

URL shortener service — specification

URL shortener service — SPEC (post-alpha, s.f0xx.org)

Table of contents

1. Purpose
2. Scope
3. Requirements
 - 3.1 Must
 - 3.2 Should
4. Architecture
5. Authentication and trust
6. REST API
 - 6.1 Shorten — GET
 - 6.2 Shorten — POST
 - 6.3 Error response
 - 6.4 Success response
 - 6.5 QR
7. Redirect and permalink behaviour
8. Slug algorithm
9. Database schema
 - 9.1 links
 - 9.2 bearer_tokens
 - 9.3 signer_keys
 - 9.4 audit_log
10. QR codes
11. OpenAPI
12. Deployment
13. Security
14. Observability and limits
15. Tests
16. Out of scope (v1)
17. Related docs

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

Document type: SPEC

Source draft: docs/drafts/20100611_3_url_shortener.txt

Design review: docs/DRs/20100611_3_url_shortener.md

Status: Post-alpha (next after alpha); priority may shift with BE capacity/load

Severity: Medium

Public host: <https://s.f0xx.org>

Repo placement: backend/url-shortener (git submodule → `git://f0xx.org/androidcast_project/url-shortener`)

1. Purpose

Provide **short, encoded URLs on demand** for long targets. The service is a **helper** for other apps and operators (Android Cast BE console, curl scripts, future mobile share flows) — not a user-facing product on its own.

High-level flow:

1. A client (any HTTP app; Android Cast not required) requests shortening of a long URL.
2. The client authenticates with a **project bearer token** (and optionally a **signer** for elevated TTL).
3. The service validates trust, creates or returns a cached short link, and responds with JSON.
4. End users open `https://s.f0xx.org/{slug}` → HTTP redirect to the original URL (while TTL valid).

2. Scope

In scope (v1)	Out of scope (v1)
Standalone opt-in service (not inside crash-reporter PHP app)	Alpha release blocker
REST API + OpenAPI 3 document	Custom domains per tenant
PHP + nginx + MariaDB stack	Click analytics dashboard
Permalink reuse for same (bearer, url)	Link preview / OG scraper
Time-limited and permanent links (signer-gated)	Federated multi-region
QR code PNG for any short URL	Mobile SDK (HTTP only v1)

Scheduling: Post-alpha by default. PO may reorder if BE load is low and another product track needs short links urgently (see DR §Estimates).

3. Requirements

3.1 Must

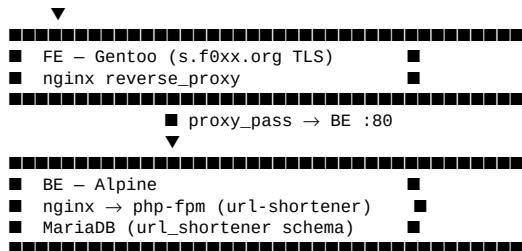
ID	Requirement
R1	Standalone service; development separate from main BE crash/tickets app
R2	REST API documented with OpenAPI 3 (<code>openapi.yaml</code> in repo)
R3	Stack: PHP 8.x , nginx , MariaDB ; OSS dependencies only (MIT/BSD/Apache)
R4	Project path: <code>backend/url-shortener</code> submodule
R5	QR code generation (PNG) for shortened URLs
R6	Permalink semantics: same bearer + same normalized URL → same slug while record valid
R7	TTL : temporary links expire; <code>t t l=0</code> permanent only for trusted signers
R8	JSON request/response contract (§6)
R9	Reverse lookup: <code>GET /{slug}</code> → 302 to stored origin (or 410 when expired)

3.2 Should

ID	Requirement
S1	Support GET (query) and POST (JSON body) shorten endpoints
S2	Rate limits per bearer token
S3	Audit log table for create/access/deny
S4	Admin token mint/revoke via BE RBAC (shared operator session) — see DR
S5	Unit tests (PHP) + curl-based smoke script in <code>scripts/</code>

4. Architecture

Client (curl, BE job, future app)
■ HTTPS shorten API



Browser / QR scan
 GET https://s.f0xx.org/{slug}
 same path → lookup → 302 Location: origin

- **Not mounted** under /app/androidcast_project/crashes/ — dedicated vhost s.f0xx.org (DR §Deployment).
- **Opt-in deploy:** absent from default BE sync until operator enables submodule + nginx fragment.

5. Authentication and trust

Normative decisions: DR §Authentication and trust.

Concept	Spec
bearer	Project-scoped API token (varchar(128)), stored hashed in DB
signer	Optional second secret proving elevated rights (permanent TTL, high quotas)
demo token	Pre-seeded read-only or low-quota token for smoke tests
App-wide session	Human operators mint/revoke bearer tokens via BE console using existing ac_crash_sess RBAC — not used for machine curl shorten calls

Trust checks (shorten):

1. bearer required → 401 if missing/unknown/revoked.
2. Lookup existing row for (bearer_id, normalized_origin):
 - If found and not expired → return cached short URL; ttl in response = **remaining seconds**.
 - If found but expired → 409 / error code TTL_EXPIRED (do not extend without new shorten).
3. If not found → require **whitelist/trusted** bearer for create.
4. If ttl=0 (permanent) → require valid **signer** with can_permanent capability.
5. If ttl>0 → bearer must have can_temporary (default for trusted tokens).

6. REST API

Base: https://s.f0xx.org/api/v1

6.1 Shorten — GET

GET /api/v1/shorten?url={URL_ENCODED}&bearer={TOKEN}&signer={SIGNER}&ttl={TTL}

Param	Required	Description
url	yes	URL-encoded long URL (max 2048 chars after decode)
bearer	yes	API token
signer	if ttl=0	Signer secret for permanent links
ttl	no	Seconds until expiry; 0 = permanent (signer required); default 86400

6.2 Shorten — POST

POST /api/v1/shorten
 Content-Type: application/json

```

{
  "url": "https://defender.net/very/long/link#with_anchors&param1=value1",
  "bearer": "4abef433d23ab5x&&83^c6",
  "signer": "optional-signer-key",
  "ttl": 3600
}
  
```

Prefer POST when URL length or sensitive query strings make GET impractical.

6.3 Error response

```

{
  "code": "ACCESS_DENIED",
  "url": "https://defender.net/...",
  "bearer": "4abef...",
  "desc": "access denied"
}
  
```

code	HTTP	Meaning
0	200	Success (success body below)
ACCESS_DENIED	403	Bearer not trusted / revoked
MISSING_BEARER	401	No bearer
INVALID_URL	400	Malformed or disallowed scheme
TTL_EXPIRED	409	Cached link expired
SIGNER_REQUIRED	403	t t l=0 without valid signer
RATE_LIMITED	429	Quota exceeded
INTERNAL	500	Server error

6.4 Success response

```
{
  "code": "0",
  "url": "https://defender.net/very/long/link#with_anchors&param1=value1",
  "u": "https://s.f0xx.org/a1b2c3d4e5f6",
  "bearer": "4abef433d23ab5x&&83^c6",
  "ttl": "3542"
}
```

Field	Description
url	Original URL (normalized canonical form)
u	Short URL
bearer	Echo of bearer (truncated in logs only)
ttl	Remaining seconds; "0" if permanent

Optional fields (v1.1 suggestion in DR): qr — URL to QR PNG.

6.5 QR

GET /api/v1/qr/{slug}.png?size=256

Returns image/png encoding https://s.f0xx.org/{slug}. 404 if slug unknown; 410 if expired.

7. Redirect and permalink behaviour

Request	Behaviour
GET /{slug}	302 Found → Location: origin; update accessed_at
Expired slug	410 Gone HTML or JSON if Accept: application/json
Unknown slug	404 Not Found
HEAD /{slug}	Same as GET without body

Permalink: For a given (bearer, normalized_origin), the **same slug** is returned on repeat shorten while the row is valid. New slug only after expiry + delete or explicit admin purge.

URL normalization (before hash/store):

- Trim whitespace
- Lowercase scheme and host
- Preserve path, query, fragment
- Reject non-http/https schemes

8. Slug algorithm

Normative (per DR): SHA-256 truncated — **not** Java String.hashCode or character interleaving from the draft.

```
normalized = normalize_url(url)
material   = bearer_id + "\x1f" + normalized
digest     = SHA256(material)
slug       = lowercase_hex(digest[0:6]) // 12 hex chars → ~48 bits
short_url  = "https://s.f0xx.org/" + slug
```

Collision handling: On insert conflict on slug, append 1...f suffix (max 4 chars) or re-hash with salt counter — see DR §Slug algorithm.

Reverse lookup: Primary key lookup on slug column (indexed unique). No recomputation at redirect time.

Draft note preserved in DR §Slug algorithm — interleave + Java hash rejected for production.

9. Database schema

Database: url_shortener (separate schema or DB — DR §Data store).

9.1 links

Column	Type	Notes
id	BIGINT UNSIGNED AUTO_INCREMENT	PK
bearer_id	INT UNSIGNED	FK → bearer_tokens.id

signer_id	INT UNSIGNED NULL	FK → signer_keys.id if used
slug	VARCHAR(16)	Unique, indexed
origin	TEXT	Normalized original URL
origin_hash	CHAR(64)	SHA-256 of origin (dedup aid)
ttl_seconds	INT UNSIGNED	0 = permanent
expires_at	DATETIME NULL	NULL when permanent
created_at	DATETIME	
updated_at	DATETIME	
accessed_at	DATETIME NULL	Last redirect
access_count	BIGINT UNSIGNED	Default 0

Unique index: (bearer_id, origin_hash) where expires_at IS NULL OR expires_at > NOW() — implementation via application check + partial unique strategy per MariaDB version (DR).

9.2 bearer_tokens

Column	Type	Notes
id	INT UNSIGNED AUTO_INCREMENT	
token_hash	CHAR(64)	SHA-256 of bearer secret
label	VARCHAR(128)	Operator label
trusted	TINYINT(1)	Whitelist flag
can_temporary	TINYINT(1)	Default 1 for trusted
can_permanent	TINYINT(1)	Default 0; needs signer anyway
rate_limit_per_hour	INT UNSIGNED	Default 1000
revoked_at	DATETIME NULL	
created_at	DATETIME	

9.3 signer_keys

Column	Type	Notes
id	INT UNSIGNED AUTO_INCREMENT	
key_hash	CHAR(64)	SHA-256 of signer secret
label	VARCHAR(128)	
can_permanent	TINYINT(1)	
revoked_at	DATETIME NULL	

9.4 audit_log

Column	Type	Notes
id	BIGINT UNSIGNED AUTO_INCREMENT	
event	VARCHAR(32)	shorten, redirect, deny, ...
bearer_id	INT UNSIGNED NULL	
slug	VARCHAR(16) NULL	
ip	VARCHAR(45)	
detail	TEXT	
created_at	DATETIME	

10. QR codes

- Library: **endroid/qrcode** (MIT) or pure-PHP equivalent — license entry required in BE licenses file before merge.
- Default size: 256×256 PNG, ECC level M.
- Content: full short URL string (<https://s.f0xx.org/{slug}>).
- Cache: optional filesystem cache keyed by slug+size (DR suggestion).

11. OpenAPI

Ship backend/url-shortener/openapi.yaml covering:

- POST /api/v1/shorten
- GET /api/v1/shorten (documented; mark POST preferred)
- GET /api/v1/qr/{slug}.png
- GET /{slug} redirect (document as GET /r/{slug} if internal router separates API — DR \$Routing)

Serve docs at GET /api/v1/docs (Swagger UI static bundle or Redoc) — opt-in on staging only.

12. Deployment

Item	Value
Submodule path	backend/url-shortener
Remote	git://f0xx.org/androidcast_project/url-shortener (when available)
BE docroot	/var/www/localhost/htdocs/apps/s/ (suggested — confirm in DR)
nginx	New fragment url-shortener.conf; FE server_name s.f0xx.org
PHP-FPM	Same pool as other BE apps or dedicated pool if load warrants
MariaDB	Schema url_shortener
Enable	Operator copies submodule, runs migration, includes nginx fragment, reload

Not required for alpha BE validation script `validate_be_services.sh` until service is enabled.

13. Security

Topic	Rule
Secrets at rest	Store only hashes of bearer/signer
TLS	Required on public FE
SSRF	Block shorten targets to 127.0.0.0/8, 10.0.0.0/8, 169.254.0.0/16, metadata hosts
Open redirect	Only stored origins issued by trusted bearers
Logging	Never log full bearer/signer; truncate to 6 chars
CSRF	N/A for token-based API; admin UI uses session + CSRF token

Optional HMAC request signing for GET — DR §Suggestions.

14. Observability and limits

Limit	Default
Max URL length	2048
Max slug length	16
Rate limit	1000 shorten/hour per bearer
Audit retention	90 days (cron purge script)

Metrics (suggestion): redirect count, 4xx/5xx rate — post-v1 Prometheus or nginx logs.

15. Tests

Layer	Location
PHP unit	backend/url-shortener/tests/ — normalize, slug, TTL math, auth
Smoke	backend/url-shortener/scripts/smoke_shorten.sh
BE validation	Optional line in <code>validate_be_services.sh</code> when deployed

16. Out of scope (v1)

- Link editing after creation
- Custom slugs chosen by user
- Analytics UI
- Integration inside Android APK (use HTTP from BE or scripts first)

17. Related docs

- DR — design review
- INFRA.md — FE → BE proxy
- 20260610_SERVICES.md — service catalog (add row when live)
- ROADMAP.md — post-alpha track