

KUBERNETES COST OPTIMIZATION PLAYBOOK

Advanced Techniques for Container Cost Management Cluster Autoscaling, Pod Rightsizing & Multi-Tenancy

EXECUTIVE SUMMARY

Kubernetes deployments often waste 60-70% of allocated resources due to poor configuration, overprovisioning, and lack of optimization. This comprehensive playbook provides advanced techniques to dramatically reduce Kubernetes costs while improving performance and reliability.

KEY OPTIMIZATION AREAS:

- â€¢ Cluster autoscaling and node optimization - 40-60% savings
- â€¢ Pod rightsizing and resource allocation - 30-50% savings
- â€¢ Multi-tenant cost allocation and governance - 20-40% savings
- â€¢ Storage and network optimization - 15-30% savings

TABLE OF CONTENTS

Chapter 1: Kubernetes Cost Fundamentals.....	Pages 3-8
Chapter 2: Cluster Autoscaling Strategy.....	Pages 9-15
Chapter 3: Pod Optimization & Rightsizing.....	Pages 16-22
Chapter 4: Resource Allocation & Limits.....	Pages 23-28
Chapter 5: Multi-Tenant Cost Management.....	Pages 29-33
Chapter 6: Monitoring & Cost Visibility.....	Pages 34-38

OPTIMIZATION STRATEGIES

1. Horizontal Pod Autoscaler (HPA) Configuration
 - Right-size pods based on actual resource consumption
2. Vertical Pod Autoscaler (VPA) Implementation
 - Automatically adjust CPU/memory requests and limits
3. Cluster Autoscaler Optimization
 - Scale nodes efficiently based on workload demands
4. Spot Instance Integration
 - Leverage spot instances for non-critical workloads
5. Storage Class Optimization
 - Use appropriate storage tiers for different workloads

RESOURCE MANAGEMENT BEST PRACTICES

- â€¢ Set appropriate CPU/memory requests and limits