

Scaling Systems, Scaling Influence

Lessons from Distributed Architecture

Mallika Rao

Engineering Manager, Netflix

The Staff+ Shift: Beyond the Code

Influencing Decisions

Moving from technical expertise to guiding strategic outcomes required a completely new approach.

Gratification Reimagined

My success was no longer measured in merged PRs, but in the success of others and the evolution of strategy.

Long-Term Vision

Shifting from sprint-to-sprint execution to envisioning and shaping the future became paramount.



The Staff+ Shift: Beyond the Code

Influencing Decisions

Moving from technical expertise to guiding strategic outcomes required a completely new approach.

Gratification Reimagined

My success was no longer measured in merged PRs, but in the success of others and the evolution of strategy.

Long-Term Vision

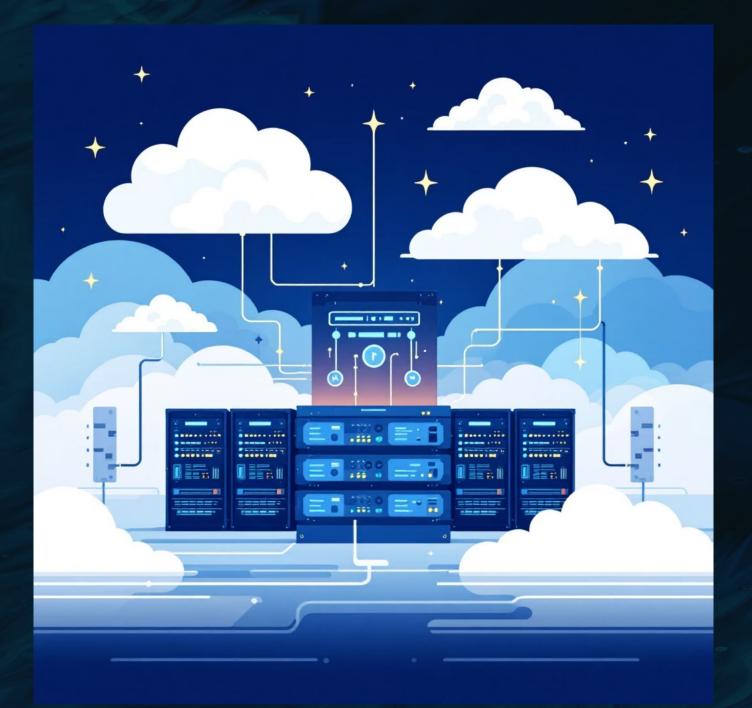
Shifting from sprint-to-sprint execution to envisioning and shaping the future became paramount.

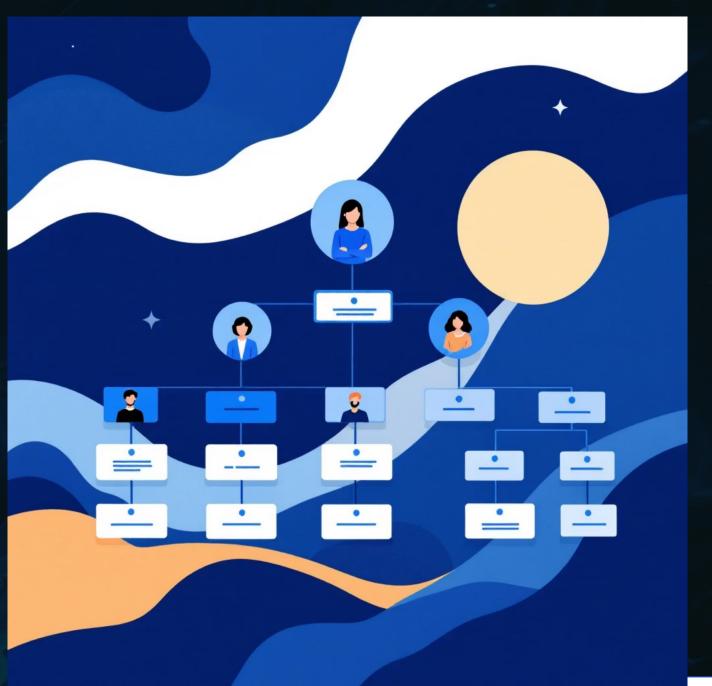




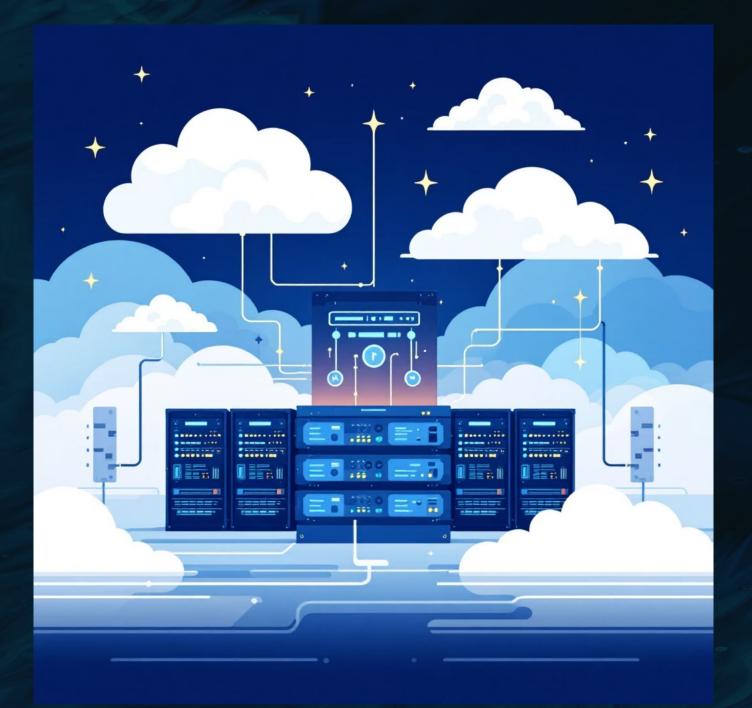
Why is scaling at Staff+levels so hard?

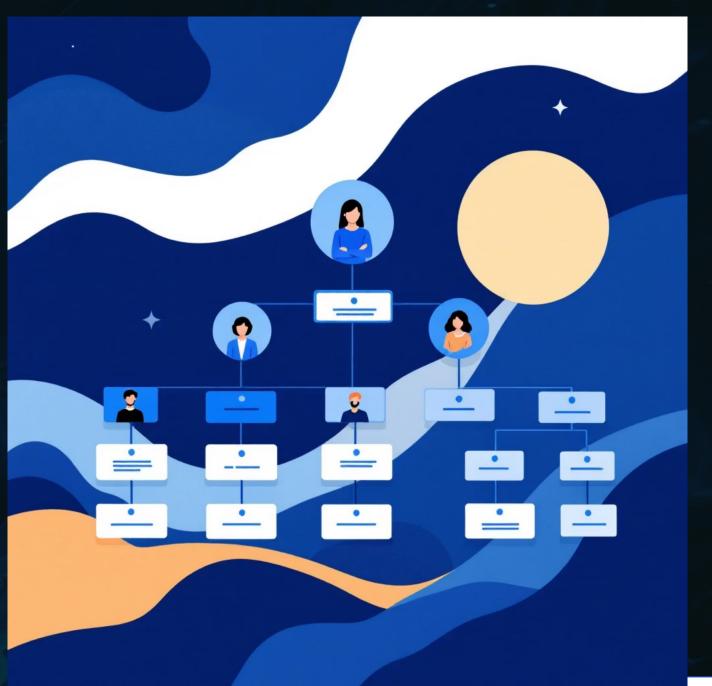
Scaling Systems vs. Scaling Influence





Scaling Systems vs. Scaling Influence

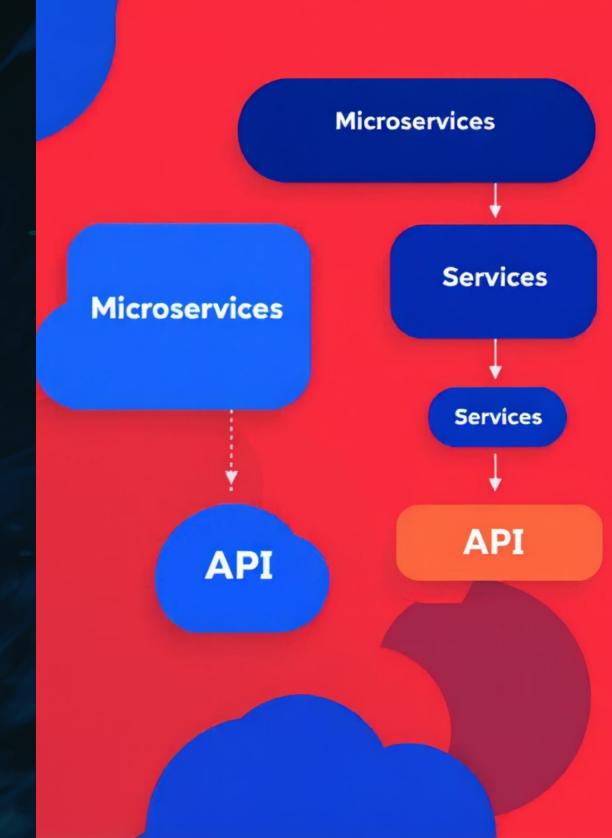




Autonomy Without Chaos

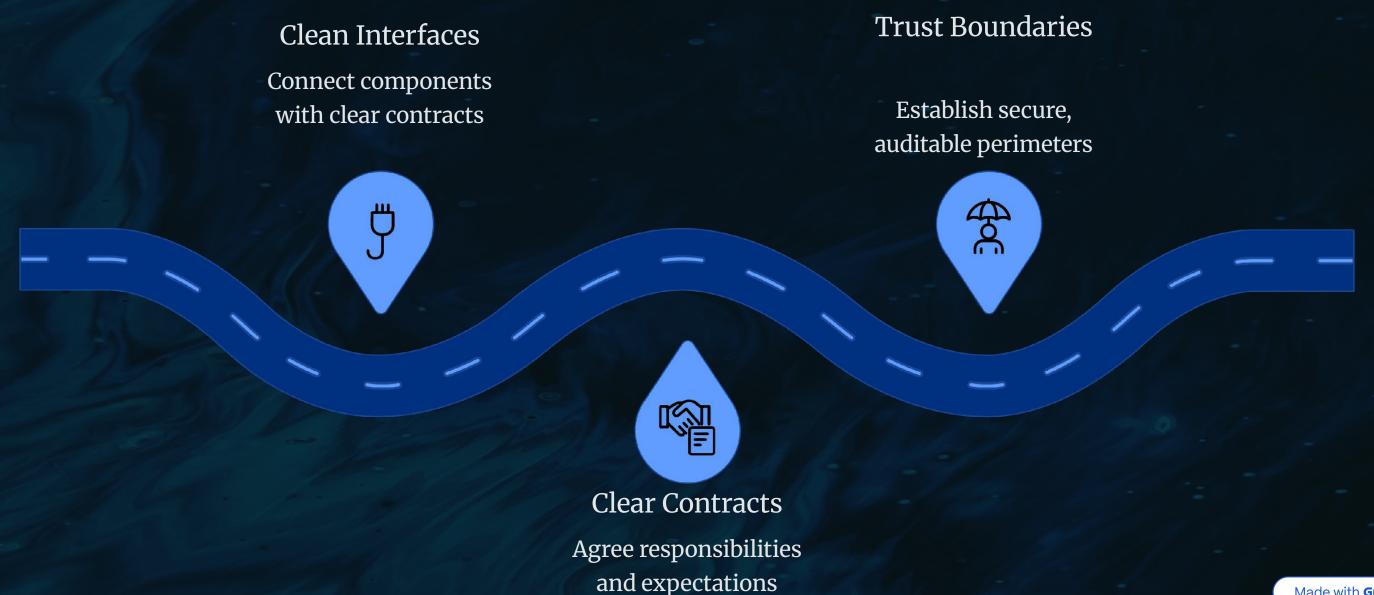


Guardrails + Interfaces > Micromanagement



Designing Decision Protocols

Your job shifts from making decisions to designing the protocols that let others make good ones.



Fault Tolerance for People



Build redundancy & graceful degradation in leadership



Building Organizational Fault

Tolorana







Delegation: Replicate
Leadership
Empower others, building
redundancy in leadership.

Deprioritize: Create
Capacity
Remove tasks to make space for others, shifting to collective capacity.

Redundancy: Code & Org Structure Design resilient systems and structures for continuous performance.

Your success at scale is measured by how gracefully your systems and teams perform when you're not in the loop.

Eventual Consistency in Alignment

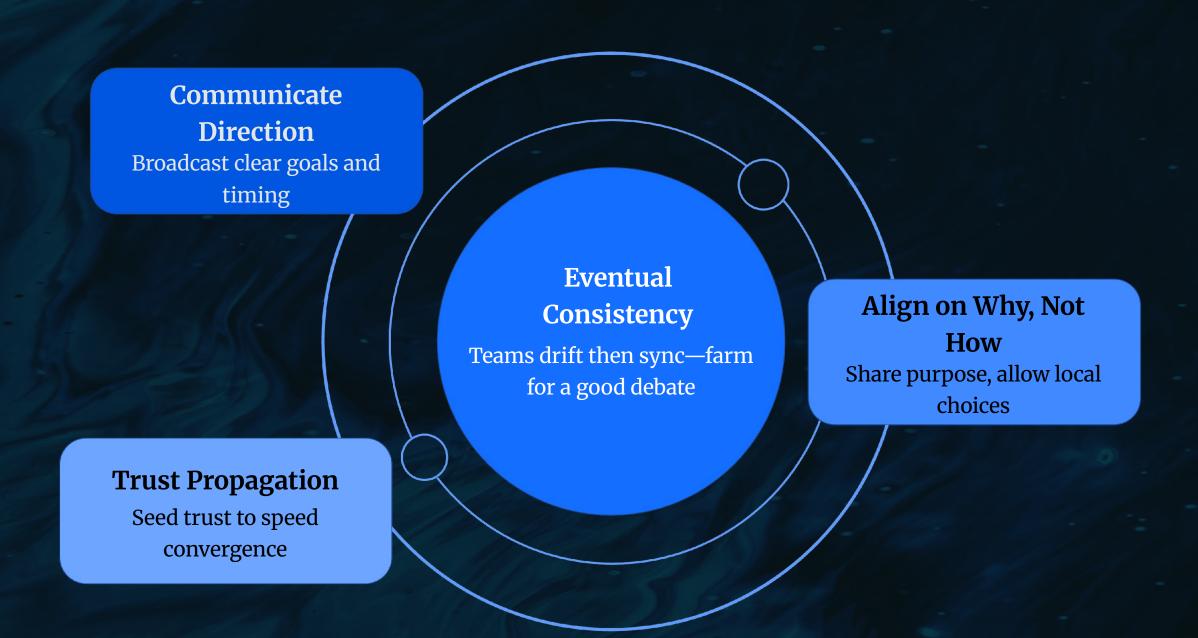


Don't block for instant consensus



Eventual Consistency in Organizations

"Silence doesn't always mean misalignment. Sometimes, it's just... network latency."



A Tale of Two Indexes: A Search Challenge

Tweet Index

- Massive scale, optimized for speed
- Trillions of tweets
- Sub-second latency

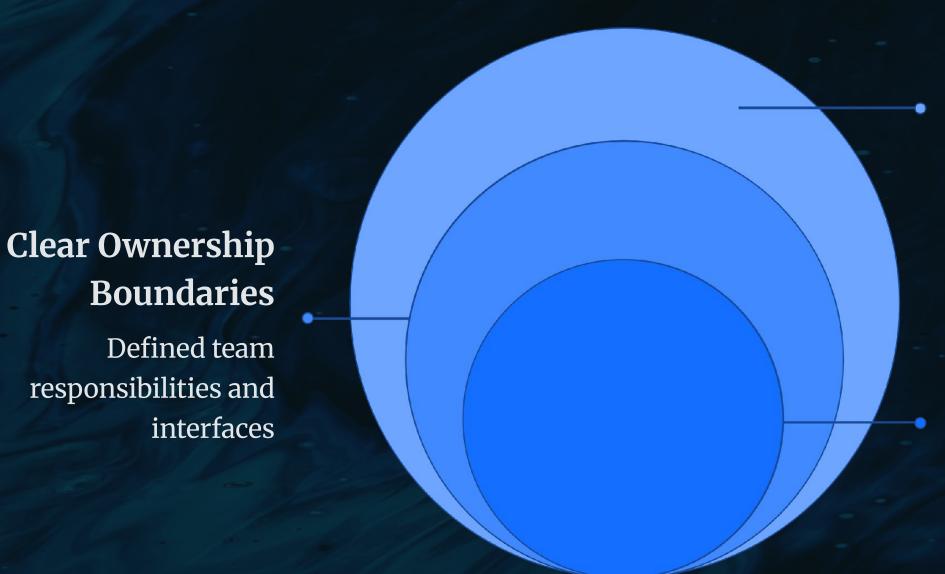
Flexible Index

- Smaller, flexible
- Media, profiles, events
- Adaptable content

How do we unify speed and flexibility without compromising reliability?



Systems Thinking in Action: The Unified Solution



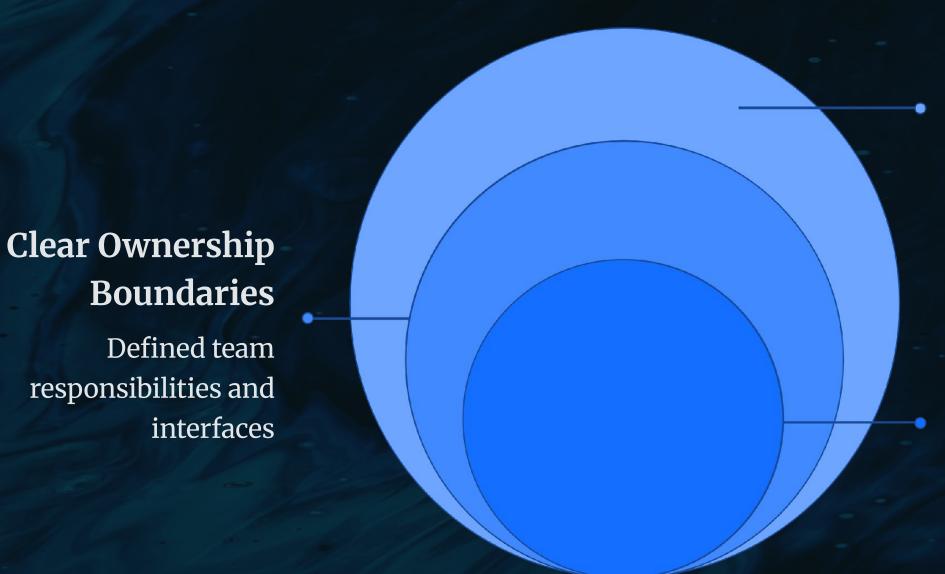
Shared ObservabilityTools

Common metrics, logs, and tracing platform

Unified Serving Layer

Single API surface for serving requests

Systems Thinking in Action: The Unified Solution



Shared ObservabilityTools

Common metrics, logs, and tracing platform

Unified Serving Layer

Single API surface for serving requests

Your Three Levers at Scale

Strategy

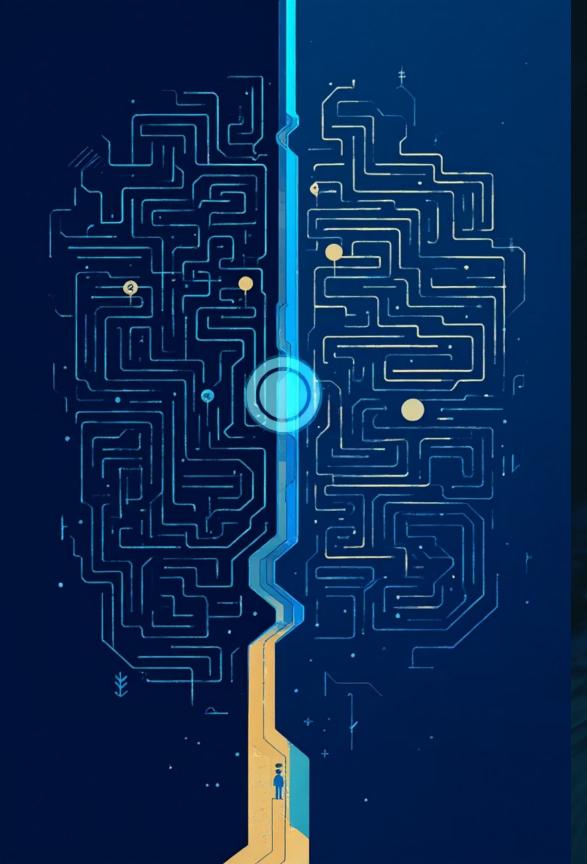
Pick the right bets and capabilities to build

Communication

Make your thinking legible so others can align

Partnerships

You can't scale alone; allies amplify your reach



From Complexity to Clarity

What system principle will you apply next week?

