EN

# FinOps Guide

## Cloud Cost Management for Canadian SMBs

Inform · Optimize · Operate — BOTUM case -47% in 60 days

**Inform** · **Optimize** · **Operate**

Mars 2026

# FinOps: Cloud Cost Management for Canadian SMBs

The scene repeats in almost every SMB migrating to the cloud. "We're finally going to save money on our infrastructure!" Then the first bill arrives. It's three times higher than expected. FinOps is the structural answer to this problem.

## What Is FinOps?

FinOps isn't a tool. It's a cultural practice that brings Finance, DevOps and Product together around a shared goal: aligning every cloud dollar spent with the business value it generates. The FinOps Foundation: "Enabling organizations to get maximum business value by helping engineering, finance and technology teams to collaborate on data-driven spending decisions."

### In practice:

- Engineers know what their code costs in production
- Finance understands why a cost spike is normal during a product launch
- Capacity decisions are based on real data, not estimates
- Everyone shares responsibility for spend

## The 3 FinOps Lifecycle Phases

### Phase 1 — Inform: Total Visibility

Before optimizing, you need to see where the money goes. Activate tagging (project, environment, team, owner), configure cost allocation reports, create budgets per team and set up alerts at 50%, 80%, and 100% of monthly budget.

### Minimum tagging taxonomy:

```
Project = "crm-v2" # Project name
Environment = "production" # dev / staging / production
Team = "backend" # Responsible team
Owner = "m.tremblay" # Individual owner
CostCenter = "IT-PROD-001" # Accounting cost center
```

### Phase 2 — Optimize: Concrete Actions

Once visibility is established, identify and implement savings: rightsizing oversized instances, purchasing Reserved Instances for stable workloads, migrating to Spot for batch workloads, deleting orphaned resources, aggressive auto-scaling.

### Phase 3 — Operate: Continuous Governance

Optimization isn't a project with an end date. Establish monthly reviews (30 min), auto-shutdown policies for dev/staging, and drift alerts when spend exceeds a baseline.

## FinOps Tools

AWS Cost Explorer: Visualization by service, tag, region. Automatic Rightsizing Recommendations. Cost: $0.01 USD per API request.

Azure Cost Management + Billing: Built into the Azure portal. Budgets by Resource Group, email/webhook alerts. Azure Advisor provides rightsizing recommendations directly in the console.

GCP Billing Reports: Native BigQuery dashboards with automatic export. Custom SQL queries on cost history — powerful for analyzing patterns over 12 months.

Infracost (open source): Integrates into your CI/CD pipeline. On every Terraform PR, calculates the cost delta and posts it as a comment. Engineers see the cost before merging.

OpenCost (CNCF): For Kubernetes environments. Allocates costs per namespace, deployment, label. Know what each microservice actually costs.

## The 5 Optimization Levers

### Reserved Instances and Savings Plans (-30% to -40%)

If your workloads have been stable for 3+ months, Reserved Instances offer up to 40% reduction for a 1 or 3-year commitment. m5.xlarge On-Demand: ~$0.192/h. Reserved 1-year No Upfront: ~$0.118/h. Over 8,760 hours = $645 in savings per instance.

### Rightsizing (-20% to -30%)

Most cloud instances are oversized. Measure P95 CPU/RAM utilization over 14 days. If your m5.2xlarge rarely exceeds 40% CPU at P95, an m5.xlarge will do — at half the price. AWS Compute Optimizer and Azure Advisor do this automatically.

### Spot and Preemptible Instances (-60% to -75%)

60-75% cheaper using excess datacenter capacity, but can be reclaimed with 2 minutes notice. Ideal for: CI/CD pipelines, batch processing, ML training, video rendering.

### Dynamic Auto-Scaling (-25% to -35%)

Dev/staging environments scale to zero at night (10 PM–8 AM) and on weekends. For 5 environments on AWS: 65% savings on those instances. On Kubernetes: Cluster Autoscaler + KEDA to scale unused namespaces to zero.

### Deleting Orphaned Resources (-10% to -20%)

Unattached Elastic IPs ($0.005/h for nothing), accumulated EBS snapshots, empty load balancers. A monthly scan with AWS Trusted Advisor or cloud-nuke reveals 5-15% unused resources in any account active for 6+ months.

## Comparison: Reserved vs On-Demand vs Spot

| Type | Savings | Commitment | Use Case | Risk |
|---|---|---|---|---|
| On-Demand | Baseline (0%) | None | Variable loads | Max cost |
| Reserved 1yr | -30% to -37% | 1 year flexible | Stable production | Low |
| Reserved 3yr | -40% to -60% | 3 years fixed | Core stable infra | Medium |
| Savings Plans | -25% to -40% | 1 or 3 years | Flexible family | Low |
| Spot/Preemptible | -60% to -75% | None | Batch, CI/CD, ML | Interruption |

## FinOps Governance for SMBs

The SMB FinOps champion: a single senior DevOps engineer dedicating 20% of their time (~8 hours/week) is sufficient for a cloud account up to $50,000 CAD/month. Enforce tags via Azure Policy or AWS SCP: no resource can be created without tags.

## Mandatory budget alerts:

- 50% of monthly budget → information
- 80% of monthly budget → investigation required
- 100% → immediate action, Slack + management email notification

## Common Mistakes That Cost You

### Paying On-Demand by default

Once a workload has been running in production for 3 months, evaluate Reserved Instances. A startup with stable production for 6 months is leaving 30-40% of its bill on the table.

### Dev environments running 24/7

A dev environment running at night and on weekends does nothing. Implement Instance Scheduler (AWS) or Azure Auto-shutdown. Savings: 60-70%.

### Accumulating snapshots

EBS snapshot: $0.05/GB/month. A 500 GB DB over 2 years = 50 snapshots = $1,250/month. Define a retention policy: 7-day daily, 4-week weekly, 12-month monthly.

### Ignoring egress fees

Outbound data transfer is billed ($0.08-0.09/GB on AWS). Can represent 15-25% of total bill. Use CloudFront or Azure CDN to cache assets and enable gzip/brotli compression.

## BOTUM Real Case: -47% in 60 Days

B2B services SMB, 28 employees, AWS stack (EC2, RDS, S3, ALB). Initial monthly bill: $8,400 USD/month.

### Initial audit (weeks 1-2):

- 6 m5.xlarge EC2 instances in production at 8-15% CPU on average
- No Reserved Instances — everything On-Demand for 14 months
- 3 dev/staging environments active 24/7 (weekends included)
- 847 EBS snapshots accumulated over 2 years, representing 12 TB
- 4 unattached Elastic IPs, 2 empty load balancers, 1 unused NAT Gateway
- No consistent tagging

### Actions implemented (weeks 3-8):

- EC2 Rightsizing: 6× m5.xlarge → 6× t3.large → -$1,200 USD/month
- Reserved Instances 1-year No Upfront (6 EC2 + 2 RDS) → -$1,650 USD/month
- Deleted 843 snapshots + zombies → -$890 USD/month
- Dev/staging auto-shutdown (10 PM–8 AM + weekends) → -$620 USD/month
- Complete tagging via SCP — new visibility baseline

Result at 60 days: $8,400 → $4,452 USD/month. Savings: $3,948 USD/month (-47%). That's $47,400 over 12 months.

## Where to Start This Week

1. Enable Cost Explorer rightsizing recommendations (AWS) or Azure Advisor — savings list in 48h
2. Inventory orphaned resources: IPs, snapshots >90 days, LBs with no targets — delete them
3. Configure a budget alert at 80% of current monthly spend — never surprised again

---

## Go Further with BOTUM

FinOps audit, cloud optimization, bill reduction — BOTUM teams support Canadian SMBs.

**Discuss your project → www.botum.ca/contact**

## Complete PDF Guide

⬇ Download the EN guide (PDF) → botum.ca/guides/