

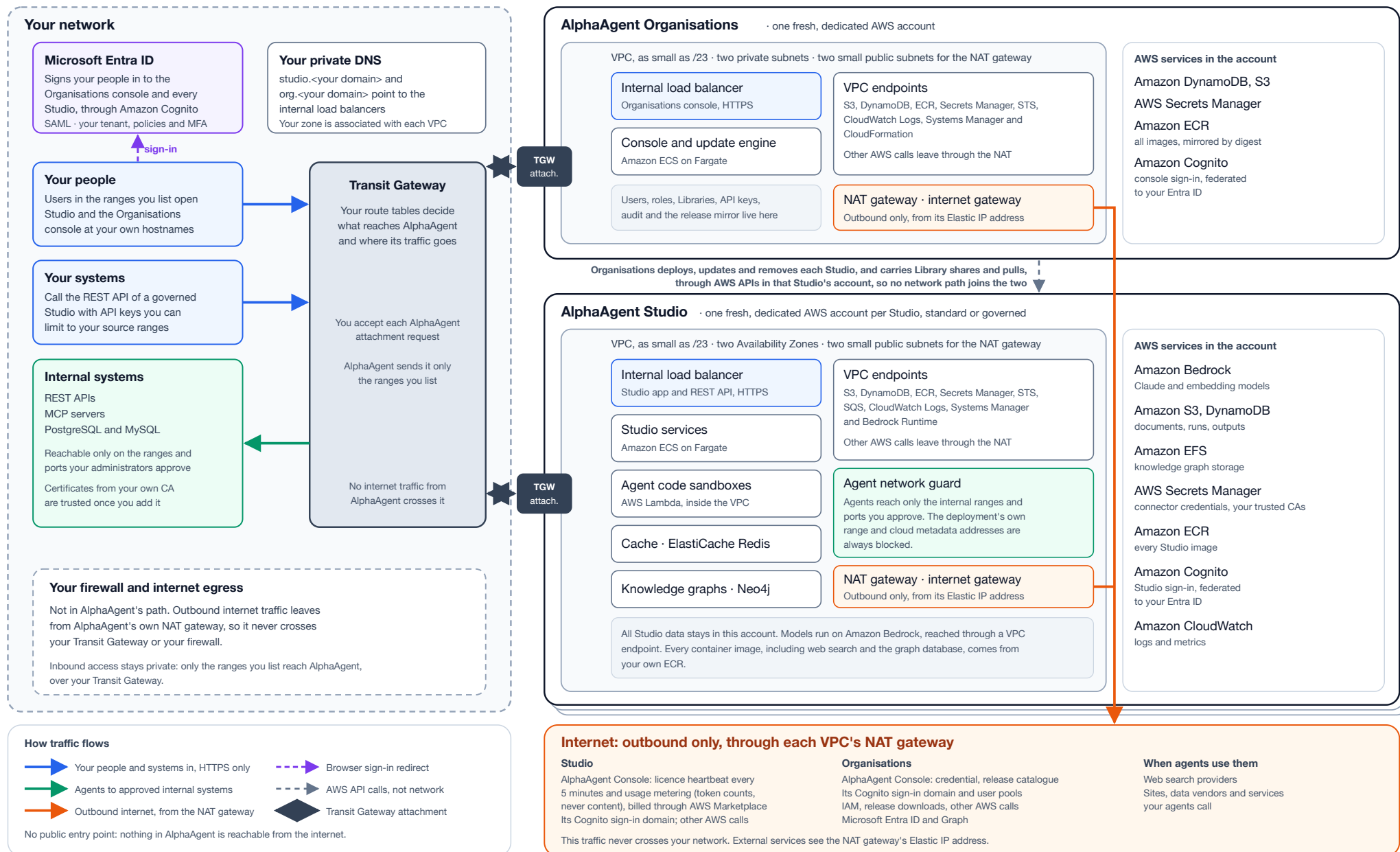
AlphaAgent private deployment on your AWS network

Studio and Organisations, reached only over your Transit Gateway, with outbound traffic through AlphaAgent's NAT gateway

AlphaAgent

Prometheus Research Labs

Studio 2.0.19 · Organisations 1.0.19



How it works, and what each side provides

The same design for every deployment: one fresh, dedicated AWS account for Organisations and one for each Studio

How traffic flows

- Your people and systems reach AlphaAgent over your Transit Gateway.** Studio and the Organisations console sit behind internal load balancers at your own hostnames, admitting HTTPS only from the user access ranges you list. Nothing is reachable from the internet.
- Sign-in uses your Microsoft Entra ID.** The Organisations console and each Studio federate to your tenant through Amazon Cognito (SAML). Organisations creates each Studio's enterprise application through Microsoft Graph. Only sign-in redirects take this path, never your data.
- Agents reach your internal systems only where you allow.** Connectors and agent code reach the internal service ranges and ports your administrators approve, and trust your own certificate authorities once you add them. Only the ranges you list are routed to your Transit Gateway.
- Outbound internet traffic leaves through AlphaAgent's NAT gateway.** Each AlphaAgent VPC holds one NAT gateway with an Elastic IP address and an internet gateway, in two small public subnets that hold nothing else. Internet traffic never crosses your Transit Gateway or your firewall, and nothing on the internet can open a connection in.
- Core AWS services stay on AWS.** VPC endpoints carry the calls to Bedrock Runtime, S3, DynamoDB, ECR, Secrets Manager and the other core services; other AWS calls leave through the NAT gateway. Every container image comes from your own ECR.
- Organisations manages Studios through AWS APIs.** It deploys, updates and removes each Studio with a cross-account IAM role, and moves Library shares and pulls the same way, so no network path joins the two.
- Licensing goes to the AlphaAgent Console.** Studio sends a signed licence heartbeat every 5 minutes and usage metering (token counts, never content), billed through AWS Marketplace. Organisations uses the same host for its credential and releases.

Design principles

- Identical units.** Organisations and every Studio each run in a fresh, dedicated AWS account with their own VPC, built the same way. Upgrades, support and removal behave the same for one deployment or a hundred.
- Compact address ranges.** A VPC can be as small as a /23: its first half holds the two private subnets, and two /28 subnets hold the NAT gateway. Organisations checks every range against your other deployments and its own, so nothing overlaps on your network.
- Clean ownership.** AlphaAgent owns everything inside its accounts. You own your network, Transit Gateway route tables and DNS, and AlphaAgent never changes them.
- Fixed, checked settings.** Network access, outbound path, VPC range and Transit Gateway are fixed for each deployment's life, and checked before anything is created.

What leaves through the NAT gateway

Destination	Used by	Why
telemetry.api.alphaagent.prometheusr.l.com	Studio, Organisations	Licence, heartbeat and metering; the Organisation credential and release catalogue
<prefix>.auth.<region>.amazoncognito.com	Studio, Organisations	Sign-in: the load balancer completes each one
AWS service endpoints	Studio, Organisations	AWS calls with no VPC endpoint in the region, calls to other regions, and release downloads
login.microsoftonline.com, graph.microsoft.com	Organisations	Microsoft Entra ID, once you connect it
Web search providers, and the sites and services agents use	Studio	Only when agents use those tools

None of this traffic crosses your network, so your firewall needs no rule for it. External services see the NAT gateway's Elastic IP address: add it where a service you connect needs an allow-list, such as a Snowflake network policy or a public database.

Who provides what

You provide	AlphaAgent creates and runs
Fresh AWS accounts: one for Organisations, one per Studio	The VPC, private subnets, internal load balancer and VPC endpoints in each account
Your Transit Gateway, shared to those accounts; someone to accept each attachment; routes back to each VPC range	The attachment request, and routes to the ranges you list and nothing else
Nothing for internet traffic: no egress VPC, firewall rule or egress IP address	The NAT gateway, its Elastic IP address, the internet gateway and two small public subnets that hold nothing else
A VPC range per deployment (a /23 is enough), user access ranges, internal service ranges and ports	Overlap and capacity checks across your Organisation, shown in the Address plan
Private DNS records and certificates for your hostnames, with the zone associated with each VPC	Sign-in through your Entra ID, and the Studio services, data stores and agent sandboxes
Your CA certificates, if your internal services use your own CA	Upgrades by Update Manager, with rollback to a restore point