

FINOPS · 90 DIAS

FinOps em 90 dias

Diagnóstico, ações rápidas e baseline de governança para reduzir custo cloud sem comprar compromisso errado.

Redgator · v1 · 2026-05

Cloud fica caro em silêncio. Uma instância nasce grande demais, um snapshot fica esquecido, um NAT gateway vira padrão, uma compra de Savings Plan tenta compensar uma base ainda inchada. Em três meses, a fatura deixa de ser assunto de engenharia e vira pergunta de board. Este guia mostra como a Redgator organiza um engajamento FinOps de 90 dias contra a cloud que você já tem. Não é uma migração. Não é um dashboard mágico. É uma sequência técnica: inventário, baseline, quick wins e governança.

1. Por que 90 dias?

Trinta dias é bom para diagnóstico. Raramente é suficiente para mudar custo com segurança. Você consegue olhar Cost Explorer, listar os 10 maiores gastos e achar desperdício óbvio. Mas ainda não valida risco de produção, não conversa com todos os owners e não mede se a economia apareceu depois da mudança.

Seis meses, por outro lado, costuma virar programa grande demais. A pauta expande para migração, replatform, multi-cloud, governança corporativa e ferramenta nova. O custo pode cair, mas a organização perde o foco do que precisava responder primeiro: **onde existe desperdício seguro de remover agora e qual baseline devemos governar daqui para frente?**

Noventa dias é o meio útil.

- **Até o dia 30:** a conta deixa de ser um bloco único. Existe inventário, dono, dashboard simples e política mínima de tags.
- **Até o dia 60:** os quick wins com evidência entram em produção em ondas controladas.
- **Até o dia 90:** budget alerts, anomaly detection, revisão mensal e handoff estão operando sem depender da Redgator para cada decisão.

A régua é simples: nenhuma economia vale se aumenta risco operacional sem dono. O corte certo é aquele que vem com evidência, rollback e owner de negócio.

O mapa dos 90 dias

Pense no programa como uma sequência de decisões, não como uma lista de tarefas.

Dias 0-7 — descobrir a fatura real

A pergunta é: **onde o dinheiro sai hoje e quem pode decidir sobre cada linha?** A saída é inventário, top 10 custos, cobertura de tags, compromissos existentes e riscos de produção.

Dias 8-30 — transformar custo em baseline

A pergunta é: **qual parte da fatura já tem dono, métrica e política?** A saída é dashboard de uma página, regra mínima de tags, showback/chargeback decidido e fila curta de oportunidades.

Dias 31-60 — aplicar mudanças reversíveis

A pergunta é: **o que conseguimos reduzir sem apostar a produção?** A saída é rightsizing, limpeza, lifecycle e revisão de commitments em ondas com evidência.

Dias 61-90 — deixar o rito operando

A pergunta é: **quem continua olhando quando a Redgator sai?** A saída é alerta, revisão mensal, handoff, backlog e responsabilidades claras.

Se uma etapa não tem owner, ela não avança. FinOps sem dono vira relatório de custo. Relatório não reduz fatura.

2. Day 0 — inventário da primeira semana

A primeira semana não começa com recomendação. Começa com inventário. Se a base está errada, a recomendação só parece precisa.

O que auditamos

- **Estrutura de contas e billing:** contas, subscriptions, compartments, organizations, cost centers, unidades de negócio, contratos e moeda de referência.
- **Higiene de tags:** cobertura por ambiente, aplicação, owner, centro de custo e criticidade. Tag que não bloqueia provisionamento vira sugestão, não governança.
- **Top 10 linhas de custo:** serviço, conta, região, tendência de 90 dias, variação semanal e owner técnico quando existe.
- **Reserved Instances / Savings Plans / compromissos:** cobertura, utilização, vencimento, compra por família, região e escopo.
- **Recursos ociosos:** compute parado, volumes desanexados, snapshots antigos, IPs elásticos, load balancers sem target, clusters de teste.
- **Pegada de data egress:** NAT gateways, VPC endpoints, inter-region transfer, tráfego para internet, tráfego cross-AZ e rotas recorrentes.
- **Risco operacional:** workloads que não podem sofrer resize simples, janelas de mudança, dependências de banco, picos sazonais e owners ausentes.

O que sai da semana 1

Uma lista priorizada por **economia provável × risco × facilidade de validar**. Não é backlog infinito. É uma fila curta para decidir o que entra nos primeiros 60 dias.

3. Day 30 — baseline que dá para governar

No dia 30, o cliente deve conseguir explicar a própria fatura sem abrir 15 abas. O objetivo não é perfeição. É uma baseline que permita decisão.

Artefatos entregues

- **Dashboard de custo de uma página**

Mostra gasto por conta, serviço, owner, ambiente e tendência. Uma página. Se precisa de treinamento para entender, está grande demais.

- **Política de tagging mínima**

Define tags obrigatórias, valores aceitos e ponto de bloqueio. O ideal é validar antes do recurso nascer: Terraform policy, OPA, SCP, Azure Policy, GCP Org Policy ou mecanismo equivalente.

- **Decisão de chargeback / showback**

Nem todo cliente deve começar com chargeback. Em muitos casos, showback mensal já muda comportamento sem criar briga contábil. A decisão precisa ser explícita.

- **Baseline escrita de governança**

Quem olha custo semanalmente. Quem aprova resize. Qual alerta chama finance. Qual mudança entra por PR. Qual exceção precisa de TL ou sponsor.

- **Mapa dos 10 maiores custos com dono**

Sem owner, a discussão vira “infra gasta demais”. Com owner, vira “este workload custa X por mês por causa de Y; vamos testar Z”.

O baseline bom não é bonito. Ele é usável em uma reunião de 30 minutos com engenharia e finanças na mesma mesa.

4. Day 60 — quick wins com evidência

Quick win não é sair desligando coisa. É uma mudança com três propriedades: evidência antes, risco conhecido e medição depois.

As faixas abaixo são **bandas de planejamento**, não promessa de economia. A porcentagem final só entra em proposta depois do Day 0, com billing e métricas reais.

Quick win	Faixa típica	Evidência que o cliente recebe	Perfil de risco
Rightsizing de compute, RDS e containers	10-25% da linha afetada; em workloads sem otimização recente, 15-25% da conta total pode aparecer como oportunidade	Export de 30 dias, candidatos por owner, antes/depois de CPU/memória/IO, plano de waves e rollback	Baixo a médio. Baixo em non-prod; médio em produção com batch, banco ou sazonalidade
Rebalance de RI / Savings Plans	10-20% da base estável coberta quando há pouca ou má cobertura	Relatório de coverage/utilization, baseline reduzida pós-rightsizing, simulação de 1 ano vs. 3 anos	Médio. Compromisso financeiro dura mais que o projeto
Políticas de lifecycle em S3, logs e backups	10-35% da linha de storage/logs elegível, dependendo de retenção e classe atual	Inventário por bucket/prefixo, regra proposta, janela de retenção, exceções regulatórias	Baixo se retenção estiver validada; alto se o dado não tem dono
Limpeza de EBS, snapshots, EIPs, load balancers e imagens antigas	5-15% em contas sem governança quando combinado com limpeza de waste	Lista de recursos, última atividade, owner, proposta de quarantine antes de deletar	Baixo quando o recurso é comprovadamente órfão; médio quando falta owner
Reroteamento de data egress / NAT / VPC endpoints	10-30% da linha de rede quando rede está no top 10	Flow Logs, mapa de rotas, endpoints redundantes, simulação de arquitetura mínima	Médio. Rede muda blast radius; validação precisa ser explícita

5. A ordem dos quick wins importa

A ordem errada é cara.

Comprar commitment antes de reduzir a baseline trava desperdício por 12 ou 36 meses. Migrar workload para outra cloud antes de entender data egress troca uma fatura por duas complexidades. Instalar ferramenta de FinOps antes de definir owner só produz dashboard que ninguém usa.

A sequência que usamos:

1. Rightsizing primeiro

Reduz a linha de base técnica. Começa por ambientes de menor risco, usa métricas de 30 dias e aplica em waves. Para banco, a análise precisa ser workload-aware: CPU média baixa não basta se existe batch de fechamento no fim do mês.

2. Commitments depois

Savings Plans, RI e reservas entram sobre a base já reduzida e estável. Compromisso parcial costuma ser mais saudável que cobertura máxima. A pergunta não é “qual desconto o vendor oferece?”, é “qual capacidade vamos de fato usar durante o prazo?”.

3. Limpeza sempre

Snapshots, logs, volumes, imagens e ambientes não-prod voltam. Se a limpeza não vira rotina, a economia do mês 2 evapora no mês 5.

4. Governança em paralelo

Tags, budgets e owners não esperam o fim do projeto. Sem eles, o time volta a discutir custo como se fosse fenômeno natural.

6. Day 90 — governança que fica

No dia 90, a pergunta muda de “quanto a Redgator achou?” para “o cliente consegue continuar achando?”.

O que o cliente deve possuir

- **Budget alerts:** alertas em 50/75/90/100%, por owner ou unidade, com canal e responsável definidos.
- **Anomaly detection:** detecção de variação anormal por serviço/conta, com regra de triagem. Alerta sem triagem vira ruído.
- **Template de revisão mensal FinOps:** pauta de 45 minutos, decisões pendentes, top variações, savings realizados, riscos e exceções.
- **Checklist de handoff:** dashboard, política de tags, owners, runbooks, exceções aprovadas, commitments com vencimento, backlog de otimização e decisões que ficaram para depois.
- **Rito com engineering:** custo aparece no PR quando possível. Infracost, policy-as-code ou comentário de plan não existem para enfeitar pipeline; existem para evitar surpresa.
- **Rito com finance:** showback ou chargeback definido, com linguagem que finance entende e engenharia consegue contestar.

A entrega final não é “um relatório”. É um sistema pequeno de decisão: quem olha, quando olha, o que faz quando encontra algo e como mede se funcionou.

7. O que este guia não é

Esta seção importa porque FinOps vira facilmente desculpa para vender outra coisa.

Não é migração cloud

Pode aparecer recomendação de arquitetura. Pode aparecer workload que deveria sair de onde está. Mas o engajamento de 90 dias não é projeto de migração. Se a resposta certa for migrar, isso vira outro escopo, com discovery, riscos e contrato próprios.

Não é estratégia multi-cloud

Usar multi-cloud para “negociar com a AWS” raramente reduz custo total. A fatura pode até baixar em um serviço, mas operação, rede, segurança e skill do time entram na conta. Multi-cloud só entra se o problema operacional justificar.

Não é substituto permanente de time FinOps

A Redgator ajuda a montar baseline, executar a primeira onda e deixar governança. Se o cliente tem escala para um time FinOps interno, ótimo. Se não tem, o rito precisa caber na operação existente.

Não é corretora de commitment

A gente não ganha por vender Savings Plan, RI ou reserva. A recomendação pode ser comprar menos, comprar depois ou não comprar.

Não é dashboard mágico

Dashboard mostra. Engenharia muda. Sem owner, PR, alerta e revisão, dashboard só deixa o desperdício mais visível.

8. Como funciona um engajamento Redgator

O primeiro contato útil é técnico. A conversa inicial de 45 minutos é com o TL da prática — não com vendedor lendo roteiro. O objetivo é entender trigger, escala, restrição e risco antes de falar de formato.

Como dimensionamos

A unidade comercial de referência é **engineering-month**: um mês-calendário de atuação de engineer sênior, TL, staff engineer ou especialista contra outcomes definidos, ritos, documentação e critérios de aceite.

Não publicamos tabela de horas. Quando procurement precisa decompor, explicamos capacidade, responsabilidade, entregáveis e marcos. O contrato e a proposta formal vivem no pacote comercial controlado; este guia não substitui esse documento.

Como começa

- **Call de 45 minutos com TL da prática** — trigger, fatura, cloud principal, risco, owner e urgência.
- **Pedido mínimo de evidência** — sem billing export sensível na primeira conversa; para discovery, precisamos de acesso controlado ou export sanitizado.
- **Escopo curto de assessment** — Day 0 e baseline antes de assumir economia.
- **Plano de 90 dias** — waves, quick wins, governança, handoff e o que fica fora.

O melhor resultado possível às vezes é um “não agora”. Se a conta é pequena, se não há sponsor técnico, se o problema é ferramenta e não operação, a gente diz cedo.

9. Prova — números do proof ledger

Os recortes abaixo usam somente números do `strategy/proof-ledger.md` do PR #25. O CEO ainda verifica a base no RED-20; se algum número for retraído, esta seção precisa de uma segunda passada rápida antes de publicação.

Case-let 1 — coorte FinOps de 90 dias

- **Indústria:** multi-indústria; contas com gasto cloud acima de R\$ 200k/mês.
- **Escopo:** engajamentos FinOps de 90 dias contra faturas cloud já existentes.
- **Resultado aprovado:** -23% custo cloud médio em 90 dias.
- **Contexto:** Use como benchmark de coorte, não promessa para toda conta; a proposta final depende do inventário Day 0.
- **Ledger:** `proof-ledger/canonical-proof-numbers/-23-average-cloud-cost-in-90-days`.

Case-let 2 — confiabilidade durante mudança de custo

- **Indústria:** contas retainer 24x7 com serviços de produção escopados em SLA.
- **Escopo:** operação dos últimos 12 meses em contas 24x7.
- **Resultado aprovado:** 99,98% uptime médio.
- **Contexto:** O corte de custo só entra quando a confiabilidade continua mensurável; economia não justifica mudar sem owner, alerta e rollback.
- **Ledger:** `proof-ledger/canonical-proof-numbers/99-98-average-uptime-12m-retainer`.

Case-let 3 — governança de incidente como guardrail

- **Indústria:** contas retainer 24x7 com incidentes Sev1/Sev2 de produção.
- **Escopo:** MTTR mediano dos últimos 12 meses em incidentes de produção.
- **Resultado aprovado:** 37 min MTTR mediano em produção.
- **Contexto:** Quick wins entram em waves porque a economia precisa preservar capacidade de recuperação; se piora MTTR, não é quick win.
- **Ledger:** `proof-ledger/canonical-proof-numbers/37-min-median-production-mttr`.

Critério de publicação: cada número precisa apontar para uma linha do proof ledger, um case público ou uma exceção sob NDA aprovada. Se não tem fonte, não entra.

10. Próximo passo

Se a fatura de cloud já virou pauta fora da engenharia, comece pelo diagnóstico. Não pela compra de commitment. Não pela ferramenta. Não pela migração.

Leve para a primeira conversa:

- gasto mensal aproximado e cloud principal;
- top 3 serviços que mais pesam;
- maturidade de tags;
- quem hoje assina a decisão de custo;
- restrições de produção, compliance e janela de mudança.

Fale com a Redgator: <https://redgator.com.br/contact/>

A resposta esperada não é um pitch. É uma triagem técnica: faz sentido seguir para assessment, precisa de outra prática ou é melhor não mexer agora.