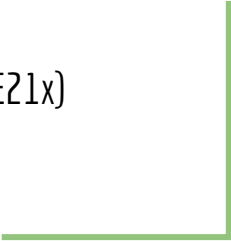




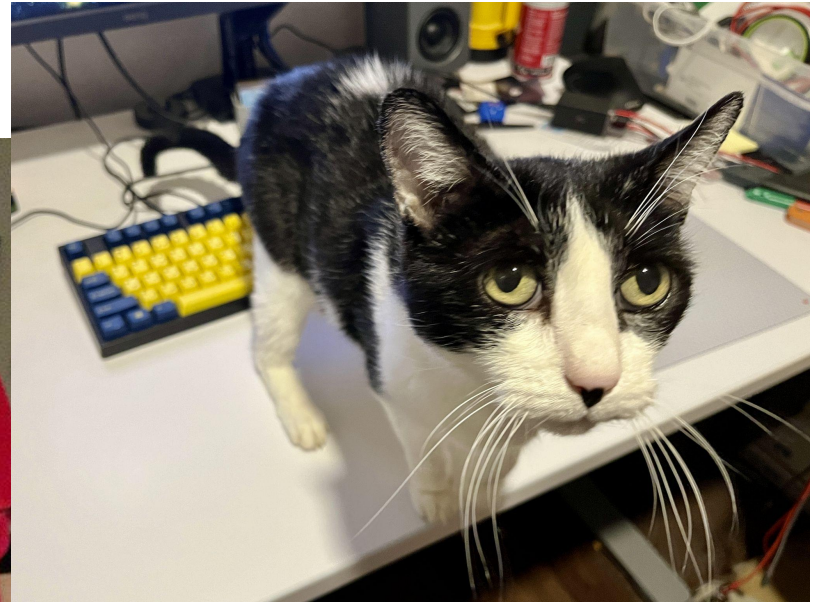
# Anti-Patterns in Tech Cost Management

Michael Gat  
For  
SoCal Linux Expo 21x (SCaLE21x)

March 16, 2024



# First... a story



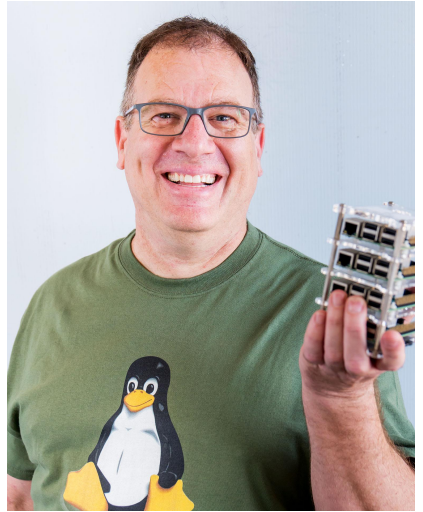
For the full story see:  
<https://michaelgat.com/2023/12/04/my-most-important-lesson/>

# Key Takeaway

- Engineering is where science meets economics
- Cost is key to every decision we make
- ZIRP has blinded many of us to this reality
  - Some of us have never worked in a non-ZIRP environment
  - Money has real value again
  - We're going to have to treat it with more respect
- This talk will address how to avoid disrespecting our costs

# A bit more about me

- Recently, Tech Program Manager at AWS
  - Managed capacity and efficiency of internal load balancing supporting over 60TB/sec, and over \$70m/year capital spend
  - Later managed onboarding to new services, and then large scale systems integration programs in the billing organization
- Previously an IT consultant, SDM, Tech strategist and engineer
  - Multiple industries and range of company sizes
- CS undergrad, MBA, certified project manager, etc.
- Serve the whims of two very demanding cats
- Originally from New York



# Why Anti-Patterns?

- There are a lot of good ideas
- There are also a lot of bad ones
- Patterns are easy to recognize and are broadly useful
- Avoiding the really big mistakes gets you 50-80% of the way there
- *It's a lot easier to talk about what you shouldn't do than figure out how to give you useful proactive advice*

# What won't I cover

- Team management/organization
- Build/Buy decisions
  - Given the SCaLE audience, I will focus on “stuff we build ourselves”
- Interplay of technology and business strategy
- Forecasting
- There's only so much I can cover in 45 minutes...

# Overall Framework

- 2 big spend buckets in any tech org
  - Code/Infrastructure
  - People
  - Conways law says they're the same thing! (Given that infra is now also code)
- 4 dimensions to consider
  - Scale of the enterprise
    - You are (probably) not Google, but what are you?
  - Strategic
  - Architectural - both infrastructure and people, remember Conway's Law
    - Infra cost and infra architecture are one and the same
  - Operational/Tactical
- We impact or are impacted by all of these
  - You will either do it, or it will be done to you!



# The Top 10 Things not to do



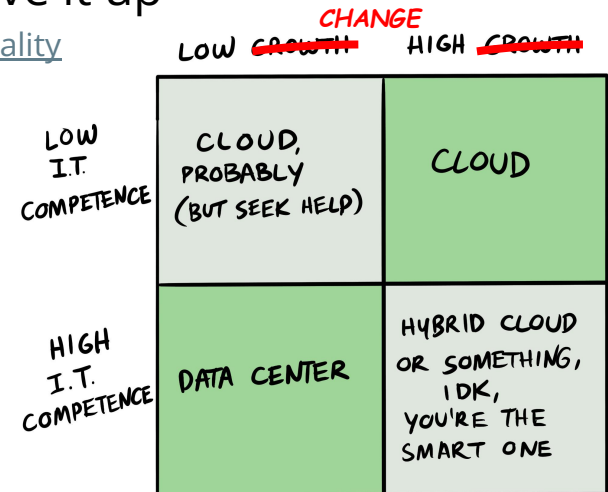
# 1. Not considering scale

- [You are not Google!](#)
- The scale you operate at will determine how closely you can/should manage costs
- A large organization can continue pursuing 1-2% gains almost indefinitely
- A small organization can often find some big things quickly, but after that, the cost is hard to justify
- Scale will determine when you should stop



## 2. Bad Cloud Strategy

- The cloud will not save you money right away, possibly never
- High “cloudiness” improves velocity, reduces fixed costs, and makes cost attribution easier
- There is no such thing as “cloud neutrality” so give it up
  - See DevOps Vancouver: [Kubernetes and the Myth of Cloud Neutrality](#)
- Consider “which cloud?”
- Running the cloud as a datacenter is mostly a bad idea, even if your initials are “DHH”
  - Forest Brazeal has a great analysis
  - “Cost analysis” [assuming you run your existing architecture in the cloud](#) is always wrong
  - But doesn’t mean the cloud is right either!



Graphic by Forrest Brazeal, see appendix

# 3. No ability to assign/attribute costs

- Often a “legacy” problem
  - It wasn’t an issue when you were a lot smaller
  - Frequently associated with failure to refactor “startup” products
- Hard to implement good attribution later
  - There are technical as well as political issues
  - The process is roughly: Nothing -> Useful directional data -> Agreed-upon numbers
- Especially hard where there are multiple layers to track:
  - Infrastructure
  - Platforms
  - Internal Services
  - Actual products/services that somebody pays to use
  - “Each language is it’s own project and different features are implemented...”
- High “cloudiness” usually correlates to easier attribution
- Data ownership/cost can be especially tricky

# 4. No metrics/bad Metrics

- Metrics are “*quantitative measurements* that provide *insight* into the *inputs* and/or *outputs* of a process”
  - Cost is usually \$\$\$/<*some other metric*> not a standalone
- Remember Goodhardt’s Law!
  - We are really, really, really good at gaming the numbers!
  - “Incentives are superpowers; set them carefully.” ~ [Sam Altman](#)
- “Informational” metrics focus people on the wrong things
  - “If you can’t measure it, you can’t manage it”  
~ [Not said by Deming](#) (he said the opposite actually)
  - “What gets measured, gets managed”  
~ [Not said by Drucker](#)
  - “Don’t measure what you don’t want to be forced to manage”  
~ Said by me (© 2023 Michael Gat)
- This is the strength of OKRs: “measure what matters”
  - Best if the people working on the problem [determine targets](#)

# 5. Not designing it in

- Some of us will sometimes have the opportunity to design something from the ground up
  - Do not waste this opportunity if you get it!
- Ideally, costs and benefits can be quantified to justify the effort
- At the very least, we can build in ability to measure and attribute cost
- Long term infrastructure cost is a direct reflection of architectures chosen early on, and can be challenging to change later
  - The choice is also a choice of your staffing model and its costs
  - Be explicit about the choices made and the tradeoffs implied
- Open source offers alternatives to optimize for cost and functionality
- This applies to any other initiative that includes architectural change

# 6. Cost Management as a standalone

- *“Align authority and responsibility”* ~ Kent Beck
- Somebody else can't be accountable for your costs!
- There can be an independent, standalone team responsible for highlighting and reporting costs
- But only the teams/organizations performing actions that can change the costs should be accountable for that
- A common mistake is to put together a standalone team and declare that costs are their problem to address
  - It's burnout inducing and leads nowhere, for the organization and the person doing it (ask me how I know!)

# 7. No ongoing review of utilization/spend

- The existence of metrics doesn't mean they're being reviewed
- Often, cost data is not shared. This is a fail.
- Anomalies will show themselves quickly if you have the right metrics
- Need to be consistent across the org, and work from the bottom up
  - There should not be a surprise at the end of the quarter
- This is a space where a FinOps team, working with each engineering team, can make a big difference
- "What can we turn off?" is the most important question
- Remember to review what data is being stored, where, and for how long
- Actual vs. Budget is an advanced topic

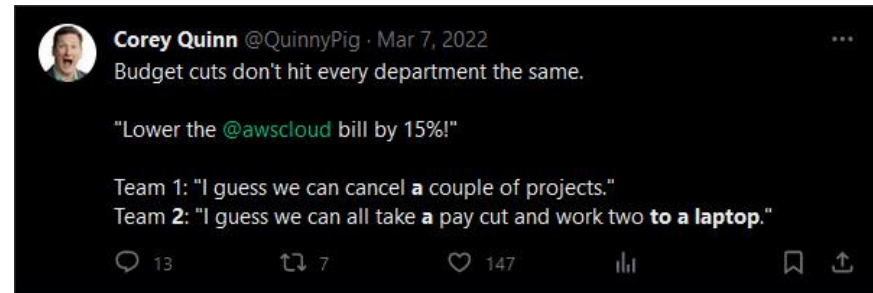
# 8. Not reviewing/benchmarking infrastructure

- New instance types
- New storage capabilities/categories
- Changes that allow you to minimize data transfer fees
- Changes that make costs more predictable
- Opportunities to use new tools to eliminate work
- How has the build/buy, cloud/on-prem math changed?
- Languages (org dependent: may be tactical, architectural or strategic)



# 9. “Across the board” targets

- Across-the-org targets for cost usually reflect a lack of understanding of where the costs are generated
- *“The reality is that even in low-performing organizations, there are always pockets of high performance. In high-performing organizations, there are always pockets of weakness.”* - The DuckBill Group
- “Everybody cut 20%” punishes teams that have been doing the right thing, and creates perverse incentives
  - A bad metric/target
- Encourages bad choices like “back pocketing” improvements



# 10. A tool will save us!

- Most cloud cost tools are still immature
- Tools cost money
  - There are open source options for instrumenting systems, that can be leveraged for cost
- Tools require additional staffing to support
- Tools may disrupt the way you are doing things now
- Make sure you understand all this
- Know what you want the tool to do, and be sure it does it
  - Consider whether anything else in your stack can give you similar information
  - Remember that metrics exist to give insight. What insight do you need?
- The salespeople/engineers won't tell you

# 11 (Bonus!). Rewards programs???

- Recommended by the people at ACloudGuru and some others
- Remember the Cobra Effect
- Introduces IBGYBG thinking
- We can all think of bad things to do if there's money attached
  - What are the odds that backup will be needed before you're gone?
  - How likely is a zonal outage in the last 2 months of the year?
  - Major cost consultancies won't take "pay for savings" jobs, for good reasons, and neither should you
  - If you want a commission, Get a sales job



The image shows a screenshot of the Urban Dictionary website. The header includes the 'urban DICTIONARY' logo and navigation links for 'Browse', 'Store', 'Blog', 'Discord', and 'Advertise'. A search bar is visible below the header. The main content area displays the entry for 'IBGYBG', which is an acronym for 'I'll Be Gone, You'll Be Gone.' The definition is annotated with handwritten red text: 'ingenious' above 'douchebag', 'tenderly embracing' above 'fucking', and 'American public' above 'global economy'.

urban  
DICTIONARY

Browse ▾ Store Blog Discord Advertise

Search

## IBGYBG

Acronym for "I'll **Be Gone**, **You'll Be Gone**."

It's what <sup>ingenious</sup> ~~douchebag~~ **hedge fund** managers say to each other when they create an investment mechanism that makes both of them rich while <sup>tenderly embracing</sup> ~~fucking~~ the ~~American public~~ **global economy**.



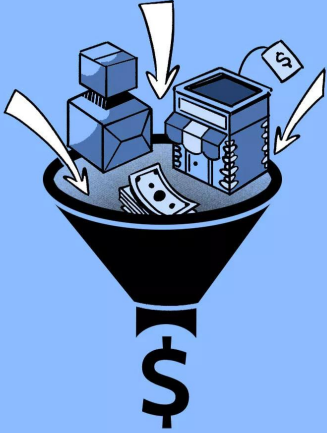
Extra:  
Three things to do right now!

# I. Start thinking about attribution

- You will either do this, or it will be done to you
  - Better that you do it yourself
  - Easier to discuss with management when you have a better idea
- Implementing attribution is hard;  
coming up with an approach may not be
- Once you know how costs are best assigned,  
a lot of other things fall into place


# II. Learn something about cost accounting

- One of the few things I learned in business school that still has value
- Engineering is where science meets economics; the economics you care about are most likely to be described by cost accounting
- Cost accounting is as much a way of thinking about the world as anything.
  - Not your classic “debits/credits” accounting!
  - It is a language and you are stronger if you know the language
- Another thing that you’ll either do, or it will be done to you
- Free classes online



**Cost Accounting**  
[ˈkɒst ə-ˈkaʊn-tɪŋ]

A form of managerial accounting that aims to capture a company's total cost of production by assessing the variable costs of each step of production as well as fixed costs, such as a lease expense.

 Investopedia

# III. Your worst enemy (sorrynotsorry)



The screenshot shows the top of the Atlantic website. At the top left is a search icon, followed by navigation links for 'Popular', 'Latest', and 'Newsletters'. The site's name, 'The Atlantic', is centered in a serif font. On the top right is a 'Sign In' link. Below the navigation is a yellow holiday-themed banner with gift icons and the text: 'A gift that gets them talking. Give a year of stories to spark conversation, plus a free tote.\*'. The main content area features the category 'TECHNOLOGY' in red, followed by the article title 'Slack Is Basically Facebook Now' in a large serif font. Below the title is the lead text: 'Slack's redesign suggests that keeping up with Slack is the only work worth doing.' and the author's name 'By Ian Bogost'.

<https://www.theatlantic.com/technology/archive/2023/09/slack-redesign-office-social-media/675324/>

# Thanks!

- Vancouver Software Development Managers
- Corey Quinn, Mike Julian, the Duckbill Group Slack members, etc.
- Tech Leaders Chat
- My colleagues at AWS, Stripe and elsewhere
- SCaLE
- These two =====>>>>>

*Slides and a writeup of the talk are available at*

*<http://michaelgat.com>*

*In the "presentations" tab*



Flash and MishMish  
without whose support this presentation would not have happened  
(or maybe it would have, but I'd be a lot less happy)

<https://linkedin.com/in/michaelgat>  
<https://michaelgat.com>

---



# Questions?

*Slides and a writeup of the talk are available at  
<http://michaelgat.com>  
In the "presentations" tab*

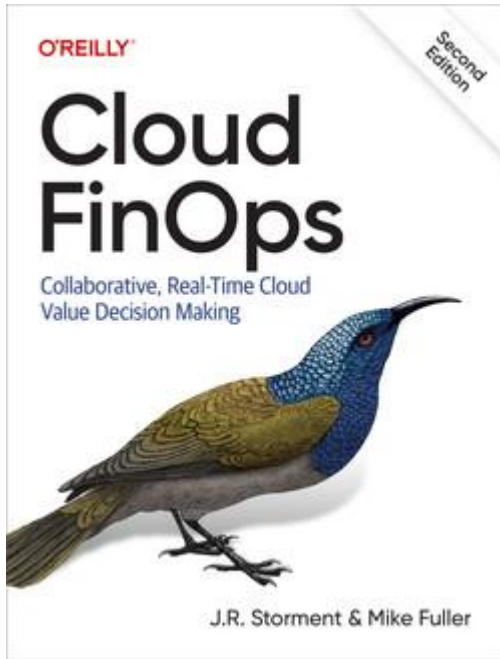


Flash and MishMish  
without whose support this presentation would not have happened  
(or maybe it would have, but I'd be a lot less happy)

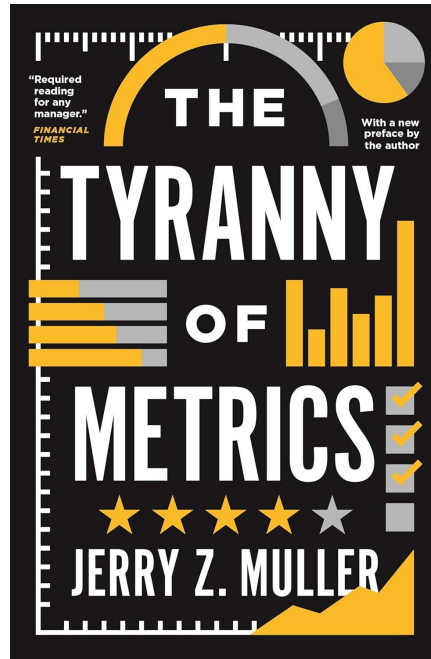
<https://linkedin.com/in/michaelgat>  
<https://michaelgat.com>

---

# Appendix: Reading



[Cloud FinOps, 2nd Edition](#)



[The Tyranny of Metrics](#)

## Forrest Brazeal: [Wait, is cloud bad?](#)

I love how he dissects the benefits and downsides of cloud vs on-prem architectures, and applies them to different types of businesses with different staffing/growth patterns.

# Appendix: More reading/viewing

## **Dale Hopkins:** [How Metrics Can Actually Help](#)

Remember Goodhardt's Law!

Addresses the power of OKRs, in which employees set the KR's within the parameters of O set by management. (And ideally, with some discussion and negotiation.) Also addresses the importance of "aspirational" KR's.

## **Federico Fregosi / DevOps Vancouver:** [Kubernetes and the Myth of Cloud Neutrality](#)

The migration journey of a demo application from AWS to GCP. Shows a lot of the pitfalls with a simple app. Can you imagine scaling this up to a real production app? Or an entire company? The unstated point is the major one: moving from one cloud provider to another is a massive, expensive, multi-year undertaking.

## **Jean Yang:** [You Are Not Google!](#)

Stop idolizing a small set of companies that have problems no one else actually has...

# Appendix: More reading (why aren't you productive)

## [Paul Graham: Maker's Schedule, Manager's Schedule](#)

Basically, why we can't schedule our engineers the way we often have to schedule ourselves. And maybe why we should try not to schedule ourselves that way either.

## [Stop the Meeting Madness \(HBR\)](#)

This one shows that there's nothing new to the problem. It's from 2017. But the fact that it's so old does make a point: the approaches it suggests, of a slow, managed, and executive-led approach to culling meetings don't work. Spotify's do.

## [Shopify exec: This is what happened when we canceled all meetings](#)

Pointless meetings are estimated to cost as much as \$25k (US) per employee at most companies. Shopify believes this will drive as much as 25% higher velocity.

“Uninterrupted time is the most precious resource of a craftsperson...”