_ Midpoint 2016.4.11

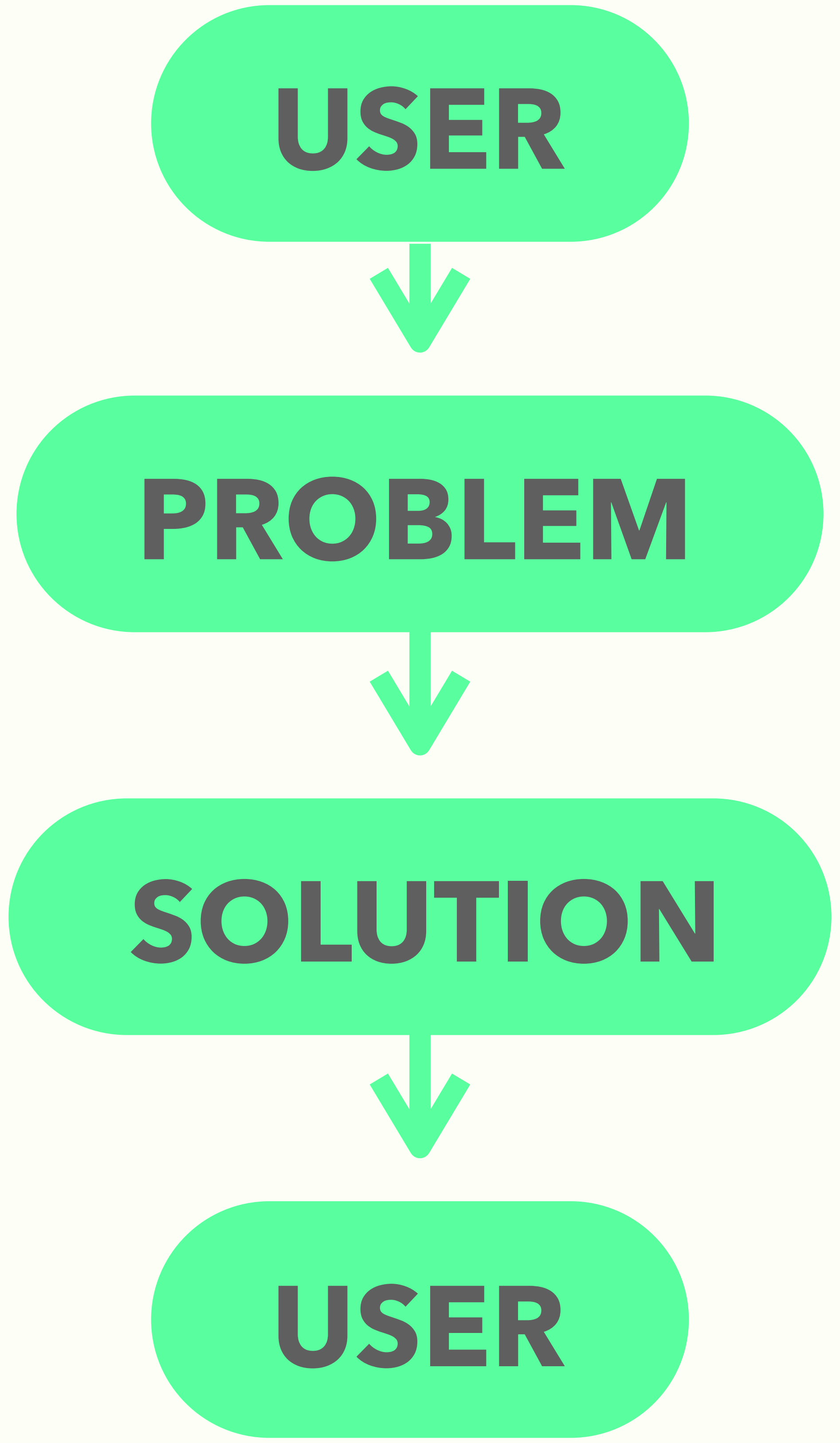
alphadoop

Doflamingo

An light-weight monitoring system for Apache Hadoop

TITLE	Kafka/ Zookeeper Monitoring Module built for Flamingo Ecosystem		
DURATION	March 13, 2016 ~ June 8, 2016		
CLIENT	EXEM	PRESENTER	ALPHADOOP

_ Midpoint



USER

PROBLEM

SOLUTION

USER

CONTENTS

OVERVIEW

USERS

PROBLEMS

SOLUTIONS

NOVELTY

SCENARIO

SCHEDULE

Overview
Users
Problems
Solutions
Novelty
Scenario
Schedule

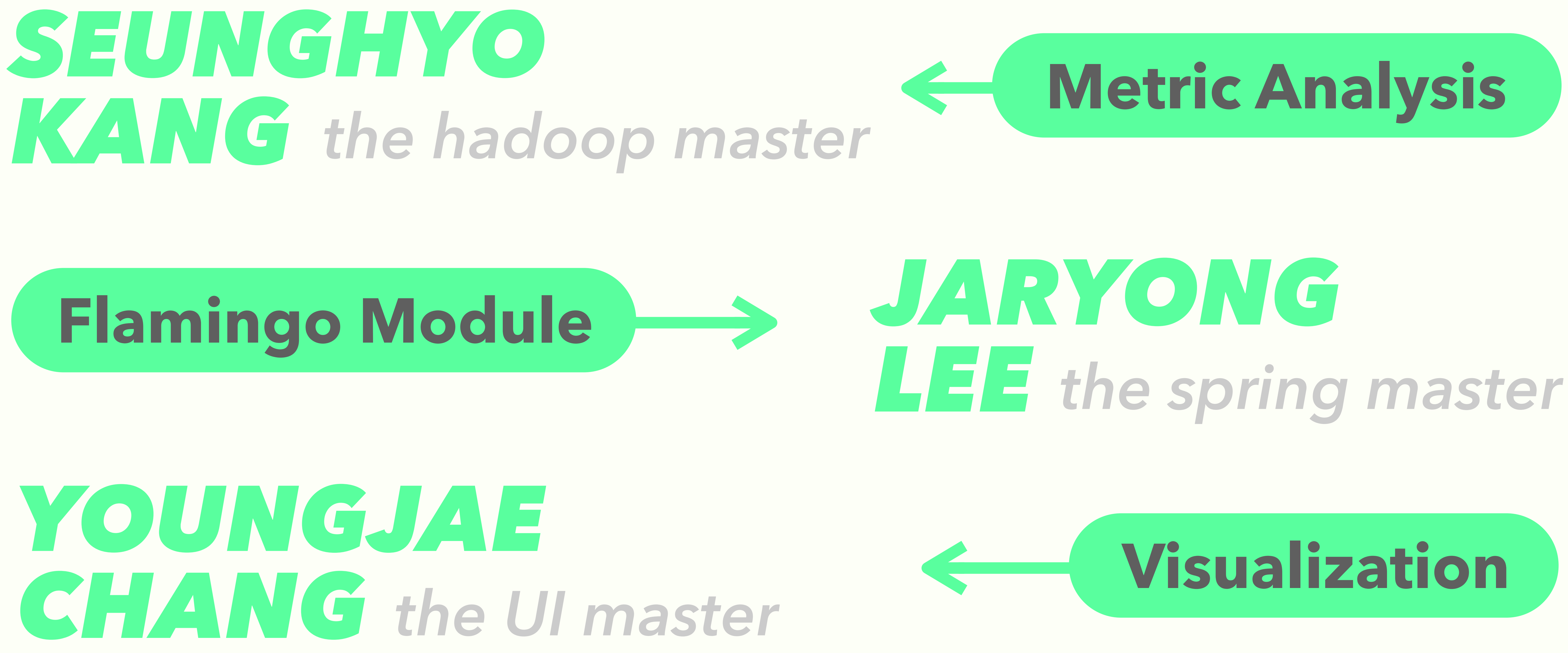
_ PROJECT GOAL

**Collect Performance Metrics,
Visualize it, and
Integrate it with Flamingo.**

- **Overview**
- Users
- Problems
- Solutions
- Novelty
- Scenario
- Schedule

_ TEAM

TEAM _ ALPHADOOP



SEUNGHYO KANG *the hadoop master* ← **Metric Analysis**

Flamingo Module → ***JARYONG LEE*** *the spring master*

YOUNGJAE CHANG *the UI master* ← **Visualization**

Overview
Users
Problems
Solutions
Novelty
Scenario
Schedule

_ REQUIREMENTS

1. Flamingo Ecosystem
2. Real-time Monitoring
3. JAVA, Spring, Sencha
4. Kafka, Zookeeper

_ SCOPE

Only Monitoring, No Management

We know nothing:

Concentrate on given metrics

Overview

Users

Problems

Solutions

Novelty

Scenario

Schedule

_ QUESTION

PHASE #1 **What is a monitoring?**

_ QUESTION

PHASE #1 **What is a monitoring?**

PHASE #2 **Why do we monitor?**

Overview

Users

Problems

Solutions

Novelty

Scenario

Schedule

_ TWO NEEDS

*To ensure the **normal operation** of the system*

*To find out the cause of **abnormal behavior***

Overview

Users

Problems

Solutions

Novelty

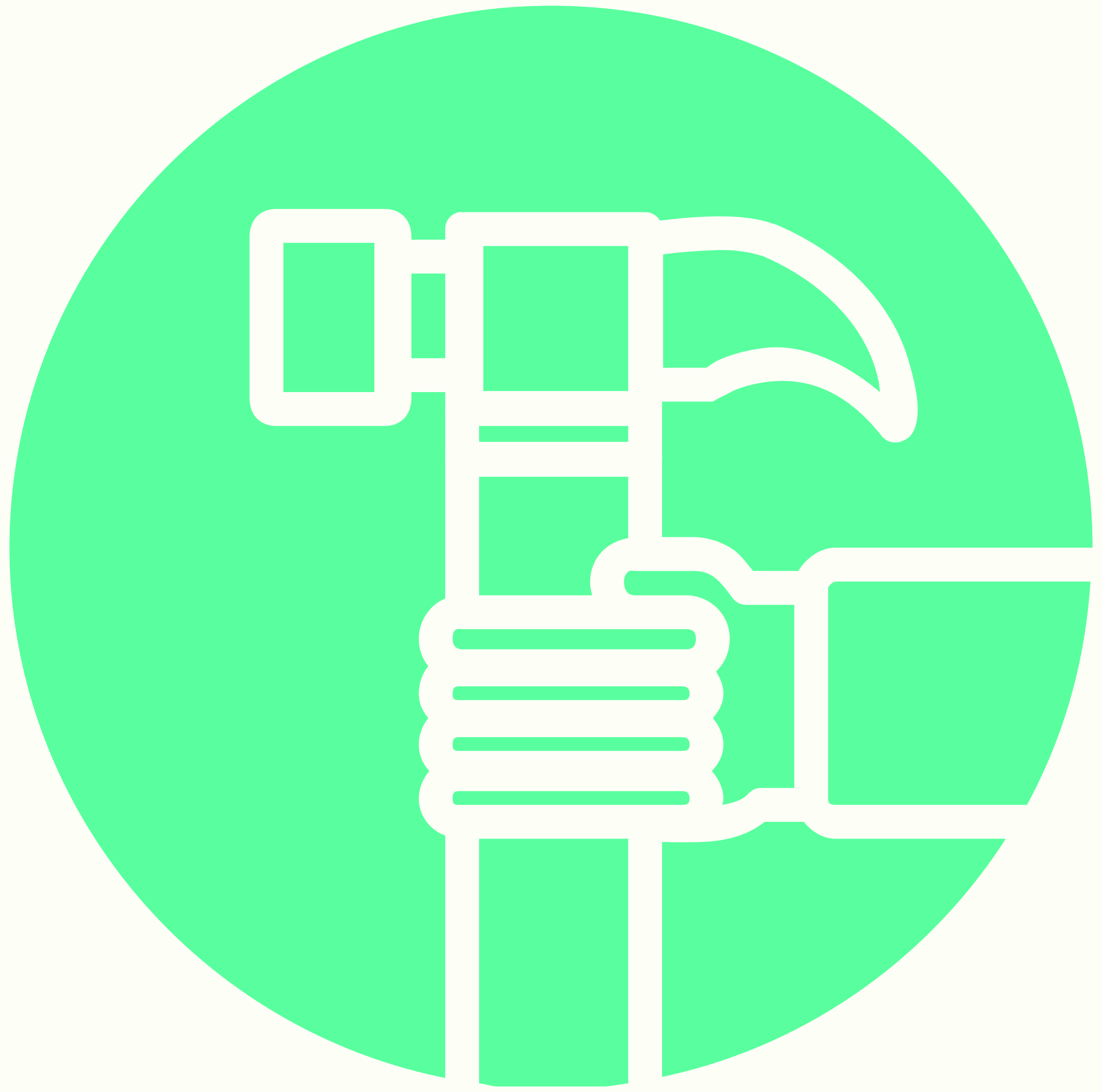
Scenario

Schedule

_ TWO USERS

USER #1

Maintainer



USER #2

Developer

Overview

Users

Problems

Solutions

Novelty

Scenario

Schedule

_ USERS

USER #1

Maintainer

A. Hope everything stays normal

B. Determine whether to put more resources or not

C. Usually maintains a volume of system

D. Focus on real-time data

USER #2

Developer

A. Fix the problem

B. Find out the cause of the problem by traveling the past data

C. Deeper understanding on whole system

D. Focus on specific events

Overview

Users

Problems

Solutions

Novelty

Scenario

Schedule

_ DIFFERENT REQUIREMENTS

USER #1

Maintainer

A. Visualize constantly changing statistics of sys.
B. At a glance view of metrics
C. Real-time update without user intervention

USER #2

Developer

A. Visualize abrupt events
B. Can travel back to the past to find the cause of event
C. Detailed analysis on changing variables during specific timeframe

Overview

Users

Problems

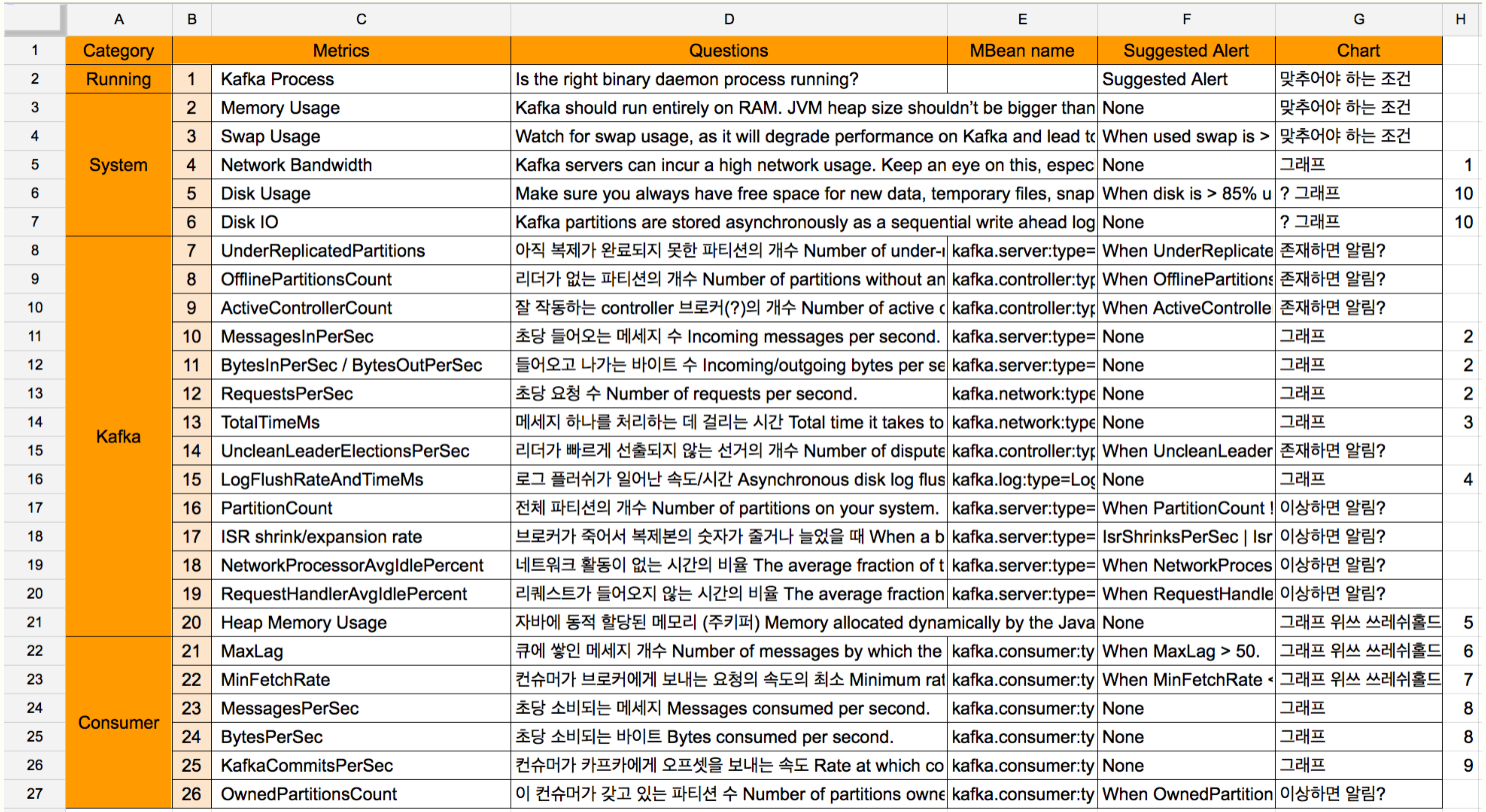
Solutions

Novelty

Scenario

Schedule

_ Metric Decision



	A	B	C	D	E	F	G	H
1	Category	Metrics		Questions	MBean name	Suggested Alert	Chart	
2	Running	1	Kafka Process	Is the right binary daemon process running?		Suggested Alert	맞추어야 하는 조건	
3	System	2	Memory Usage	Kafka should run entirely on RAM. JVM heap size shouldn't be bigger than		None	맞추어야 하는 조건	
4		3	Swap Usage	Watch for swap usage, as it will degrade performance on Kafka and lead t		When used swap is >	맞추어야 하는 조건	
5		4	Network Bandwidth	Kafka servers can incur a high network usage. Keep an eye on this, espec		None	그래프	1
6		5	Disk Usage	Make sure you always have free space for new data, temporary files, snap		When disk is > 85% u	? 그래프	10
7		6	Disk IO	Kafka partitions are stored asynchronously as a sequential write ahead log		None	? 그래프	10
8	Kafka	7	UnderReplicatedPartitions	아직 복제가 완료되지 못한 파티션의 개수 Number of under-	kafka.server:type=	When UnderReplicate	존재하면 알림?	
9		8	OfflinePartitionsCount	리더가 없는 파티션의 개수 Number of partitions without an	kafka.controller:typ	When OfflinePartitions	존재하면 알림?	
10		9	ActiveControllerCount	잘 작동하는 controller 브로커(?)의 개수 Number of active c	kafka.controller:typ	When ActiveControlle	존재하면 알림?	
11		10	MessagesInPerSec	초당 들어오는 메세지 수 Incoming messages per second.	kafka.server:type=	None	그래프	2
12		11	BytesInPerSec / BytesOutPerSec	들어오고 나가는 바이트 수 Incoming/outgoing bytes per se	kafka.server:type=	None	그래프	2
13		12	RequestsPerSec	초당 요청 수 Number of requests per second.	kafka.network:type	None	그래프	2
14		13	TotalTimeMs	메세지 하나를 처리하는 데 걸리는 시간 Total time it takes to	kafka.network:type	None	그래프	3
15		14	UncleanLeaderElectionsPerSec	리더가 빠르게 선출되지 않는 선거의 개수 Number of dispute	kafka.controller:typ	When UncleanLeader	존재하면 알림?	
16		15	LogFlushRateAndTimeMs	로그 플러쉬가 일어난 속도/시간 Asynchronous disk log flus	kafka.log:type=Log	None	그래프	4
17		16	PartitionCount	전체 파티션의 개수 Number of partitions on your system.	kafka.server:type=	When PartitionCount !	이상하면 알림?	
18		17	ISR shrink/expansion rate	브로커가 죽어서 복제본의 숫자가 줄거나 늘었을 때 When a b	kafka.server:type=	IsrShrinksPerSec | Isr	이상하면 알림?	
19		18	NetworkProcessorAvgIdlePercent	네트워크 활동이 없는 시간의 비율 The average fraction of t	kafka.server:type=	When NetworkProces	이상하면 알림?	
20		19	RequestHandlerAvgIdlePercent	리퀘스트가 들어오지 않는 시간의 비율 The average fraction	kafka.server:type=	When RequestHandle	이상하면 알림?	
21		20	Heap Memory Usage	자바에 동적 할당된 메모리 (주키퍼) Memory allocated dynamically by the Java		None	그래프 위쓰 쓰레쉬홀드	5
22	Consumer	21	MaxLag	큐에 쌓인 메세지 개수 Number of messages by which the	kafka.consumer:ty	When MaxLag > 50.	그래프 위쓰 쓰레쉬홀드	6
23		22	MinFetchRate	컨슈머가 브로커에게 보내는 요청의 속도의 최소 Minimum rat	kafka.consumer:ty	When MinFetchRate <	그래프 위쓰 쓰레쉬홀드	7
24		23	MessagesPerSec	초당 소비되는 메세지 Messages consumed per second.	kafka.consumer:ty	None	그래프	8
25		24	BytesPerSec	초당 소비되는 바이트 Bytes consumed per second.	kafka.consumer:ty	None	그래프	8
26		25	KafkaCommitsPerSec	컨슈머가 카프카에게 오프셋을 보내는 속도 Rate at which co	kafka.consumer:ty	None	그래프	9
27		26	OwnedPartitionsCount	이 컨슈머가 갖고 있는 파티션 수 Number of partitions owne	kafka.consumer:ty	When OwnedPartition	이상하면 알림?	

_ EXTERNAL INTERFACE

FUNC #1

Overview

A. Dashboard

B. ~~Configuration~~

FUNC #2

Timeline

A. Event Timeline

B. Timemachine

_ Midpoint

Overview

Users

Problems

Solutions

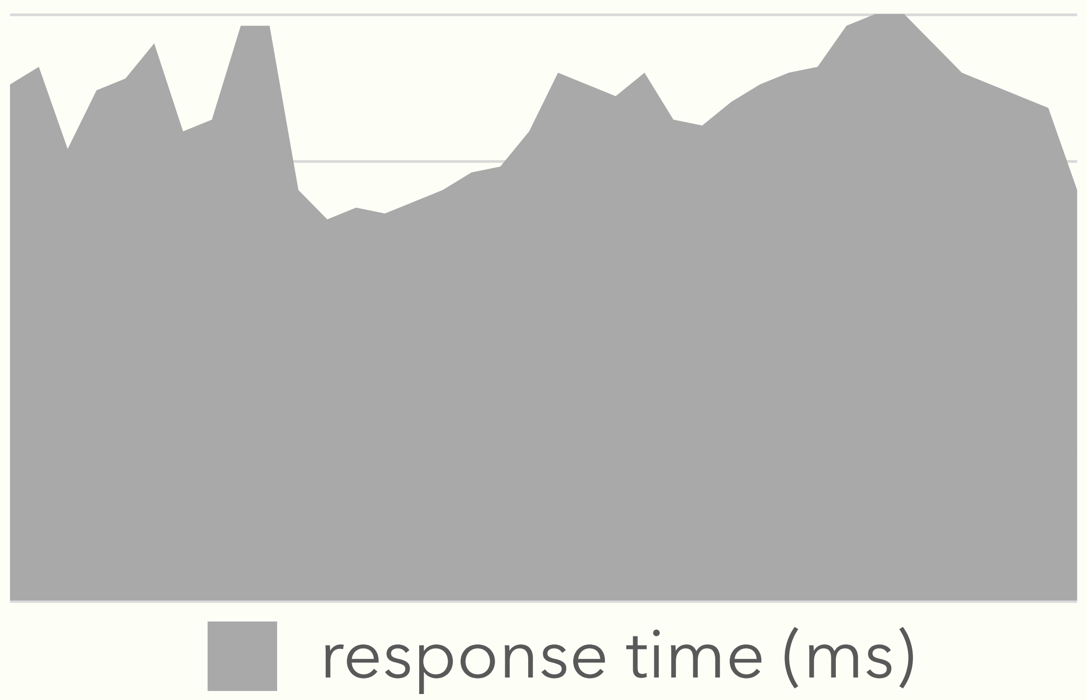
Novelty

Scenario

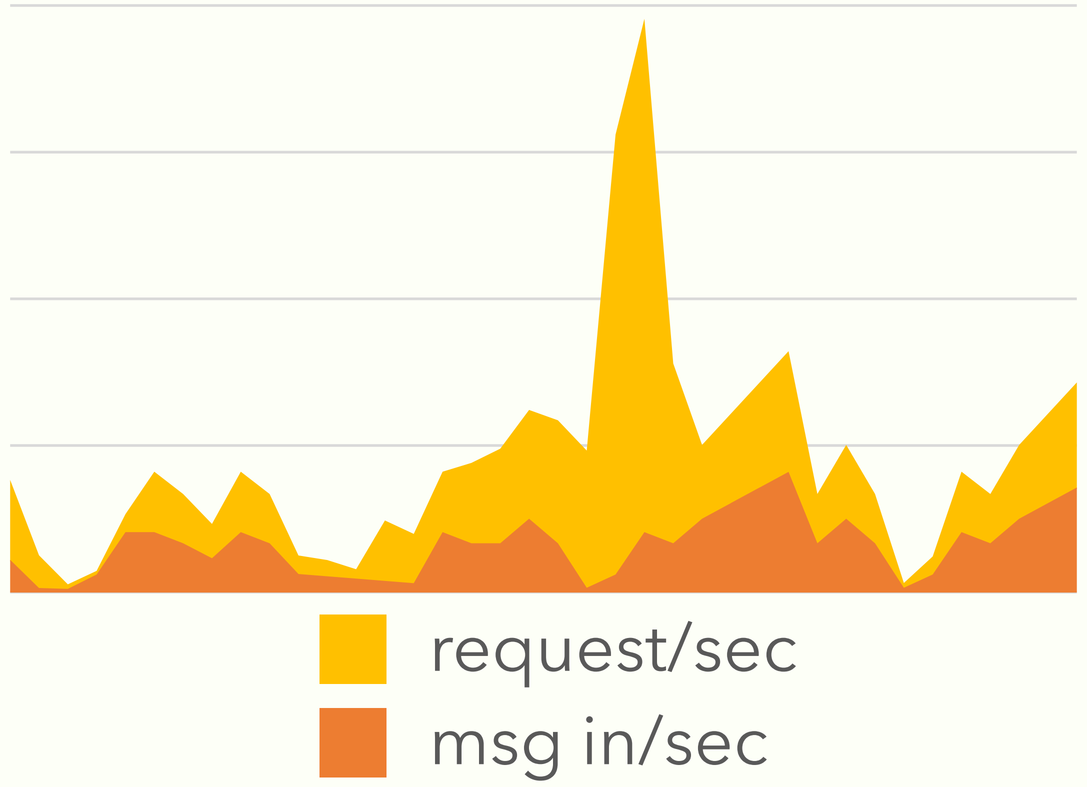
Schedule

FUNC #1 Overview

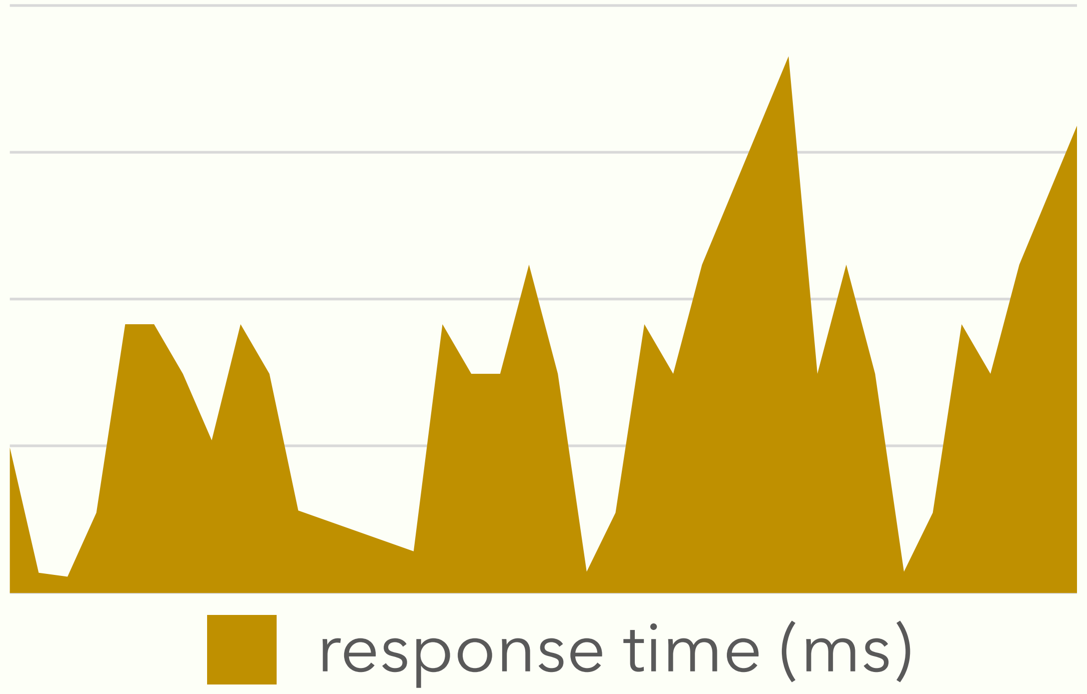
Heap memory usage



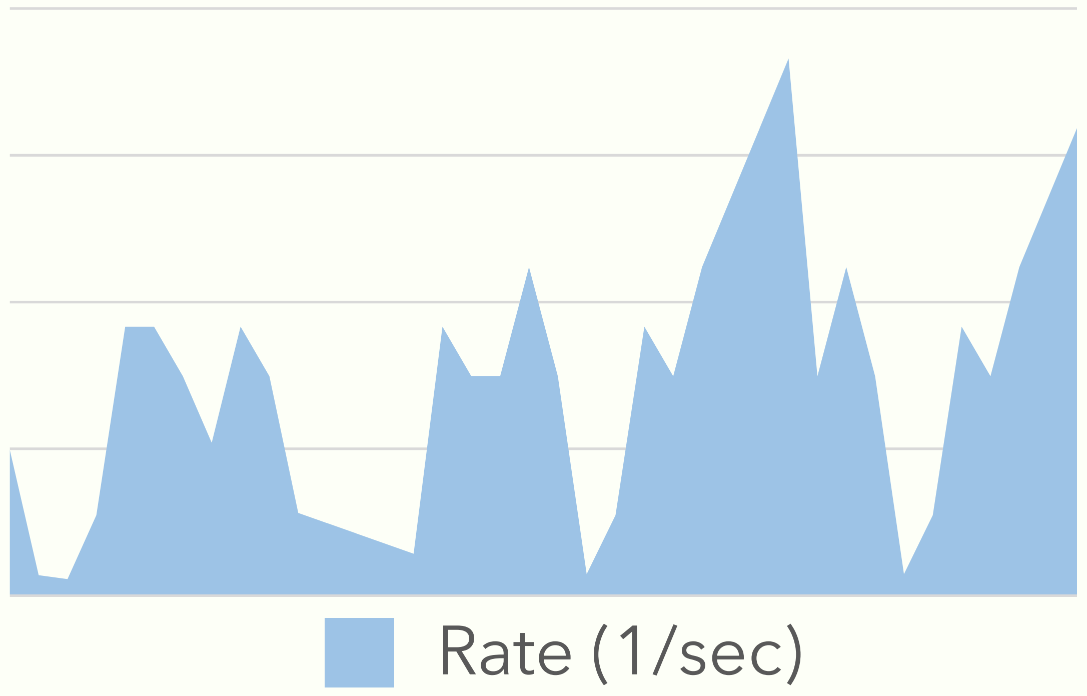
Message Condition



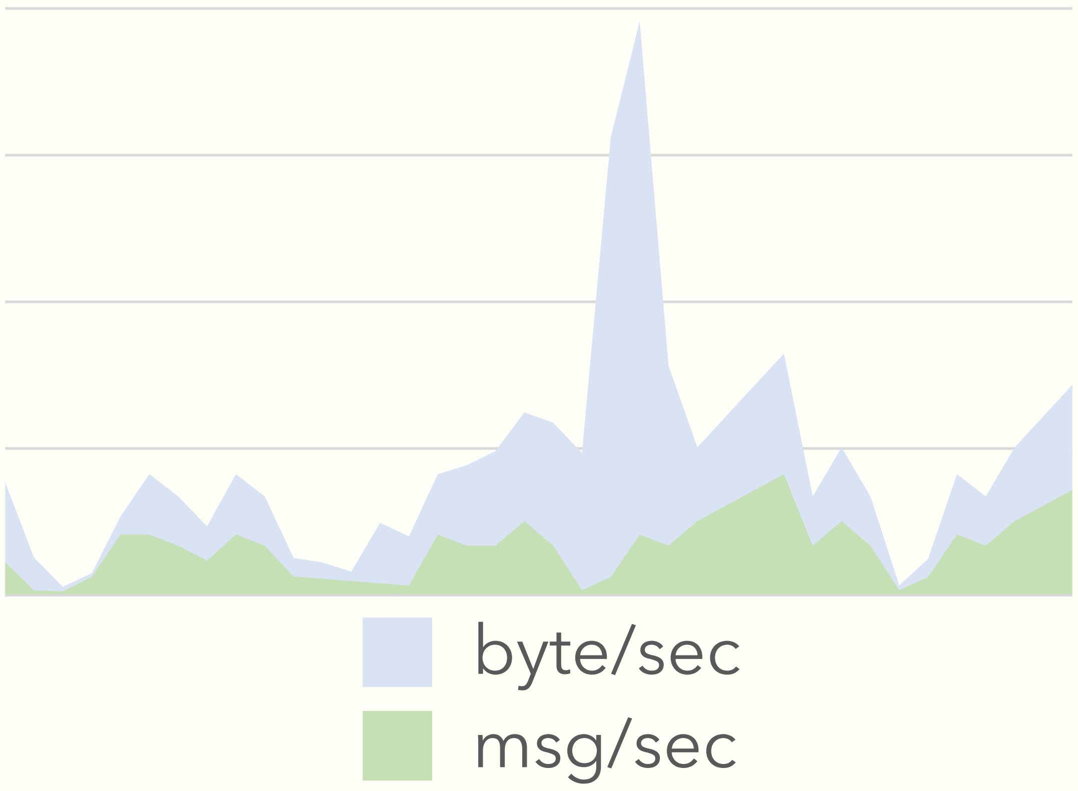
Response time



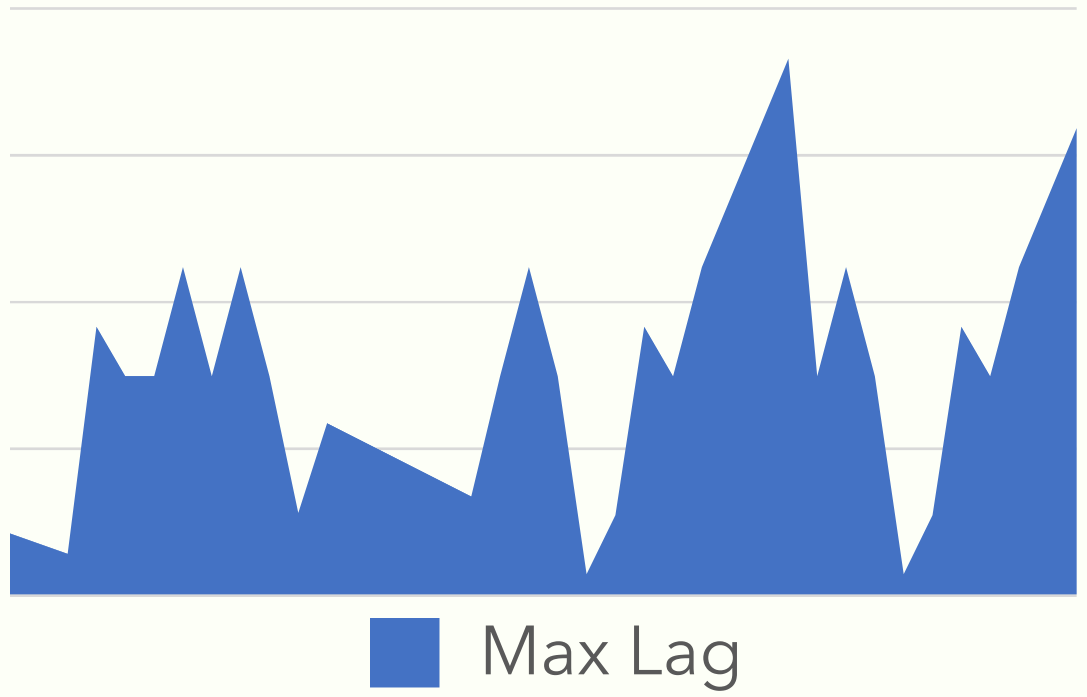
Minimum Fetch rate



Message Consumed



Max Lag



Overview

Users

Problems

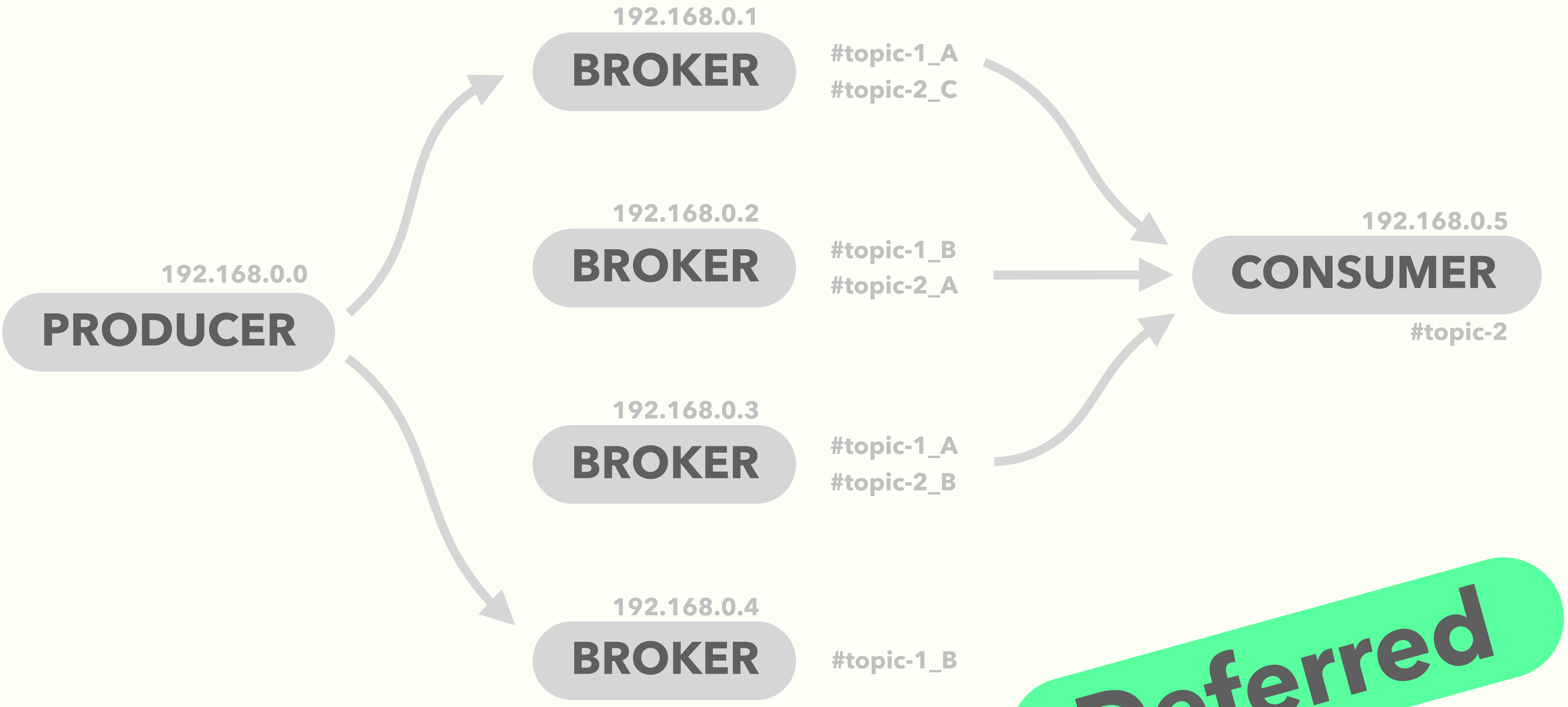
Solutions

Novelty

Scenario

Schedule

FUNC #1 Overview



Deferred

_ Midpoint

Overview

Users

Problems

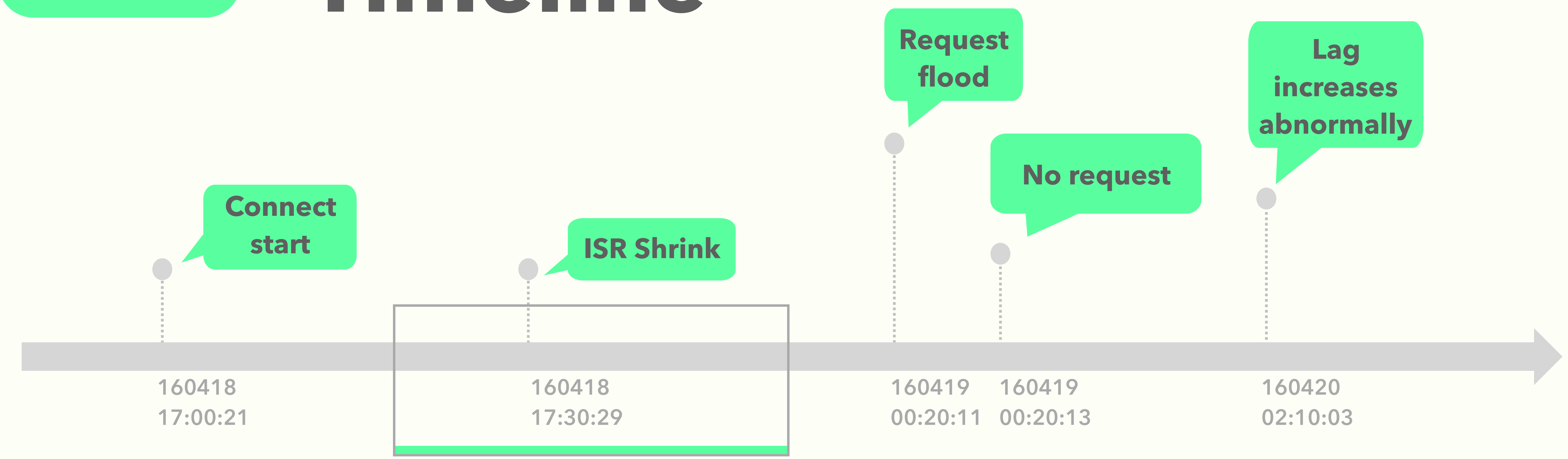
Solutions

Novelty

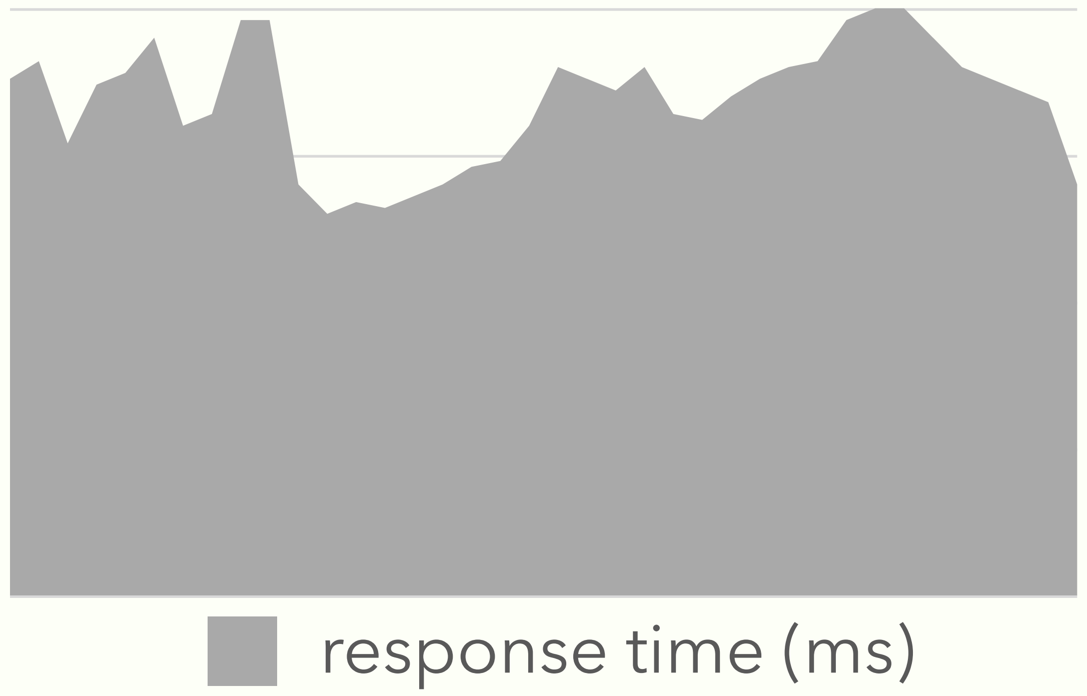
Scenario

Schedule

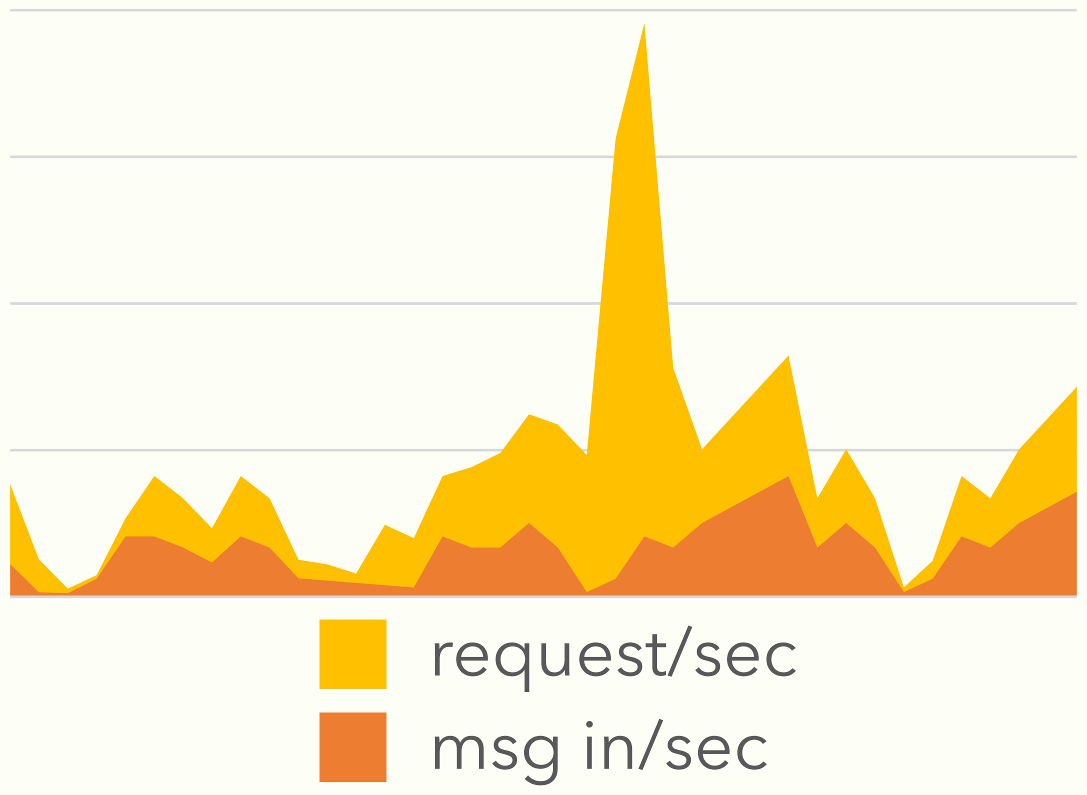
FUNC #2 Timeline



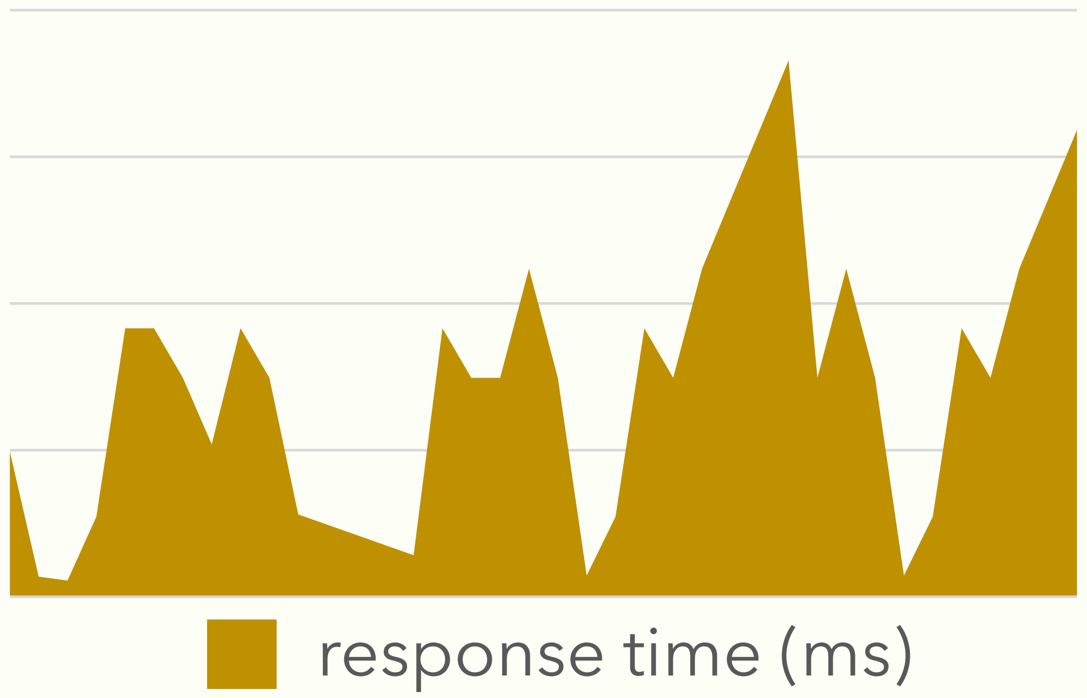
Heap memory usage



Message Condition



Response time



_ Midpoint

Overview

Users

Problems

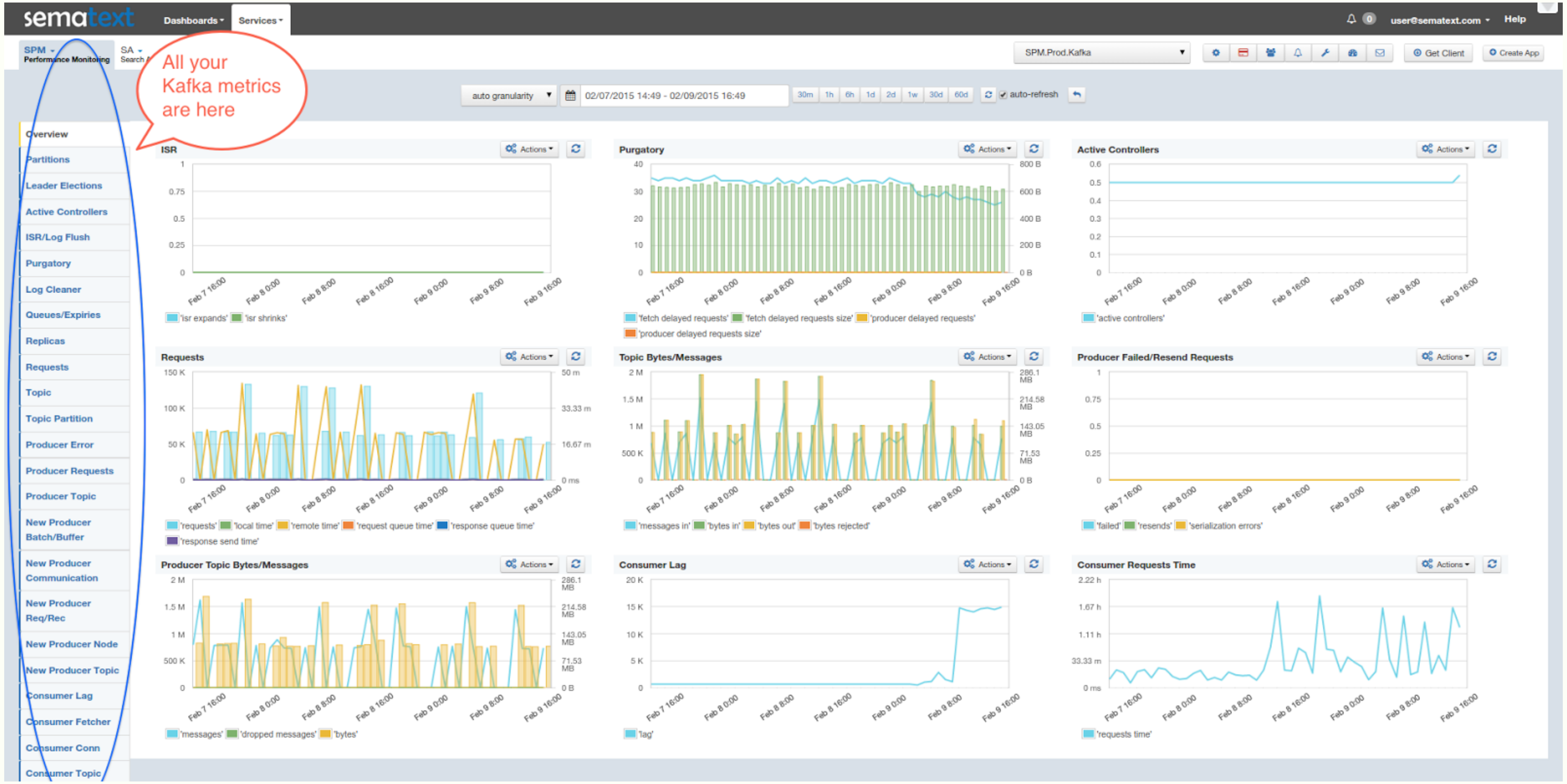
Solutions

Novelty

Scenario

Schedule

_ WHAT'S NEW?



All your Kafka metrics are here

Overview

Users

Problems

Solutions

Novelty

Scenario

Schedule

_ WHAT'S NEW?

Clear division of monitoring task

Further implication to BM

Overview

Users

Problems

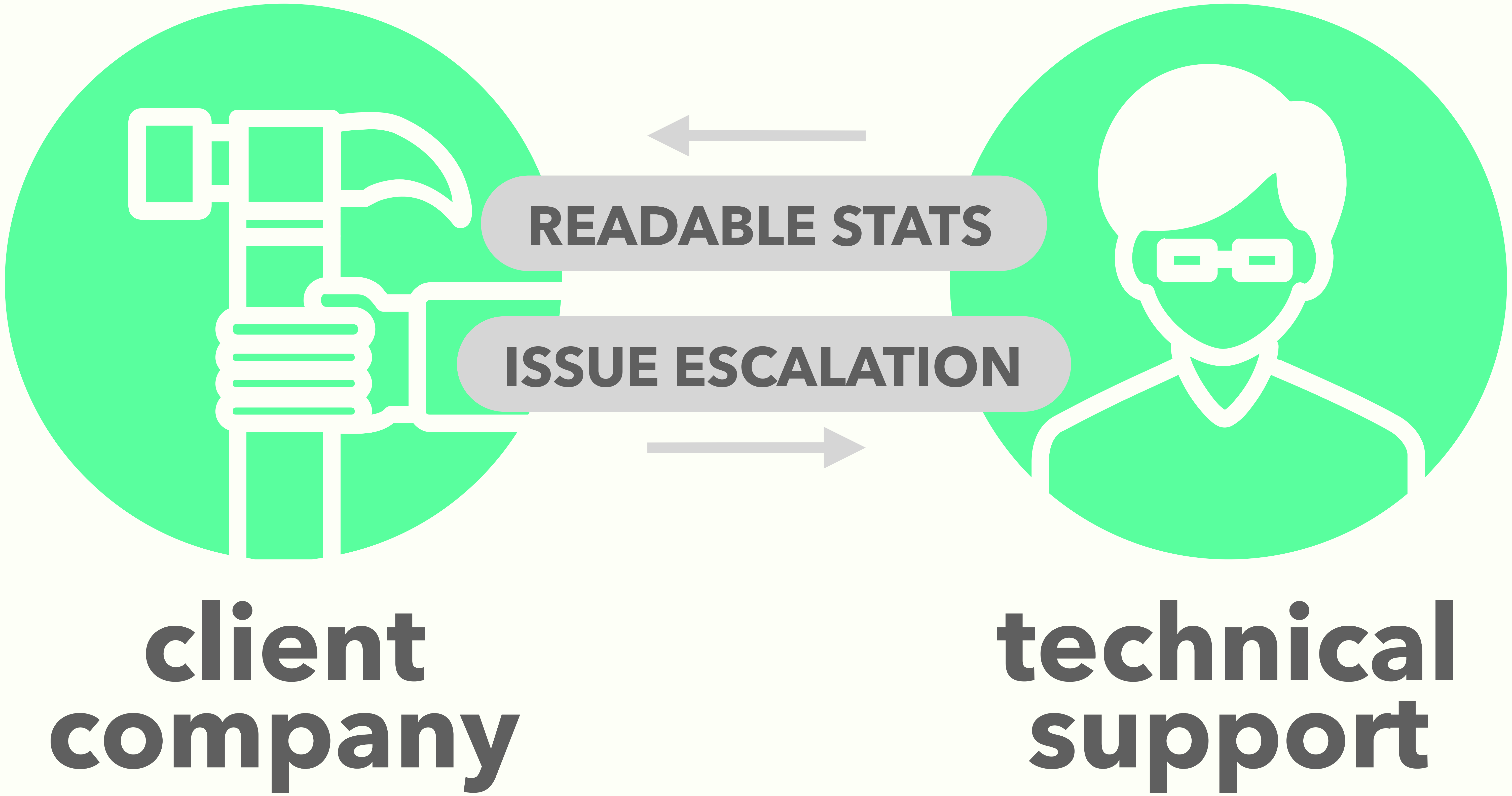
Solutions

Novelty

Scenario

Schedule

_ USER SCENARIO



READABLE STATS

ISSUE ESCALATION

client company

technical support

Overview

Users

Problems

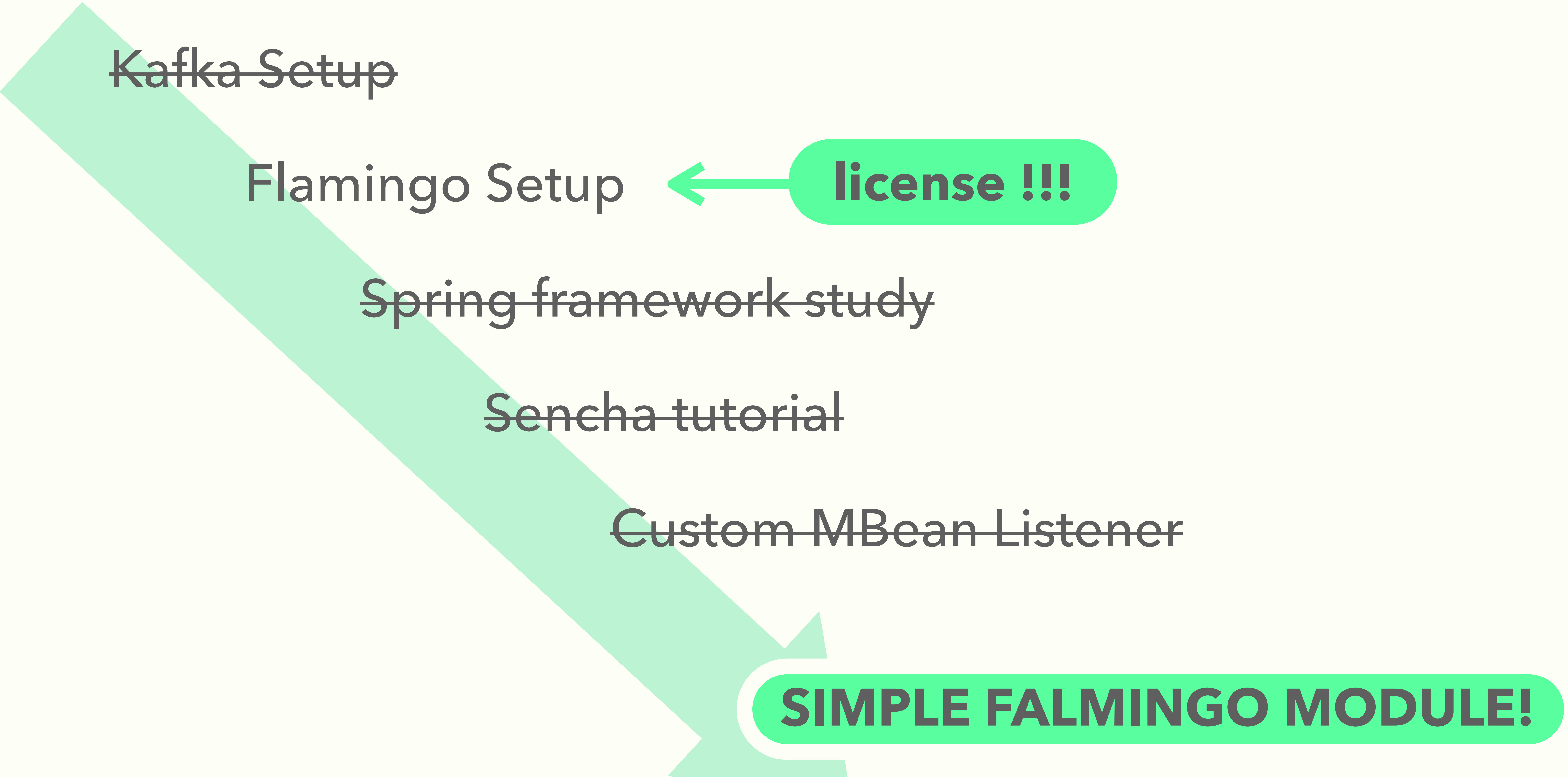
Solutions

Novelty

Scenario

Schedule

_ ON THE WAY



~~Kafka Setup~~

Flamingo Setup ← **license !!!**

~~Spring framework study~~

~~Sencha tutorial~~

~~Custom MBean Listener~~

SIMPLE FALMINGO MODULE!

END