

FinOps: como diretorias de TI podem controlar custos em ambientes cloud

Em grandes empresas que já operam em cloud há alguns anos, a conversa sobre custos costuma ter um padrão previsível. A fatura mensal chega mais alta do que o esperado, alguém pergunta o que aconteceu, engenharia responde que provisionou novos ambientes para um projeto legítimo, finanças questiona por que ninguém avisou antes, procurement descobre que existem três contratos paralelos com o mesmo hyperscaler negociados por áreas diferentes, e o mês seguinte começa com determinação de "controlar melhor". Meses depois, o mesmo padrão se repete com outro projeto, outro sistema, outro pico de consumo. Esse ciclo se repete porque cloud é fundamentalmente diferente de infraestrutura tradicional: em vez de compra planejada com investimento fixo, é consumo variável com fatura reativa. É exatamente esse cenário que a disciplina de **FinOps** existe para resolver, oferecendo modelo operacional estruturado que conecta engenharia, finanças, negócio e procurement em torno de gestão contínua do valor da tecnologia. Para gerentes e diretores de TI em ambientes corporativos sérios, essa disciplina deixou de ser "boa prática opcional" para se tornar capacidade crítica com impacto direto em margem, previsibilidade orçamentária e capacidade de escalar iniciativas modernas.

A urgência cresceu por motivo concreto. O gasto global em cloud público deve superar 830 bilhões de dólares em 2026, segundo projeções da Gartner, e análises da Flexera indicam que 32% desse gasto é desperdiçado por over-provisioning, recursos ociosos, commitments não otimizados e falta de visibilidade. Ao mesmo tempo, 76% das grandes empresas já gastam mais de 5 milhões de dólares por mês em cloud, e a chegada de cargas de IA agêntica com consumo baseado em tokens introduziu uma classe de custo que budgets tradicionais não conseguem prever. Esse cenário conversa diretamente com o que abordamos no conteúdo sobre [otimização estratégica de TI](#), porque desperdício em cloud é hoje uma das fontes mais volumosas de custo evitável em grandes operações.

O que é FinOps, e por que sua definição mudou em 2026

A definição precisa importa. **FinOps** é um framework operacional e prática cultural que combina disciplinas de gestão financeira com engenharia e operações de cloud para maximizar o valor de negócio extraído dos gastos com tecnologia. A palavra é abreviação de "Cloud Financial Operations", cunhada pela FinOps Foundation, hoje hospedada pela Linux Foundation, que mantém o framework de referência adotado globalmente. Vale registrar com honestidade que o objetivo declarado da disciplina não é reduzir custo. É maximizar valor de negócio por dólar investido, com controle e previsibilidade adequados.

A mudança de escopo em 2025 e 2026 é o ponto mais relevante para uma diretoria de TI que está avaliando o tema. Historicamente, FinOps significava "controlar contas do cloud público". O

framework moderno redefiniu a disciplina como "advancing people who manage the value of technology", com escopo expandido para o que a FinOps Foundation chama de Cloud+: cloud público, SaaS, data center, licenciamento e agora, com força crescente, gastos com IA e agentes autônomos. Essa expansão reflete uma realidade concreta: para uma empresa com 100 milhões de dólares em gasto anual em tecnologia, cloud público representa tipicamente 30% a 40% do total. O resto se fragmenta em SaaS, infraestrutura própria, contratos legados de licenciamento e consumo de IA generativa. Tratar essas categorias em silos é receita para desperdício estrutural.

Os três pilares do FinOps Framework moderno

Para uma diretoria de TI que pretende estruturar **FinOps** com seriedade, é útil entender que o framework de 2025 e 2026 é construído sobre três pilares interdependentes. O primeiro é **Domains**, que são áreas funcionais aplicadas transversalmente. Seis domínios formam a base: Plan & Forecast (desenvolvimento de orçamento, previsão de uso, planejamento financeiro), Inform (visibilidade de custo, relatórios, benchmarking, chargeback), Optimize (identificação de desperdício, rightsizing, otimização de descontos, gestão de commitments), Govern (aplicação de política, detecção de anomalia, workflows de aprovação, compliance), Manage (ferramenta FinOps, operações do time, workflows de processo) e Build (tagging, coleta de dados, modelagem de custo em infraestrutura como código).

O segundo pilar é **Capabilities**, que são competências técnicas e organizacionais específicas que sustentam cada domínio. Uma das adições mais relevantes do framework 2026 é a Capability de Executive Strategy Alignment, que formaliza a conexão entre FinOps e tomada de decisão executiva. Essa evolução reflete o reconhecimento de que valor de tecnologia virou conversa de board, e que a prática FinOps precisa operar como parceira estratégica da liderança, não apenas como função de reporte reativo. O terceiro pilar é **Scopes**, que substitui a categoria única de "cloud" por múltiplas categorias aplicáveis: Public Cloud, SaaS, Data Center, AI, entre outras, cada uma com dinâmica própria de precificação e governança.

Esse desenho conversa diretamente com o que discutimos em [gestão de capacidade em TI](#), porque dimensionamento de recursos é gêmeo natural do controle de custos, e as duas disciplinas se reforçam quando desenhadas em conjunto.

As três fases do ciclo FinOps: Inform, Optimize, Operate

A prática de **FinOps** organiza a jornada de maturidade em três fases que se realimentam continuamente e definem o nível de sofisticação da empresa. A primeira fase é **Inform**. Aqui o objetivo é estabelecer visibilidade granular sobre onde o dinheiro está sendo gasto, com atribuição correta a times, projetos, produtos e centros de custo. Sem essa base, otimização é adivinhação. Empresas que pulam essa fase acabam otimizando o que é fácil de medir em vez do que traz maior

retorno, e produzem economia superficial. Nesta fase, entram as capacidades de tagging rigoroso, chargeback, showback e visibilidade unificada entre múltiplos provedores.

A segunda fase é **Optimize**. Com a informação confiável em mão, começam as ações concretas: rightsizing de instâncias sobredimensionadas, identificação de recursos ociosos, otimização de commitments (Reserved Instances, Savings Plans, contratos com desconto por volume), eliminação de armazenamento órfão, negociação com fornecedores. Análises de mercado apontam que compute ocioso e instâncias superdimensionadas representam cerca de 60% de todo o desperdício em cloud, tornando-os os alvos de maior retorno inicial. A terceira fase é **Operate** (também chamada Run), onde a disciplina se torna cultura contínua embutida nos processos diários. Aqui, engenharia toma decisões considerando custo no momento de provisionar, times de produto acompanham unit economics de suas features, procurement negocia com base em uso real projetado e finanças opera com forecasts baseados em modelagem preditiva.

Esse ponto conversa com o que abordamos em [observabilidade em TI](#), porque observabilidade e FinOps compartilham base de dados operacionais, e maturidade em uma acelera maturidade na outra.

A expansão para Cloud+ e o desafio novo da governança de custos de IA

A frente que mais tem mudado a prática de **FinOps** em 2026 é a incorporação de IA e agentes autônomos ao escopo. Pesquisa State of FinOps 2026 aponta que 98% dos praticantes FinOps agora gerenciam gasto com IA, contra 63% no ano anterior. Essa migração se explica pela natureza específica dos custos de IA: consumo baseado em tokens, invocação de ferramentas, consultas a vector databases, GPUs com pricing variável, agentes autônomos que executam workflows sem supervisão humana entre invocações.

O padrão de risco é qualitativamente diferente do cloud tradicional. Um agente mal configurado pode gerar custos em horas que levariam meses para acumular sob provisionamento tradicional, e casos publicados relatam picos de faturamento em ordem de grandeza vezes maior que o esperado. A assimetria entre o ciclo lento de procurement e a velocidade de execução de agentes autônomos criou exposição financeira nova que a maioria das empresas ainda não sabe quantificar. Esse tema dialoga diretamente com o que discutimos em [governança de IA nas empresas](#), porque governança de IA e FinOps de IA são duas faces do mesmo problema. Quem responde por decisões, riscos e custos precisa ter inventário, ownership e trilha de auditoria formais. E conversa com o que abordamos em [SAP Joule](#), porque agentes corporativos em escala exigem monitoramento de consumo com granularidade que dashboards de custo genéricos não entregam.

Adicionalmente, 2026 marca a consolidação do FOCUS (FinOps Open Cost and Usage Specification), padrão aberto que normaliza dados de billing entre provedores. À medida que mais fornecedores

adotam o padrão, empresas ganham fundação consistente para aplicar princípios de FinOps sobre todo o gasto tecnológico, sem depender de mapeamentos manuais entre formatos incompatíveis.

Os erros mais comuns em programas de FinOps

Falar de **FinOps** sem nomear honestamente os erros comuns seria desserviço. Cinco armadilhas concentram a maior parte dos casos onde o programa decepciona. A primeira é tratar como iniciativa exclusivamente financeira, sem coautoria substantiva da engenharia. Sem engenheiros que entendam o impacto de decisões técnicas em custo e finanças que entendam a natureza variável de consumo cloud, o programa vira relatório mensal que ninguém aciona. O modelo operacional mais bem-sucedido, adotado por 60% das empresas maduras segundo State of FinOps 2026, é o de enablement centralizado com times distribuídos que executam otimização em suas próprias áreas.

A segunda armadilha é comprar ferramenta e esperar que ela resolva o problema. Análises de mercado convergem em que ferramenta sozinha não reduz gasto. Ela viabiliza recomendações, mas a governança de decisão, a responsabilidade dos times de engenharia, a disciplina de tagging e o SLA de resposta a anomalias é que reduzem a fatura. Empresas que instalam plataforma sem desenhar processo terminam com dashboards bonitos e mesma trajetória de custo. A terceira armadilha é ignorar a governança de commitments. Reserved Instances, Savings Plans e contratos com desconto por volume podem entregar redução de 30% a 45% em compute quando bem gerenciados, mas viram passivo financeiro caro quando mal dimensionados. Empresas maduras miram cobertura de commitment entre 60% e 70% da baseline estável de workloads.

A quarta armadilha é desconectar FinOps de arquitetura. Decisões arquiteturais tomadas hoje definem o perfil de custo de amanhã. Escolhas sobre serverless versus container, banco gerenciado versus autogerenciado, região primária, estratégia de multi-cloud, tudo isso tem impacto plurianual em conta cloud. Esse ponto conversa com o que discutimos em [SAP BTP](#), porque plataformas modernas com modelo de consumo exigem disciplina FinOps desde a primeira decisão de desenho, não como remediação depois. A quinta armadilha é ignorar SaaS e AI. Empresas que definem escopo apenas em cloud público estão cobrindo cada vez menor parcela do gasto tecnológico total, e o custo evitável nas categorias ignoradas costuma ser tão relevante quanto o do cloud.

Para referência de framework atualizado e completo, vale acompanhar diretamente o site da [FinOps Foundation](#), onde o framework oficial e pesquisas anuais como State of FinOps estão disponíveis publicamente.

Indicadores que mostram se o programa de FinOps está entregando valor

Programas sérios medem progresso em quatro dimensões. A primeira é **eficiência de gasto**: custo unitário por transação, por usuário, por request de API, por token consumido. Essa métrica, acompanhada em série temporal, mostra se o gasto está escalando de forma saudável com o negócio

ou se está gerando desperdício estrutural. Empresas maduras aplicam disciplina de unit economics em cloud, tratando cada workload como produto com margem apurável.

A segunda dimensão é **cobertura e qualidade de tagging**. Sem taxonomia rigorosa que atribua cada gasto a time, produto e centro de custo, chargeback e accountability são impossíveis. A terceira é **tempo médio de resposta a anomalias**. Quando um recurso começa a consumir fora do padrão, quanto tempo a organização leva para detectar, investigar e corrigir? Em ambientes maduros, isso cabe em horas ou dias. Em ambientes imaturos, leva o ciclo de fechamento mensal e o problema explode antes. A quarta é **retorno realizado versus recomendado**. Quantas das recomendações geradas pela ferramenta foram efetivamente executadas no último trimestre? Essa taxa distingue programa que age de programa que apenas reporta. Esse tema dialoga com o que abordamos em [gestão de indicadores empresariais](#), porque a hierarquia de indicadores precisa cobrir desde saúde operacional até performance estratégica.

Como uma diretoria de TI deveria estruturar o programa

A pergunta útil não é "vamos adotar FinOps?", mas "como organizar a prática para que ela entregue valor cumulativo e sustentável?". Quatro disciplinas costumam ser determinantes. A primeira é estabelecer papéis formais com ownership claro. FinOps opera com pelo menos cinco personas conforme o framework: praticante FinOps (o especialista dedicado), engenharia (que executa provisionamento), finanças (que aplica disciplina orçamentária), procurement (que negocia contratos) e executivo (que patrocina e aprova trade-offs). Sem papéis nomeados com tempo dedicado, o programa vira responsabilidade difusa que ninguém assume.

A segunda disciplina é o modelo de time. As duas estruturas mais eficazes são centralized enablement (time central que capacita e ferramenta as áreas, adotado por 60% das empresas maduras) e hub-and-spoke (time central com pontos focais em cada unidade de negócio, mais comum em grandes empresas com 21% de adoção). A escolha depende do porte, da distribuição geográfica e da maturidade da empresa. A terceira disciplina é o roadmap por fase, priorizando Inform primeiro, expandindo para Optimize com foco em quick wins de rightsizing e commitments, e evoluindo para Operate com cultura contínua embutida em processos.

A quarta disciplina é a integração com estratégia mais ampla. FinOps não vive isolado. Ele se conecta com estratégia de cloud, com arquitetura, com governança de IA, com sustentabilidade e com IT Financial Management. Esse tema dialoga com o que abordamos em [gestão de processos empresariais](#), porque disciplina de processo estruturado se aplica integralmente à evolução da prática FinOps.

Em última análise, **FinOps** moderno é uma das disciplinas mais relevantes que uma diretoria de TI pode desenvolver quando a empresa atinge porte em que gasto tecnológico se tornou fatia crítica do orçamento corporativo. Quando inserido em modelo bem desenhado, com ownership claro e

integração com engenharia, ele transforma cloud de fonte de surpresa em vetor de valor previsível, libera capital para reinvestimento em iniciativas estratégicas e prepara a empresa para governar a próxima onda de custos vindos de IA agêntica. Quando tratado como iniciativa financeira isolada, vira relatório mensal que ninguém aciona e que não muda a trajetória de gasto.

Se a sua empresa quer estruturar uma prática de **FinOps** para controlar custos em ambientes cloud, melhorar previsibilidade orçamentária e construir base sólida para escalar iniciativas de IA e transformação digital sem surpresas financeiras, a Simple pode apoiar esse movimento com Arquitetura de Soluções, Mapeamento de Requisitos, Arquitetura de Software, Desenho de Soluções Completas, projetos com CELONIS, Consultoria e Execução SAP, Análise de Aderência, Implementação do S/4 Hana, Soluções Customizadas SAP, Integrações SAP com outros fornecedores e Terceirização de Tecnologia, incluindo busca, avaliação, alocação de profissionais e formação de squad. Entre em contato com a Simple para avaliar o estado atual da sua governança de custos cloud e desenhar o caminho de evolução que melhor se ajusta à realidade do seu negócio.