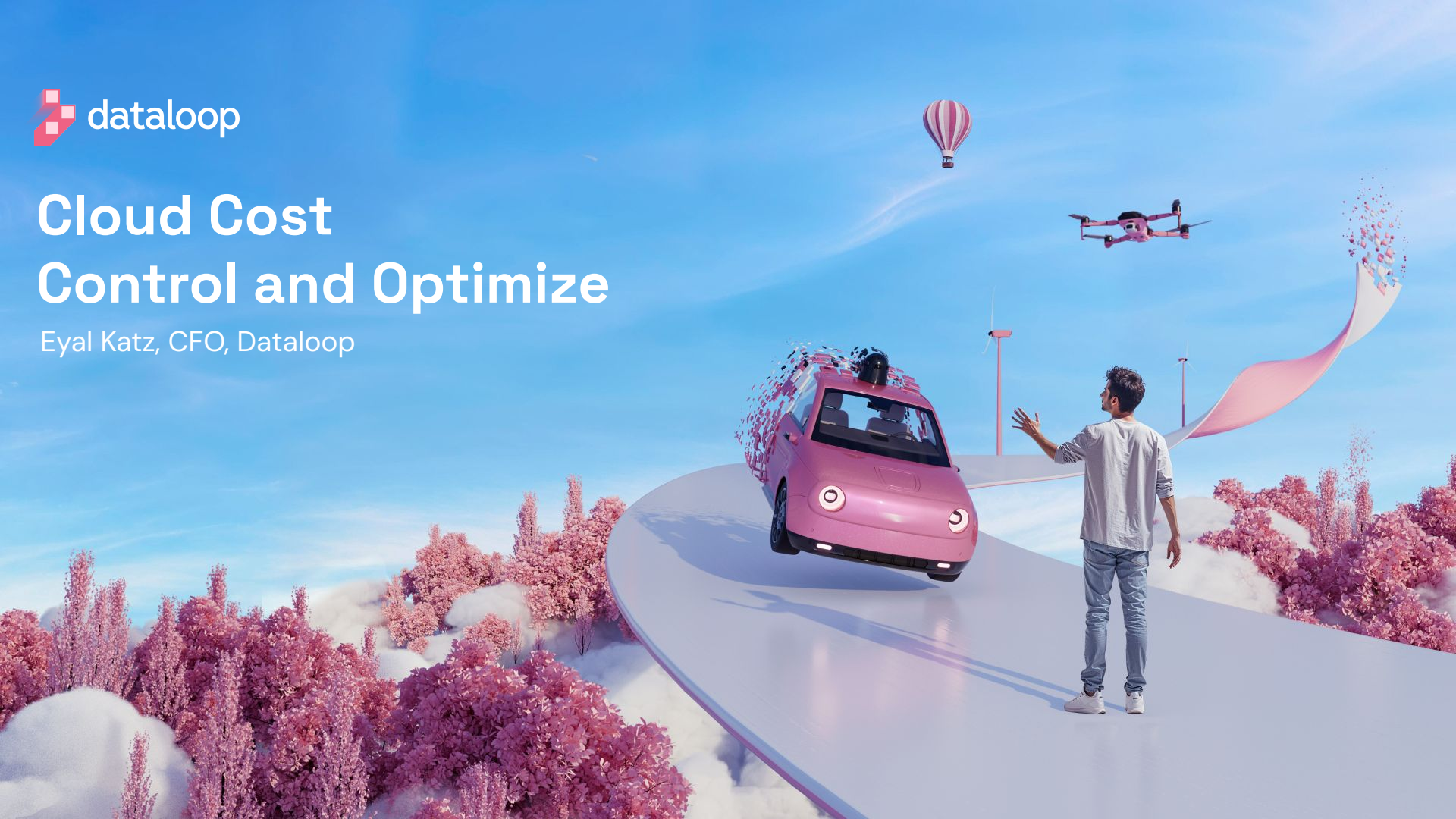
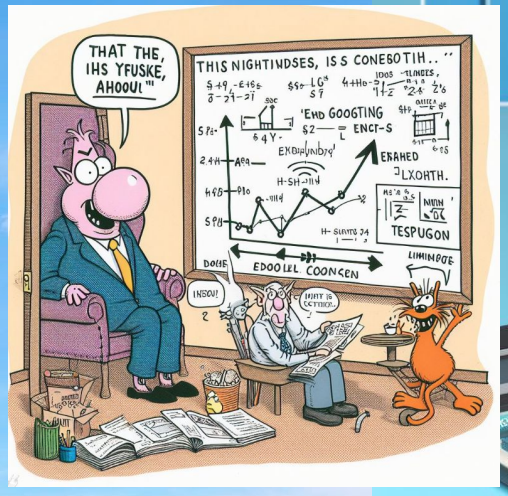
dataloop

Cloud Cost Control and Optimize

Eyal Katz, CFO, Dataloop




THAT THE, IHS YFUSKE, AHOOUL'"
TESPUGON
EDOOLEL COONCEN


$4/.90
PRICING
$19

Agenda

01.
Understanding Cloud Cost

02.
Strategies for Cost Optimization

03.
Tools and Best Practices

dataloop

End-to-End AI Development Platform

Build any Gen AI Application

About Dataloop

2017
Founded

Pioneered the first and most advanced end-to-end AI development Platform

80
Teams

HQ in Tel Aviv, Israel. With 3 additional remote sites in the US, EU and UAE

70%
R&D

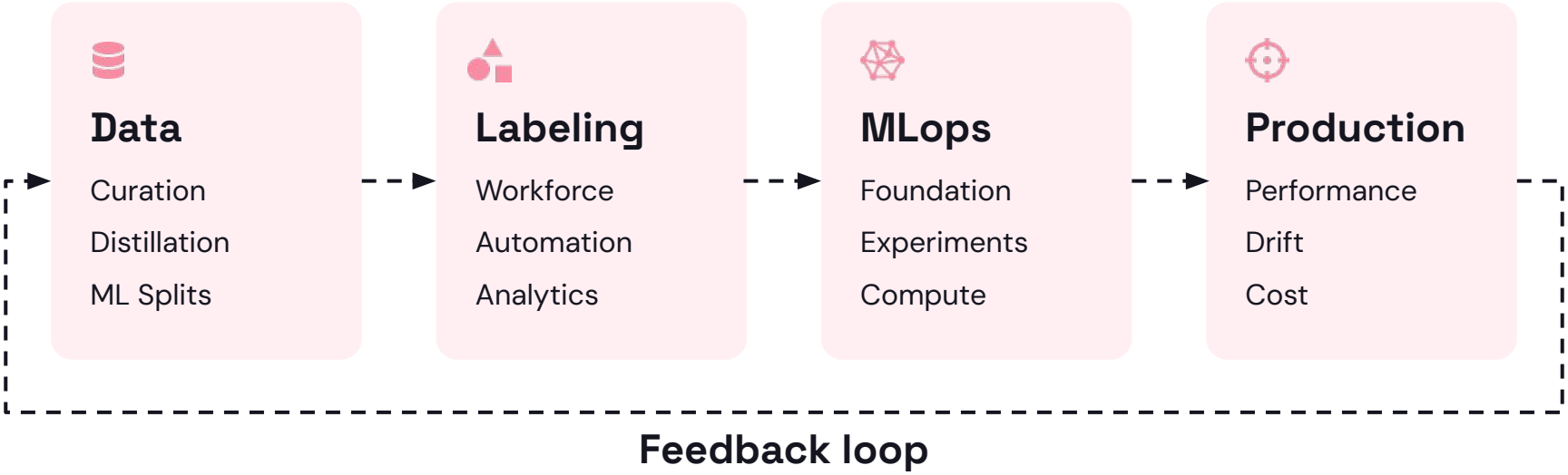
Top R&D, engineering, ML, & data science product with decades of experience.

$50M
Funding

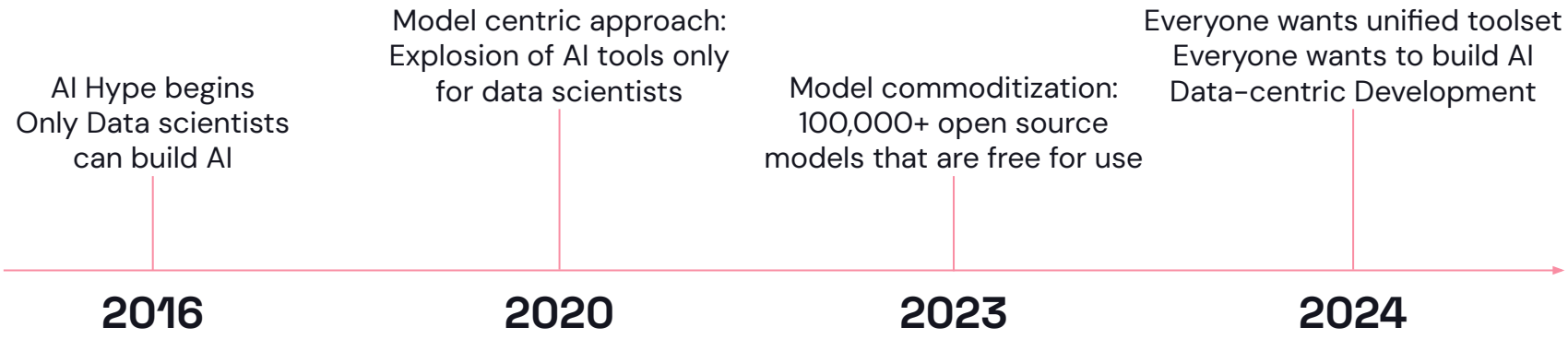
Raised by leading VCs such as Alpha Wave Global, NGP Capital and others.

Our customers

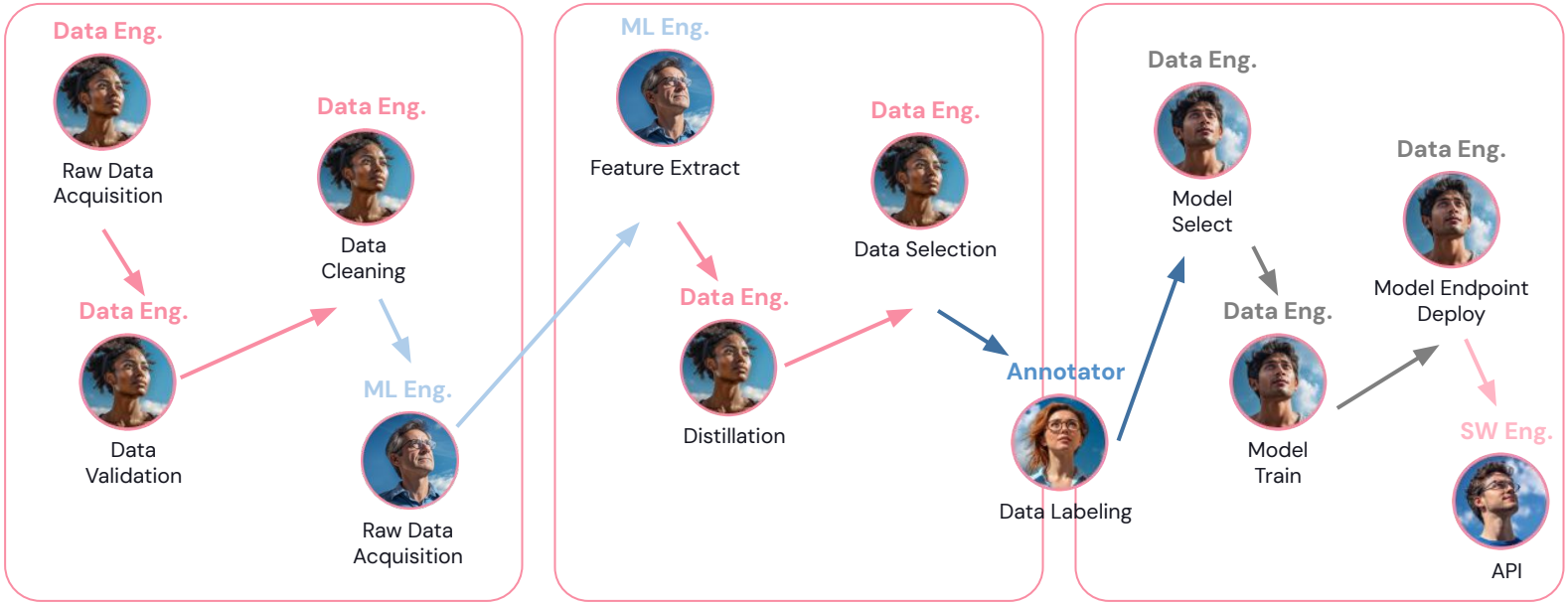
Great AI applications are created through continuous workflows



AI models have become commoditized, shifting the development cycle from model-centric to data centric



Home-grown AI development has become a series of disconnected steps with too many tools & people



Gen AI requires higher data quality & flexible data infrastructure to adopt latest AI technologies

Enterprise AI teams saw

2/3

of projects fail due to poor data quality.

While Hugging Face added

540,000

new models since 2021.

Locking into a single cloud leaves compute resources unused and wastes time rebuilding infrastructure

Companies are overspending by

40%

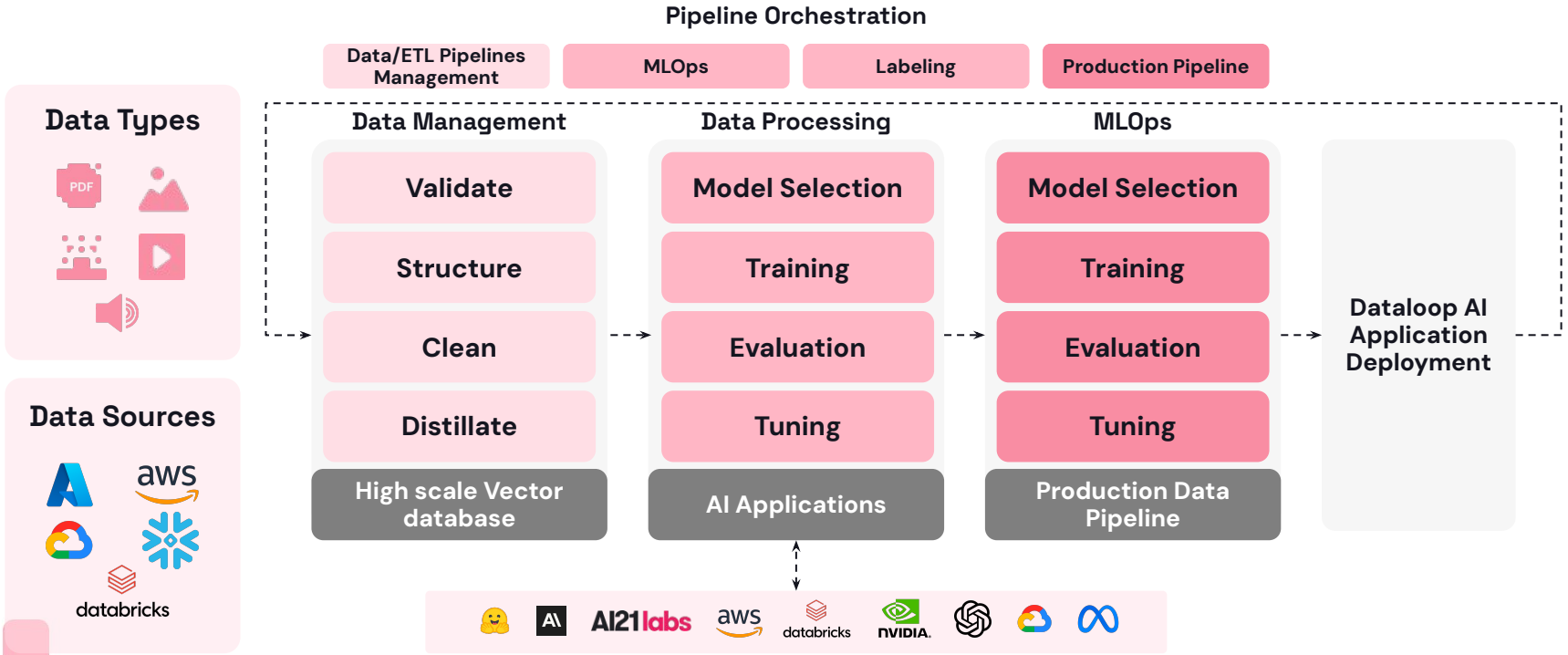
by not optimizing their cloud compute resources.

While engineering teams spend

3x

more time rebuilding AI infrastructure on each public cloud.

The fastest way to build robust AI applications



How Dataloop makes a difference

70%

Time saved

Silos don't scale. Get your entire data org working together.

10x

Higher quality

By using easy-to-build feedback loops between humans and machines.

95%

Automation

Across the average Dataloop pipeline. Reduce manual work to near-zero.

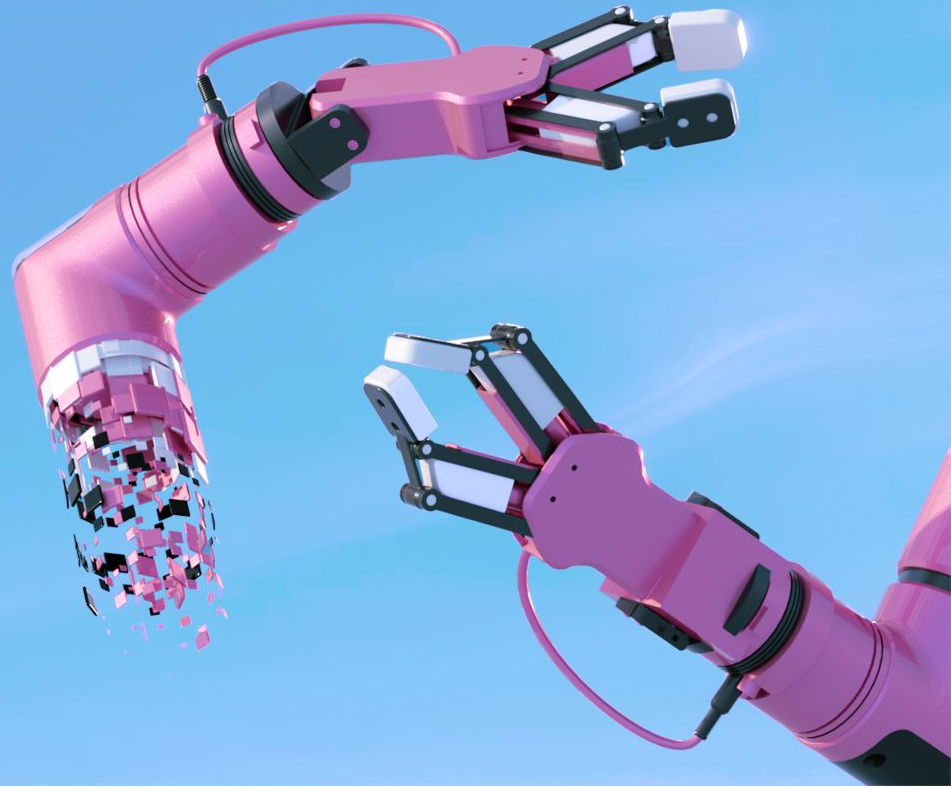
20x

Faster development

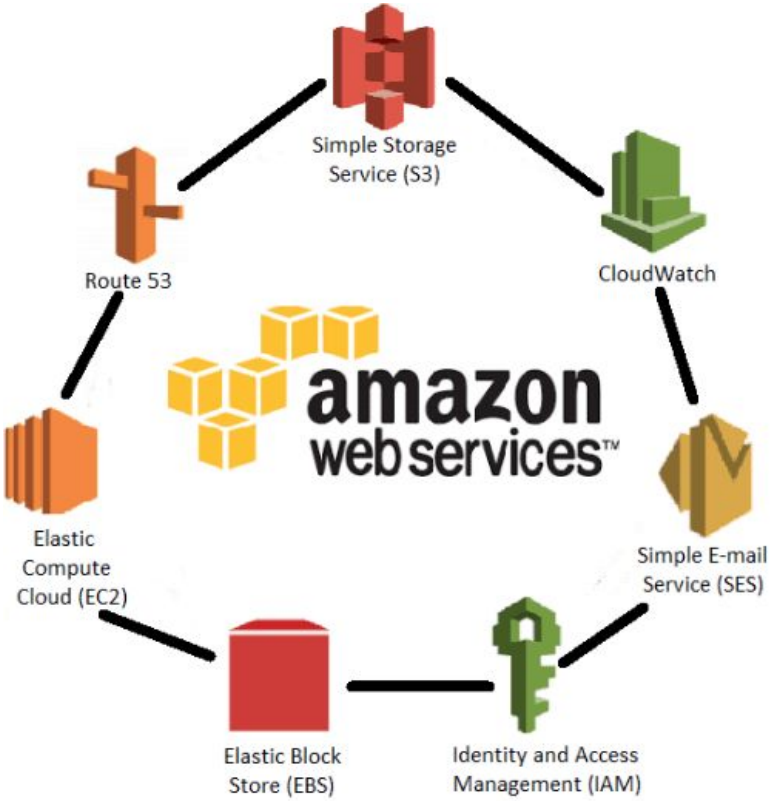
Stay ahead of the AI Curve and meet market demand.

Claudine

Understanding Cloud Costs



Overview of Cloud Compute Resources



Cost Factors in Cloud Infrastructure

How do you pay for AWS?

Pay-as-you-go

Pay-as-you-go allows you to easily adapt to changing business needs without overcommitting budgets and improving your responsiveness to changes. With a pay-as-you-go model, you can adapt your business depending on need and not on forecasts, reducing the risk of overprovisioning or missing capacity.

Read more »

Save when you commit

For AWS Compute and AWS Machine Learning, Savings Plans offer savings over On-Demand in exchange for a commitment to use a specific amount (measured in $/hour) of an AWS service or a category of services, for a one- or three-year period.

Read more »

Pay less by using more

With AWS, you can get volume based discounts and realize important savings as your usage increases. For services such as S3, pricing is tiered, meaning the more you use, the less you pay per GB. AWS also gives you options to acquire services that help you address your business needs.

Read more »

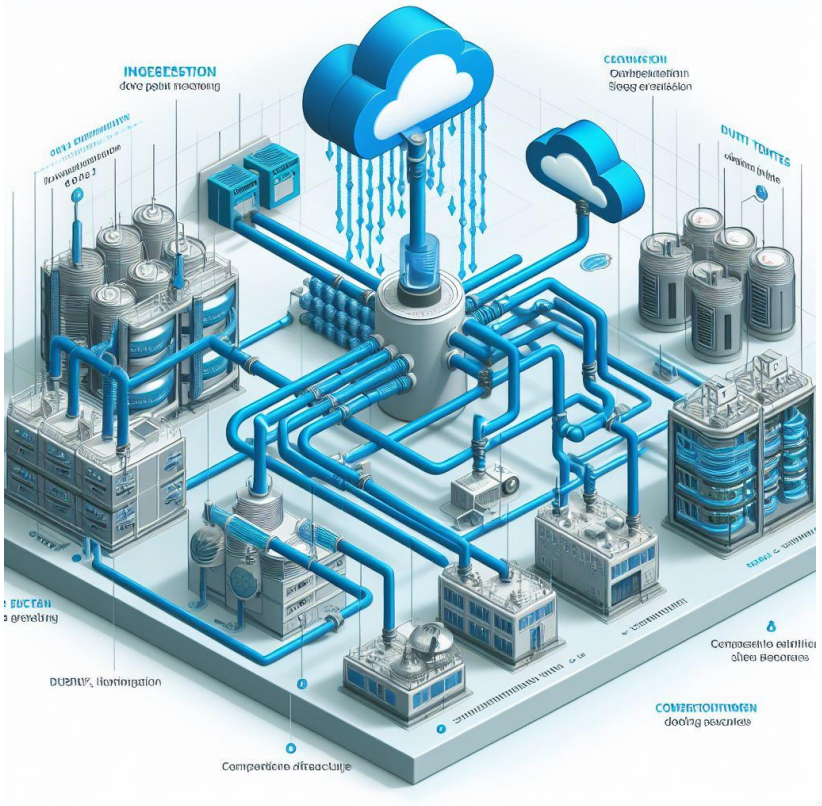
Common Challenges in Managing Costs

- Reporting
- Tagging
- Control Cost
- Planning
- Optimize Utilization
- Oops!!! I forgot to turn off the machine over the weekend

Strategies for Cost Optimization

Right Sizing Compute Resources

- Understand the Business
- Understand Unit Economics
- GM Target
- What's the Infrastructure Looks Like?
- What are the Cost Drivers?
- Reporting
- Auto Scaling



Leveraging the Commitment

- Reserved Instances
- Savings Plans
- EDP
- Spot Instances
- Other commitments

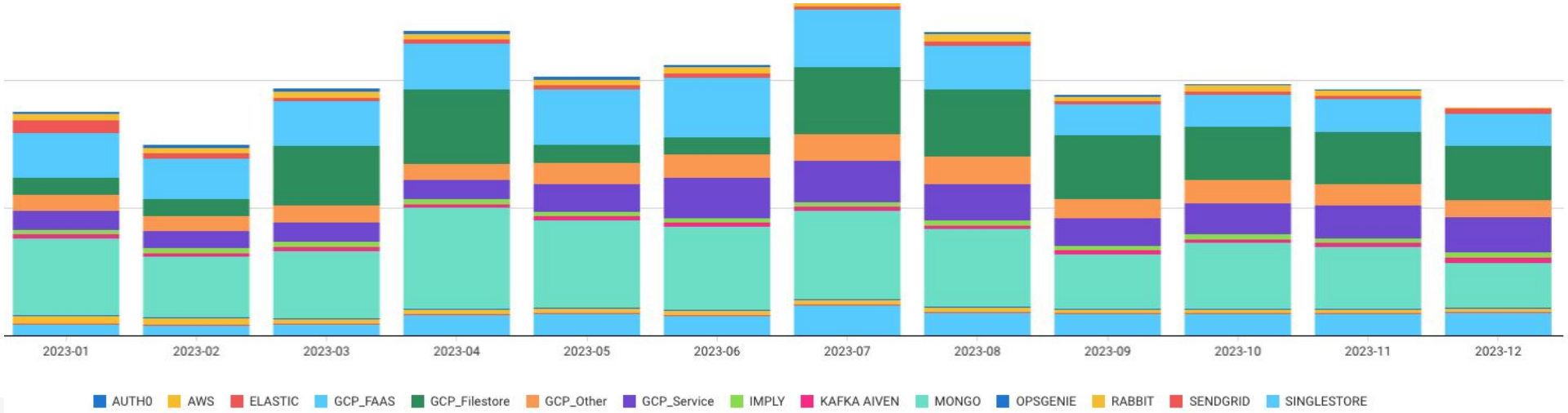
The Best Strategy of All

- Be Friends with the Relevant Stakeholder

Tools and Best Practices

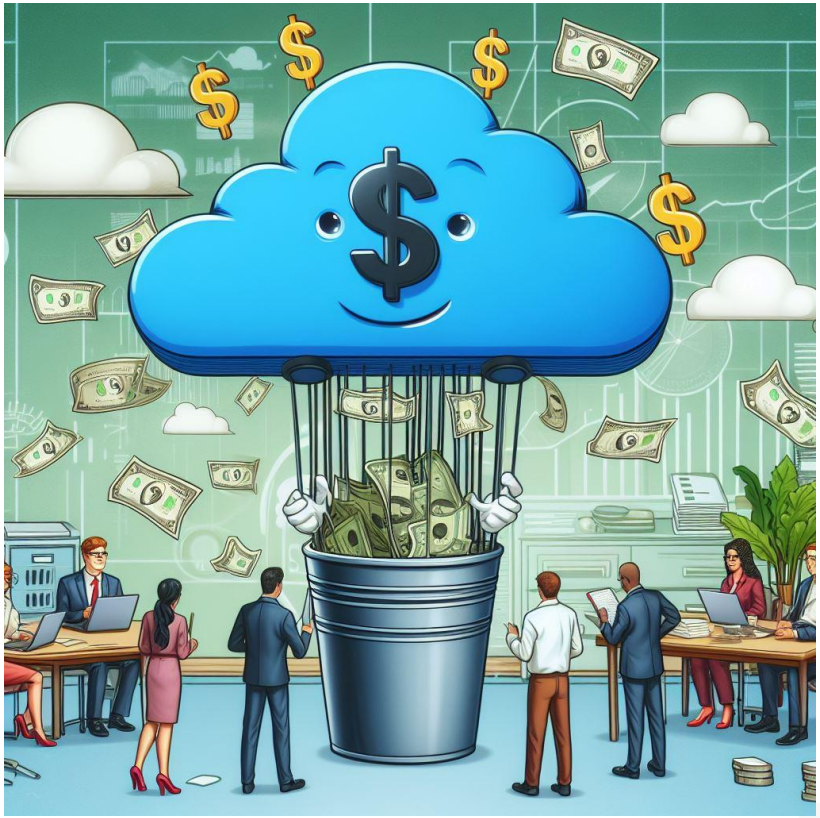
Cost Monitoring and Reporting Tools

- Billing Tools
- 3rd Party Monitoring Tools (Cloudwatch)
- Open Source Monitoring Tools
- Your BI



Implementing Policies

- Lead Without Authority!!!
- Agree on a Plan and Targets
- Set a Weekly Check Up Meeting
- It’s STRATEGIC



Case Study

- Cloud is the 2nd Highest Expense
- GM Target and Timeline was set and Agreed with All Stakeholders
- A Weekly Meeting is Held to Track the Progress
- Result – Improved GM and Unit Economics
- Strategic Takeaway – The Org Trusts the Data and the Finance Team Takes a much more Strategic Part as It Continues to Deliver Value

p	Owner	current price (monthly)	timated price decrese (mont	due date	Jira
Closing Dev clusters	Aharon			29 May 23	DAT-47720
Services preemptible adoption	Naor			21 Aug 23	DAT-47801
Reduce rc compute in clusters over night & weekends	Aharon			22 Aug 23	
ETL lazyness and inside Rubiks - Image Pre processing	Shachar			29 Aug 23	DAT-49251
Logs level reduction	Koby/Naor			29 Aug 23	
FlexSave EOL - Move to GCP commit	Naor			01 Sep 23	
Move chatGPT, GitHub, 365	Eyal			06 Sep 23	
Services utilisation top 3	Naor			15 Sep 23	DAT-44709
Storage without owner (11.2 TB)	Shachar			02 Oct 23	DAT-11741
use doit as aws supplier	Eyal			15 Oct 23	RPP-38
imply (from Droid) to managed service	Guy			01 Nov 23	DAT-51183
Mongo replaced by S2 RC (Rubiks)	Shachar			31 Mar 24	
RabbitMQ resize (RC)	Naor			15 Jan 24	DAT-62614
Mongo replaced by S2 PROD (Rubiks)	Shachar			01 Apr 24	
Mongo replaced by S2 Ford (Rubiks)	Shachar			15 Apr 24	
Mongo replaced by S2 Syngenta (Rubiks)	Shachar			30 Apr 24	
~~AWS S3 storage - Cleanup~~	~~Naor~~			~~01 Feb 24~~	~~DAT-62613~~
Ramp down from TestRail	Boaz			10 Feb 24	DAT-57969
Services node pools optimizations (Preserve pod)	Naor			31 Mar 24	DAT-49455
Move global FaaS to Azure (using existing credit)	Koby			16 Feb 24	
Buy more GCP CUD's	Naor			13 Mar 24	DAT-63022
Services utilisation - Rubiks PROD (Ram, CPU, Replicas)	Naor			15 Apr 24	DAT-49317
Phase1: inner cluster (KubeDB) RC	Naor			01 May 24	DAT-47913
~~Elastic servers (FaaS Logs)~~	~~Aharon~~			~~29 Feb 24~~	
Catalog refinement	Aharon			30 Nov 24	DAT-47433
Billing monitoring (RC - off, Production - hourly)	Guy			1 Mar 24	
Reduce min Rubiks scale (API stream uploads)	Shachar			01 Mar 24	DAT-63024
Reduce Rubiks API, Geisha (post Async bulk)	Shachar			30 Apr 24	DAT-63025


Thank you