

URL shortener service — design review (DR)

URL shortener service — design review

Table of contents

1. Executive summary
2. Resolved open questions
3. Authentication and trust
 - 3.1 Bearer token
 - 3.2 Signer
 - 3.3 Shared session (operators)
 - 3.4 Trust flow (normative)
4. Slug algorithm
 - 4.1 Decision: SHA-256 truncated
 - 4.2 Collision policy
 - 4.3 Draft example (preserved for comparison)
5. API and flow corrections
6. Data store
7. Deployment and infra
8. QR codes
9. Security model
10. Integration with androidcast
11. Estimates
12. Task dependency graph
13. Suggestions (PO / dev discretion)
14. Remaining open items
15. Changelog

Source: 20100611_3_url_shortener.md — section links work in PDF viewers that support internal anchors.

Document type: DR (Design Review)
Source draft: docs/drafts/20100611_3_url_shortener.txt
Specification: docs/specs/20100611_3_url_shortener.md
Date: 2026-06-11
Status: Reviewed — ready for post-alpha implementation planning
Severity: Medium

1. Executive summary

The draft defines a **standalone URL shortener** at <https://s.f0xx.org>, separate from the crash-reporter BE app, with bearer/signer trust, TTL/permalinks, QR support, and a JSON shorten API.

DR outcome:

Area	Decision
Priority	Post-alpha ; medium severity; may move up if BE load stays low
Slug hash	SHA-256 truncated (12 hex) — reject draft interleave + Java hashCode
Auth	Bearer tokens for API; signer for permanent TTL; operator session RBAC only for token admin UI
Placement	backend/url-shortener/ scaffold in repo; dedicated nginx vhost; opt-in deploy
Admin UI	Hub Short links card (not separate admin skin) — SPEC §20
Infra	Git remote + FE TLS live ; nginx snippets in repo
Stack	PHP + nginx + MariaDB per draft
OpenAPI	Required in v1

2. Resolved open questions

Questions from draft §“TBD and open questions list” and inline TBD markers.

#	Draft question / gap	DR answer
Q1	Token signing, trust models, sessions, security	See §3 and §9. Machine clients use bearer (+ optional signer). Humans use existing BE session only to mint bearers. Optional HMAC for GET left as suggestion .
Q2	... TBD — open part	Filled in SPEC: audit log, rate limits, SSRF blocklist, normalization rules, collision policy, admin tables, smoke tests.
Q3	Database *... TBD columns	Full schema in SPEC §9: origin_hash, expires_at, access_count, FK tables for bearer/signer.
Q4	Java hash vs SHA-256	SHA-256 (see §4). Java hash is 32-bit, non-cryptographic, JVM-stable but poor for URL slugs.
Q5	Interleaved URL+TOKEN string	Rejected for production. Deterministic SHA256(bearer_id ■ normalized_url) instead. Draft interleave kept as historical note only.
Q6	Authentication: “app-wide shared session, project-wise token”	Split roles : curl/services → project bearer token ; console operators → shared ac_crash_sess to create/revoke tokens (same RBAC as crashes/tickets). Shorten API does not accept PHP session cookies.
Q7	Submodule remote	Done — git://f0xx.org/androidcast_project/url-shortener; pinned at backend/url-shortener/ submodule (next).
Q8	OTA-style “invalid URL” on unrelated services	N/A here; shorten service validates URL scheme/host and returns INVALID_URL.
Q9	Step 9.1.1 / 9.1.2 trusted sub-capabilities	can_temporary on bearer; can_permanent on bearer and valid signer required for ttl=0.
Q10	Cached hit when TTL “exceeds given limits”	If stored link expired → error TTL_EXPIRED, not silent renew. If valid → return same slug with remaining ttl in response (draft step 8).
Q11	Estimations (draft empty)	See §11.

3. Authentication and trust

3.1 Bearer token

- Opaque string (≥32 chars random), stored as **SHA-256 hash** in bearer_tokens.
- trusted=1 required to **create** new links (draft steps 6–7).
- **Demo token**: single trusted=0 token that can only resolve existing slugs in docs/examples — cannot create (403).

3.2 Signer

- Separate long-lived secret for **elevated** operations.
- Required when ttl=0 (permanent permalink).
- Validated against signer_keys where can_permanent=1 and not revoked.

3.3 Shared session (operators)

- **Frozen**: hub **Short links** card (SPEC §20) — same nav shell as Remote access / Builder.
- Actions: create bearer, revoke bearer, list links for bearer, purge expired.
- Uses same RBAC roles as REMOTE_ACCESS_IMPL pattern (admin, operator).

3.4 Trust flow (normative)

Matches draft low-level flow with SPEC clarifications:

```
Request → bearer present? ■no→ 401 MISSING_BEARER
        → lookup (bearer, url)
            → hit & valid TTL → 200 cached (remaining ttl)
            → hit & expired → 409 TTL_EXPIRED
        → miss → bearer trusted? ■no→ 403 ACCESS_DENIED
            → ttl=0 → signer ok? ■no→ 403 SIGNER_REQUIRED
            → compute slug, insert, return 200
```

4. Slug algorithm

4.1 Decision: SHA-256 truncated

Chosen:

```
slug = hex(SHA256(bearer_id + "\x1f" + normalized_url))[0:12]
```

Criterion	Java hashCode + interleave (draft)	SHA-256 truncated (chosen)
Collision resistance	Poor (~32 bits)	~48 bits with 12 hex chars
Determinism	JVM-specified but weak	Standard crypto
Reversibility	No (good)	No (good)
Debuggability	Opaque interleave	Simple test vectors

4.2 Collision policy

- Unique index on slug.
- On duplicate insert: append -1...-f (4 chars max) **or** re-hash with : \x1e{counter} appended to material.
- Log collision to audit_log at WARN level.

4.3 Draft example (preserved for comparison)

Draft interleaved https://t.co + token → http5bpsf:4/3/3td.2c3oa → Java hash or SHA slice 4623b436f.

Do not implement interleave — kept only to show original intent (per-token namespacing). Namespacing is achieved via bearer_id in hash input instead.

5. API and flow corrections

Draft item	Correction
GET https://s.f0xx.org?url=...	Use /api/v1/shorten path so root /{slug} stays redirect-only
Response ttl as string	Keep string in JSON for draft compatibility; document integer seconds
signer on GET	Optional query param; omit from success body (security)
Fail body echoes bearer	Truncate in response to first 6 + ... in production logs; full echo in JSON only for client debugging (SPEC)

Routing suggestion: nginx sends /api/ → PHP front controller; slug paths match regex (quoted on BE: location ~ "^[a-f0-9]{6,16}\$") → redirect handler.

6. Data store

Option	Pros	Cons	DR pick
A. Same MariaDB instance, new schema url_shortener	Simple ops, one backup	Shared failure domain	Recommended v1
B. Separate MariaDB database	Blast radius isolation	More admin	If PO wants hard isolation

C. SQLite on BE	Zero MariaDB migration	Poor fit for BE pattern	Reject
-----------------	------------------------	-------------------------	---------------

Retention cron: scripts/purge_url_shortener.php (mirror purge_remote_access.php style) — delete expired rows + audit >90d.

7. Deployment and infra

Per INFRA.md topology:

Internet → FE (s.f0xx.org TLS) → proxy_pass BE :80 → url-shortener php-fpm

Step	Owner	Notes
DNS s.f0xx.org → FE	PO / infra	Done (with TLS)
FE TLS cert	PO / infra	Done — /etc/letsencrypt/live/s.f0xx.org/
FE nginx server_name s.f0xx.org	DEV	nginx.fe-s.f0xx.org.snippet
BE vhost fragment	DEV	nginx.be-url-shortener.fragment
Git remote	PO	Done — git://f0xx.org/androidcast_project/url-shortener
Licenses	DEV	Add QR library to BE license inventory when impl starts

Safe deploy: nginx fragment only; no topology changes to router/FE↔BE path. Validate with nginx -t before reload.

Not alpha-blocking: omit from validate_be_services.sh until first deploy.

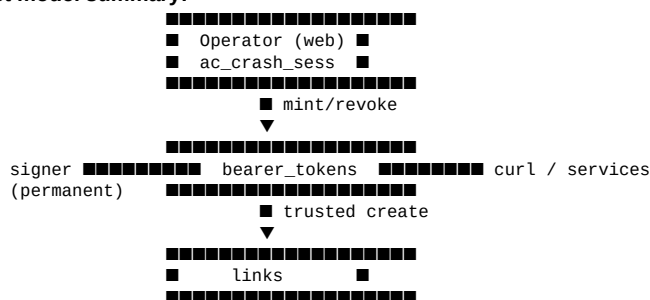
8. QR codes

Decision	Detail
Must-have	Yes (draft)
Format	PNG via /api/v1/qr/{slug}.png
Library	endroid/qr-code (MIT) — suggestion ; PO approves license addition
Inline in JSON	Optional qr URL field in shorten success — suggestion for v1.1

9. Security model

Threat	Mitigation
Token leak	Revoke bearer; rotate signer; rate limit
SSRF via shorten	Block private/metadata IPs at validation
Slug enumeration	48+ bits entropy; rate limit redirects (suggestion)
Open redirect abuse	Only whitelisted bearers create links
Session fixation	N/A — API is stateless

Trust model summary:



10. Integration with androidcast

Consumer	Use case	When
BE remote-access UI	Short "open session" links for operators	Post-alpha
OTA / docs	Stable short links in printed QR	Post-alpha
Android app	None v1 — use HTTPS from BE	Deferred
Crash reporter	No embed — standalone service	Normative

Submodule lives beside examples/crash_reporter/backend/ conceptually but **not** inside it.

11. Estimates

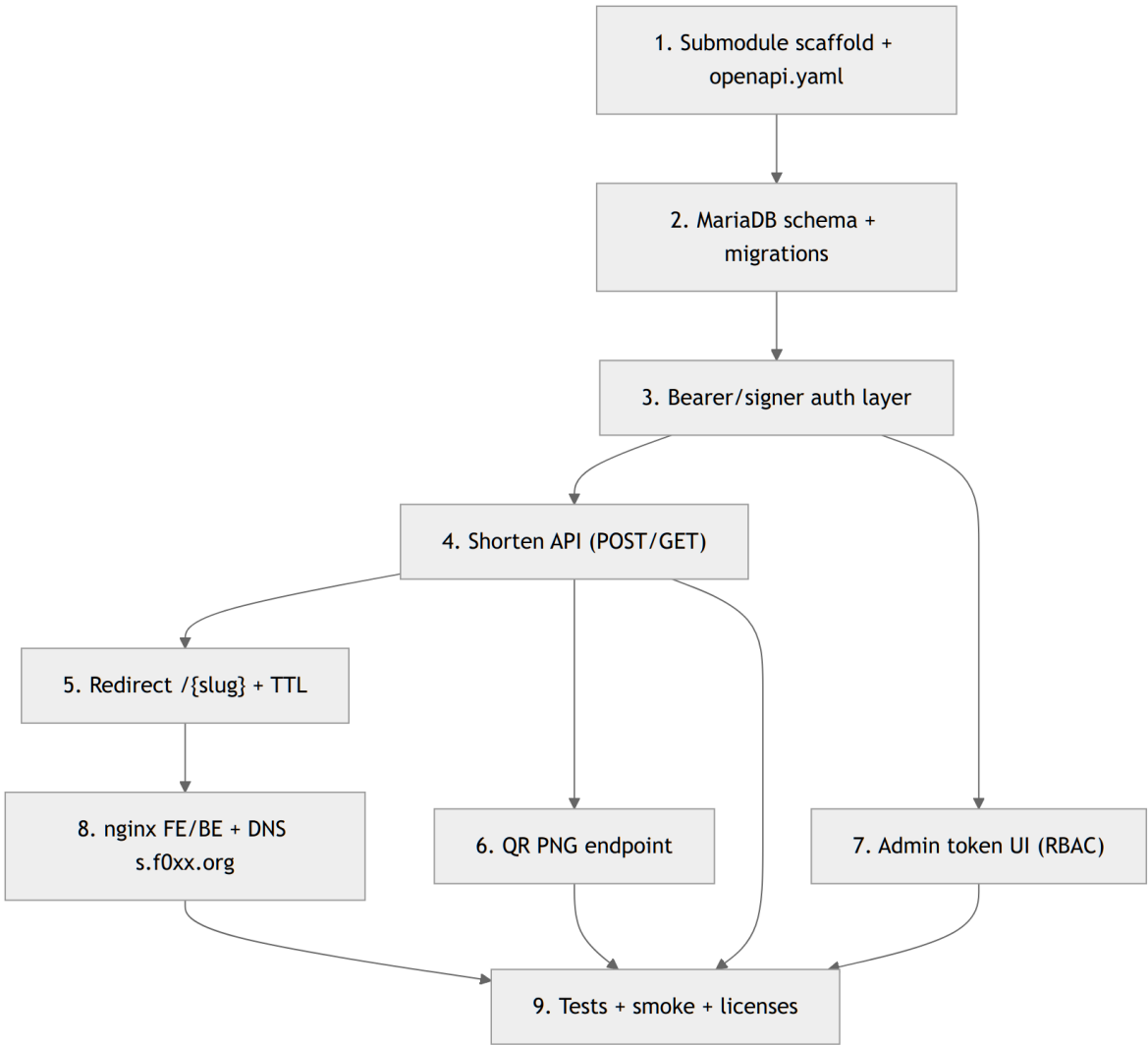
Engineering person-days (single DEV, familiar with BE stack):

Task	Days
Repo scaffold + OpenAPI + DB migrations	1.5
Shorten API + auth + TTL logic	2
Redirect handler + permalink cache	1
QR endpoint + license	0.5
Admin token UI (minimal)	1.5
nginx FE/BE fragments + deploy doc	1
PHPUnit + smoke script	1
Total	~8.5 days

Severity: Medium — no alpha path dependency.
Reorder triggers (PO): high demand for shareable debug links; BE CPU <50%; remote-access operator UX needs short URLs.

12. Task dependency graph

Legend: **1** = highest priority within this feature.



#	Task	Blocker	Assignee
1	Submodule + OpenAPI	—	DEV
2	Schema	1	DEV
3	Auth	2	DEV
4	Shorten API	3	DEV
5	Redirect	4	DEV
6	QR	4	DEV
7	Admin UI	3	DEV
8	Infra deploy	5, PO DNS	DEV + PO
9	Tests/licenses	4–7	DEV

13. Suggestions (PO / dev discretion)

Items intentionally **not** frozen in SPEC — recommended but optional:

1. **HMAC signed GET** — `&sig=HMAC-SHA256(url|ttl, bearer_secret)` to prevent query tampering on shared links; skip for v1 if all clients use POST.
2. **Console integration** — **Done (spec)**: hub **Short links** card (SPEC §20).
3. **Response field** `qr` — include ready-made QR URL in shorten JSON.
4. **Click analytics** — increment `access_count` only (already in schema); defer graphs to Grafana/nginx logs.
5. **Custom slug** — `&alias=` for operators; increases collision/abuse risk — post-v1.
6. **Separate PHP-FPM pool** — if redirect traffic is high; share pool initially.
7. **FE rate limit** — `limit_req` on `/api/v1/shorten` at nginx.
8. **Draft demo token** — ship `examples/url-shortener/demo_curl.sh` with documented read-only token for CI smoke.
9. **PDF export** — run `python3 docs/_pdf_build/md_to_pdf.py` on SPEC+DR when PO wants signed-off copies.
10. **Atomic commits on next** — per `bottomline_reminder.txt`: one commit per task row above for easy revert.

14. Remaining open items

Requires PO / infra input (hard blockers for **prod** only):

ID	Item	Owner	Status
O1	Git remote <code>git://f0xx.org/androidcast_project/url-shortener</code>	PO	Done
O2	DNS + TLS for <code>s.f0xx.org</code> on FE	PO / infra	Done
O3	BE docroot <code>/var/www/localhost/htdocs/apps/s/</code>	DEV + PO	Default in SPEC; confirm on first sync
O4	Approve QR library license addition	PO	Open (before QR impl)
O5	First production bearer token issuance policy	PO	Open
O6	PHP shorten/redirect handlers + hub Short links UI	DEV	Open (scaffold only)

Soft blockers cleared for BE nginx + MariaDB local dev. Hard blocker for prod traffic: **O6**.

15. Changelog

Date	Rev	Change
2026-06-11	1	Draft → SPEC + DR; resolved hash, auth, schema, API paths; task graph + estimates
2026-06-11	2	Scaffold backend/ <code>url-shortener/</code> ; <code>openapi.yaml</code> ; nginx snippets; curl appendix; O1/O2 done per draft

Source linkage: `docs/drafts/20100611_3_url_shortener.txt`