



March 21-25, 2022
San Francisco, CA

Real-time anomaly monitoring in multiplayer online games

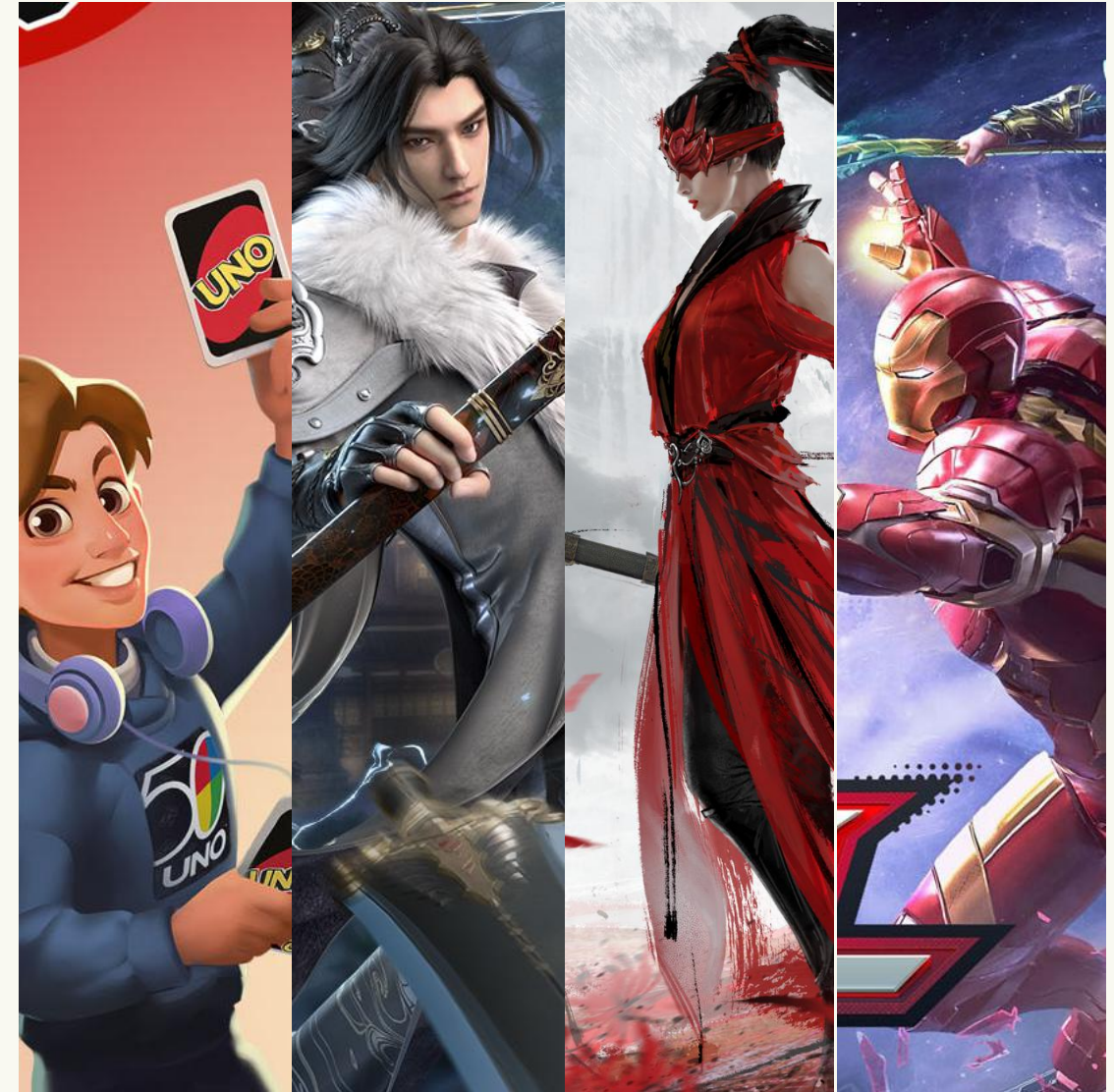
Lei Xia
ThunderFire UX
Senior Data Mining Engineer

#GDC22



ThunderFire UX

- ThunderFire UX User Experience Center (ThunderFire UX) aims at becoming the first-class user experience insight department in the video game industry. We are committed to deepening player understanding and getting actionable findings from our players to help various games of NetEase and ThunderFire achieve the best user experience.



Outline

- Background
- Practice in ThunderFire UX
- Takeaway

Part.1

Background

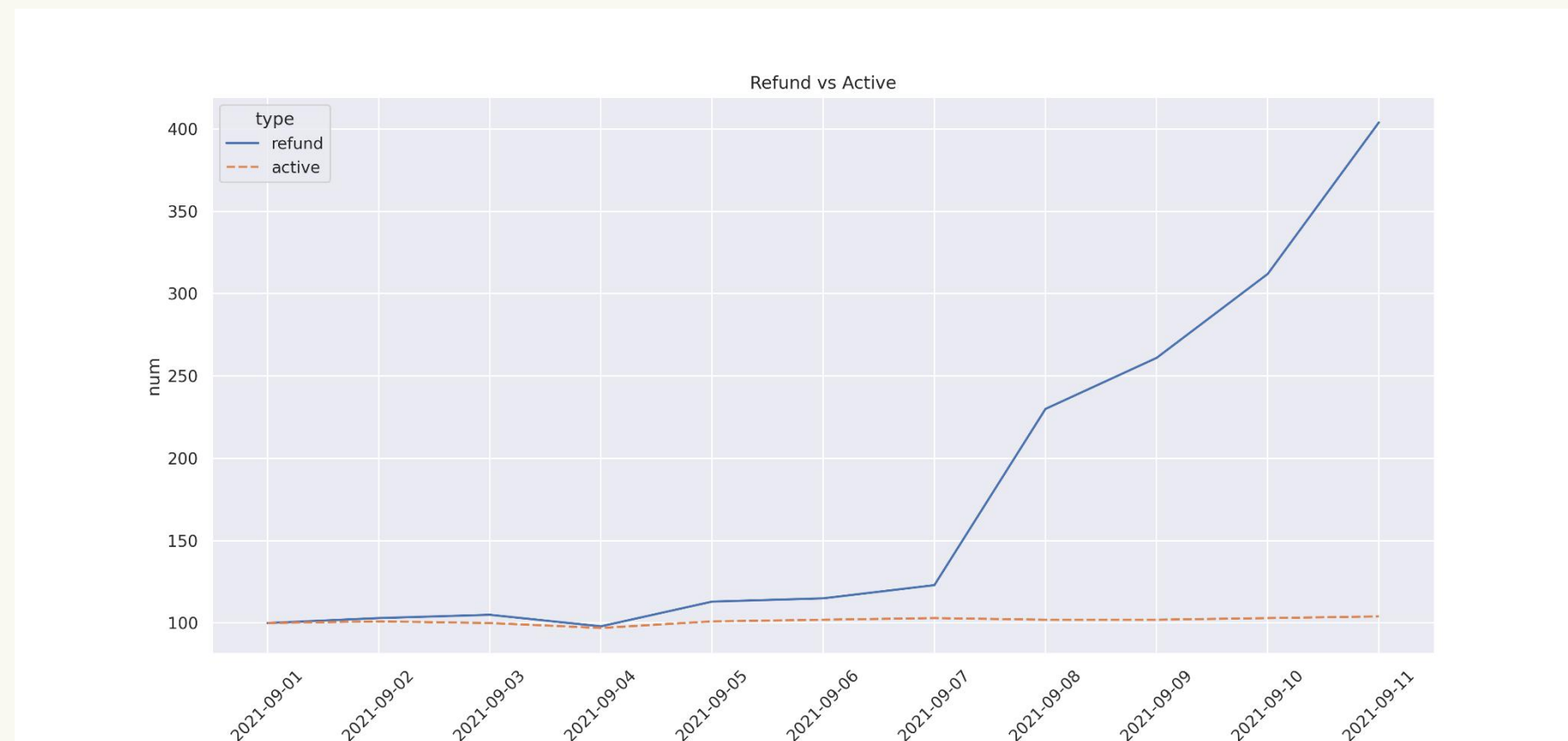
The Steam Refund Story

- In September, our game “NARAKA: BLADEPOINT” saw a significant increase in the number of in-game refunds (The values in the chart are **fake data**, just for illustration).



The Steam Refund Story

- By comparing the trends of active and refund number, we can determine that there are abnormal events in our game, possibly game cheating.



The Steam Refund Story

- When we found this problem, it has been some time since malicious refund. Is there any way to troubleshoot the problem in time?

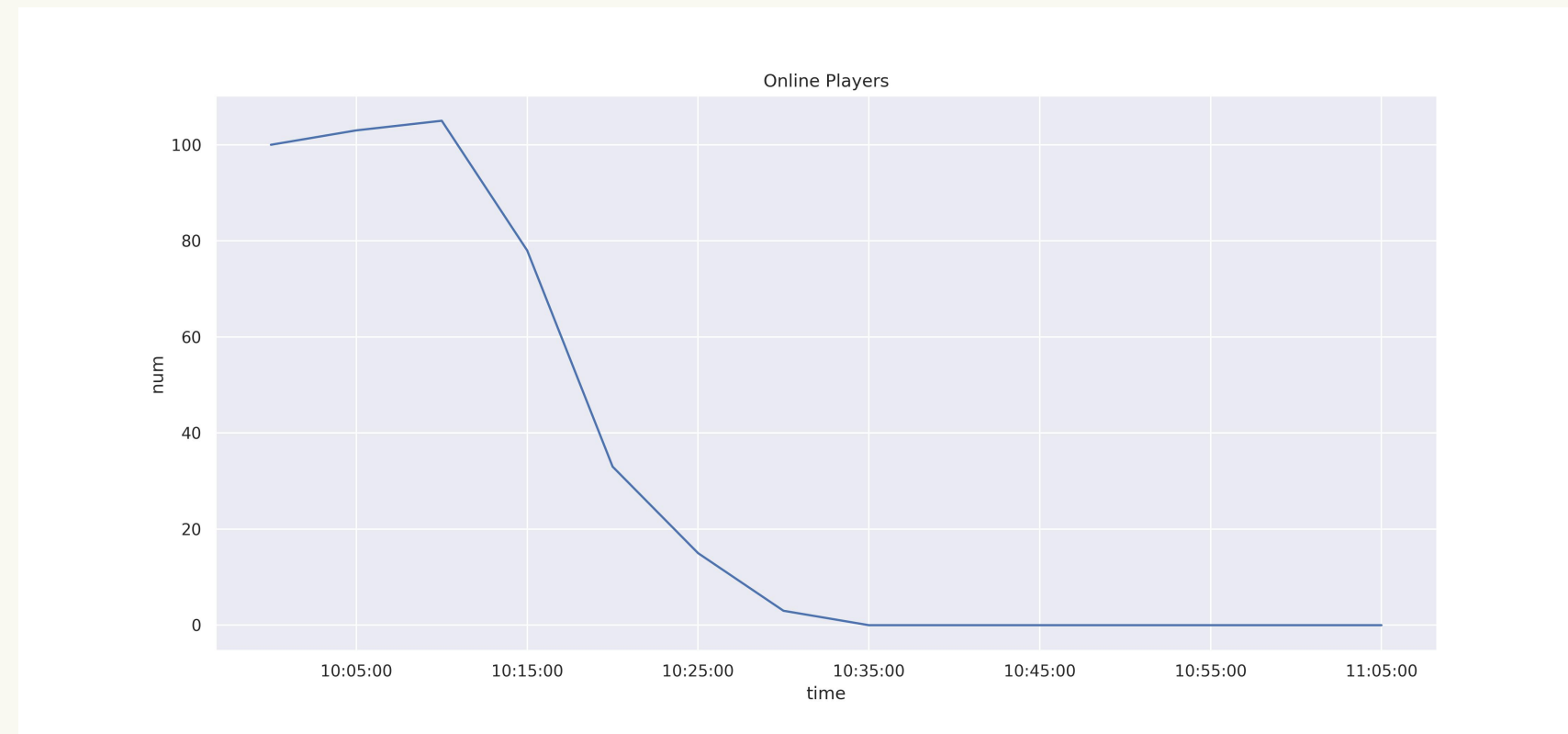


What is Anomaly

- Anomaly, also referred to as outlier, is generally understood to be the identification of rare items, events or observations which deviate significantly from the majority of the data and do not conform to a well defined notion of normal behaviour.

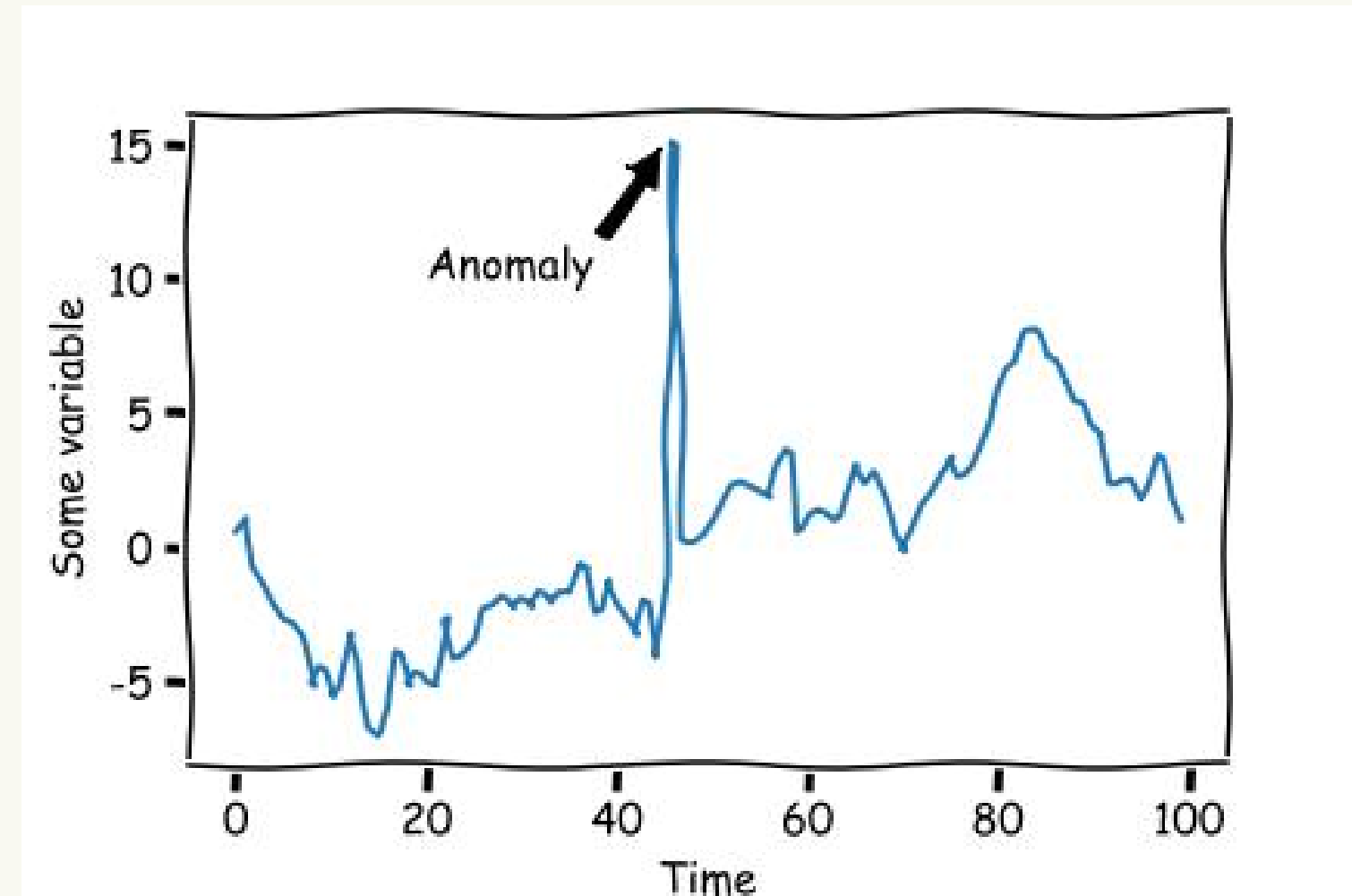
Normal Does Not Means Constant

- For example, the number of online players during the shutdown maintenance of a game is zero.If the number is not 0 at this time, it is an anomaly.



Anomaly Detection for Time Series Data

- Time series data is composed of a sequence of values over time. Time series data anomaly detection first create a baseline for normal behavior. With that baseline, time series data anomaly detection systems can track seasonality.



Know What vs Know Why

Anomaly
Detection

V
S

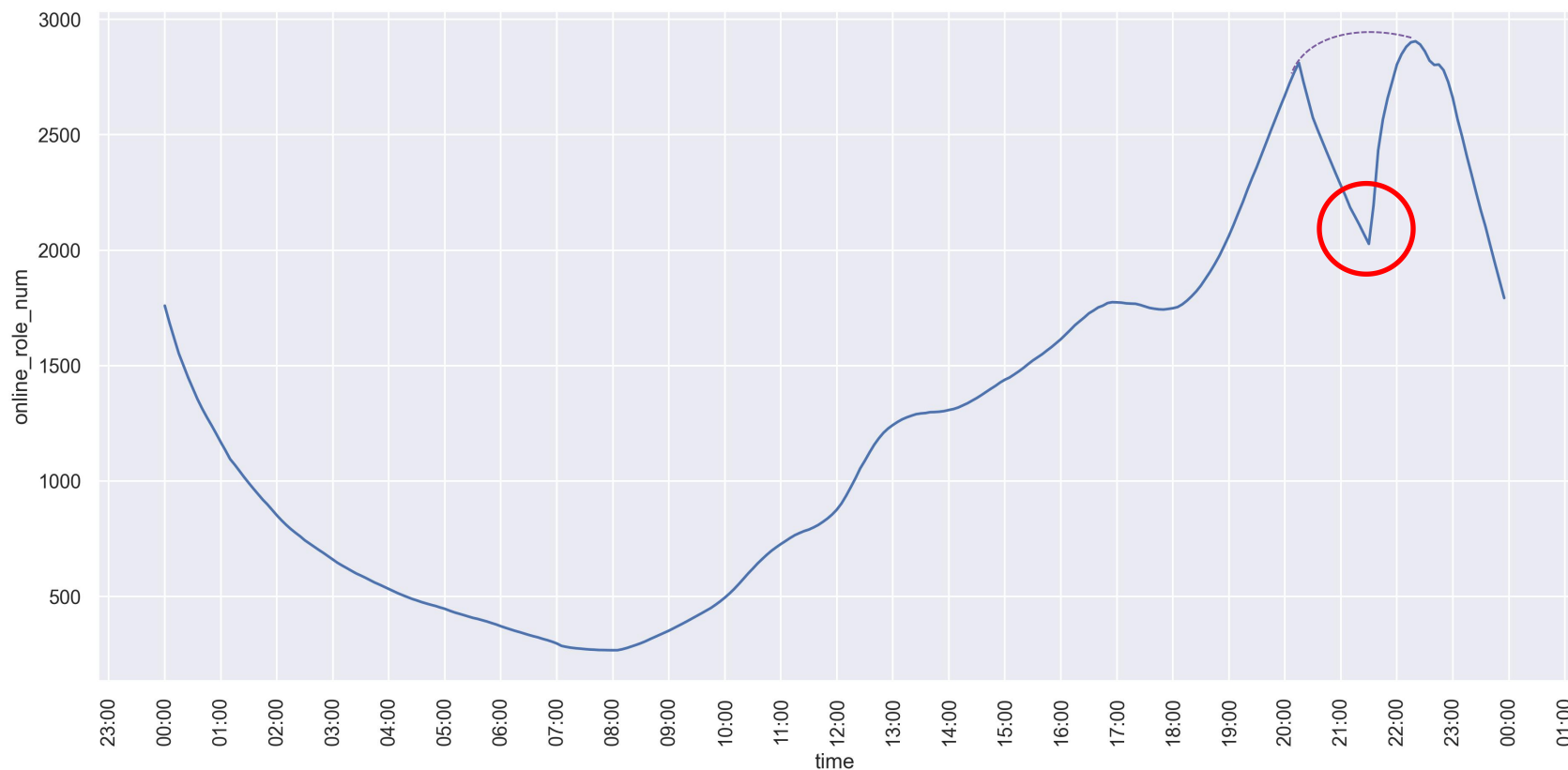
Anomaly
Attribution

What's Next?

- We will discuss all kinds of online player number anomalies, because the online player number curve is clear and easy to understand.
- We will analyze the infrastructure anomalies behind these cases, because the whole pipeline of infra is relatively clear.

A Real-Time Anomaly Case

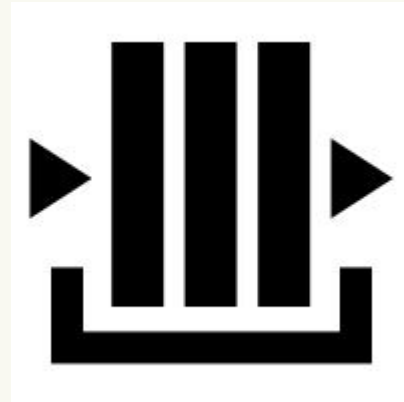
- When we see abnormal fluctuations in real-time data statistics, can we know what happened right now?



Real-time Data Pipeline



Game
Server



Message
Queue



Data
Processing



Dashboard

Three Sources of Anomalies

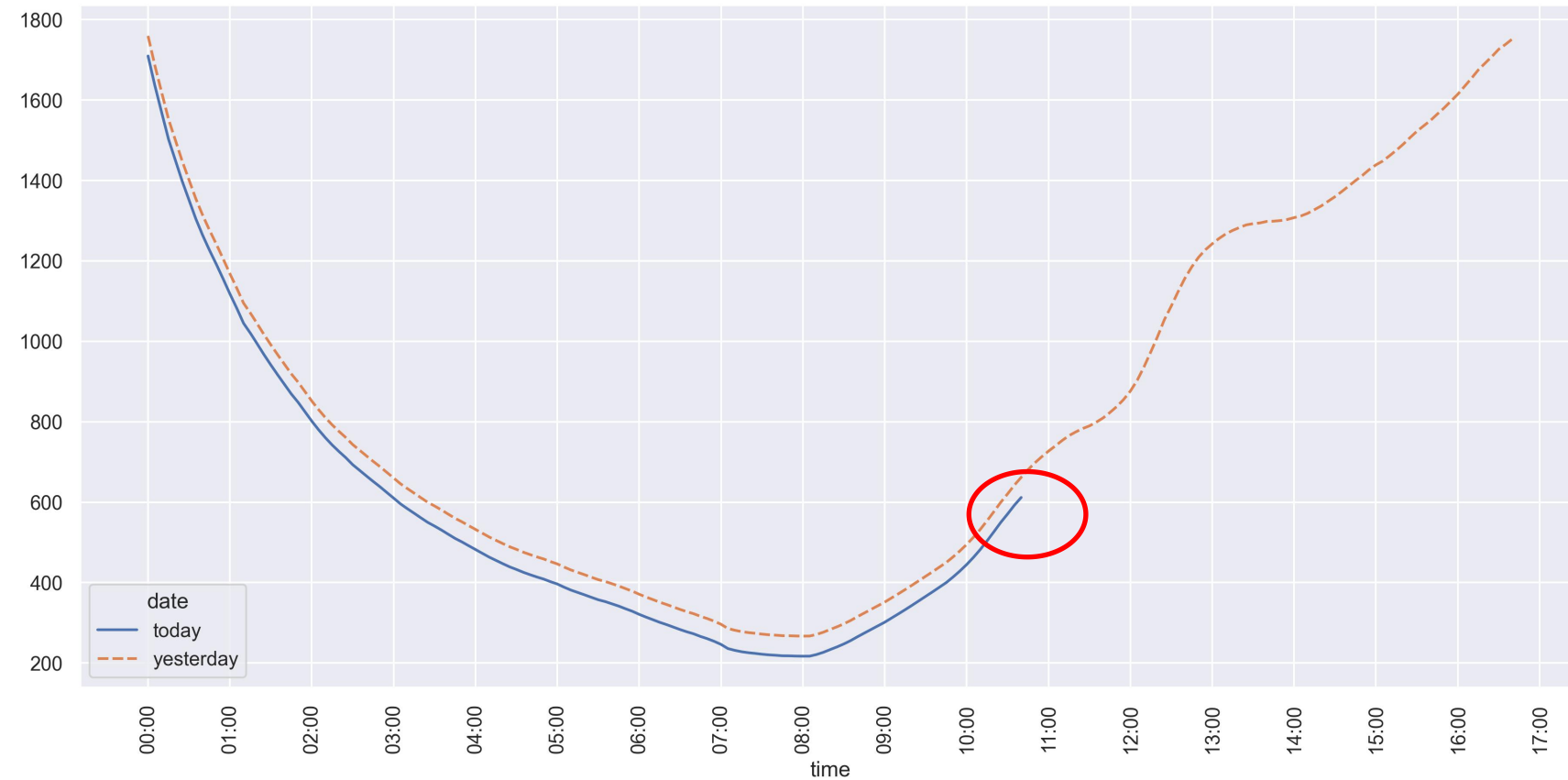
- Game Server Crashes
- Data Collection & Transmission Failure
- Errors in Data Processing

Common Anomalies in Dashboard

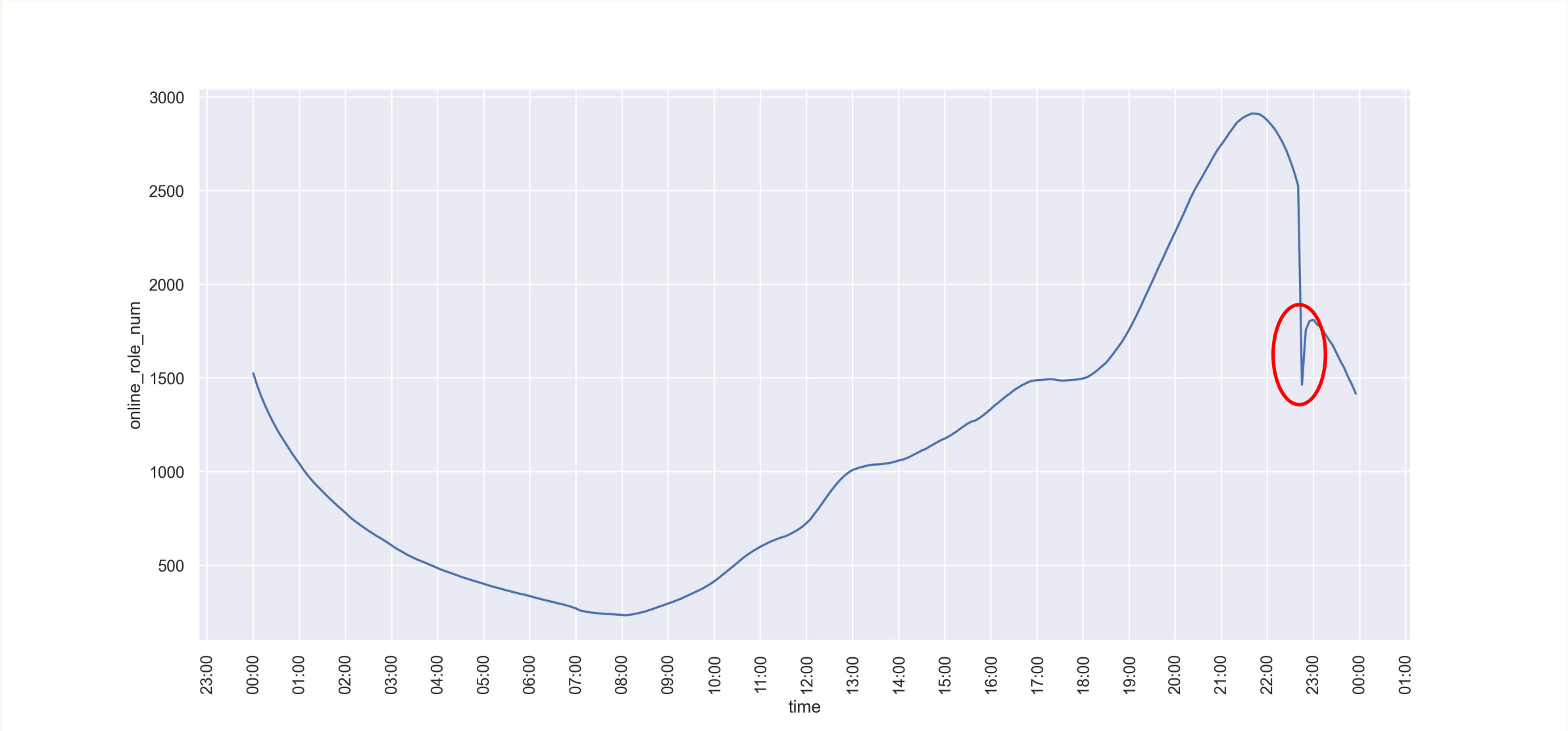
- Delay
- Jitter
- Fall to Zero

Delay

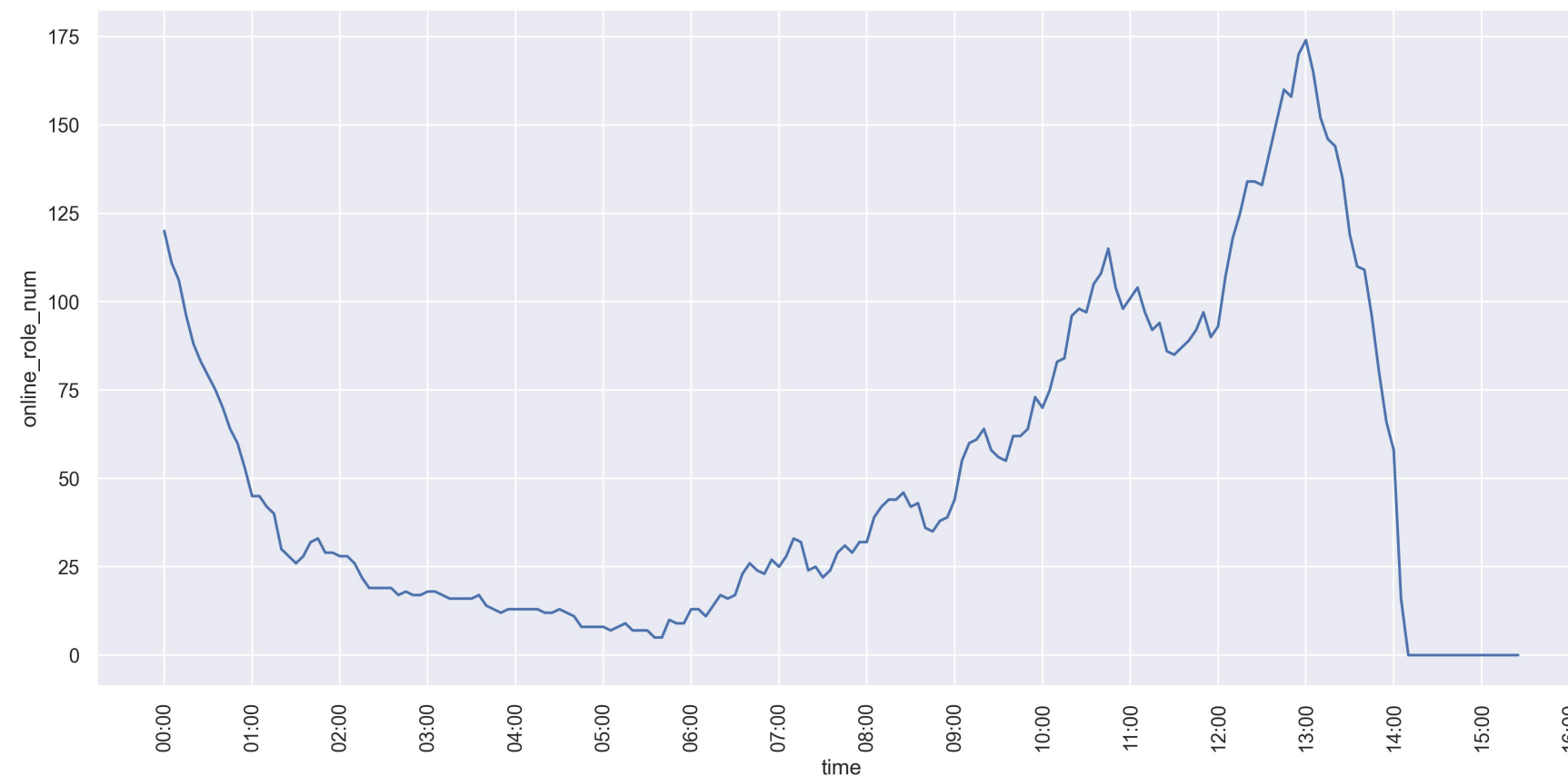
Current Time is 11 o'clock



Jitter



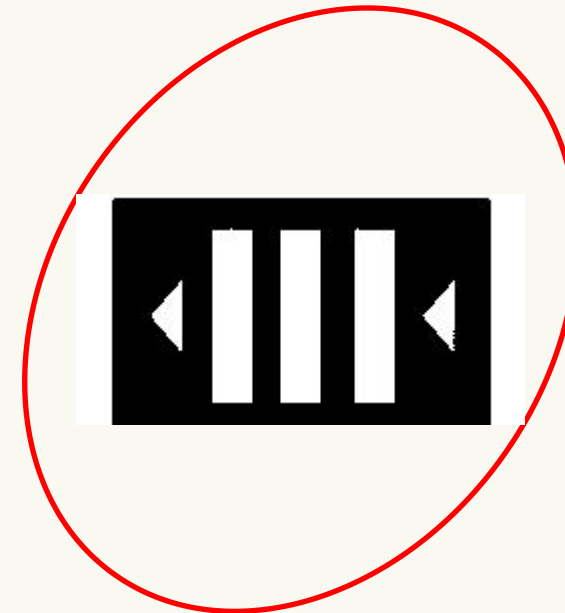
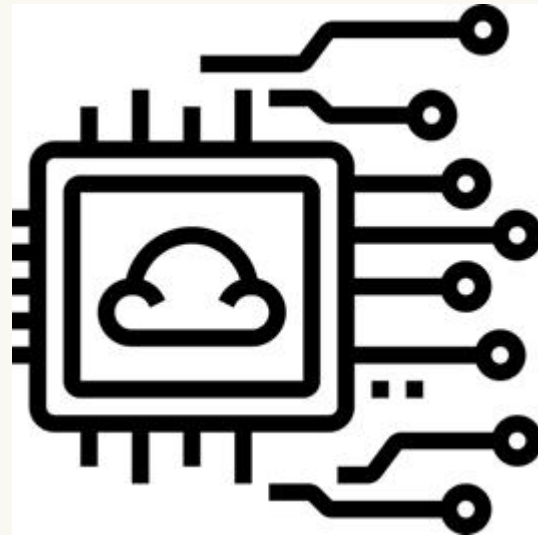
Fall to Zero



When you see the various abnormal dashboards mentioned above, how can you find out at which stage the problem occurred in the real-time data pipeline?

Part.2 Practice in ThunderFire UX

Troubleshooting the Anomalies



Game
Server

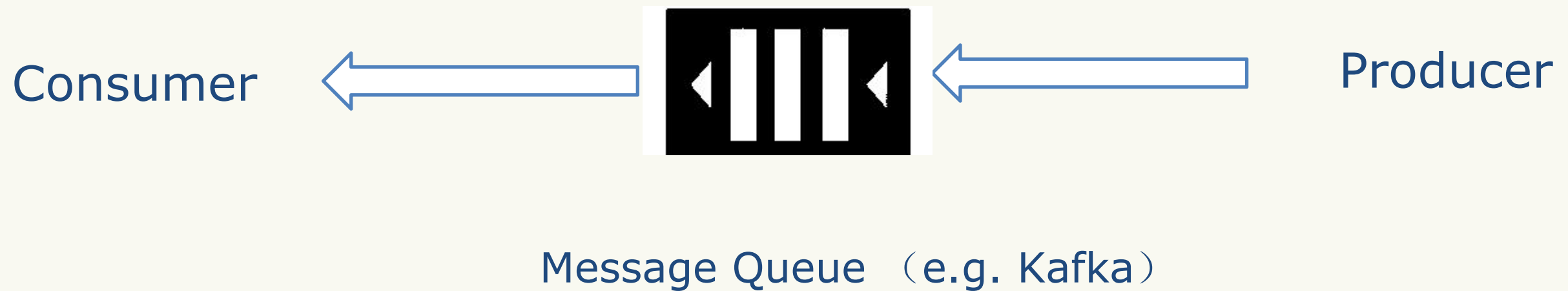
Data
Processing

Message
Queue

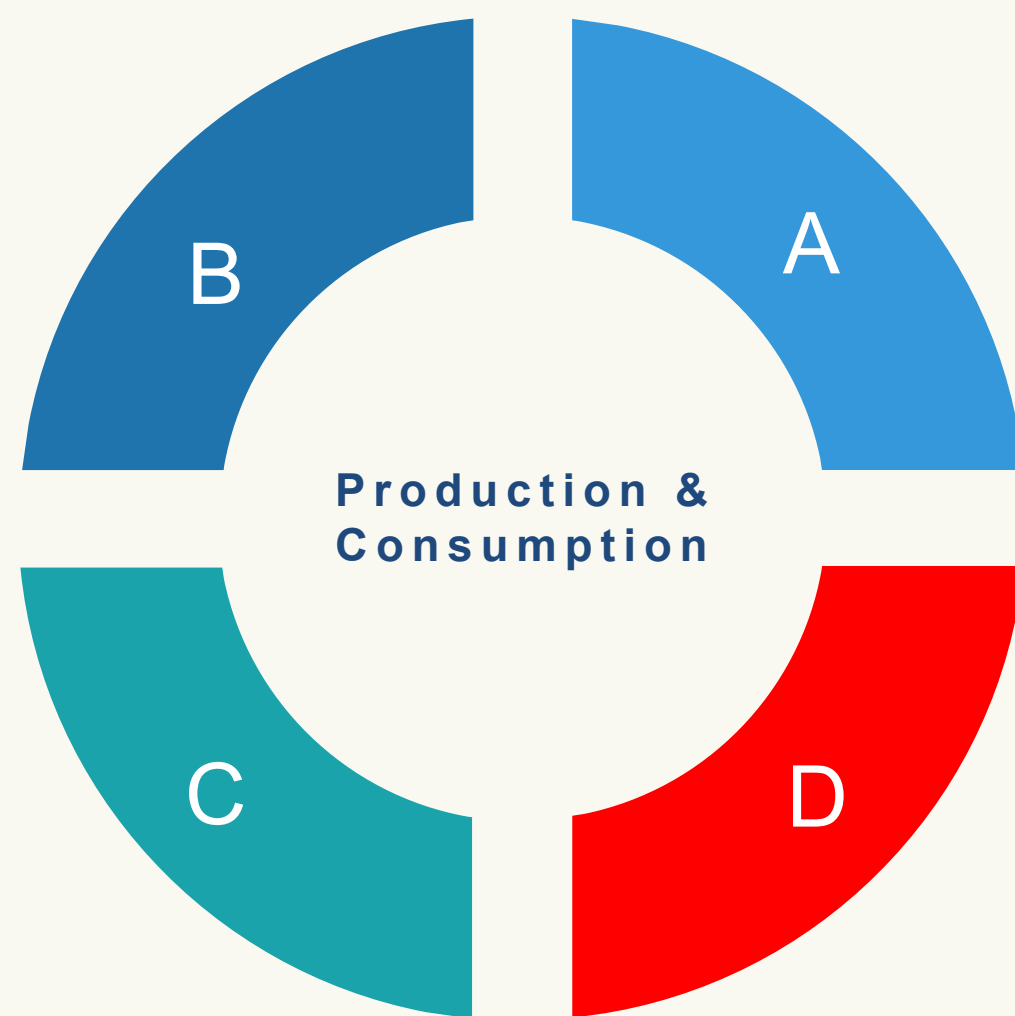
Dashboard

Monitoring the Message Queue

- Connecting upstream data production and downstream data consumption



Four Quadrants of Production and Consumption



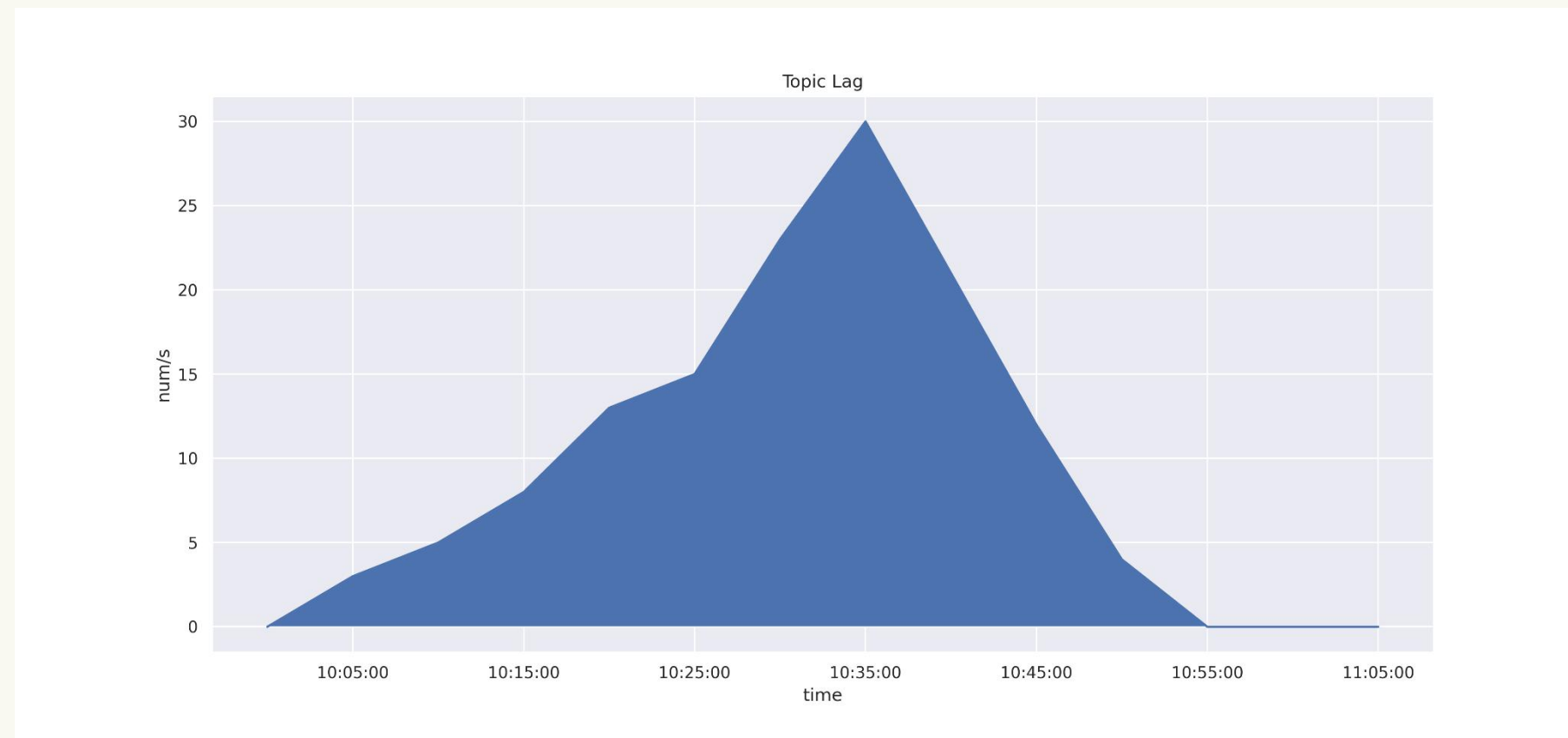
- A** Normal production & normal consumption
- B** Normal production & abnormal consumption
- C** Abnormal production & normal consumption
- D** Abnormal production & abnormal consumption

Key Monitoring Indicators

- Topic Lag
- Consumer Rate
- Producer Rate

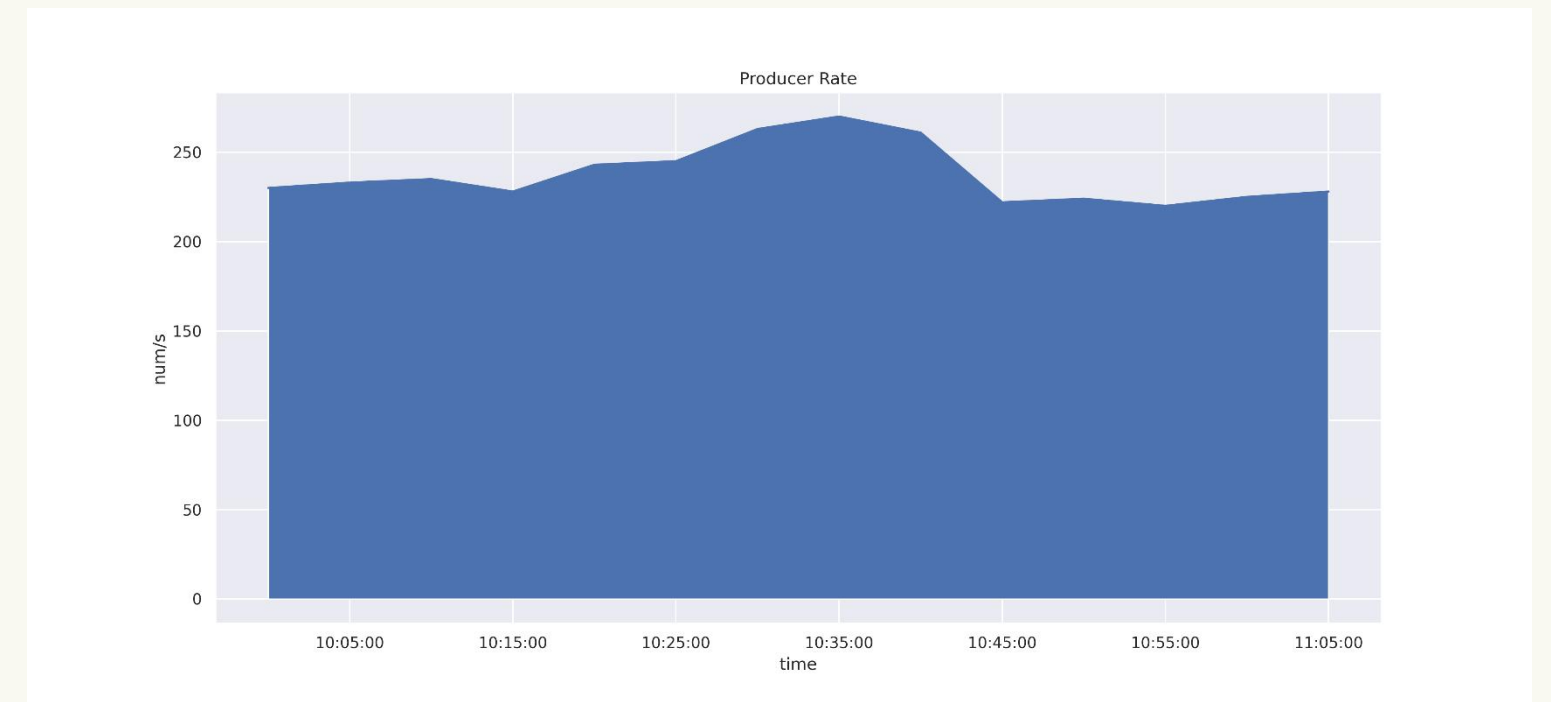
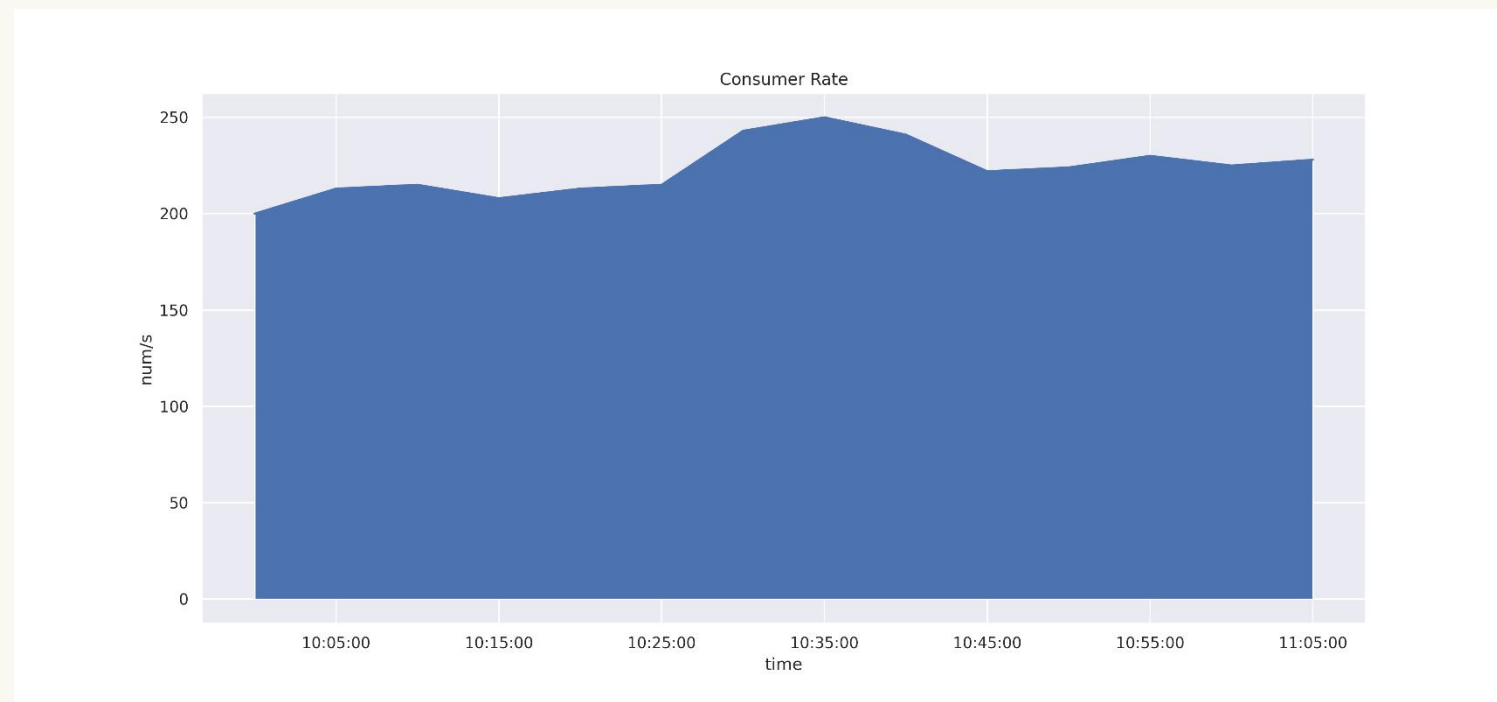
Topic Lag

- *Lag* is simply the delta between the last produced message and the last consumer's committed offset.



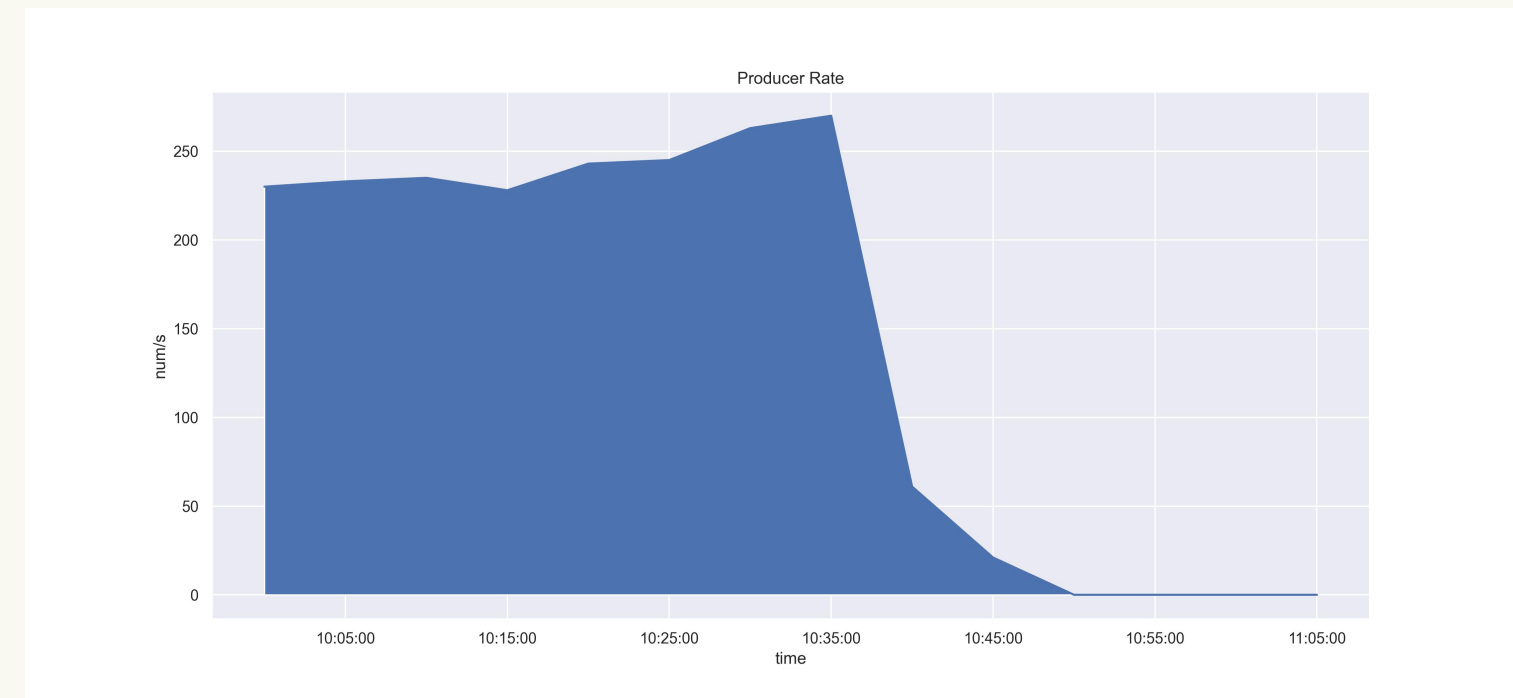
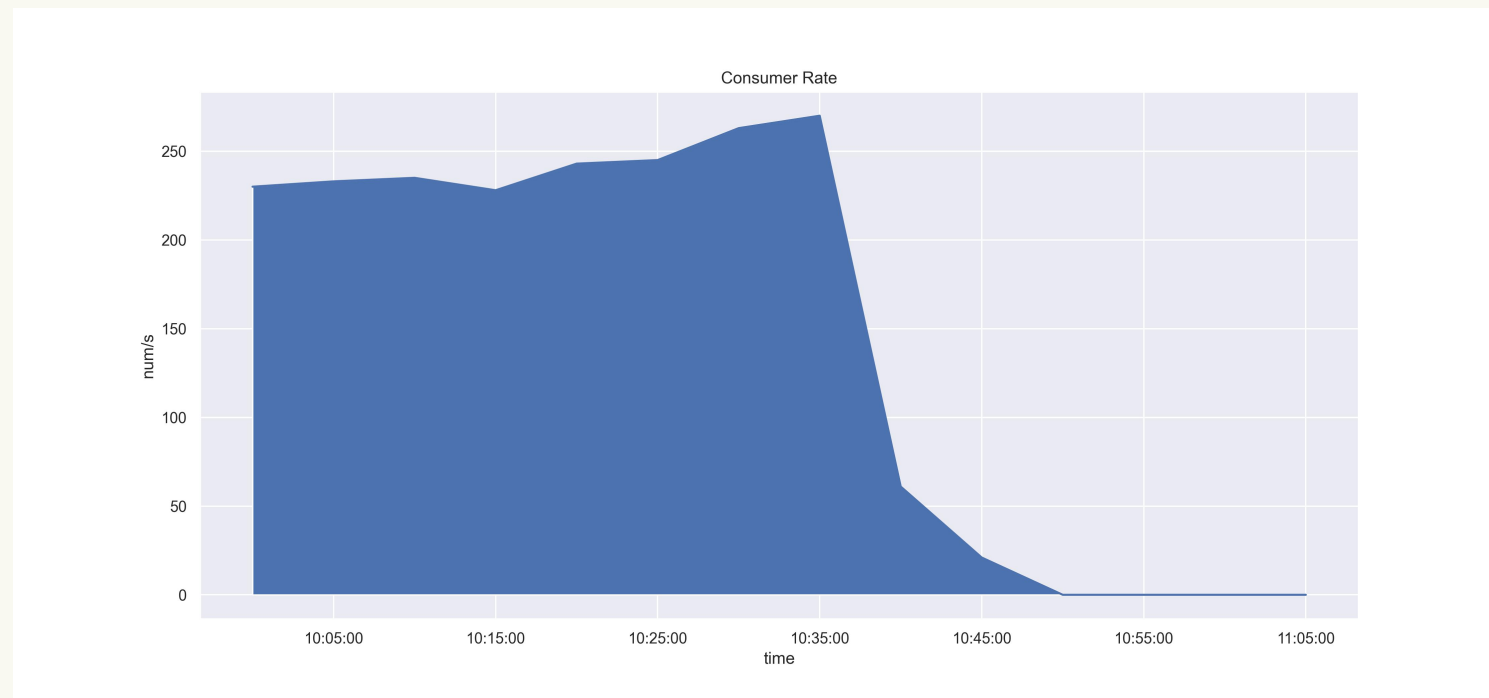
Consumer Rate and Producer Rate

- The rate of records consumed per second
- The rate of records produced per second



Cases in Practice

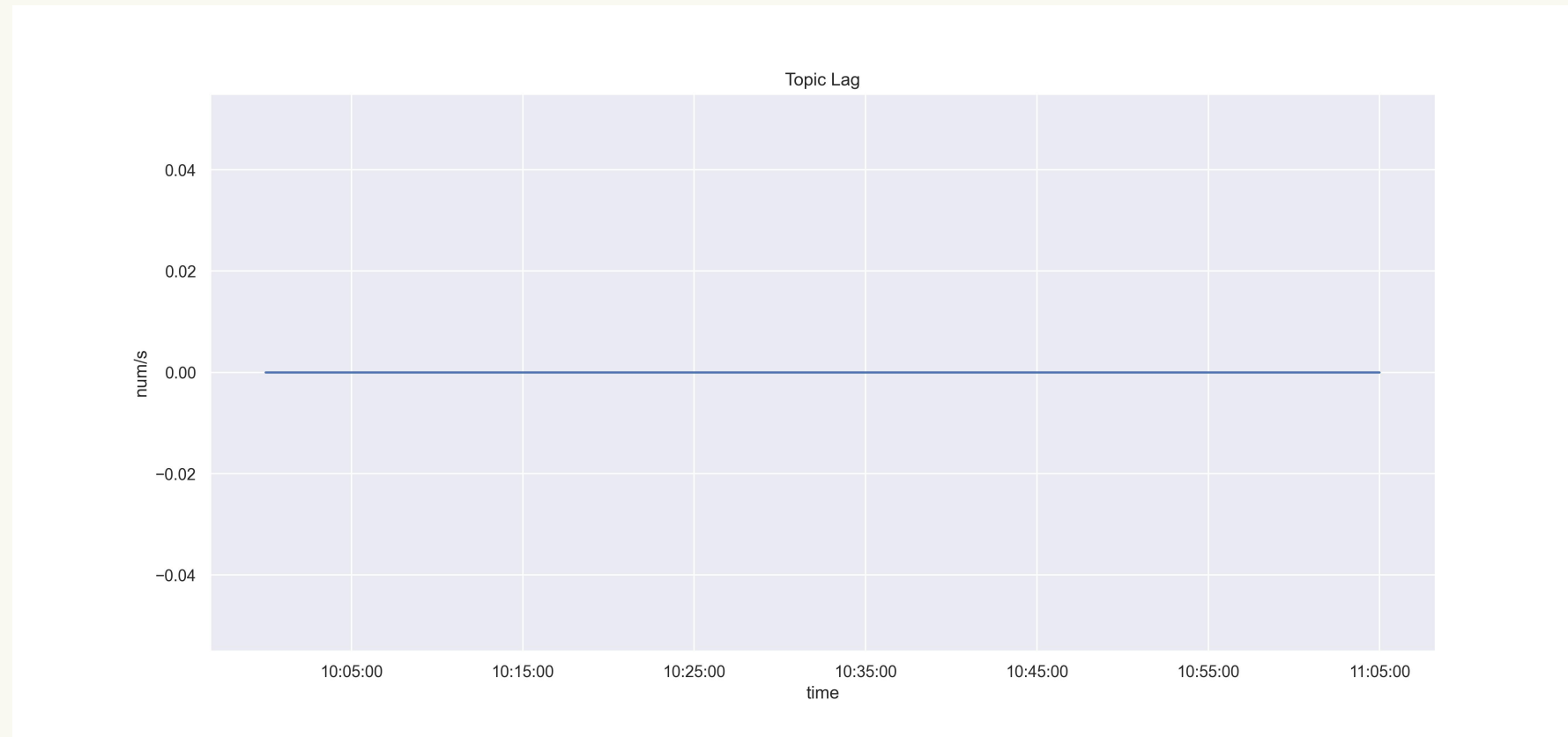
- stop producing messages (normal consumption)



Similar Trends

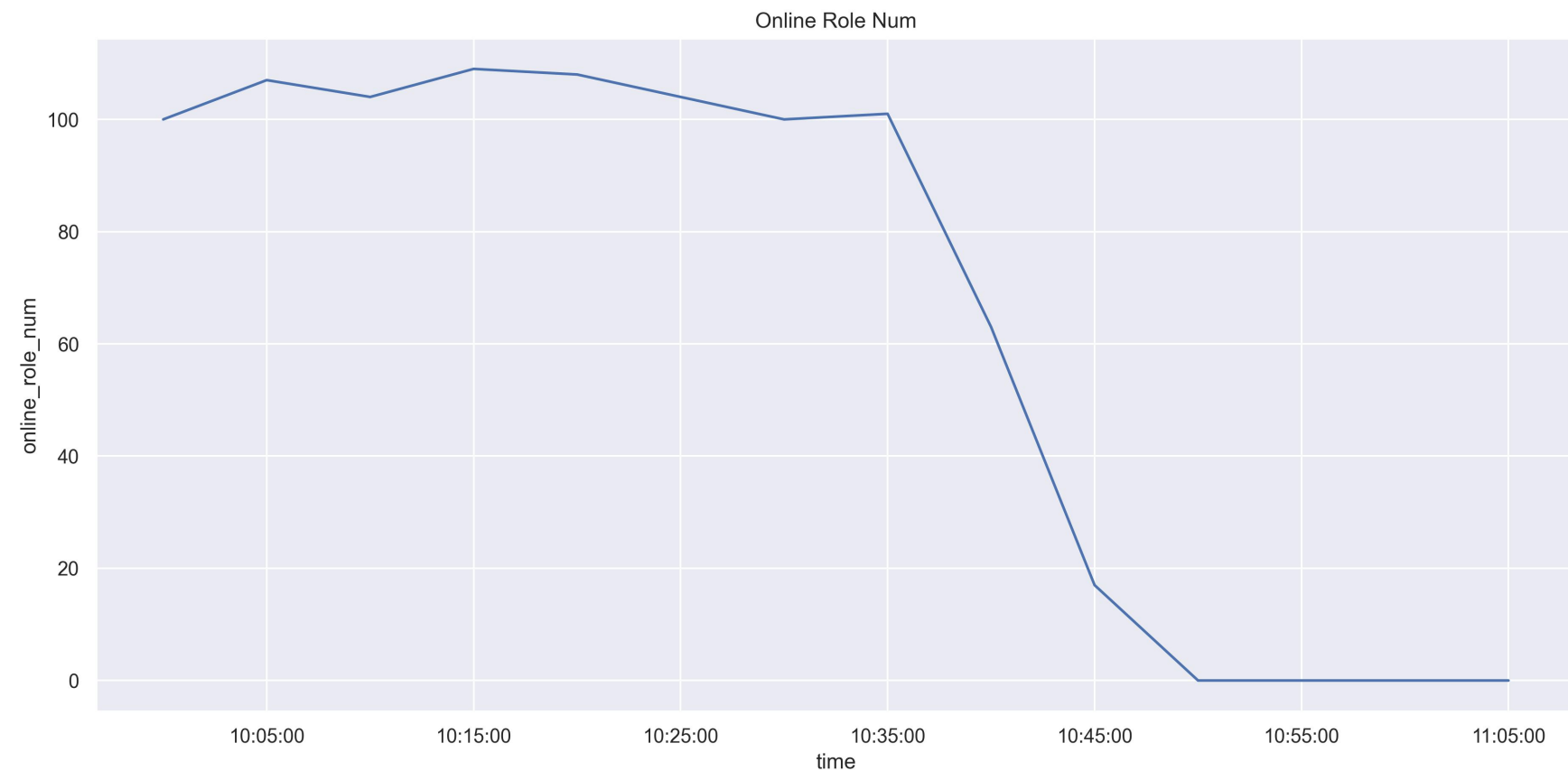
Case1——Stop Producing Messages

- No Topic Lag



Case1——Stop Producing Messages

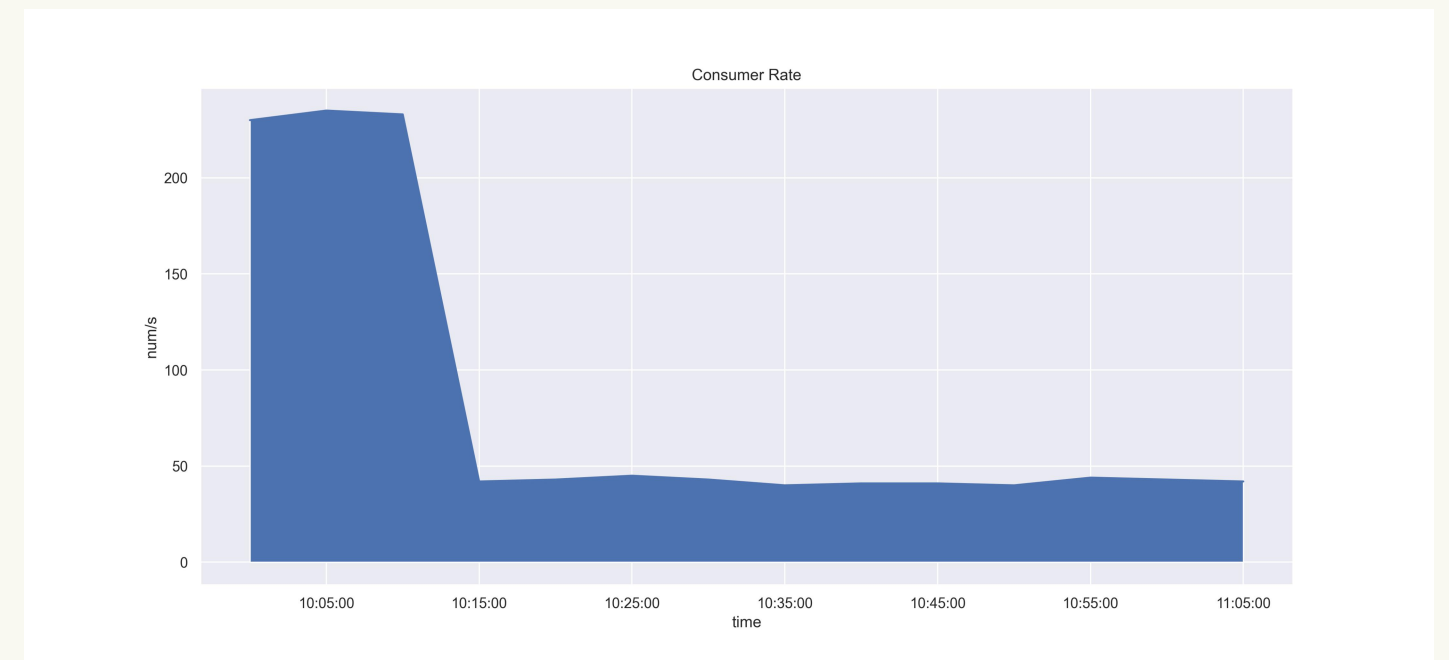
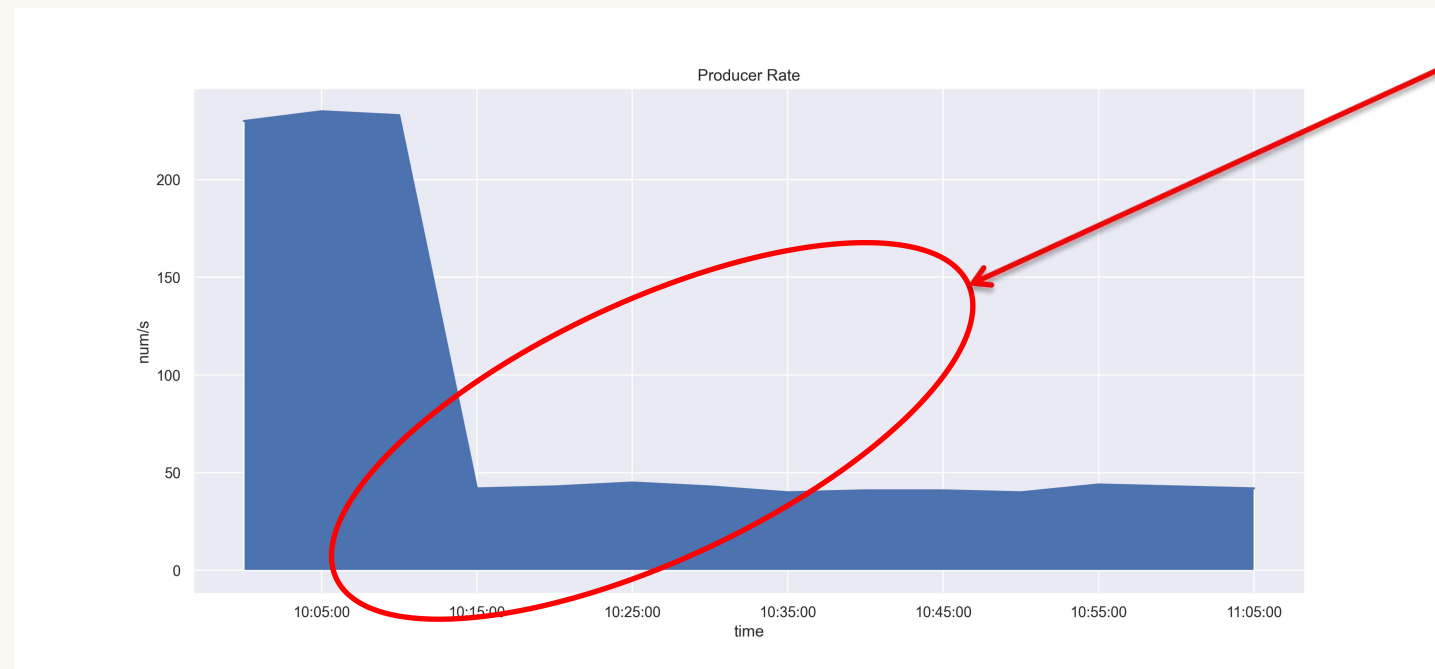
- Fall to Zero Dashboard



Case2——Delayed Messages Production

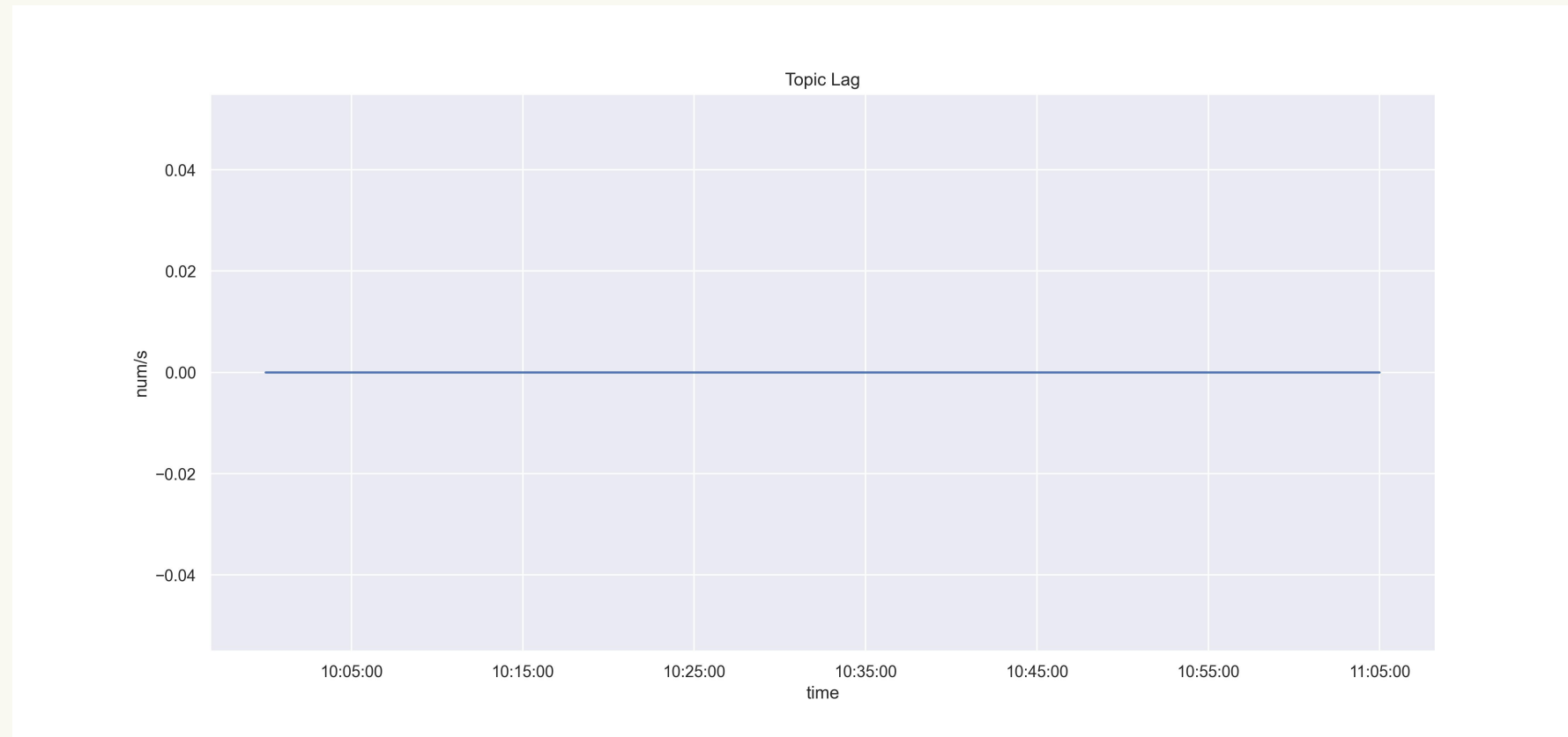
- Similar Trends of Consumer and Producer

network latency



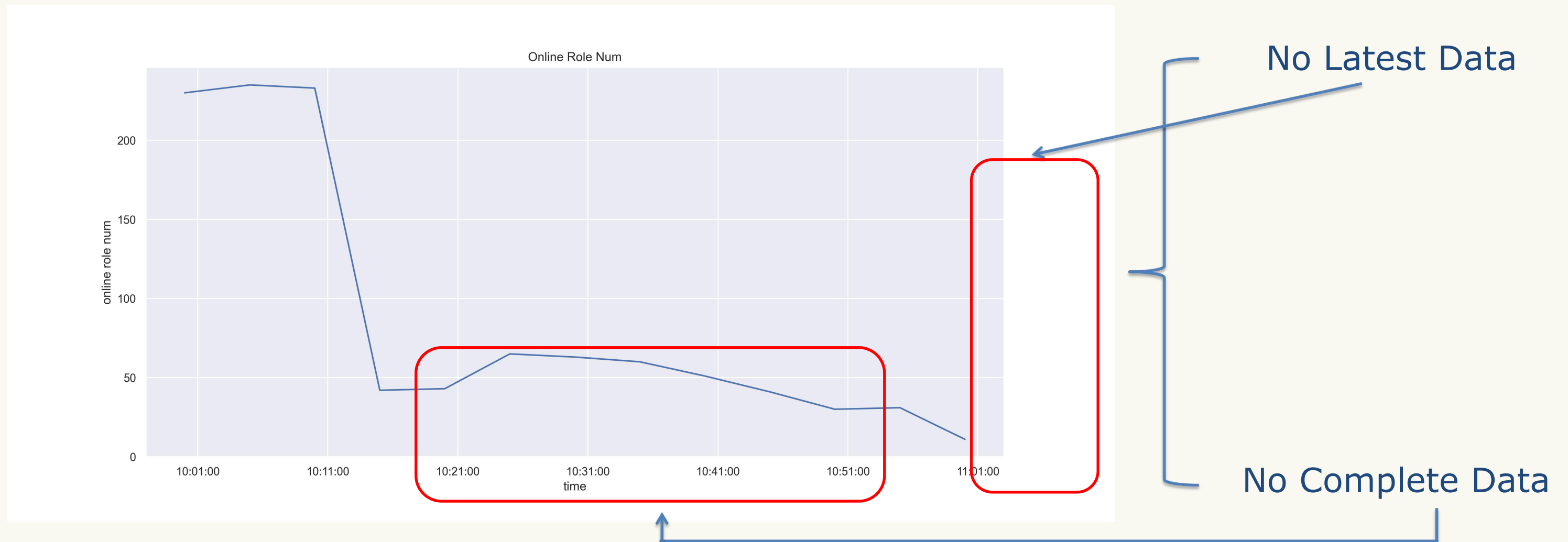
Case2——Delayed Messages Production

- No Topic Lag



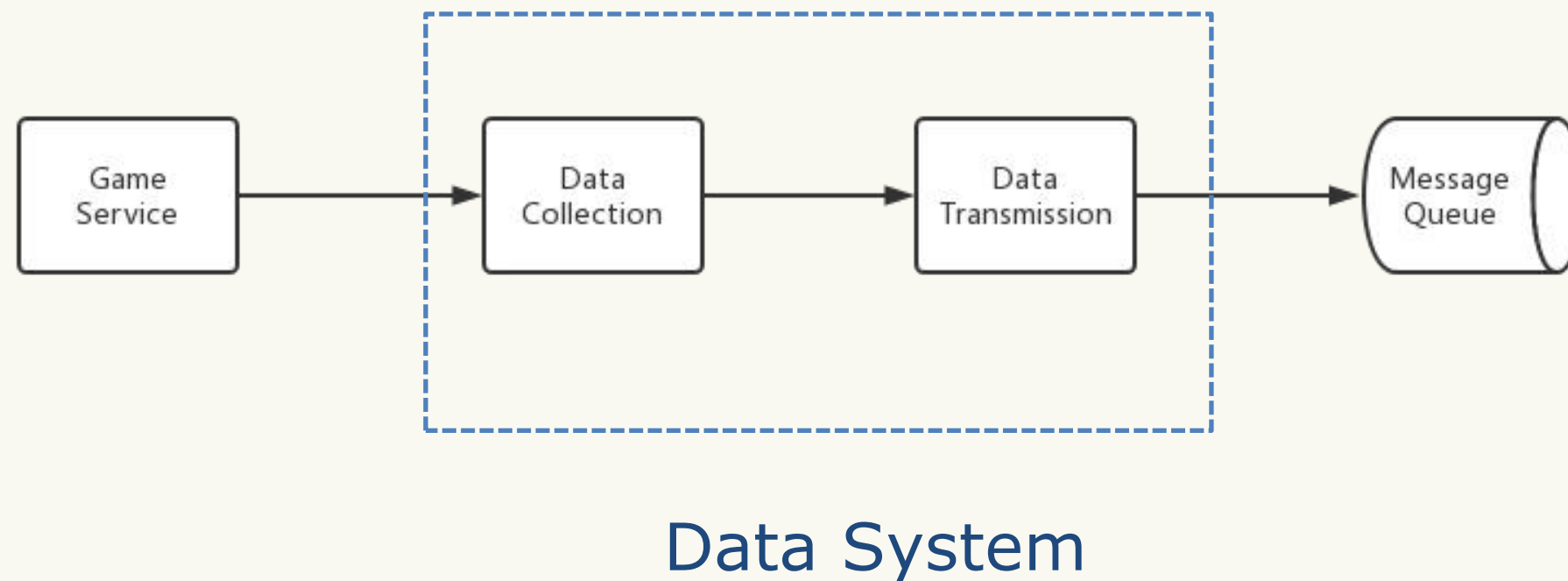
Case2——Delayed Messages Production

- Delayed Pattern Dashboard



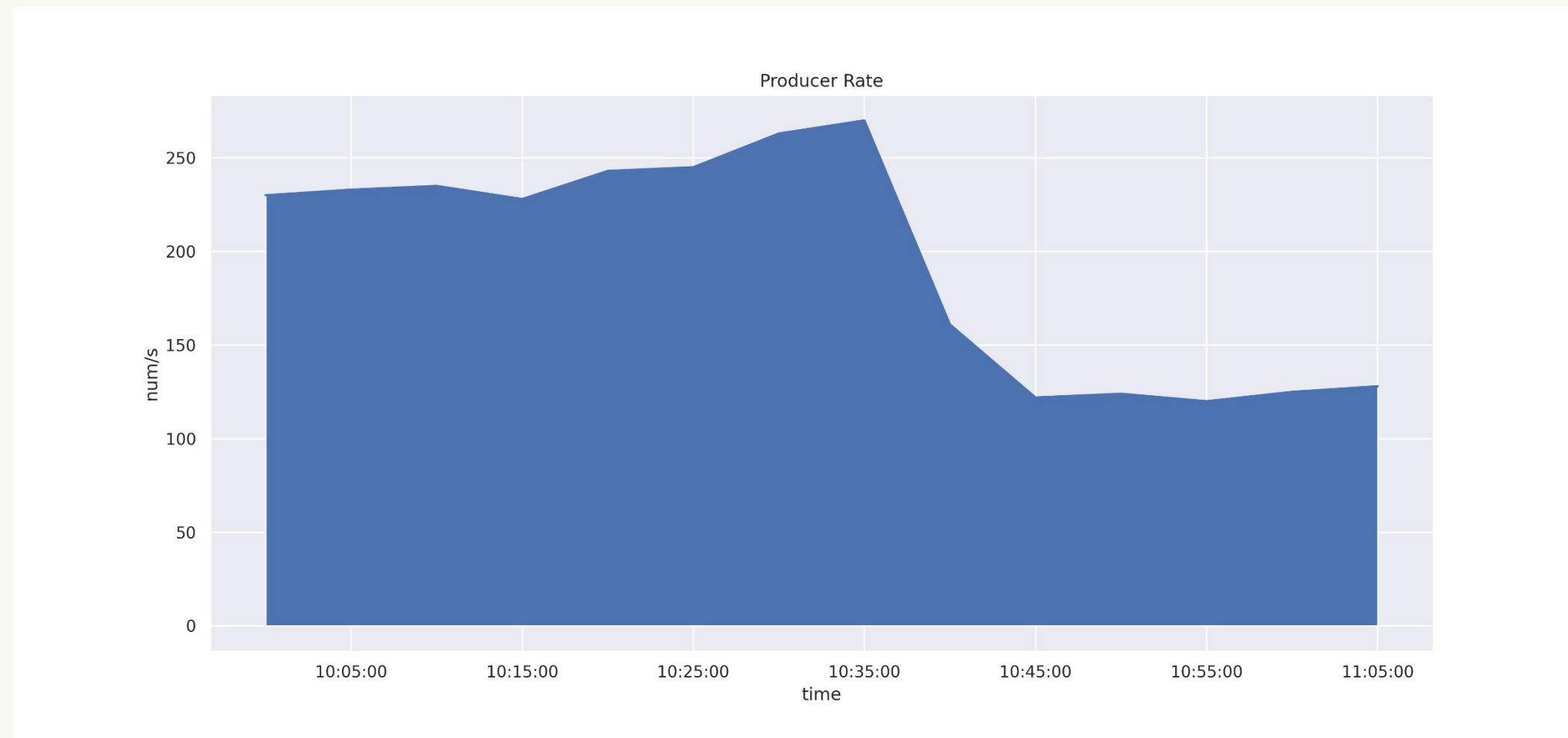
Dive Deeper into Producer Side Anomaly

- Can we find out whether the anomaly occurs in game service or data system?



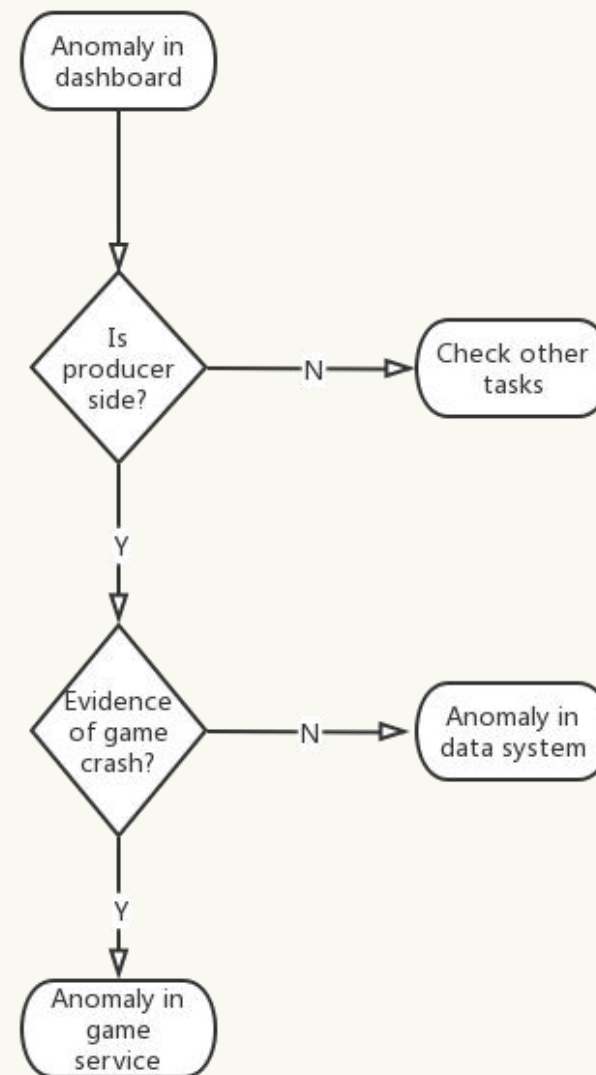
Anomaly in Game Service or Data System

- When monitoring message queue, it is impossible to distinguish 2 kinds of anomalies and will get the same hints.



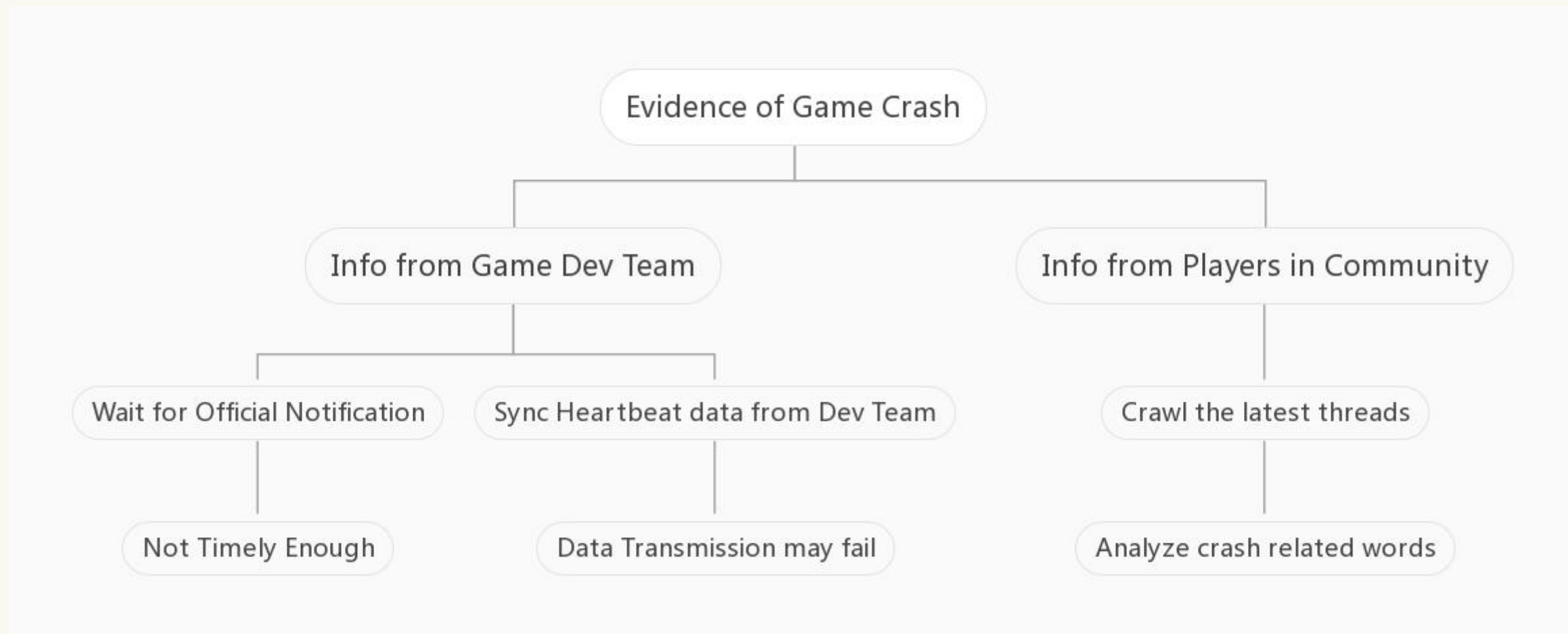
Anomaly in Game Service or Data System

- Process of Anomaly Detection



Anomaly in Game Service or Data System

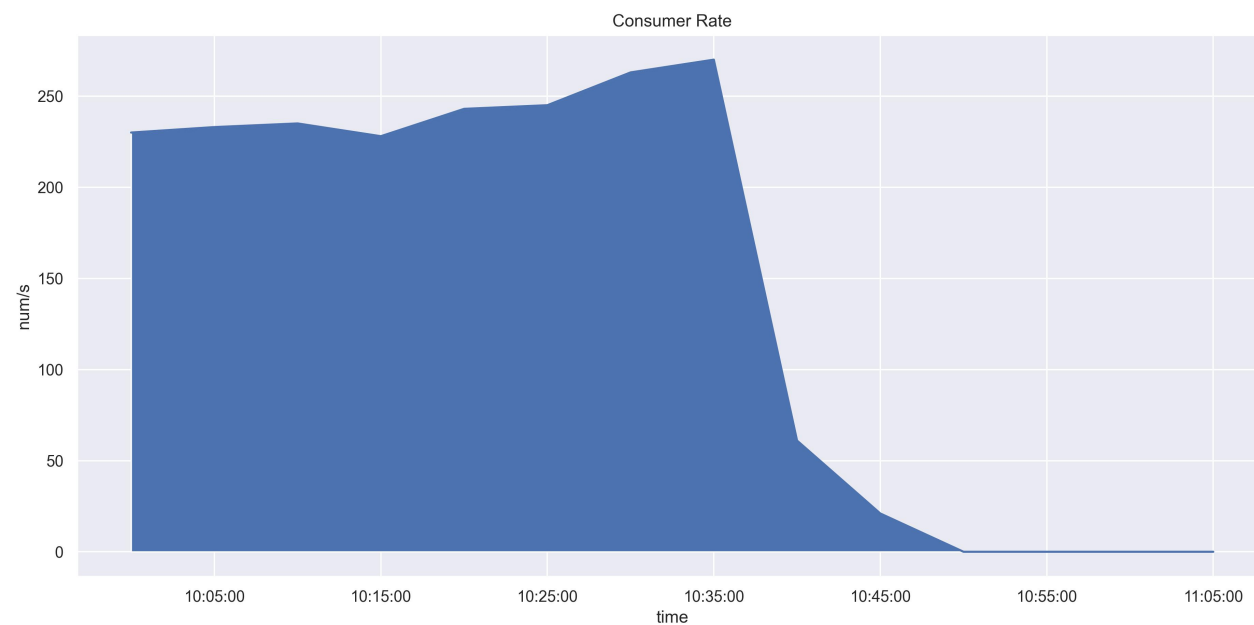
•Evidence of Game Crash



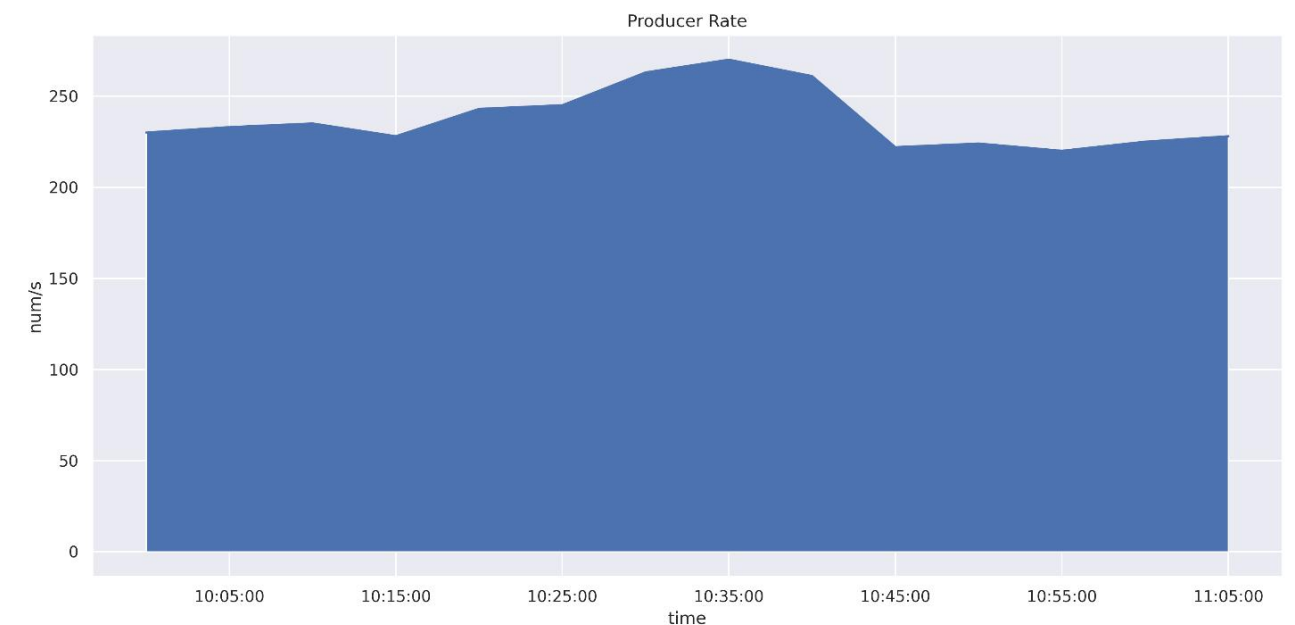
Case3——Abnormal Messages Consumption

- Different Trends of Consumer and Producer

Abnormal

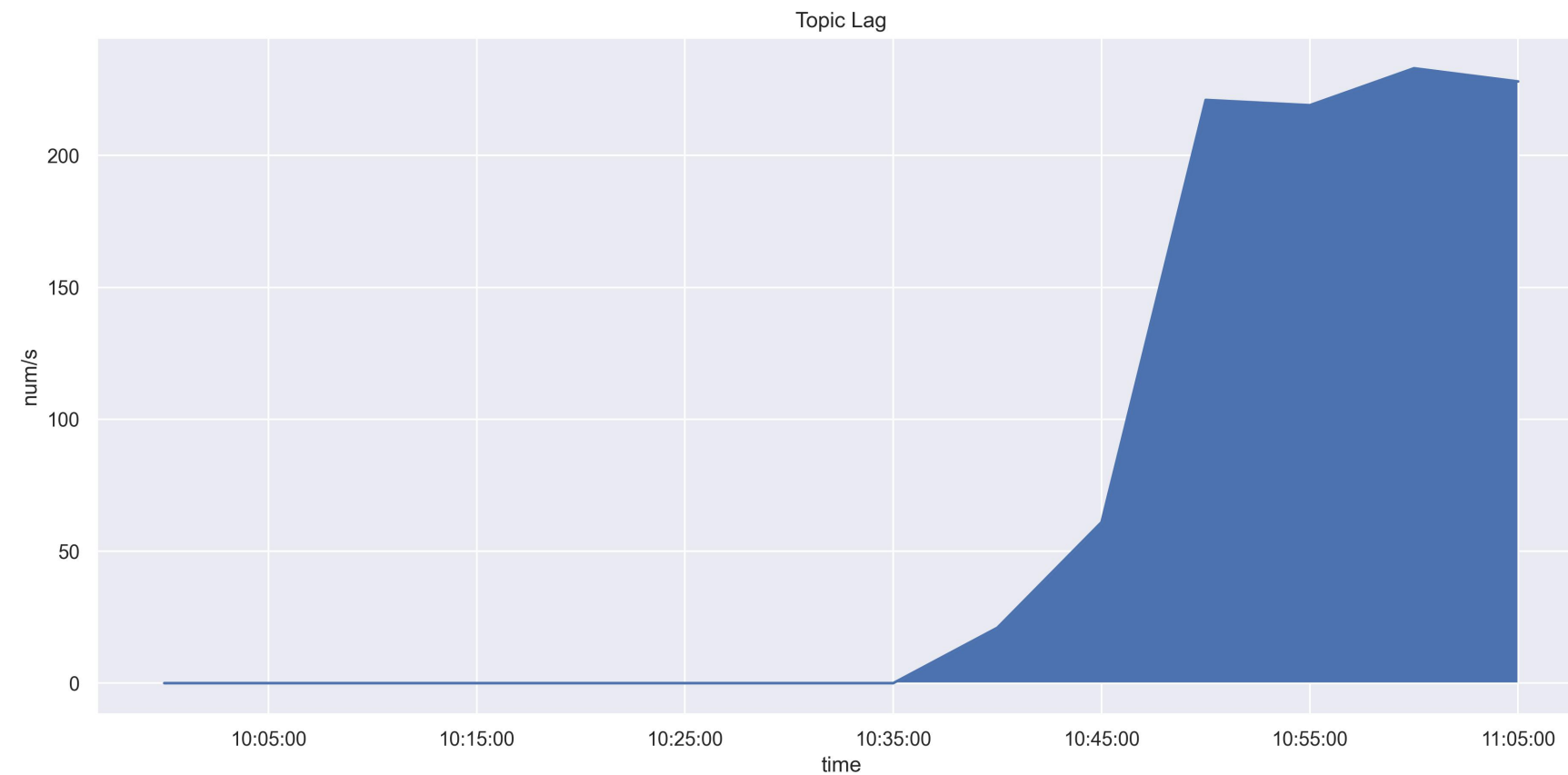


Normal



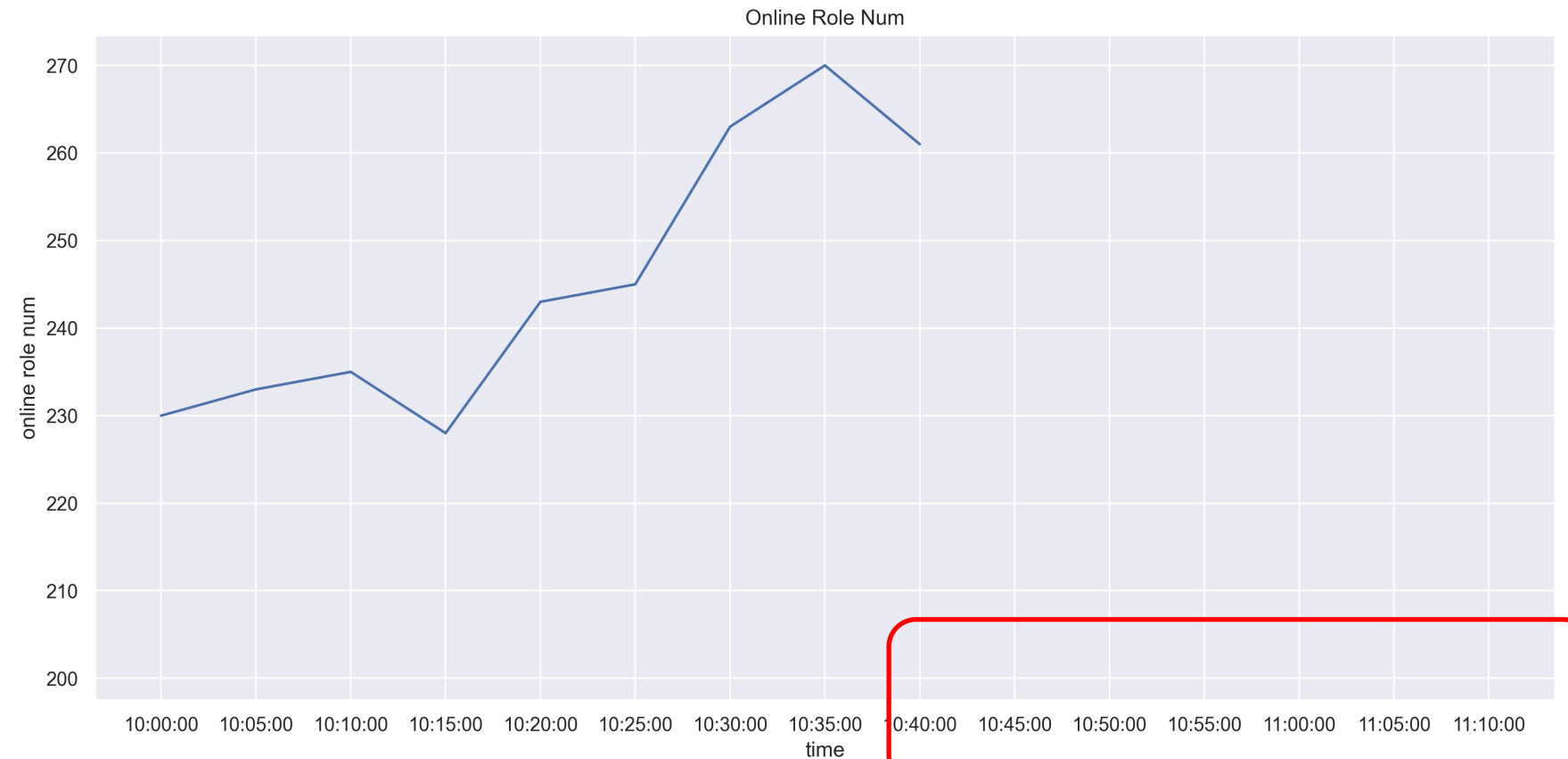
Case3——Abnormal Messages Consumption

- Topic Lag



Case3——Abnormal Messages Consumption

- Dashboard——No Latest Data

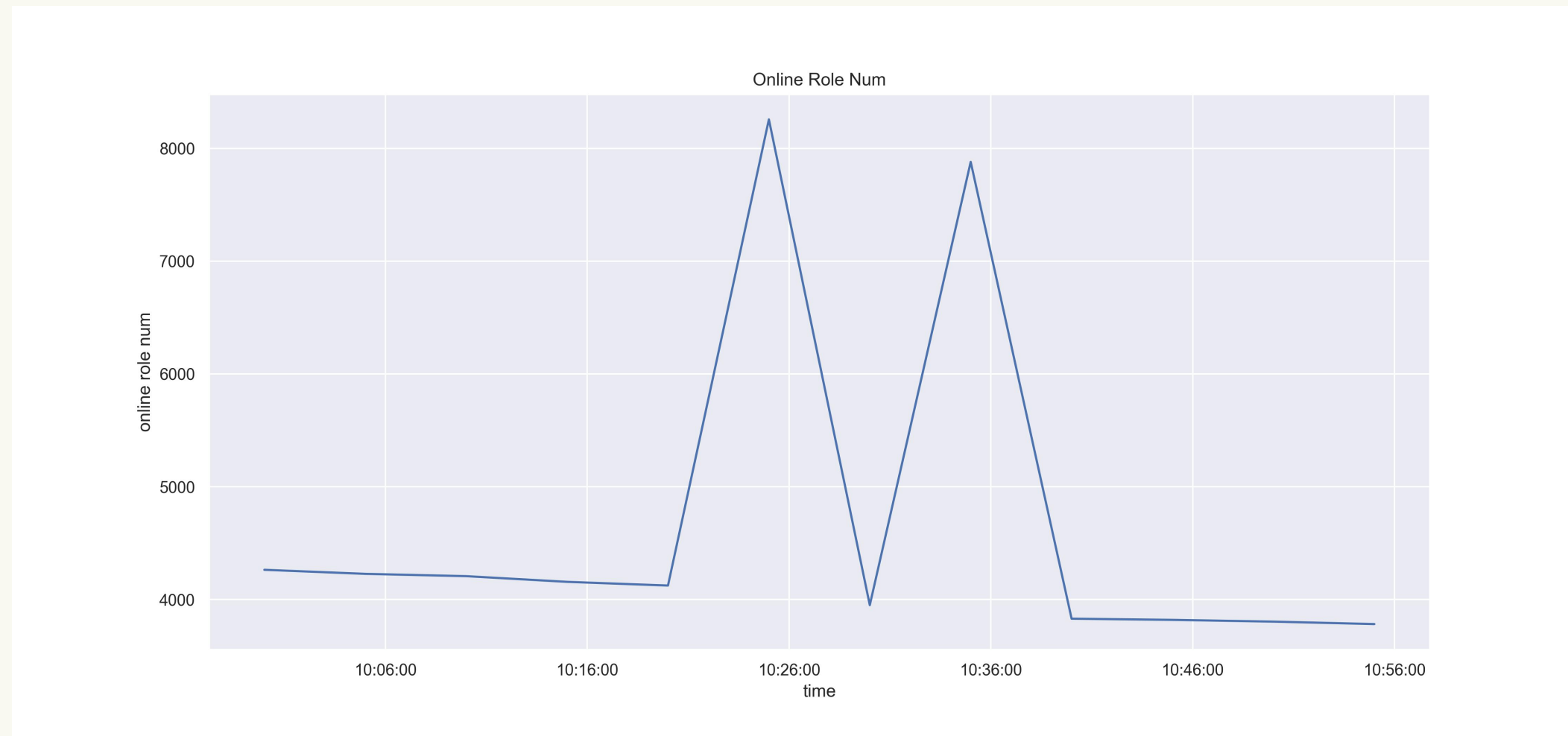


Summary for 3 Cases

By monitoring the message queue, we can find out whether the anomaly comes from the consumer side or the producer side, that is, upstream or downstream.

Special Case——Normal Production & Consumption

- Two obvious jitters in the dashboard



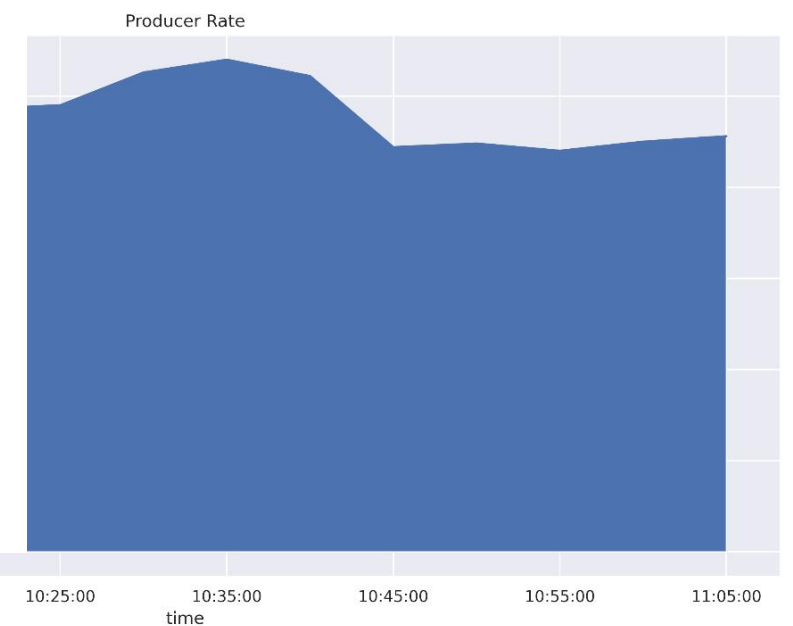
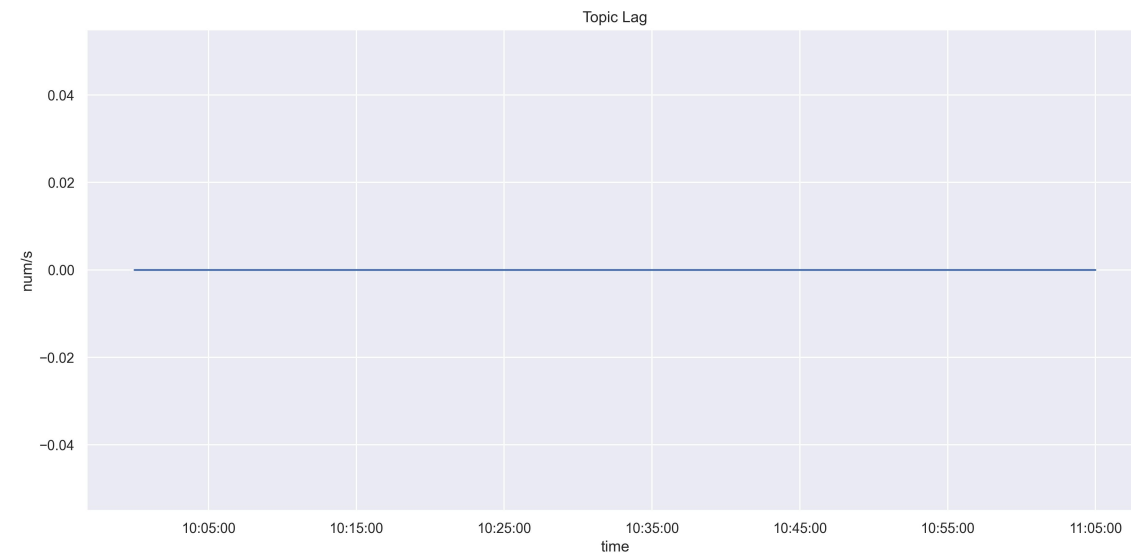
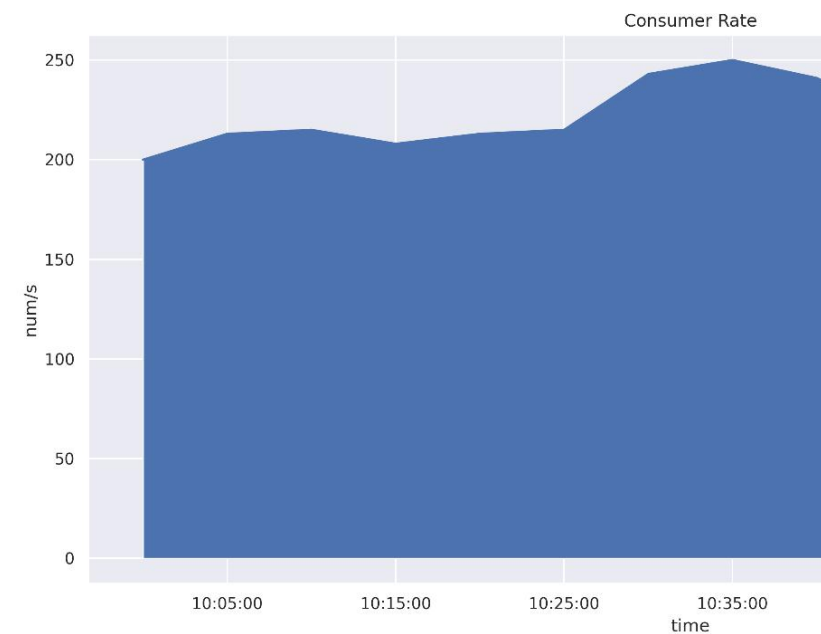
Special Case——Normal Production & Consumption

- But Everything Seems Okay?

Normal

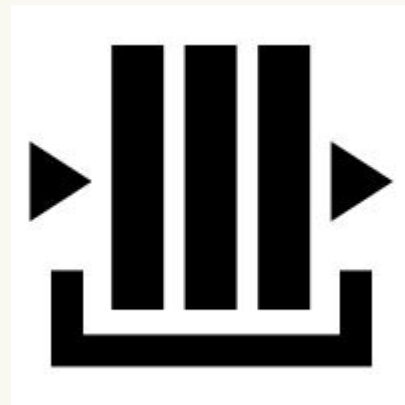
Normal

Normal



Detailed Analysis of Real-time Data

- The data pipeline needs to be updated. The data should be persistent in OLAP DB (we use Apache Kudu here) and can be queried quickly (Apache Impala used)



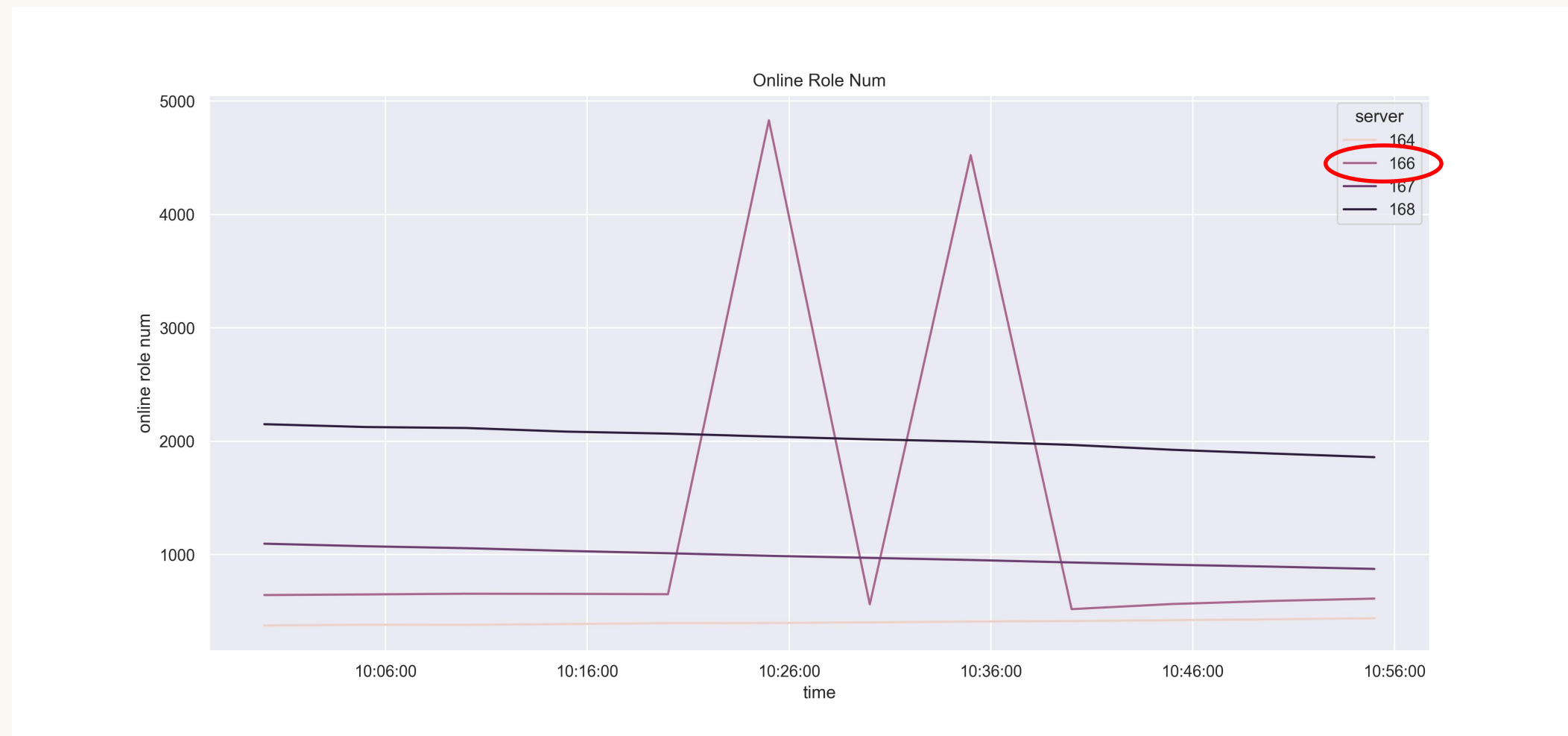
Game
Server

Message
Queue

OLAP
Database

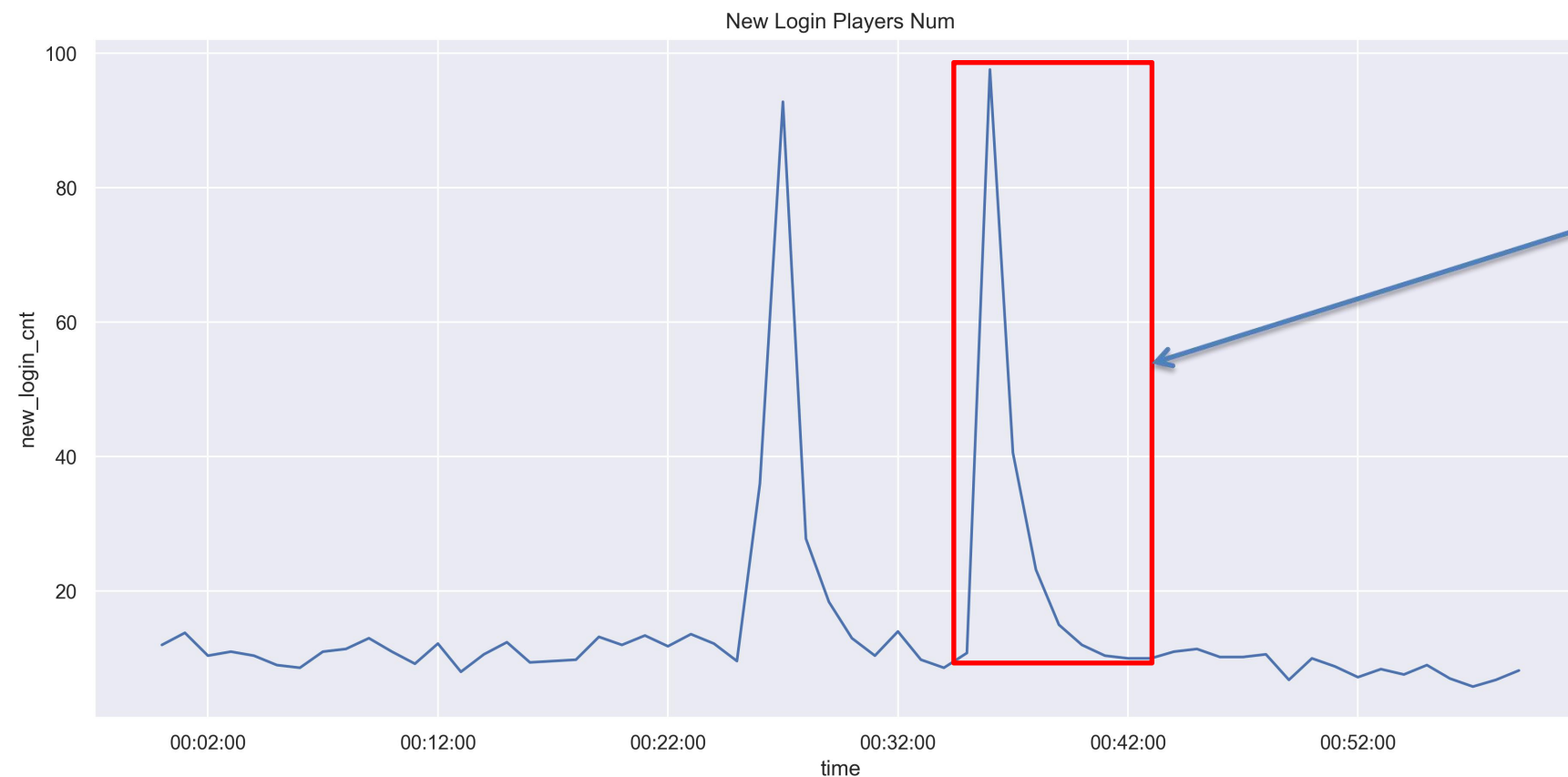
Special Case——Normal Production & Consumption

- Analyze every server in the online role num



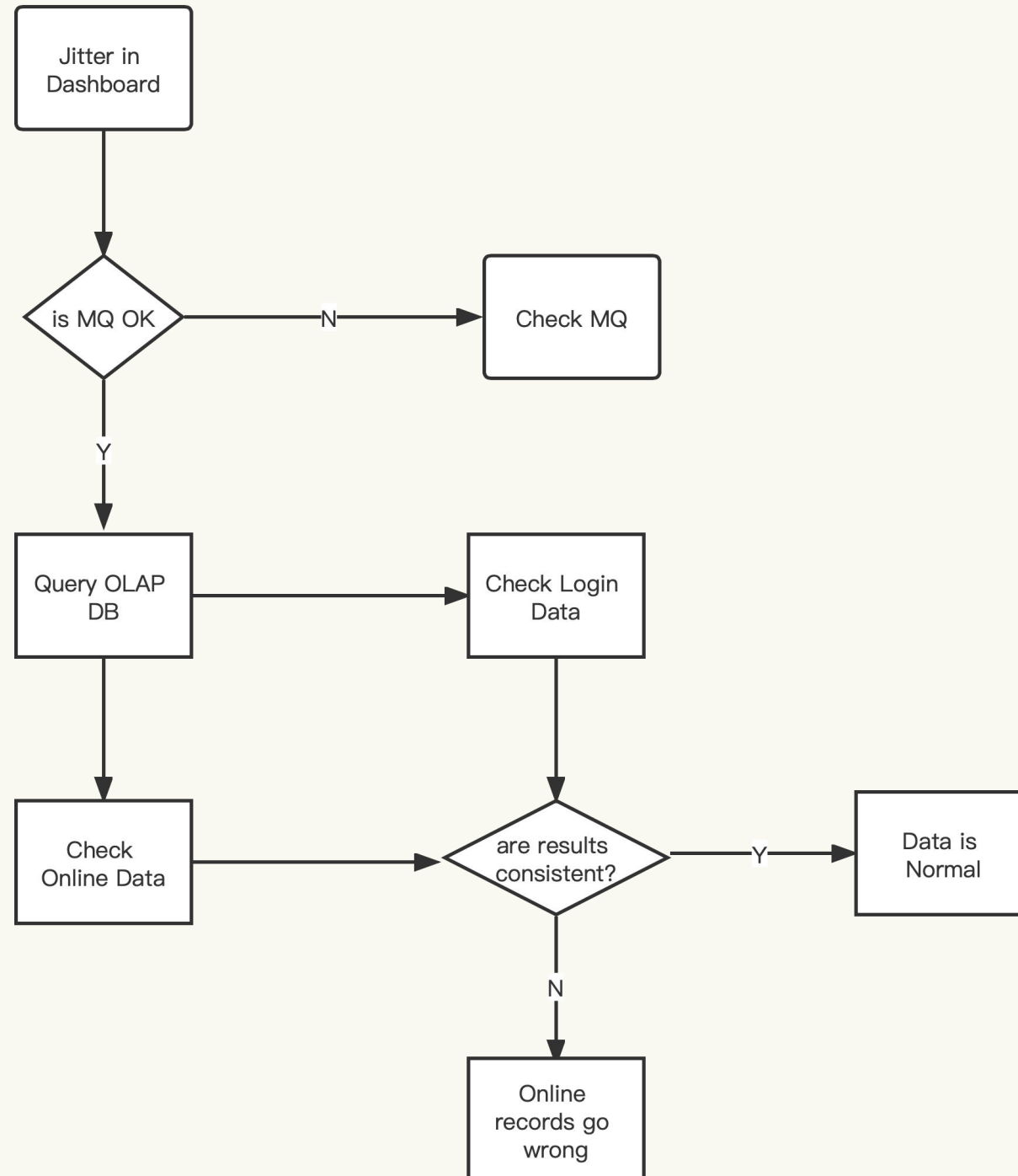
Special Case——Normal Production & Consumption

- Is it true that a large number of players have logged in during this period? **NO!**



Only less than 100
new login players
in server 166

Review the Special Case



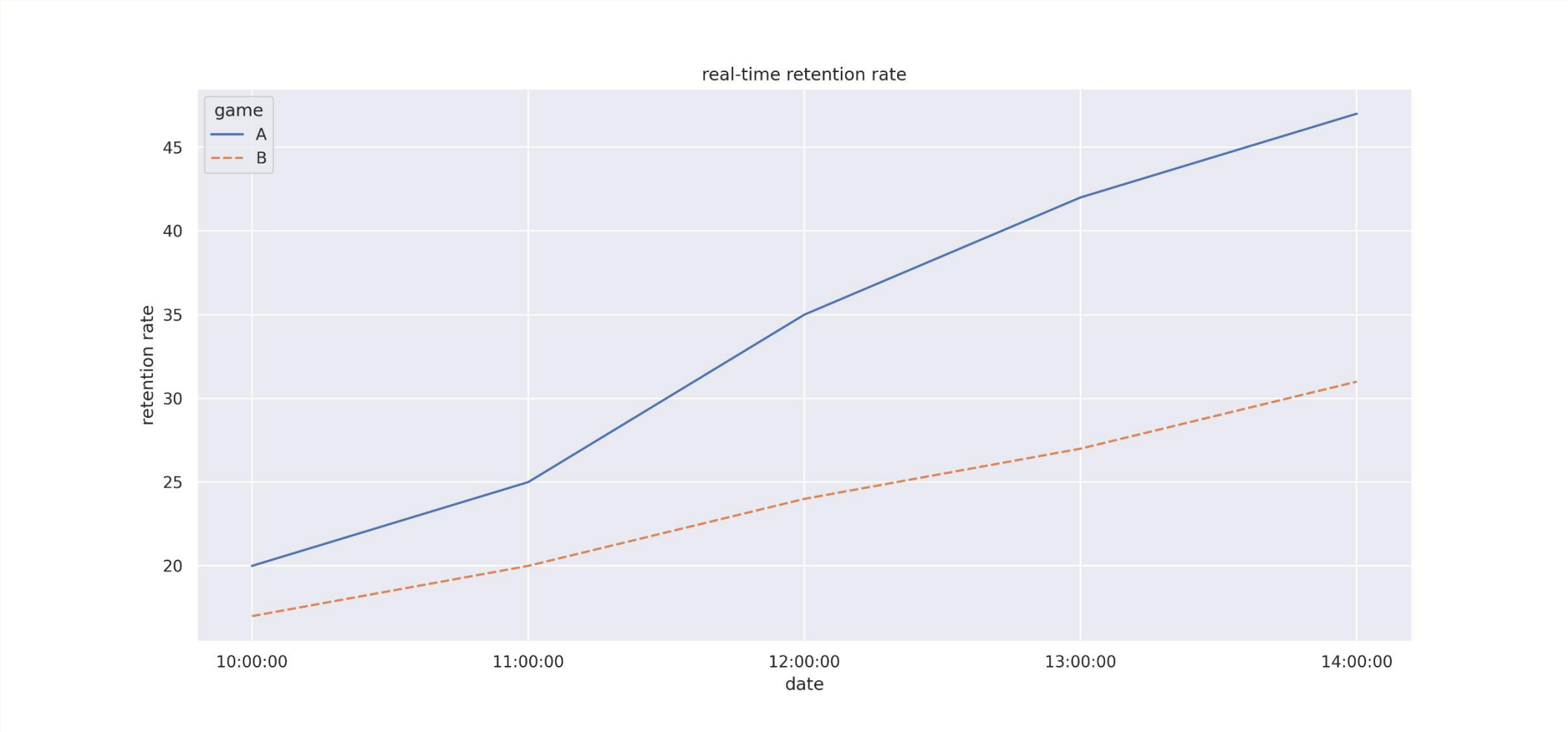
Summary for Special Case

- Data record itself may go wrong
- Errors in data record cannot be detected by rough message queue monitoring
- Subtle data anomalies require a real-time data analysis architecture

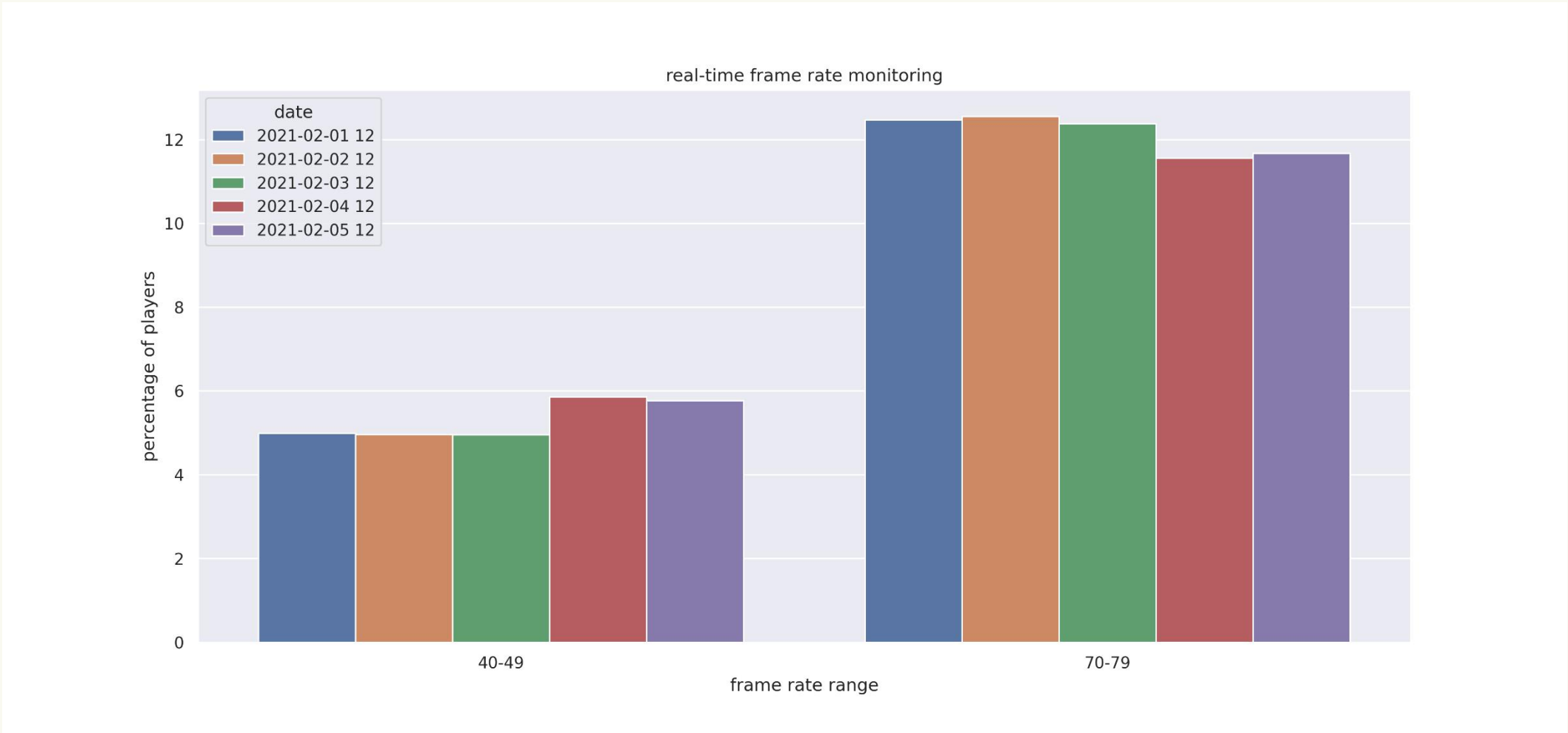
Two More Cases of Real-time Data Analysis Architecture

- Evaluate the Retention Rate of Different Games
- Monitor FPS of Different Versions of Game Clients

Evaluate the Real-time Retention Rate of Different Games



Monitor FPS of Different Versions of Game Clients



Takeaway

- Understand common data anomaly characterization
- Learn how to use message queue to discover upstream and downstream anomalies
- Understand the value of real-time analysis architecture in detailed monitoring

Thanks for your attention!