CODE CREATOR

Advanced Analytics Server

powered by Apache Superset

Customer Launch and Administration Guide

AWS Marketplace AMI

Secure-by-default launch: unique credentials and SSH host keys are generated on each new instance.

This guide applies to the Code Creator AWS Marketplace version shown above. Commands must be run on the launched EC2 instance unless stated otherwise.

Quick facts

Item	Value
Web address	http://<EC2-public-IP-or-DNS>/ until HTTPS is configured
SSH user	**ubuntu**
Customer-facing ports	TCP 22 for SSH; TCP 80 for HTTP; TCP 443 after HTTPS is configured
Internal-only ports	8088 (Superset), 5432 (PostgreSQL), 6379 (Valkey)
First-login file	/home/ubuntu/FIRST_LOGIN.txt
Management command	sudo cc-superset <command>
Application directory	/opt/codecreator/superset
Runtime secrets	/etc/codecreator/superset/runtime.env — root only, mode 0600

Do not expose internal services

Never open TCP 5432, 6379, or 8088 to the internet.
Access Superset through Nginx on TCP 80 or through a production HTTPS endpoint on TCP 443.

1. Product overview and architecture

Code Creator Advanced Analytics Server powered by Apache Superset is a single-instance analytics appliance for Amazon EC2. It packages Apache Superset with a PostgreSQL metadata database, Valkey cache and task broker, Celery worker and scheduler, and Nginx reverse proxy.

What the appliance provides

- Browser-based analytics, dashboards, charts, SQL Lab, datasets, and role-based access controls.
- Database drivers for PostgreSQL, MySQL-compatible systems, Amazon Redshift, Amazon Athena, Snowflake, and Trino.
- Background task processing and scheduled report infrastructure using Celery and Valkey.
- Chromium/Playwright support for chart thumbnails and report rendering.
- A one-time customer initialization service that creates unique secrets and administrator credentials.
- A management CLI for status, health, restart, logs, version, and validation checks.

2. Launch prerequisites

Prepare these AWS resources

- An AWS account with permission to subscribe to AWS Marketplace products and launch EC2 instances.
- An EC2 key pair. AWS injects the selected public key into the ubuntu account during launch.
- A VPC and subnet with routing appropriate for your users and connected data sources.
- A security group that follows the inbound rules in this guide.
- Sufficient EBS capacity for the application, metadata, logs, and expected growth. Do not reduce the listing default root volume.

Reference instance sizing

The release was validated on a t3.xlarge instance with 4 vCPUs and approximately 16 GiB of memory. Actual capacity depends on concurrent users, query complexity, chart rendering, scheduled reports, and database latency. Validate smaller or larger instance types against your workload before production use.

Recommended initial security-group rules

Inbound rule	Recommended source
SSH — TCP 22	Source: your administrator public IP or a controlled bastion/security group. Never use 0.0.0.0/0 for routine administration.
HTTP — TCP 80	Temporary evaluation: trusted user IP ranges only. Production: allow only from the Application Load Balancer security group, or replace with direct HTTPS.
HTTPS — TCP 443	Open to approved user networks when HTTPS termination is configured.
All other inbound ports	Do not add rules for 5432, 6379, or 8088.

3. Launch from AWS Marketplace

1. Open the existing AWS Marketplace product listing
2. Accept the product terms and continue to the EC2 launch configuration.
3. Select the target Region, VPC, subnet, instance type, and root EBS settings.
4. Select the security group prepared in Section 2.
5. Select the EC2 key pair that administrators will use for SSH access.
6. Launch the instance and record its instance ID, private IP, and public IPv4 address or public DNS name, if assigned.

> **Public addresses can change**
>
> A public IPv4 address assigned automatically by EC2 may change after a stop/start cycle. Use an Elastic IP or DNS record when a stable endpoint is required.

Connect by SSH

Linux or macOS:

```
chmod 400 /path/to/key.pem
ssh -i /path/to/key.pem ubuntu@<public-ip-or-dns>
```

Windows PuTTY: Use the public IP or DNS name, user **ubuntu**, and the private key corresponding to the EC2 key pair selected at launch. Convert a PEM key to PPK only when required by the installed PuTTY version.

4. First boot and administrator sign-in

On a new customer instance, cloud-init injects the selected EC2 public key and generates new SSH host keys. After cloud-init finishes, the Code Creator first-boot service creates unique PostgreSQL, Valkey, Superset, and administrator secrets; initializes the metadata database; starts the application stack; and writes the first-login file. Be Patient as it initiates.

Check first-boot status: Run these preflight commands:

```
sudo systemctl status codecreator-superset-firstboot.service --no-pager --full
```

Successful completion is shown as **active (exited)**. The completion log contains FIRSTBOOT_RESULT=PASS.

```
sudo grep -E 'FIRSTBOOT_STARTED|FIRSTBOOT_COMPLETED|FIRSTBOOT_RESULT' \
  /var/log/codecreator/superset/firstboot.log
```

Retrieve the generated credentials

```
cat /home/ubuntu/FIRST_LOGIN.txt
```

> **Protect the first-login file**
> It contains the initial administrator password. Do not email it, paste it into tickets, place it in shared storage, or include it in an AMI or snapshot intended for another party.

Sign in

1. Open the Web address shown in **FIRST_LOGIN.txt**
2. Sign in with the generated administrator username and password.
3. Change the administrator password immediately through Superset user administration.
4. Create named administrator accounts for routine use and reserve the original admin account for controlled recovery.
5. Confirm that the welcome page, SQL Lab, datasets, charts, and dashboards are accessible as expected.

ADDITIONAL & HELPFUL INFORMATION

5. Security and production deployment

Required security actions

- Change the generated administrator password and create named user accounts.
- Restrict SSH to approved administrator addresses or a bastion host.
- Configure HTTPS before transmitting production data or credentials over untrusted networks.
- Keep PostgreSQL, Valkey, and Superset application ports internal-only.
- Use least-privilege database accounts for every connected data source.
- Apply AWS IAM policies, security groups, routing, and network ACLs appropriate to the data being accessed.
- Take a recovery snapshot before significant configuration changes or upgrades.

Recommended HTTPS architecture Optional

For AWS production deployments, the preferred pattern is:

```
Users --HTTPS/443--> Application Load Balancer + ACM certificate
                          |
                          +--HTTP/80--> EC2 security group --> Nginx --> Superset
```

- Attach an AWS Certificate Manager certificate to the load balancer HTTPS listener.
- Redirect HTTP to HTTPS at the load balancer.
- Allow EC2 TCP 80 only from the load balancer security group.
- Do not expose the EC2 instance directly to the internet when the load balancer is the public endpoint.
- After HTTPS is working, enable secure session-cookie behavior in the application configuration and restart the appliance.

HTTP is for controlled evaluation only

The default first-login URL uses HTTP. Treat it as an initial access path on a restricted network, not as a finished production endpoint.

Secrets and sensitive files

Asset	Handling requirement
/etc/codecreator/superset/runtime.env	Generated service secrets and initial administrator credential. Root-owned, mode 0600.
/home/ubuntu/FIRST_LOGIN.txt	Initial access instructions and administrator password. Owned by ubuntu, mode 0600.
/opt/codecreator/superset/config/superset_config.py	Application configuration. Do not place plaintext production credentials here unless the deployment has an approved secret-management process.
EC2 key pair private key	Remains with the customer. The AMI contains no seller SSH private key.

Host firewall and network exposure

The primary network control is the EC2 security group. The appliance publishes only TCP 80 on the host for the application. Docker-internal services remain on isolated networks. Do not add host port mappings or security-group rules for the internal ports.

6. Daily operations and management commands

Command	Purpose
sudo cc-superset status	Show the Compose service state.
sudo cc-superset health	Check Nginx and application HTTP health.
sudo cc-superset validate	Run the complete product validation suite.
sudo cc-superset restart	Restart the application stack.
sudo cc-superset start	Start the application stack.
sudo cc-superset stop	Stop the application stack.
sudo cc-superset logs 100	Show the latest 100 log lines.
sudo cc-superset version	Show product, host, and Superset version information.

Common operating checks

```
df -h /
free -h
sudo docker compose -f /opt/codecreator/superset/docker-compose.yml ps
sudo systemctl --failed --no-pager
```

Restart after configuration changes

```
sudo cc-superset restart
sudo cc-superset validate
```

Do not run first boot manually on an initialized system

The first-boot service is protected by a completion marker. Do not delete the marker or runtime.env on a live customer system unless intentionally rebuilding the appliance from a clean state.

7. Data persistence, backup, and recovery

What must be protected

- Superset metadata: users, roles, dashboards, charts, datasets, saved queries, and database connection definitions.
- The Superset secret key in runtime.env, which is required to decrypt stored connection credentials.
- Application configuration and any customer-created certificates or reverse-proxy files.
- The root EBS volume containing Docker named volumes and the Superset home directory.

Back up the secret key with the database

A PostgreSQL backup without the matching Superset secret key may leave stored database passwords undecryptable. A full root-volume snapshot captures both items together.

Recommended full-appliance backup

1. Schedule an approved maintenance window.
2. Confirm the product is healthy with sudo cc-superset validate.
3. Stop the application with sudo cc-superset stop, or stop the EC2 instance for the strongest consistency boundary.
4. Create an EBS snapshot of the root volume. Record the instance ID, product version, snapshot ID, and backup date.
5. Start the instance or application again and run sudo cc-superset validate.
6. Retain snapshots according to organizational recovery and data-retention policies.

Recovery approach

For full-instance recovery, create a replacement EBS volume or AMI from an approved snapshot and launch it in a controlled environment. Confirm security groups, IAM roles, DNS, HTTPS certificates, and database network access before returning the restored appliance to users.

Before terminating an instance

- Create and verify a current root-volume snapshot.
- Export any required dashboards or configuration records according to your governance process.
- Confirm whether the root volume is configured for deletion on termination.
- Record any Elastic IP, DNS, load balancer, IAM role, and security-group dependencies.

8. Connecting data sources and using Superset

Supported driver families in this release

Driver	Use
PostgreSQL	PostgreSQL-compatible databases and services.
MySQL	MySQL-compatible databases, including common managed-service configurations.
Amazon Redshift	Redshift SQLAlchemy dialect using the PostgreSQL-compatible driver.
Amazon Athena	Athena connectivity through PyAthena; requires AWS access and an S3 query-results location.
Snowflake	Snowflake SQLAlchemy connectivity.
Trino	Trino SQLAlchemy connectivity.

Connection checklist

- Confirm network routing, DNS resolution, and security-group rules from the EC2 subnet to the database endpoint.
- Create a dedicated, least-privilege database account for Superset.
- Require TLS for database connections whenever the target supports it.
- Test the connection before saving it, then grant access only to the intended Superset roles.
- Avoid enabling DML, file upload, or advanced SQL features unless the business requirement and risk have been reviewed.

Typical workflow

1. Add and test a database connection in Superset.
2. Open SQL Lab and validate a read-only query.
3. Create or register a dataset.
4. Build a chart and save it to a dashboard.
5. Assign dashboard and dataset permissions to the required roles.

Alerts and reports

The worker, scheduler, and browser-rendering components are installed and active. Email delivery is not ready until the customer configures SMTP and disables report dry-run behavior through an approved secret-management process. Validate sender identity, recipients, schedules, and screenshot rendering before production use.

9. Maintenance

Ubuntu security updates

```
sudo apt-get update
sudo DEBIAN_FRONTEND=noninteractive apt-get -y full-upgrade
sudo apt-get -y autoremove
sudo apt-get clean
test -f /var/run/reboot-required && cat /var/run/reboot-required.pkgs || true
```

Create a current recovery snapshot before applying significant updates. If a reboot is required, reboot during a maintenance window and run sudo cc-superset validate after the instance returns.

Routine capacity checks

- Monitor root-volume free space and expand the EBS volume before capacity becomes critical.
- Review memory pressure and CPU utilization during peak queries and report rendering.
- Review application logs for repeated database timeouts, failed tasks, authentication errors, or browser-rendering failures.
- Use CloudWatch metrics and alarms appropriate to the instance and business workload.

10. Troubleshooting

Symptom	Action
Website does not open	Run sudo cc-superset health and sudo cc-superset status. Confirm TCP 80/443 rules, route tables, public/Elastic IP assignment, load balancer health, and DNS.
FIRST_LOGIN.txt is missing	Check sudo systemctl status codecreator-superset-firstboot.service and the first-boot log. The file is created only after initialization completes.
First boot failed	Read sudo tail -n 200 /var/log/codecreator/superset/firstboot.log. Do not paste runtime.env or the generated password into a support ticket.
Container is unhealthy	Run sudo cc-superset logs 200, then sudo cc-superset restart and sudo cc-superset validate.
Database connection fails	Check endpoint, port, DNS, route, TLS settings, database credentials, source security group, and least-privilege grants.
Charts or reports fail to render	Check worker and beat status, available memory, browser-related log errors, and destination/report configuration.
Disk is nearly full	Inspect df -h / and Docker disk usage. Expand the EBS volume or remove approved obsolete data; do not delete named volumes blindly.

Useful diagnostic commands

```
sudo cc-superset status
sudo cc-superset health
sudo cc-superset logs 200
sudo cc-superset validate
sudo systemctl --failed --no-pager
sudo journalctl -u codecreator-superset-firstboot.service --no-pager
df -h /
free -h
```

Informational log message

“fastmcp is not installed, skipping MCP initialization” is expected in this release. MCP support is not included and the message does not indicate a Superset startup failure.

Collect information for support

- AWS Region, instance ID, instance type, and product version.
- The failing user action and the approximate event time with timezone.
- Output from cc-superset status, health, and validate.
- Relevant log excerpts with passwords, secret keys, database URLs, tokens, and private data removed.
- Recent infrastructure or configuration changes.

11. Support handoff and reference information

Support channel

Use the seller support contact and support terms shown on the AWS Marketplace product listing. AWS account, billing, EC2 platform, and Marketplace subscription questions may require AWS Support. Product configuration and appliance behavior should be directed to the seller support channel. See: https://superset.apache.org/developer-docs/

Key paths

Path	Purpose
/opt/codecreator/superset	Product files and Docker Compose definition.
/opt/codecreator/superset/docker-compose.yml	Container service definition.
/opt/codecreator/superset/config/superset_config.py	Superset application configuration.
/etc/codecreator/superset/runtime.env	Generated runtime values and secrets.
/etc/codecreator/superset/nginx/conf.d/superset.conf	Nginx application reverse-proxy configuration.
/var/lib/codecreator/superset/home	Superset home and application-managed files.
/var/log/codecreator/superset	First-boot and product logs.
/home/ubuntu/FIRST_LOGIN.txt	Generated initial access information.
/usr/local/sbin/cc-superset	Management CLI.
/usr/local/sbin/cc-superset-firstboot	One-time customer initialization script.