

Platform scaling and multi-environment architecture — design review (DR)

Platform scaling — DR (rev.2: bandwidth + GEO)

Table of contents

1. Executive summary
2. Problem statement
3. Architecture options
 - 3.1 Option A — Dev cluster (draft default)
 - 3.2 Option B — Tighter cost (minimal split)
 - 3.3 Option C — Cloud-first hybrid
 - 3.4 Comparison matrix
4. DR decision — recommended path
5. Current vs target diagrams
 - Today (single BE)
 - Target phase B (minimal split + HA web)
6. Service split rationale
7. Environment promotion model
8. RSSH and connection capacity
 - Assumptions
 - Estimates
- 8b. Network bandwidth (DR)
 - Baseline risk (today)
 - DR sizing decisions
 - Planning numbers (single-site prod, until CDN)
- 8c. Global availability and GEO segments (DR)
 - Policy model (DR decision)
 - Phasing
 - Example use cases
9. Future capabilities — priority weighting
10. Simulation lab design
11. Cloud provider notes (GCP / Azure)
12. RBAC evolution
13. Risks and mitigations
14. Estimates
15. Task dependency graph
16. Resolved open questions
17. Remaining open items (PO)
18. Suggestions (PO / dev discretion)
19. Changelog
- Related docs

Source: 20100612_1_scaling.md — section links work in PDF viewers that support internal anchors.

Document type: DR (Design Review)
Source draft: docs/drafts/20100612_1_scaling.txt
Specification: docs/specs/20100612_1_scaling.md
PDF: 20100612_1_scaling.pdf · Regenerate: bash scripts/build-all-docs-pdf.sh (Mermaid → PNG diagrams; requires npx)
Date: 2026-06-12
Status: Reviewed — planning / brainstorming (no implementation sprint yet)
Severity: Strategic

1. Executive summary

The project outgrew **one Alpine BE VM** that hosts consoles, APIs, MariaDB, Gitea, OTA, WireGuard, RSSH, URL shortener, and Docker builds. The draft asks for **clusters**, **dev/staging/prod**, **zero-downtime deploys**, optional **cloud**, and headroom for **RT cast**, **VPN commercialization**, **payments**, and **ffmpeg** pipelines.

DR outcome:

Area	Decision
Timing	Post-alpha strategic track; does not block alpha soak
Architecture	Option B → A phased: minimal VM split first, then full dev cluster simulation
Edge	Keep Gentoo FE as TLS/DNAT anchor; add LB upstream to N BE web nodes
Data	Dedicated MariaDB VM before scaling anything else CPU-heavy
Build	First split candidate — Docker off shared PHP VM
Cloud	Hybrid — object store + staging dress rehearsal before prod migration
RSSH capacity	Plan 100 concurrent SSH / 5k heartbeats per bastion pair before commercial VPN
Bandwidth	Document WAN / FE→BE / per-service Mbps & GB-month budgets (SPEC §12.3); CDN before horizontal web scale for OTA
Global + GEO	Worldwide reach default; deny segments per content (e.g. no serve in ZM) via layered enforcement (SPEC §14)
Future slots	Weighted backlog in §9 ; SFU before ffmpeg farm; F10 GEO with F3/F5

2. Problem statement

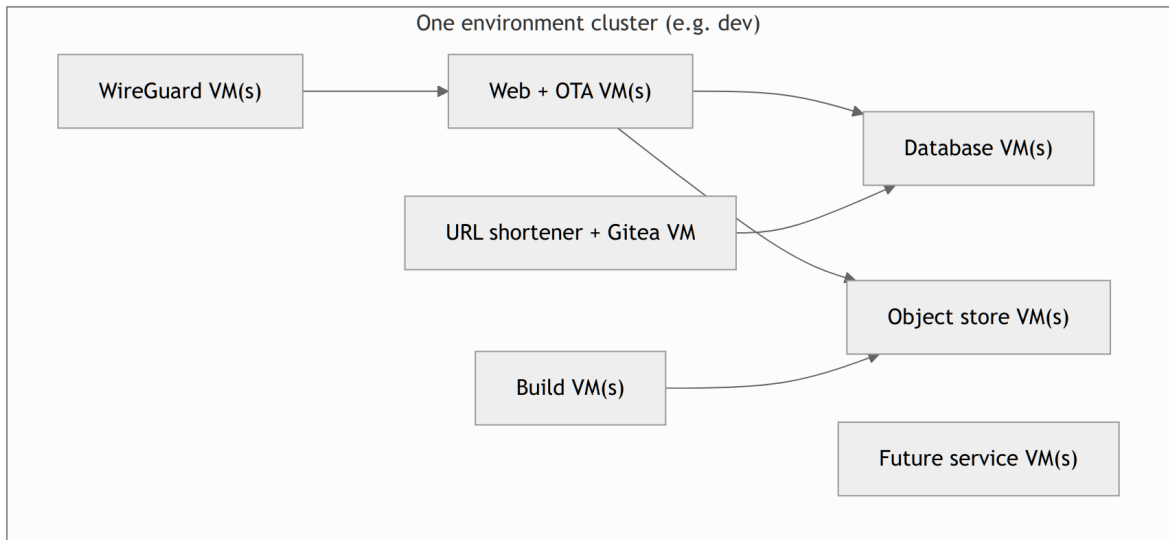
Pain	Evidence today
Blast radius	Submodule sync wiped gitignored configs; one BE restart affects all services
Resource contention	Docker build spikes CPU/RAM during PHP console use
No promotion path	Single prod-like env; staging OTA path still pending (20260608_BE_SERVICES_and_infra.md)
Scaling ceiling	RSSH + WG + future SFU on same VM as MariaDB
Bandwidth ceiling	Single FE WAN + FE→BE path; OTA/APK egress competes with consoles (SPEC §12.3)
GEO gap	No per-content region policy yet; required before global recordings/commerce (SPEC §14)
Operational risk	Topology changes on FE/router are hard to roll back (bottomline_reminder.txt)

Success criteria: PO can deploy unstable work to **dev**, beta-test on **staging**, and promote to **prod** with rollback per service — without rewriting mobile apps or public URLs each time.

3. Architecture options

3.1 Option A — Dev cluster (draft default)

Draft proposal: **N VMs per environment**, each role class (web+OTA, DB, object store, WG, url shortener+Gitea, build, +N for new features).

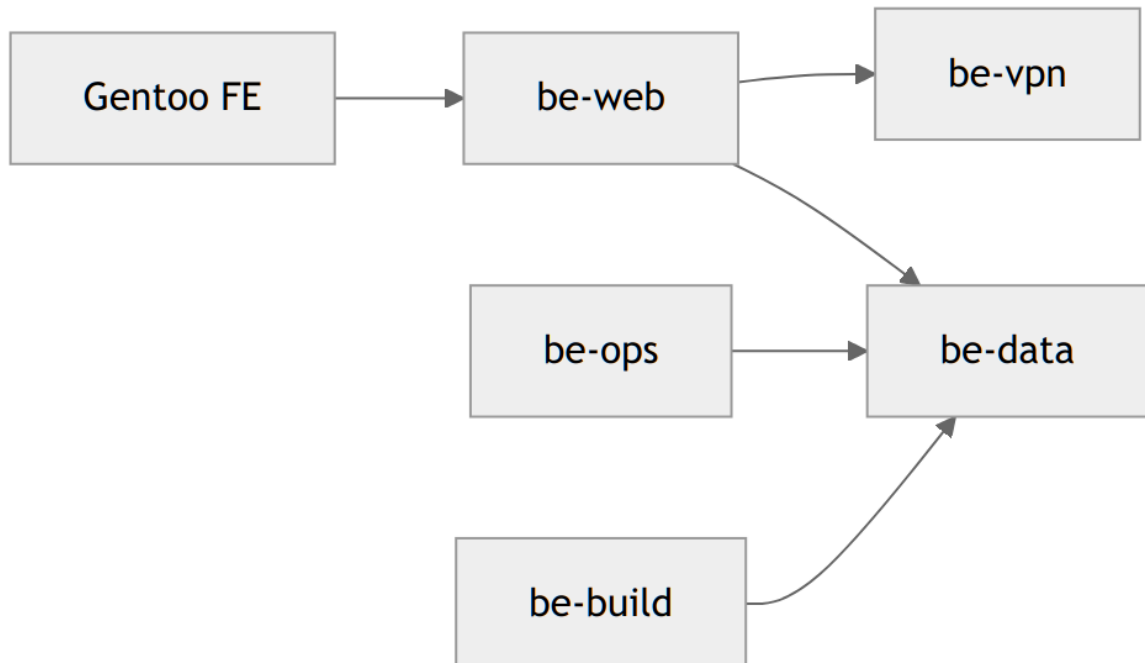


Pros	Cons
Clear isolation; matches future SFU/ffmpeg	Highest VM count / cost
Easy to simulate on FE XEN	Ops overhead for small team
Aligns with cloud MIG later	Overkill before alpha exit

3.2 Option B — Tighter cost (minimal split)

Phase 1 prod shape (recommended start):

VM	Bundled services
be-web	nginx, PHP (hub, crashes, graphs, builder UI, OTA metadata)
be-data	MariaDB only
be-ops	Gitea, url-shortener
be-build	Docker builder
be-vpn	WireGuard + RSSH bastion



Pros	Cons
~5 VMs vs 7+; fits "money-wise" draft	Gitea + shortener still coupled
Unblocks build/DB contention immediately	Second web node needed later for true HA
Low risk to current URLs	WG still single point until pool

3.3 Option C — Cloud-first hybrid

Move **staging** entirely to GCP/Azure; keep **prod** on FE→BE until hybrid LB proven.

Pros	Cons
Managed LB, DB, object store day one	Egress cost; UDP/WG complexity
Dress rehearsal for full cloud	PO learning curve; vendor lock-in mild
ffmpeg/SFU elastic scale	Violates "safest topology" unless staged

3.4 Comparison matrix

Scores 1 (poor) – 5 (excellent) for **current team size** and **alpha proximity**:

Criterion	A — Full cluster	B — Minimal split	C — Cloud-first
Cost now	2	4	3
Time to first value	2	5	3
Zero-downtime readiness	4	3	5
Topology safety	4	5	2
SFU/ffmpeg fit	5	3	5
Simulation fidelity	5	4	2
Global CDN / GEO	3	3	5
Weighted total	25	27	25

Weights used: safety and cost now > cloud elasticity pre-alpha.

4. DR decision — recommended path

Now ■■■ Option B (minimal split) on-prem HVMS

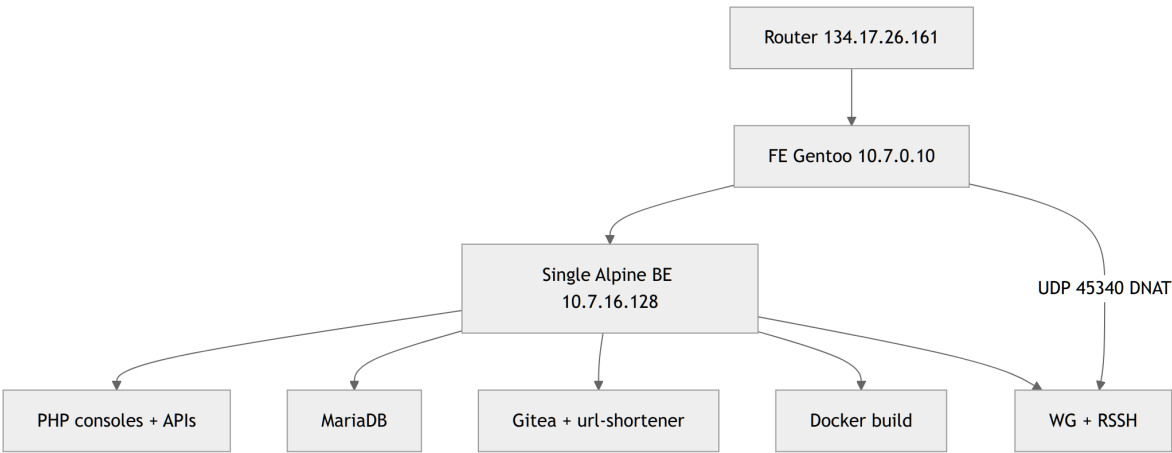
- Simulate Option A as DEV cluster (3-7 Alpine HVMS)
- Add be-web-2 + FE upstream when metrics say so
- Option C for staging only → then hybrid prod

Step	Action	Owner
1	Document service boundaries (this SPEC/DR)	Agent / PO review
2	Move Docker builder to be-build-1	Sysop
3	Move MariaDB to be-data-1 (hardest cutover)	Sysop
4	Stand up dev cluster simulation on intra network	Sysop
5	Staging promotion runbook + DNS prefix	PO + DEV
6	Cloud object store for OTA/APK	Sysop
7	CDN + bandwidth dashboards (SPEC §12.3)	Sysop
8	SFU SPEC + media workers (separate track)	PO spec
9	GEO policy schema + edge rules (SPEC §14)	PO + DEV

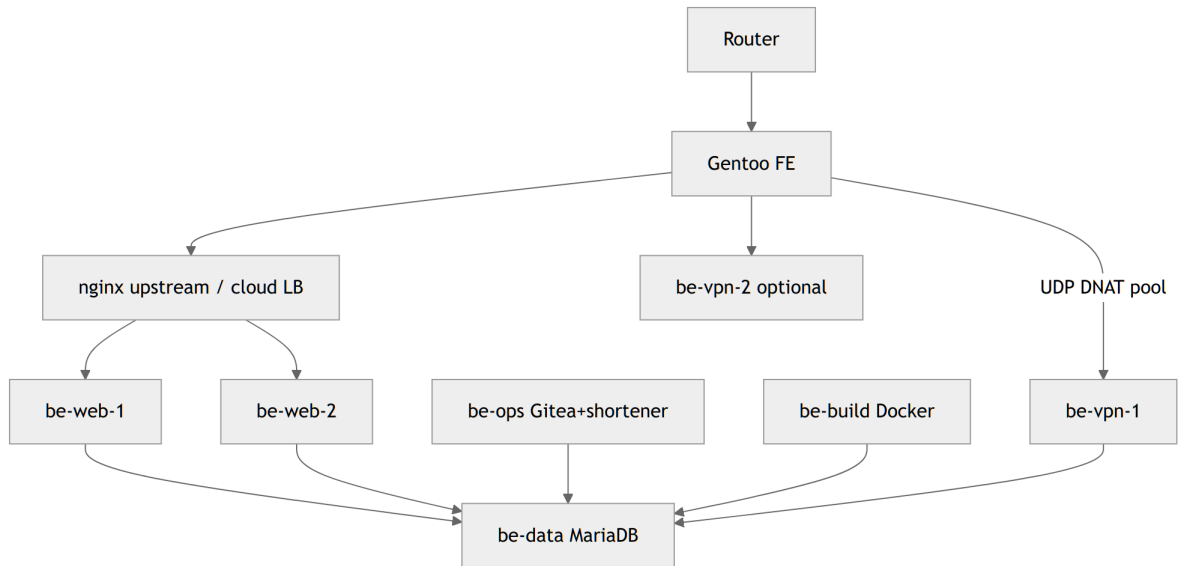
Not chosen now: full Option A in prod; full Option C prod cutover.

5. Current vs target diagrams

Today (single BE)



Target phase B (minimal split + HA web)



6. Service split rationale

Draft asked to balance **cost vs CPU/RAM needs**. DR sizing rules:

Workload	CPU	RAM	Disk	Pin to
PHP consoles	Medium	2–4 GiB	Small	be-web*
MariaDB	Medium	8+ GiB	Fast SSD	be-data
Gitea mirrors	Low	2 GiB	Medium	be-ops
URL shortener	Low	512 MiB	Tiny	be-ops
Docker build	Burst high	16 GiB	Large ephemeral	be-build
WireGuard	Low	1 GiB	Tiny	be-vpn
RSSH sshd	Low–med	2 GiB	Tiny	be-vpn
ffmpeg (future)	Sustained high	8–32 GiB	Temp SSD	be-media*
SFU (future)	High	8+ GiB	Low	be-media*
Recordings (future)	Low	—	Large	object store

Anti-pattern rejected: MariaDB + Docker build + ffmpeg on one VM (draft implied bundle risk).

7. Environment promotion model

Gate	dev → staging	staging → prod
Automated tests	<code>./gradlew :app:testDebugUnitTest + BE smoke</code>	+ soak checklist
Manual	PO feature acceptance	PO sign-off + tag
Data	Synthetic	Anonymized copy optional
Rollback	Re-deploy previous image	Git revert + DB restore runbook

Align app releases with GIT_FLOW.md; infra releases may cadence independently with **config version pins** per environment.

8. RSSH and connection capacity

Draft question: *how many simultaneous connections per instance?*

Assumptions

- Heartbeat poll **1–7 min** (REMOTE_ACCESS_IMPL)
- Average SSH session **15 min** operator work
- Alpine **be-vpn-1**: 2 vCPU, 2 GiB RAM

Estimates

Resource	Limiting factor	Conservative	With tuning
Concurrent SSH	sshd MaxSessions / CPU	50	100
Devices polling heartbeat	PHP-FPM + MariaDB writes	2 000	5 000
New SSH/min	MaxStartups	10	30
WG peers (lab)	kernel + UDP	200	500 / 2 nodes

Tuning knobs: separate **heartbeat** PHP pool to **be-web** (off bastion); keep sshd on **be-vpn** only; connection pooling to MariaDB; read replicas for heartbeat read-heavy phases.

DR recommendation: Instrument **active SSH** and **heartbeat QPS** in graphs before selling commercial VPN (draft F4).

8b. Network bandwidth (DR)

Draft rev.2 adds **throughput** alongside connection counts.

Baseline risk (today)

Path	Risk
Single WAN uplink	OTA release or viral short link saturates all services
FE→BE proxy	All HTTPS doubles through FE NIC
No CDN	Every APK download hits origin BE
Media future	SFU + ffmpeg on same uplink = hard fail without PoP/CDN

DR sizing decisions

Decision	Rationale
CDN before be-web-N for OTA/static	Fixes egress Mbps cheaper than more PHP VMs
Separate media NIC/subnet when SFU lands	Isolate UDP/WebRTC from MariaDB LAN
Egress budget alerts at 80% ISP/cloud quota	PO notified before hard cap
Simulate release day on staging	e.g. 1000 concurrent OTA downloads

Planning numbers (single-site prod, until CDN)

Traffic	Peak Mbps (conservative)	Mitigation
Steady HTTPS (API + consoles)	5–20	Normal
OTA drop (1000 devices × 40 MB APK)	200–400 burst 10 min	CDN required
Recording download (F3)	50–200 per hot file	Object store + CDN signed URLs
SFU fan-out 100 viewers	50–500+	Regional PoP; not on FE WAN

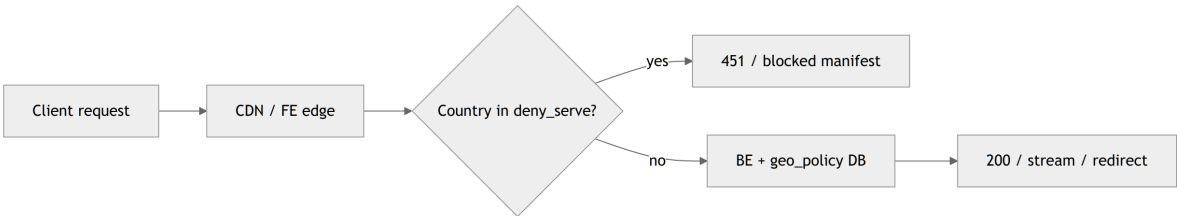
Full table: SPEC §12.3.

8c. Global availability and GEO segments (DR)

Draft rev.2: **global reach** + **segment isolation** (example: deny publish/serve in **Zambia ZM**).

Policy model (DR decision)

Rule	Choice
Default	Allow worldwide for control plane and public APIs
Content objects	Attach optional geo_policy with deny_publish / deny_serve / allow_only
Enforcement order	CDN/WAF → BE API (trusted geo header) → object signed URL → app entitlement
Operator UI	Hub/console editor for geo segments when F3/F10 ship — not alpha
Sanctions	PO-maintained deny lists; not auto-synced from external API in v1



Phasing

Phase	GEO capability
Now	Implicit worldwide; no per-content rules
CDN phase	Edge geo block for coarse country bans
F3 storage	Per-recording geo_policy in MariaDB + object metadata
F5 payments	Entitlement includes geo_allowed
Multi-PoP	GeoDNS routes to nearest allowed PoP only

Example use cases

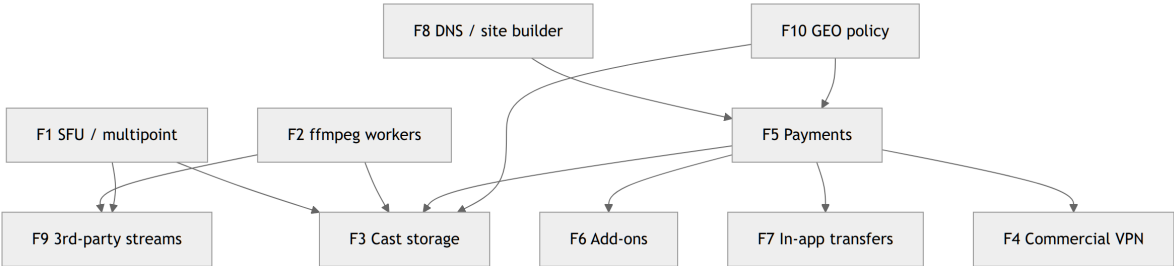
Use case	Policy
OTA stable globally	default: allow
Beta APK EU-only	allow_only: ["DE", "FR", ...]
Recording licensed EU/US only	deny_serve: ["ZM", ...] + allow_only optional
Short link to blocked catalog	BE returns 451 at redirect time

Open for PO: whether VPN exit in denied country circumvents policy (DR: log + ToS; technical block on VPN product separate from HTTP geo).

9. Future capabilities — priority weighting

Weighted for platform dependency and revenue enablement (PO may override):

Rank	Slot	Weight	Rationale
1	F1 Multipoint cast (SFU/MCU)	Critical path	Unblocks WAN cast; already on OPEN_TASKS_GRAPH.md
2	F2 ffmpeg transcode	High	Shared worker pool for F3, F9
3	F3 Cast storage	High	Needs F1+F2 + object store
4	F5 Payments	Medium	Enables F3, F4, F7
5	F4 Commercial VPN	Medium	RSSH/WG scale must be proven first
6	F9 3rd-party streaming	Medium	Uses F2 cast handlers
7	F6 Mobile add-ons	Lower	Can ship on-device first
8	F7 In-app transfers	Lower	Compliance heavy
9	F8 DNS / subdomains	Exploratory	Largest scope creep
10	F10 GEO enforcement	Medium	Required for F3/F5 global commerce; draft rev.2



10. Simulation lab design

Sysop simulates Option A under dev without WAN changes:

HVM	Alpine packages	Role
dev-web-1	nginx, php81-fpm	Web tier
dev-data-1	mariadb	DB
dev-ops-1	git, gitea (optional)	Gitea + shortener
dev-build-1	docker	Builder
dev-vpn-1	wireguard-tools, openssh	VPN lab

Network: 10.7.0.0/8 only; FE routes dev.apps.f0xx.org to dev-web when ready.
Deliverable suggestion: docs/SIMULATION_CLUSTER.md + scripts in phase 1 — not blocking this DR.

11. Cloud provider notes (GCP / Azure)

Concern	GCP	Azure
Managed LB	ILB + NEG	App Gateway / Front Door
Object store	GCS	Blob
Managed MariaDB	Cloud SQL	Azure Database for MariaDB
UDP / WG	harder	harder — keep on-prem FE DNAT longer
ffmpeg burst	MIG (CPU)	VMSS
Egress cost	Watch OTA/APK	Same
CDN + geo WAF	Cloud CDN / Cloud Armor	Azure Front Door + rules
Multi-region	GCS dual-region; Cloud CDN PoPs	Blob + AFD PoPs

DR: Use cloud for **staging cluster first** (Option C partial) — validates IaC without prod risk. **CDN geo rules** prototype on staging before prod F3.

12. RBAC evolution

Stage	Who	Access
Alpha (now)	Developers = operators	Max privileges; shared session
Beta	Split developer vs support operator	Tickets/RA without build admin
Commercial	Customer admin + platform admin	Tenant isolation; audit export

New roles MUST NOT require mobile app changes until F5 entitlements API exists.

13. Risks and mitigations

Risk	Impact	Mitigation
MariaDB migration downtime	Prod outage	Replica + cutover window; practice on dev
Config loss on sync	503 (seen url-shortener)	ensure_ * scripts; secrets outside tree
FE upstream misconfig	Partial outage	nginx -t; weighted rollout
WG UDP cloud move	VPN breaks	Keep DNAT path until PO signs UDP plan
Over-scoping F5–F9	Never ship SFU	S9 order
WAN saturation on OTA	Global outage perception	CDN + S8b budgets
Geo policy bypass via VPN	Compliance / licensing	Separate VPN egress policy; audit logs
Wrong deny list	Legitimate users blocked	Staging test + X-Debug-Country; PO approval workflow
Simulation sprawl	Sysop time	Cap dev HVMs at 7; destroy scripts

14. Estimates

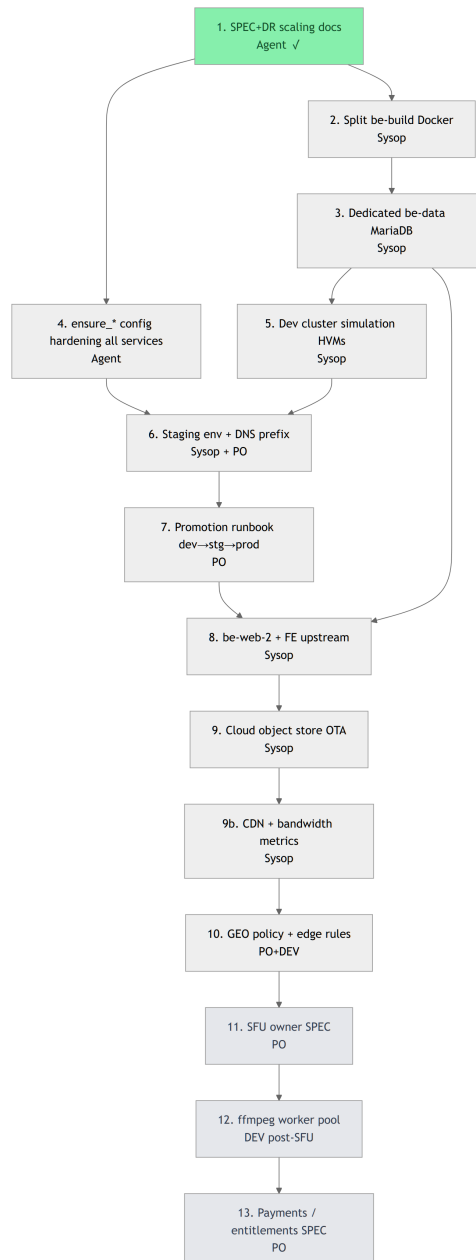
Person-days (**single sysop + DEV support**), planning accuracy ±40%:

Phase	Work	Days
P0	SPEC/DR + simulation design (this doc)	2
P1	Split Docker build VM	2
P2	MariaDB dedicated VM + migration	3
P3	be-ops split (Gitea + shortener)	1
P4	Dev cluster HVM simulation (5 VMs)	4
P5	Staging env + promotion runbook	3
P6	be-web-2 + FE upstream HA	2
P7	Cloud object store for OTA	2
P7b	CDN + bandwidth metrics on graphs	2
P8	GEO policy schema + staging CDN geo test	3
P9	SFU/media (separate SPEC)	TBD
Total (through P7b)		~21 days

Severity: Strategic — parallel to alpha; no alpha gate.

15. Task dependency graph

Legend: **1** = highest priority (PO: post-alpha infra). Assignee **Sysop** = PO today; **Agent** = documentation/automation; **DEV** = app/config changes.



ID	Task	Assignee	Blocks
1	Scaling SPEC/DR	Agent	—
2	Docker → be-build	Sysop	DB split
3	MariaDB → be-data	Sysop	HA web
4	Post-sync config scripts (all services)	Agent	staging
5	Dev HVM simulation	Sysop	staging realism
6	Staging cluster	Sysop	promotion
7	Promotion runbook	PO	prod multi-node
8	Second web node	Sysop	cloud OTA
9	Object store OTA	Sysop	CDN
9b	CDN + bandwidth dashboards	Sysop	GEO test
10	GEO policy + edge	PO+DEV	F3 commerce
11	SFU SPEC	PO	media workers
12	ffmpeg pool	DEV	commercial storage
13	Payments SPEC	PO	F6–F8

Parallel now (non-blocking): alpha soak, url-shortener ops, Gitea mirrors, ticket **SVC-OTA**.

16. Resolved open questions

#	Draft question	DR answer
---	----------------	-----------

Q1	How many VMs in a cluster?	Option B: 5 prod VMs to start; dev simulation: 3–7 (SPEC §6.3)
Q2	"Other" architecture approaches?	Option C hybrid for staging; not prod-first (§3.3)
Q3	Cost vs CPU/RAM balance?	Split by §6 profile table
Q4	Seamless scaling / LB?	FE upstream → N PHP nodes; cloud managed LB in phase 5 (SPEC §9)
Q5	Simultaneous connections / RSSH?	§8 — 100 SSH, 5k heartbeats per bastion pair (tuned)
Q6	Alpine simulation?	§10 — intra 10 . 7 . 0 . 0 / 8 HVMs only
Q7	Future MCU/SFU/ffmpeg/storage/VPN/payments?	Platform slots F1–F9; priority §9
Q8	RBAC model change?	§12 three-stage
Q9	Keep current topology safe?	Yes — SPEC §18
Q10	Network bandwidth metrics?	§8b + SPEC §12.3 — CDN before web horizontal scale
Q11	Global availability + GEO deny (e.g. Zambia)?	§8c + SPEC §14 — default allow, per-content deny

17. Remaining open items (PO)

ID	Question	Needed for
O1	Target monthly infra budget for phase B (5 VMs)?	Option A vs B timing
O2	Staging URL pattern: staging.apps.f0xx.org vs path prefix?	DNS + TLS
O3	Cloud vendor preference GCP vs Azure vs none for 12 months?	P7–P9
O4	Accept maintenance window for MariaDB cutover (minutes of read-only)?	P2 schedule
O5	Commercial VPN — same WG infra or separate product stack?	F4 design
O6	SFU — Janus on :8089 vs greenfield mediasoup?	F1 SPEC (20260608_BE_SERVICES_and_infra.md)
O7	Payments — Stripe-only vs multi-provider?	F5 SPEC
O8	ISP/cloud egress cap (GB/month) on FE WAN?	§8b budgets
O9	CDN provider (Cloudflare vs cloud-native vs none 12 mo)?	OTA + GEO WAF
O10	GEO deny lists — PO-owned static JSON vs admin UI first?	F10 rollout
O11	Recordings 451 vs silent 404 for denied regions?	UX + compliance

When blocked on O1–O7, append to tmp/YYYYMMDD_implementation_summary.pdf per bottomline_reminder.

18. Suggestions (PO / dev discretion)

#	Suggestion
S1	Add infra smoke to validate_be_services.sh per VM role after split
S2	Version nginx upstream fragments in examples/crash_reporter/backend/deploy/
S3	Use Docker for PHP-FPM pools only after MariaDB split — avoid double complexity
S4	Keep Citea on be-ops until mirror count > 10 repos
S5	Document connection pool sizes when RSSH device count > 500
S6	Re-read this DR when SFU SPEC lands — revise §9
S7	Add WAN Mbps graph series on FE before next OTA wide release
S8	Staging test matrix: X-Debug-Country: ZM for geo deny paths

19. Changelog

Date	Change
------	--------

2026-06-12	Initial DR from draft 20100612_1_scaling . txt; options matrix; task graph
2026-06-12	Rev. 2 — bandwidth (§8b), global/GEO (§8c), F10, task graph + open items

Related docs

- SPEC — Platform scaling
- INFRA.md
- OPEN_TASKS_GRAPH.md
- 20260608_ALPHA_PRIORITIES.md