Lecture HSE · May 23, 2026

Cloud Infrastructure for AI Business

How managers choose a platform, govern complexity and evaluate infrastructure economics

Alexander Polomodov

Speaker Context

Technical Director & Fellow, T-Technologies

- Architecture and engineering R&D.
- AI adoption in SDLC.
- Focus: cloud platform choice.

Introduction | After Cloud Basics

Here cloud is a management choice

Options, criteria and regular governance

No basic recap

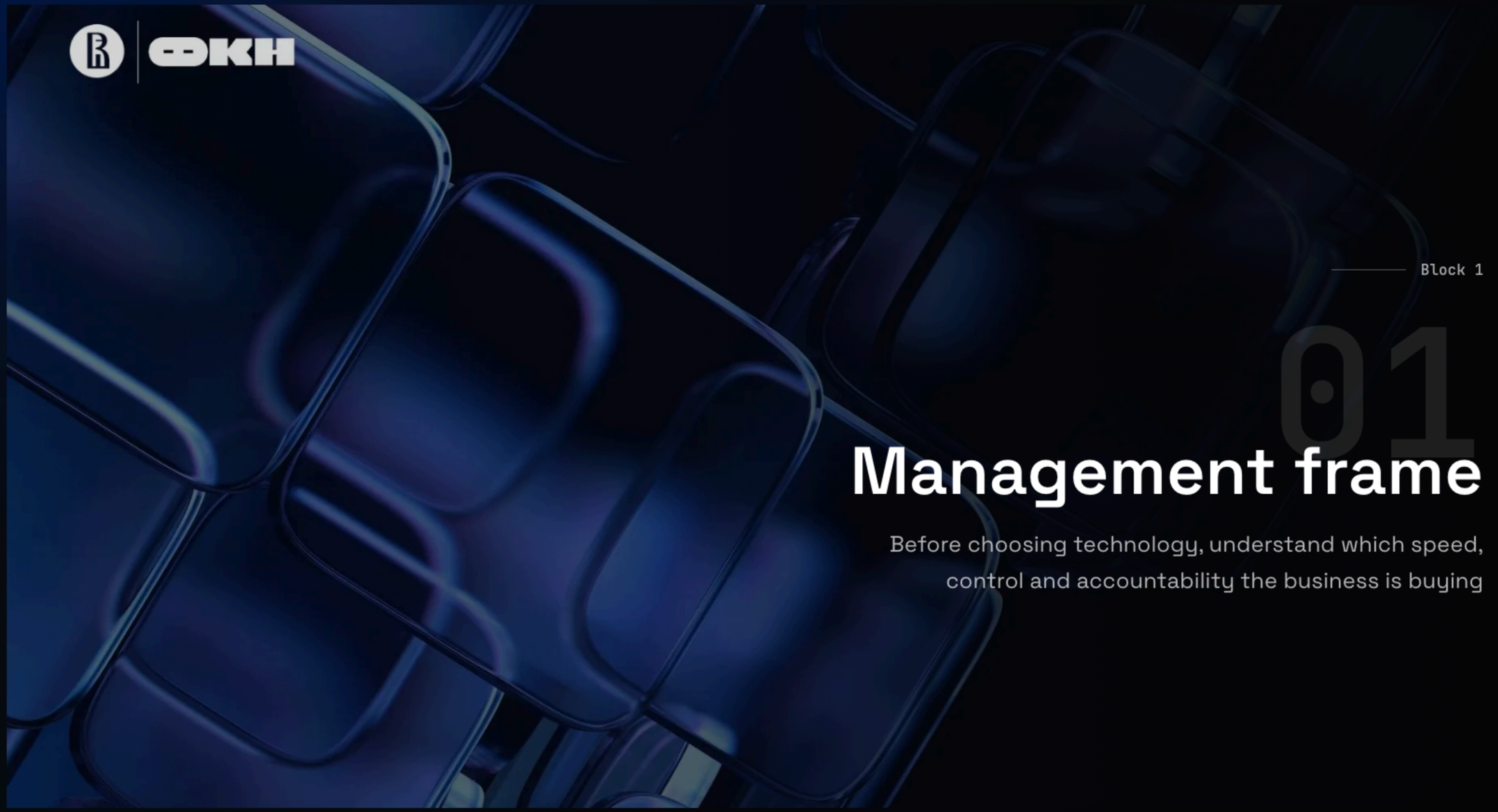
Cloud, IaaS/PaaS/SaaS — in notes.

We look like managers

Options, acceleration, risk.

We calculate economics

We keep the resource, hardware, payback and FinOps examples.



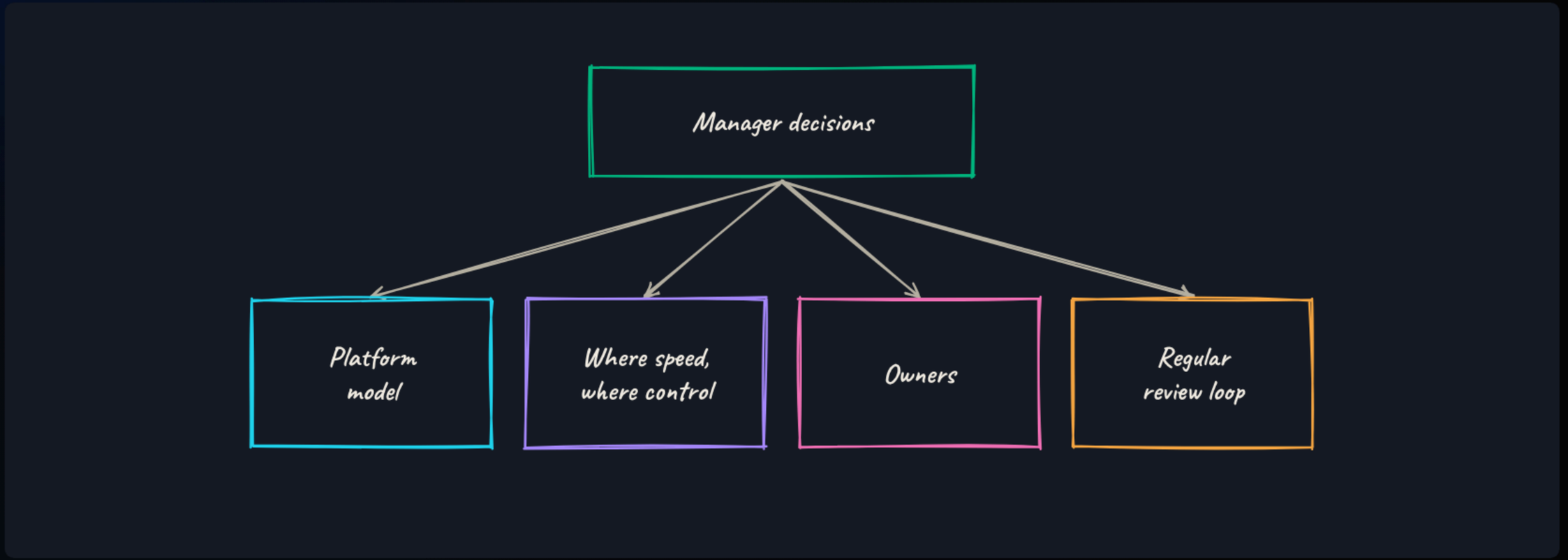
Block 1

01

Management frame

Before choosing technology, understand which speed, control and accountability the business is buying

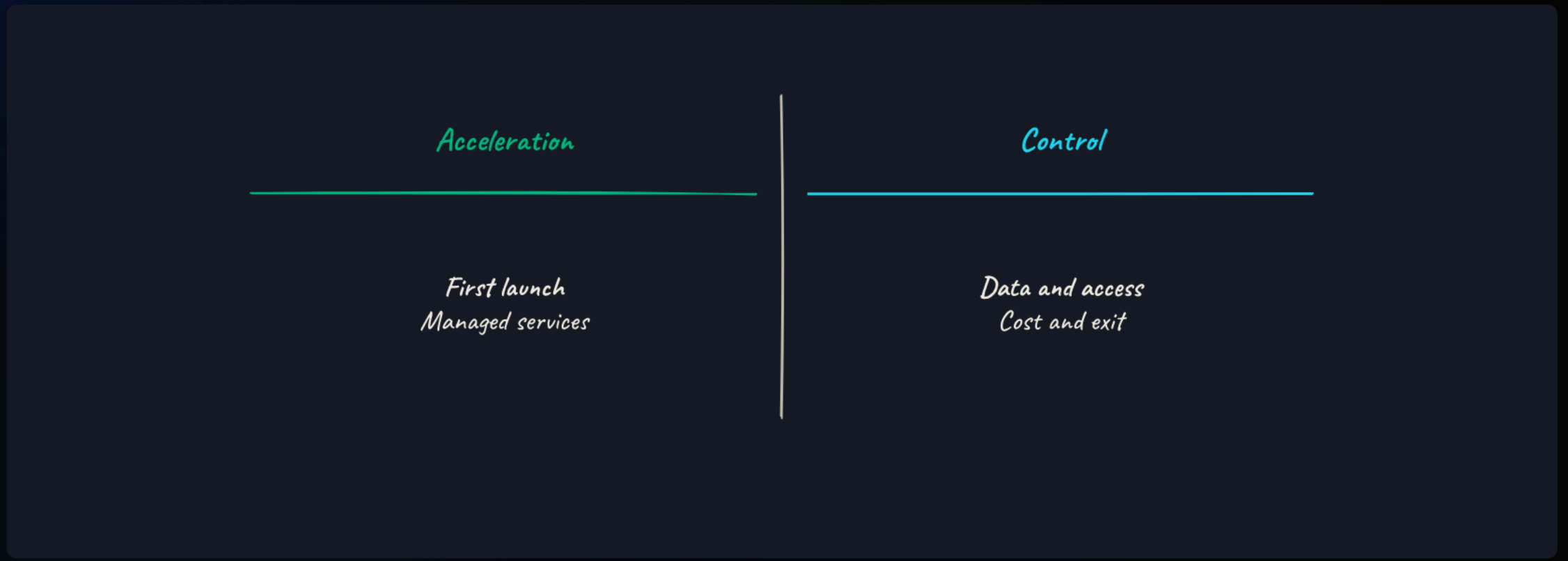
Which decisions managers actually make



A platform portfolio beats one answer

Public / managed	Private platform	Hybrid strategy
Speed and services Provider-rule dependency	Scale and compliance Predictable load	Core under control Spikes stay in cloud

Evaluate speed against control





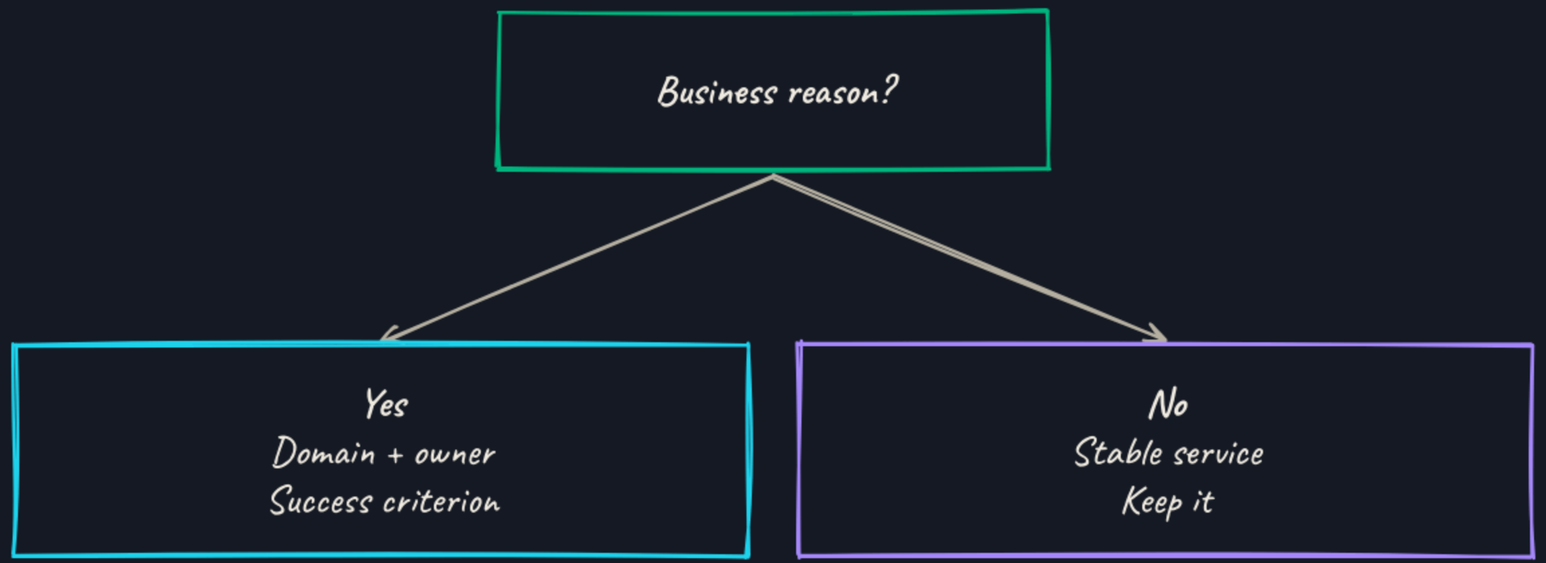
Block 2

02

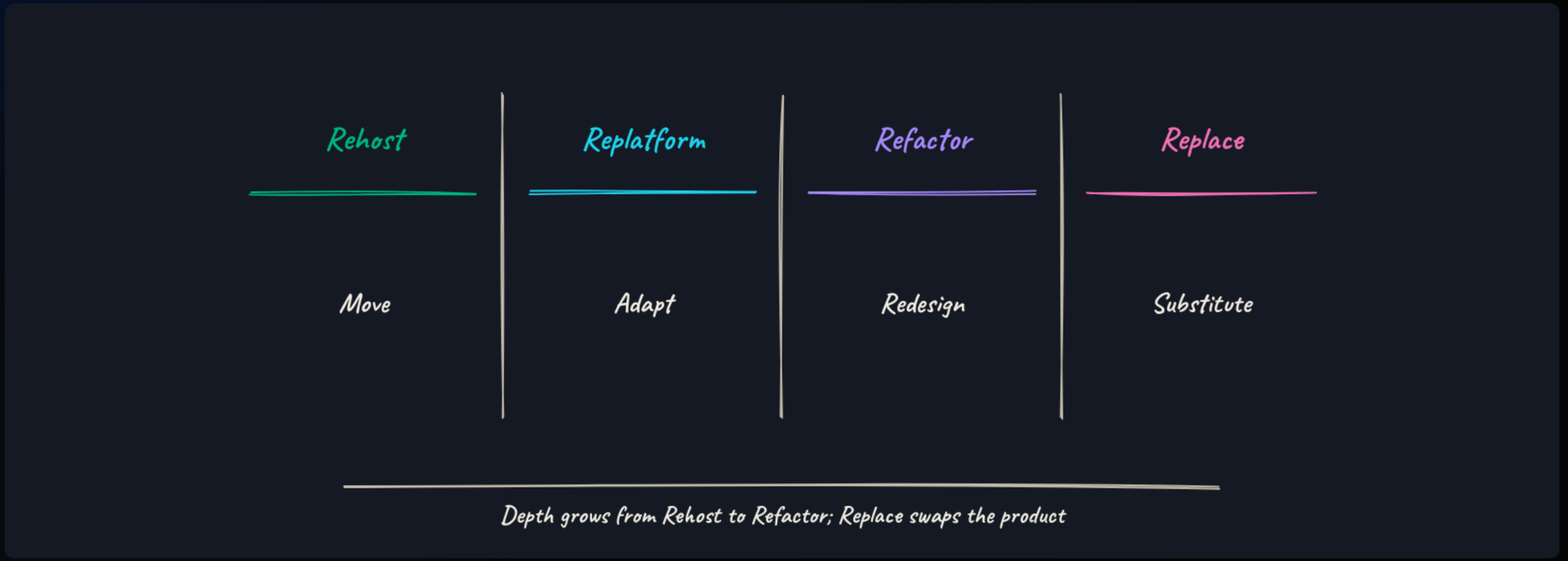
Migration as a business program

Move not servers, but the model of speed, risk, accountability and change governance

Migrate only for business pull



Choose migration depth for the goal



Executive anti-patterns

- Migration without priority and owners.
- Move itself beats speed/reliability/economics.
- Unclear: on-call, changes, risk, bill.
- Old delivery process in new cloud.



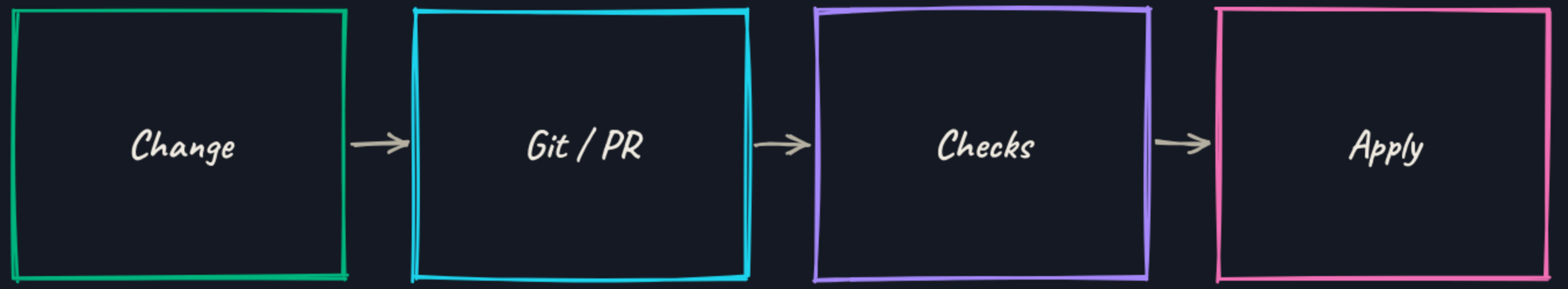
Block 3

03

Platform governability

Repeatable safe infrastructure change

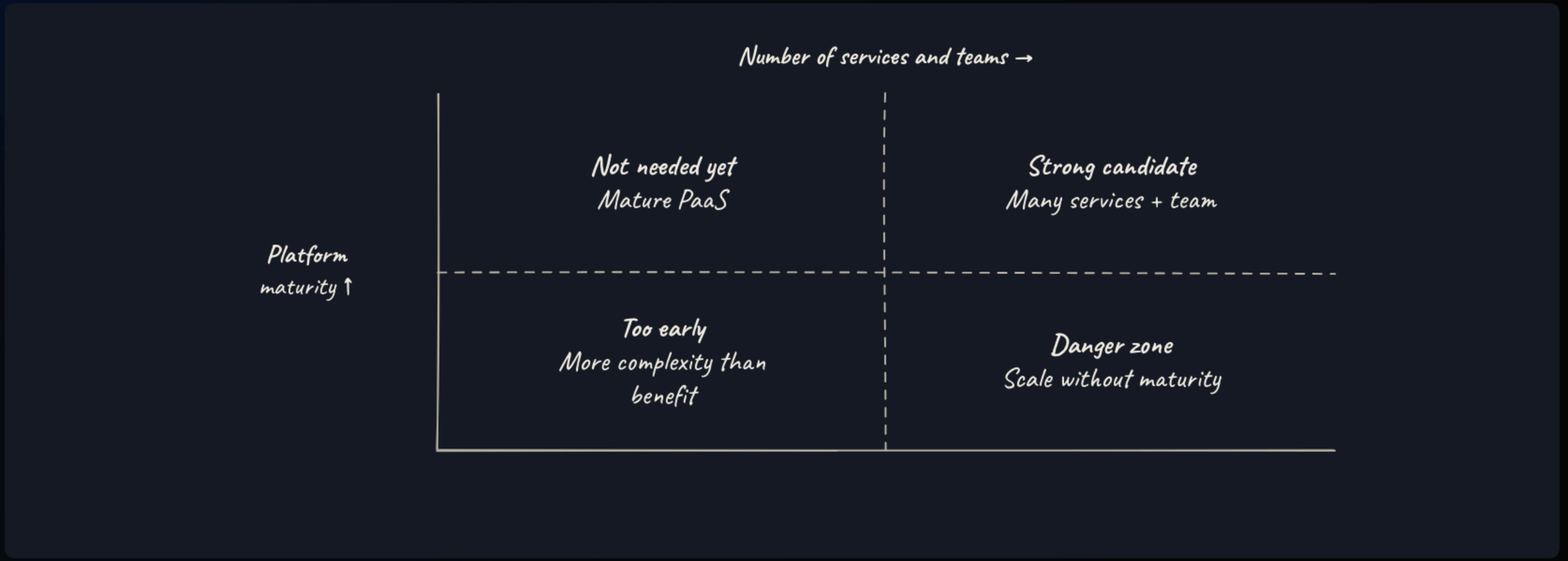
IaC is management control



IaC makes changes reproducible and auditable

Deliberate choice

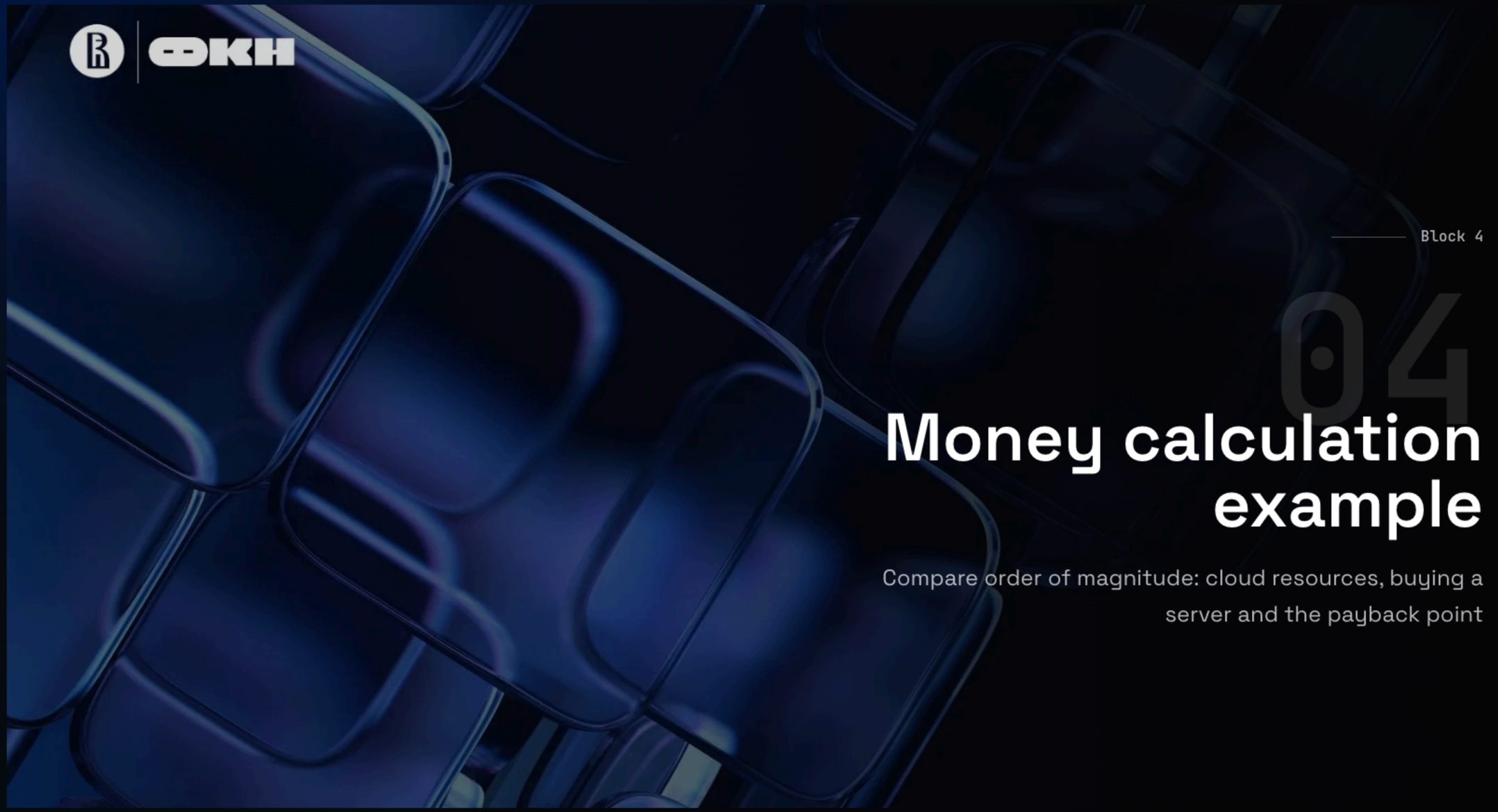
Kubernetes must earn its complexity



Who is responsible for what

Operating model

Platform	Product	Governance
Templates Reliability	Value Workload	Risk Access and TCO



Block 4

04

Money calculation example

Compare order of magnitude: cloud resources, buying a server and the payback point

May 2026 prices

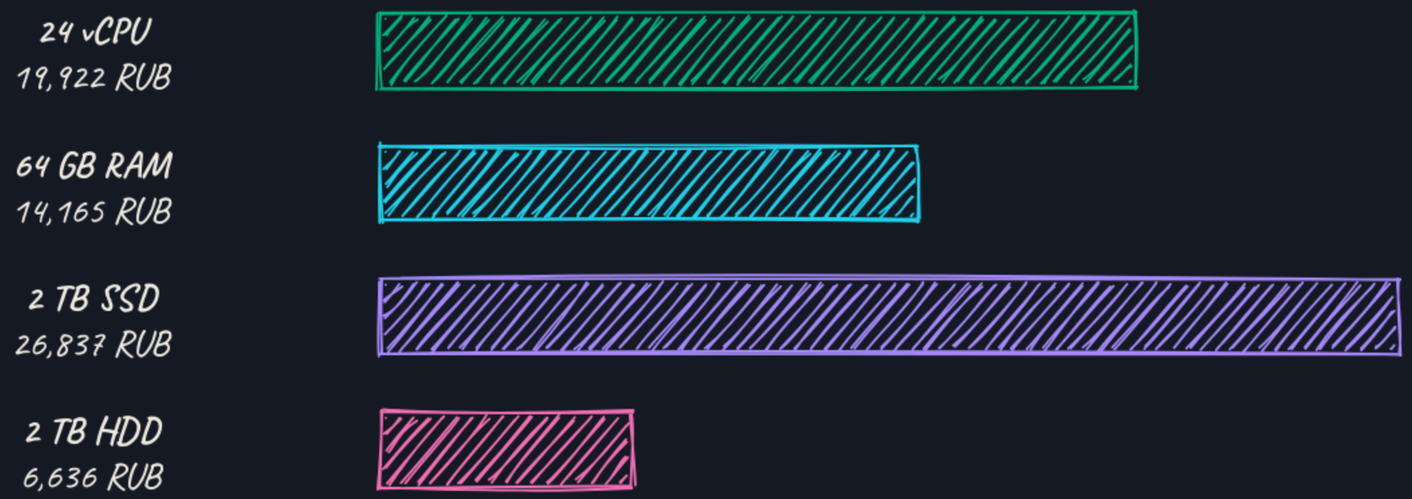
Price per cloud resource unit

vCPU 100%	RAM, GB	SSD, GB	HDD, GB
1.1529 RUB/h ≈830 RUB/mo	0.3074 RUB/h ≈221 RUB/mo	0.0182 RUB/h ≈13.1 RUB/mo	0.0045 RUB/h ≈3.24 RUB/mo

Rate × 720 hours × quantity; Yandex Cloud rates as of May 2026

Configuration

Cloud equivalent of a regular server



Total: SSD 60,924 RUB/month or HDD 40,723 RUB/month

Cost of similar hardware

Catalog example

- Dell PowerEdge R750: 24 physical cores.
- 64 GB RAM in base configuration.
- Catalog price: from 566,050 ₽.

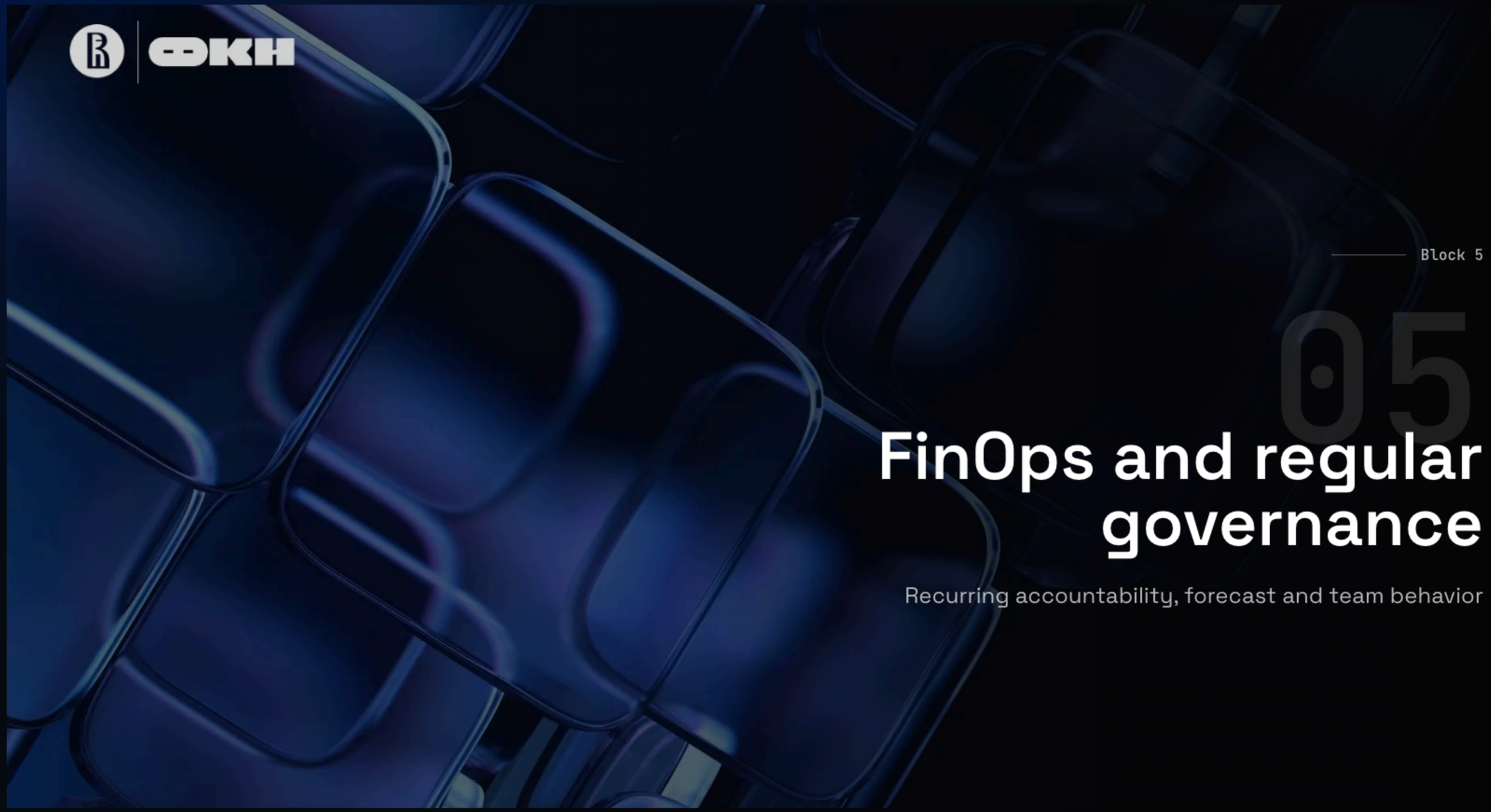
What is not included

- noHDD: price disks separately.
- No DC, power, network, licenses, ops.
- Educational order-of-magnitude estimate.

Payback point and private cloud at scale

Hardware	CPU + RAM	With SSD
566,050 RUB	34.1k RUB/month	60.9k RUB/month
No disks or ops	≈16.6 months	≈9.3 months

Rough: vCPU ≠ physical core; full TCO includes data center, disks and team



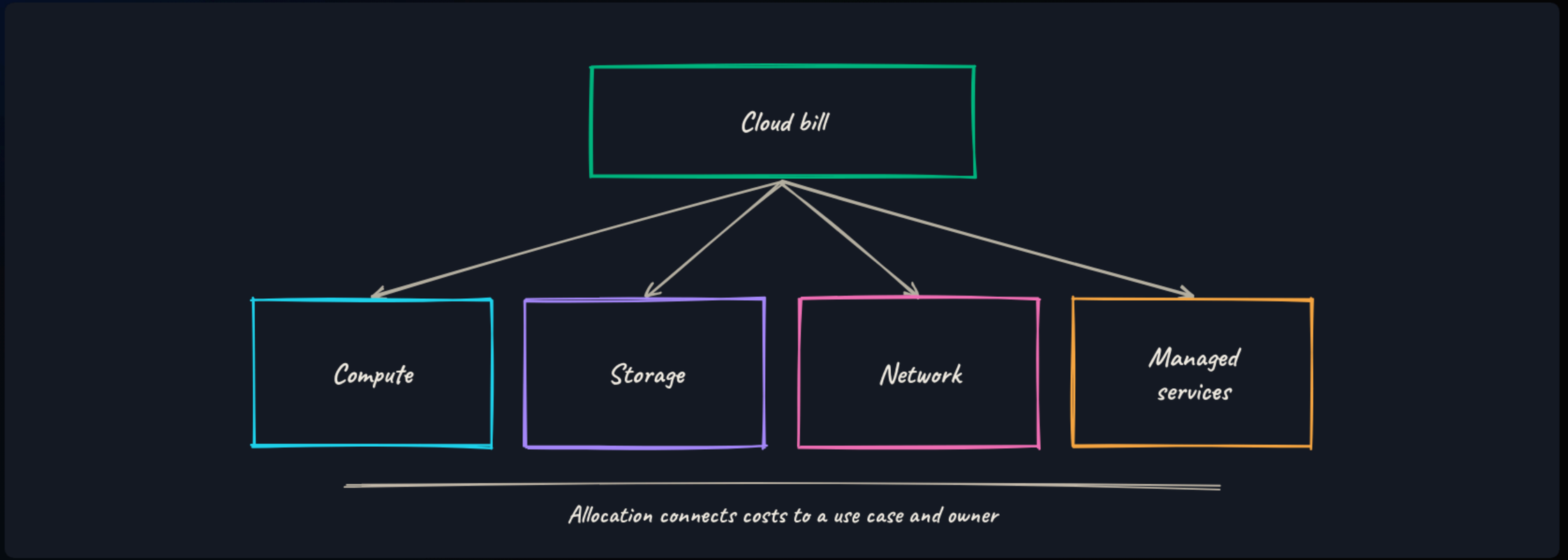
Block 5

05

FinOps and regular governance

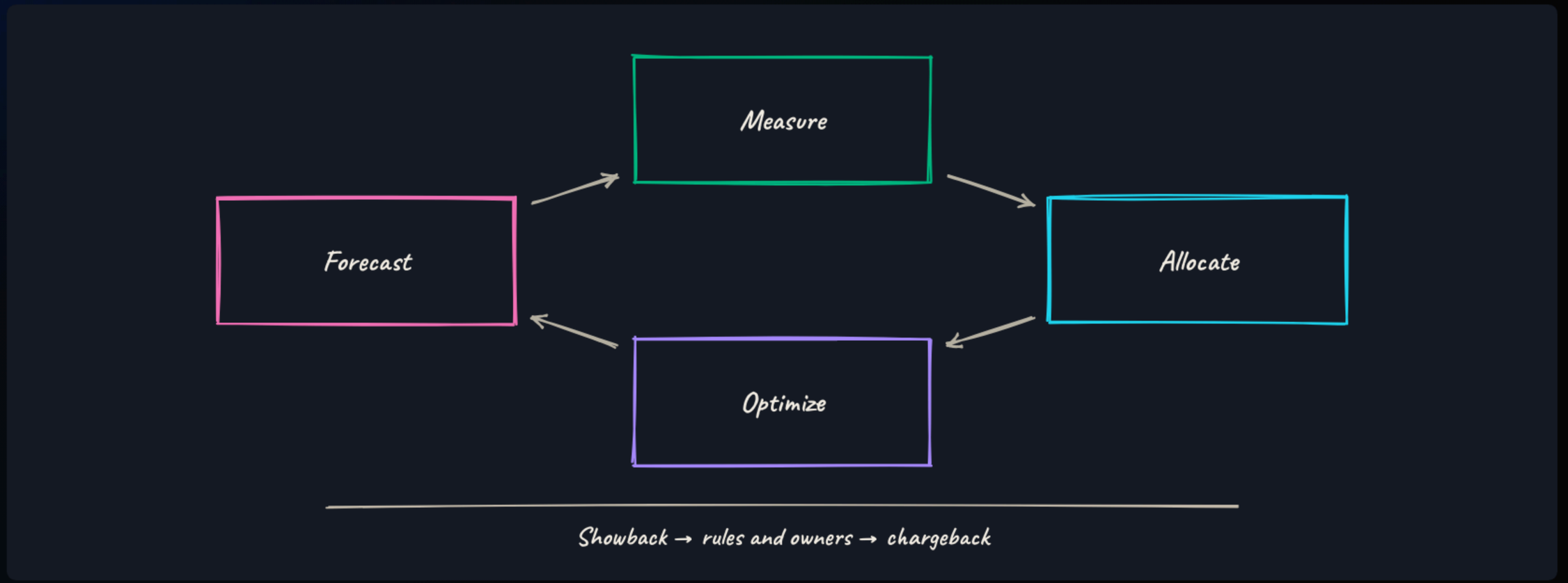
Recurring accountability, forecast and team behavior

Cost drivers and expense allocation

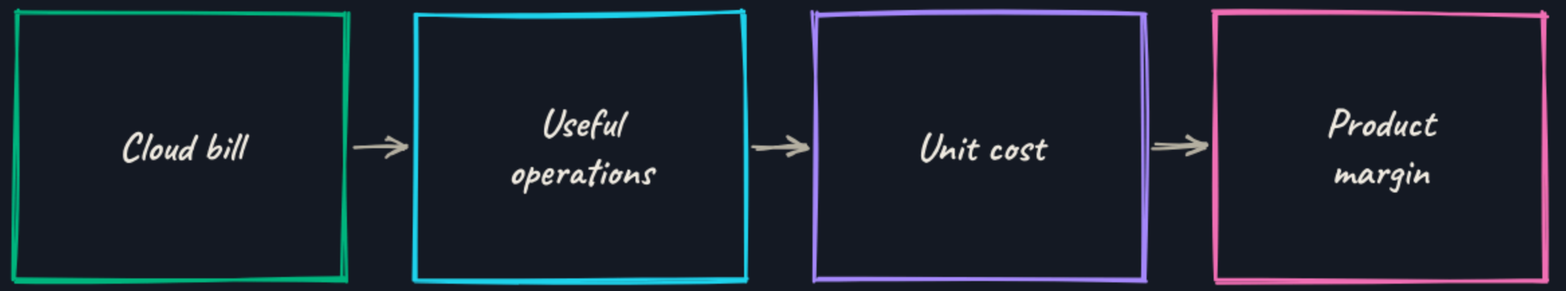


FinOps loop: cost allocation, optimization, forecast

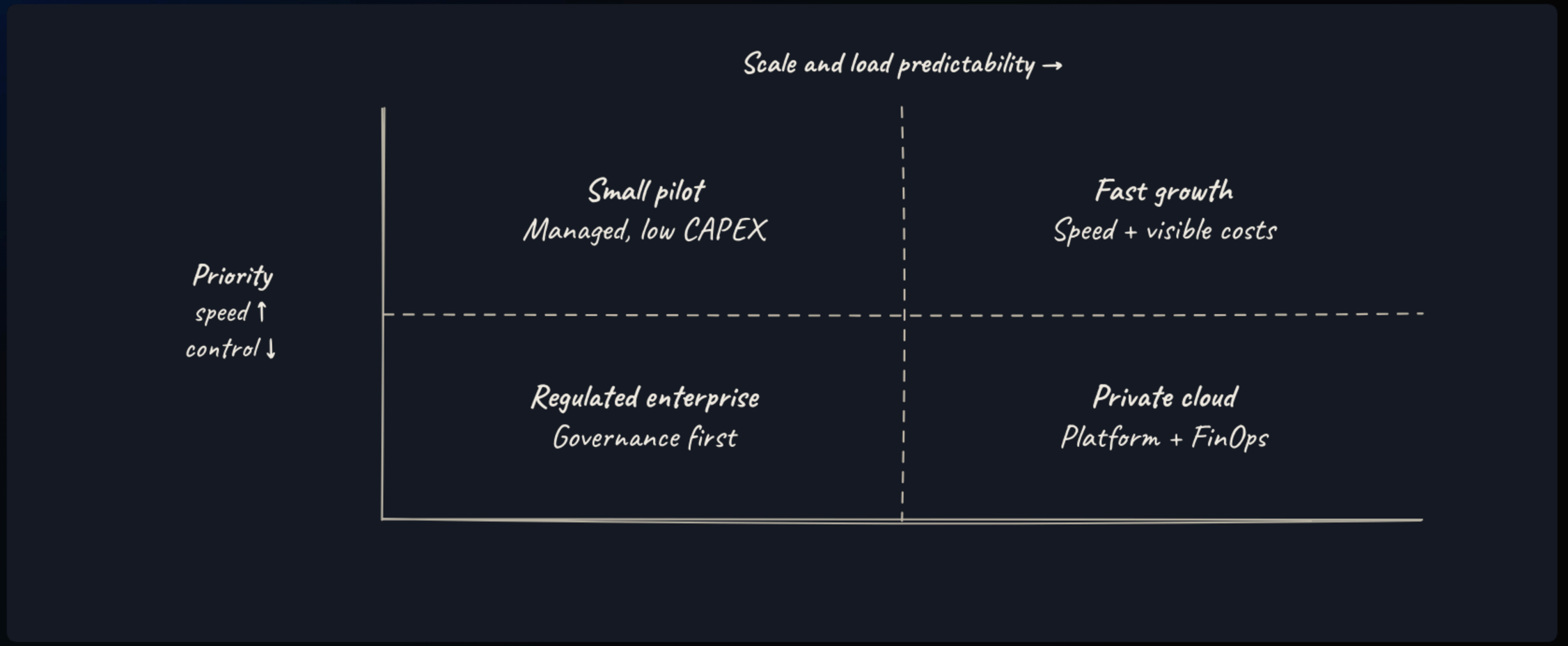
Block 5 · Loop FinOps



Unit economics instead of total invoice size



Final decision matrix



References and materials

Sources | Materials

Core sources

- FinOps Foundation - FinOps Framework: ↗finops.org
- CNCF Kubernetes docs: ↗kubernetes.io · ↗cncf.io
- Cloud architecture frameworks.
- Yandex Cloud pricing and Servermall R750.