Next Generation of DevOps
AIOps in Practice @Baidu

Xianping Qu & Jingjing Ha

Baidu 百度

About Baidu
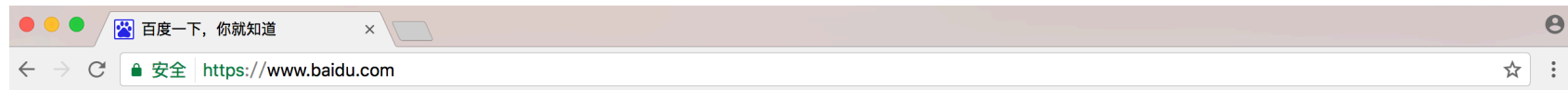
Agenda

- History of Baidu SRE team
- Next generation of DevOps – AIOps
- Best practice based on AIOps
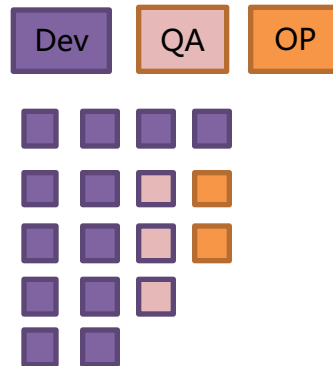- Future

Bai度百度

Tools (2007-2009)

- DevOps
  - Deployment
  - Monitoring
  - Budgeting
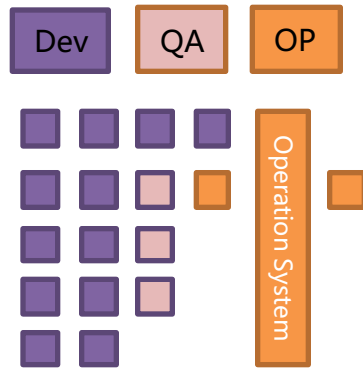 - Consulting
  - …
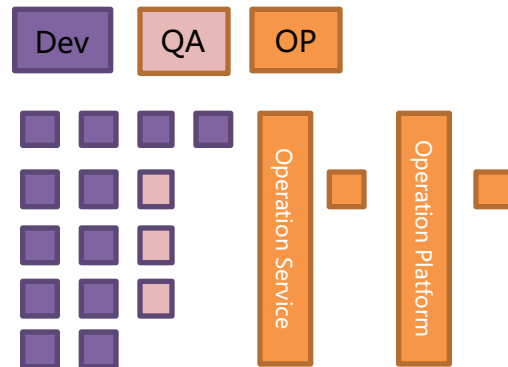- Problems
  - Human labor

Systems (2009-2012)

- Building operation systems
  - Service management system
  - Monitoring system
  - Deployment system
  - Traffic scheduling system
  - Naming service
  - …
- Problems
  - Human labor（GUI, configure）

Platforms (2012-2014)

- Building operation platforms
  - API
  - Configurable
  - Executable
  - …
- Problems
  - Reusability
  - Scalability

Standardization (2014-)

- Building operation standards
  - Unified language
  - Unified method
  - Unified solution
  - …
- Problems
  - Need a brain

齐 楚 燕 韩 赵 魏 秦

马 （小篆）

马 （隶书）

Services, IDCs
Clusters, Servers
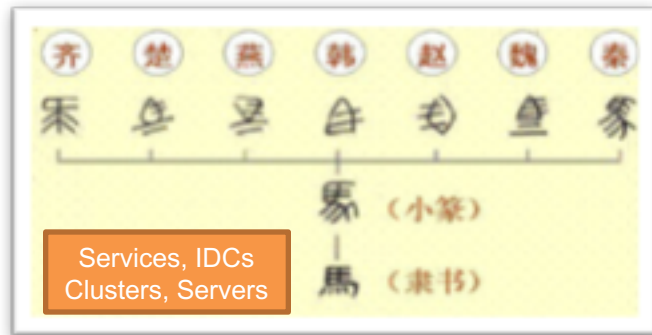
AIOps (2014-)

- Intelligent Operation Platforms
  - Development framework
  - Big data
  - Algorithm
    - Data mining, machine learning…

Development framework
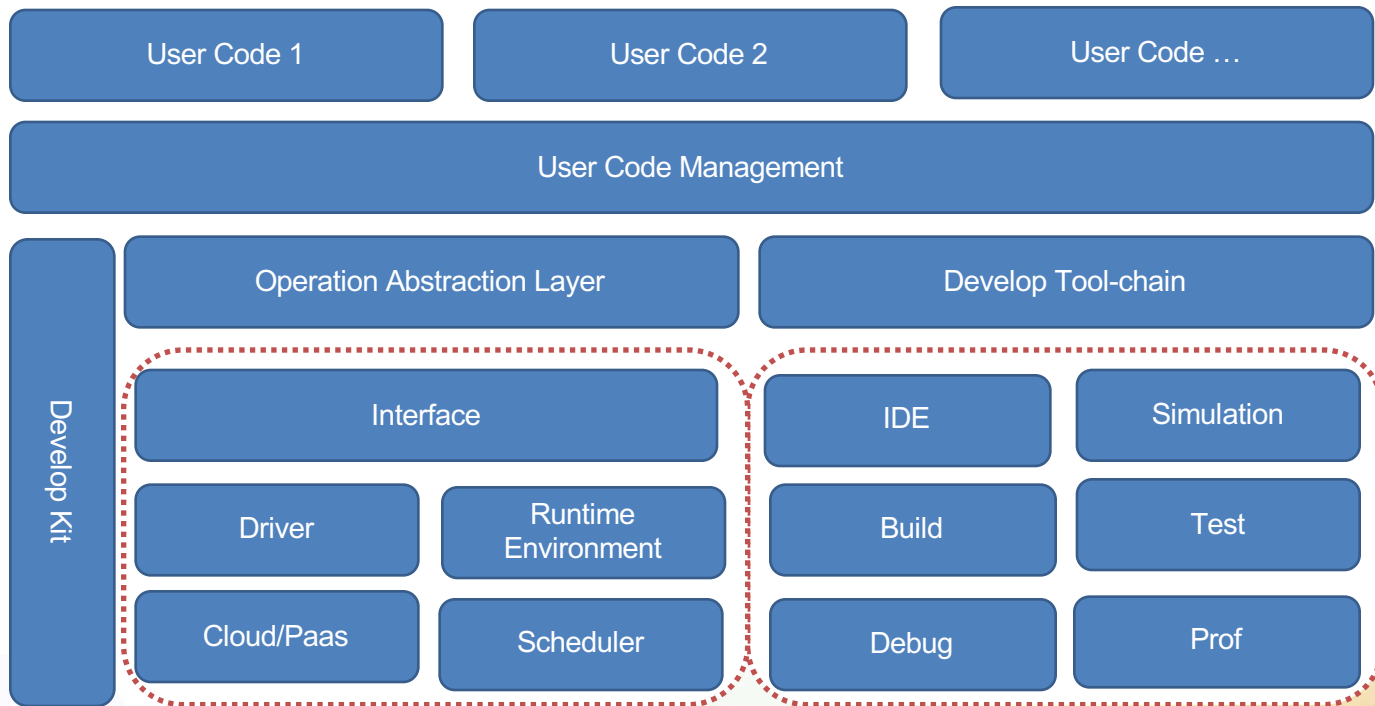
Operation Knowledge Database

Solution

Anomaly detection

Traffic scheduling

Root cause analysis

Trend forecasting

Other data mining & machine learning algorithms

Solution Development

Deployment Management

Incident Management

…

Ops Algorithm Development

Ops Platform Development

Operation Development Framework（SDK, RE）

Operation Knowledge Database

General Components and Tools（transmission、 storage、 scheduling…）

Best practice based on AIOps

- Incident Management
  - Single cluster stop-loss by traffic shifting

- Deploy Management
  - Unattended deployment with automated checker

- Consulting
  - ChatBots do Consultation

Baidu百度

When will a failure occur?

- Infrastructure issue

- Program defects
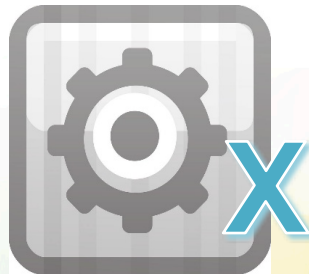
- Change exception

- Dependent service unavailable

How to stop loss?

- Limited Failure in one cluster

  – Deployment isolation

  – Dependency decoupling

  – Reduce global risk

- Capacity redundancy

  – Availability and cost trade-off

  – N+M redundancy

  – Service degradation

When single cluster fails do perform traffic shifting
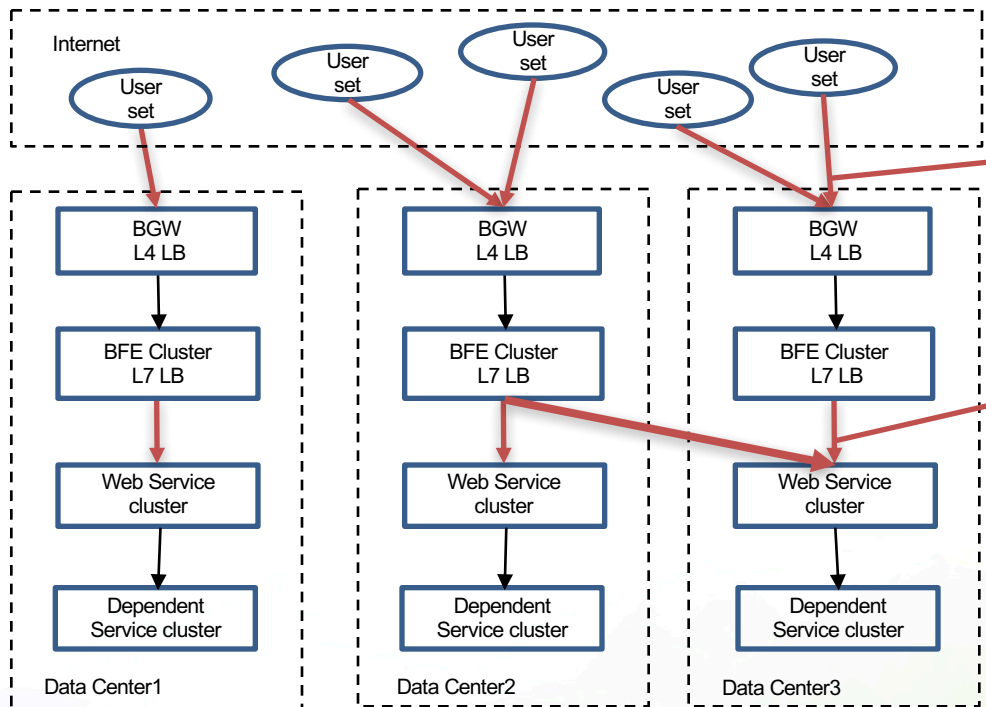
Bai百度

Two layer traffic shifting @Baidu

Internet

User set

User set

User set

User set

User set

User set

shift traffic between user set and the edge node

BGW
L4 LB

BGW
L4 LB

BGW
L4 LB

BFE Cluster
L7 LB

BFE Cluster
L7 LB

BFE Cluster
L7 LB

shift traffic between front end and back end

Web Service cluster

Web Service cluster

Web Service cluster

Dependent Service cluster

Dependent Service cluster

Dependent Service cluster

Data Center1

Data Center2

Data Center3

BGW: Baidu Gate Way, layer-4 load balancer

BFE: Baidu Front End, layer-7 load balancer

Baidu 百度

Two layer traffic shifting @Baidu

- shift traffic between user set and edge node

  – 10 minute to shift 80% traffic to the healthy edge node because of DNS caching in the client side and ISP side

- shift traffic between front end and back end

  – 10 second to shift 100% traffic to the healthy backend by changing BFE's routing configuration

Bai**du**百度

Shift traffic between front end and back end

Concerns:
- Service Capacity
- Intranet bandwidth

Scenarios：
- Web service cluster
- Dependent service cluster
- Internal network switch

Internet

User set

User set

User set

User set

User set

BGW L4 LB

BGW L4 LB

BGW L4 LB

BFE Cluster L7 LB

BFE Cluster L7 LB

BFE Cluster L7 LB

Web Service cluster

Web Service cluster

Web Service cluster

Dependent Service cluster

Dependent Service cluster

Dependent Service cluster

Data Center1

Data Center2

Data Center3
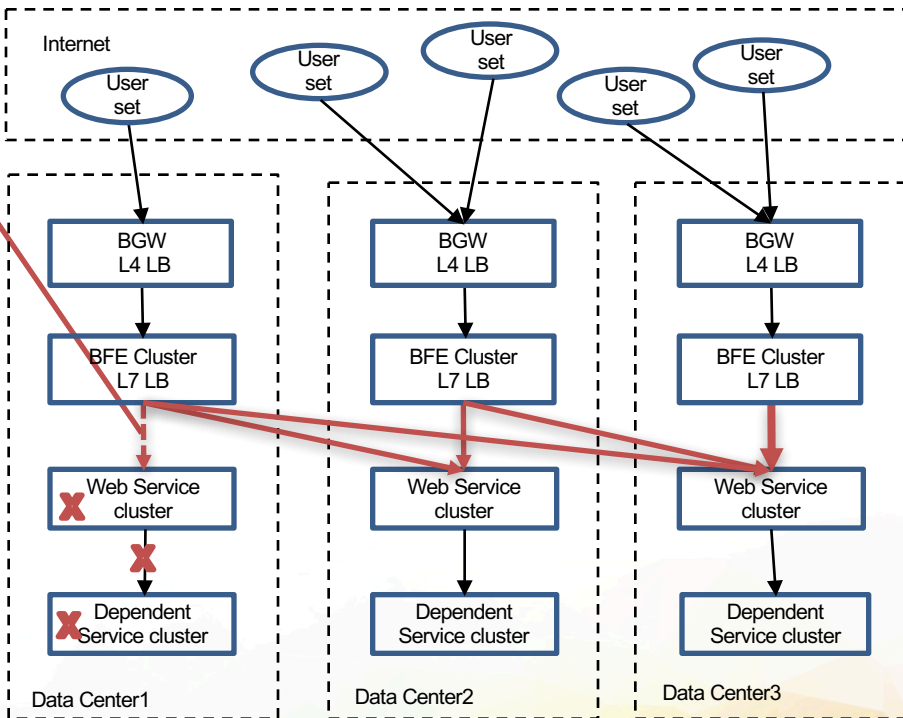
Shift traffic between user set and the edge node

Concerns:
- Bandwidth
- BGW/BFE Capacity
- Delay

Scenarios :
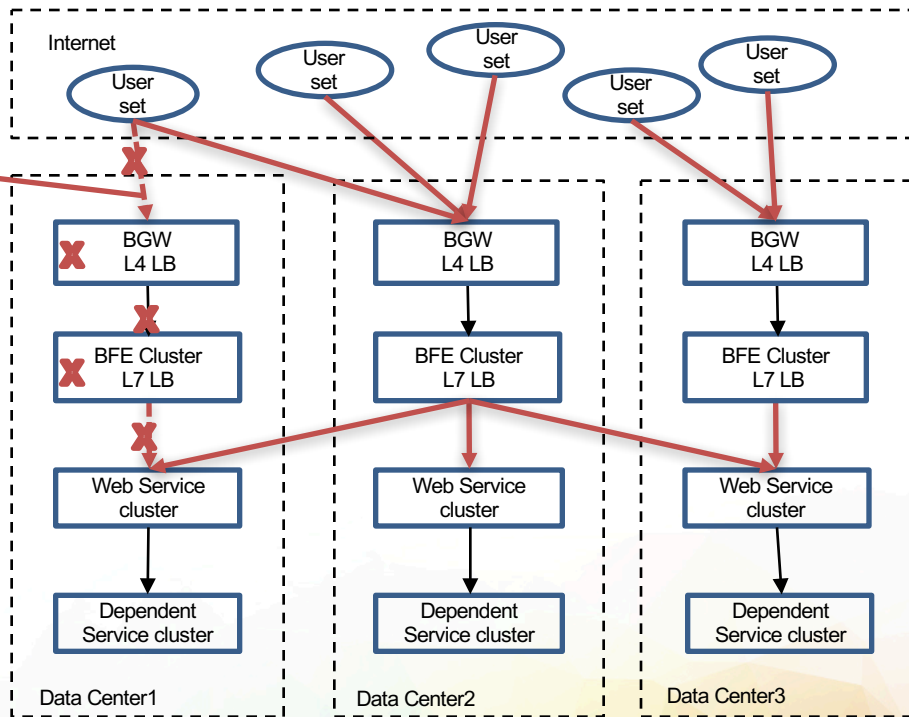- BGW & BFE
- External network switch

Internet

User set
User set
User set
User set
User set

BGW L4 LB
BGW L4 LB
BGW L4 LB

BFE Cluster L7 LB
BFE Cluster L7 LB
BFE Cluster L7 LB

Web Service cluster
Web Service cluster
Web Service cluster

Dependent Service cluster
Dependent Service cluster
Dependent Service cluster

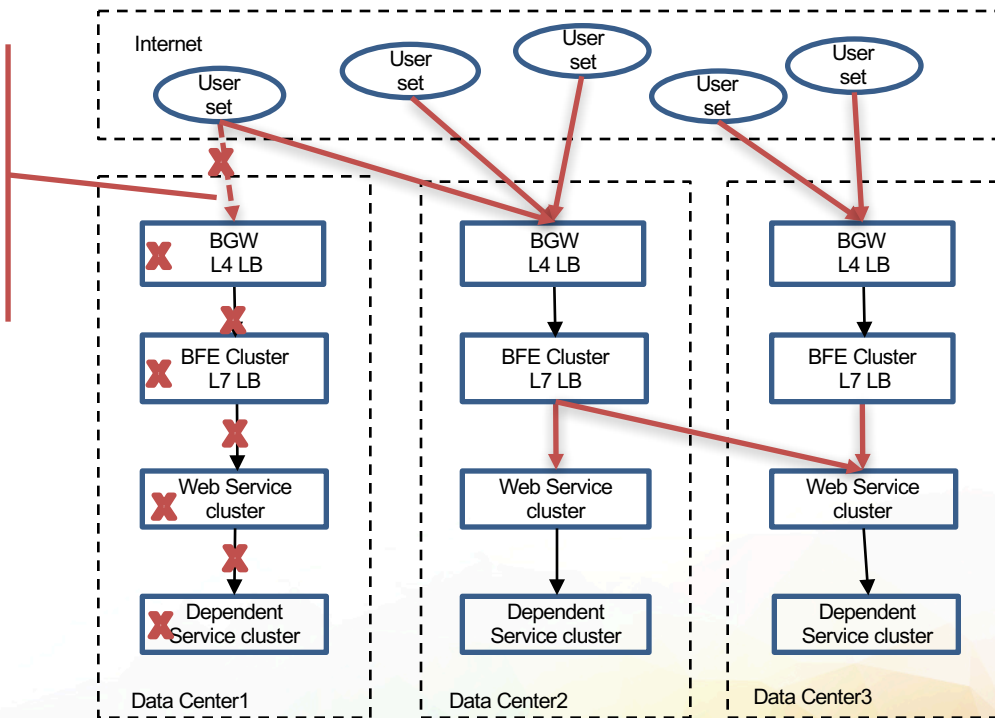Data Center1
Data Center2
Data Center3

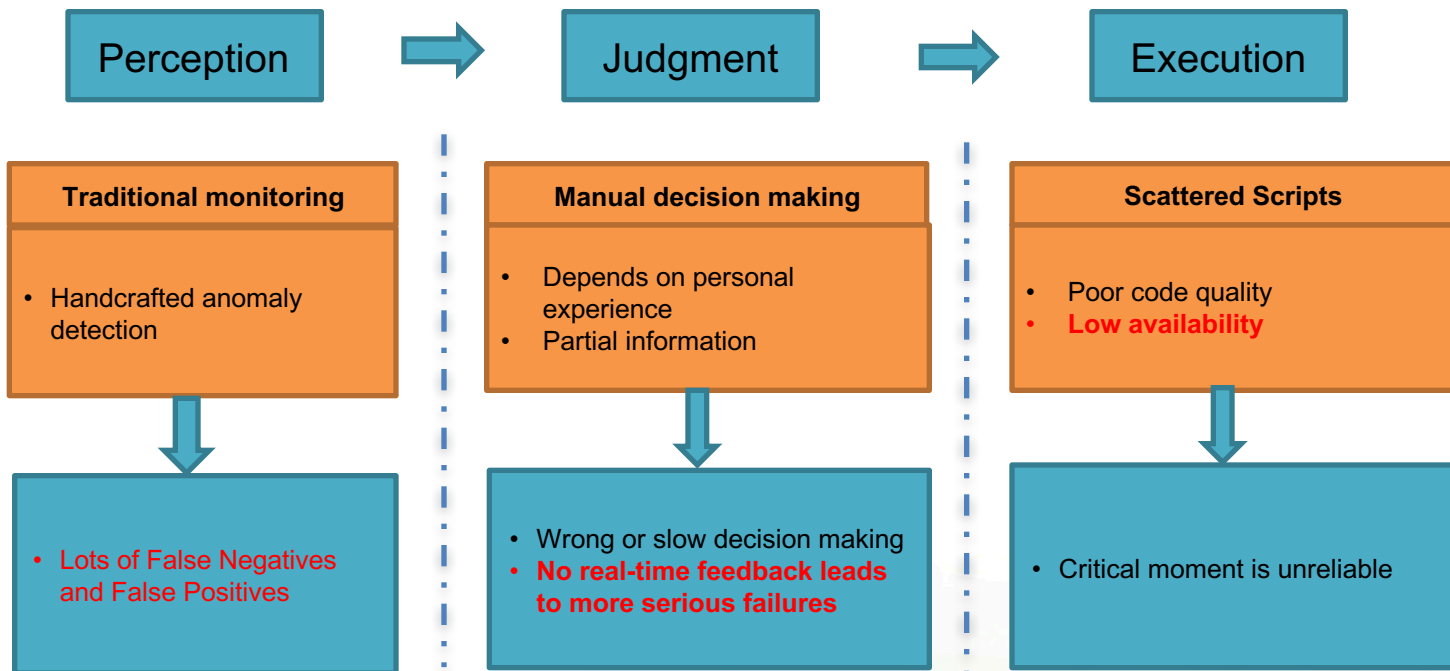# Shift traffic between user set and the back end

Concerns:
- Bandwidth
- BGW/BFE Capacity
- Delay
- Service Capacity
- Intranet bandwidth

Scenarios::
- Entire data center failure

Single cluster stop-loss before AIOPs

Perception	→	Judgment	→	Execution

Traditional monitoring

- Handcrafted anomaly detection

- Lots of False Negatives and False Positives

Manual decision making

- Depends on personal experience
- Partial information

- Wrong or slow decision making
- No real-time feedback leads to more serious failures

Scattered Scripts

- Poor code quality
- Low availability

- Critical moment is unreliable

Bai<u>du</u>百度

Single cluster stop-loss after AIOPs

Perception	→	Judgment	→	Execution

Intelligent monitoring

- Intelligent anomaly detection
- Abnormal event stored in **Operation Knowledge Database**

- High precision and recall

Automated decision making

- Rely on **Algorithm Platform**
- Global information

- Accurate and quick decision making
- Feedback control

Standard framework

- **Development framework**
- Deployment framework

- High development efficiency
- High availability

The architecture of single cluster stop-loss

Unattended deployment with automated checker

Person

- begin
- confirm
- confirm
- confirm
- confirm
- end

Robot

- pre-check
- pipe line
  - deploy
  - check
  - ...
  - deploy
  - check
- post-check
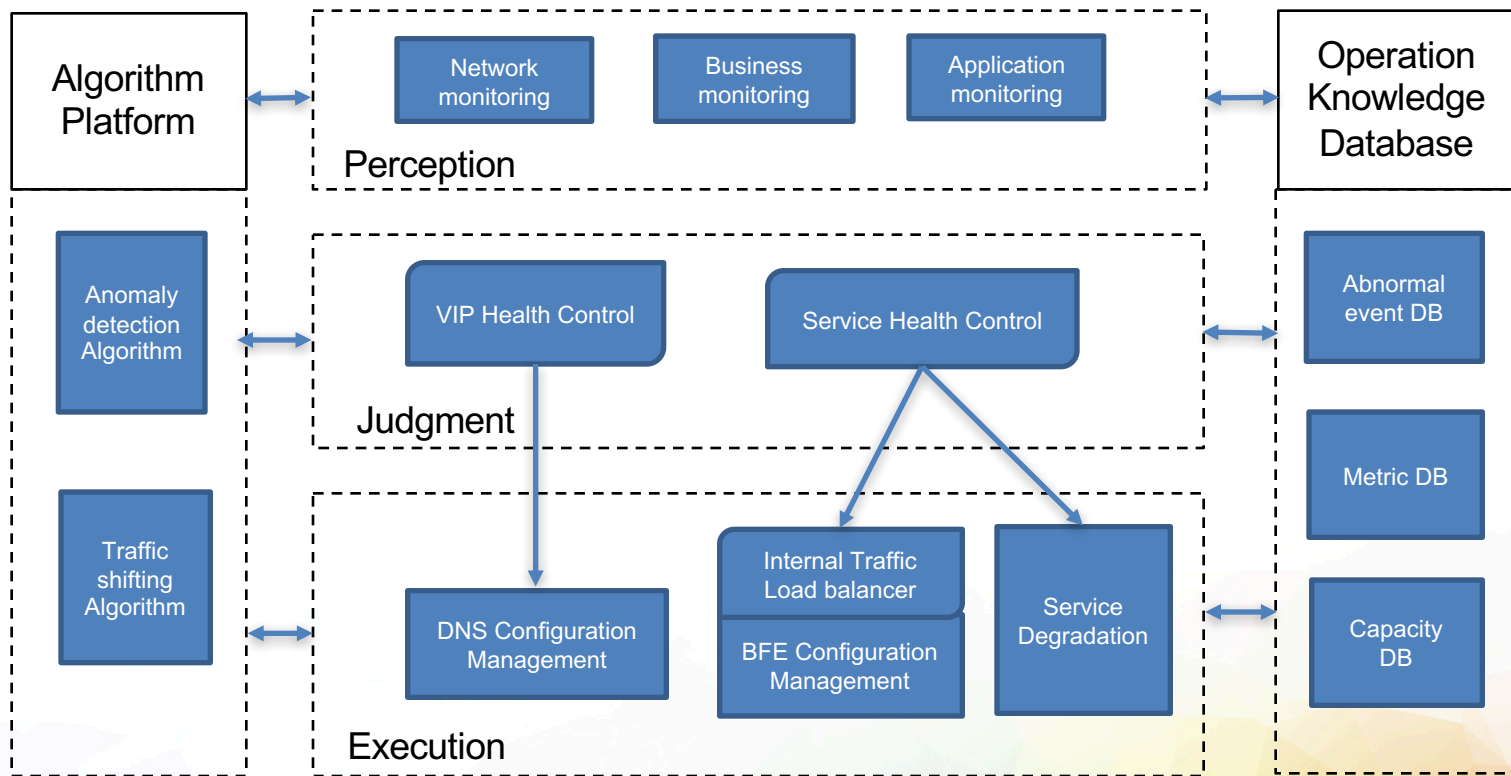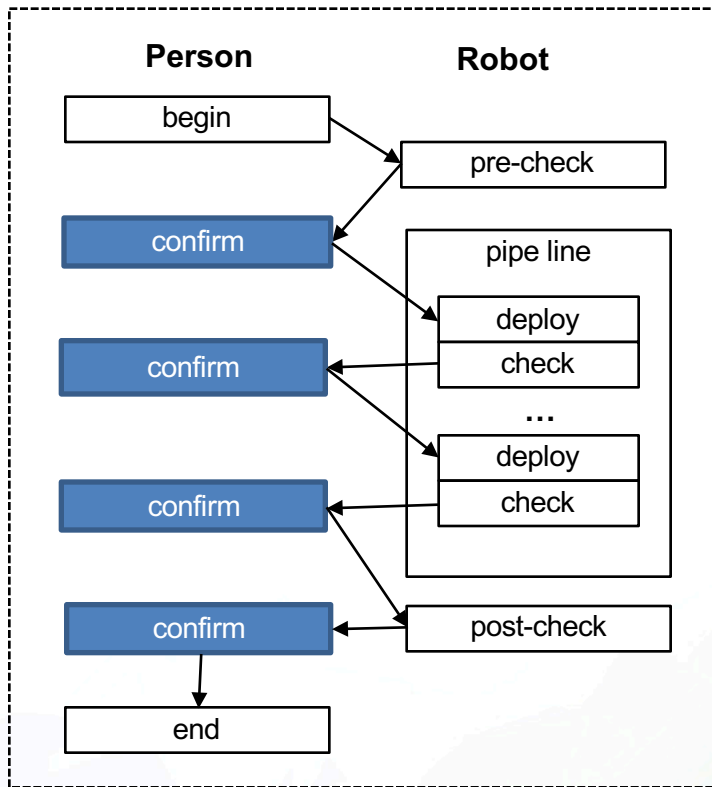
Check process optimization

Manual multiple metric dashboards check

Intelligent Anomaly detection

Manual single anomaly event dashboard check

Operation Knowledge Database

Automated API check and notify
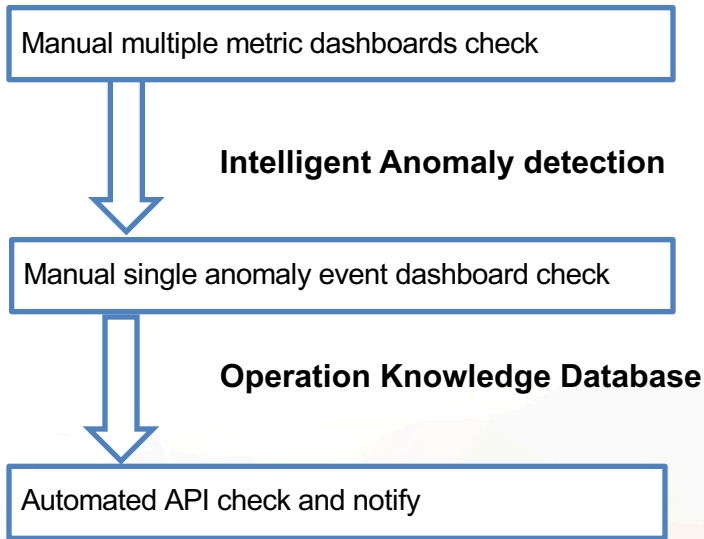
ChatBots do consultation

- Change consultation scenario

query example	intention	slots
xx模块从昨晚到现在有上线么？	Change query	Module：xx；Time：从昨晚到现在
今天xx模块有全流量上线么？	Change query	Module：xx； Time：今天 ;Stage：全流量
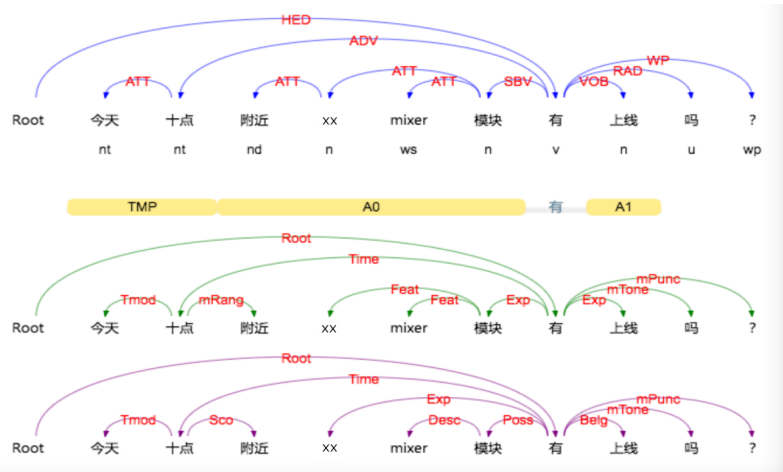
- Key points of building ChatBots

1. Accumulate manually labeled query

2. Train an NLP model to understand the questions offline

3. Translate natural language questions into structured questions

4. Query operation knowledge database

5. Display results on SRE Service desk

Future

- Dynamic resource allocation
- Capacity management
- Identification of performance problems
- …

THANKS

quxianping@baidu.com
hajingjing@baidu.com

Baidu 百度