Problem Statement 13
RPKI Observatory

Reference: IETF RFCs 6480 (RPKI arch), 6482 (ROA), 6487 (cert profile), 6486 (manifests), 6811 (ROV states), 8182 (RRDP), 8210 (RTR), 8630 (TAL), 7115 (operational guidance), and 9319 (ROV/maxLength clarifications).

Objective

Build an **RPKI Observatory** that continuously ingests RPKI and BGP data, computes **ROV (Route Origin Validation)** status (Valid / Invalid / NotFound), detects anomalies (e.g., **ROA misconfigs, leaks, hijacks, collateral damage** under partial ROV), and presents real-time **dashboards, alerts, and reports** for operators and researchers.

Core Tasks

1. **Data Ingestion & Normalization**
    - Pull **VRPs** from RPKI repositories/validators (e.g., Routinator/octoRPKI/rpki-client) via **RRDP/rsync**; cache by TAL per RFC 8630.
    - Stream **BGP upd**ates (e.g., RIS Live / RouteViews / local GoBGP/FRR feed).
    - Normalize to a common schema: {prefix, maxLength, ASN, time, source, tal}.
2. **ROV Classification & State Tracking**
    - For each BGP announcement, compute **Valid / Invalid / NotFound** per RFC 6811 and record **reason** (AS mismatch, maxLength exceeded, no covering ROA).
    - Track **state transitions** over time (e.g., NotFound→Valid after ROA publication, Valid→Invalid after ROA change).
3. **Anomaly & Collateral-Damage Analytics**
    - **Detect surges in Invalids**, suspicious new origins, or prefix length anomalies.
    - **Partial ROV simulation:** model diverse operator policies (drop/mark-down/ignore) to estimate **reachability impact** and potential **collateral damage** to bystanders.
4. **Propagation & Timeliness Studies**
    - Measure **ROA time-to-effect**: from ROA publish → VRP availability → observed routing change.
    - Compare **RRDP vs rsync** freshness and validator sync intervals.

5. **Observatory UI & Alerts**
    - Real-time **dashboard**: time-series of validity states, top Invalids by ASN/prefix, TAL breakdown, heat-map by region/ASN.
    - **Alerting**: threshold or rule-based (e.g., “Invalids > X for ASN Y in 5m”).
    - **Drill-down** views: show the exact VRP and BGP paths supporting each classification.

Deliverables

- **Pipeline** (containerized) that ingests data, classifies, stores, and serves APIs.
- **Web UI** with live charts + searchable incident explorer.
- **Reproducible dataset** (sample day of MRT/JSON + VRPs) and **README** with deploy/run steps.
  Short report: notable incidents found, collateral-damage estimates, ROA hygiene insights, and recommendations.

Evaluation Criteria

- **Correctness & Standards Alignment**: ROV logic matches RFC 6811; repository/TAL handling consistent with RFCs 8182/8630/6480/8210.
- **Observability & UX**: clear visuals, fast drill-downs, helpful alerts.
- **Scale & Robustness**: handles sustained update rates; resilient to validator/stream hiccups.
- **Insightfulness**: quality of anomaly detections, collateral-damage analysis, and operational recommendations.
- **Reproducibility**: clean code, docs, configs, and sample data.