Chapter 15 — Capacity & Scaling

15.1 Philosophy and Scope

Capacity planning at this fleet size is not autoscaling — it is **knowing the numbers, naming the thresholds, and pre-writing the procedures** so scaling is a scheduled maintenance action rather than an incident response. The Platform's load profile helps: Compartment agent workloads are queue-driven and tolerate minutes of latency (a traffic-log reconciliation that waits 5 minutes is not an incident), so the design scales for *throughput over time*, not burst absorption. The only latency-sensitive paths are the human web surfaces and the audit write path. Scope: droplet sizing, per-Compartment resource budgets, storage growth, the thresholds that trigger action, and the procedures per scaling event. Growth triggers named in earlier chapters (Vault HA §8.2, remote-write metrics §12.2, hot-standby DR §14.4, private DNS §3.3) consolidate here.

15.2 Baseline Sizing and Budgets

As-built fleet totals: **22 GB RAM / 11 vCPU / 500 GB disk** across five droplets (§2.2.1, incl. `prod-vault` at creation). Per-droplet budgets:

`sbsdash-server-prod` (8 GB / 4 vCPU / 160 GB) — the contended host. Budget under the §6.5 interim (data services colocated):

Allocation	RAM	Notes
OS + Docker + nginx + agents	1.0 GB	
PostgreSQL	1.5 GB	`shared_buffers` 512 MB, capped connections via per-Compartment pool
MongoDB	1.0 GB	WiredTiger cache capped 512 MB
Redis	0.5 GB	`maxmemory` 384 MB, `allkeys-lru` on cache keyspaces
Audit writer + buffer	0.5 GB	
Compartment containers	**3.5 GB**	≈ 1.0–1.2 GB per Compartment stack (app + MCP sidecar) at §6.4 limits
Headroom	~0.5 GB	Below 1 GB free is itself a warning signal

Consequence, stated plainly: **the 8 GB prod droplet supports ~3 concurrent Compartment stacks in the interim topology.** That matches the engagement's minimum-Compartment starting footprint, but full ten-Compartment operation on this host requires either the data-service offload (Managed DBs, §7.2) or the resize/split ladder (§15.4) — this is a designed checkpoint, not a surprise.

Other hosts: `monitor-servers` (4 GB) — Prometheus 30 d TSDB ≈ 2–6 GB disk at current cardinality, Loki 90 d on its Volume; RAM adequate until Loki ingest grows with Compartment count (threshold below). `prod-vault` (1 GB) — trivial load, no growth coupling. `sbs-wiki` (2 GB) — static. `Dev-SBS-server` — mirrors prod budgets by design (§2.4).

15.3 Growth Thresholds (Named Triggers → Named Actions)

Thresholds are Prometheus alert rules (§12.4 `warning` tier unless noted), each mapped to a §15.4 procedure — a threshold without a pre-decided action is just anxiety:

Signal	Threshold	Action
Prod RAM sustained (15 min)	≥ 85 %	P1: resize droplet
Prod free RAM	< 1 GB at Compartment deploy time	Deploy blocked; P1 or P2 first
Compartment count	4th Compartment Workorder received	**P2: data-service offload decision forced** (Managed DB check, item 36, must be closed by now)
Per-Compartment container at its §6.4 limit	throttling/OOM-kill events > 0	Raise that Compartment's budget via PR (visible, reviewed) — never silently unlimit
Any disk	≥ 80 % / ≥ 90 %	warning: P3 volume grow / critical: immediate P3
PG connections	≥ 80 % of max	Pool tuning, then P2
Loki ingest	> 5 GB/day sustained	P4: monitor resize or Loki retention drop to 60 d (decision, not drift)
Prometheus TSDB	> 40 GB or scrape latency alerts	P4 + revisit remote-write (§12.2 deferred item)
Audit writer queue age	> 5 min sustained under normal ops (distinct from §13.3 outage alert)	P5: writer I/O tuning / dedicated Volume IOPS
Fleet size	> 10 droplets	Private DNS zone (§3.3 trigger); Vault HA evaluation (§8.2)
Any single droplet	> 16 GB resize on the table	P6: split evaluation instead — vertical scaling stops being the answer

15.4 Scaling Procedures

All procedures are IaC changes (§10.4 workflow) executed in maintenance windows; none is novel at execution time because each is pre-written as a wiki runbook:

- **P1 — Droplet resize (vertical).** tofu size change → PR → apply = powered-off resize (**~1–3 min downtime**, disk+CPU+RAM resize; reserve the disk-inclusive variant so storage grows too). Order of preference: resize prod before adding hosts — vertical is operationally free until §15.3's 16 GB ceiling. Tier-1 backup taken pre-resize (fast rollback = restore + revert PR).
- **P2 — Data-service offload.** The designed inflection at Compartment #4: Managed DBs available in NYC2 (item 36) → migrate per §7.2 (dump/restore per engine, maintenance window, Change Order) → frees ~3 GB on prod → Compartment ceiling rises to ~7–8 on the same droplet. If Managed DBs unavailable → dedicated `prod-data` droplet (8 GB) in the Prod VPC running the §6.5 compose stacks — same isolation model, same entitlement networks, one hop away.
- **P3 — Storage grow.** DO Volumes resize online (grow-only); filesystem expand follows in the same window. Root-disk pressure on droplets without Volumes → P1 disk-inclusive resize. Spaces is elastic — no procedure, only the cost line.
- **P4 — Observability scaling.** Monitor resize (P1 pattern) or retention reduction — an explicit reviewed trade, never silent data loss.
- **P5 — Audit path tuning.** Volume IOPS/size bump, fsync batching review (§13.3 semantics preserved — Guardian stream stays per-event), segment-close interval tuning.
- **P6 — Horizontal split.** When prod outgrows vertical: split along the already-drawn lines — data tier out first (P2's droplet variant), then Compartments 6–10 to a second app droplet (`net-edge` spans hosts via VPC; nginx upstreams update by IaC variable). The compose-per-Compartment design (§6.4) makes Compartment placement a *scheduling* decision, not a re-architecture — this is the payoff for refusing the mega-compose.
- **Scale-down** is the same procedures reversed, gated on 30 days below 50 % of the triggering threshold — flap-damping for capacity decisions.

15.5 Capacity Review Cadence

Quarterly (aligned with the §12.5/§13.5 quarterly review artifacts): dashboard-sourced trend review of RAM/CPU/disk/ingest slopes per host, Compartment-count forecast against the P2 checkpoint, cost line (droplets + Volumes + Spaces + DO Backups usage-based charges + Anthropic consumption trend from egress/billing data), and a written one-page capacity position filed to the wiki. The review's only mandatory output: **confirmation that the next scaling event is known, named, and scheduled** — or explicitly "none within horizon."

15.6 Verification (Build Gates)

Gates (LH-SBS-INST-001): §15.3 thresholds present as alert rules in the repo (promtool-validated, mapped action in annotation); §6.4 resource limits present on every Compartment service and OOM/throttle events exported as metrics; P1 resize drill executed once on Dev (timed, rollback path proven); Compartment-deploy RAM precondition check wired into the deploy pipeline (blocked-deploy negative test); P2 decision documented and item 36 closed before 4th-Compartment Workorder acceptance; capacity runbooks (P1–P6) published to wiki; first quarterly capacity position filed.

Revision #1
Created 2026-07-18 12:07:38 UTC by SBS Admin

Updated 2026-07-18 12:09:23 UTC by SBS Admin