



# Monitoring as a Service(MaaS)

May Mei, Xingze Gao  
PM, Engineer

---

# Agenda

---

Why Monitoring

Monitoring for M365

Monitoring Framework

Demo

Tips

# Why Monitoring



Product/Service quality is one of the key indicator for the customers, especially the enterprise customers.

The availability/performance of services are highly relevant with the user amount, the reputation of product, even the company image and stock price.

# Why Monitoring

## **01 AWS : 3 Outages in Dec 2021**

Between 2010 and 2019, AWS outage 2.4 times per year. While in the last month of 2021, AWS outage 3 times.

## **02 Azure : Windows virtual machines outage**

On Oct 23<sup>rd</sup> 2021 Azure Virtual Machines outage for 6 hours which impact American, Europe, Middle East, Africa and Asia regions.

## **03 IBM Cloud : 2 outages in 5 days**

On May 22<sup>nd</sup> to 26<sup>th</sup> 2021, IBM outage twice in 5 days. And the incident on May 22<sup>nd</sup> was rated as severity one and a lot of regions globally are impacted.

## **04 Google Cloud: Outage in new region after release**

Google Cloud outage twice in 2021, Google Cloud Networking, Google Cloud Functions, Google Cloud Run, Google App Engine, Google App Engine Flex, Apigee and Firebase can't work; lots clients are impacted.

## **05 Facebook: outage for 7 hours globally.**

Facebook is down for 7 hours globally in Oct, 2021. 2 billion global active users can not use their services, and this incident caused 5% decrease of Facebook stock price(around 50 billion USD in total).

# What does monitoring means for M365

- **Our mission is to *empower every person and every organization on the planet to achieve more.***
- Microsoft 365 Consumer subscribers increased to 51.9 million.  
(According to Annual Report, 2021)
- **Office 365 SLA** in China Mainland(Via 21 Vianet Blue Cloud)

### 3. 服务级别承诺。

一项服务的“月正常运行时间百分比”按照下列公式计算：

$$\frac{\text{当月的总分钟数} - \text{故障期}}{\text{当月的总分钟数}} \times 100$$

如果某一月的月正常运行时间百分比低于 99.9%，您可能有资格获得下列服务费抵扣：

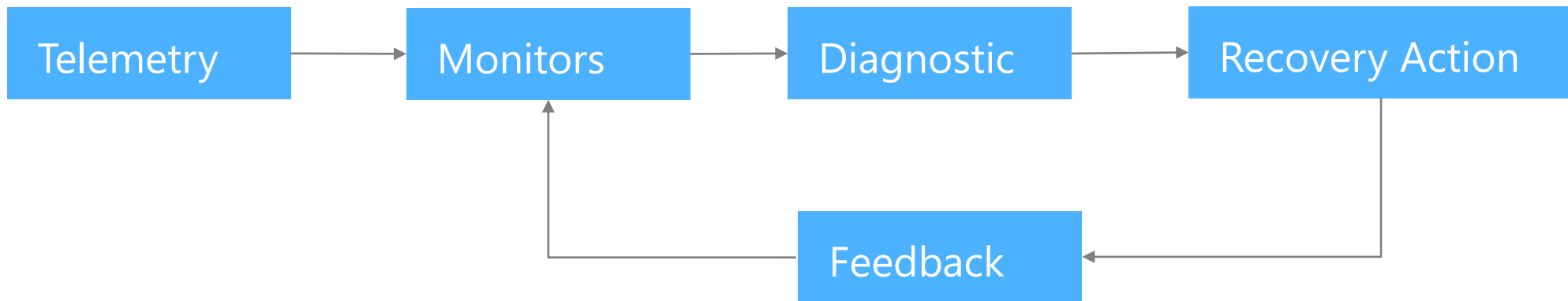
月正常运行时间百分比	服务费抵扣
< 99.9%	25%
< 99%	50%
< 95%	100%

图片来源：[世纪互联在线服务的服务级别协议 – 21Vianet Blue Cloud \(21vbluecloud.com\)](https://www.21vianet.com/zh-cn/bluecloud/)

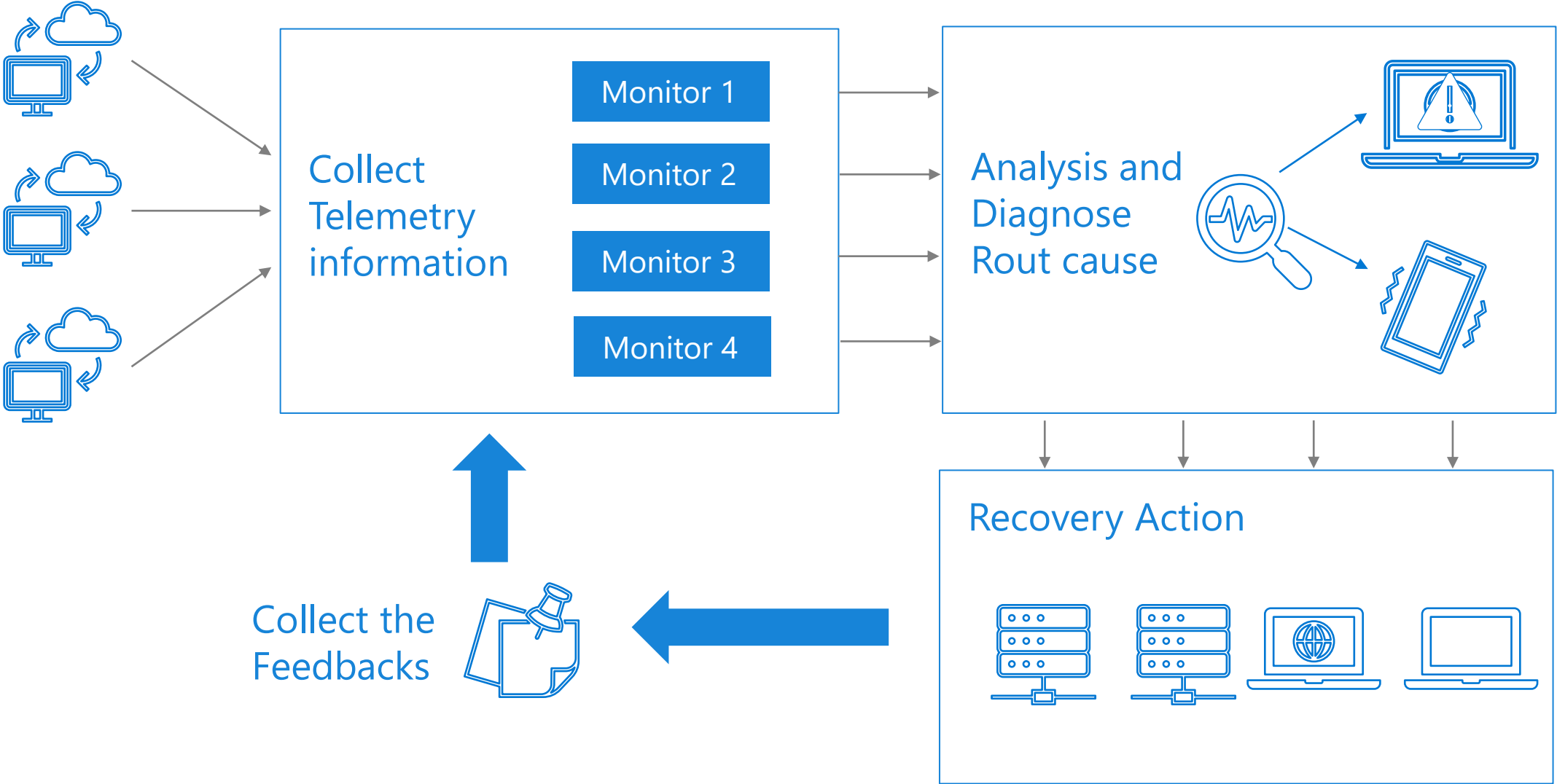
# Monitoring Framework



What if.....we could prevent it or mitigate before impacting more



# Monitoring Framework

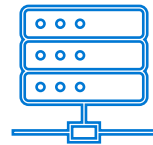


# The Monitoring Framework-Telemetry



## SDK/Metric

Using SDK to collect the metrics, using open telemetry technology to collect the logs.



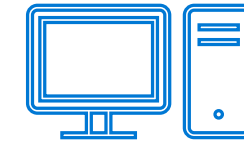
## Data

Aggregate the metrics/logs in different dimension for the next step.



## Signal

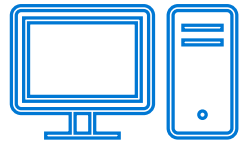
The signals you want to monitor including Passive signals, Active signals(Mimic user behaviors), Infra signals



## Monitors



# The Monitoring Framework-Monitoring & Diagnostic



## Monitors

Set up monitors for Availability, Latency or other metrics you care.



## Alerts

Define a threshold or calculate a threshold through historic data and when it lower/higher than the threshold it will fire an alert.



## Analysis

Analyze the root cause and send the alert to the owner/team with troubleshooting information.  
Upgrade/downgrade

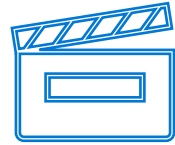


# The Monitoring Framework-Recovery Action



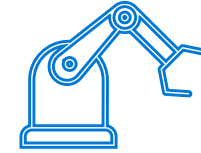
## Collect historic data

Before automating the recoveries, we can also use auto-fill the logs, troubleshooting tools/guide and dashboard to help debug. Collecting the troubleshoot activities to generate the common actions.



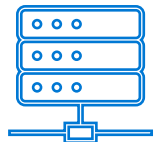
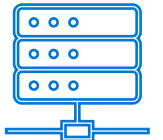
## Define recovery action

Generate the common recovery actions through the historic data. For example, exception xxx map to the action of xxxxxx.



## Automate recovery actions

Halt deployment  
Rollback deployment  
Memory dump  
Reboot  
Auto-mitigate;  
Capture other recovery actions in troubleshooting guide/tools



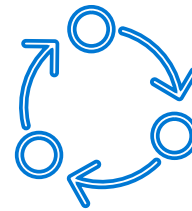
# The Monitoring Framework-Feedback



## **Alert/Diagnosis**

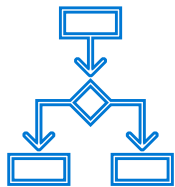
### **accuracy:**

Threshold setting,  
responsible owner, root  
cause.



## **Postmortem:**

Check monitor configuration;  
Provide/update troubleshoot  
guide; update recovery action  
and others.



## **Recovery action coverage and accuracy:**

The percentage of recovery  
and the accuracy of  
recovery actions.

# The Monitoring Framework-Supported Tools



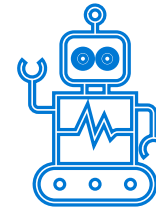
## **Dashboard of monitors**

Display the monitors health status in dashboard to help engineers easily view and locate the issue.



## **Communication channel**

Provide communication channel to better plan the deployment and prepare for the change and minimize the impact.



## **Oncall Assistant**

Build the oncall assistant to grab/summarize the key information to reduce the incidents TTM(Time to mitigate).

Demo

# Demo

Tips

Thank You.