REFERENCE SPECIFICATION

# On-Premises Deployment Specification

Topologies, sizing, OS and database support, network and security, validation, backup and disaster recovery, install runbook, and the support model for on-premises and private-cloud deployments of **V5 Ultimate**.

| | |
|---|---|
| **Document** | On-Premises Deployment Specification |
| **Product** | V5 Ultimate |
| **Version** | v2026.06 |
| **Last updated** | June 23, 2026 |
| **Audience** | Customer IT, Quality, Procurement |
| **Owner** | S.G. Systems, LLC |
| **Companion document** | On-Premises Licence Addendum (commercial terms) |

# Contents





**S.G. Systems, LLC · Confidential — for customer evaluation**

## 1 Scope and intent

This document describes the reference architecture and minimum requirements for installing and operating V5 Ultimate inside customer-controlled infrastructure. It is intentionally conservative — most deployments comfortably exceed these numbers. Final sizing for a given site is confirmed in writing during the design workshop, taking account of plant count, concurrent operator sessions, line throughput, retention period, and any HA / DR posture.

Commercial terms (licence grant, support, escrow, liability) are governed by the **On-Premises Licence Addendum** to the MSA. This specification is technical only and does not amend the contract.

## 2 Supported deployment topologies

| Topology | Description |
|---|---|
| Bare-metal / VM (Linux) | App + Postgres on customer-owned RHEL 9 / Ubuntu 22.04 LTS / Rocky 9. Single-host or split app/db. Suits one-site plants and validated GxP environments where IT prefers traditional package management. |
| Docker / Docker Compose | Single-host container deployment using the published OCI images and a Compose file we provide. Recommended for small/mid sites and pilot environments. Air-gap capable when run without outbound telemetry. |
| Kubernetes (Helm) | Helm chart for multi-site, HA deployments. Targets EKS / AKS / GKE / OpenShift / vanilla k8s 1.28+. Supports horizontal scaling of the app tier, externalised Postgres, and customer-managed ingress. |
| Private cloud (AWS / Azure / GCP tenancy) | Deployed inside the customer's own cloud account/VPC. Managed Postgres (RDS / Azure Database / Cloud SQL), customer KMS, customer object storage for attachments and backups. We never hold credentials. |
| Hybrid (on-prem + private cloud) | Mix on-prem and cloud tiers — e.g. app/kiosks on the plant floor with managed Postgres in the customer's cloud, or on-prem primary with a warm cloud DR standby. See §2a for the supported shapes. |

## 2a Hybrid deployment patterns

"Hybrid" means any deployment that spans on-premises infrastructure and the customer's private cloud tenancy. V5 Ultimate supports four common patterns out of the box. All four use the same binaries, the same Helm chart or Compose file, and the same IQ/OQ pack — the difference is where each tier runs and how the link between them is sized.

- **Split-tier hybrid (app on-prem, data in cloud).** App and operator kiosks run on-premises for low latency to PLCs, scales, and printers. Postgres is the customer's managed service (RDS / Azure Database / Cloud SQL) and attachments live in S3-compatible object storage in the same cloud account. Requires a stable site-to-cloud link (≥ 50 Mbps, ≤ 30 ms RTT) — falls back to read-only mode on WAN loss.
- **Edge + control-plane hybrid (kiosks on-prem, app in cloud).** A thin edge node on the plant network terminates kiosks, scales, label printers and OPC-UA/Modbus sources, with store-and-forward so the floor keeps running through WAN outages. The main app, Postgres, and admin UI run in the customer's private cloud. Recommended where IT prefers to centralise the control plane across multiple sites.
- **On-prem primary + cloud DR.** Production runs on-prem; a warm standby (app replicas idle, Postgres streaming replica) runs in the customer's cloud account for disaster recovery. Cuts RTO to ≤ 30 minutes and RPO to ≤ 15 minutes without duplicating on-prem hardware.


S.G. Systems, LLC · Confidential — for customer evaluation

- **Multi-site hybrid.** Some plants on-prem (regulated / air-gapped / high-latency sites), others in the customer's private cloud (greenfield or smaller sites), unified by a single SAML/SCIM identity plane and a shared reporting tier. Per-site data residency is preserved.

Hybrid deployments are scoped in the same design workshop as any other topology. The SOW identifies which tier lives where, who owns the link between them, and how IQ/OQ is run across both halves.

## 3 Reference hardware sizing

Numbers below assume one production site, a 7-year retention window for regulated records, and standard Postgres on local NVMe or equivalent. Add ~25% per additional concurrent plant. Sizing is per environment (production / validation / sandbox) and includes both the application and database tier unless split.

| Tier | Concurrent operators | vCPU | RAM | Storage (yr 1) |
|---|---|---|---|---|
| Small (single line / pilot) | ≤ 25 | 4 | 16 GB | 200 GB SSD |
| Medium (one plant) | 25–100 | 8 | 32 GB | 500 GB SSD |
| Large (multi-line plant) | 100–300 | 16 | 64 GB | 1 TB NVMe |
| Enterprise (multi-site, HA) | 300+ | 2×16 (app) + 16 (db) | 128 GB+ | 2 TB+ NVMe, replicated |

## 4 Operating system, runtime and database matrix

- **OS (bare-metal / VM):** Red Hat Enterprise Linux 9, Rocky Linux 9, AlmaLinux 9, Ubuntu Server 22.04 LTS / 24.04 LTS, SUSE Linux Enterprise Server 15 SP5.
- **Container runtime:** Docker 24+ or Podman 4+; Kubernetes 1.28+ (EKS, AKS, GKE, OpenShift 4.14+, Rancher, vanilla k8s).
- **Database:** PostgreSQL 15 or 16 (self-managed, AWS RDS, Azure Database for PostgreSQL, Google Cloud SQL, Crunchy Data, EDB). Postgres 14 supported during migration windows only.
- **Object storage (attachments, backups):** S3-compatible — AWS S3, Azure Blob (S3 gateway), MinIO, Wasabi, Backblaze B2, or local filesystem on smaller single-host installs.
- **Browser support (operator kiosks & office users):** evergreen Chromium, Edge, Firefox, Safari — last two major versions.
- **Reverse proxy / TLS termination (customer-provided):** NGINX, HAProxy, Traefik, F5, Azure Application Gateway, AWS ALB. TLS 1.2 minimum, TLS 1.3 recommended.

## 4a Can I run V5 Ultimate on Windows Server?

Short answer: **yes.** What changes between options is the validation story and the support scope — and how much that matters depends on whether the site is regulated (GxP — pharma, medical devices, food under FSMA, etc.) or non-regulated (industrial, contract manufacturing, internal use, R&D, pilots).

V5 Ultimate's OCI images, Helm chart, systemd units, backup scripts, and the shipped IQ/OQ evidence are built and validated against Linux (RHEL 9, Ubuntu 22.04/24.04 LTS, Rocky 9, SLES 15). That is the only configuration covered by the standard IQ/OQ pack. Everything else still *runs* — it just isn't pre-validated for you.



S.G. Systems, LLC · Confidential — for customer evaluation

| Option | How it runs | Regulated (GxP) | Non-regulated |
|---|---|---|---|
| **Linux VM on Hyper-V / VMware / Nutanix / Proxmox** | Customer provisions an Ubuntu 22.04 or RHEL 9 guest on their existing Windows Server / virtualisation host. V5 installs into the Linux guest as documented. | Recommended. Covered by the shipped IQ/OQ pack, no extra validation cost. | Recommended. Fastest path; same support coverage as any other Linux deployment. |
| **Docker Desktop on Windows Server with WSL2** | Containers run on the WSL2 Linux kernel hosted by Windows. Works with the published Compose file. | Pilot / sandbox only. Not covered by the standard IQ/OQ pack. | **Fully usable for production.** Standard M&S support applies — no validation pack needed because no regulator is asking for one. |
| **Windows Server as the direct host (no Linux guest)** | Possible only via Docker Desktop / WSL2 as above — there is no Windows-native build of the V5 app tier. Postgres can run natively on Windows; the app cannot. | Custom validation, chargeable. Bespoke IQ/OQ rewrite and a written variance to the support term, quoted per engagement. | Supported under the standard support term. No extra cost beyond the normal Order Form. |

**For regulated plants** the recommendation is unambiguous: use the Linux VM on Hyper-V (or VMware / Nutanix / Proxmox). It keeps Windows IT in charge of the host while preserving the shipped validation pack at no extra cost. A Windows-native validated install is available, but it's a chargeable professional-services engagement with its own IQ/OQ evidence and support variance.

**For non-regulated plants** the validation question doesn't apply — there is no regulator asking for an IQ/OQ pack. Pick whichever option fits your IT team: Linux-guest-on-Hyper-V is still the smoothest path, but Docker Desktop + WSL2 directly on Windows Server is fine for production and is covered by the same Maintenance & Support term as any other on-prem install.

Postgres is a separate question. PostgreSQL 15/16 runs natively on Windows Server and is supported as the database tier in any of the three options above — regulated or not.

## 5 Network and ports

- **Inbound to app:** 443/tcp (HTTPS) from operator kiosks, office users, and integrated systems on the customer LAN/WAN.
- **App → database:** 5432/tcp (Postgres), inside the deployment subnet.
- **App → object storage:** 443/tcp to the configured S3 endpoint.
- **Egress (optional):** 443/tcp to the customer's chosen identity provider (Okta, Entra ID, Google Workspace SAML, JumpCloud, OneLogin), SMTP relay for notifications, and any ERP / MES / LIMS integration endpoints. Egress can be disabled entirely for air-gapped sites — V5 Ultimate does not require outbound calls to S.G. Systems.
- **Hardware on the floor (scales, label printers, scanners, sensors):** on the same VLAN as the application or routed through a customer-controlled gateway.

## 6 Identity and access



S.G. Systems, LLC · Confidential — for customer evaluation

- SAML 2.0 SSO against any compliant IdP. SCIM 2.0 user/group provisioning supported on container and Kubernetes deployments.
- Role-based access control with roles stored in a separate `user_roles` table (never on the user/profile row) to prevent privilege-escalation classes of attack.
- 21 CFR Part 11 §11.300 — unique user IDs, password complexity and expiry enforced at the application layer, with optional delegation to the IdP for both.
- Operator-only accounts are locked to the kiosk surface and have no admin UI access.

## 7 Data security and cryptography

- **In transit:** TLS 1.2+ on every external connection, mTLS available for integration endpoints on request.
- **At rest:** AES-256 — Postgres TDE where supported by the customer's distribution, plus filesystem / volume encryption (LUKS, EBS encryption, Azure Disk Encryption, GCP CMEK). Customer-owned encryption keys via the customer's KMS — S.G. Systems never holds them.
- **Audit trail:** every signature, override, and regulated-record write lands in an immutable audit table with operator, timestamp, IP, reason, and before/after hash. Append-only at the database level — not editable through the application.
- **Secrets:** sourced from environment variables, Docker/K8s secrets, AWS Secrets Manager, Azure Key Vault, or HashiCorp Vault. Never written to disk in clear.

## 8 Validation (IQ / OQ / PQ)

Each release ships with an IQ/OQ documentation pack and OQ test scripts. S.G. Systems delivers the documentation; execution inside the Customer Environment is the customer's responsibility, with remote assistance available under an active Maintenance & Support term. The pack includes:

- Installation Qualification (IQ) checklist covering OS prerequisites, package versions, file permissions, and database schema verification.
- Operational Qualification (OQ) scripts covering authentication, role gating, signature workflow, audit-trail immutability, backup/restore drill, and ERP/MES integration handshakes (where licensed).
- Performance Qualification (PQ) template the customer completes against representative production traffic.
- Change-control procedure for applying minor and major upgrades, including revalidation guidance.

## 9 Backup, disaster recovery and high availability

- **Backups:** nightly full + 15-minute WAL archive to customer-controlled S3-compatible storage. Encrypted with customer KMS keys. Retention is customer-defined (typical: 35 days hot, 7 years cold for Part 11 records).
- **Recovery objectives (reference):** RPO ≤ 15 minutes with WAL shipping, RTO ≤ 4 hours for single-host restore, RTO ≤ 30 minutes on a HA topology with a warm standby.
- **HA:** Kubernetes deployment with ≥ 2 app replicas behind the customer ingress + Postgres in primary/replica (e.g. Patroni, RDS Multi-AZ, Crunchy PGO).
- **DR drills:** a documented restore-from-backup runbook ships with the install; an annual drill is recommended and supported.



S.G. Systems, LLC · Confidential — for customer evaluation

## 10 Observability and logging

- Structured JSON application logs to stdout — consumable by the customer's log stack (Splunk, Elastic, Datadog, Loki, CloudWatch, Azure Monitor).
- Prometheus `/metrics` endpoint for app and database health, queue depth, signature throughput, and integration error rates. Grafana dashboard JSON shipped with the release.
- Healthcheck endpoints (`/healthz`, `/readyz`) for load balancer and Kubernetes probes.
- No usage telemetry is sent to S.G. Systems from on-premises deployments. Annual licence self-attestation per the addendum, §9.

## 11 Integrations supported on-premises

- **ERP / MRP:** SAP S/4HANA & ECC (IDoc / OData / RFC), Microsoft Dynamics 365 BC & F&O, NetSuite, Oracle EBS & Fusion, Sage X3, Epicor, Infor CloudSuite, IFS, Acumatica, Odoo.
- **LIMS / QMS:** LabWare, LabVantage, STARLIMS, MasterControl, Veeva Vault, Greenlight Guru — via REST / SOAP / SFTP per the vendor's published interface.
- **Floor:** OPC-UA, Modbus TCP, MQTT for sensors and PLC tags; direct drivers for the supported scales, label printers, and barcode/RFID readers listed on our hardware page.
- **Identity:** SAML 2.0 + SCIM 2.0 against any compliant IdP, plus optional LDAP / Active Directory bind for legacy environments.

## 12 Support model

- Maintenance & Support is annually renewable. Within an active term: updates, patches, new minor and major releases, security patches, and regulatory updates to maintain conformity with the standards on the Order Form (e.g. 21 CFR Part 11, EU Annex 11).
- Response-time targets and severity definitions are set on the Order Form. Critical security patches are released within the timelines stated there.
- Source-code escrow is available at customer expense via a mutually agreed third-party agent (NCC Group, Iron Mountain) — see addendum §8.

## 13 Install options (self-install vs assisted)

V5 Ultimate on-prem can be installed by the customer's own IT team or with hands-on help from S.G. Systems. Both paths use the same binaries, the same Helm chart or Compose file, and the same IQ/OQ pack — the difference is who runs the keyboard and who signs off the OQ.

| Path | Who installs | Who runs IQ/OQ | Best for |
|---|---|---|---|
| **Self-install** | Customer IT, using the published Compose file or Helm chart and a licence key we issue. | Customer Quality executes and signs the OQ scripts; remote Q&A available under M&S. | Strong IT teams, non-GxP or low-risk GxP sites, sandbox / validation environments, pilots. |



S.G. Systems, LLC · Confidential — for customer evaluation

| Path | Who installs | Who runs IQ/OQ | Best for |
|---|---|---|---|
| **Assisted install** | S.G. Systems engineer drives the install in a shared session against the customer's hosts. | S.G. Systems executes OQ alongside Customer Quality, who countersigns. | First production site, regulated GxP (Part 11 / Annex 11), tight go-live deadlines. |
| **Managed handover** | Assisted install for production; self-install template for any additional sandbox / DR environments. | Mixed — production validated with us, secondary environments validated by the customer. | Multi-environment or multi-site rollouts where only the primary needs hands-on. |

**What customers get either way:** the signed Compose file or Helm chart, a per-environment licence key, the IQ/OQ pack and OQ test scripts, the restore-from-backup runbook, and the Grafana dashboard JSON.

## 14 Install runbook (for the setup engineer)

The ordered checklist the setup engineer works through on the day of install. It maps one-to-one onto the IQ section of the validation pack — tick each step in order, capture the evidence (screenshot, log line, or signed checkbox), and don't skip ahead. Each step should take minutes, not hours; if one is fighting you, stop and check the prerequisite step before brute-forcing through.

1 **Confirm server specs**
Verify CPU, RAM, disk and OS match the sized line in §3 for this site's concurrent operator count. Check NTP is configured and the host clock is within ±1 s of UTC — Part 11 audit trails depend on it.

2 **Install Docker or Podman**
Docker 24+ or Podman 4+ on the Linux host. Enable the service at boot. On Windows-hosted non-regulated sites, install Docker Desktop with the WSL2 backend and confirm the Linux distro is running.

3 **Install or connect PostgreSQL**
Use PostgreSQL 15 or 16. Either install locally, point at the customer's existing cluster, or wire up the managed service (RDS / Azure Database / Cloud SQL). Confirm TLS is enabled on the connection.

4 **Create the database and role**
Create a dedicated database and a non-superuser role owning it. Grant CONNECT, USAGE and the standard CRUD on the schema. Record the connection string for the env file — do not paste it into chat.

5 **Load the provided Compose file (or Helm chart)**
Drop the signed docker-compose.yml (or helm values.yaml) into /opt/v5 and pull the pinned image tags. Air-gapped sites: side-load the OCI tarball from the release bundle before this step.



S.G. Systems, LLC · Confidential — for customer evaluation

6 **Add the customer licence key**
Place the per-environment licence key issued with the Order Form into the secret store. The app refuses to start without a valid key and will warn 30 days before expiry.

7 **Configure environment variables and secrets**
Set DATABASE_URL, OBJECT_STORAGE_*, SMTP_*, SAML/SCIM endpoints, and the encryption key reference. Use Docker secrets, Kubernetes secrets, AWS Secrets Manager, Azure Key Vault or HashiCorp Vault — never a plain .env on disk in production.

8 **Configure HTTPS and the reverse proxy**
Terminate TLS at the customer's chosen proxy (NGINX, HAProxy, Traefik, F5, ALB, App Gateway). TLS 1.2 minimum, 1.3 recommended. Point the proxy at the app's internal port; restrict inbound to 443 only.

9 **Start the services**
Bring up the stack (compose up -d or helm upgrade --install). Watch logs for the migration banner — schema migrations run automatically on first boot and are idempotent on subsequent restarts.

10 **Confirm /healthz and /readyz pass**
Both endpoints must return 200 before traffic is cut over. /healthz proves the process is alive; /readyz proves the DB, object storage and licence checks all passed. Wire these into the load balancer probes.

11 **Log in through a browser**
Sign in as the bootstrap admin, confirm SSO redirects against the customer IdP, and check that the first SCIM sync brings in expected groups. Rotate or disable the bootstrap admin after the first real admin is provisioned.

12 **Test a kiosk / operator login**
Pull up the kiosk surface on a representative floor device. Confirm the operator-only role lands on the kiosk view with no admin UI exposed, and that the session-timeout matches the site SOP.

13 **Test one printer, scanner and scale (if applicable)**
Print a label, scan it back, and capture a weight from one scale per line. Confirms the floor VLAN and device drivers are wired correctly before go-live, not during it.

14 **Run a backup**
Trigger an on-demand backup to the customer's S3-compatible target. Verify the object lands encrypted with the customer KMS key and that the WAL archive is shipping on its 15-minute cadence.



S.G. Systems, LLC · Confidential — for customer evaluation

| | |
|---|---|
| 15 | **Run a restore test**<br>Restore the backup into a scratch database and confirm row counts and audit-trail hashes match. This is the single most important step — sign it off in the IQ/OQ pack before declaring the site live. |

On an assisted install an S.G. Systems engineer drives this checklist with your IT and Quality teams on a shared screen. On a self-install your team works through it independently, with remote Q&A available under M&S. Either way, the completed checklist becomes the IQ evidence in the validation pack.

## 15 Acceptance criteria

Deployment is deemed accepted on successful completion of the OQ scripts in the Customer Environment, an end-to-end signature workflow against a representative master record, and a successful restore-from-backup drill. The acceptance certificate template is included in the IQ/OQ pack.

## 16 What this document is not

- It is not a Statement of Work. The SOW is produced after the design workshop and forms part of the Order Form.
- It is not a commercial offer. Pricing and term are on the Order Form.
- It is not a certification. See the Security & Trust page on v5ultimate.com for V5 Ultimate's current security posture and the live certification roadmap.

## 17 Next steps

To request a sized Order Form, escrow agreement, or a design workshop for your environment, get in touch via **v5ultimate.com/resources/on-prem-specification** — we'll reply within one working day. The companion legal document is the On-Premises Licence Addendum.

**S.G. Systems, LLC** · V5 Ultimate · On-Premises Deployment Specification · v2026.06 · Last updated June 23, 2026
Maintained for sharing with customer IT, Quality, and procurement teams during evaluation. Read alongside the On-Premises Licence Addendum (commercial terms).



S.G. Systems, LLC · Confidential — for customer evaluation