Bilko Authentication -- Entra External ID (CIAM)

Overview

Bilko uses **Microsoft Entra External ID (CIAM)** as its sole identity provider. Entra authenticates users; Bilko authorises them. Roles and permissions live exclusively in the Bilko database — no role claims are read from Entra tokens.

Decision anchor (ADR): "Entra authenticates, Bilko authorises; single-role v1; multi-org deferred (MC #103089)."

Stage status: Live on stage (branch stack WP1–WP4, feat/rbac-wp4-retire-legacy-auth commit 3ac1388). Production cutover pending consolidated PR.

Tenant Configuration

Field	Value
Tenant ID	`20bb17de-9be5-4143-a7e5-8c1ddae6a064`
Display name	Bilko CIAM
Domain	`bilkociam.onmicrosoft.com`
Type	Entra External ID (CIAM), EU data residency, Norway
Billing	MAU — free tier from 2026-06-07
Authority (MSAL)	`https://bilkociam.ciamlogin.com/20bb17de-9be5-4143-a7e5-8c1ddae6a064`
Issuer (exact, from OIDC discovery)	`https://20bb17de-9be5-4143-a7e5-8c1ddae6a064.ciamlogin.com/20bb17de-9be5-4143-a7e5-8c1ddae6a064/v2.0`
JWKS URI	`https://bilkociam.ciamlogin.com/20bb17de-9be5-4143-a7e5-8c1ddae6a064/discovery/v2.0/keys`
OIDC discovery	`https://bilkociam.ciamlogin.com/20bb17de-9be5-4143-a7e5-8c1ddae6a064/v2.0/.well-known/openid-configuration`

Issuer note: OIDC discovery returns the issuer with the tenant-ID subdomain (`20bb17de-...ciamlogin.com`), NOT the named subdomain (`bilkociam.ciamlogin.com`). The Kotlin backend

`ENTRA_EXTERNAL_ID_ISSUER` env var MUST match the discovery value exactly. See evidence: `/tmp/evidence-103076/phase0-config.md`.

App Registrations

App	Client ID	Flow	Notes
Bilko API (resource)	`fe39e0f5-513e-40af-93f0-c3ee624df56c`	Exposes scope	Scope: `access_as_user`; full scope string: `api://fe39e0f5-513e-40af-93f0-c3ee624df56c/access_as_user`; audience for token validation = client ID; no client secret (resource app)
Bilko Web SPA	`c2902239-ea63-41bd-8619-6cf096d7d45a`	PKCE auth code (SPA)	Redirects: localhost:3000, stage Cloud Run URL, bilko-demo.alai.no, app.bilko.cloud, app.bilko.io, app.bilko.company (+ /auth/callback variants); no client secret
Bilko Mobile (native)	`916bb9f3-658d-4729-b5a0-64b1f157c8c2`	PKCE auth code (native/public)	Redirect: `com.alai.bilko://auth`, `msauth.com.alai.bilko://auth`; `isFallbackPublicClient=true`; Expo Go workaround documented in Phase 0 config

Secrets (none exist — all public clients). Non-secret configuration is stored in GCP Secret Manager (`bilko-entra-issuer`, `bilko-entra-audience`, `bilko-entra-jwks-url`) and Bitwarden ("Bilko CIAM Tenant Config").

Token Claims

- `oid` — object ID, immutable cross-app anchor; **mandatory identity key** (built-in in CIAM, always present)
- `sub` — pairwise pseudonymous per app; NOT used as identity anchor (changes on app re-registration). The backend logs a warning when `sub != oid` (expected in CIAM) and uses `oid` exclusively. Confirmed live: E2E test showed `sub=053nt0lk` vs `oid=3b53a25a`.
- `email` / `preferred_username` — informational only; mutable; NOT used for identity resolution in JWT claims issued by Bilko
- `name`, `family_name`, `given_name` — optional claims

Bilko JWTs issued after exchange use the internal Bilko user UUID as the subject claim — email is not the identity anchor in issued tokens.

Authentication Flow — Web (MSAL Direct Bearer)

1. Browser opens login page → MSAL browser (`@azure/msal-browser` + `@azure/msal-react`) initiates PKCE auth code flow
2. Redirect to `bilkociam.ciamlogin.com` → user authenticates (email/password or social) → Entra issues auth code
3. MSAL exchanges code for tokens (PKCE, in memory — NOT localStorage)
4. MSAL acquires access token for scope `api://fe39e0f5.../access_as_user`
5. Web sends `Authorization: Bearer <entra-access-token>` to Kotlin API
6. Kotlin `EntraExternalIdService.verifyIdToken()` validates: RS256 signature via live JWKS, issuer exact match, audience = `fe39e0f5...`, `oid` claim present, JWKS URL domain-pinned to `ciamlogin.com`/`microsoftonline.com`
7. JIT provisioning or email-match link (see JIT section below) → Bilko session returned
8. Session cookie (`SameSite=Lax`, httpOnly) established; subsequent requests use Bilko refresh token
9. Sign-out calls Entra logout endpoint to invalidate Entra session + clears local cookie

Authentication Flow — Mobile (Token Exchange)

1. Expo native app initiates PKCE via `expo-auth-session` / `useEntraAuthRequest`
2. Redirect to Entra → auth code returned to `com.alai.bilko://auth`
3. MSAL exchanges code; **id_token** (not access_token) sent to `POST /auth/entra/session`
4. Kotlin backend verifies id_token, runs JIT provisioning or link, returns `{ accessToken, refreshToken }`
5. Tokens stored in SecureStore; TTL aligns with Bilko 7-day refresh token window

JIT Provisioning + Identity Linking

Implemented in `AuthService.createSessionFromEntraIdToken()` (SERIALIZABLE transaction — martin-kleppmann race-prevention mandate):

1. Lookup `entra_external_identities` by `issuer + oid`. If found: return session.
2. If not found: email-match lookup in `users` (case-normalised, lowercase both sides). If unique match and `email_verified`: insert `entra_external_identities` row, log audit event

`entra_jit_link`, return session.
3. If no match: call `UserProvisioningService.provisionNewUserForEntra()` — creates a new org + user with role `viewer` + inserts `entra_external_identities` row. New user must be promoted by an admin.

Design dissent on record (martin-kleppmann + bruce-momjian): email-match JIT is risky if email is mutable or duplicate. Pre-provision by OID (via admin invite or MS Graph export script) is the safer path. JIT email-match is constrained with a serializable transaction as a partial mitigation. The pre-provision script path (D8 in MC #103075) is the recommended path for production migration.

JWKS Cache + Key Rotation

`EntraExternalIdService` maintains a time-bound JWKS key cache:

- TTL: 12 hours per key (stored as `Pair<RSAPublicKey, Instant>`; evicted on read if age > 12h)
- On `kid` miss: force re-fetch regardless of other cached keys
- JWKS URL domain-pinned: must match `^https://([a-z0-9-]+\.)*ciamlogin\.com/` or `^https://login\.microsoftonline\.com/`
- Startup fail-closed: if `ENTRA_EXTERNAL_ID_ISSUER` set but any config absent or URL fails domain assertion → `IllegalStateException` at Ktor module init (not lazy 503)
- JWKS verification: live E2E test confirmed 6 RSA keys, all `kty=RSA`, TLS valid 2026-11-22

Refresh Token Revocation (Known Limitation)

Bilko refresh tokens are 7-day HMAC validated locally. A disabled Entra account remains valid in Bilko for up to 7 days. **Open CEO/Securion decision (OC#4 from MC #103075):**

- Option A (not yet implemented): revalidate Entra account status on every refresh (~50ms latency)
- Option B (current default): 7-day window; immediate revocation requires an admin to also disable the user in the Bilko DB. Documented as a risk-acceptance decision in `AuthService.kt` (code comment references MC #103075).

Legacy Email/Password — RETIRED (410 Gone)

As of branch `feat/rbac-wp4-retire-legacy-auth` (commit 3ac1388), the following endpoints return `HTTP 410 Gone` with body `{"code":"ENDPOINT_RETIRED"}`:

- `POST /auth/register`
- `POST /auth/login`
- `POST /auth/forgot-password`
- `GET /auth/reset-password`
- `POST /auth/reset-password`

Kept active: `POST /auth/entra/session`, `POST /auth/refresh`, `POST /auth/mobile/refresh`, `POST /auth/2fa/challenge`.

Web login page: email/password form removed; Entra primary CTA only. Self-serve register page removed; shows "contact your admin" message. Forgot/reset password redirects to Entra SSPR portal.

Break-glass (first-admin bootstrap): run `./gradlew :apps:api:bootStrapAdmin` with `BOOTSTRAP_ADMIN_EMAIL` + `BOOTSTRAP_ADMIN_PASSWORD` env vars. Calls `AuthService.register()` directly; no HTTP endpoint exposed.

Phase 0–4 Deployment Facts

Phase / WP	Scope	Branch	Status
Phase 0 (MC #103076)	CIAM tenant provisioning, 3 app registrations, JWKS verification	FlowForge standalone	DONE — stage live
WP1 (MC #103141)	RBAC permissions catalog V67, PermissionService, BilkoPrincipal, requirePermission, 204 matrix tests	feat/rbac-wp1-permissions-catalog	DONE — Proveo PARTIAL (integration test fix applied post-verification)
WP2 (MC #103142)	JIT provisioning V68, UserProvisioningService, admin/invite API, role-assign endpoint	feat/rbac-wp2-user-provisioning	DONE — Proveo PASS
WP3 (MC #103143)	Web: Entra primary CTA, register retired, forgot/reset SSPR, RBAC admin UI	feat/rbac-wp3-web-entra-ui	DONE — Proveo PASS
WP4 (MC #103144)	Retire legacy endpoints (410), web login Entra-only, break-glass documented	feat/rbac-wp4-retire-legacy-auth	DONE — Proveo PASS
WP5 (MC #103145)	E2E: live CIAM token, OID anchor, JIT provision, RBAC enforcement, invalid token rejection	feat/rbac-wp3-web-entra-ui	DONE — PASS (browser MSAL flow deferred to Proveo pre-prod)

Evidence bundles: `/tmp/evidence-103141` through `/tmp/evidence-103145`, `/tmp/evidence-103076/phase0-config.md`.

Revision #1
Created 2026-06-08 07:39:57 UTC by John
Updated 2026-06-08 07:39:57 UTC by John