SCOPING EXAMPLE · ROUTER ROLLOUT CONTROL PLANE · FEDRAMP HIGH

Software, service bus and deployment costs for 3,579 sites

Cloudflare for Government gates every change, AWS GovCloud keeps the evidence, and the site enforces. Built in Rust, so per-device SaaS fees and SaaS data holds never enter the path.

Yoonsun Lee · YSL150 LLC · ysl@ysl150.com · November 19, 2026 · Scoping example: example costs from public sources.

Scoping example, public sources only. This document scopes example pricing for 3,579 router migrations. It uses only public information: news reports, government reports and vendor list prices. It contains no confidential, customer or NDA material, and it is not a quote or a description of any system in operation. The example costs are set against two public figures for FAA air traffic control modernization: the $12.5 billion Congress appropriated in 2025, and the $50.1 billion Transportation Secretary Sean Duffy asked Congress for in November 2026, days after the nationwide ground stop. The FAA Administrator put Phase 1 at $16 billion, and GAO reports no full lifecycle cost estimate yet. Year one in this example is about $2.55 million: about 0.02% of the $12.5 billion, and about 0.005% of the $50.1 billion ask. The router upgrade is very few units: 3,648 routers in all, about $713 per site in year one, while the ask also funds up to 140 towers ($10 billion), air route control centers and the digital architecture.

$2.55M	$33,595	$2,995	$53,685
year one, under $3 million: $2.15M of gateways plus 12 months	a month for data operations, typical ($9.39 per site)	of that is the cloud backend; the rest is the cellular plan	a month avoided: per-host SaaS monitoring at $15 list

Bottom line. Year one costs about $2.55 million: $2.15 million for 3,579 site gateways and $403,000 for twelve months of data operations. Once the gateways are installed, the FedRAMP High backend for all 3,579 held sites costs about $3,000 a month. The cellular plan for those sites costs ten times that. A SaaS stack priced per device would add more than the whole data operations bill. It would also keep the evidence in a system the AI review can't reach inside the authorization boundary. The software that prevents fan-out disruptions adds $0: it is router configuration and open-source Rust (section 8).

1 Deployment

Retrofit first, then the rest. The 69 routers already installed get the gateway first, in cohorts of 1, 4, 16 and 48 sites. Each cohort is watched before the gate opens. The other 3,579 sites stay held at the gate. They come online in later cohorts only when the certification record is complete and the previous cohort was quiet.

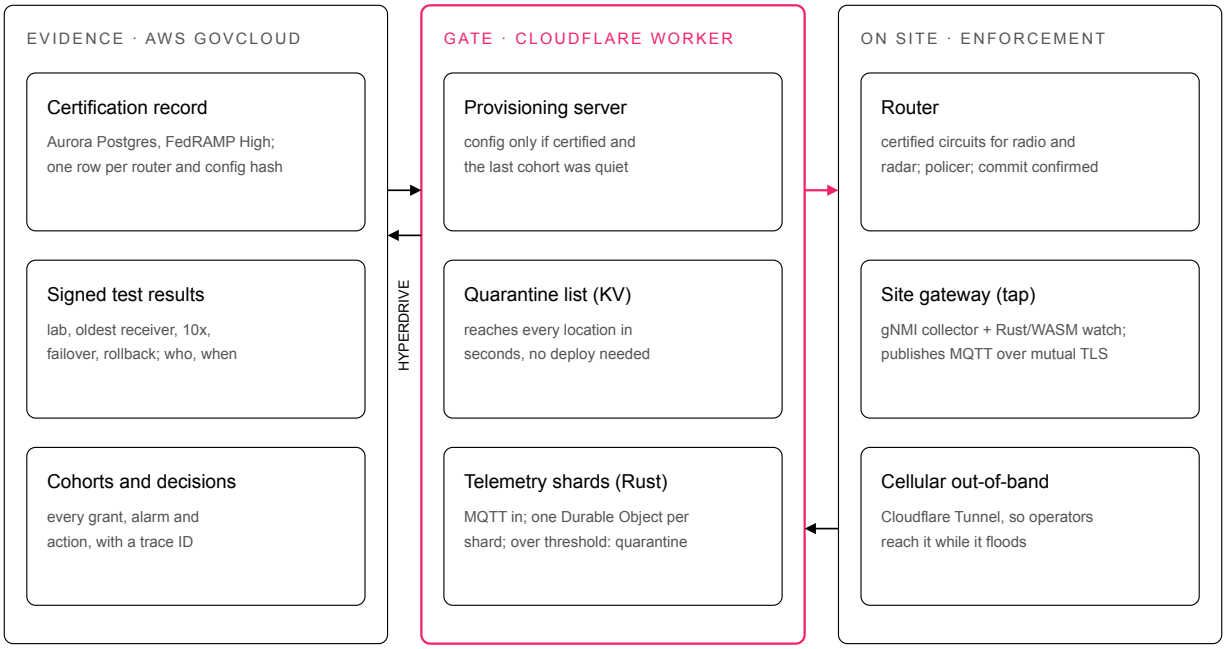
Two pre-approved actions. On site, the gateway can do two things and nothing more: quarantine the site, or load the certified rollback. Air-traffic radio and radar stay on their certified circuits and never pass through Cloudflare. This is a control plane, not a data path.

CONTROL PLANE FOR ROUTER ROLLOUT · TESTING, BOUNDARY, TRUST

Cloudflare and GovCloud decide; the site enforces

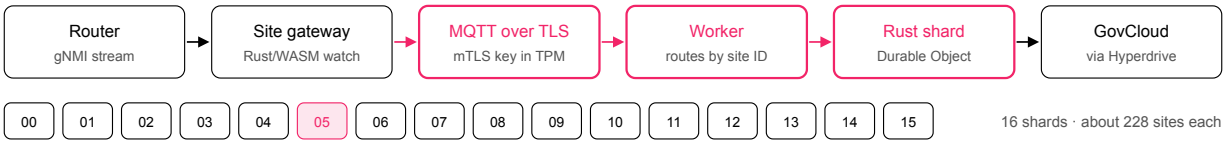
The control plane picks the configuration and keeps the evidence. Radio and radar stay on their circuits.

THREE PLACES, THREE JOBS



The Worker reads GovCloud and writes decisions back through Hyperdrive. Configuration goes out only through the gate (pink); telemetry comes back as MQTT.

TELEMETRY PATH: MQTT IN, RUST SHARDS, EVIDENCE IN GOVCLOUD



Each shard owns a fixed slice of sites, keyed by site ID. A noisy site trips only its own shard (05), which writes the quarantine to KV.

16 is an example; size shards to the message rate. Gateways send MQTT over WebSocket on 443, since Workers don't accept raw MQTT on 8883.

FORWARD DEPLOY: RETROFIT THE ROUTERS ALREADY INSTALLED, THEN THE REST



cohorts of 1, 4, 16 and 48 sites: 69 installed, each watched before the gate opens

3,648 planned in all

Hold the gate
no new installs until the retrofit cohorts pass; the Worker refuses

Two pre-approved actions
the gateway can quarantine and load the certified rollback, no more

Control plane, not data path
air-traffic radio and radar never pass through Cloudflare

Example costs. FedRAMP High: Cloudflare for Government (Aug 2026), Aurora in GovCloud; confirm Durable Objects are in scope.
Counts: 3,648 per senators citing L3Harris; more than 60 installed per the FAA; 69 from one report.

Retrofit steps and sources in the runbook section "Forward deploy: failover for the routers already installed".

files.ysl150.com/runbooks/data-layer

Control plane for the router rollout. The full-size diagram is at files.ysl150.com/runbooks/data-layer/control-plane.svg.

2 FedRAMP High service bus

Every message follows one path. Each hop runs on a FedRAMP High service, or on hardware the operator owns at the site.

1. Router to gateway. The router streams gNMI telemetry to the gateway every 10 seconds over the mirror port. Raw samples stay on the gateway's SSD for 30 days.
2. Gateway to Cloudflare. The Rust/WASM watch reduces the samples to one 4 KB MQTT batch a minute. It publishes over WebSocket on 443, using mutual TLS with a key kept in the gateway's TPM.
3. Cloudflare Access. Access checks the gateway's certificate before anything reaches the Worker.
4. Worker to Rust shard. The Worker routes each site to a Rust shard, a Durable Object keyed by site ID: 16 shards of about 228 sites each. A shard that trips writes the quarantine to KV, and every location sees it within seconds.
5. Shard to GovCloud. Through Hyperdrive, the shards write decisions and 1-minute rollups to Aurora PostgreSQL in GovCloud. Batches are archived to S3 with Object Lock, so they can't be changed.
6. AI on exceptions. A shard trip or an operator request queues a review. A Rust Lambda in GovCloud reads the evidence from Aurora, calls Claude on Bedrock in GovCloud, and writes the verdict back with a trace ID.
7. Gated updates. Signed configuration bundles in R2 go only to cohorts the Worker has opened.

HOP	SERVICE	FEDRAMP STATUS	NOTES
Gate, KV, Access	Cloudflare for Government	High	Certified Aug 6, 2026; Moderate since Dec 2022. High-impact data is processed only at US locations.
Rust shards	Cloudflare Durable Objects	Confirm	Not named in Cloudflare's published scope. Ask Cloudflare's federal team, or check the Marketplace listing FR2000863987.
Link to GovCloud	Cloudflare Hyperdrive	In scope	Added to Cloudflare's FedRAMP scope in May 2025.
Update bundles	Cloudflare R2	In scope	Named in Cloudflare's 2025 scope expansion.
Evidence	Amazon Aurora PostgreSQL, GovCloud	High	Serverless v2 is available in GovCloud (US-West).
Archive	Amazon S3, GovCloud	High	Object Lock compliance mode for write-once evidence.
AI review	Claude on Amazon Bedrock, GovCloud	High, IL4/5	The direct Anthropic API is not FedRAMP authorized, so use Bedrock.
Certificates, keys, audit	AWS Private CA, KMS, CloudTrail, GuardDuty, Security Hub, Lambda (GovCloud)	Confirm each	Check each one against AWS's FedRAMP services-in-scope list.
Data modeling, not FedGov system design	Xano	None found	Models the data and writes the mapping rules only. It isn't part of the federal system: the record is Aurora in GovCloud.

3 Software components

Every component on the gateway, in the gate and in GovCloud. Open-source licenses carry no fee. Cloud lines are billed by usage and appear in section 5.

COMPONENT	WHERE	SOFTWARE	LICENSE	JOB
Base OS	Gateway	Minimal Linux (Debian or Yocto), secure boot	Open source	Runs only the agent; no shell for remote users
Host agent	Gateway	Rust on the tokio runtime	MIT	Bounded queues and backpressure; nothing blocks the watch
Watch module	Gateway	WASM on wasmtime, signed	Apache-2.0	Thresholds and rollups; can be replaced without reflashing
gNMI collector	Gateway	Rust tonic gRPC client, OpenConfig models	MIT	Reads router telemetry from the mirror port
MQTT client	Gateway	rumqttc, MQTT 5 over WebSocket	Apache-2.0	One 4 KB batch a minute, retried with backoff
TLS	Gateway, Lambda	rustls with the aws-lc-rs FIPS provider	Apache-2.0 / ISC	FIPS 140-3 validated cryptography
Key storage	Gateway	TPM 2.0 through tss-esapi	Apache-2.0	The client key never leaves the chip
Local buffer	Gateway	SQLite on SSD	Public domain	30 days of 10-second samples; survives outages
Gate	Cloudflare	Rust Worker (workers-rs), KV	Apache-2.0 / MIT	Provisioning, quarantine list, cohort checks
MQTT bus	Cloudflare	Rust shards on Durable Objects, MQTT codec from rumqtt	Apache-2.0	Ends MQTT sessions; one shard owns a slice of sites
Evidence schema	GovCloud	PostgreSQL schema, append-only triggers, sqlx	MIT / Apache-2.0	Certification, tests, cohorts and decisions
AI review runner	GovCloud	Rust on AWS Lambda, calling Bedrock	MIT	Reads evidence and writes a verdict with a trace ID

Certificates and keys

CERTIFICATE OR KEY	ISSUED BY	LIFETIME	COST	PROTECTS
Router device identity (IDevID)	Juniper, at the factory	Life of the device	Included	Proves the router is genuine before provisioning
Gateway client certificate (mutual TLS)	AWS Private CA, short-lived mode, GovCloud	7 days, renewed automatically	$50 a month + $0.058 each	Only enrolled gateways reach the gate
Edge server certificate	Cloudflare	Managed	Included	The gateway connects only to the real gate
Hyperdrive to Aurora TLS	Amazon RDS CA	Managed	Included	Evidence encrypted in transit
Update and module signing key	AWS KMS, ECDSA P-256	Rotated every year	About $1 a month a key	Gateways reject bundles or WASM that aren't signed

CERTIFICATE OR KEY	ISSUED BY	LIFETIME	COST	PROTECTS
Operator access	Agency identity provider (PIV/CAC) through Cloudflare Access	Session	Part of the Cloudflare contract	Every operator action is tied to a person

Open-source alternative for the gateway certificates: step-ca with its key in KMS, on a small GovCloud instance, at roughly $40 a month. That puts the certificate authority inside your own approval package. AWS Private CA keeps it managed.

---

4 Rust vs. SaaS fees

At one batch a minute, the messaging bus is cheap whichever way it's built. SaaS gets expensive where it charges per device or per host.

NEED	BUILT (RUST + FEDRAMP HIGH SERVICES)	MONTHLY	SAAS OPTION	SAAS FEE BASIS	MONTHLY AT 3,579
MQTT bus	Rust shards on Durable Objects	$50	AWS IoT Core	$1 per million messages, $0.08 per million connection-minutes, $0.15 per million rules (commercial list)	about $190
Gateway monitoring	Rust/WASM watch with Aurora rollups	in Aurora	Per-host monitoring (Datadog Infrastructure)	$15 per host Pro, $23 Enterprise, annual list	$53,685 to $82,317
Register (CMDB)	Aurora PostgreSQL, GovCloud	$655	ServiceNow CMDB / ITOM	Contract pricing, not published	not public
Certificates	step-ca, self-hosted	about $40	AWS Private CA (managed)	$50 per CA + $0.058 per certificate	$1,295
AI review	Claude on Bedrock GovCloud, on demand	$340	Vendor AI add-ons (Now Assist, Bits AI and similar)	Per seat or per use, not published; runs on the vendor's model endpoints	not public

> What the comparison leaves out. SaaS fees recur every month. The Rust route replaces them with build and maintenance effort, which this proposal doesn't price. Check every SaaS option's FedRAMP level before comparing. SaaS prices here are commercial list prices; Government pricing differs.

5 Data operations, monthly

FEDERAL DATA COST · AI BACKEND · 3,579 SITES · FEDRAMP HIGH

The cloud costs less than the cellular plan

Data is sized at the gateway, updates go only to open cohorts, and AI runs only on exceptions.

LINE	FEDRAMP HIGH SERVICE	SIZED FOR 3,579 SITES	TYPICAL / MO	SURGE / MO
Site link	Cellular data plan, 250 MB site hardware, not cloud	1 MQTT batch a minute, 4 KB 173 MB per site a month	$30,600	$30,600
Gate	Cloudflare for Government Workers, KV, Durable Objects*	155M MQTT messages a month 7.7M billed requests (20:1)	$50	$60
Fan-out protection	Router configuration and open-source Rust: no license	policers, token buckets, budgets, retry limits, circuit breakers	$0	$0
Gated updates	Cloudflare R2 signed configuration bundles	5 MB × 3,579 = 18 GB a rollout shipped only to an open cohort	$1	$1
Link to GovCloud	Cloudflare Hyperdrive pooled Postgres connections	reads and writes from the Worker	incl.	incl.
Evidence	Aurora PostgreSQL Serverless v2 AWS GovCloud, Multi-AZ	5.2M rows a day 6 ACU typical, 10 in a surge	$655	$1,075
Archive	Amazon S3, GovCloud Object Lock (write once)	618 GB a month of 1-minute batches one-year retention	$145	$145
AI on demand	Claude Opus 5.5 on Bedrock AWS GovCloud	1,573 reviews a month (surge 11,237) 30k tokens in, 3k out each	$340	$2,427
Security on demand	KMS, CloudTrail, GuardDuty, Security Hub (GovCloud)	keys, audit logs, findings	$500	$900
Certificates	AWS Private CA, GovCloud short-lived mode, 7-day certificates	about 21,500 certificates a month key stays in the gateway TPM	$1,295	$1,295
Egress	GovCloud to Cloudflare Hyperdrive reads	50 GB a month (surge 150 GB)	$8	$23
Monthly total	$9.39 per site typical, $10.21 in a surge		$33,595	$36,525
Cloud backend only	without the cellular plan: $0.84 per site typical, $1.66 in a surge		$2,995	$5,925

One time: 3,579 site gateways at about $600 (Teltonika RUTX50, as in the router cost diagram)	$2.15M

Sized at the gateway
10-second gNMI stays on the site SSD for 30 days; only 1-minute batches leave

Gated updates
bundles ship only to a cohort the Worker opens; the other sites download nothing

AI only on exceptions
a shard trip or an operator request starts a review: 1% of sites a day, 10% in a surge

Prices: Cloudflare public list (the Government contract price is not public). Aurora and S3 GovCloud rates are estimates, about 1.2× commercial; Opus 5.5 on Bedrock GovCloud estimated at $4.80 / $24 per million tokens, the same 1.2× as Sonnet 5.5's listed $2.40 / $12. Security line is rough.
*Confirm Durable Objects are inside Cloudflare's FedRAMP High boundary. Excludes the ATO assessment, staff and router licenses.

Same sites as the control plane diagram: 3,579 held at the gate. Sizing: one 4 KB MQTT batch a minute per site.

files.ysl150.com/runbooks/data-layer

Monthly data operations for 3,579 sites. A typical month assumes AI reviews on 1% of sites a day; a surge assumes 10%.

PERIOD	TYPICAL	SURGE	BASIS
Monthly data operations	$33,595	$36,525	Table above
Yearly data operations	$403,140	$438,300	12 months
One time: site gateways	$2,147,400	$2,147,400	3,579 × $600
Year one total	$2,550,540	$2,585,700	Gateways + 12 months; under $3 million
Five-year total	$4,163,100	$4,338,900	Gateways + 60 months
Added by per-host SaaS monitoring	+$3,221,100	+$4,939,020	60 months at $15 (Pro) or $23 (Enterprise) a host

6 AI blockers when SaaS holds the data

A review is only as good as the evidence it can read. When a SaaS product is the system of record, these are the things that stop AI from using it:

BLOCKER	WHAT HAPPENS	HOW THIS DESIGN AVOIDS IT
Outside the boundary	A SaaS without FedRAMP High (Xano, Redis Cloud, commercial brokers, a commercial ServiceNow tenant) holds the evidence. Sending it to Bedrock in GovCloud moves it across the boundary.	The system of record is Aurora and S3 in GovCloud, and the AI runs in the same boundary.
Contract terms	Terms of service can restrict bulk export, automated access, or processing with a third-party model. Some steer you to the vendor's own AI add-on.	Your data, in your database, under your agency approval.
Vendor model endpoints	Built-in SaaS AI may call a model API that isn't FedRAMP authorized. The direct Anthropic API, for example, is not.	Claude runs only through Bedrock in GovCloud.
Rate limits and paging	API quotas mean an AI review can't read the full history while an incident is happening, only after a nightly export.	The review runner queries Aurora directly, with indexes on site and time.
Retention holds	Default retention (often 30 to 90 days for logs) deletes evidence before a review or audit. Legal hold costs extra.	S3 Object Lock keeps one year of batches, write once; Aurora keeps every decision.
Changeable records	An admin or help-desk user can edit a CMDB row, so an AI verdict built on it can't be defended.	Append-only tables, trace IDs, and records signed by the gate.
Residency and staff access	A commercial-region tenant or vendor support staff outside the US can touch CUI.	US-only processing at Cloudflare High locations and in GovCloud.
Exit	Proprietary schemas and export fees make it hard to move the evidence or rebuild the AI pipeline on another platform.	Plain PostgreSQL and S3 objects, exportable at any time.

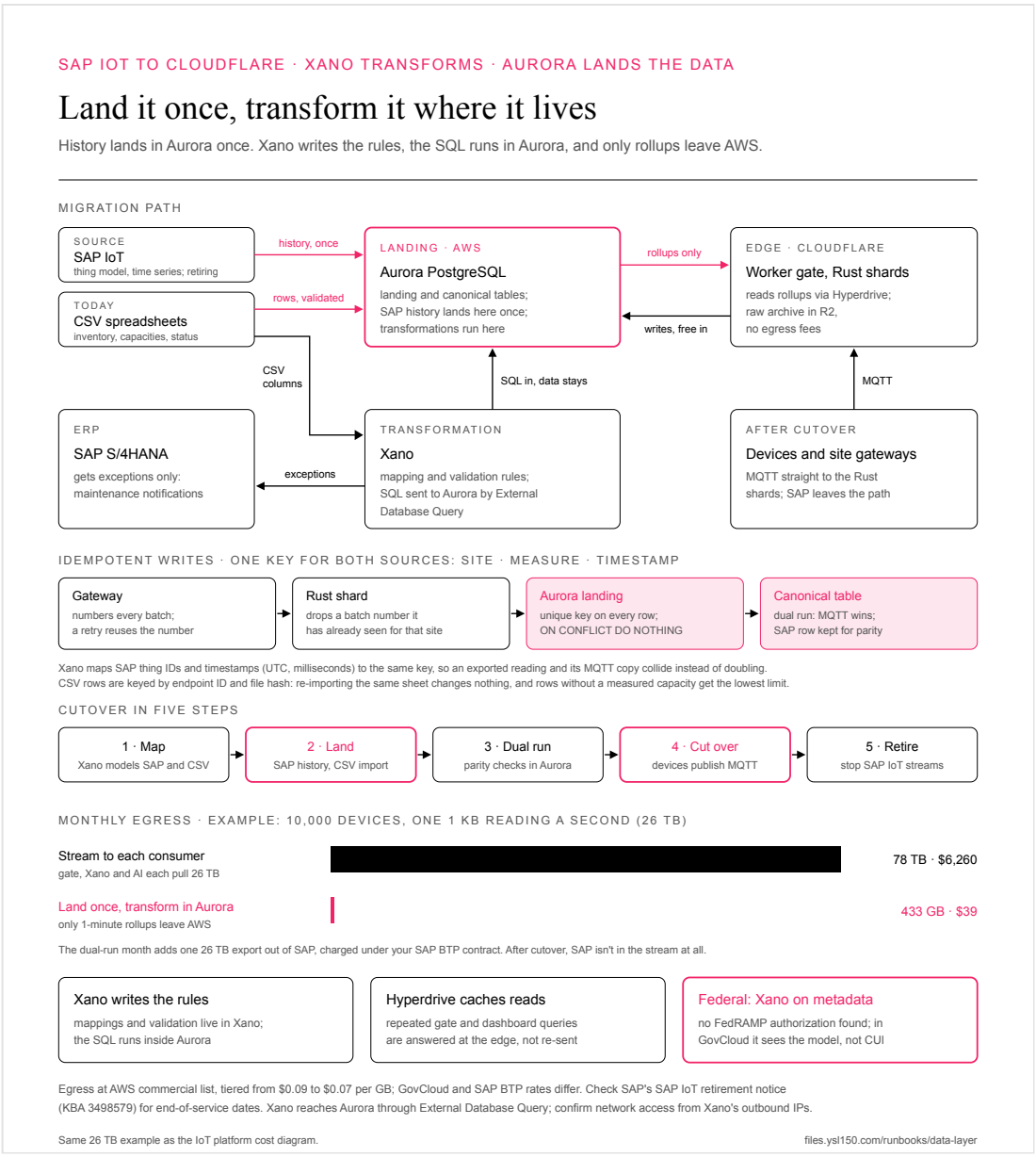
SaaS can still be useful downstream. Give a SaaS tool read-only copies for tickets or dashboards, but never make it the gate or the record the AI reads from.

7 Migrations: SAP IoT, SAP streaming to AI, and ECC

> Not specific to this project. Turning CSV files and manual data pushes into automated feeds is a prerequisite for AI everywhere: SAP, Shopify, IoT. It applies to any data an AI will read, not only this router rollout, so it lives in the runbook as general guidance: First migration on the path to AI: automate every CSV and manual push.

Migrating from SAP IoT

SAP is retiring its SAP IoT service. To move to Cloudflare without paying egress for every consumer, land the SAP history in Aurora PostgreSQL once and let Xano transform it in place. Xano holds the mapping and validation rules and sends the SQL to Aurora, so the telemetry never leaves AWS. Cloudflare reads only rollups, through Hyperdrive, which caches repeated reads. After cutover, devices publish MQTT straight to the Rust shards and SAP IoT leaves the stream. S/4HANA keeps getting what it needs from Xano: exceptions such as maintenance notifications.



Migration path, idempotent writes, cutover steps and monthly egress. The egress example is 10,000 devices sending one 1 KB reading a second, the same as the runbook's IoT cost diagram.

Idempotency and dedupe between data sources

Every place a duplicate can come in has a key and a rule. Writes are safe to repeat: a replayed message, export or action changes nothing.

WHERE	WHERE DUPLICATES COME FROM	KEY	RULE
Gateway to Rust shard	MQTT QoS 1 redelivers a message; the gateway retries after an outage	site ID + batch number	Each shard keeps the last batch number per site in Durable Object storage and drops anything at or below it. One shard owns a site, so there's no race.
SAP export and MQTT during the dual run	The same reading arrives from both sources	site · measure · timestamp (UTC, milliseconds)	Aurora's landing table has a unique index on the key and uses `ON CONFLICT DO NOTHING`. In the canonical table the MQTT reading wins. The SAP row is kept, and any difference is logged as a parity finding.
Rerun of an export or transform job	A job is replayed after a failure	job ID + row key	Xano's job writes are upserts. Aurora marks a finished job as done, so running it again changes nothing.
Gate actions	A quarantine or rollback is sent twice	decision trace ID; configuration bundle hash	The KV quarantine entry is set per site, so a second write is harmless. The router's commit-confirmed uses the bundle hash, so loading the same hash again does nothing.
AI reviews	The same shard trip fires more than once	site + rule + 15-minute window	One Bedrock call per incident. Repeats attach to the open review instead of starting a new one, which keeps the AI line on budget.
Exceptions to S/4HANA	The sender retries a send (Xano on a commercial network, the GovCloud runner on a federal one)	trace ID as the external reference	The sender looks for an existing notification with that reference before creating one.

Is SAP FedRAMP rated?

Only through SAP NS2, SAP's separate US national-security subsidiary. Commercial SAP BTP, which SAP IoT runs on, shows no FedRAMP authorization.

SAP OFFERING	AUTHORIZATION	WHAT IT MEANS HERE
SAP NS2 secure cloud	FedRAMP Moderate (JAB)	A federal S/4HANA can run in a FedRAMP environment, at Moderate, not High.
SAP NS2: S/4HANA Cloud Private Edition and SAP BTP	DISA IL4; IL5 provisional (June 2026)	DoD-level authorizations. Confirm which BTP services, including AI services such as Joule, are inside the approved scope.
Commercial SAP BTP, including SAP IoT	None found	The source of the migration is outside the boundary. Its history comes into GovCloud once, and nothing flows back.

> Federal deployments. I found no FedRAMP authorization for Xano. In GovCloud, keep Xano on the thing model and the mapping rules, which aren't CUI, and let the SQL it writes run in Aurora inside the boundary. Exceptions to an SAP NS2 S/4HANA should go from the Rust runner in GovCloud, not from Xano. A commercial SAP customer can let Xano run the full transformation and send the exceptions.

AI-secure SAP data streaming: approved pipelines that scale without fees

Stream SAP data to AI through pipelines that are approved once and then scale by cohort, with no fee that grows with volume. "Approved" means three things: an interface SAP permits, services inside the authorization boundary, and a contract with a version. "Without fees" means no per-message, per-GB or per-document meter in the streaming path.

Callout: the SAP data barriers, and the Cloudflare and AWS workarounds. SAP data doesn't get stuck because it's hard to read. It gets stuck behind licensing, fees, network position and load limits. Each barrier has a workaround that keeps the data in your boundary:

SAP DATA BARRIER	WHAT SAP DOES	CLOUDFLARE WORKAROUND	AWS WORKAROUND
Extraction API restricted	SAP Note 3255746 bars third-party extraction over ODP-RFC	The Worker runner calls released OData and ODP over OData	AWS Glue SAP OData connector, in GovCloud since July 2026; the AppFlow SAP OData connector in commercial Regions (no GovCloud listing found)
Middleware fees	Integration Suite is sold by message allowances	The runner pulls directly, at Cloudflare's flat request pricing	Glue jobs are billed by run time, not per message
Network reach	SAP runs on premises or in a private cloud behind firewalls; BTP needs SAP Cloud Connector	Cloudflare Tunnel: an outbound-only connector from the SAP network, with no inbound ports, behind Access and mTLS	Private connectivity in the same Region as RISE with SAP on AWS or SAP NS2 on GovCloud: VPC peering, PrivateLink or Direct Connect, with no internet path
System load	Gateway and work-process limits cap concurrent OData calls; large reads slow production	One Durable Object per SAP system caps calls and requests in flight; deltas only	Scheduled delta pulls off peak, with Glue job concurrency capped
Egress and document fees	Data transfer out of SAP-hosted clouds; Digital Access counts documents created from outside	Land once; R2 has no egress fee; Hyperdrive caches reads; streams are read-only	Land in the same Region as the SAP system, so the transfer stays inside AWS at intra-Region rates; write back rarely
Data shape	Abbreviated field names (`MATNR`, `WERKS`), leading zeros, internal units and codes	One Xano shape per destination (commercial)	The SQL shape runs in Aurora (federal); schemas in the Glue Data Catalog
Authorization level	SAP NS2 is FedRAMP Moderate (IL4/IL5 for DoD); commercial BTP has none found	Cloudflare for Government (High) is the only endpoint the sites and runner call	GovCloud (High) holds the High-impact copy; AI runs on Bedrock in GovCloud
Replication steered to SAP's platform	SAP points replication to Datasphere and Business Data Cloud	Your copy, in your boundary, through released APIs	Glue zero-ETL into your own AWS account; AI reads your copy, not SAP's

None of these workarounds touch SAP's own system or licensing. They use the interfaces SAP releases, and they move the copy, the meters and the AI to infrastructure you control. Confirm your SAP contract terms and each AWS service's FedRAMP scope before you rely on them.

Approved: three checks before a pipeline scales

CHECK	WHAT PASSES	WHAT DOESN'T
The interface SAP permits	Released OData APIs, including ODP over OData, and business events	ODP-RFC from a third-party tool. SAP Note 3255746 bars it, with stricter enforcement reported from June 2026
The authorization boundary	Federal: SAP NS2 as the source (FedRAMP Moderate; DISA IL4, provisional IL5), Cloudflare for Government (High), AWS GovCloud and Bedrock (High). Commercial: your own Cloudflare account and Aurora or Xano	A pipeline hop with no authorization at the level the data needs
The contract	One Xano shape per destination, versioned, with each field classified: which fields may reach a model, and which never may	Column positions, files edited by hand, and fields nobody classified

Without fees: take the meters out of the streaming path

METER THAT GROWS WITH VOLUME	WHERE IT CHARGES	PATH WITH NO VOLUME FEE
Integration middleware messages	SAP Integration Suite is sold by monthly message allowances, with extra blocks of messages; one report says each 250 KB over counts as another message	The runner pulls OData deltas (changes since the last timestamp) directly. Cloudflare's request pricing is flat and has no egress fee
Documents created in SAP by outside systems	SAP Digital Access counts document items created in nine document types; viewing, updating and deleting aren't counted	Stream reads, which aren't counted. Keep writes back to SAP rare, deduplicated by trace ID, and approved: each maintenance notification created from outside may count. Check your contract
Egress per GB	Hyperscaler data transfer out	Land once in Aurora; only rollups leave; R2 for raw data, with no egress fee; Hyperdrive caches repeated reads
Messages per device	IoT and eventing platforms that bill per message	MQTT on the Rust shards, batched once a minute
Seats and hosts	Per-host monitoring and vendor AI add-ons	The Rust watch and Aurora rollups; Claude on Bedrock in GovCloud, called on exceptions only

AI-secure: what the model may and may not touch

1. The model never connects to SAP. It reads shaped rollups in Aurora through scoped tools. The tools hold the credentials; prompts and logs never do.
2. Fields are classified at the shape. Only fields marked as allowed reach a model. CUI, personal data and prices under contract stay out.
3. SAP is a legacy endpoint. The runner rate-limits its calls to each SAP system (rate limiting legacy endpoints), so an AI backlog can't flood an ECC or S/4HANA.
4. AI proposes, the runner writes. At the Walk stage a person approves every write back to SAP. The runner creates the SAP document once, keyed by trace ID.
5. Everything is logged: the shape version, the rollup the model read, its answer, and who approved the action.

Steps

1. List today's SAP extraction paths: ODP-RFC tools, Integration Suite flows, CSV exports, manual uploads.
2. Move each one to a permitted interface: ODP-RFC to OData or ODP over OData, CSV to scheduled delta pulls (first migration).
3. Shape and classify each destination in Xano. On a federal network, the SQL runs in Aurora.

4. Approve the pipeline once: contract, boundary, rate limits and a log-only week, then a canary cohort.
5. Scale by cohort with no new fees: more SAP systems, plants or sites run through the same pipeline at the same flat rate.
6. Open AI read access at Walk, then scoped writes at Run.

SAP ECC: data requirements, breaking changes and complications for approved FedRAMP pipelines

ECC is the part of SAP with an end date, not OData or Fiori. ECC 6.0 enhancement packages 0 to 5 left mainstream maintenance on December 31, 2025, with no extended option. Packages 6 to 8 leave it on December 31, 2027, with optional extended maintenance to December 31, 2030 at a surcharge, and only customer-specific maintenance after that. A pipeline that reads ECC today has to survive the move to S/4HANA without being approved again.

Requirements: what ECC needs before it can feed an approved pipeline

REQUIREMENT	WHAT IT TAKES ON ECC
A permitted interface	SAP Gateway on NetWeaver 7.40 or later. ECC has few of the released APIs S/4HANA has (such as `A_Product`), so you build Gateway services or expose extractors through ODP over OData. Third-party ODP-RFC is barred (SAP Note 3255746). Connectors such as the AWS Glue and AppFlow SAP OData connectors also need paging support (`$top`, `$skip`, `$count`)
Its own authorization	An on-premises or private-cloud ECC is covered by the agency's own ATO, not by FedRAMP. The link into FedRAMP High services needs a documented interconnection (an interconnection security agreement, per NIST SP 800-47) in both system security plans
Encryption on the ECC side	TLS 1.2 or later with FIPS-validated cryptography. Older kernels and crypto libraries need updates first; check SAP's CommonCryptoLib versions
A read-only technical user	No dialog users. Authorizations limited to the Gateway services the pipeline reads. Writes go only through the runner, keyed by trace ID
A supported system	After 2027, or 2030 with extended maintenance, security patches narrow. Each unpatched finding becomes an open item on the plan of action and milestones (POA&M). Keep ECC behind an outbound-only Cloudflare Tunnel and rate-limit it as a legacy endpoint

Breaking changes: what moves when ECC becomes S/4HANA

ECC	S/4HANA	WHAT BREAKS	WHAT THE XANO SHAPE DOES
Material number, 18 characters	Up to 40 characters, when the extension is switched on (off by default)	Interfaces and keys sized for 18 characters	Key on a 40-character field from day one; strip and restore leading zeros in one place
Customer and vendor masters	Business Partner, required	Every customer and supplier reference	Map both to one party key; keep the old customer and vendor numbers as aliases
Finance documents in BKPF, BSEG and many index tables	The Universal Journal (ACDOCA)	Direct table reads and reports built on them	Read through released APIs or views, never tables; one finance shape for both
Material documents in MKPF and MSEG	MATDOC	Inventory feeds	One inventory shape, scoped like an EDI 846
Pricing conditions in KONV	PRCD_ELEMENTS	Price and discount feeds	One price shape, scoped like an EDI 832

ECC	S/4HANA	WHAT BREAKS	WHAT THE XANO SHAPE DOES
Non-Unicode code pages, on some older systems	Unicode only	Characters in names and descriptions	Normalize to UTF-8 in the shape, before JSON
Dates and times without a time zone (DATS, TIMS)	The same fields, plus some UTC timestamps	Idempotency keys collide or split at daylight-saving changes	Convert to UTC milliseconds in the shape, so the dedupe key holds
Custom Gateway services and BW extractors	Released APIs and CDS views; not every ECC extractor is released for S/4HANA	Pipelines bound to ECC-only services	Swap the source connector; the shape, Aurora tables and AI tools stay the same

Complications: what makes an approved FedRAMP pipeline harder with ECC

1. Two boundaries, one stream. ECC sits under the agency's ATO; the pipeline is in FedRAMP High services. Both plans have to describe the same connection, data flow and impact level, and a change on either side can reopen the other.
2. The data's impact level wins. The pipeline must be authorized at least at the level of the data it carries, whatever level the ECC system itself was assessed at.
3. SAP NS2 is Moderate. Moving ECC to RISE through SAP NS2 brings FedRAMP Moderate (JAB) and DoD IL4 or IL5, not FedRAMP High. A High-impact copy still belongs in GovCloud.
4. Third-party extractors can stop working. Any tool that still uses ODP-RFC is exposed to SAP's reported enforcement from June 2026. List them before they break.
5. Dual running doubles the sources. During the move, ECC and S/4HANA both produce the same business objects. One shape per destination, a shared key and the parity checks from the SAP IoT migration keep the AI reading one version of each record.
6. Continuous monitoring covers the link too. Logs for the pipeline, its runner and the ECC technical user go to CloudTrail and Security Hub in GovCloud, so the evidence for the ATO is collected as the data moves.
7. Write-backs can be both licensed and audited. A document created in ECC from outside can count under SAP's indirect-use licensing. It needs a person's approval, a trace ID and a log entry.

---

8 Rate limiting legacy endpoints: software, evaluation and event functions

The oldest receiver sets the limit, not the newest sender. To rate-limit the data flowing into legacy endpoints (old radar and radio receivers, serial circuits, SOAP and OData services), three things are needed: software that enforces the limit at every hop, an evaluation that measures what each endpoint can really take, and management of the event functions that react when a limit is hit.

Starting point: CSV spreadsheets. Assume today's inventory, capacities and certification status live in CSV spreadsheets: no versions, no owner per row, and nothing a gate can read. The first job is to import them into tables the gate can enforce from, and to keep the spreadsheets as exports from then on, not the source.

This builds on the throttle formula in the runbook and the staged deploy for mTLS rules.

1. Software: enforce the limit at every hop

HOP	COMPONENT	WHAT IT ENFORCES
Router	Policer on new connections in the firewall filter	Hard cap per receiver, in hardware; holds even if everything else fails

HOP	COMPONENT	WHAT IT ENFORCES
Site gateway	Rust token bucket per receiver (governor) and a Tokio semaphore	Rate and messages in flight per receiver; for serial links, a minimum gap between messages
Edge	Workers rate limiting, keyed by the client certificate	Each machine is limited as itself, in 10- or 60-second windows
Per legacy endpoint	One Durable Object per receiver, holding its total budget	The sum across every sender stays under the receiver's capacity; one object per receiver, so the count is exact
CSV import	Upload to a staging table, validated against the schema Xano models from the spreadsheet's columns, then merged into the canonical tables	Bad rows are rejected with a reason, not loaded. Idempotent by endpoint ID and file hash, so re-importing the same sheet changes nothing
Store	Capacity and throttle tables in Aurora (federal) or Xano (commercial), published to KV	Every limit has a source measurement, a version and an owner
Retries	Exponential backoff with jitter, a retry budget (at most 10% of calls), `retry-after` honored, idempotency keys	Retries can't become a second flood, and a repeated call changes nothing
Circuit breaker	Per receiver: opens on errors or latency over threshold, then a single probe	A struggling receiver gets rest instead of more load

Fan-out protection costs $0 in software. Fan-out is one source multiplying into many downstream calls: one router opening connections to many receivers, many senders reconnecting to one legacy receiver at once, one alarm triggering many AI reviews or SAP documents. Every control that stops it is router configuration or open-source code you own, so none carries a license or subscription fee:

FAN-OUT RISK	CONTROL	SOFTWARE	SOFTWARE COST
One router floods the receivers it feeds	Policer on new connections, per receiver	Junos firewall filter: configuration	$0
A gateway sends faster than a receiver can take	Token bucket and in-flight cap per receiver	governor and tokio, MIT and Apache-2.0	$0
Many senders converge on one legacy receiver	Shared budget per receiver	Your Rust in a Durable Object	$0
Every sender reconnects at once after an outage	Reconnect ramp with jitter	Your Rust on the gateway	$0
Timeouts turn into a retry storm	Backoff with jitter, a retry budget, idempotency keys	Your Rust; keys in Aurora	$0
One alarm fans out to many AI reviews or SAP documents	A budget per event function, dedupe by trace ID, a kill switch	Your Rust; the switch is a KV flag	$0
A struggling receiver keeps getting load	Circuit breaker per receiver	Your Rust	$0

"$0" means no license, subscription or per-unit fee. Adding these controls doesn't add a cost line: the Durable Object and KV usage is already inside the gate's monthly cost, and the gateway hardware is already counted. The real cost is the engineering time to build and test them, which this example doesn't price.

2. Evaluation: measure each legacy endpoint before setting its limit

1. Import today's CSV spreadsheets of endpoints, capacities and senders. Each row needs an endpoint ID, protocol, owner and the date it was last measured. Rows without a measured capacity are flagged as untested, and they get the most conservative limit until they're tested.
2. List the endpoints from the data dependency map, with protocol (serial, TCP, HTTP, SOAP, OData), age and every sender that feeds each one.
3. Ramp test a lab copy of the oldest receiver, headless: raise new connections a second and messages a second in steps (for example 10% every two minutes), and record errors, p99 latency and queue depth. The knee is where p99 latency doubles or errors pass 0.1%.
4. Test the reconnect storm: every sender reconnecting at once after an outage. This is the case that floods old receivers. Set a reconnect ramp, for example at most 10% of senders a second, with jitter.
5. Set the limit: knee × headroom (50%) × this sender's share, the runbook's throttle formula.
6. Run log-only in production for two weeks, then enforce on one canary site, then by cohort.
7. Record the result in the machine certification record with its test ID. Evaluate again when firmware, sender count or receiver hardware changes, and recompute nightly as a check.

PER ENDPOINT, THE EVALUATION PRODUCES	EXAMPLE
Capacity knee	40 new connections a second
Safe limit for all senders together	20 a second (50% headroom)
Maximum in flight	8 sessions
Minimum gap between messages (serial)	50 ms
Recovery time after overload	90 seconds
Reconnect ramp	10% of senders a second, with jitter

3. Event function management: the functions that react to a limit

Each event function (alarm, quarantine, rollback, AI review, exception to SAP) is managed like a release, not a script:

CONTROL	WHAT IT MEANS
Registry	ID, version, owner, trigger, allowed actions and signed WASM hash, stored with the evidence
Budget per function	Maximum invocations per site and per minute: a quarantine at most once per site every 15 minutes, one AI review per incident, one SAP exception per trace ID
Idempotency key	Every invocation carries a trace ID, so a repeat does nothing
Concurrency cap and queue	Bounded per receiver; overflow goes to a dead-letter queue, never to the endpoint
Kill switch	A KV flag turns a function off fleet-wide in seconds
Change control	A new or changed function goes log-only, then canary, then by cohort. Only the two pre-approved actions run without an operator.
Audit	Every invocation is logged with its trace ID, input hash and decision
AI	Advisory only, with its own budget: it can recommend a quarantine, never apply one

Would idempotency alone have stopped El Dorado? No. Its 200 to 250 connection attempts a second were new connections, not duplicates, and an old receiver pays for the handshake before it can read a key. Idempotency makes retries and configuration pushes safe to repeat. The policer, the reconnect ramp and the per-receiver budget cap the rate. Background: timeouts, retries and backoff with jitter (Amazon Builders' Library).

Assumptions and sources

- Sizing: 3,579 sites held at the gate, out of 3,648 planned (per senators citing L3Harris). Gateways send one 4 KB MQTT batch a minute (173 MB per site a month). AI reviews run on 1% of sites a day typical and 10% in a surge, plus 500 operator requests a month, at 30k tokens in and 3k out each.
- Prices: Cloudflare public list; the Government contract price isn't public. Aurora and S3 GovCloud rates are estimated at about 1.2× commercial. Opus 5.5 on Bedrock GovCloud is estimated at $4.80 / $24 per million tokens, applying the same 1.2× seen in Sonnet 5.5's third-party GovCloud listing ($2.40 / $12). The security line is rough.
- Cellular: a 250 MB plan at about $513 over five years per site, the same rate used in the runbook's router cost diagram. Gateway: Teltonika RUTX50 at about $600.
- Scope: a scoping example with example pricing from public sources, set against the $12.5 billion Congress appropriated in 2025 and the $50.1 billion the Transportation Secretary asked Congress for in November 2026. No confidential, customer or NDA material is used.
- Excluded: the ATO assessment, staff, build and maintenance effort, and router and Junos licenses.
- FedRAMP: Cloudflare for Government High (Nextgov, Aug 2026; Cloudflare blog). Claude on Bedrock GovCloud High and IL4/5 (Anthropic; AWS). EKS, Aurora and ElastiCache in GovCloud (AWS services in scope). Hyperdrive scope (Cloudflare changelog, May 2025). Check every listing on fedramp.gov/marketplace before you submit.
- SAP IoT migration: SAP knowledge base article 3498579 (SAP IoT retirement); Xano External Database Query documentation; egress at AWS commercial list, tiered from $0.09 to $0.07 per GB. GovCloud and SAP BTP egress rates were not available.
- Data today: inventory, capacities and certification status are assumed to live in CSV spreadsheets, imported into Aurora before any limit is enforced.
- Every source is linked in section 9, References.

---

9 References

Links checked November 19, 2026. FedRAMP listings and prices change: confirm each before you submit.

Scope and public cost projections

1. US DOT: administration calls on Congress to fully fund air traffic control modernization ($50.1 billion) transportation.gov/briefing-room/…-501
2. ABC News: Duffy asked for $50B for air traffic control after ground stop abc7news.com/story/duffy-asked-50b-air-traffic-control…
3. Aviation Week: Trump administration seeks $50B for ATC overhaul aviationweek.com/…/trump-administration-seeks-50b-atc-overhaul
4. GAO-26-107992: FAA's Brand New Air Traffic Control System files.gao.gov/reports/GAO-26-107992
5. FAA head says billions more needed to modernize air traffic control (Sept 2026) fox7austin.com/news/faa-head-billions-modernize-air-traffic-control

FedRAMP status: Cloudflare

1. Nextgov: Cloudflare to seek DoD IL4 after new FedRAMP, GovRAMP designations (Aug 2026) nextgov.com/acquisition/2026/08/cloudflare-seek-dod-il4-…
2. Cloudflare blog: Cloudflare for Government achieves FedRAMP Class D (High) blog.cloudflare.com/fedramp-class-d-certification
3. Cloudflare for Government cloudflare.com/cloudflare-for-government
4. Cloudflare: FedRAMP architecture brief cfl.re/Cloudflares_Unique_FedRAMP_Architecture
5. Cloudflare blog: public sector scope (R2, API Shield, Hyperdrive) blog.cloudflare.com/cloudflares-commitment-to-advancing-public-s…
6. Cloudflare changelog: Hyperdrive FedRAMP (May 2025) developers.cloudflare.com/changelog/2025-05-14-hyperdrive-fedramp
7. Cloudflare changelog: Containers and Sandboxes GA (Apr 2026) developers.cloudflare.com/changelog/2026-04-13-containers-sandbox-ga
8. Cloudflare changelog: regional placement, fedramp jurisdiction (Apr 2026) developers.cloudflare.com/changelog/2026-04-05-regional-placement
9. FedRAMP Marketplace fedramp.gov/marketplace

FedRAMP status: AWS GovCloud and Claude

1. AWS: services in scope by compliance program (FedRAMP) aws.amazon.com/compliance/services-in-scope/FedRAMP
2. AWS: Aurora MySQL and PostgreSQL FedRAMP High in GovCloud aws.amazon.com/about-aws/whats-new/2019/11/amazon-aurora-my…
3. AWS: Aurora Serverless v2 supported Regions docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Concepts.Au…
4. AWS: Amazon EKS FedRAMP High aws.amazon.com/about-aws/whats-new/2021/04/amazon-eks-is-no…
5. AWS: ElastiCache for Redis FedRAMP in GovCloud aws.amazon.com/about-aws/whats-new/2018/09/amazon-elasticac…
6. AWS: Bedrock models FedRAMP High and DoD IL4/5 in GovCloud aws.amazon.com/about-aws/whats-new/2025/05/amazon-bedrock-m…
7. Anthropic: Claude in Amazon Bedrock, FedRAMP High anthropic.com/news/claude-in-amazon-bedrock-fedramp-high
8. Anthropic: Claude for Government claude.com/government
9. Docker: FedRAMP compliance with Docker Hardened Images docker.com/blog/fedramp-compliance

Prices

1. Cloudflare: Workers rate limiting binding developers.cloudflare.com/workers/runtime-apis/bindings/rate-limit
2. Amazon Builders' Library: timeouts, retries and backoff with jitter aws.amazon.com/builders-library/timeouts-retries-and-backoff…
3. Cloudflare: Durable Objects pricing developers.cloudflare.com/durable-objects/platform/pricing
4. AWS: Private CA pricing aws.amazon.com/private-ca/pricing
5. AWS: IoT Core pricing aws.amazon.com/iot-core/pricing
6. AWS: Bedrock pricing aws.amazon.com/bedrock/pricing
7. Future AGI: Claude Sonnet 5.5 on Bedrock, US GovCloud (third-party) futureagi.com/llm-cost-calculator/bedrock/us-gov-anthropic…
8. Withorb: how Datadog prices (Nov 2026 capture) withorb.com/pricing-atlas/datadog
9. CloudZero: AWS egress costs cloudzero.com/blog/aws-egress-costs
10. Enterprise Storage Forum: S3 pricing by region (GovCloud row from 2018) enterprisestorageforum.com

Migrations: SAP IoT, SAP NS2, ECC and Xano

1. SAP KBA 3498579: SAP Internet of Things retirement userapps.support.sap.com/sap/support/knowledge/en/3498579
2. SAP: migration guide for SAP IoT legacy services help.sap.com/doc/7eb381fcf3b44745b845153c8d2eefb3/2206b/e…
3. SAP Community: SAP IoT retirement and SAP APM groups.community.sap.com/t5/supply-chain-management-blogs-by-sap/sap-…
4. SAP NS2: solution suite GA in FedRAMP+ Impact Level 5 sapns2.com/news/sap-ns2-announces-general-availability…
5. ERP Today: SAP NS2's FedRAMP+ IL5 authorization erp.today/sap-ns2-fedramp-il5-defense-erp
6. SAP NS2: S/4HANA Cloud FedRAMP authorized prnewswire.com/news-releases/sap-ns2-announces-sap-s4hana…
7. AWS: run SAP workloads at DoD IL5 with SAP NS2 on GovCloud aws.amazon.com/blogs/publicsector/run-sap-workloads-at-dod…
8. SAPinsider: SAP Note 3255746, the ODP-RFC deadline sapinsider.org/blogs/sap-note-3255746-odp-rfc-deadline-2026
9. Theobald Software: data integration without ODP-RFC theobald-software.com/en/blog/sap-note-325574
10. SAP: Digital Access sap.com/assetdetail/2024/08/9ed76003…
11. Flexera: Digital Access document types and multipliers docs-sap.flexera.com/…/indirect-usage/digital-access
12. SAP: Integration Suite pricing sap.com/canada/products/technology-platform/integration-suite/pricing
13. AWS: Glue SAP OData connector and zero-ETL in GovCloud (July 2026) aws.amazon.com/about-aws/whats-new/2026/07/aws-glue-sap-zero-etl…
14. AWS: AppFlow SAP OData connector docs.aws.amazon.com/appflow/latest/userguide/sapodata
15. AWS for SAP: ODP-based change data capture over OData aws.amazon.com/blogs/awsforsap/sap-odp-based-change-data…
16. Cloudflare: Tunnel developers.cloudflare.com/cloudflare-one/connections/connect-networks
17. SAPinsider: SAP end-of-maintenance plans for Business Suite sapinsider.org/blogs/sap-details-end-of-maintenance-plans…
18. Rimini Street: no extension to the 2025 and 2027 ECC dates riministreet.com/…/no-extension-to-ecc-support-2027-deadline
19. SAP Community: the long material number in S/4HANA community.sap.com/…/the-long-material-number-in-sap-s-4hana…
20. NIST SP 800-47 Rev. 1: managing the security of information exchanges csrc.nist.gov/pubs/sp/800/47/r1/final
21. Xano: External Database Query docs.xano.com/the-function-stack/functions/database-reques…

Open-source components

1. tokio (async runtime) tokio.rs
2. wasmtime (WASM runtime) wasmtime.dev
3. tonic (gRPC for gNMI) github.com/hyperium/tonic
4. rumqtt (MQTT client and codec) github.com/bytebeamio/rumqtt
5. rustls (TLS) github.com/rustls/rustls
6. aws-lc-rs (FIPS cryptography provider) github.com/aws/aws-lc-rs
7. rust-tss-esapi (TPM 2.0) github.com/parallaxsecond/rust-tss-esapi
8. workers-rs (Rust on Cloudflare Workers) github.com/cloudflare/workers-rs
9. sqlx (Rust and PostgreSQL) github.com/launchbadge/sqlx
10. governor (rate limiting for Rust) github.com/boinkor-net/governor
11. SQLite sqlite.org
12. step-ca (open-source CA) github.com/smallstep/certificates

This runbook's diagrams

1. Control plane for router rollout files.ysl150.com/runbooks/data-layer/control-plane.svg
2. Federal data cost, 3,579 sites files.ysl150.com/runbooks/data-layer/federal-data-cost.svg
3. SAP IoT to Cloudflare files.ysl150.com/runbooks/data-layer/sap-iot-migration.svg
4. Router cost per branch site files.ysl150.com/runbooks/data-layer/router-cost.svg
5. IoT platform cost files.ysl150.com/runbooks/data-layer/iot-cost.svg
6. Data Layer Runbook files.ysl150.com/runbooks/data-layer