LOGZILLA DOCUMENTATION

Kubernetes Scaling and Tuning

Scale LogZilla Kubernetes StatefulSets for ingest, storage, and query workloads, tune CPU and memory requests, and align SM_API_ADDRESSES ordinals

Kubernetes Deployment Overview · Generated May 3, 2026 · logzilla.ai/docs/kubernetes-deployment/scaling-and-tuning

Replica scaling

Increase or decrease replicas for StatefulSets to match ingest and query load:

```
# Ingest (syslog receivers, parser, http receiver)
kubectl scale statefulset ingest --replicas=8

# Storage (ensure ordinal references align, e.g., storage-0..N)
kubectl scale statefulset storage --replicas=7

# Query module (usually scale with storage)
kubectl scale statefulset querymodule --replicas=3
```

When scaling `storage`, review ordinal references used by other modules:

- `SM_API_ADDRESSES` in Query Module (e.g., `http://storage-{{0-4}}.storage:81`).
- `SM_INGEST_URLS` in Ingest ParserModule (e.g., `http://storage-{{1-4}}.storage:81/ingest`).

Adjust these ranges to reflect the desired replica counts.

Requests and limits

Manifests define CPU and memory requests/limits for each container. Increase requests to provide scheduling guarantees; adjust limits to prevent eviction. Observe actual usage via the platform metrics (e.g., `kubectl top pods`).

Guidance by component:

- Ingest `syslogng`: raise `--worker-threads` and CPU/memory for higher EPS.
- Ingest `parsermodule`: increase `PARSER_WORKERS` and resources as needed.
- API `gunicorn`: tune `--workers` (default 10 in the manifest command) and allocate CPU accordingly.
- Celery worker: adjust autoscale args and CPU/memory.
- Storage/InfluxDB: ensure disk performance and memory are adequate; consider faster StorageClass for production.

Rolling updates

```
# Check rollout status
kubectl rollout status statefulset/api
```

```
# Restart pods to pick up new config or image tag
kubectl rollout restart statefulset/api
```

Readiness and liveness probes

Probes are defined in all manifests. If a pod flaps between Ready/NotReady:

- Describe the pod and review the probe configuration and last failure:

```
kubectl describe pod <pod-name>
```

- Examine container logs:

```
kubectl logs <pod-name> -c <container-name>
```

StorageClass and PVC sizing

- Replace `storageClassName` with the provider's class as needed.
- Right-size PVC requests (e.g., `sm-data`, `sm-archives`, `influxdb-data`).

Ingress and external exposure

- For GKE, use the provided Ingress example and NEG annotations.
- For other providers, adapt annotations and class to the installed controller.

Image tags

- Use the `stable` tag by default for production deployments.
- For controlled rollouts, pin to a specific release tag (for example, `vX.Y.Z`) and update via `kubectl set image` or `kubectl rollout restart`.