**Total Suggested VM Count: 10-11 VMs**

**Breakdown**

1. **Control Plane Nodes** (3 VMs):
   * Each with ETCD and Kubernetes API server, scheduler, and controller-manager components.
   * Purpose: Ensures high availability for the Kubernetes control plane.
   * Specs: 4 vCPUs, 8 GB RAM, 50 GB SSD.
2. **Worker Nodes** (3 VMs, scalable based on workload needs):
   * These will run the application pods, handling workloads from the Kubernetes scheduler.
   * Purpose: Dedicated to running application pods and scaling as required by workloads.
   * Specs: 4 vCPUs, 8 GB RAM, 50 GB SSD (can scale based on application demand).
3. **Monitoring Server** (1 VM):
   * Runs Prometheus, Grafana, and Alertmanager to monitor the Kubernetes cluster and application workloads.
   * Purpose: Observability for metrics and alerting; can include advanced monitoring like Node Exporter.
   * Specs: 4 vCPUs, 8 GB RAM, 50 GB SSD.
4. **Logging and Tracing Server** (1 VM):
   * Runs Loki, Promtail, and Jaeger for centralized logging and tracing.
   * Purpose: Aggregates logs and enables tracing for troubleshooting across the cluster.
   * Specs: 4 vCPUs, 8 GB RAM, 50 GB SSD.
5. **GitLab CE Server** (1 VM):
   * Provides CI/CD pipelines and can integrate with Vault for secrets management.
   * Purpose: Manages CI/CD for application deployment to Kubernetes.
   * Specs: 4 vCPUs, 16 GB RAM, 100 GB SSD (storage may need to scale with CI/CD demands).
6. **Secrets Management Server (Vault)** (1 VM):
   * Runs HashiCorp Vault to securely manage sensitive information for applications and CI/CD workflows.
   * Purpose: Centralized secrets management for secure access to sensitive data.
   * Specs: 4 vCPUs, 8 GB RAM, 50 GB SSD.

**Optional VM for Backup Storage and Disaster Recovery** (1 VM):

* Purpose: Stores backups from ETCD and Velero; ideally has expanded storage and high I/O performance.
* Specs: 2 vCPUs, 4 GB RAM, scalable storage (500 GB+ SSD or network-attached storage).

Summary

* Core Setup: 10 VMs (3 control planes, 3 workers, 1 monitoring, 1 logging, 1 GitLab CE, 1 Vault).
* Optional for Backup: 1 additional VM for off-cluster storage.

This layout ensures redundancy, high availability, security, and observability, scalable to production needs.