Scaling Workflows Automation Platform for Intuit Peak

Nivedita Nayak, Principal Engineer, QuickBooks Advanced, Intuit
Raghav Agarwal, Software Engineer 2, QuickBooks Advanced, Intuit

September 2021

Agenda

Introduction
About Us

Workflow Landscape/Architecture Recap
Workflow Use cases @ Intuit
Architecture
Maturing as we are Expanding

Journey so far
Accomplishments across 2 Peaks

Challenges
DB Bottlenecks
Horizontal vs Vertical Scale
Model Challenges
Definition Challenges

How we did it?
Testing and more Testing
Observability

Learnings
History
DB Management
Single Definition
Multi Tenancy

Q&A

About Us

Nivedita Nayak

- Principal Engineer at Intuit. Been with Intuit for 5 plus years.
- Tech lead for Workflow Automation Platform which provides Workflow as a Service across Intuit.
- Passionate Full Stack Developer worked on adding new/revamping/scaling features in QBO. Backend developer, image processing in my previous stints.

Raghav Agarwal

- Software Engineer 2 at Intuit with a stint of 2+ years at Intuit.
- Workflow Automation Platform Team which provides Workflow as a Service across Intuit.
- Enthusiastic Problem solver and Backend Developer with multiple developments /scales on various QBO features and Workflow Platform.


INTUIT MISSION
Powering Prosperity
Around the World

Intuit Strategy

AI-Driven Expert Platform

Quickbooks Advanced

For mid-market customers

QuickBooks is an accounting software geared mainly toward small and medium-sized businesses and provides on-premises accounting applications as well as cloud-based versions that accept business payments, manage and pay bills, and payroll functions.

QuickBooks Online Advanced provides critical features powerful to automate business processes like Workflow Automation, Batch Transactions, Business Insights, etc.

Intuit’s Workflow Landscape - Recap

Workflow Use cases @ Intuit

Small Business Processes

Approvals

Reminders

Tasks

Reviews

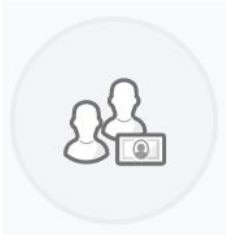
Auto Updates

Assisted Services by Intuit

Experts

Onboarding

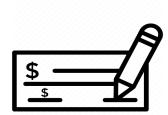
Accounting Automation



Bookkeeping

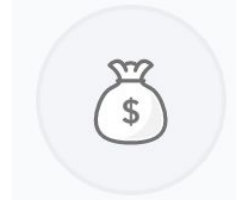
Internal Processes

Training



Automation

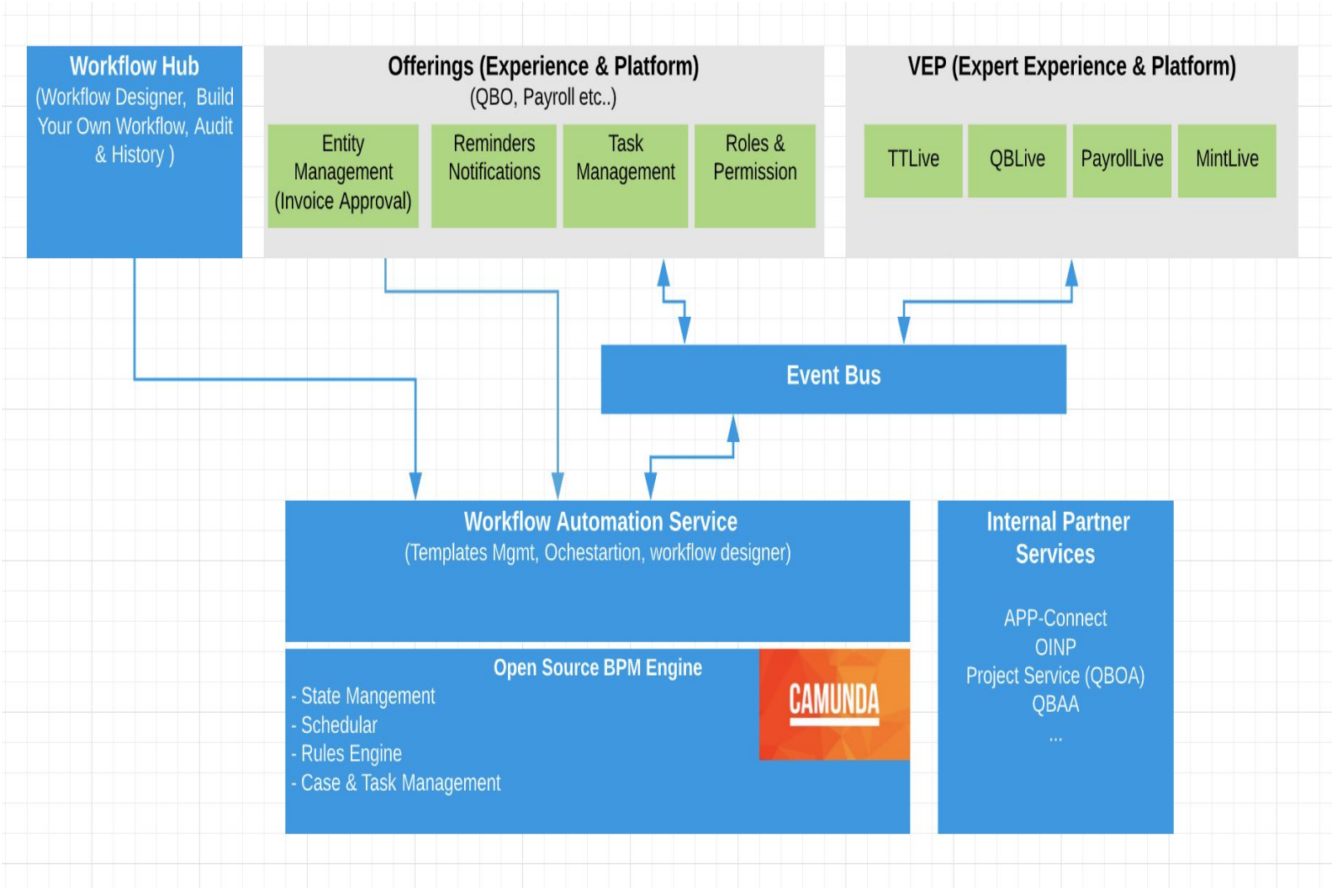
Hiring



Capital Processing

Intuit Workflow Automation Platform

- Workflow as a service for customers providing Abstraction, Configuration, Customization
- Async architecture, plug and play task adapters, schedule management
- Templates, recommendations, integrations with services, workflow experience



Workflows at Scale

Infrastructure

Different swimlanes

Multi-tenant architecture

Multi-Region HA/DR deployment

Non-blocking external task workers

Separate worker pools per workflow type

Testing

Karate based automation tests

Gatling tests for performance tests

Automated FMEA tests using Chaos Monkey

Production War Games/Load Testing

Capacity Planning

E2E/Perf on par with Prod with all the monitoring

Observability

Micrometer and Prometheus based metrics

Wavefront for dashboards and alerts

Real-time alerting using PagerDuty

Jaeger based traces using OpenTracing

AppDynamics for JVM monitoring

Maturing as we are Expanding

01 Explore

- Understand the business use cases.
- Identify Capabilities
- Evaluating technology options.
- POCs and solution validation.
- Decision sign off

02 Unit Of One

- Prove architecture through unit of one.
- Building the foundational platform (Core functional capabilities, Hosting, Opex, Monitoring, FMEA)
- Learning from the feedback and making amends

03 Expand

- Broaden the functional capabilities to support all use cases.
- Scaling the platform for higher traffic and more offerings.
- Onboard more use cases
- Well defined Availability, HA/DR, FMEA strategy in place.

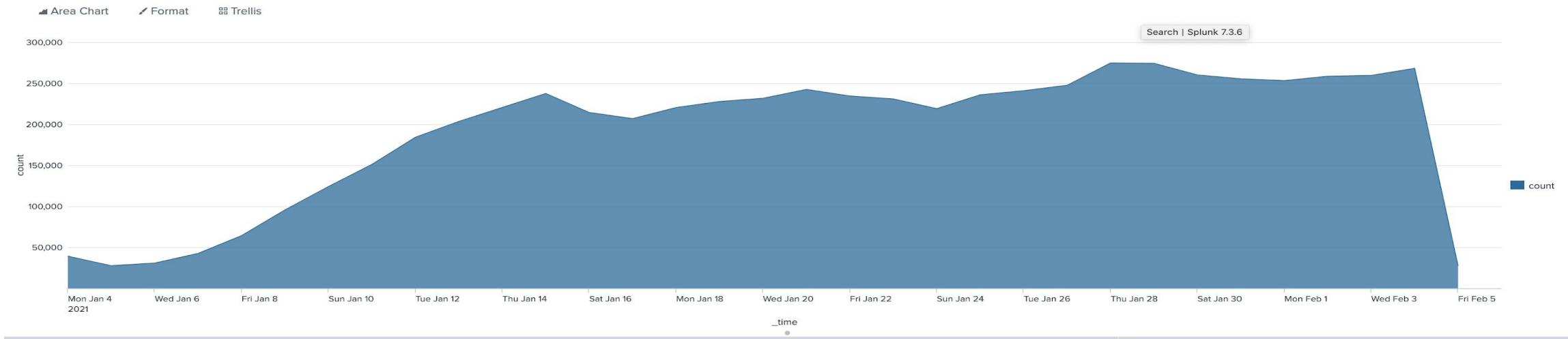
04 Mature

- Evolving functional capabilities to increase efficiency
- Easy & automated developer tooling for open contribution, enhanced self serve
- Mature process on opex, monitoring, alerts etc.
- Continuous improvements

Journey So Far

What is Peak?

Peak customer needs, business growth and the highest surge in traffic!



Accounting Peak

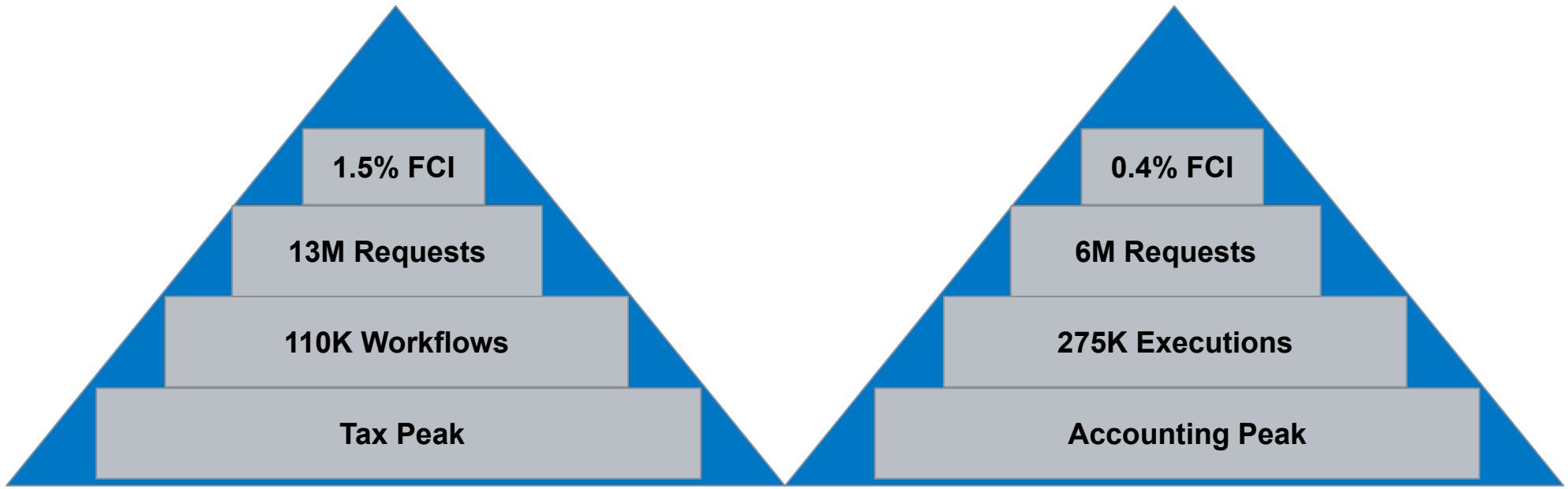
- Start of the year is a new start for accounting
- New subscriptions, more transactions
- 5 fold growth in Platform use

Tax Peak

- In line with tax filing (seasonal)
- 100% traffic on Workflows for assisted flows
- 4 fold growth in Platform use

Peak Statistics

100% Availability, Max TPS of 396



Challenges

The Internet was done so well that most people think of it as a natural resource like the Pacific Ocean, rather than something that was man-made. When was the last time a technology with a scale like that was so error-free?

- *Alan Kay*

Database Bottlenecks

Underperforming Queries

- Rise in Latency with Load
- Inefficient querying of BLOB storage (model)
- Cascading effects

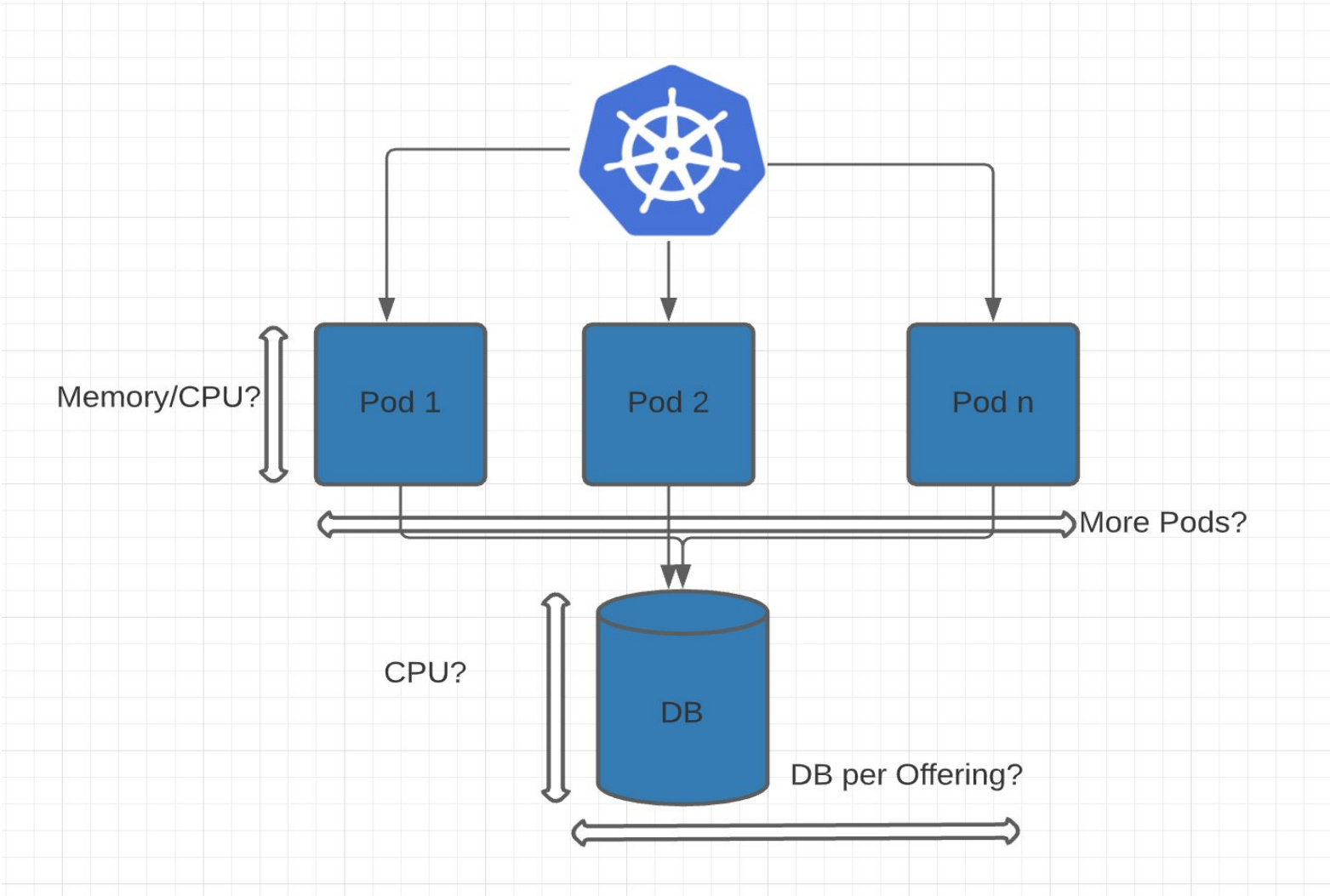
Exponential DB Growth

- Volume grew 10 folds in a month
- History tables grew to more than a TB
- Dead Tuple/Max TxId growth to near 200M triggering bookkeeping DB operations

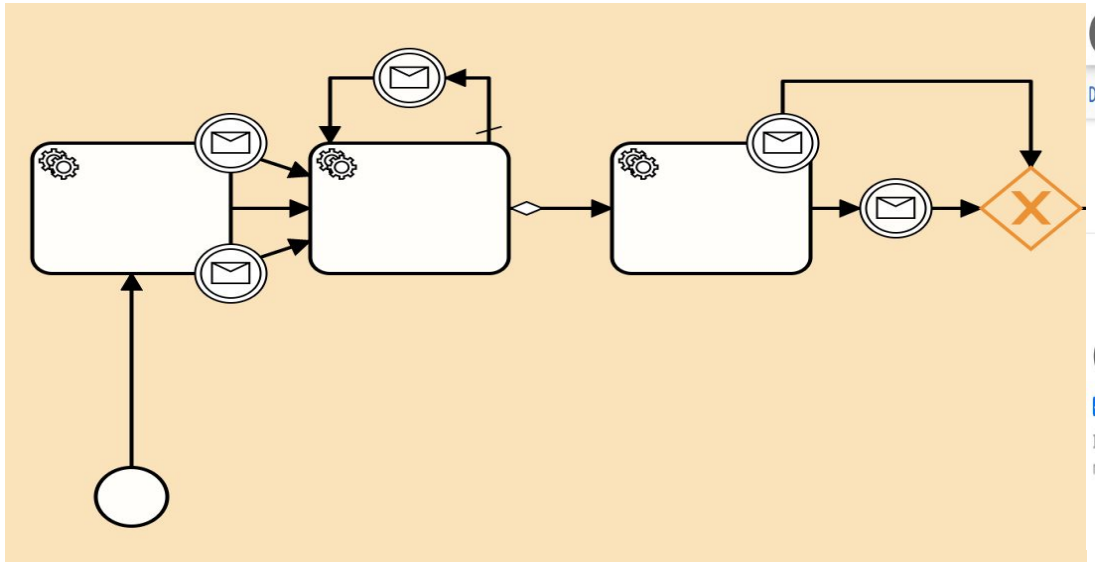
High IOPS/CPU

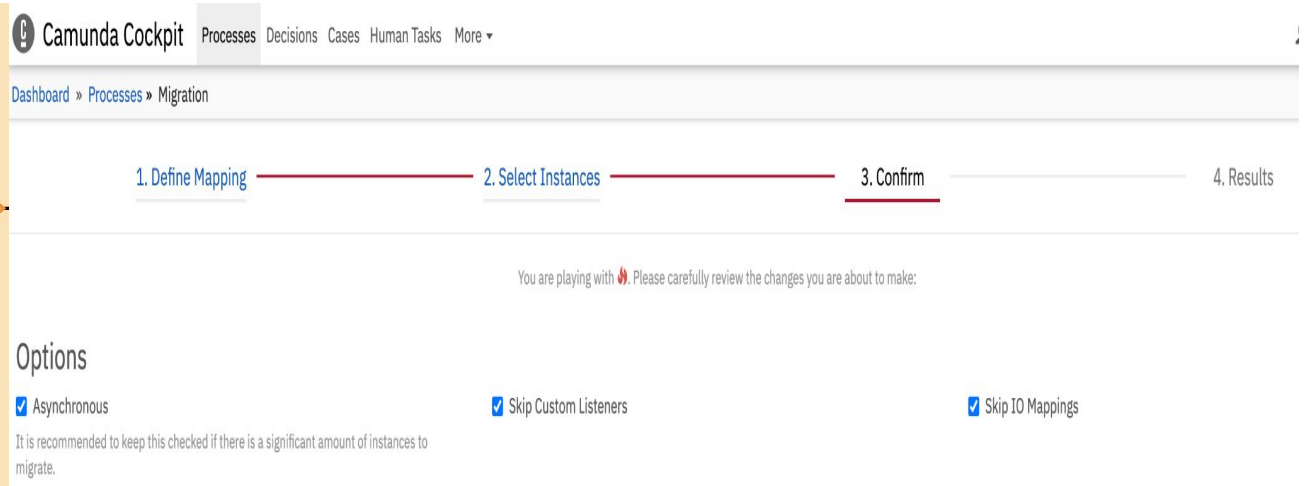
- As TPS increased IOPS increased
- CPU almost at max
- Cascading effects

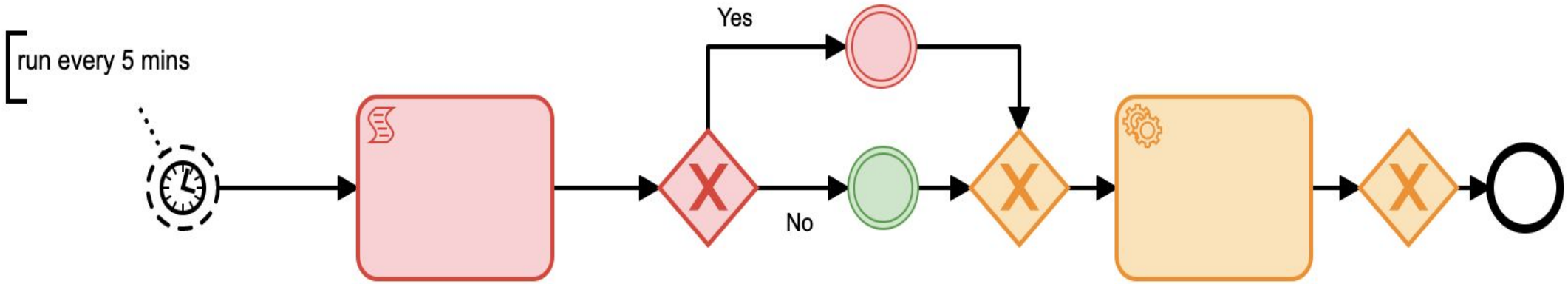
Horizontal vs Vertical Scale



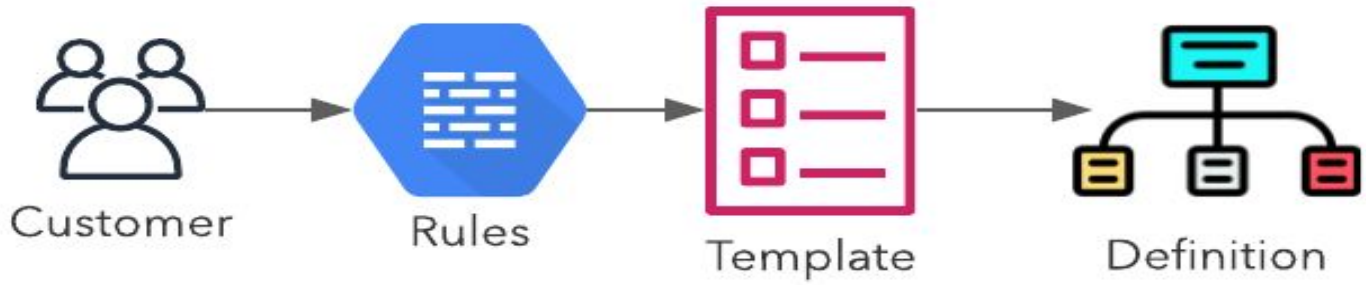
Model Challenges







Definition Challenges



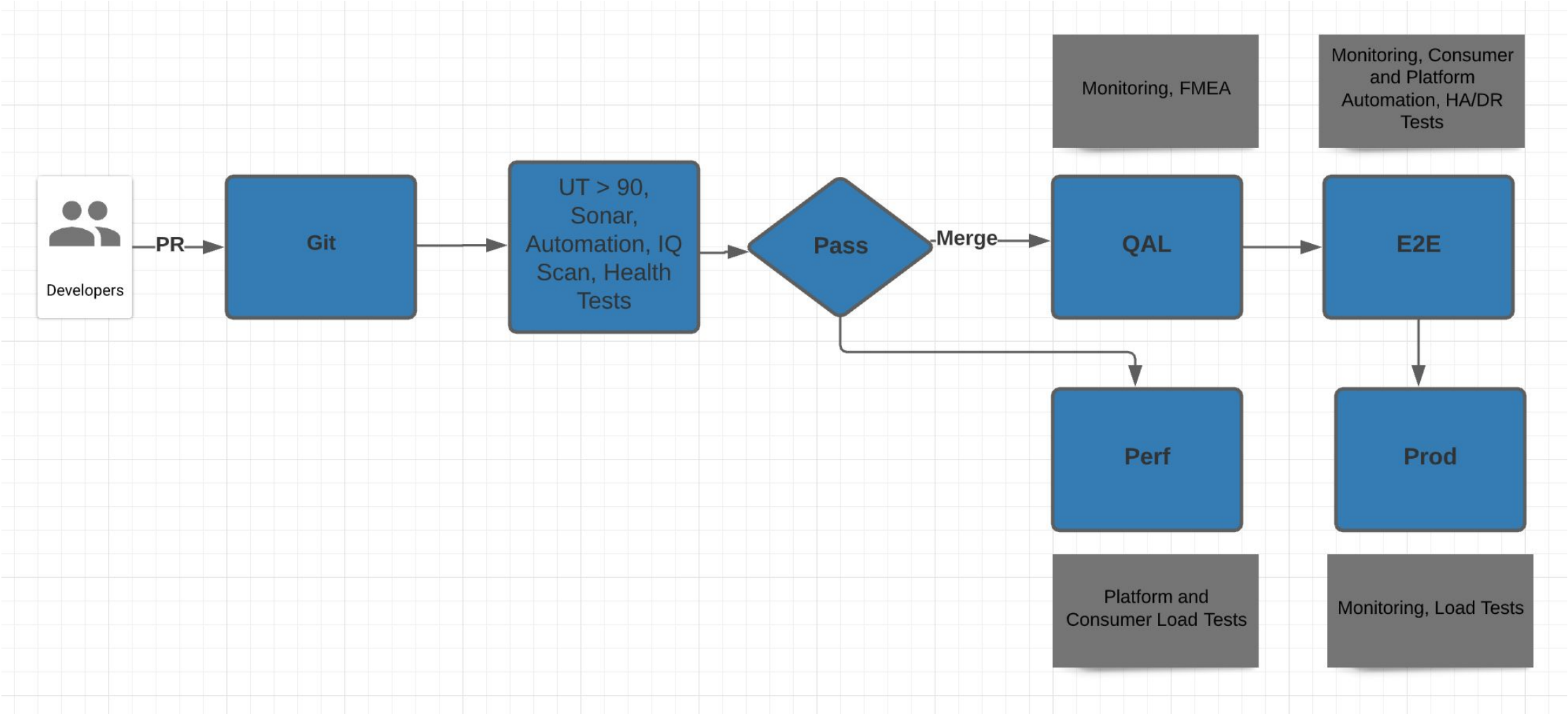
- Higher memory needs in systems
- Limited capabilities in Cockpit, Optimize
- Redundancy is maintainability nightmare

How we did it?

All code is guilty until proven innocent!

Testing and More Testing

Unit Tests, Automation, FMEA, Load Tests in all Environments

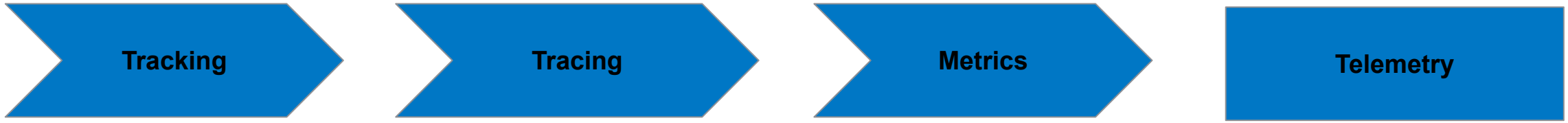


Right Capacity Plan

- Scaling policies upon memory threshold breach, queue buildup
- Right limits for CPU, memory on every pod/DB based on request capacity
- Keep headroom for restarts, termination
- Right threadpool for workers, event publisher/consumers and connection pool

High Performing Systems are Highly Scalable

Observability



- End to end tid tracing/workflow tracing
- Comprehensive metrics/logs dashboards for trends
- Alerts on anomalies reduced time to detect
- Dedicated on call

Resiliency

- Retries and timeouts (Resilience4j)
- Request throttling per consumer
- HA/DR Strategy, Reader DB, Dedicated cockpit
- Async executions and DLQs

Idempotency

- Idempotency keys in events
- Checks based on workflow entities
- Idempotent task adapters
- Process engine is idempotent

Learnings

What works at scale may be different from scaling what works

Lean DB/Clean DB

Revisiting Queries

- Optimizing queries with right index
- Optimizing fetch with custom queries to exclude model data
- Query Insights

Controlling DB Growth

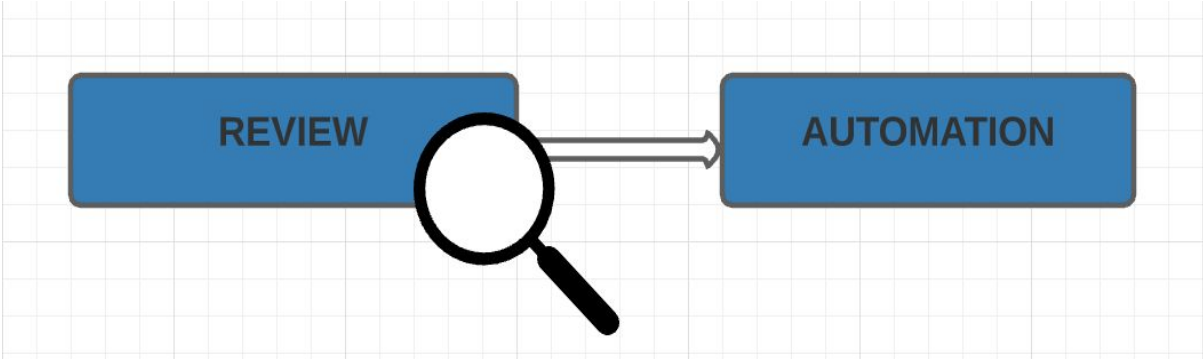
- Mandatory low history TTL. Optimize as Backup.
- Custom History handlers
- Sequenced history cleanups
- DB alerts, controlled bookkeeping operations

Capacity Planning

- Change capacity leaving headroom
- Failover downtime reduction/playbooks
- DB alerts across environments

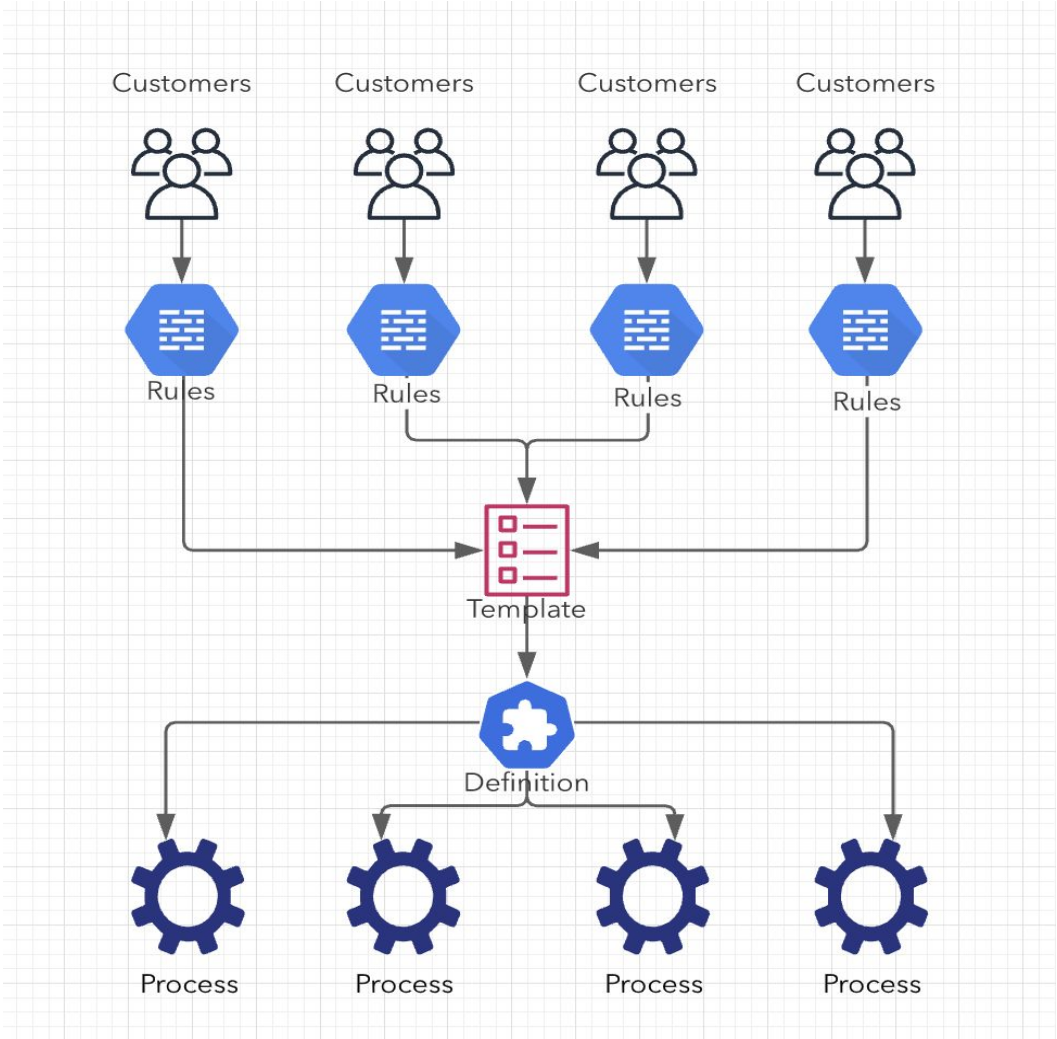
Learnt the Hard Way to Model

- Race conditions thwart system resources
- Detect anomalies in model especially concurrent executions
- Workarounds for timer
- Early signals around task/job pile up



Single Definition - Less is More

- Uber definition with placeholders for user configurations
- One definition (BPMN, DMN) for a workflow type
- Provides maintainability, flexibility and scalability
- Reduced memory footprint, faster user experience



Multi Tenancy/Self Serve

- Provides isolation and ensures fair share of resources including Process Engine
- Insulation against Noisy Neighbors
- Easy debugging for offerings
- Cockpit access segregation for offerings

Q&A

#slido-name