Alibaba Group
阿里巴巴集团

Capacity Planning

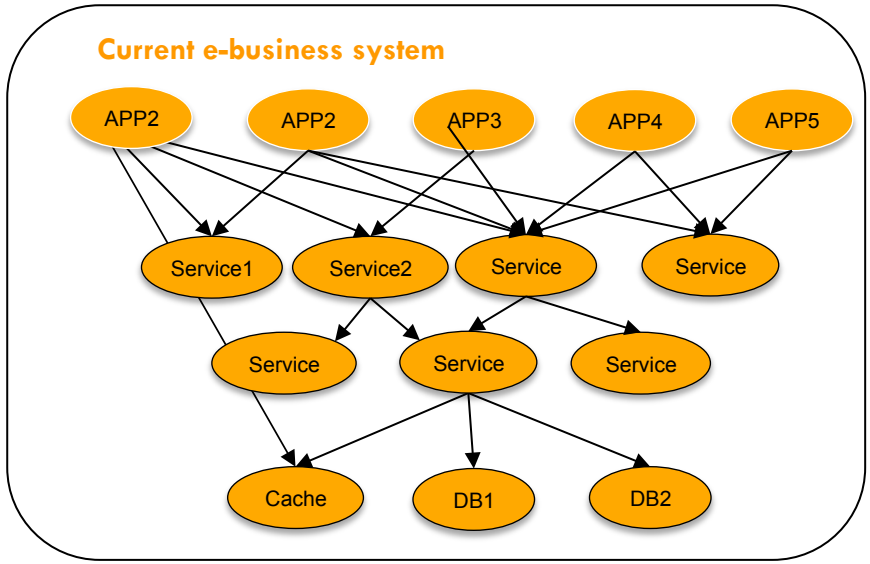
Alibaba High Availability Architecture Team
Jun Zhang, Julia Lin

Alibaba Group 阿里巴巴集团

Capacity Planning

Business Requirements

- Business growth
- New business
- Promotions



Questions

- When scalability is needed?
- How many machines?
- Correct ratio for each server?

Capacity planning is the process of determining the capacity needed by a complexity distributed system to meet the workload with guarantee on certain level of performance

Alibaba Group 阿里巴巴集团

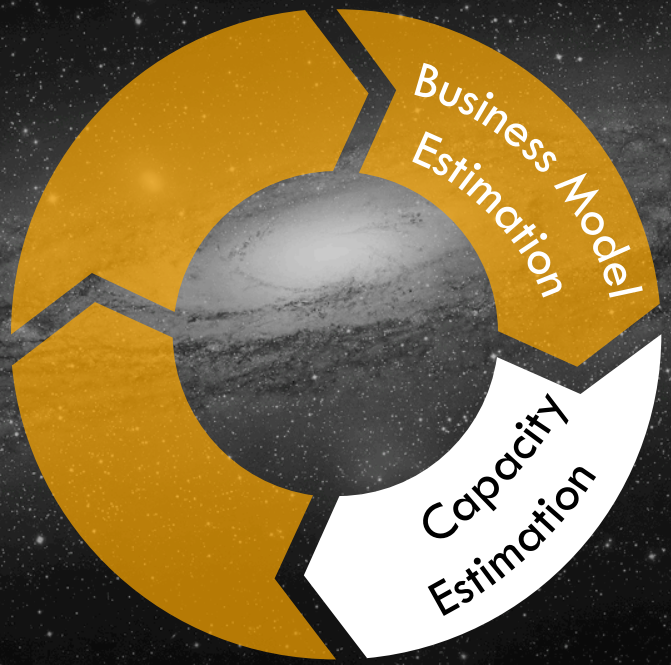
Capacity planning



- Traffic Model
- Historical Data
- Prediction Algorithm

Alibaba Group 阿里巴巴集团

Capacity planning



- Traffic Model
- Historical Data
- Prediction Algorithm

- Single Machine
- App Capacity

Alibaba Group 阿里巴巴集团

4 Ways of Stress test for single machine



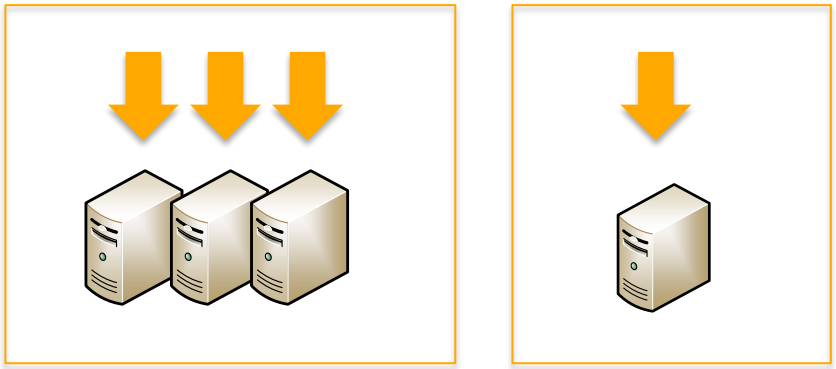
Simulation

Replication

Redirect

Load Balance

Simulation



Production environment

Test environment

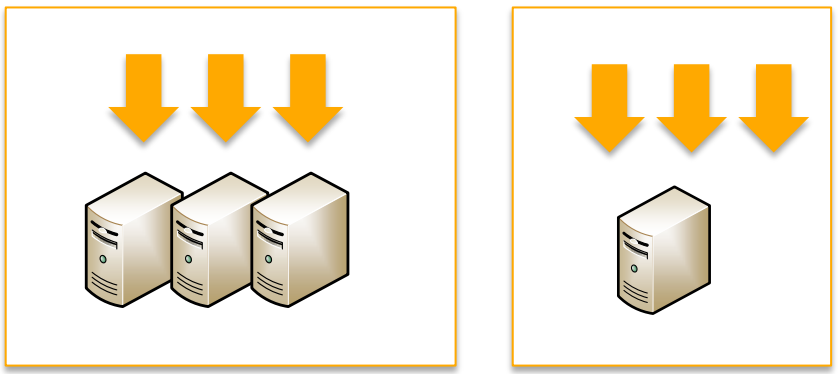
Advantage

- Easy implementation
- Fits for new applications

Disadvantage

- Not real production traffic
- Additional dirty data

Replication



Production environment Test environment

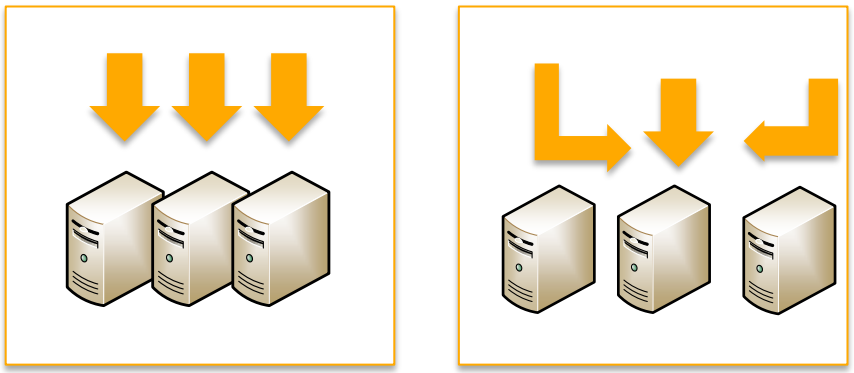
Advantage

- Close to the production environment
- For application with few traffic, we can multiple the flow.

Disadvantage

- Need additional machine
- Dirty data as well

Redirect Traffic



Production environment

Production environment

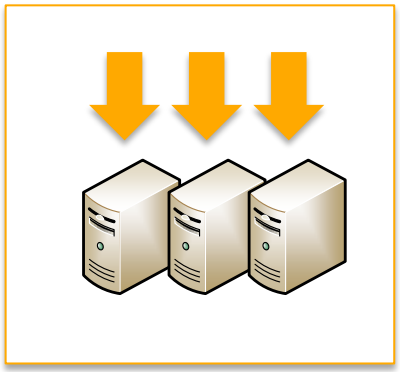
Advantage

- All traffic from production, no dirty data

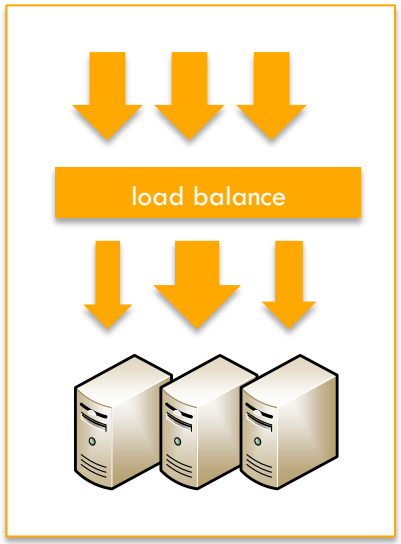
Disadvantage

- Not work for new application, or application with limited traffic

Load Balance



Production environment



Production environment

Advantage

- All traffic from production, no dirty data

Disadvantage

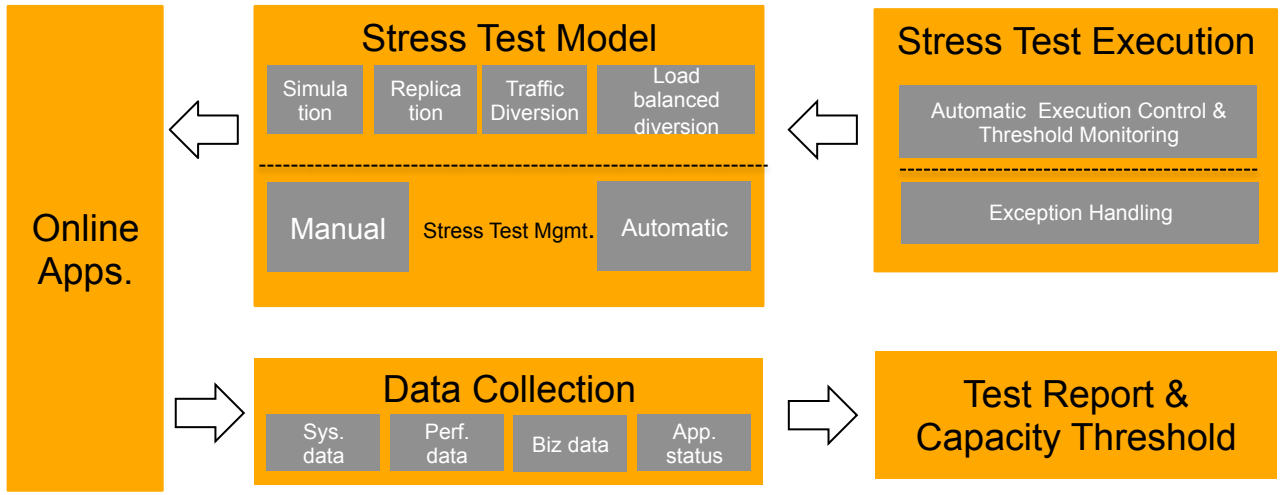
- Not work for new application, or application with limited traffic

Alibaba Group 阿里巴巴集团

Architecture of Stress Testing Platform for single machine

- Automatic start/stop execution, automated stress testing
- Automatic report and capacity threshold generation

- Routine process: **5000+** stress test/month
- Criteria for capacity expansion/ reduction decisions





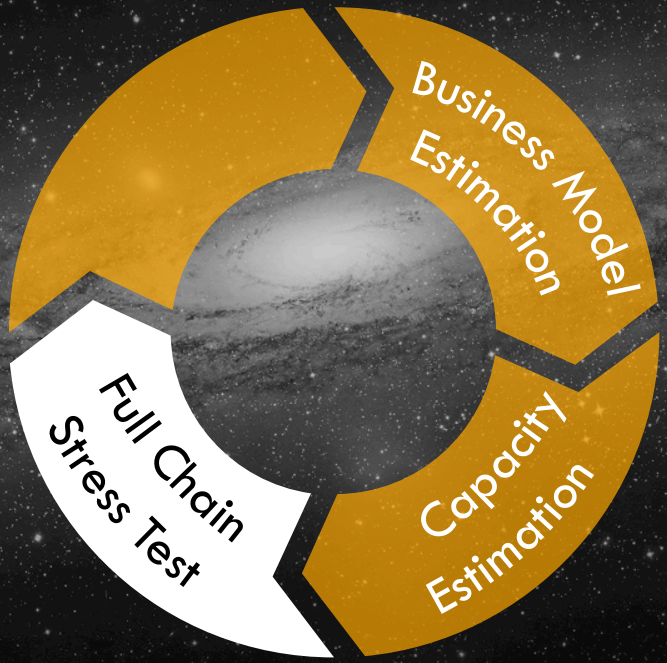
Machine Needed

=

Business estimation / Single machine capacity + Buffer

Alibaba Group 阿里巴巴集团

Capacity planning



- Traffic Model
- Historical Data
- Prediction Algorithm

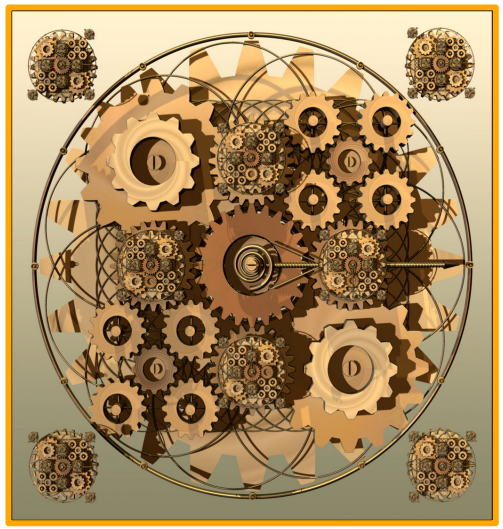
- Single Machine
- App Capacity

- Verify
- Fine Tune

Alibaba Group 阿里巴巴集团

Why Full Chain Stress Test

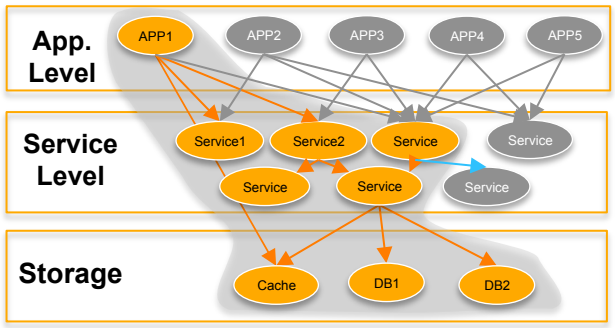




“Single point” approach is totally different from “scenarios”

Full Chain Stress Test

By simulating the same scenario as Big Promotion, connect all scenarios for verification



Goal

- Verify the capacity and performance
- Find out bottle necks
- Verify and finetune the rationality of resource allocation for application
- Verify and finetune the network device and cluster capacity
- Drill for 11/11

Difficulties in Full Chain Stress Test



Super high volume

- 10,000,000 requests/sec

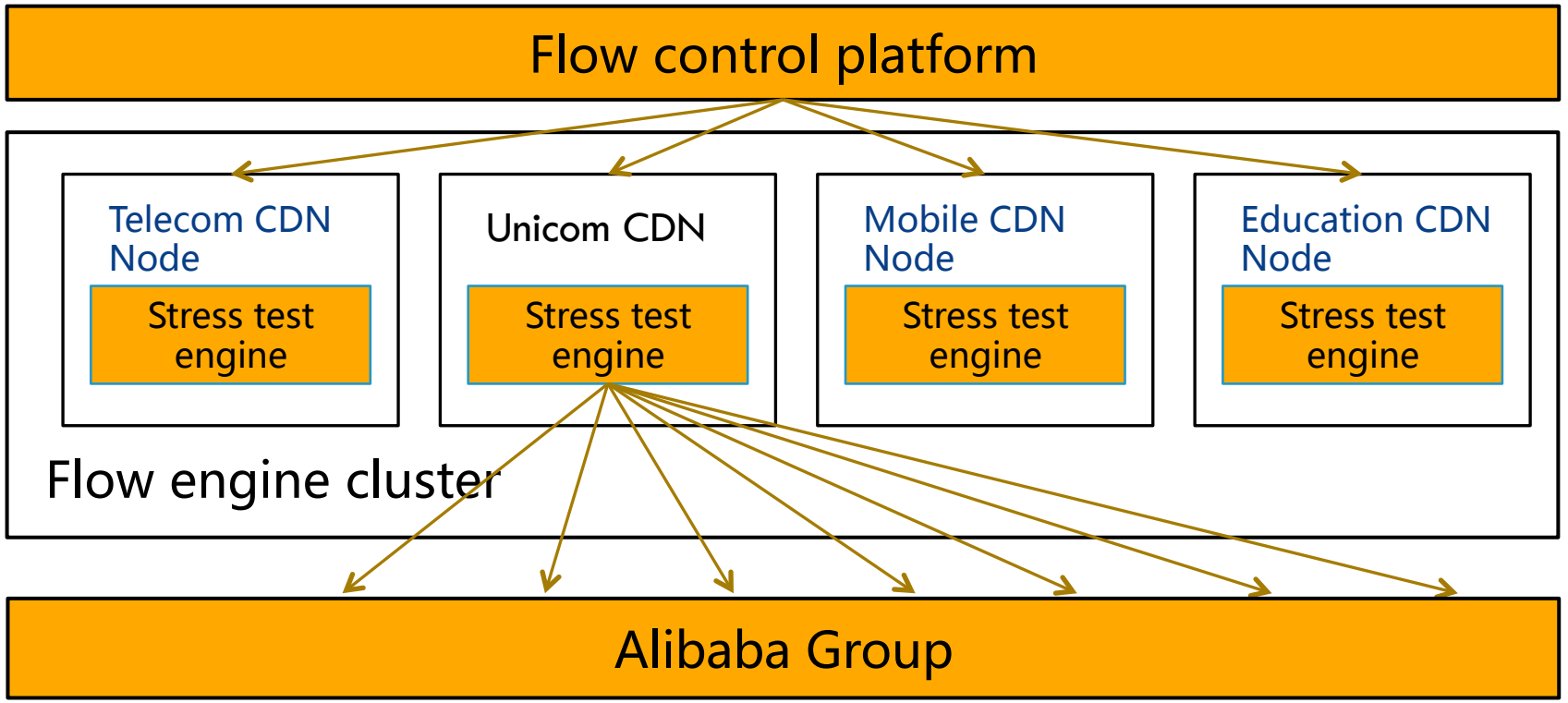
Close to real user behavior

- Discount goods selection in big promotion

No dirty data pollution

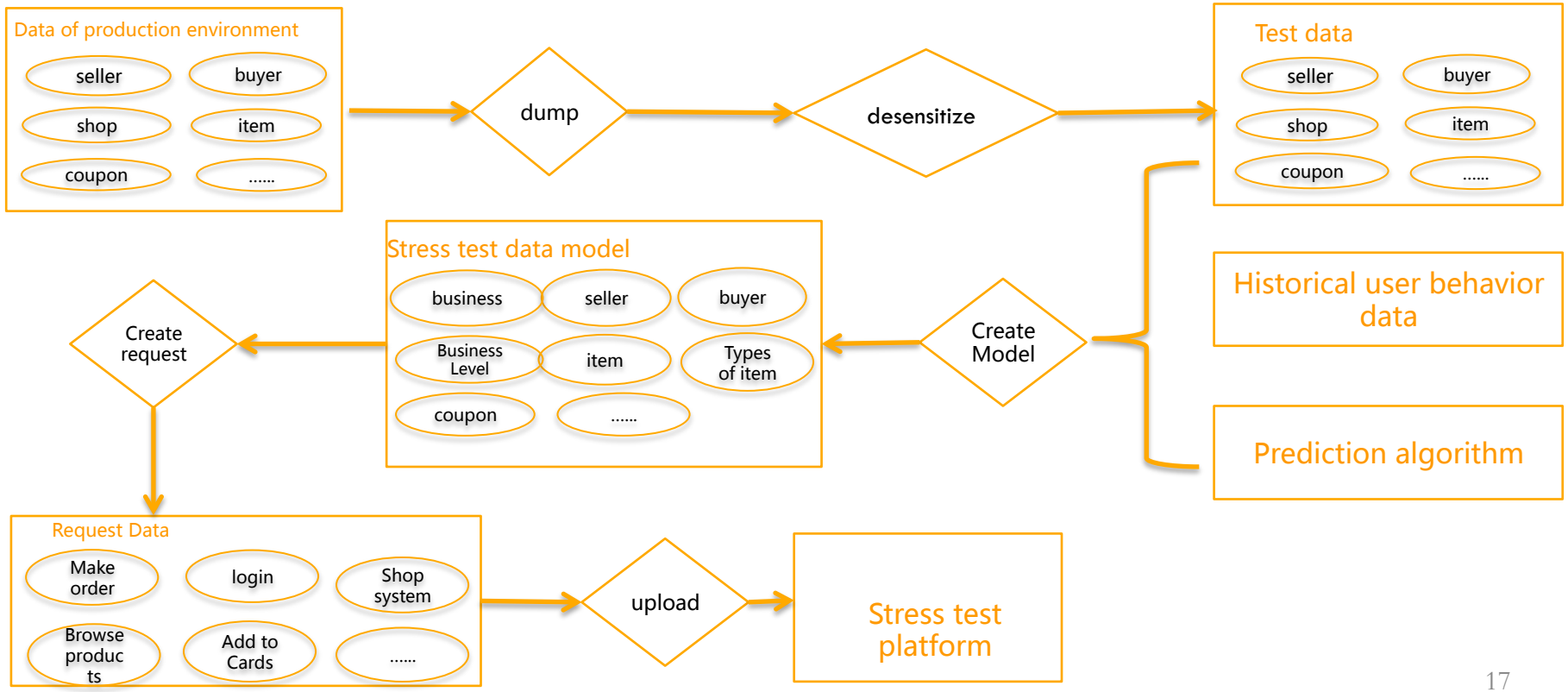
Alibaba Group
阿里巴巴集团

Simulate same level traffic flow



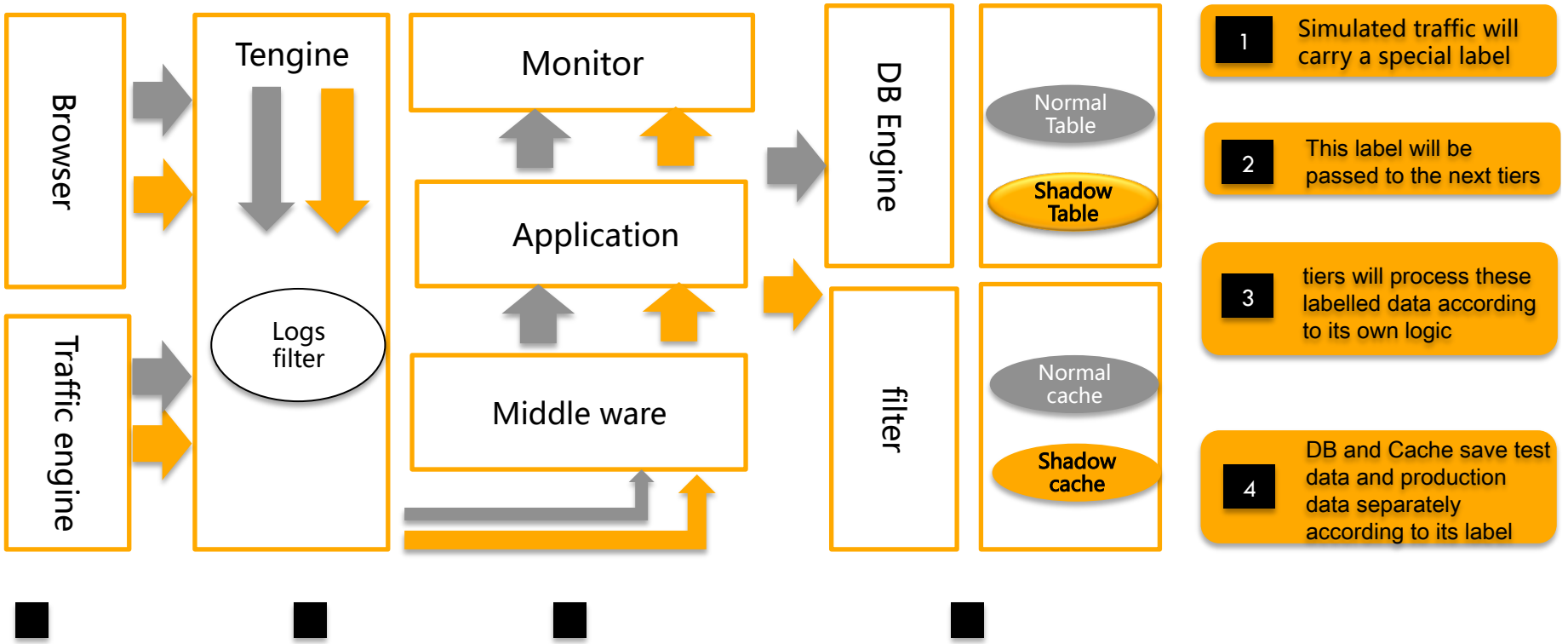
Alibaba Group 阿里巴巴集团

Create data model



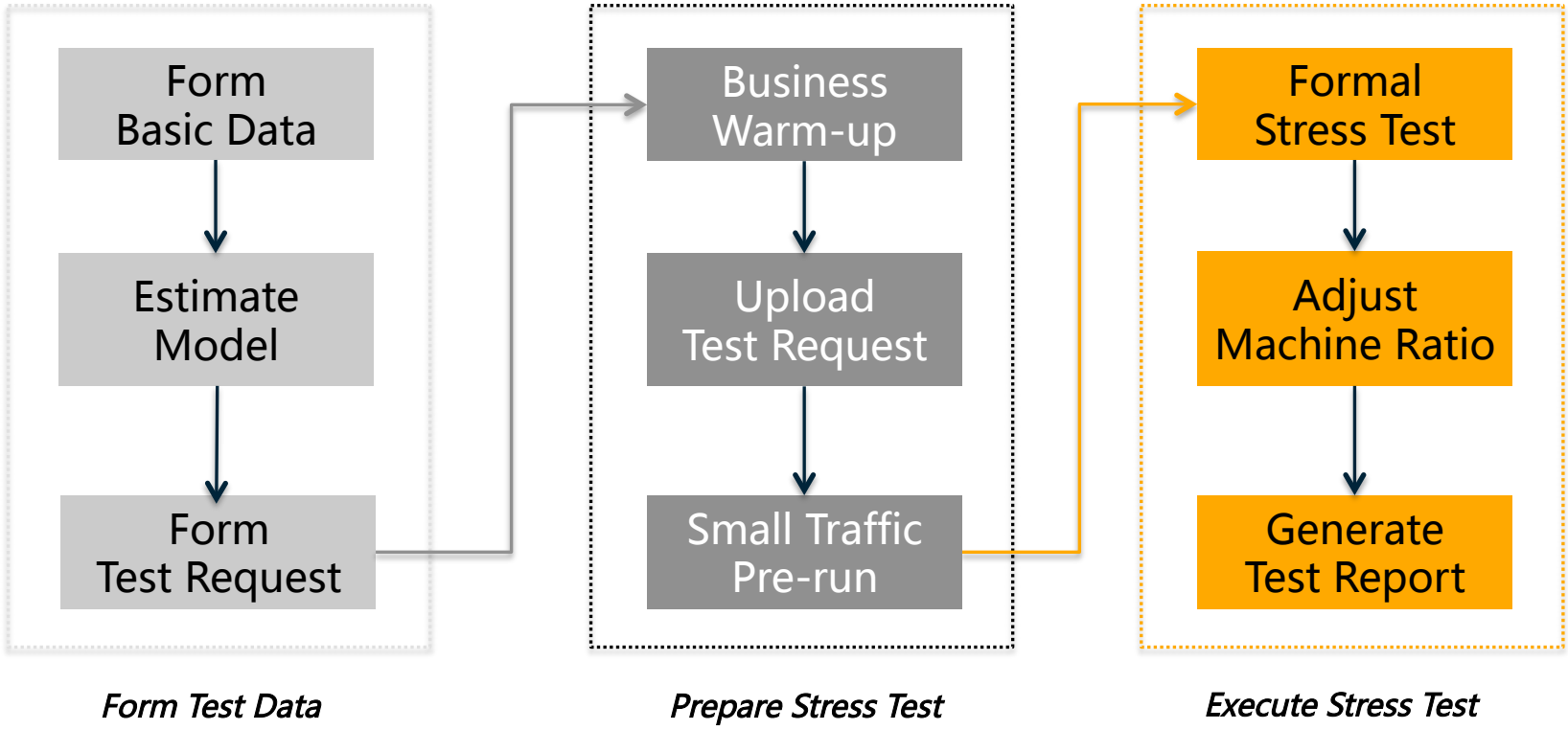
Alibaba Group 阿里巴巴集团

Isolate test data

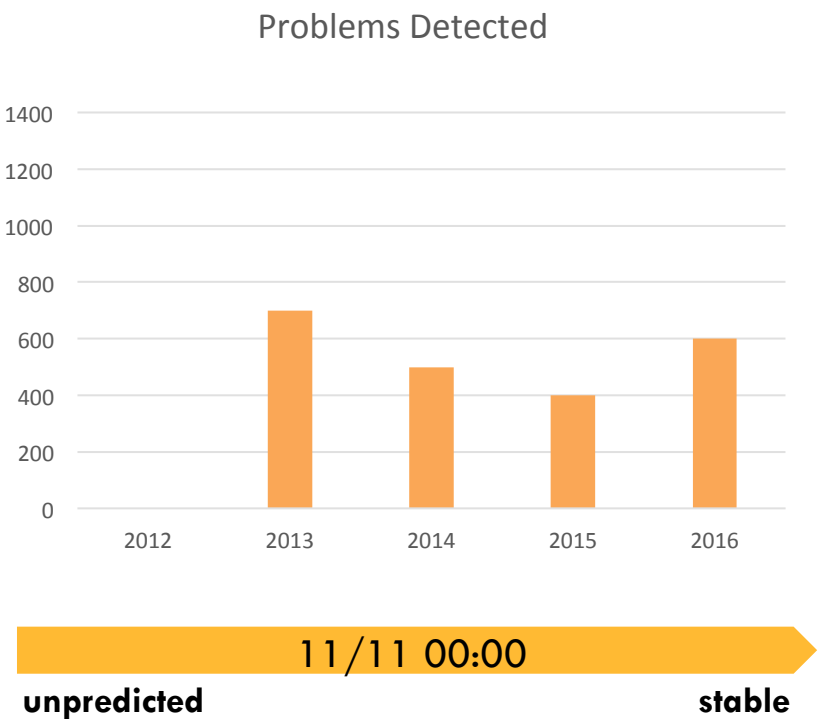


Alibaba Group 阿里巴巴集团

Key Process



Outcome



Launch in 2013

- Mar. 8 Promotion
- Jun.18 Promotion
- Sept. 9 Promotion
- Nov. 11 Promotion (5 rounds)

Progress between 2013 to 2015

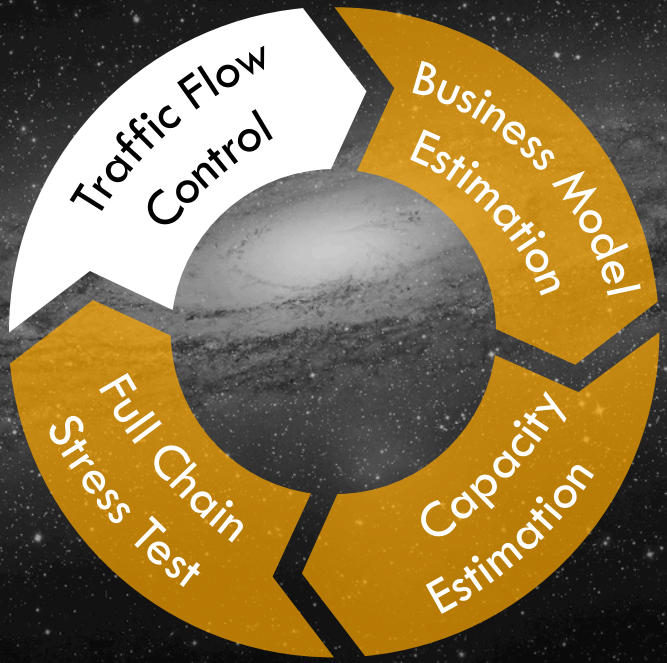
Achievements in 2016

- 4000+ stress tests made
- Embracing various product lines

Alibaba Group 阿里巴巴集团

Capacity planning

- Flow Control
- Fast Recovery

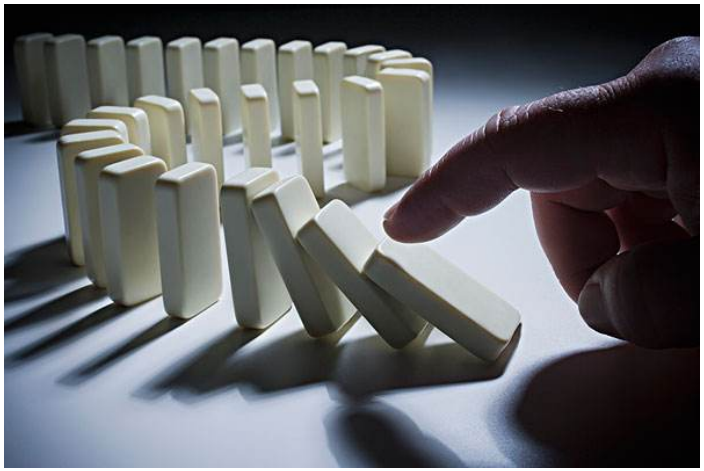


- Traffic Model
- Historical Data
- Prediction Algorithm

- Verify
- Fine Tune

- Single Machine
- App Capacity

Why flow control is necessary



Stuck in Queue

- Lower down server performance,
- Increase response time,
- Affect application and user experience.

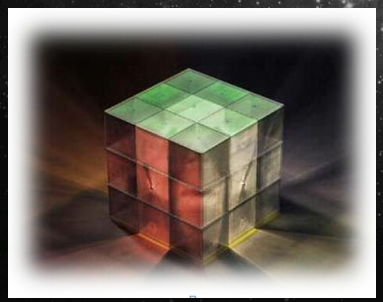
Chain Reaction

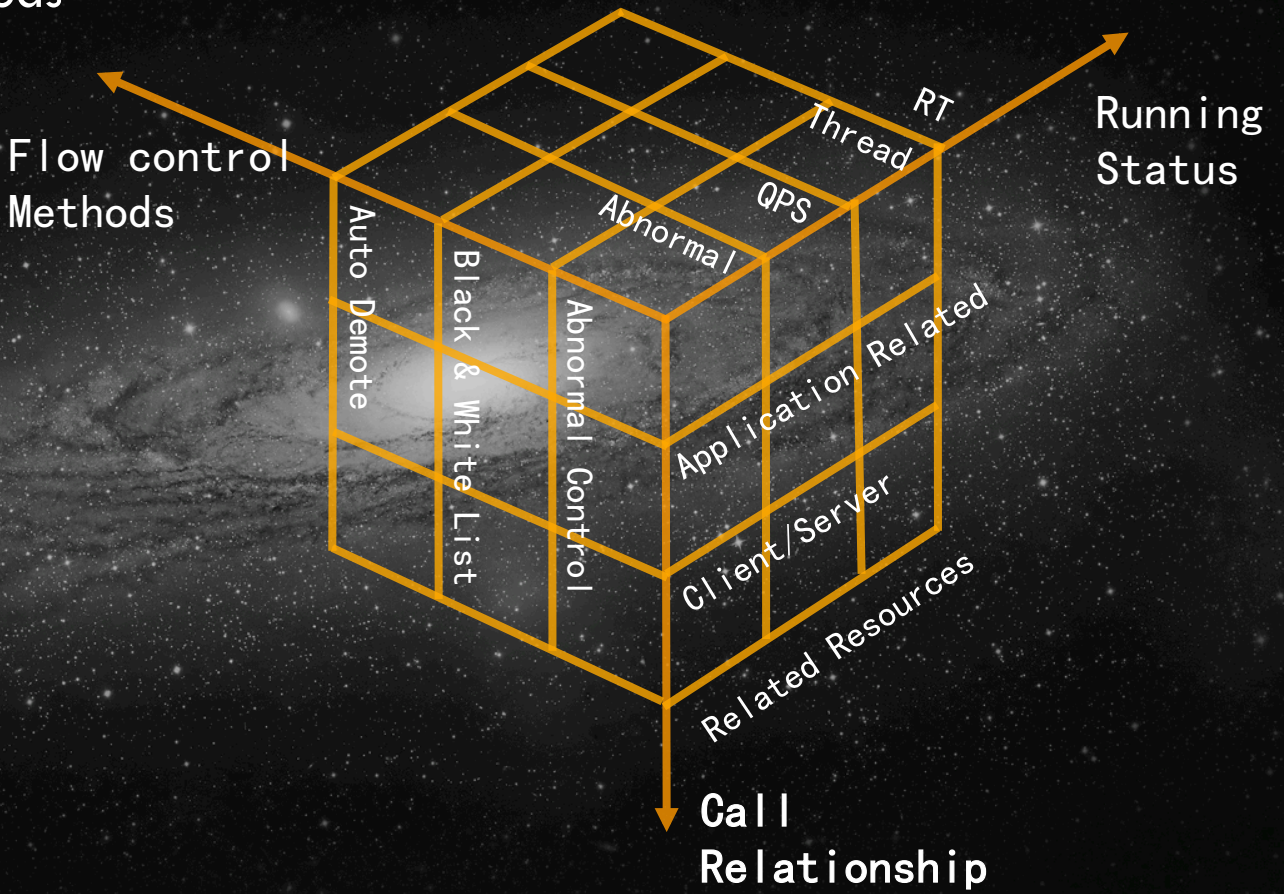
Flow Control

To restrict overloaded traffic.

Alibaba Group 阿里巴巴集团

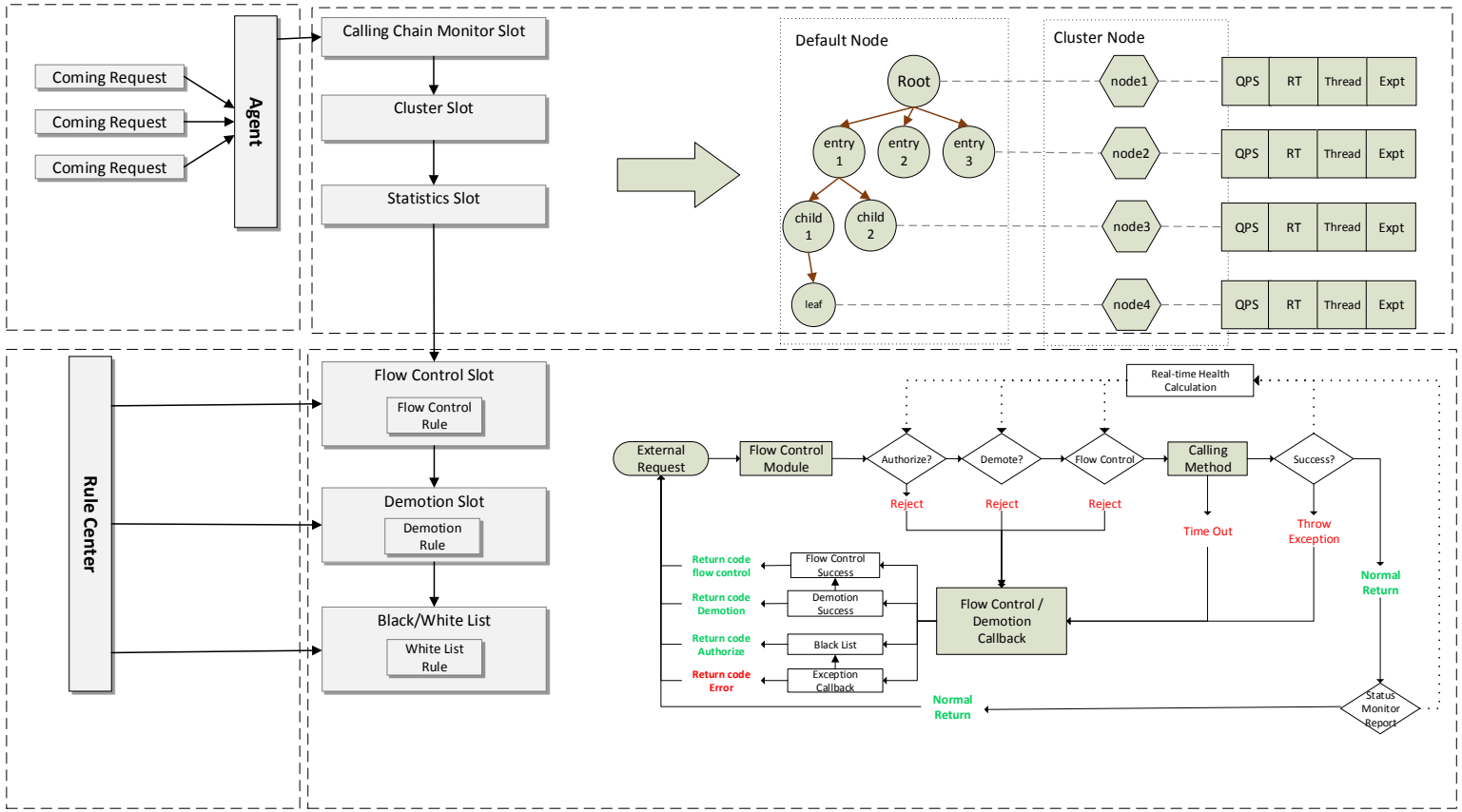
Flow Control Methods



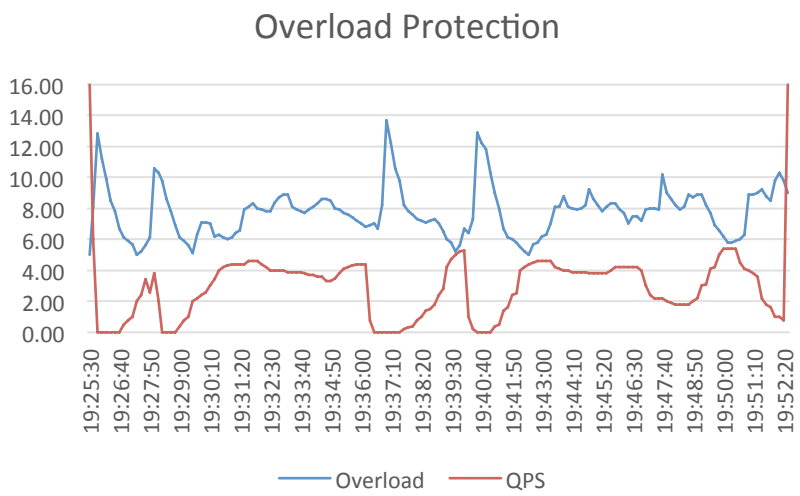


Alibaba Group 阿里巴巴集团

Flow Control Principle



System overload protection (traditional)



Control Principle

- Prohibit/reduce incoming volume when system overload exceeds threshold.
- Allow/recover incoming volume when overload is healthy

Use hard bar to control

Problems Related (affecting user experience)

- QPS fluctuation
- Slow recovery

System overload protection (New)



Basic Calculation

- Entrance size = volume * RT
- Request = constants * Load * RT

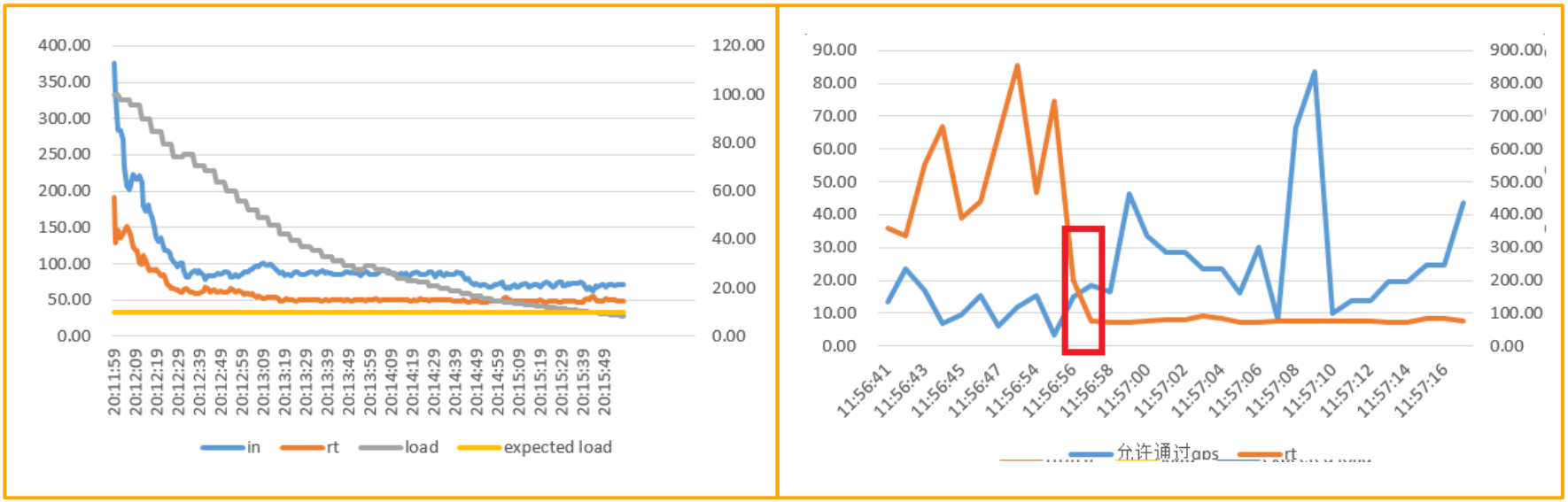
Control Principle

- Prohibit/reduce incoming volume when system overload exceeds threshold.
- Allow/recover incoming volume when overload is healthy

Use *Dynamic threshold* to control

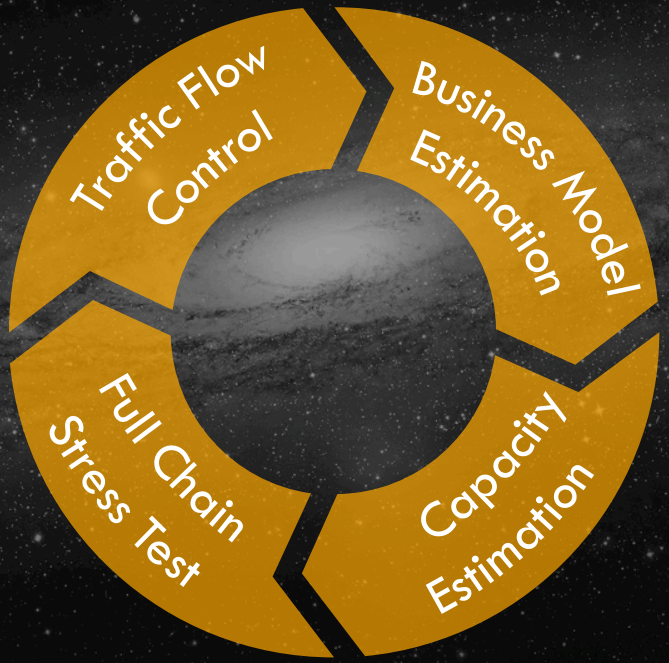
Alibaba Group
阿里巴巴集团

Stable Result



Alibaba Group 阿里巴巴集团

Consolidated Plan




Alibaba Group
阿里巴巴集团
Thank you!