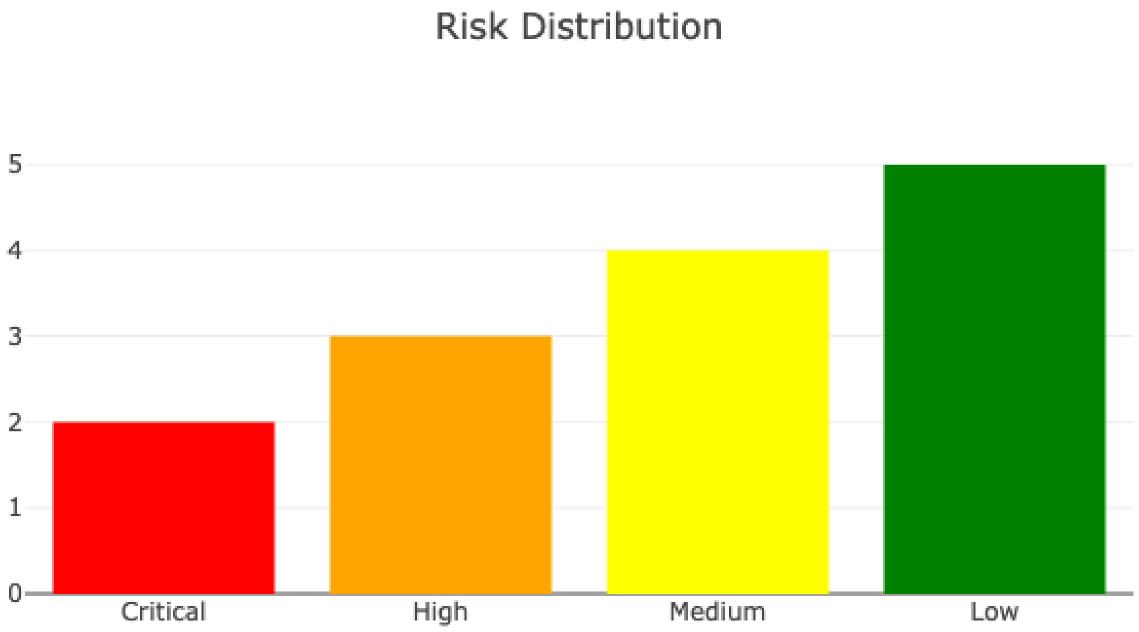
EKS Operational Review Report

Generated on: 2025-04-22 12:44:03

Executive Summary

This report provides a comprehensive review of the EKS cluster operations and identifies key areas for improvement. The assessment covers cluster health, cost optimization, security, monitoring, CI/CD, and other critical aspects of the EKS infrastructure.

Risk Assessment Overview



Cluster Health

Cluster Status:

EKS cluster is running in us-west-2 with 3 availability zones. Current control plane health is good with no reported issues. API server response time averages 200ms. etcd cluster is healthy with no leader elections in past 30 days.

Node Health:

Running 15 nodes across 3 node groups:

- 8 x m5.2xlarge (Production workloads)
- 4 x c5.xlarge (Batch processing)
- 3 x t3.large (Development workloads)

2 nodes reported kubelet connectivity issues last week. Memory pressure observed on 3 production nodes during peak hours.

Pod Scheduling Issues:

- 15% pods experiencing scheduling delays due to resource constraints
- 5 pods stuck in Pending state due to PersistentVolume binding issues
- Occasional pod evictions observed due to node memory pressure
- Resource quotas hitting limits during deployment peaks

Cost Optimization

Resource Utilization:

- Average CPU utilization: 45%

- Average Memory utilization: 78%

- 30% of PersistentVolumes underutilized

- Identified 5 idle EBS volumes

- Spot instances not currently utilized

Cost Allocation:

- Monthly EKS costs: $2,500

- EC2 instances: $8,000/month

- EBS volumes: $800/month

- No cost allocation tags implemented

- Missing chargeback mechanism for teams

Optimization Opportunities:

- Right-sizing potential for 6 nodes

- Spot instance adoption possible for non-critical workloads

- Implement automatic scaling for dev environments

- Storage class optimization needed

- Consider Graviton instances for cost reduction

Security

IAM Configuration:

- IRSA (IAM Roles for Service Accounts) partially implemented

- 5 shared IAM roles identified

- Pod security policies not enforced

- Root account access detected in audit logs

- AWS Security Hub integration missing

Secret Management:

- Using AWS Secrets Manager for 60% of secrets

- Some secrets still in plain ConfigMaps

- External Secrets Operator not implemented

- No secret rotation policy

- Key management using AWS KMS

Network Policies:

- Default deny policies missing

- No microsegmentation implemented

- Calico network policies partially configured

- Public endpoints exposed without WAF

- Security groups need tightening

Monitoring

Monitoring Tools:

- Prometheus/Grafana stack deployed

- AWS CloudWatch Container Insights enabled

- X-Ray tracing implemented for 40% of services

- Custom metrics pipeline using Prometheus Operator

- Logging via EFK stack

Alert Configuration:

- Node-level alerts configured

- Pod-level resource alerts active

- Missing alerts for PV capacity

- SLO/SLI monitoring needed

- No alert correlation system

Metric Collection:

- Custom metrics for business KPIs

- Standard kubernetes metrics collected

- Missing some network flow metrics

- Retention period: 15 days

- Storage optimization needed

CI/CD

Pipeline Setup:

- GitLab CI/CD with ArgoCD
- Image scanning with Trivy
- Automated testing coverage: 75%
- Manual approval gates for production
- Jenkins legacy pipelines still active

Deployment Strategy:

- Mix of rolling updates and blue/green
- No canary deployments implemented
- Average deployment frequency: 8/day
- MTTR (Mean Time to Recovery): 45 mins
- Change failure rate: 12%

Rollback Process:

- Manual rollback procedures
- No automated rollback triggers
- Average rollback time: 15 minutes
- Version control for all deployments
- Missing automatic health checks

Others

EKS Version:

- Currently on EKS 1.24

- Planning upgrade to 1.27

- Add-ons require updates

- Custom admission controllers need compatibility testing

- CNI version: 1.12.0

Cluster Architecture:

- Multi-AZ deployment

- Private networking with VPC endpoints

- Transit Gateway integration

- Direct Connect hybrid connectivity

- Running on EC2 with managed node groups

Special Requirements:

- PCI compliance requirements

- 99.99% uptime SLA

- DR RPO: 15 minutes

- DR RTO: 4 hours

- GPU nodes needed for ML workloads

Best Practices & References

Security

https://docs.aws.amazon.com/eks/latest/userguide/security.html

AWS EKS Security Best Practices

Cost Optimization

https://aws.amazon.com/blogs/containers/cost-optimization-for-kubernetes-on-aws/

Cost Optimization for Kubernetes on AWS

Operations

https://aws.github.io/aws-eks-best-practices/

EKS Best Practices Guide

Networking

https://docs.aws.amazon.com/eks/latest/userguide/network_reqs.html

EKS Networking Best Practices

Recommendations

Short Term (3 months)

[High] - Implement automated node health checks

[High] - Configure cluster autoscaling

[Critical] - Enable container vulnerability scanning

Medium Term (6 months)

[Medium] - Implement GitOps practices

[High] - Set up cross-region disaster recovery

[Medium] - Implement cost allocation tags

Long Term (>6 months)

[Medium] - Migrate to newer EKS version

[Low] - Implement service mesh

[Medium] - Set up multi-cluster management