Next Generation of DevOps
AIOps in Practice @Baidu

Speaker: Xianping Qu & Jingjing Ha
About Baidu
Agenda

• History of Baidu SRE team
• Next generation of DevOps – AIOps
• Best practice based on AIOps
• Future
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
AIOps (2014-)

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

- User Code 1
- User Code 2
- User Code …

User Code Management

- Operation Abstraction Layer
- Develop Tool-chain

- Develop Kit
  - Interface
  - Driver
  - Cloud/Paas
  - Runtime Environment
  - Scheduler

- IDE
- Simulation
- Build
- Test
- Debug
- Prof
Operation Knowledge Database

Unified data model
- Meta data
  - product
  - person
  - app
  - service
  - instance
  - host
  - IDC
  - network
  - ... 
- Metric data
  - throughput
  - latency
  - error
  - cpu
  - mem
  - io
  - disk
  - bandwidth
  - rtt
  - ... 
- Event data
  - anomaly
  - change
  - root cause
  - remediation
  - ... 

Data production process
- raw data
  - mapping
  - cleaning
- structural data
  - calculating
  - controlling
  - mining
- resulting data
  - feedback
  - authority, quota

Service management model
- metaDB
- TSDB
- eventDB

Data source
- Management platforms
- Monitoring platforms
- Operation platforms
Solution

General Components and Tools (transmission, storage, scheduling...)

Ops Algorithm Development

Ops Platform Development

Deployment Management

Incident Management

Operation Development Framework (SDK, RE)

Operation Knowledge Database

Anomaly detection

Traffic scheduling

Root cause analysis

Trend forecasting

Other data mining & machine learning algorithms

Solution Development
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
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
Two layer traffic shifting @ Baidu

User set

Internet

Data Center1

Data Center2

Data Center3

shift traffic between user set and the edge node

shift traffic between front end and back end

BGW: Baidu Gate Way, layer-4 load balancer
BFE: Baidu Front End, layer-7 load balancer
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
Shift traffic between front end and back end

Concerns:
- Service Capacity
- Intranet bandwidth

Scenarios:
- Web service cluster
- Dependent service cluster
- Internal network switch
Shift traffic between user set and the edge node

- Concerns:
  - Bandwidth
  - BGW/BFE Capacity
  - Delay

- Scenarios:
  - BGW & BFE
  - External network switch
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
- Traditional monitoring
  - Handcrafted anomaly detection
  - Lots of False Negatives and False Positives

Judgment
- Manual decision making
  - Depends on personal experience
  - Partial information
  - Wrong or slow decision making
  - No real-time feedback leads to more serious failures

Execution
- Scattered Scripts
  - Poor code quality
  - Low availability
  - Critical moment is unreliable

Manual decision making
- Perception
- Execution
Single cluster stop-loss after AIOPs

Perception
- Intelligent monitoring
  - Intelligent anomaly detection
  - Abnormal event stored in Operation Knowledge Database
  - High precision and recall

Judgment
- Automated decision making
  - Rely on Algorithm Platform
  - Global information
  - Accurate and quick decision making
  - Feedback control

Execution
- Standard framework
  - Development framework
  - Deployment framework
  - High development efficiency
  - High availability
The architecture of single cluster stop-loss

- **Algorithm Platform**
  - Anomaly detection Algorithm
  - Traffic shifting Algorithm

- **Perception**
  - Network monitoring
  - Business monitoring
  - Application monitoring

- **Judgment**
  - VIP Health Control
  - Service Health Control
  - Internal Traffic Load balancer
  - BFE Configuration Management

- **Execution**
  - DNS Configuration Management
  - Service Degradation

- **Operation Knowledge Database**
  - Abnormal event DB
  - Metric DB
  - Capacity DB
Unattended deployment with automated checker

- **Person**: begin -> pre-check -> confirm
- **Robot**: pipe line -> deploy -> check -> confirm
- **Robot**: deploy -> check -> confirm
- **Robot**: post-check -> confirm
- **Robot**: end

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

![Diagram of ChatBot consultation]

- 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

<table>
<thead>
<tr>
<th>query example</th>
<th>intention</th>
<th>slots</th>
</tr>
</thead>
<tbody>
<tr>
<td>xx模块从昨晚到现在有上线么？</td>
<td>Change query</td>
<td>Module : xx ; Time : 从昨晚到现在</td>
</tr>
<tr>
<td>今天xx模块有全流量上线么？</td>
<td>Change query</td>
<td>Module : xx ; Time : 今天 ;Stage : 全流量</td>
</tr>
</tbody>
</table>
Future

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

THANKS

quxianping@baidu.com
hajingjing@baidu.com