did we know before we signed?

07 Is remaining 2026 budget allocated across hardware, network, protection, engineering & skills — or 100% media?

FOR THE TECHNICAL READER

08 For each array: have we validated the exact Cinder driver, code level, transport, multipath and failure behaviour — or only the compatibility matrix?

09 Which workloads did we exclude from EC before FastEC, and which have we rebenchmarked on Tentacle 20.2.2?

10 Do our EC benchmarks include degraded, backfill and 99th percentile tail latency — or only healthy-cluster throughput?

11 Could we migrate an RBD image between clusters today, live, with a rollback plan? Have we rehearsed it?

12 In a shared Ceph design: are per-cloud pools, CephX identities, quotas, CRUSH rules and an emergency isolation procedure actually defined, or assumed?

13	Which Ceph client versions exist across our estate, and which would block Tentacle features or Umbrella's anticipated pool migration?
14	Are our monitor, configuration and encryption-key recovery procedures tested this year — not written, **tested**?
15	Can we upgrade OpenStack and Ceph independently, or has coupling crept into our deployment tooling?
16	Have we tested Nova's parallel live migration against our storage fabric's actual headroom, or only read the release note?
17	Have we planned the Eventlet-to-native-threading transition for Hibiscus as a separate change from the version upgrade?
18	For every proposed new silo: do the five decision-gate questions from section 07 have written, evidence-based answers?
19	Does every service in our catalogue have a documented and tested exit path?
20	If our storage architect resigned tomorrow, is the rationale for the current architecture written down anywhere they aren't?

> No organisation we've met answers all twenty with evidence. That's not a criticism — it's the honest starting point. The interesting information is **which** questions you couldn't answer, because they cluster.

14 NEXT PHASE MAP

Which gaps map to which next phase.

Architectures are solutions to problems. The next step depends on where your gaps sit. This is how we'd read your questionnaire — and what SIXE does in each case.[13]

YOUR GAP	NEXT PHASE
Q 1–4 · You can't describe what you own	Current-state storage map
Q 5 · Tiers describe vendors, not services	Workload-to-service mapping
Q 7, 18 · Decisions rest on preference, not comparison	Competing target architectures (scored)
Fear check · You have never tested what you fear	Two-week de-risking pilot in a lab
Q 8–12, 16 · Nothing measured on your workloads	Representative PoC (failures + recovery + migration)
Q 3, 6, 11, 15, 17, 19 · Exit and execution exist only as intention	Procurement & migration plan
Q 13–14, 20 · Knowledge lives in one head, or nowhere	Knowledge transfer + production support

The desired result is never permanent dependence on a consultant. It's an internal team that understands why the architecture looks the way it does — and knows how to change it safely. Whether the first conversation is a one-day workshop or a full PoC, the questionnaire will already have told you. Bring your answers; we'll bring the lab.

99 CONCLUSIONS

H2 2026 — six lines to remember.

01

Shortage isn't over.

Plan around uncertainty. Don't wait for an imaginary reset.

02

New cloud ≠ new storage.

Qualify what you already own before signing anything.

03

Test the fear, don't pay for it.

Name the scenario, reproduce it in a lab, act on the result.

04

Shared Ceph = default candidate.

For new pooled capacity — governed, segmented, operated as critical infra.

05

Dedicated only for real boundaries.

Regulation, geography, ownership, blast radius, scale. Not "new control plane."

06

Buy architecture, not inventory.

Optionality — network, backup, spares, automation, observability, docs, training.

> Buy storage when you need storage. Build a silo only when you need a **boundary.**

Because — as we said in January —

you decide. This is the only important thing.

Our job is to make sure you decide with your questions answered.

REFERENCES

References.

1. Ceph Tentacle release notes — docs.ceph.com
2. TrendForce — NAND Flash contract price forecast, Q1 2026
3. OpenStack release series and schedules
4. TrendForce — NAND Flash contract price forecast, Q2 2026
5. OpenStack Cinder — multiple storage back ends
6. Ceph — block devices and OpenStack
7. Ceph blog — v20.2.0 Tentacle released
8. Ceph blog — transparent pool migration / CephFS transcoding
9. Ceph blog — monitor recovery from OSDs
10. Nova release notes — 2026.1 Gazpacho
11. Cinder release notes — 2026.1 Gazpacho
12. Nova release notes — unreleased (Hibiscus)
13. SIXE — Private Cloud services
14. SIXE — Ceph vs MinIO 2026 distributed storage guide (Jan 2026)