**Runbook: Incident (Detect, Triage, Contain, Recover, Learn)**

**Project:** MCPX-KendoBridge Admin Portal  
**Runbook ID:** TJ-MCPX-RB-03  
**Version:** 1.1.0  
**Status:** Draft  
**Date:** 2025-09-27  
**Owner:** Release Manager / DevSecOps (Technijian)  
**Confidentiality:** Technijian Internal

**Document Control**

| **Version** | **Date** | **Author** | **Change Summary** | **Status** |
| --- | --- | --- | --- | --- |
| 1.0.0 | 2025-09-26 | SRE | Initial incident handling procedure | Draft |
| 1.1.0 | 2025-09-27 | SRE | Added SSE TTFB/heartbeat, parity gate checks, envelope codes | Draft |

**Approvals**

| **Role/Team** | **Name** | **Signature/Date** | **Comment** |
| --- | --- | --- | --- |
| Director of Software Eng. |  |  |  |
| Systems Architect |  |  |  |
| Security & Compliance |  |  |  |
| QA Lead |  |  |  |

**1) Purpose**

Provide a **repeatable, auditable** process to handle incidents impacting the Admin Portal. This covers **detection, triage, containment, recovery, communications, and postmortem**, aligned with our error envelope, SSE budgets, RTM↔Prod parity gates, and RBAC/CORS/rate-limit policies.

**2) Scope**

* Environments: **Alpha → Beta → RTM → Prod** (Prod has stricter thresholds)
* Systems: Admin API (.NET 8), Web (React + KendoReact Fluent 2), SQL Server (SP-only), AAD/MSAL, Child MCP (STDIO), monitoring/alerts

**3) Severity Matrix**

| **Sev** | **Definition (Prod)** | **Examples / Triggers** |
| --- | --- | --- |
| **P1** | Broad outage or critical function broken; regulatory/compliance risk; customer-visible major degradation. | /ready red >5 min; 5xx >1% 5 min; p95 JSON >600ms 10 min; **SSE TTFB >200ms** median 10 min; **heartbeat gap >15s**; CORS denies from allowed origins; security exposure; parity shows **critical diffs** in RTM gating Prod |
| **P2** | Partial outage, repeated errors for a feature, performance materially degraded, limited scope. | Feature path 4xx/5xx spikes; single-node flaps; rate-limit anomalies; auth failure spikes |
| **P3** | Minor bug, cosmetic or intermittent; no SLO breach; no customer impact. | UI glitch; non-blocking warning logs |

Default to **higher** severity if in doubt; downgrade later.

**4) Roles & Responsibilities**

* **Incident Commander (IC):** Release Manager (backup: SRE on-call) — drives decisions, timeboxes steps
* **Deputy / Ops Lead:** SRE on-call — executes containment/recovery, coordinates tooling
* **Comms Lead:** PM or IC delegate — stakeholder updates (internal)
* **Scribe:** QA — timeline, evidence links, decisions
* **SMEs:** Dev Lead, Systems Architect, DBA, Security as needed

**Engagement target:** **P1** engage ≤ **15 min**, MTTR ≤ **60 min**.

**5) First 15 Minutes (Checklist)**

1. **Acknowledge alert**; open **#incident** thread with env, SHA, time.
2. **Assign IC / Deputy / Scribe**; set severity.
3. **Snapshot health:**
   * /healthz, /ready
   * **SSE**: first-event (TTFB), **heartbeat gaps**
   * Error envelope top codes (origin\_forbidden, rate\_limited, missing\_session\_id, etc.)
4. **Isolate change:** what shipped (tag, diff, migrations, flags)?
5. **Decide containment path** (Section 7) within **10 minutes** of start.
6. **Notify stakeholders** (brief: impact, actions, next update time).

**6) Rapid Triage Matrix**

| **Signal / Symptom** | **Likely Area** | **Quick Checks / Commands** |
| --- | --- | --- |
| /ready red / flapping | Dependency / config / DB | Check component tiles; verify RTM→Prod parity (RTM); recent config changes (audit) |
| 5xx spike; envelope feature\_disabled | Feature flag regression | Ensure legacy endpoints remain OFF; audit recent toggles |
| 4xx spike; envelope origin\_forbidden | **CORS allow-list** | Verify intended origins exist for env; revert CORS edit |
| SSE slow start (TTFB) or no progress; gap >15s | Child MCP / load / network | Restart child; check queue depth; run **k6** smoke |
| Auth failures spike | AAD/MSAL or RBAC drift | Login synthetics; token validation; role claims present |
| Latency p95 breach | DB hot path / scale | Query Store; SP timings; consider scale out |
| Rate-limit surge / 429s | Abuse or bug | Inspect client/IP; temporarily raise/targeted block |

**7) Containment Options (choose one or combine)**

* **Rollback to LKG** (preferred for P1) — Run **Runbook: Rollback (RB-02)**.
* **Feature flag flip** — Disable risky feature (e.g., EnableLegacyHttpSse), or temporarily relax a non-critical check.
* **Traffic shaping** — Rate-limit class tightened for abusive IP/principal; circuit breaker on failing dependency.
* **Scale-out** — Add instances; increase DB resources (temporary) — see **Runbook: Scale-Out (RB-05)**.
* **Hot fix config** — Revert config keys (SP-only), re-run /ready.

**Timebox:** If containment does not re-stabilize in **15 min**, **rollback**.

**8) Recovery Steps (after containment)**

1. **Verify green:** /healthz, /ready, key pages.
2. **Run k6 SSE smoke:** ensure **TTFB ≤ 200 ms**, **heartbeat ≤ 10 s**.
3. **Contract smoke:** a few critical endpoints and a UI path (login → dashboard → jobs).
4. **Monitor 30–60 min:** p95, 5xx, SSE metrics; no alert flaps.

**9) Evidence to Capture (for release & postmortem)**

* Timeline (UTC) with actors/decisions
* Screenshots: Executive, **SSE Health**, Readiness & Parity panels
* k6 results, error envelope samples (with requestId)
* Config/Flag **AuditEvent** diffs (who/what/when/before→after)
* If DB changes: migration IDs, **SP signature snapshot** diffs
* Slack thread link, incident ticket, rollback/runbook execution IDs

Retention: **≥ 1 year** with release tag.

**10) Communications Templates**

**Incident declared (internal):**

🚨 **P{1|2} INCIDENT** — {env} — {summary}. Impact: {api/web/jobs}. Actions underway: {containment}. Next update at {time}.

**Stabilized:**

✅ Stabilized — {env}. Health/Ready ✅. SSE TTFB {ms}, heartbeat {s}. Monitoring for {30–60} min. RCA to follow.

**Rollback executed:**

↩️ Rolled back to {tag\_prev} due to {reason}. Metrics recovered. Evidence attached. RCA scheduled.

**11) Postmortem (within 3 business days)**

**Template (attach to incident ticket):**

* **Summary & Impact** (duration, affected users)
* **Timeline (UTC)** (detection → decision → containment → recovery)
* **Root Cause** (code/config/data/infra)
* **Contributing Factors** (process/tooling)
* **What Worked / What Didn’t**
* **Action Items** (owner, due date)
  + Tests to add/update (E2E/contract/perf/a11y)
  + Monitoring/alert improvements
  + Runbook/doc updates (FR/NFR/CI/CD/Monitoring)
  + ADR or policy changes (rate-limit, CORS workflow)
* **Evidence Links** (release artifacts, dashboards, logs, audits)

No-blame, learning-first culture. Track actions to closure.

**12) Quick Commands & Queries (reference)**

**SSE Smoke:**

k6 run tests/perf/k6\_sse\_ttfb.js -e BASE\_URL=https://{env-host}/api -e TOKEN={bearer}

**Health & Config:**

curl -fsS -H "Authorization: Bearer $TOKEN" https://{env-host}/api/healthz

curl -fsS -H "Authorization: Bearer $TOKEN" https://{env-host}/api/config/effective

**Recent Flag/Config Changes (SQL, example):**

SELECT TOP 20 \* FROM dbo.AuditEvent

WHERE CreatedAt > DATEADD(hour,-12,sysutcdatetime())

ORDER BY CreatedAt DESC;

**Disable risky legacy endpoints (feature flag):**

EXEC dbo.sp\_Feature\_Set @FlagKey=N'EnableLegacyHttpSse', @IsEnabled=0, @Scope=N'{Alpha|Beta|RTM|Prod}',

@TargetRole=NULL, @Description=N'Incident containment', @Actor=N'incident-bot', @RequestId=N'{reqId}';

**13) Acceptance Criteria (Incident Runbook)**

* IC/roles assigned within **15 min**; severity set; comms started.
* Containment chosen within **10–15 min**; rollback used when in doubt.
* Recovery verified (health/ready, perf, SSE metrics) and monitored **30–60 min**.
* Evidence compiled; postmortem scheduled; action items captured with owners/dates.

**14) Related Runbooks & Docs**

* **RB-01 — Deploy**
* **RB-02 — Rollback**
* **RB-04 — Rotate Telerik/Kendo License**
* **RB-05 — Scale-Out** *(next)*
* **04 System Context**, **05 NFRs**, **09 Test Strategy**, **10 CI/CD**, **11 Monitoring**, **12 Evidence Pack**, **13 Compliance**, **17 Threat Model**

**End of Runbook — TJ-MCPX-RB-03 v1.1.0**