Deterministic Operations 2.0 Cases

Security and Stability | Resource Efficiency | Service Agility | Intelligent Operations |

XoE: Revolutionizing Enterprise Operations and Unleashing Productivity on the Cloud

Building AZ Fault Detection and Switchover Capabilities to Achieve Minute-level Self-healing

Refined Management with Landing Zone for a Secure Cloud Ecosystem

A New Journey Driven by Intelligent R&D

Pangu Models Drive Innovation in Medical Technology, Fostering a Smart Healthcare Ecosystem with **Enhanced Doctor-Patient Connection**

Issued By

SRE Dept, Huawei Cloud Computing Technologies Co., Ltd.

Consultants

Gao Jianghai, An Yu, Xue Ying

Editor-in-Chief

Lin Huading

Deputy Editor-in-Chief

Wang Qingyuan

Authors

Chen Jie, Chen Zidong, Shen Ningkun, Cheng Ling, Deng Yuqiao, Gan Xin, Gao Wenke, Han Yi, Han Linlin, Han Xinyi, Kang Zhen, Li Guoqiang, Li Zhenxiang, Mu Yong, Qu Chen, Wang Gang, Wang Xinyang, Wang Fuqiang, Mao Xuepeng, Xie Qiong, Yin Gezhen, Yuan Zhou, Zhang Xi, Zhang Yibo (by alphabetical order)

Editors

He Liping, Wu Wenhui, Wu Shunqun, Lu Sujuan, Meng Miao, Xu Xinan, Zhang Fei (by alphabetical order)

Contact

To obtain the electronic publication, to request access to these articles, to contribute an article, or to provide suggestions or opinions, please contact the SRE Dept of Huawei Cloud Computing Technologies Co., Ltd.

Email

snzx1@huawei.com

Address

Xiliubeipo Village, No. 9, Huanhu Road, High-tech Industrial Development Zone, Songshan Lake campus, Dongquan, China

Postal Code

523830

General Disclaimer

The content in this magazine is for reference only. Huawei Cloud Computing Technologies Co., Ltd. does not provide any express or implied warranty of the content of this magazine, including but not limited to the warranty of merchantability or use for a specific purpose. To the extent permitted by applicable law, in no case shall Huawei Cloud Computing Technologies Co., Ltd. be liable for any special, indirect, or consequential damages, or lost profits, data, business goodwill, or anticipated savings arising out of or in connection with any use of this magazine

(internal release, free of charge).

Copyright © Huawei Cloud Computing Technologies Co., Ltd. All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Cloud Computing Technologies Co., Ltd.

Contents

	Security and Stability	01
	Building AZ Fault Detection and Switchover Capabilities to Achieve Minute-level Self-healing	02
	4+2+1 Secure O&M for Even More CDN PoPs	07
	Quickly Identifying Service Architecture Risks Through Chaos Drills, Improving System Resilience	12
	Shangluo Government Archived Secure, Efficient Cloud Transformation with Resilient Migration Services	18
	Cloud-based Protection Earned Full Marks in a Cybersecurity Drill	23
	Resource Efficiency	27
	Qingting FM: Application Modernization Is a New Industry Chapter	28
	Big Data Migration Enhances Cost Efficiency, Multi-Cloud Flexibility, and SLA Compliance in Finance	33
	Refined Management with Landing Zone for a Secure Cloud Ecosystem	38
	Service Agility	44
	A New Journey Driven by Intelligent R&D	45
Per	Intelligent Operations	50
7	AI Powers a Smart Revolution in AGV Efficiency and Service Excellence	51
	Pangu Models Drive Innovation in Medical Technology, Fostering a Smart Healthcare Ecosystem with I Doctor-Patient Connection	Enhanced 56
	RPA for Al-powered Government Services	62

Industry Voices

Cui Haojun

CEO, Shenzhen Irobotbox Tech Co., Ltd

Having explored the *Huawei Cloud Deterministic Operations 2.0 Cases*, I am struck by the transformative power of AI in the operations field. As an IT professional, I have witnessed firsthand how AI enhances efficiency by automating repetitive tasks, reducing both time and costs. Beyond automation, AI enables predictive risk analysis, allowing proactive fault prevention. These advancements ensure more reliable and streamlined enterprise operations, driving superior business outcomes. For those intrigued by AI's role in operations, this is a must-read.

Liu Yixin

CTO, Quark International Holdings

In the FinTech sector, we recognize the critical importance of data stability and security in enterprise digital transformation. The case study Big Data Migration Enhances Cost Efficiency, Multi-Cloud Flexibility, and SLA Compliance in Finance showcases Quark's collaboration with Huawei Cloud to successfully migrate its data platform services, boosting service quality while also cutting costs. Faced with challenges like stringent data accuracy requirements and massive operations tasks, Quark adopted a hybrid approach combining full migration, incremental migration, real-time synchronization, and scheduled migration. This strategy ensured data integrity and consistency, accelerated migration, and enhanced system compatibility. Offering practical insights for operations professionals in the finance industry, the article highlights strategies for achieving flexible data deployment and management while maintaining uninterrupted operations. It serves as an inspiring guide for building resilient financial big data platforms and driving enterprise digital transformation forward.



Lian Yong

Operations Director, Glodon Company Limited

This compilation of case studies showcases operations strategies and experience of enterprises across diverse industries. Highlighting key areas — security and stability, resource efficiency, business agility, and intelligent operations — it illustrates how operations professionals gain deep insights into prevalent issues and leverage cutting-edge technologies to address challenges effectively. Offering practical guidance, it helps operations teams navigate common pitfalls and adopt innovative approaches. By embracing the advanced principles and proven methods outlined here, practitioners can boost efficiency, foster growth, and achieve significant progress in their fields.

Yang Jinguan

R&D VP, Beijing Networkbench Co., Ltd (also called Tingyun)

In today's digital era, driven by rapid advancements in technologies such as cloud computing, big data, and AI, enterprise operations have evolved beyond IT functions to become central to innovation and business sustainability. The *Huawei Cloud Deterministic Operations 2.0 Cases* serves as an essential guide for professionals across diverse industries.

Focusing on the four pillars of Deterministic Operations 2.0 — security and stability, resource efficiency, business agility, and intelligent operations — this compilation delves into the specific demands of sectors like finance, automotive, and healthcare. Through meticulously curated and analyzed case studies, it illustrates practical applications of Deterministic Operations 2.0 while highlighting key challenges and solutions encountered during enterprise operations transformations. This offers actionable insights for enterprises navigating cloud migration, use, and management.

I highly recommend this collection to anyone invested in advancing enterprise operations. Whether you are leading organizational changes or seeking to deepen your expertise, this book is a vital resource.





Security and Stability

Faults Prevented, Controlled, and Resolved with Industry-Leading Deterministic Operations; Cloud Safeguarded with Automated Security Operations and Management



Building AZ Fault Detection and Switchover Capabilities to Achieve Minute-level Self-healing

Abstract

This case uses an equipment room fault as an example to analyze the mid- and long-term advancing policy of the OWS architecture and implement the HA optimization solution. In this case, OWS achieves automatic switchover in minutes upon the occurrence of an AZ fault, significantly shortening the fault self-healing duration and improving system availability.

Background

OWS is a tool platform developed by Huawei GTS. It is deployed on cloud to offer services to projects and customers around the world, with an availability of over 99.99%. Faults in equipment rooms and data centers have significantly negative impacts

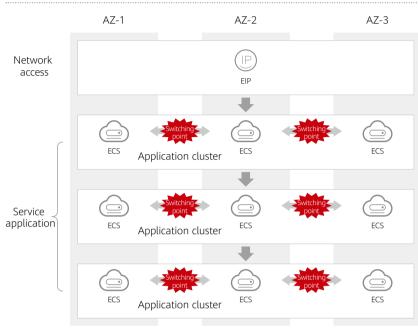
on the reliability of cloud tenant applications. Leading cloud service providers have experienced various issues with equipment rooms and data centers in the past, resulting in disruptions to customer service. Addressing such issues requires building an HA system architecture considering the application deployment modes, microservice dependencies, resource redundancy, and fault detection and switchover capabilities to improve service HA capabilities.



Service Status-Quo

OWS services are deployed in many regions around the world, with each region operating its own independent OWS environment. In each region, OWS microservices are deployed in a cluster across multiple AZs. For example, the GAM user authentication service of OWS is delivered through a three-node cluster, deployed in three AZs. According to historical data analysis, 80% of OWS outages are caused by single-AZ faults. Therefore, it is critical to improve service availability in single-AZ fault scenarios.

One OWS system



Guaranteeing the reliability of the AZ cluster deployment solution faces the following challenges:



Upper-layer applications fail to detect the subhealth state of cloud resources.

ECS availability depends on the status of underlying resources, such as network, storage, and computing resources. Underlying resource faults (particularly those in subhealth state), however, are not visible in the ECS status, and therefore cannot be detected by upper-layer applications. As a result, the upper-layer applications mistakenly consider a faulty ECS to be healthy and available and transfer service requests to it, causing the requests to time out or the responses to fail.



Faults can spread from a single AZ to the entire system.

Microservices are deployed on cluster nodes and called across AZs. When some hosts in an AZ are in subhealth state but not isolated, service requests are not responded to after being transferred through multiple layers, leading to a global fault.



Assessing resource redundancy can be challenging.

It is difficult to effectively assess resource redundancy of an AZ. In a deployment scenario with three AZs, if an ECS in one AZ fails, 66% of the resources in other AZs have to handle all service loads. Insufficient resources can lead to system overload and performance degradation. Therefore, it is essential to develop an effective method for assessing resource redundancy.

Solution Implementation

Optimization Solution Overview

Based on a systematic analysis, the OWS product line has set a long-term goal of achieving an annual availability of \geq 99.999%. This goal is driven by the strategic approach of "building a five-nines product availability on a three-nines cloud infrastructure stability" and will be achieved through continuous development of product capabilities and SRE capabilities.

The product line aims to reach the goal in three phases:

Phase 1: Enhance the cross-AZ cluster reliability to quickly detect single-AZ faults and perform switchover within a cluster, achieving 99.95% system availability.

Phase 2: Implement the cross-AZ multi-active capability to quickly detect single-AZ faults and perform switchover across clusters, achieving 99.99% system availability.

Phase 3: Implement the cross-region DR capability to quickly detect single-

The subsequent parts illustrate the solution details.

1. Identification of single-AZ deployment risks

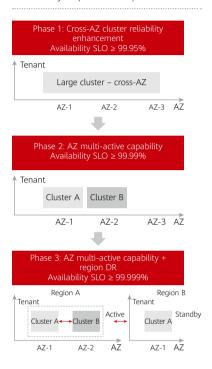
In the microservice deployment phase, single-AZ deployment may occur due to low resource availability and manual configuration errors. In subsequent reconstruction, single-AZ deployment may not be identified due to insufficient visualization, resulting in single-AZ risks. The OWS's O&M platform, iDOE, establishes mappings between the environments, microservices, hosts, and AZ labels through CMDB to manage and visualize the cross-AZ HA of the environments and microservices, effectively identifying risks in single-AZ microservice deployment.

region faults and perform switchover across clusters, achieving 99.999% system availability.

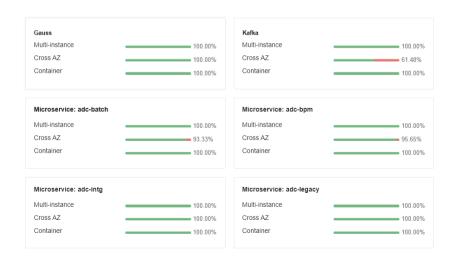
Currently, phase 1 has been completed. The detailed solution is as follows:

- (1) Risk identification of single-AZ deployment: Multiple deployment modes are manageable and visible through an HA panel.
- (2) Balanced distribution of microservices across AZs:
 Microservices are evenly distributed across AZs by utilizing K8s labels, ensuring optimal resource usage.
- (3) Single-AZ fault detection and isolation: When a single-AZ fault occurs, the microservice cluster can detect the faulty node and isolate it.
- (4) Clarified capacity management rules: Cluster resources are properly planned based on rules to ensure stable service running after faulty AZs are isolated.

Availability improvement plan



Balanced distribution of microservices across AZs



2. Balanced distribution of microservices across AZs

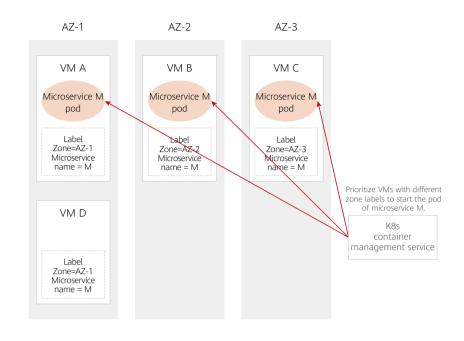
Multiple microservices are deployed in a large ECS resource pool and scheduled based on loads to avoid unbalanced resource usage. Assume that there are nine ECS nodes in an ECS resource pool evenly distributed in three AZs, meaning that each AZ has three ECSs. Further, assume that a microservice can meet service requirements with only two pods started. If all pod nodes of the microservice are running in one AZ, a failure in the AZ equipment room will cause all the nodes to become unavailable. To address this issue, the Topology Spread Constraints feature of K8s is used to schedule the distribution of pods in AZs.

- (1) Zone labels (e.g., zone=AZ-1) are added to nodes in the data zone when the nodes are managed. K8s divides AZs based on the zone label values on the nodes.
- (2) The **Topology Spread Constraints** field in the pods guides the distribution of pods in the AZ. Pods are evenly distributed to AZs based on the app labels, meaning that the difference between the number of pods in each AZ is less than or equal to 1.

In the following figure, microservice M is preferentially deployed on VMs A, B, and C with different zone labels (AZ-1 to AZ-3). Because VM D has the same zone label as VM A, the priority of VM D is lower than that of VMs B and C for microservice distribution (cross-AZ scheduling is preferred).

3. Single-AZ fault detection and automatic isolation

When a single-AZ equipment room is faulty, the cloud vendor conducts fault detection and switchover of cloud services such as networks, middleware, and databases. Meanwhile, the service side (OWS) is

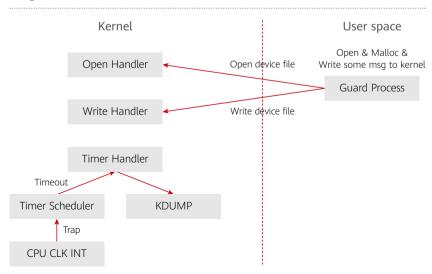


responsible for ECS fault rectification and conducts cross-AZ cluster switchover, because the cloud provides only single-node capabilities.

Utilizing the Hang-Handler mechanism at the OS layer, OWS enhances the cluster's capability to proactively isolate faults on host CPUs, memory, disks, and I/O. The

Hang-Handler sets a kernel timer. If the timer times out without any updates, KDUMP or LKCD is triggered to restart the OS. After the OS restarts, the service pod running on the ECS is removed from the service queue, enabling fault detection and isolation. After the fault is rectified, the service pod is reassigned for fault self-healing.

Hang-Handler mechanism



4. Clarified capacity management rules

In a cluster deployed across three AZs, if the equipment room in a single AZ is faulty, all ECS nodes in the faulty AZ become unavailable, and the service load is redistributed across the other two AZs. If the two AZs do not have sufficient resources for redundancy, service overload occurs, leading to additional service disruptions.

To effectively manage the computing resource capacity configurations of the platform runtime environment, the specifications on GTS cloud service computing resource capacity management have been formulated,

has been formulated, which clarifies the computing capacity requirements in the deployment planning phase and runtime phase and aims for the following objectives:

- (1) Properly plan service computing resources to maintain service functions and prevent performance degradation in multi-AZ deployment scenarios following a single AZ failover.
- (2) Effectively manage the computing resources used for O&M and security to prevent them from preempting service computing resources and ensure that exceptions in O&M or security processes and threads do not affect service functions.

Some rules are as follows:

- [Rule] The total number of CPU requests on the container processes of the host machine cannot exceed (Host machine CPU specification Number of CPU resources reserved for non-service processes) × 60%.
- [Rule] If the P95 value of the CPU usage on the host machine or container, and that of CPU load/CPU cores on the host machine are greater than 70% within 24 hours, take measures for mitigation. Scaleout is preferred.
- [Rule] If the P95 value of the memory usage on the host machine is greater than 90% within 24 hours, take measures for mitigation.

Business Improvement

To ensure that service applications can be switched over as expected in the event of a single AZ failure, it is essential to conduct chaos engineering drills on the live network and identify problems in advance for rectification. Based on fault pattern identification, faults in an AZ equipment room are classified into several scenarios, and all scenarios are covered in the drills.

With the support of Huawei Cloud technical staff, the GTS SRE team has conducted five large-scale AZ fault injection drills in a production-like environment, involving 200+personnel from Huawei Cloud, GTS SRE, and OWS R&D and test teams. The drills have identified several key reliability improvement points and optimized them as follows:

Drill category	Subcategory	Scenario	Fault injection tool	Expectation
Single-AZ power-off	Single-AZ power-off	Host, database, or cloud service suspension	Management console	Fault detection and switchover within <i>X</i> minutes
	Host cloud service subhealth	CPU, memory, disk and I/O subhealth	CAST service on the Huawei Cloud tenant plane	
Single-AZ cloud service	Database cloud service subhealth	CPU, memory, disk and I/O subhealth	CAST service on the Huawei Cloud tenant plane	
subhealth	Cloud middleware service subhealth	CPU, memory, disk and I/O subhealth Single-node request delay	CAST service on the Huawei Cloud tenant plane CloudScope on the Huawei Cloud management plane	

- (1) Even distribution of microservice resources: 90% -> 100%
- (2) Host disk fault detection capability: 0% -> 100%
- (3) Database fault perception capability: 80% -> 100%
- (4) AZ failover time of OWS: several hours -> several minutes

Summary

To address single-AZ equipment room issues and improve the system availability of OWS products, this case analyzes the mid- and long-term system architecture evolution and short-term system architecture optimization, and specifies the following

improvement strategies for optimization: (1) Risk identification of single-AZ deployment; (2) Balanced distribution of microservices across AZs; (3) Single-AZ fault detection and automatic isolation; (4) Clarified capacity management rules.

In addition, multiple rounds of chaos engineering drills have been conducted on the live network based on fault pattern identification. The drilling result shows that a single-AZ failover can be achieved within minutes.

4+2+1 Secure O&M for Even More CDN PoPs

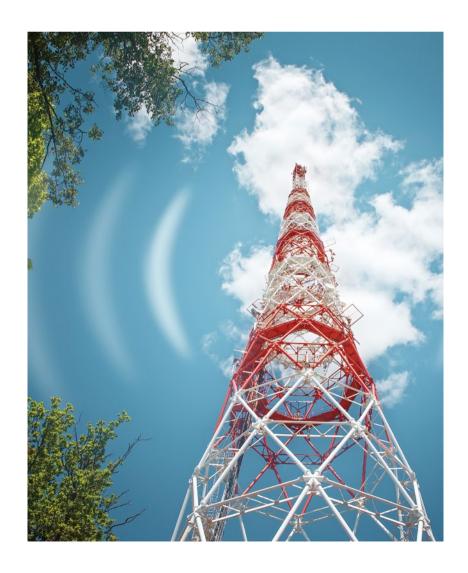
Abstract

Group Z faced tough competition in the content delivery network (CDN) market. They partnered with Huawei Cloud to create a 4+2+1 operations and maintenance (O&M) service system on their existing CDN. This partnership boosts efficiency, enhances competitiveness, improves experience, and powers growth.

Background

The fast growth of the Internet, cloud computing, Internet of Things (IoT), and 5G has led to a sharp rise in digital content consumption, especially online videos. By 2025, over 55 billion IoT devices will be connected and produce huge amounts of data. As a result, businesses urgently require highquality content services with high bandwidth, low latency, and heavy connection support.

Z is a corporate group that specializes in network technology. Its proprietary CDN offers fast, reliable, convenient, and secure Internet access for users worldwide. However, CDN providers now offer similar basic technical features, so CDNs with only costeffective bandwidth distribution are no longer competitive. The key challenge now lies in combining basic technologies with strong O&M skills to stay ahead.



Challenges

Z faced these main challenges in developing its CDN services:



Upgrading the service system

CDN services run in real time. They require coordination between professional O&M teams, service processes, and assurance measures. Specialized O&M skills are crucial during unexpected traffic surges, and professional teams are needed to analyze service indicators and refine operational strategies. An upgraded system would improve O&M service stability, reliability, and efficiency.



Improving the system architecture

Z struggled to update CDN technology, analyze changes to service models, and plan and adjust system architecture. These limitations affected system performance.



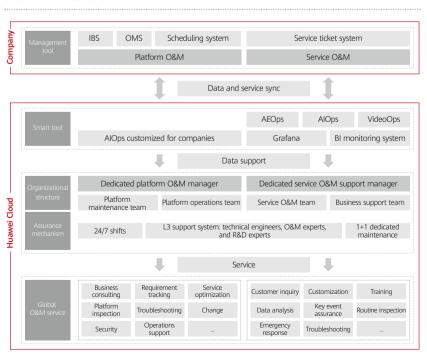
Automating O&M

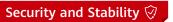
CDN services involve hundreds of key indicators and generate a large amount of data. Real-time analysis is required for prompt policy adjustment. Z initially depended on manual efforts by the O&M team and basic tools with limited automation, leading to poor efficiency and high management costs.

Solution

Z leveraged Huawei Cloud's proven expertise in CDN to create a 4+2+1 O&M service system designed for industry developments and its current services. This system centers on CDN service delivery, maintenance, and operations. It includes the organizational structure, service processes, management tools, and assurance mechanisms.

Solution architecture





The 4+2+1 O&M service system is so named for four service activities, two service plans, and one smart tool platform.

Four service activities

Proactive inspection

Inspection templates and configurations are customizable to service requirements and system status. The system regularly checks more than 200 inspection items and pushes inspection results to service owners in real time. 24/7 shifts ensure timely handling of problems found during inspections.

Emergency response

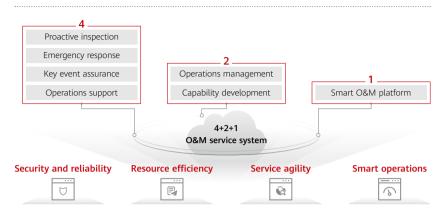
Such measures ensure service stability by covering common CDN service incidents, such as faults on points of presence (PoPs), network interruptions, traffic surges, content hijacking, PoP attacks, service faults, and hardware faults.

Key event assurance

Z partnered with Huawei Cloud to handle traffic surges during key events such as product launches and live sports. They formed a dedicated expert team and tailored a solution with pre-event evaluation, in-event assurance, and post-event summary.

- (1) Pre-event evaluation: Assesses services, develops assurance solutions, evaluates and verifies solutions, and provides technical consulting and support.
- (2) In-event assurance: Responds to customers quickly with dedicated measures, such as global unified scheduling, intelligent inspection, and hierarchical emergency prevention. The smart O&M platform accurately monitors over 20 key service indicators, such as CPU and bandwidth, and optimizes services in real time.

4+2+1 O&M service system



Emergency response



Key event assurance process



(3) Post-event summary: Provides reports to collate assurance experience.

Operations support

A daily monitoring system checks routine service access, inspection, troubleshooting, change, requirement tracing, and security management. Team members collaborate better by reporting platform needs and risks identified every week, sharing updates on fixes and improvements for management plane issues, and updating the PoP construction progress and service traffic details each month.

Two service plans

Operations management

The O&M service system prioritizes the quality and development of service operations. It analyzes daily data to identify service requirements and trends and optimize performance accordingly.

For customer operations, the smart tool platform helps customers manage their services more efficiently with value-added services, such as configuration customization, domain conflict handling, log management optimization, and automatic data analysis.

For service operations, the smart tool platform automatically analyzes key service indicators such as domain quality, PoP quality, traffic ratio of the local province, service volume change, and PoPs with little or no traffic. It generates reports and schedules traffic based on CDN service needs and rules for maximum system service value.

Capability development

Z collaborated with Huawei Cloud on comprehensive CDN courses and phase-by-phase training. Courses cover market analysis, service assessment, planning, design, product selection, solution configuration, system maintenance, troubleshooting, system operations, service operations, and R&D enhancement. Together, these efforts form a structured service system with architecture design, performance tuning, and service assurance capabilities.

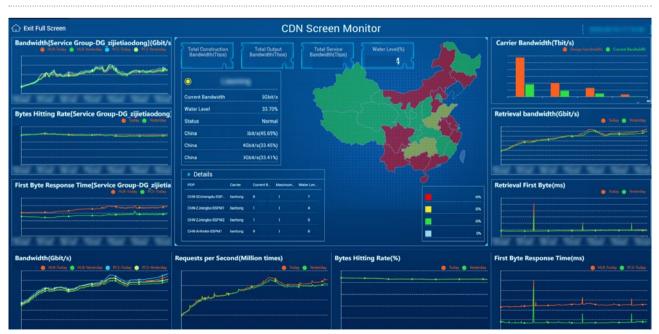
One smart tool platform

Z worked with Huawei Cloud on a specialized smart management platform that tracks live system status, displays key service indicators and trends in real time, allows digital data management, and boosts O&M efficiency.

The smart management platform offers these key features:

- Systematic configuration management: Automates configurations, such as adding or removing PoPs and adjusting parameters. This makes operations easier and ensures consistent, accurate setups.
- Log management and analysis:
 Collects CDN logs and creates
 analysis reports. These help the
 O&M team understand traffic
 patterns and user actions and
 make decisions.
- Global smart O&M: Monitors and analyzes key service indicators such as water level, traffic, and bandwidth in real time. Based on smart scheduling policies, if a PoP fails, the platform alerts the team and switches services to normal PoPs to maintain service stability. It lowers traffic sent to overloaded PoPs and balances traffic across PoPs to improve performance. It also automatically selects the best delivery routes.

Smart data analysis platform



Service Improvement

Reliable processes

A specialized O&M team developed standard processes for service monitoring and analysis, fault detection and alerts, emergency handling, and incident and change grading. Team members work in 24/7 shifts to maintain service reliability.

Optimized architecture

To handle more requests for HTTPS access, dynamic content acceleration, and smaller files, Z adjusted the quantity and hardware configurations of CDN PoPs. This prevents server overloads and PoP failures under high query rates.

Stable support for burst traffic

During sudden traffic surges, the O&M team maintained system stability and resource availability at peak hours. This improved customer satisfaction, enhanced Z's reputation, and increased demand for key event assurance services by 420%.

Summary

This new CDN service provider quickly entered the CDN service market with Huawei Cloud's 4+2+1 O&M service system. On top of steady growth, Z developed efficient service capabilities in terms of organizational structure, processes, systems, services, and assurance mechanisms. Focusing on high-availability O&M, Z helped businesses optimize costs, improve efficiency, enhance user experience, and unleash efficiency. This differentiated service edge set the company apart in a competitive market. Using the 4+2+1 O&M service system, Z built a strong industry reputation, attracting more customers. Requirements for on-demand service acceleration grew by 240%, while livestreaming acceleration needs increased by 260%. The number of domain names using its CDN expanded by 50%, enabling further business growth.



Quickly Identifying Service Architecture Risks Through Chaos Drills, **Improving System Resilience**

Abstract

This case illustrates how an aviation enterprise enhances system resilience by leveraging the key functions of Cloud Operations Center (COC) when dealing with challenges such as high system complexity, huge fault blast radius, limitations of conventional test methods, and rapid version iterations. These functions include application architecture analysis, risk identification, chaos drill scenario design, deterministic drill plan development, and result analysis and service improvement.

Background

Company G is a large-sized, stateowned enterprise with significant influence in the Chinese air transportation sector. Given the particularity and complexity of air transport, the company attaches great importance to system stability and reliability. The system's high performance standards not only elevate the passenger travel experience but also guarantee the

safety and timeliness of goods transportation, thereby enhancing the company's competitiveness. As digital transformation progresses, Company G is systematically migrating its onpremise operations to Huawei Cloud. To better maintain continuous service running and verify system resilience following service migration to the cloud, Company G urgently needs a method for proactively identifying potential system issues and risks.

Following a comprehensive evaluation of Huawei Cloud's professional services, Company G finds that COC delivers a one-stop, automated drill function. This function covers the entire drill process from risk identification to drill review and service improvement, suiting the company's needs.



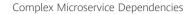
Pain Points

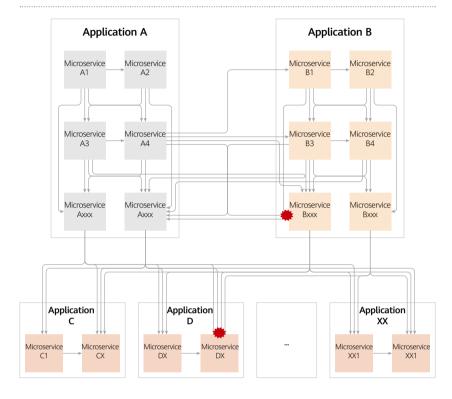
Company G operates on an immense scale with constantly-growing business volume. The application architecture is complex and involves collaboration between multiple systems and services. This means high requirements for system resilience, necessitating the ongoing optimization of related technical architectures and management policies to ensure the efficient and stable running of services. Currently, Company G faces the following four business pain points:



High System Complexity

The application architecture consists of various microservices, and the dependencies between these microservices are complex, like an elusive network. This makes it difficult for the O&M team to gain a comprehensive understanding of system behavior and potential risks, increasing the difficulty of troubleshooting and risk prevention.







Limitations of Conventional Test Methods

Conventional test methods are mainly used to test known faults and service scenarios. However, they fall short in terms of identifying potential system issues or simulating complex actual service scenarios.



Derived Risks in Rapid Iteration

To better meet market demands, the service system requires rapid iteration and updates. However, frequent introduction of new features comes with risks. Such risks affect the stability and reliability of the system to some extent.



Large Fault Blast Radius

In the aviation industry, due to the importance and extensive interconnectivity of services, even minor system malfunctions can lead to issues such as flight delays or cancellations. These disrupt the normal travel plans of countless passengers, causing serious economic losses and reputational damage.

Solution Implementation

Leveraging the scenario-specific troubleshooting capabilities of the Huawei Cloud deterministic operations solution, COC can effectively address such pain points of Company G. Huawei Cloud offers a one-stop, automated drill function designed to proactively identify potential risks that your applications may encounter in a cloud environment. When utilizing the chaos drill function, you can effectively handle emergencies with contingency plans while simulating various real-word fault scenarios with fault injection. This approach ensures that you are well-informed about any vulnerabilities or threats before they become issues.

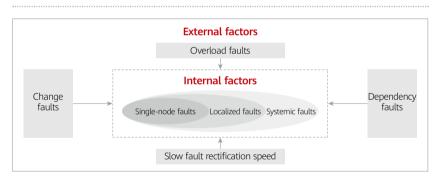
1. Application Architecture Analysis and Risk Identification

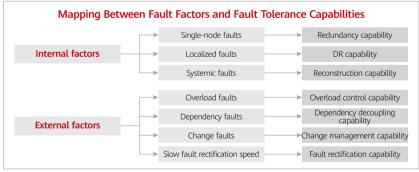
- 1) Through in-depth communication with Huawei Cloud, COC predicts the capacity requirements based on actual service traffic and architecture design objectives, with a focus on key performance indicators (KPIs). Additionally, COC clearly defines the relationship between system functions and services.
- 2) The service team reviews the requirement, design, and interface documents and specifies the information about the internal and external interfaces of the system. Additionally, COC monitors the actual interaction status and identifies potential issues with these interfaces using methods such as static code analysis, walkthroughs, integration tests, real-time metric monitoring, and log analysis.
- 3) Using Huawei Cloud's FT-FMEA risk analysis method, the service team analyzes potential risks from the perspective of service fault tolerance by conducting both forward development and backward improvement analyses based on emergency measures. This approach uses a comprehensive failure mode library that covers seven types of faults: single-node faults, localized faults, systemic faults, overload faults, dependency faults, change faults, and slow fault rectification speed. Based

Huawei Cloud SRE Chaos Drill Practice



FT-FMEA Risk Analysis Method



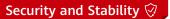


on this library, COC has developed a series of fault contingency plans.

2. Chaos Drill Scenario Design

To verify the high availability of Company G's cloud services, COC conducts chaos drills in seven key dimensions: fault detection, fault localization, fault impact evaluation, expected fault rectification duration estimation, fault handling methods, fault rectification implementation, and rectification confirmation. The key drill scenarios are listed as follows:

» Faulty RDS primary node: Simulate a misoperation that causes the host of the RDS primary node to go offline, resulting in an abnormal instance. This fault will lead to intermittent



disconnections between the service and the database. In this case, COC verifies whether the RDS primary/ standby switchover operates normally and assesses the actual service disruption.

- » High-latency fault in the NIC of a single-node DMS: Simulate a scenario where an increase in service traffic causes high network latency due to insufficient performance of the DMS node's NIC. In this case, COC verifies the availability of the instance under these conditions.
- » DCS process restart: Simulate a situation where a spike in service traffic causes the DCS workloads to reach the upper limit, leading to a node failure. In this case, COC tests the master/standby switchover capability of the DCS service when the master node becomes unavailable.
- » CCE node overload due to spike in CPU usage: Simulate a node exception caused by a sudden increase in service traffic and request stacking. In this case, COC checks whether the service is properly deployed in HA mode.

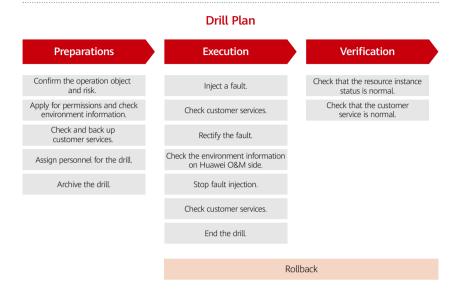
3. Drill Plan Development

- Create a drill plan for each scenario. The plan includes pre-drill preparations (permission request, environment check, and service check), drill execution, post-drill verification, and rollback plans.
- Specify the drill team roles. Huawei acts as the operator. Company G acts as the coordinator, monitoring owner, and service verification owner.

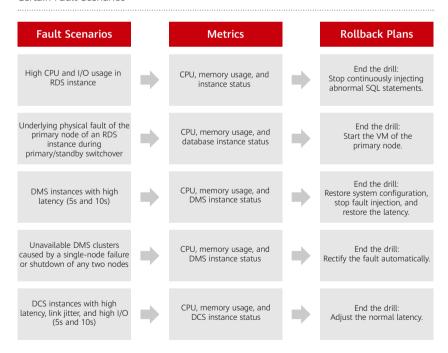
4. Drill Execution and Monitoring

- 1) Before drill execution: Given that the chaos drills aim to fully identify potential system risks, a specific contingency plan should be precisely matched to each fault scenario, and corresponding steady-state and monitoring metrics should be set. This ensures that even if unexpected issues arise during the drills, they can be handled in an orderly manner, safeguarding system stability and reliability, and effectively achieving the expected goals of the chaos drills.
- 2) During drill execution: The drill

Drill Plan



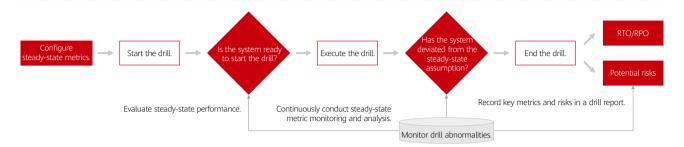
Certain Fault Scenarios



team initiates a drill task based on the pre-determined failure mode. It then properly configures the monitoring metrics to comprehensively monitor system statuses and KPIs such as response duration, throughput, and error rates in real time. Once the system exhibits abnormal behavior, the drill execution personnel can quickly

terminate the drill process by clicking the "immediate stop" button on the page and immediately perform fault rollback operations. This ensures the stability and recoverability of the system during the drill execution, highlighting the critical value and rigor of chaos drills in testing system risk detection and emergency handling capabilities.

Drill Process



5. Result Analysis and Continuous Improvement

After the drill is complete, the drill team performs in-depth analysis on the drill result from the following aspects: fault detection, fault localization, fault rectification, service blast radius, and O&M team cooperation. The key findings are as follows:

In the DCS high-latency drill scenario, some services reported access errors. The Huawei Cloud expert team recommended using service degradation strategies in cases of multi-instance and multireplica deployments to reduce dependency on Redis. When the backend ECS associated with an ELB encounters abnormalities, an excessive number of health check detections configured on the service side results in prolonged wait time for fault detection. O&M personnel were recommended to reduce the number of health checks or shorten the check interval.

In this drill, four potential high-risk items were identified in Company G's service system, representing 25% of the drill scenarios. Based on the identified issues the Huawei Cloud drill team produced a detailed drill report and provided optimization recommendations.



End-to-End Chaos Drill Capabilities

End-to-End Chaos Drill Capabilities Risk Analysis Fault Drills Improvement Tracking Resilience score Risk analysis using FT-FMEA Automated Steady-state fault injection Drill report Built-in 300+ failure modes Safe termination Fault rectification Improvement tracking Contingency plan Feed back drill review results to the risk management module, and execute continuous drills to mitigate risks. Steady-state metric Failure mode Contingency plan Fault orchestration Review and improvement monitoring Analyze application architecture to Develop contingency plans targeted Orchestrate attack tasks to support Automatically score resilience and Monitor steady-state metrics to identify failure modes. for the applied failure mode. complex scenarios. track improvement items. ensure safety and controllability.

Service Improvement

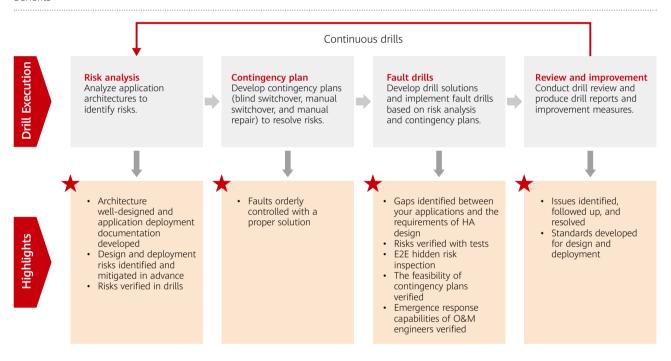
This chaos drill effectively enhances the resilience of Company G's cloudbased services. Specifically:

Improved system resilience: The system's fault tolerance and fault

rectification mechanisms have been effectively verified and optimized. In various fault situations, the system can quickly return to normal operation, reducing service downtime and data loss.

Enhanced risk awareness: Through this drill, Company G's development and O&M teams have gained a deeper understanding of the system's potential risks, improving their risk awareness and validating the feasibility of the routine O&M contingency plans.

Benefits



Summary

As a large-scale, state-owned enterprise in the domestic air transport market, Company G has exacting requirements for system stability due to its specialized business needs. Following the migration of services to Huawei Cloud during

its digital transformation, ensuring stable service operation and verifying system resilience have become top priorities. This drill encompasses the entire process from risk identification to review and improvement.

Through this drill, Company G has optimized the fault tolerance and fault rectification mechanisms of its systems, enhanced risk awareness within the team, and validated the feasibility of the routine O&M plan.

Shangluo Government Archived Secure, **Efficient Cloud Transformation with Resilient Migration Services**

■ Abstract

In keeping with China's "East-to-West Computing Resource Transfer" strategy, Shangluo city has been continuously promoting the development of "Qinling Smart Valley and Digital Shangluo" to drive high-quality transformation and growth. Following the completion of phase 1 of the Qinling Cloud Computing Big Data Center, Shangluo smoothly and securely migrated its government cloud to the center with the support from Huawei. Using scientific migration methodologies, efficient execution, and extensive cross-industry expertise, our migration team delivered a 40% faster transition while maintaining complete system stability for Shangluo.

Background

In recent years, China has been actively advancing the "East-to-West Computing Resource Transfer" initiative to promote balanced regional development in the digital economy. This initiative aims to build a new computing network that integrates data centers, cloud computing, and big data to transfer data and processing tasks from the eastern regions to the western regions, where compute resources can be provided more efficiently. In line with this national strategy,

Shangluo has been enhancing its digital information infrastructure across six major application sectors: industry, environment, agriculture, cultural tourism, healthcare, and logistics. This effort has been driving the development of the "Qinling Smart Valley and Digital Shangluo" project and has been accelerating the transformation of traditional industries in the city.

As a key part of Shangluo's smart city development, the Qinling Cloud

Computing Big Data Center serves as an information infrastructure hub. Once completed, it will provide a unified cloud platform for government agencies and various industries, enabling centralized and shared data resources to empower smart city initiatives, public services, and industrial growth. Additionally, it will extend its reach to surrounding areas, laying a solid foundation for achieving the strategic goal of data computing and backup in Shangluo.



Challenges

To enhance the reliability, security, and scalability of the government cloud platform while fostering innovation in government operations and governance, Shangluo officially launched its government cloud migration plan following the completion of phase 1 of the Qinling Cloud Computing Big Data Center. The plan involved migrating the Shangluo government cloud platform from its

temporary data center to the Qinling Cloud Computing Big Data Center.

Service migration is a complex and systematic project, where any misstep can result in service disruptions, data loss, or data breaches. Service migration can involve significant and unforeseen risks. The Shangluo government cloud platform had been operating in the

temporary data center for over four years. This platform supports the municipal and district-level core government information systems. These systems are deeply integrated with government functions, urban management, public services, and industrial development. Given its critical role, this migration demands exceptional service continuity, security, and success rates. The specific challenges are as follows:



Large migration scope

- xx service systems of xx government agencies on leased and hosted devices
- Numerous servers and technical components, and massive volumes of data



Demanding technical requirements

- Consistent Internet and government extranet addresses, with unchanged IP configurations
- Migration plans tailored to the needs of each commission or bureau
- A smooth migration, no service interruptions
- Security compliance to meet graded cyber security protection requirements



Tight project timeline

Migration completion deadlines regulated in the government work plan

Solution

The Huawei government cloud migration team ensured a secure, stable, and efficient migration of the Shangluo government cloud platform by leveraging the Cloud Migration Framework (CMF) as well as deterministic O&M capabilities and solutions. The team took a structured approach, consisting of five standard phases (survey and evaluation, planning and design, environment deployment, migration implementation, and migration acceptance) and three key measures to facilitate a seamless transition.

Three measures







1. CMF: a structured and proven approach for stable, secure, and efficient cloud systems

Huawei Cloud Migration Framework (CMF) is a trusted blueprint for driving intelligent digital transformation in cities. It is the crystallization of over 1,000 cloud migration projects and more than 500 global cloud experts. Through scientific planning and resilient design, the CMF systematically breaks down the cloud migration process into seven stages and twelve steps, ensuring a structured and orderly transition to the cloud.

2. Scientific planning: expertise enablement for more efficient migration

The team thoroughly analyzed the system architecture and service requirements of the Shangluo government cloud platform, the technical feasibility, the implementation challenges, and the resource needs. They also considered the live network environment, migration risks, and emergency rollback plans. After multiple rounds of refinement, the team developed a solution that

met the requirements for migration performance, efficiency, security, and reliability.

1) Migration scope

Internet zone and government extranet zones: xx service systems of xx government agencies, involving xx cloud servers and disks

Leased and hosted devices: xx leased devices and xx hosted devices, totaling xx devices

 Migration feasibility assessment Migration policy: Cloud servers would be rehosted. Service reconstruction was not required.

Migration solutions and tools: There were mature migration solutions available, along with the RDA migration tool.

Cloud service compatibility: There were no compatibility issues since the source and the destination were both Huawei Cloud Stack Online (HCSO) environments.

Exceptions:

 The security services were rebuilt on the new platform based on the existing policy. Leased and hosted devices were relocated.

Overall migration principles Careful planning and scientific verification: clearly defined work content, steps, and leadership

coordinated execution: systematic migration with a refined schedule **Security:** optimized backup and rollback mechanisms to ensure data integrity

Phased implementation with

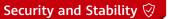
4) Refined migration strategies

Preparations: Network connectivity setup, platform deployment, and migration surveys were conducted simultaneously.

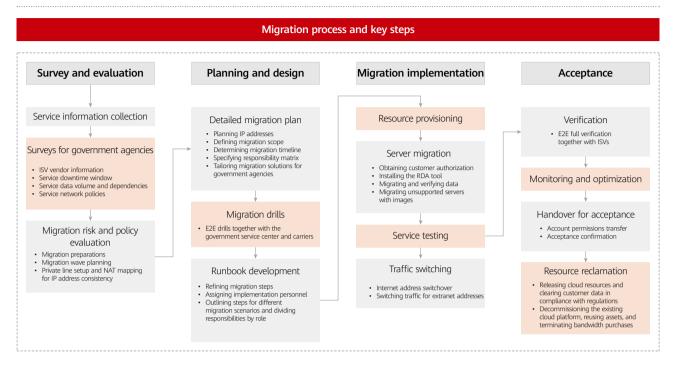
Migration wave planning: Internet zone > government extranet zone > leased and hosted devices

Service continuity measures:

- Private line connections and NAT mapping ensured unchanged server IP addresses.
- The migration scope was refined based on survey findings, with services migrated in waves and the cutover performed in batches, based on IP addresses.



Migration implementation process



3. Technical enablement: combination of experts and tools for service stability, security, and reliability

The migration solution was executed under the technical guidance of experienced cloud migration experts. Peripheral resources were strategically allocated, and key technical risks were proactively addressed to ensure stable and efficient post-migration operations. The following key technologies and tools were used:

1) Private lines and NAT mapping for IP address consistency

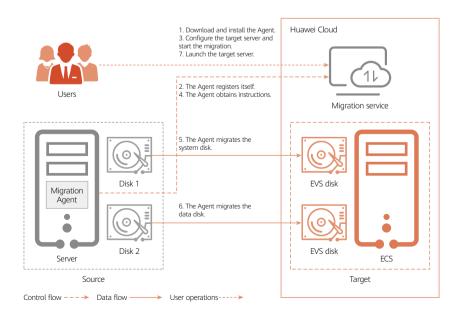
Target servers on the new platform had the same IP addresses as the source servers, so the network configurations and policies did not need to be updated after the migration.

The migration tool RDA was deployed in the VPC of the new platform to maintain connectivity with source and target servers.

2) SMS for secure, streamlined migration

Huawei Cloud SMS simplified the

migration of servers with a three-step process: selecting and configuring source and target servers, followed by confirming task settings to initiate the migration. Users just needed to install the Agent on source servers, configure target servers and start migration, and launch the target servers. SMS handled the actual migration process. Services undergone multiple incremental synchronizations, so there was only brief downtime before the final cutover. After that, services



could be started on target servers immediately. This one-stop migration solution significantly improved migration efficiency while minimizing service downtime.

For security, AK/SK pairs were used to authenticate the migration Agent. SSL encryption secured the transmission channel, ensuring data integrity. Dynamically generated SSL certificates and keys enhanced security and ensured compliance with government cybersecurity requirements.

4. Quality and efficiency improvement: E2E management for effective execution

To meet project deadlines and avoid operational bottlenecks, our migration team proactively identified risks, defined an operation mechanism, and ensured timely issue resolution throughout the migration process. Additionally, a comprehensive cost assessment was conducted to ensure investment

efficiency. Through upstreamdownstream collaboration and E2E management, the project duration was reduced by 40%, ensuring seamless execution.

Risk identification and management						
Risk	Description	Solution				
Migration failure risks	The service systems were highly complex, operations were intricate, and the data volume was substantial. Additionally, the migration had to be successfully completed in a single attempt, significantly increasing the risk of failure.	A kick-off meeting was held before the migration. At the meeting, a lead team and department were designated to coordinate resources and assign rights and responsibilities, ensuring an organized execution.				
Technological risks	Potential data loss during the migration Irrecoverable services if the migration fails	Before the migration, service systems were temporarily halted to minimize risks. The RDA tool was used to push source server details to target servers without modifying service systems. Rollback plans were developed to ensure that service systems could be restored at any time.				
Potential delays	The migration workload was heavy, and the timeline was tight. Timely completion was going to be a challenge.	Huawei migration experts provided onsite training and hands-on support to optimize migration efficiency and quality as well as prevent delays.				
Organizational coordination risks	Difficulties in determining and verifying migration methods	Dedicated customer contacts were assigned to streamline communications. Pre-migration notifications were issued to ensure collaboration between service owners and technical teams for smooth coordination.				

Service Improvement

Smooth migration

There were zero service interruptions during the migration, and services were stable after the migration. There were efficient custom migration solutions for each government agency, which significantly reduced workload across departments.

Infrastructure enhancement

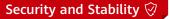
Compliance with graded security requirements and reduced O&M costs support long-term sustainability and development.



Summary

With the successful migration of the final service system from the temporary data center, the Shangluo government cloud migration was completed. The Huawei government cloud migration team achieved this milestone through a proven migration methodology, efficient execution, and extensive cross-industry expertise. All service systems are now stable, providing robust support for the digital transformation of Shangluo and the strategic development of the "Qinling Smart Valley and Digital Shangluo" project.

Source: Huawei Cloud "One City, One Cloud"



Cloud-based Protection Earned Full Marks in a Cybersecurity Drill



Abstract

In a city-level cybersecurity drill in 2024, company W worked with Huawei Cloud to enhance security capabilities through 24/7 monitoring, full-traffic threat detection, and asset sorting. The company lost no points in the drill and won the title "Excellent Defender". The company took this drill as an opportunity to improve routine security operations, shifting from DevOps to DevSecOps.

Service Background

Company W is a top-notch vendor providing efficient international warehouses for businesses of all types.

In the city-level cybersecurity drill in 2024, company W utilized Huawei Cloud's professional security services

to achieve a perfect score and won the title "Excellent Defender", concluding the drill with outstanding results.



Service Status

The service systems of company W are deployed in four regions: the Huawei Cloud, international cloud A, international warehouse, and local IDC. In this cybersecurity drill, the company and Huawei Cloud jointly set up an assurance team to analyze service security conditions.

Company W had deployed the firewall, intrusion prevention system (IPS), antivirus, log audit, and host

security services to protect the infrastructure of core service assets, but the analysis identified areas for improvement in its international cloud A and data centers:

- » The basic security capabilities were weak. Network isolation needed to be improved.
- » The cloud and on-premises networks needed to be divided

- into different zones based on service security levels.
- » Some high-risk ports were exposed to the Internet.

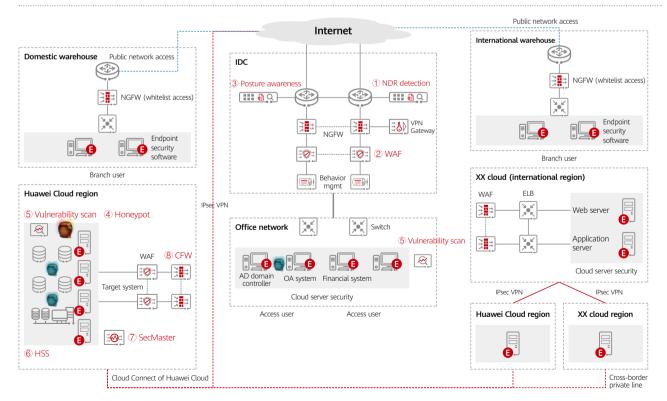
The company aimed to earn full marks in the drill. Although the time between notification and the drill was short, the assurance team managed to resolve all the security issues through research, verification, and rectification.

Solution Implementation

The assurance team for this drill had nine members: two project managers, five senior engineers, a mid-level engineer, and a junior engineer. The team provided 24/7 assurance and developed a full-traffic threat detection solution. The following figure shows the architecture design.

To enhance the network security capabilities of the IDC, we made the following enhancements:

Full-traffic threat detection solution



Global network threat awareness and analysis system

- » To compensate for the limited threat traffic analysis capabilities in remote regions, we deployed the global network threat posture awareness and analysis system in the IDC
- » We strengthened the defenses against east-west scanning and penetration on intranet assets in the event of a server compromise.
- » Web Application Firewall (WAF) was deployed to deliver dynamic protection for application systems, effectively defending against internal and external web application attacks.

Server and border security detection

» We installed the agent on cloud servers and certain on-premises

- servers to achieve comprehensive security coverage, protect servers from intrusions, and report alarms on abnormal behaviors.
- » We enhanced the detection of external attacks targeting Huawei Cloud EIPs and closely monitored abnormal outbound connections initiated by internal servers, which could indicate a possible server compromise.
- » We defended against lateral movement of attacks from a compromised node.
- » A blacklist of up to 50,000 IP addresses was provided to enable quick attack detection and IP blocking.

Honeypot system deployment

Five honeypot systems were deployed across the cloud and on-premises environments to detect and trace

lateral movement attacks within the intranet

Sorting assets for faster response

During the network assurance, the assurance team sorted out assets in advance to provide a clear asset inventory, which helped to quickly locate attacked assets and respond to incidents.

We comprehensively identified and addressed risks from multiple perspectives, including reducing exposure, removing unnecessary assets, improving protection policies, eliminating known risks, deploying honeypots, conducting vulnerability scans, and performing penetration tests. These efforts effectively prevented external attacks, enhanced monitoring capabilities and methods, and improved source tracing and analysis capabilities.

Service Improvement

During the cybersecurity drill, we performed comprehensive security monitoring and analysis on the local IDC.

Our specific measures and their outcomes were as follows:

Traffic monitoring and analysis

- » We performed traffic mirroring to monitor both north-south and east-west traffic, and implemented comprehensive threat analysis.
- » Traffic analysis helped to capture more than 6,000 north-south exceptions and more than 100 east-west exceptions.
- » Core systems such as Office Automation (OA), finance, and domain controller systems were specifically monitored, and more

than 20 abnormal traffic events were recorded.

Web attack detection and blocking

- » We performed serial detection and analysis on north-south web service traffic.
- » Web attack traffic was blocked more than 1,000 times. None of the core systems were intruded.
- » More than 300 attacker IP addresses were blocked

Server protection

- » We provided comprehensive protection for both cloud servers and the servers in international data centers.
- » Trojans, reverse proxies, and other forms of malware triggered

- alarms immediately, and were blocked and isolated upon detection.
- » More than 50 abnormal behavior alarms were reported, but all were confirmed to be normal service behaviors.

Intrusion prevention and virus detection

- » IPS and antivirus provided thorough protection.
- » More than 20 attack EIPs were blocked, and more than 400 abnormal network behaviors were detected.

Web attack detection and blocking

» We performed serial detection and analysis on north-south web

- service traffic of the target systems on the cloud.
- » Web attack traffic was blocked more than 5,000 times, protecting core systems from intrusions.
- » More than 20 attacker IP addresses were blocked.

Network honeypot deployment

- » Network honeypots were used to lure and capture attackers, analyze their behaviors, and report countermeasures.
- » In this cybersecurity drill, the honeypot system did not report

critical alarm events, and no horizontal scanning was detected on the intranet.

These measures ensured the security and stability of the network, effectively defending against a wide range of attacks and protecting core service systems.

Case Summary

This cybersecurity drill was an important validation of our regular security operations. During the drill, the security team's primary task was to implement 24/7 monitoring and emergency response, rapidly detecting and handling security threats. To meet the assurance requirements of the drill, we developed a detailed plan consisting of three phases: prevention, response, and review. We also assigned assurance personnel to work on-site in three groups: attack monitoring, analysis and investigation, and source tracing and incident response. The groups were on duty in two shifts around the clock. The following key points were gathered from this drill:

- Routine security operations and risk governance are critical: Deploying security devices, fixing vulnerabilities, and improving policies require significant time, and are not one-time efforts. Routine operations can significantly improve security, helping to prepare for key events such as cybersecurity drills.
- Asset exposure management: Reducing the asset and interface exposure to the Internet can help to mitigate risks. All the exposed assets must be protected and monitored by security devices.
- » Vulnerability scan and penetration test: Periodic vulnerability scans and penetration tests help to detect

vulnerabilities, such as improper service logic and information leakage. Conduct in-depth penetration tests as needed to proactively identify vulnerabilities.

In this cybersecurity drill, we found that many service systems were running outdated versions and had Fastjson, SQL injection, and XSS vulnerabilities, indicating a lack of security personnel involvement, security development standards, and system security development tools. Based on these findings, we plan to systematically shift security left and transition from the DevOps process to DevSecOps, enhancing security awareness and practices in the development phase.





Resource Efficiency

Redefining Cost Engineering from Multiple Dimensions: Resources, Applications, and Operations



Qingting FM: Application Modernization Is a New Industry Chapter

Abstract

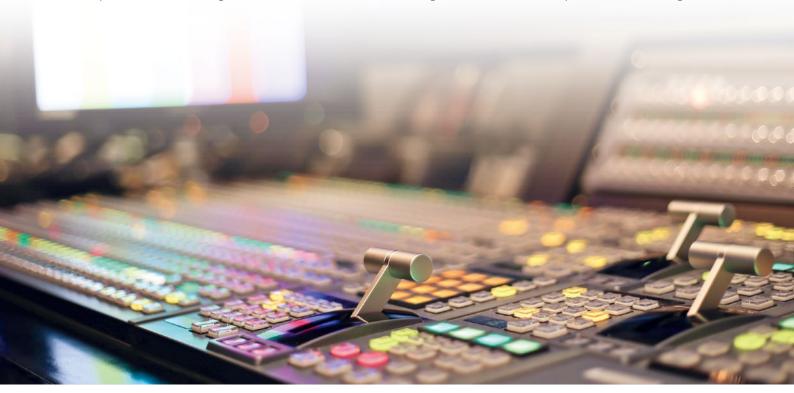
Qingting FM used Huawei Cloud application modernization to address the challenges in technical architecture iteration. It is a story of how containerized infrastructure, optimized service architecture, and technology stack upgrade help reduce costs, improve efficiency, and optimize architectures and technologies.

Service Background

Fueled by its massive market, intricate ecosystem, diverse trends, and dense geographic footprint, the audio-video industry is a pivotal internet segment with substantial growth prospects. However, that growth causes escalating technical debt, a major impediment to more progress. Initial speed-to-market strategies favored

basic architectures, overlooking scalability for high concurrency and quality, but this hindered system scalability after business expansion. Frequent technology iterations worsened technical lag and compatibility challenges. The data explosion strained early data architectures, leading to inefficient

processing, high costs, and security vulnerabilities. Expanding infrastructure complicated operations, making manual maintenance costly and inefficient due to a lack of automation. As a result, technical debt was further compounded as diversified business spurred different technology and frequent architecture changes.





Service status

As China's first online audio application. Qingting FM promises its listeners that "You can hear a better Life". Its collaborative ecosystem for users and content producers brings together high-quality audio IP from radio and copyrighted sources. It has 100 million subscribers, with 100 million monthly active users and millions of daily active users. The content spans from culture, finance, technology, music, to audiobooks. However, when current technical architecture and development approach could no longer support the platform's expanding business modes and rapid user growth, Qingting FM needed an immediate upgrade. Qingting FM faced the following challenges:



Poor application scalability

Application routing was tied to scheduling at the VM level, so new VMs needed rescheduling.



Long version release cycles

Version iteration methods, capabilities, and tools needed upgrading



Faster troubleshooting

With a large cluster hosting multiple channels per VM, diagnosing issues required first pinpointing the affected VM, followed by a cumbersome investigation, resulting in slow fault resolution and potential cascading impacts. Slow fault isolation for individual radio channels



Poor resource utilization

Direct VM calls to heterogeneous resources made host resource allocation inefficient.

Solution Implementation

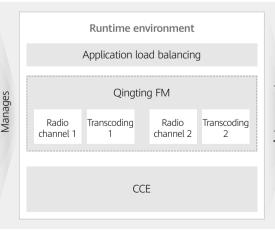
Addressing current business challenges, Qingting FM focuses on business innovation, utilizing Huawei Cloud's application modernization including ServiceStage, CCE, LTS, and AOM – to upgrade its architecture.

The overall solution migrates applications to ServiceStage, which manages the running environment and application lifecycle. Specific applications include containerization, cloud-hosted application runtime environment, CI/CD automatic deployment, auto scaling, unified log collection and analysis, and monitoring and alarms.

Modernization of Qingting FM









1. Architecture reconstruction

- » Containerization: Establishment of a foundation for scalability with over 10 systems for granular resource scheduling
- » Communication optimization: Transition from static IPs to service discovery
- » Stateless migration: Decoupling radio station creation from firmware VMs

Before:

- 1) Manual maintenance of Transcode and SRS node IP lists
- 2) Transcode and SRS could only be scaled up to add radio channels due to scale-out complexities.

3) Costly maintenance of dynamic retrieval of SRS livestream addresses

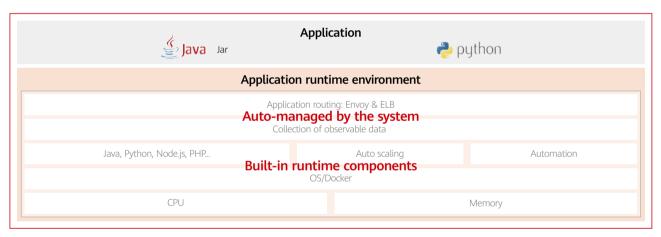
After:

- 1) Radio channel allocation is decoupled from fixed nodes, live streams are no longer dependent on static SRS IPs, and hardcoded IP communication between components is eliminated.
- 2) Operations and maintenance are streamlined and efficient, with comprehensive monitoring.
- 3) Resources are efficiently utilized, allowing for rapid, on-demand scaling.

2. Cloud-hosted runtime environment

ServiceStage is a managed runtime environment with integrated container runtime (CCE), elastic scaling, and application routing. Radio channel application container images run within this environment, allowing for flexible scaling and ensuring channel isolation.

ServiceStage supports the management of computing, network, and diverse platform services. By shielding underlying technical complexities, its streamlined and optimized interface lowers the barrier to entry and boosts development and operations efficiency.



3-step deployment: environment creation, application deployment, route advertisement

Solution:

- ① Fully managed running environment, multi-format deployment, auto connection to observability services
- 2 Automated OS and software updates with zero business disruption

Value:

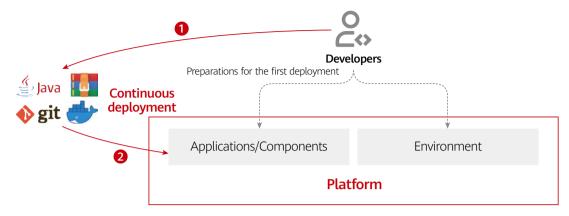
- Cloud hosting and streamlined application deployment (3-step)
- Automated OS kernel and other software updates with zero downtime and transparent user experience

3. Automatic CI/CD deployment

ServiceStage offers integrated continuous deployment. Qingting

FM simply provided the code and enabled direct release to test and production environments, eliminating the overhead of setting up Jenkins and managing base images. In addition, ServiceStage automatic build of common technology stacks from source code to images means users do not need to compile Docker files. This further facilitates getting started and quick containerization.





Preparations (developers):

Create an environment

Create an application or a component, choose a code repository and a branch, enable auto-deployment and complete the release



Continuous deployment (developers):

① Commit code and merge it to the release branch 2 Auto-deployment

4. Auto scaling

Qingting FM's CPU utilization dynamically adjusts based on the number of connected users. The variable concurrency across channels is set up with a scheduling mechanism with a maximum CPU limit per channel and CPU preemption. Exceeding the maximum CPU triggers instance scale-out, optimizing resource utilization by 20%.

5. Unified log collection, analysis, monitoring, and alarm reporting

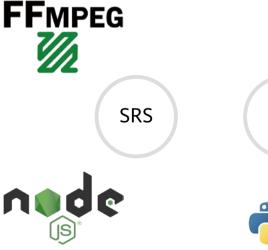
During application deployment, ServiceStage automatically connects application logs and metrics to Huawei Cloud LTS and AOM. Logs for individual radio channels are easily viewed, and structured analysis enables the extraction of business metrics for real-time operation monitoring.

Compared with AOM and LTS, viewing monitoring details and logs on ServiceStage is more user-friendly and more efficient.

6. Technical upgrade: boosting system performance

Initial research identified outdated performance and security of the legacy radio channel business technology stack, framework, and third-party dependencies. The following enhancements addressed this issue:

- Upgraded the runtime SDK to the most secure version.
- Upgraded FFMPEG, the core thirdparty component, to improve transcoding performance.
- » Upgraded the open-source live streaming components SRS and ICECast and built updated images.
- Refactored the business to decommission the Ruby technology stack to minimize future maintenance and learning costs.









Business Improvement

The radio channel partnered with Huawei Cloud to architect and implement business containerization, hosting on ServiceStage for modernization, lower costs, and better R&D and operations efficiency. This core business secured a first-mover advantage in a competitive market. Key business enhancements include:

1. Enhanced scalability

» Radio channel allocation is decoupled from fixed VMs, eliminating hardcoded IP

- communication and enabling stateless services.
- » The robust scalability of the radio channel system allows seamless expansion of channel count without business being affected by elastic resource scaling.

2. Improved O&M and deployment efficiency

» Automatic access to AOM provides streamlined and efficient monitoring, accelerating problem identification by 50%. » Service Stage improves service deployment efficiency. Rapid issue pinpointing via the LTS console doubles O&M efficiency.

3. Cost optimization

» Containerization and resource sharing optimize resource utilization. The platform's ability to rapidly scale resources in response to business needs results in a 20% reduction in platform costs.

Summary

We faced challenges stemming from the misalignment between our growth and technical architecture and development model. To overcome these, we collaborated with Huawei Cloud for a transformative upgrade leveraging application modernization professional services. Through containerization design and implementation, architecture optimization, and related measures, we bolstered scalability, enhanced operations and deployment efficiency, and optimized costs. This strategic partnership not only resolved existing technical hurdles but also laid a resilient groundwork for future business advancement.





Big Data Migration Enhances Cost Efficiency, Multi-Cloud Flexibility, and SLA Compliance in Finance



Abstract

This case focuses on the big data migration project of a financial company, including the migration survey and process, and how the compatibility with built-in functions was resolved. It also analyzes how the cloud migration solution helped the company improve migration efficiency, reduce costs, and achieve flexible multi-cloud deployment.

Background

Quark International Holdings is one of Asia's fastest-growing financial technology groups. It is dedicated to developing mobile financial applications that allow people to intuitively find financial solutions that comply with regulatory rules. Quark used to run their workloads on different clouds. To enhance cost efficiency, multi-cloud flexibility, and SLA compliance, Quark worked with Huawei Cloud to migrate workloads from Holo, MC, and DataWorks on another cloud to GaussDB(DWS), Data Lake Insight (DLI), and DataArts Studio on Huawei Cloud.

Product mapping











Challenges

There are four challenges to big data migration:



High requirements for data accuracy

The company's business covers core data fields such as financial transactions, and data must be 100% accurate.



A large number of tasks

There were more than 6,000 logarithm tables to be migrated, and more than 5,000 tasks were required.



First application of the Kyuubi component in migration

Kyuubi was used for data parsing in migration for the first time.



Compatibility

Adaptation to the SQL syntax of BI applications needed to be addressed.

Common issue adaptation

No.	Database Type	Source Expression	Huawei Cloud Expression	Description
1	GaussDB(DWS)	UNIQ(expr)	COUNT(DISTINCT COLUMN···)	COUNT(DISTINCT COLUMN) for row-store tables, and UNIQ() for column-store tables
2	DLI	Select transform(for i in seq 1 1825; do echo \$i; done) using 'sh'as (num)	Select explode(sequence(1,1825)) as num	The Transform() function must be converted to explode().
3	DLI	Partitioning field dt is not null	Partitioning field dt= or dt!=	Dt is used as a partitioning field. It must be converted to "=" or "!=".
4	GaussDB(DWS)	DATEADD (date, number, datepart)	ADDDATE (date, INTERVAL val unit)	DATEADD must be converted to ADDDATE. unit indicates the unit of the difference between two dates.
5	GaussDB(DWS)	DATEDIFF (end_date, start_date)	TIMESTAMPDIFF(unit, timestamp_expr1, timestamp_expr2)	DATEDIFF must be converted to TIMESTAMPDIFF. unit indicates the unit of the difference between two dates.
6	GaussDB(DWS)	:: TIMESTAMPTZ	:: TIMESTAMP	
7	GaussDB(DWS)	COUNT(*) FILTER (WHERE)	SUM(CASE WHEN)	Conditional statistics must be converted to SUN(CASE WHEN) or other syntax.
8	DLI	TOLOWER(string)	LOWER(string)	
9	DLI	Chinese comma ", "	English comma ","	Chinese commas must be converted to English commas.
10	DLI	Chinese parentheses " () "	English parentheses "()"	Chinese parentheses must be converted to English parentheses.



Solution

Huawei Cloud's extensive experience in big data migration projects led to the creation of both a compatibility survey and a mature process covering migration, deployment, and cutover. A detailed solution, as well as technical support, was available in each phase to keep project implementation efficient.

1. Preparations

- Network connectivity: Establish a VPN between the source cloud and Huawei Cloud for data synchronization.
- Job layer sorting: Sort out the tables to be migrated and the corresponding job, laying a foundation for migration, data verification, and recovery.

2. Platform deployment

Cloud instance deployment: Subscribe to Data Lake Insight (DLI), DataArts Studio, Cloud Data Migration (CDM), GaussDB(DWS), and Object Storage Service (OBS).

The source service architecture consists of three parts:

Databases: RDS for MySQL/

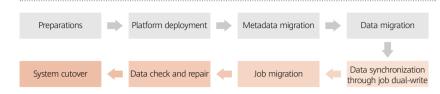
PostgreSQL.

Data platform: Holo and MC from the source cloud provide real-time data warehouses for computing and storage. DataWorks functions as a big data platform.

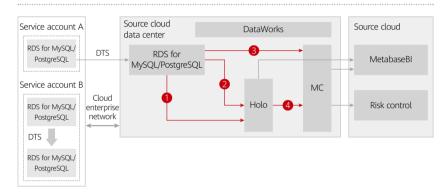
Applications: MetabaseBI, a cloudbased business intelligence tool for risk control

Among the three parts, retain databases and applications on the source platform, migrate the data platform and real-time data warehouses from Holo to Huawei

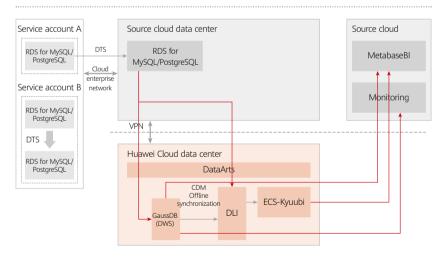
Overall migration process



Cloud instance deployment and source architecture



Final architecture



Cloud GaussDB(DWS), and migrate computing and storage workloads from MC to Huawei Cloud DLI. DLI can be used for the analysis of massive amounts of logs, the federated analysis of heterogeneous data sources, and big data extract, transform, and load (ETL). It has advantages including pure SQL operations, decoupled storage and compute, and enterprise-class multitenancy. Workloads on DataWorks should be migrated to DataArts Studio.

As a data governance center, DataArts Studio provides a one-stop data development environment in which users can complete data processing and analysis. This includes obtaining, cleaning, transforming, analyzing, mining, and visualizing data.

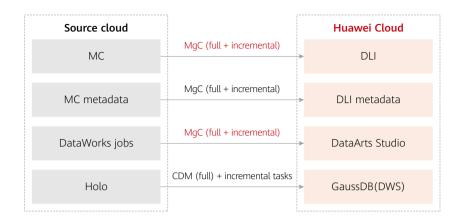
- » Adjustment of cluster parameter definitions and the source framework: Compile, debug, and deploy user-defined functions (UDFs).
- » Building and connecting with on-premises services: Deploy and commission jobs in DataArts Studio. Connect DataArts Studio with the data lake, and connect Kyuubi with DLI. The following figure shows the deployment, connection, and migration tool.

MC data migration

- » Historical data: Configure one or more jobs in the MgC service for each database table to migrate historical data.
- » T-1 incremental data: Query changed tables from the MC metadata database using MgC, and configure Spark jobs to migrate the table data.

MC metadata migration

» Use MgC to export table creation statements at the source, convert the statements, and import them



to the DLI metadata database on Huawei Cloud.

» Use MgC to compare metadata differences and synchronize incremental metadata.

DataWorks job migration

- » Use the DataWorks migration assistant to export submitted jobs, use MgC to convert the jobs into those suitable for DataArts Studio, and import them to DataArts Studio.
- » The DataWorks O&M center identifies the jobs modified and submitted every day, and synchronizes them to DataArts Studio.

Holo data migration

- » Migrate all historical data using CDM.
- » Incremental data is generated

from on-premises Hive data and imported to GaussDB(DWS) by a scheduled task every day.

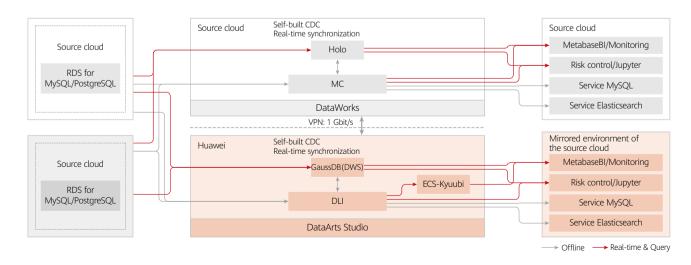
3. Metadata migration

Migration and table creation: Use the metadata migration tool to export the table creation statements of MC, which are used to create tables in Huawei Cloud DLI.

4. Data migration

- » Inventory data migration: Migrate inventory data to DLI using MgC.
- » Dual-write of incremental data: Write all types of data in the source cloud to both MC and Huawei Cloud DLI.

The following figure shows the parallel running structure.





In big data migration, the parallel running architecture ensures data security and continuity. In this architecture, data is stored both in an on-premises environment and a cloud. This prevents data loss or service interruptions caused by a fault. If the on-premises environment is faulty, enterprises can still obtain data and services from the cloud environment without service interruptions.

Data warehouse job layers: ODS, DWD, DWS, and ADS. At these layers, jobs can be scheduled, executed, and migrated. The job execution SLA is met.

6. Data check and repair

- Use a comparison tool to compare data in batches by file or service.
- Use CDM or MqC to re-migrate data for inconsistent data partitions.

7. Traffic switchover and system cutover

- Run big data services for more than one week and check the results.
- Connect data applications to Huawei Cloud big data clusters.
- Disconnect the data source from the source cloud data platform. Huawei Cloud data platform pushes data and provides services.

5. Job migration

Operation Data Store (ODS)

- It stores raw data.
- It cleanses and transforms data from service systems. For example, it checks the raw data and partitioned table settings, such as the processing of frontend tracing point logs.
- Generally, data at the ODS layer is only processed simply to ensure data integrity and accuracy.

Data Warehouse Detail (DWD)

- · It stores cleansed and transformed data.
- · It processes data from the ODS laver. For example, it cleanses core fields and degenerates dimensions, such as product table dimensions.

Data Warehouse Summary (DWS)

- · It stores aggregated and summarized data.
- It summarizes and aggregates data from the DWD layer. For example, it creates wide tables and analyzes table composition and metrics. such as the fields in user

behavior wide tables

Application Data Store (ADS)

- · It provides data services and supports service applications.
- It processes data from the DWS layer. For example, it checks and analyzes the accuracy and integrity of metrics, such as daily active users and new users

Service Improvement

In this migration project, service improvement is reflected in the following aspects:

- Efficient migration: Full, incremental, and scheduled migration, as well as real-time synchronization, was used to securely migrate data from the source system to the destination system, without compromising on data integrity or consistency. In this project, the one-time success rate of the data migration tool was over 95%.
- The project duration was shortened by about 15.22%, saving about 36.35% of total person-days.
- Compatibility assurance: The new environment was fully tested and evaluated before the migration. This was to ensure that it was compatible with the existing service system and that compatibility risks were controllable.
- Performance assurance: Before the migration, the performance of the new environment was tested and
- optimized to ensure that performance requirements were met. During the migration, close attention was paid to system performance indicators, such as the response time and throughput, to resolve performance problems in a timely manner.
- Multiple clouds: Cloud service providers can be selected on demand for flexible data deployment and management.

Summary

To enhance cost efficiency, multicloud flexibility, and SLA compliance, Quark migrated its big data platform to Huawei Cloud. Huawei Cloud's

extensive experience in big data migration projects helped Quark successfully migrate its big data platform through a mature process.

This included preparations, platform deployment, and migration of metadata, data, and jobs.

Refined Management with Landing Zone for a Secure Cloud Ecosystem

Abstract

This case describes how company Z used Huawei Cloud Landing Zone professional services to create a well-architected, secure, compliant, and scalable multi-account environment. Based on this environment, the company established a lean IT governance system to manage people, finance, resources, permissions, and security compliance in an organized and unified manner.

Background

Company Z is committed to becoming one of the most innovative Chinese agricultural science and technology companies in the world. In today's allcloud era, they decided to use Huawei Cloud Landing Zone professional services to optimize management, ensure security compliance, and improve cost-effectiveness. Landing Zone helped this company set up a multi-account management structure, improve permissions management policies, and optimize financial management processes. The

enhanced resource utilization and strengthened system security made it easier for company Z to securely and stably migrate their business to the cloud and efficiently complete digital transformation.





Service Status

Company Z faced the following challenges in digital transformation:



Centralized Account Management System



Strict Permissions Management Policies



A Stable High-Performance Network Architecture

A centralized account management system needs to be established to centrally manage cloud accounts of all business departments and subsidiaries, so as to ensure proper resource allocation and efficient resource utilization.

Strict policies are required to manage user permissions based on their roles and responsibilities while always enforcing the principle of least privilege (PoLP).

The network architecture needs to deliver high performance and stability to maintain low latency and high bandwidth during peak hours, or service continuity and user experience could be impacted.



Refined Financial Management Planning



Multi-layer Security

The financial management of its subsidiaries is scattered and needs to be centralized and optimized to improve management efficiency.

Many aspects, including boundaries, applications, and data, need to be protected.

Solution

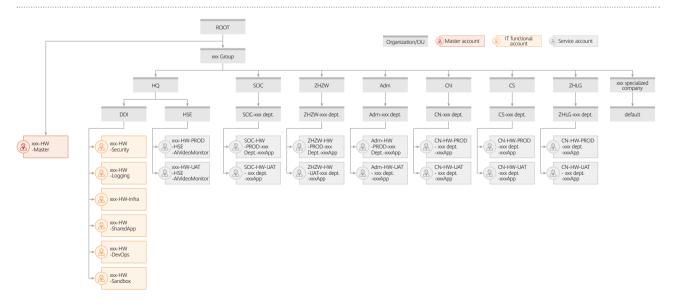
The core capabilities provided by Huawei Cloud Landing Zone professional services have accelerated their cloudbased digital transformation. The details are as follows:

1. Organization Accounts

Landing Zone provides multi-account management capabilities that can be adapted to specific company requirements. Organizational units

(OUs) and accounts are designed for different dimensions such as the business architecture, operating environment, and IT functions. The account design is as follows:

Organization Account Framework



The company set up an organizational structure based on functions and businesses. They created one functional OU and multiple subsidiary OUs. The functional OU is managed by corporate personnel. It consists of child OUs for DDI and HR departments. The DDI OU has a security account, logging account, infrastructure account, public account, pipeline account, and sandbox account. Each subsidiary OU is managed by a subsidiary account.

2. Identities and Permissions

The Huawei Cloud IAM Identity Center enables service users to interconnect with third-party identity sources, create users and groups, assign login passwords, and manage user groups. The identities and permissions are designed as follows:

» Identity and Access Management

Multi-source identity authentication: Some employees use a third-party identity provider (IdP) for single sign-on (SSO), and others use IAM Identity Center to log in.

Identity federation: IAM Identity Center is federated with third-party IdPs, which allows users from both identity sources to seamlessly access the system.

Identity synchronization: User information is periodically synchronized from third-party IdPs to IAM Identity Center to ensure user data consistency and accuracy.

» Permissions Management

- Centralized permissions management: IAM Identity Center centrally manages all user permissions, ensuring consistent and secure permission assignment.
- Role-based permission assignment: Permissions are assigned based on employee roles and responsibilities,

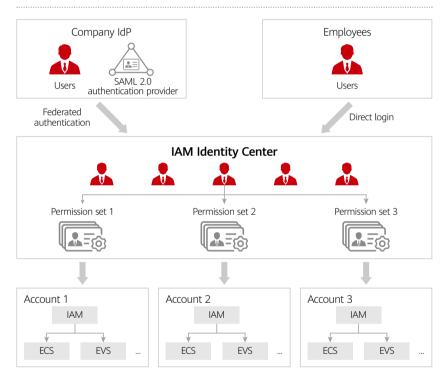
- which simplifies permissions management.
- Fine-grained permission control: Fine-grained permissions are configured for different resources and operations by following the PoLP.

3. Platform Deployment

Cloud instance deployment: Cloud services such as Data Lake Insight (DLI), DataArts Studio, Cloud Data Migration (CDM), GaussDB(DWS), and Object Storage Service (OBS) are enabled.

The source-end service architecture consists of three parts.

Permissions Management Planning



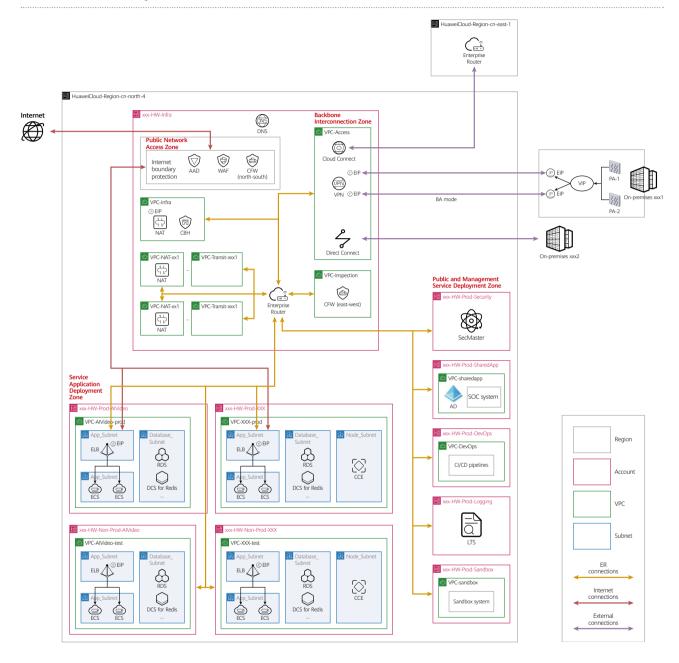
4. Network Architecture

A company's network architecture is crucial. The company used Landing Zone to build a stable, reliable, and scalable network infrastructure. This infrastructure enabled central

management of various network resources and network security resources. It also helps ensure secure and reliable communications between service accounts and IT functional accounts. The network architecture is as follows:



Network Architecture Design



Backbone Interconnection Zone

- VPCs of service accounts and IT functional accounts are connected.
- · On-premises data centers and offices are connected to Huawei
- · Third-party clouds are connected to Huawei Cloud.
- Service systems deployed in multiple Huawei Cloud regions are connected.
- East-west firewalls are used to control access for east-west
- Service Application Deployment Zone
- laaS and PaaS resources are provided for service systems.
- Isolated VPCs are provided for the different service systems under a service account.
- Each VPC is divided into application and database subnets based on the application architecture.

» Public and Management Service Deployment Zone

- Public services like VM image repositories and container image repositories are deployed.
- A centralized security operations management system is established.
- A centralized O&M monitoring system and patch update servers are deployed.
- Centralized DevOps pipelines are provided.

» Public Network Access Zone

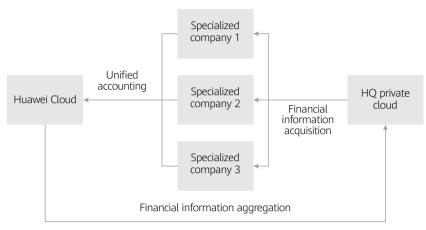
- Internet traffic is monitored in real time so that DDoS attacks can be detected and scrubbed in a timely manner.
- A north-south firewall is available for access control, intrusion prevention, traffic analysis, and log auditing.
- Web applications are protected against application-layer attacks, such as SQL injection and XSS, from the Internet.
- Public network access resources are deployed in service accounts.
 EIPs and ELBs are centrally deployed and managed on boundary subnets of VPCs in service accounts for Internet access.
- Internet ingress is managed by the boundary subnets of VPCs in service accounts.

5. Financial Management

Currently, company Z uses unified accounting on Huawei Cloud. It calls Huawei Cloud APIs to perform secondary development and centrally manage billing, ensuring accurate splitting and management of financial data.

The administrator account pays for the cloud resources used by the member accounts of the organization. Using Huawei Cloud APIs to perform secondary development, the company can control and aggregate its financial Cloud Financial Management Solution

Unified Financial Management Solution



Huawei Cloud API calling

details. Its subsidiaries can obtain bill information from the private cloud for accurate cost splitting and management.

6. Security

The company sets up a dedicated security system to protect user accounts, networks, hosts, applications, and data based on its own security status and the Huawei Cloud security system.

The security system is as follows:

» Functional OUs

- Data security: Data can be encrypted, backed up, and restored to ensure data integrity and confidentiality.
- Host security: Host security configurations, vulnerability scans, and patches are managed to prevent malware and unauthorized access.
- Network security: Network firewalls, anti-DDoS, and security groups are configured and managed to ensure network

- stability and security.
- SecMaster: Real-time monitoring and response to security events are achieved through intelligent security analysis and threat detection, improving security intelligence.

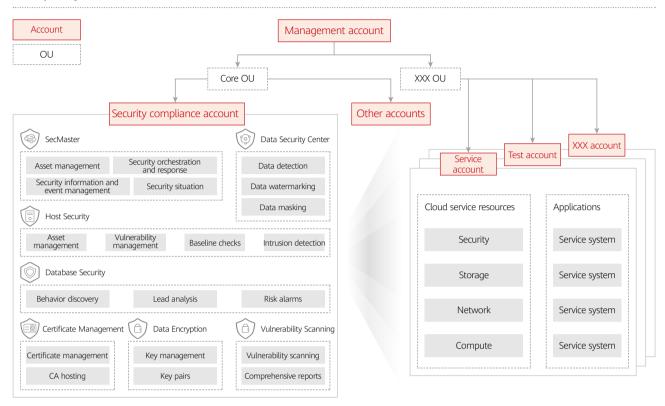
» Service OUs

- Service security: Service systems are monitored so that exceptions can be detected and handled in a timely manner to ensure service continuity and stability.
- Application security: Security tests, web application firewalls (WAFs), and vulnerability scans are used to ensure application security and reliability.

Hierarchical management clarifies responsibilities and improves security management and response efficiency to protect user accounts, networks, hosts, applications, and data on the cloud.



Security Design



Service Improvement

Huawei Cloud Landing Zone provides a secure, compliant, and efficient operations environment where company Z can manage finances, costs, and authorization more efficiently.

- Security events eliminated: Unified account management has reduced monthly account leaks from 5 to 0.
- Authorization efficiency 10 times higher: Unified authorization management eliminates the need to authorize each account for login, shortening the authorization time from 20 minutes to 2 minutes.
- Network resource utilization 30% higher: Idle resources have been significantly reduced. Network resource utilization was brought up from 60% to 90%.
- Labor cost 75% lower: Financial management reduces repetitive work, reducing labor costs by
- Security events 80% lower: The Huawei Cloud dedicated security system has reduced the number of attacks by 80%.

Summary

Huawei Cloud Landing Zone professional services helped company Z address five challenges in cloud resource management. The company

centrally manages accounts, simplifies O&M, manages finances in a refined manner, and performs automated compliance checks. This enhances its management efficiency, reduces O&M costs and security risks, and ensures the security of data assets.



Service Agility

Tackling Service Challenges with Digital SRE: Fast Iteration and Scalability





A New Journey Driven by Intelligent R&D



Abstract

Facing stringent market demands for product quality, innovation, and delivery speed, an intelligent vehicle company leveraged CodeArts to establish a tailored agile development process. Thanks to ASPICE standards compliance, Huawei Cloud's R&D toolchain empowered an efficient and flexible R&D team to boost project delivery period, product quality, and software R&D efficiency.

Service Background

S is an intelligent automobile enterprise and offers digital map, navigation, and location services. Its traditional R&D models and processes struggled to catch up with changing market requirements, so S introduced Huawei Cloud CodeArts to reshape their R&D with agile practices. S not only accelerated product delivery but also improved product quality, laying a solid foundation for sustainability.





Service status

It faced six challenges:



Requirement management

Departments and subsidiaries needed integrated requirement management tools. There was inconsistent granularity of requirement breakdowns across teams and lack of unified methodology for design and review.



Process control

The current waterfall development model resulted in variable development and test cycles. Long development iterations make process risk management challenging in the face of increasing business demands. Furthermore, there lacked a standardized software development process with clear input and output requirements.



Configurations

Siloed development of version control tools across organizations led to inconsistencies, increasing the challenges of centralized management. Furthermore, the lack of a unified code repository complicated asset management.



Build process

The absence of unified build environment management compromised the quality of build artifacts. The lack of a standardized build tool resulted in output inconsistencies. Finally, the lack of standardized build process management, with numerous developers performing independent builds, elevated version control risks.



Software deployment

Manual software deployment is inefficient. The absence of standardized deployment procedures, such as deployments bypassing checklists, elevated live network risks.



R&D process tracing

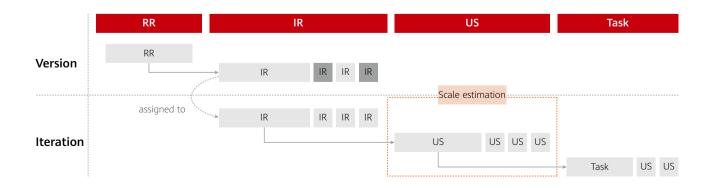
The lack of unified software R&D toolchain and process standards made it difficult to manage R&D data and trace in-process and final deliveries.

Solution Implementation

Enterprise S adopted CodeArts combined with agile development to integrate R&D methodologies with software R&D toolchains for a feasible solution for its anchor products. It developed in five directions:

1. Structured software requirements

A coordinated and consistent multilayer requirement management system (RR-IR-US) specified the responsibilities, management processes, and quality standards of product lifecycle management (PLM), R&D, and functional domains.





 Establish the Program Increment (PI) version planning mechanism at the Release Requirement (RR) - Iteration Requirement (IR) layer and make customeroriented release plans.

Establish a PI version planning at the RR - IR layer to decide the appropriate process for the release plans. Integrate requirements with the enterprise product roadmap by strategically planning PI versions to define deliverable features within specific timeframes. For example, a specific PI version based on market research and customer feedback might include adding real-time traffic updates and optimizing map display for a car navigation system to meet customer needs for accuracy and convenience.

 Establish an iterative continuous delivery mechanism at the IR-US layer for better R&D adaptability and faster closure of more valuable requirements.

This aims to improve the team's R&D adaptability. More

valuable requirements are often prioritized during iterations. For example, when developing car digital map features, the map search function improves value (e.g., quicker and more accurate search potentially increasing product competitiveness), so this feature will be prioritized in the iteration process. Therefore, users experience the most valuable features as the product iterates and adapts to market feedback.

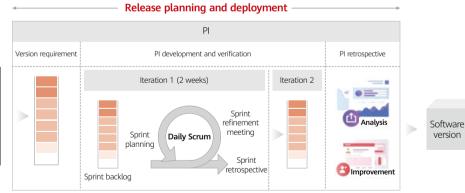
3) Establish a scale estimation at the US layer to improve the team's understanding of requirements, eliminate requirement defects, improve requirement certainty, and streamline the statistics of team and overall throughput.

Such a mechanism calibrates the team's understanding of requirements, eliminate requirement defects, and improve requirement certainty. In practice, the team analyzes the workload, technical difficulties, and service logic of each user story (quicker and more accurate search) in detail to estimate the scale. In this way, all teams reduce development issues by aligning in requirements. Besides, resources are optimized with the throughput statistics overall and from each team. For example, the number and scale of user stories completed by each team can help them better understand the efficiency and progress of the entire R&D process and thus optimize their resources.

2. Standardize the R&D process

Standardize the version operation to fit the original R&D process. Optimize the process with established practices such as product-level agile PI version planning, agile sprint, daily Scrum meetings, Sprint refinement meeting, and sprint retrospective to standardize the input and output requirements and the operation mechanism of one version every month and one sprint every two weeks.

Unified entry point (product backlog): continuous prioritization Leature tednicement Mou-functional reduirement alignment alignment areagony distribution Application of the product of

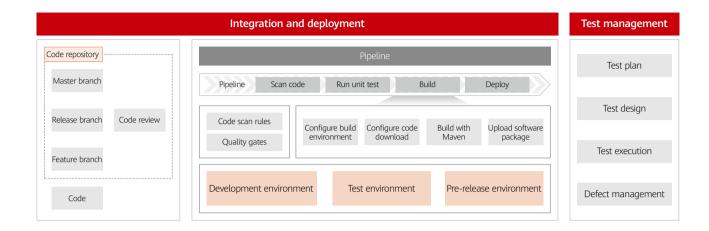


3. Standardize engineering practice

Design solutions and specifications for the engineering practice are fourfold:

- Software repository: Unify code branch management specifications to standardize the code submission process.
- Software build: Unify the build environment and provide suggestions on build environment management.
- **3) Test management:** Standardize the test process and introduce defect management specifications.
- **4) Environment management:** Clarify the environment management solution and requirements in each phase.

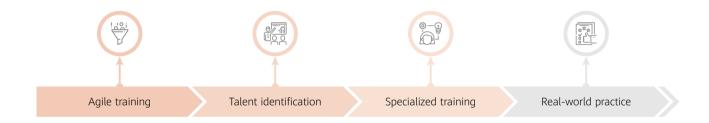




4. Cultivate agile coaches

This has four steps:

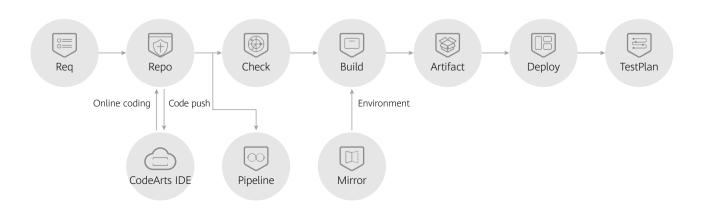
- 1) Talent identification: Identify potential candidates through nominations or applications, followed by a formal organizational
- announcement to clarify goals and responsibilities.
- **2) Specialized training:** Design tailored training programs to individual roles, equipping participants with foundational practical skills aligned with Aqile principles.
- 3) Dedicated coaching: Receive
- ongoing guidance throughout the process, adapting coaching and training plans based on individual progress.
- **4) Real-world practice:** Train within the team and continuously summarize best practices into lessons for future improvement.



5. Unify R&D toolchain

By having an end-to-end software

R&D toolchain from requirements to tests, this enables cross-team requirement distribution and association between requirements, code, and test cases for tracking software iterations and changes.















Build



Artifact types:

Conan, NuGet,

Generic









Req

Jira work item data Git-based MR and import collaboration

> Code export from Gitl ab and SVN

Repo

Check

Open-source engines: SonarQube, ESLint, Pylint, Cppcheck

Artifact

Tools: Maven, Maven, npm, Go, PyPl, RPM, Debian, Grunt, npm, Go

Deploy/ Release

Host- and container-based deployment: ArGo and Ansible

TestPlan

More functions integrated than Postman, xMind, and JMeter

Pipeline

Jenkins tasks

Board

Analysis (>200 indicators) compatible with third-party data

Service Improvement

Building on Huawei Cloud's CodeArts experience, company S has greatly improved its R&D capabilities in iteration, requirement, compilation, and deployment.

1. Requirement management improvement

S unified its requirement management specifications. A standardized document

on requirement specifications guides R&D personnel to directly participate in development. Through standard requirement breakdown and workload evaluation, the iteration plan is more reasonable, the team management level is greatly improved, and the requirement rate of a single iteration is improved by 30% with monthly version release based on two-week iteration.

2. R&D toolchain optimization

R&D toolchain: The unified R&D toolchain facilitates process management. The toolchain comes with code scanning that helps improve software quality by 50%.

Automatic build, deployment, and rollout pipeline boost software release efficiency by 50%.

Summary

Enterprise S leverages Huawei's R&D toolchain CodeArts into its own agile R&D process, addressing challenges such as requirement management, process control, and incomplete configuration. Through this agile transformation project, S has built an efficient and flexible R&D team, significantly shortening the product delivery period and improving

product quality to cope with changing market requirements and laid a solid foundation for sustainable service development.

Intelligent Operations

Unleashing Cloud Value to Drive Business Success



Al Powers a Smart Revolution in AGV Efficiency and Service Excellence

Abstract

A mobile robotics company has partnered with Huawei Cloud to create an Al brain for the automated guided vehicle (AGV) industry. By leveraging AI customer service and AI data insight applications, they have achieved rapid fault diagnosis, interactive data analysis, and real-time human-machine interaction. This effectively lowers after-sales costs, improves fault localization efficiency, and enhances user experience, boosting productivity across the enterprise.

Background

Company A is at the forefront of mobile robot technology research and innovation. As a high-tech enterprise integrating the design, manufacturing, and sales of AGVs and AGV systems, its non-reflective laser guidance technology stands unmatched in China, successfully disrupting longstanding foreign market monopolies. However, with over 1,200 AGVs operating online across 600+ projects globally, the company faces challenges such as geographically dispersed factories, difficulties in fault localization, and high maintenance costs. To address these issues.

Company A has partnered with Huawei Cloud to leverage Huawei Cloud's AI technology, enabling the development of AI customer service and AI data insight applications. These solutions not only cut after-sales service costs but also meaningfully improve productivity across the enterprise.



Challenges

The company primarily faces six key challenges:



High requirements for fault localization

Company A's unmanned factories require special expertise in problem identification, making it difficult to upskill existing factory personnel. Problem localization is challenging and inefficient, resulting in extended machine downtime and slow fault recovery.



Data analysis bottlenecks

The digital twin platform has fixed data indicators and does not support custom queries. This results in inflexible business analytics and an inability to quickly query multidimensional statistical data. There is also a lack of data mining capability.



Need to reduce operational costs

Managing such a vast number of AGVs across hundreds of global projects presents numerous operational issues. The widespread distribution of industrial sites puts significant cost pressure on field support and employee travel.



Lack of support for analytical topics

The extensive system operations required for operational report analysis result in lengthy development cycles for reporting and large-screen monitoring functions.



Inconvenient humanmachine interaction

Current human-machine interfaces rely on small screens, making efficient operation difficult for personnel working in noisy transportation environments.



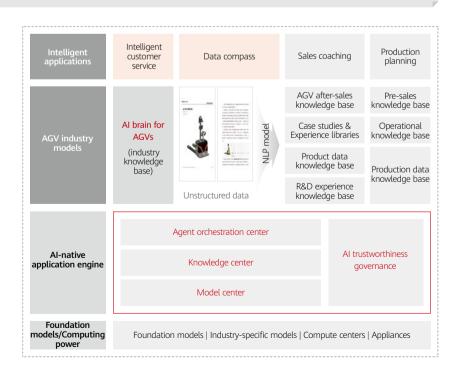
Need for faster efficiency improvement

A wealth of IoT sensor data from production and AGVs has been collected, but there is an urgent need to link this data with operational coordination data to optimize enterprise planning and scheduling (APS).

Solution

To address these six key business challenges, Company A adopted Huawei Cloud's AI-native application engine as the platform foundation, seamlessly integrating AI customer service and AI data insight applications into both AGVs and the digital twin platform. The architectural design is illustrated below.

The architecture is anchored in models and computing power, incorporating both large and small models within the Al-native application engine. The model center interconnects multiple mainstream large models. Model evaluation and routing ensure the selection of the most suitable models for diverse use cases to deliver optimal business outcomes. The knowledge center codifies industry-specific knowhow, enhancing the performance of large models through knowledge engineering to drive further innovation

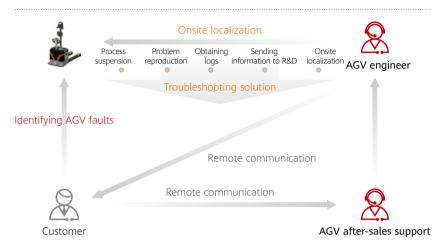


in industry applications. The agent orchestration center uses agents to enhance digital twins of people, events, and objects, enabling a "10-millionagent" architecture that fosters new production paradigms where agents and humans collaborate seamlessly. We combine multiple knowledge bases: AGV post-sales, pre-sales, case studies & experience, operational expertise, product data, R&D experience, and production data. These are combined with the AGV industry knowledge base to develop natural language processing (NLP) models.

1. Al customer service

When unmanned forklifts experience issues like stalling or incorrect pathing at customer production sites, the production process is directly impacted. Customers typically contact the AGV manufacturer's after-sales service promptly, but traditional customer service often cannot resolve these technical issues, which need intervention from R&D personnel. This not only strains R&D resources but also drives up operational costs. The company urgently needs to reduce after-sales costs and enhance efficiency through technology. This is because it manages over 600 projects and more than 1,200 AGVs globally in parallel. It also has cost pressures from geographically dispersed site support and personnel travel.

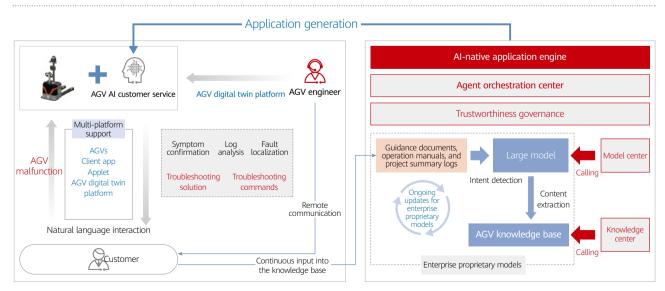
Traditional customer service



To address these challenges, Huawei Cloud collaborated with Company A to launch an AI customer service application. Built on an AI-native application engine, this application integrates AGV product manuals, maintenance guides, and system documentation, enabling large models to proactively understand customer needs. The AI customer service application can intelligently analyze logs, localize faults, and provide troubleshooting strategies to help technicians quickly pinpoint issues.

When unmanned forklifts malfunction at customer production sites, users can immediately interact with the AI customer service with their voice. The Al customer service can offer fault causes, solutions, product explanations, and answers to common questions. Furthermore, the large model can break down complex instructions to allow AGVs to automatically execute repair tasks. For complex issues, R&D engineers can use the digital twin platform for remote fault localization and realtime progress visualization. Thus, the Al customer service can relieve R&D engineers from tedious O&M tasks or reduce their workloads.

Al customer service system



2. Al data insight

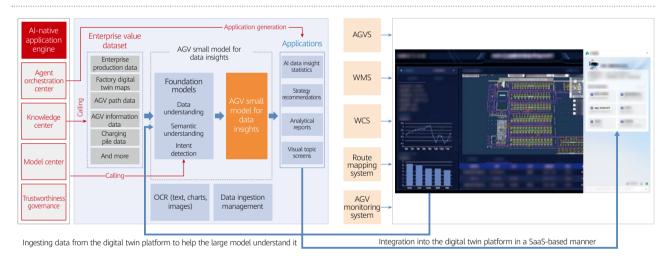
The AGV industrial digital twin platform fully integrates customers' production. management, and service operations. It enhances digital capabilities for deeper, autonomous insights into operational data, intelligent decision optimization, precise robot control, and more. Built on an Al-native application engine, the platform meticulously organizes customer business logic and systematically arranges AGV operation data indicators to define clear objectives and create standardized data tables. Large models gain a deep understanding of operational process data, allowing them to precisely track

various state data—from work volume statistics to AGV operation statuses. The large model offers open APIs that are closely linked to digital twin screens, making queries and calls convenient and efficient. An embedded data insight dialog box rapidly outputs detailed statistical and analytical results by understanding and organizing diverse data types.

Company A and Huawei Cloud have joined forces to co-develop an AGV industry model. This model consolidates scattered general knowledge, valuable engineer expertise, and complex unstructured data within the industry into a cohesive AGV industry Al brain.

By harnessing the powerful capabilities of large language models, a series of Al applications has emerged. The Al customer service uses NLP technology and intelligent reasoning algorithms for immediate, precise problem-solving, and for technical customer support. The data compass, based on big data analytics and visualization techniques, provides clear, data-driven insights for corporate decision-making. These initiatives not only deeply empower AGV operations with AI but also drive the entire AGV industry towards greater intelligence. efficiency, and innovation. They position the industry at the forefront of the digital transformation wave, reshaping both the industry ecosystem and the competitive landscape.

Al data insight system



Service Improvement

Company A leverages Huawei Cloud's powerful Al-native application engine as a solid platform foundation, successfully introducing Al customer service and Al data insight applications into AGVs and their digital twin platform. These achievements have yielded significant benefits, notably reducing AGV aftersales costs and substantially boosting enterprise productivity.

- Intelligent fault diagnosis: Moving beyond traditional high-barrier approaches to problem identification,
- semantic parsing technology allows for rapid and precise fault identification, even with ordinary semantics. With 24/7 Al experts on hand, after-sales service costs have been cut by 50%.
- 2. Intelligent log analysis: Advanced AI log parsing technology delivers swift fault resolution for efficient troubleshooting, acting like a team of expert AI R&D specialists. The efficiency of decision support has grown over five times, while fault
- localization efficiency has surged beyond tenfold.
- 3. Real-time human-machine interaction: Customers can perform troubleshooting operations directly through text selection or voice commands. Using voice interaction to quickly pinpoint issues optimizes user experience.
- **4. Visualization presentation:** Deep integration with the digital twin platform provides intuitive, clearly visualized topic screens. Specialized

operational reports can be generated with just a few clicks.

- 5. Constructing enterprise-level industry models: Large models have a deep understanding of key AGV operation data. Examples include work volume statistics, mileage, task execution counts, runtime, fault states, charging conditions, and idle time. This supports end users in
- making scientific production decisions which optimize their production processes and significantly boost their capacity.
- **6. Interactive data query and analysis:**The model performs in-depth data analysis in the digital twin process.
 Through natural language questionand-answer interactions, operational data statistics and analysis reports can

be quickly generated, delivering what users need.

7. Aggregation of enterprise value data: OCR technology is used to aggregate diverse data sources, including enterprise production data, AGV data, and twin map data. This helps uncover new value in multidimensional correlated data, efficiently transforming it into actionable insights.

Summary

Company A uses the Huawei Cloud Al-native application engine as the platform foundation. The Al customer service application uses product documentation to enable the large model to understand customer business, intelligently analyze logs, and accurately locate faults. Customers can ask questions via voice interaction, and the large model can decompose complex instructions to allow AGVs to automatically repair themselves. For more complex issues,

the digital twin platform can be used for remote fault localization and visualized troubleshooting. The Al data insight application uses Huawei Cloud's engine to organize business logic and compile data indicators. The large model understands operational data and provides open APIs to connect to digital twin screens, generating statistical and analytical results. Together, they form an industry Al brain with a wide range of Al applications.

Thanks to the empowerment mechanisms of the large AI model and the AI-native application engine, Company A has had a comprehensive technological innovation. Next-generation products will feature greater autonomy and adaptability, capable of precisely analyzing complex operational data. Additionally, the enhanced intelligence of these products reduces after-sales service costs from a technical standpoint and provides customers with more convenient and efficient service guarantees.



Pangu Models Drive Innovation in Medical Technology, Fostering a Smart Healthcare Ecosystem with

Fostering a Smart Healthcare Ecosystem with Enhanced Doctor-Patient Connection



Abstract

This case study describes how Company R used Huawei Cloud's Pangu models to power a range of intelligent healthcare applications, such as medical knowledge Q&A, test result interpretation, medical image structuring, and Al-supported diagnostics and treatment. While ensuring data security, compliance, and high system availability, the solution improves diagnostic accuracy, reduces labor costs, and creates a more efficient, intelligent, reliable healthcare platform. It helps Company R improve service quality and become more competitive, while also promoting the use of Al in healthcare services.

Background

Company R is a comprehensive medical laboratory service provider based in China, offering services such as In vitro diagnostics (IVDs), technical support, and lab operations management. As the world enters a new AI era, Company R has been actively leveraging AI technology, particularly large AI models,

to enhance the quality and efficiency of healthcare services. Working with Huawei Cloud, Company R has found promising ways to apply AI to diagnostics and treatment processes. It leverages the cloud's high performance and agility to power AI applications, enabling more efficient and accurate

support for both doctors and patients. To this end, it has developed an industry-leading medical AI model that powers a range of intelligent healthcare applications, including knowledge Q&A, test result interpretation, image structuring, and AI-supported diagnostics.



Challenges

Digital and AI transformation presents both opportunities and challenges for the healthcare industry. The opportunities primarily lie in leveraging data and AI to boost efficiency in

diagnosing and treating patients and improving service quality. Company R has been an early mover in digital transformation, but it faces many challenges in terms of data processing and storage performance, data security and privacy, regulatory compliance, and system availability and reliability.



Data processing and storage performance

Massive amounts of medical data requires structured storage and efficient computation on the cloud to guarantee real-time performance and accuracy under high concurrency.



Data security and privacy protection

Medical data contains highly sensitive information and is therefore subject to the most stringent security regulations. To prevent breaches and unauthorized access, robust measures, such as local encrypted storage and strict isolation protocols, must be implemented.



Regulatory compliance

Relevant systems must comply with applicable laws and regulations for the healthcare industry, ensuring compliance during data storage, transmission, and use.



High availability and reliability

Healthcare services have exceptionally high standards for system continuity and stability. To guarantee service continuity during emergencies, robust disaster recovery (DR) and fault recovery mechanisms are essential.

Solution

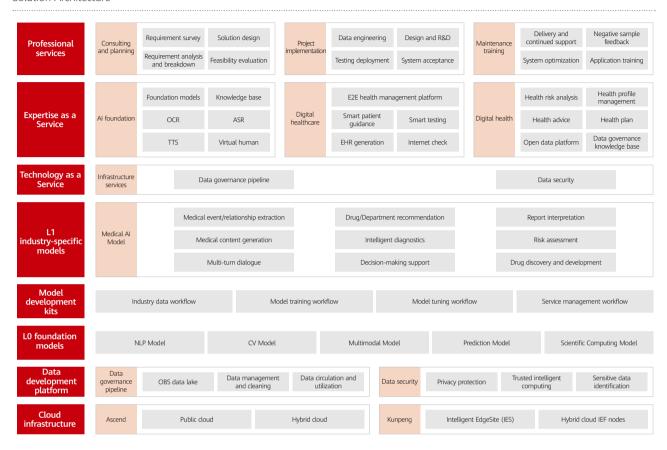
Through cooperation with Huawei Cloud, Company R not only made significant breakthroughs in technology, but also improved services and market competitiveness in many aspects.

Working with Huawei Cloud, Company R built a multi-layer healthcare solution framework with AI at it core. This framework consists of professional

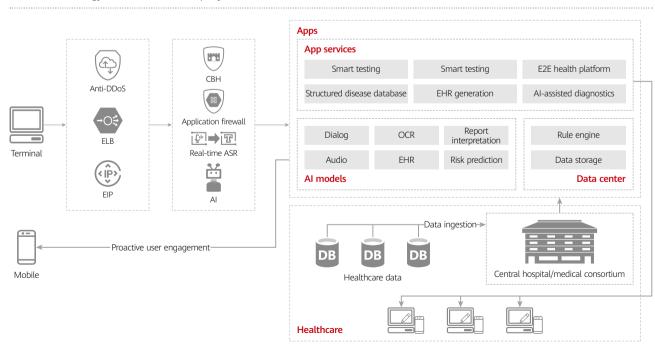
services, expertise as a service (EaaS), domain-specific models, and a unified data platform. It covers the entire lifecycle of healthcare services from requirement analysis and model development, to service deployment and maintenance. Huawei Cloud provides comprehensive support in terms of AI for medical technology, data management, and model

application. They help Company R develop intelligent healthcare services such as intelligent diagnostics and AIassisted electronic health record (EHR) generation. While attaching top priority to data privacy and system stability, the overall architecture focuses on medical data management and the structured processing of medical records, supported by AI models like OCR and speech recognition.

Solution Architecture



Medical Technology Architecture of Company R



Application 1: R Medical AI (Knowledge Q&A + Test Result Interpretation)

R Medical AI provides intelligent Q&A on medical lab testing as well as test result interpretation, while the Pangu model powers its intent recognition and Q&A capabilities. This application significantly improves efficiency and cuts the labor costs in interpreting lab test results for patients.

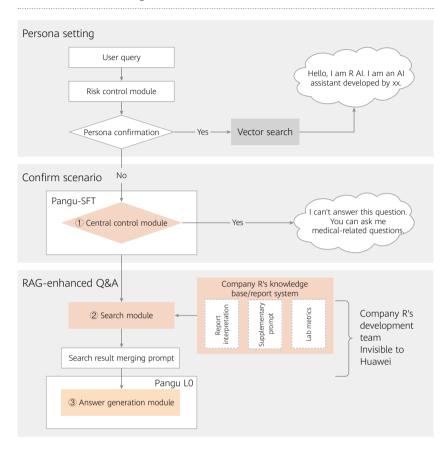
1. Solution design

- » Intelligent central control + intent recognition: Huawei Cloud supported the solution design by providing an Al consulting service. The Pangu N1 model supports accurate user intent recognition, ensuring that the Al assistant can accurately understand patient questions while limiting the scope of these questions, and generate high-quality answers.
- » Risk control and content security: A risk control module automatically blocks non-medical related questions. Persona setting, an efficient search module, and AI enablement services from Huawei Cloud helped ensure that the medical AI assistant generates reliable, professional answers to medical questions at low latency while delivering a good user experience.
- » OCR + medical knowledge base: A cloud-based OCR service converts patient-uploaded test reports to structured data. Then, the knowledge Q&A module interprets these reports with high accuracy, enabling quick response to patient questions.

2. Service improvement

» Reduced costs: R Medical AI allows doctors to spend less time reading patients' MRI reports and answering their questions, so they can concentrate on more demanding tasks, like thinking and diagnosing.

R Medical AI solution design



- Faster test result interpretation: R
 Medical AI can accurately interpret
 test results uploaded by patients,
 including locations, features, values,
 and units. Patients don't have to
 wait in long queues for someone to
 explain their test results to them.
- » Reliable healthcare services: The Pangu N2-supported intelligent Q&A service, coupled with an efficient search module, ensures that AI generates reliable, trustworthy replies.

3. Key success factors

» Accurate intent recognition: According to evaluation results for the Pangu N1 model, the intent recognition module allows the

- model to accurately understand and respond to user questions related to medical lab testing.
- » Multi-level risk control and content security assurance: The joint use of risk control, search, and persona setting ensures that Al generates reliable, professional answers in response to patient questions.
- » Integration of OCR and knowledge Q&A: OCR converts patients' test results into structured data, and then the knowledge Q&A module immediately interprets them. No human intervention is required. The whole process is seamless, ensuring a good user experience.

Application 2: Imaging Report Structuring

1. Solution design

- » Standard labeling process: Huawei Cloud helped Company R developed a standardized data labeling process through a consulting service. This helped ensure training data quality for its medical AI model.
- » Data-driven model training: Specialized doctors were recruited to perform the data labeling. This, coupled with the powerful performance of the Pangu N2 model, reduced the need for subsequent engineering and processing, and enabled efficient model training and deployment.
- » Hierarchical evaluation criteria: A 5-point scale was used to accurately evaluate the model's performance in imaging report structuring based on carefully defined criteria.

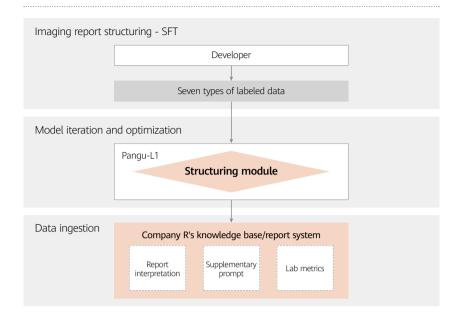
2. Service improvement

- » Reduced labor costs and accelerated project cycle: Manual labeling is reduced. High-quality datasets can be developed within a short period.
- » Wider adoption and more use cases: With the combination of efficient, standardized data labeling and a powerful NLP model, the service is expected to be adopted by more hospitals. With an overall accuracy and recall rate above 90%, it will come to support a wider range of medical use cases.

3. Key success factors

- » Professional team collaboration: Guided by a set of cross-function collaboration rules, the expertise of medical specialists, coupled with a standardized labeling process, ensured training data quality for the medical AI model.
- » Efficient data labeling: Unified labeling rules improve data labeling efficiency, thus accelerating model development.
- » Continuously enhanced model evaluation mechanisms: Models

Medical data structuring



are evaluated based on carefully defined criteria and continuously optimized based on feedback data. This helps to enhance model performance, making them more reliable and stable for cloud-based deployment.

Application 3: Al-Supported Diagnostics

1. Solution design

- » Hierarchical data management: Huawei Cloud offered a hierarchical data management consulting service. Patient information is divided into fixed information and timeline-based information. This facilitates data management and ensures that the model's input includes crucial clinical course information that describes patients' medical conditions over time.
- » Optimized processing of overlong tokens: In the case of overlong tokens, the input length is reduced by keeping only the latest clinical course information plus an abstract of the remaining information. This enables the model to handle long input text.

Efficient data aggregation and diagnostic support: Based on comprehensive clinical course information in the input, the N2 model supports decision-making during various types of diagnostic and treatment processes, while also raising the level of automation during such processes.

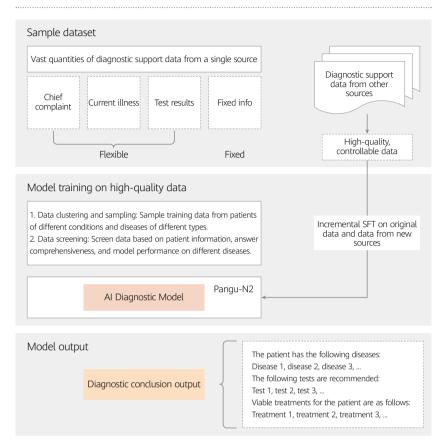
2. Service improvement

- » Improved diagnostic accuracy: Based on comprehensive clinical course information in the input, the N2 model helps doctors diagnose patients more efficiently and with higher accuracy.
- » More efficient model computation: With efficient processing of overlong tokens, the computational load is reduced, enabling faster responses from the model.
- » Effective management of complex information: Hierarchical information processing and efficient input construction ensure that complex clinical/disease course information is effectively parsed and utilized, improving the adaptability of the diagnostic AI model.

3. Key success factors

- » Data structuring and standardization: Effective data structuring and classification help the AI model extract useful information more efficiently, so it can make more accurate diagnostic decisions.
- » Optimized processing of overlong tokens: In the case of overlong tokens, the input length is reduced by keeping only the latest clinical course information plus an abstract of the remaining information. This keeps the input length under limit while preserving comprehensive information.
- » Model selection and finetuning: The N2 model has proven performance in the medical domain, and fine-tuning can further improve its specialization. The model's capabilities can be further validated in a wider range of highly specialized tasks in the realm of healthcare.

Model training on vast quantities of data



Service Improvement

Based on extensive requirement surveys and research, Company R chose intelligent diagnostics as the primary AI use case for healthcare that it would focus on. By providing first-class intent recognition and OCR capabilities, Pangu

models help Company R boost efficiency in interpreting medical test results by around 50%. Intelligent medical Q&A also delivers a high accuracy and good patient satisfaction. In the case of imaging report structuring,

the re-trained N2 model achieves an accuracy of over 95% in parsing cardiac ultrasound and thyroid images. In Alassisted diagnostics, the medical AI model helps doctors diagnose their patients more accurately and efficiently.

Summary

Working with Huawei Cloud, Company R developed a range of intelligent healthcare applications, such as medical knowledge Q&A, medical test result interpretation, medical image structuring, and Alsupported diagnostics. The joint solutions meet high standards for data security and privacy protection, availability, and reliability involved in processing and storing medical data.

In medical knowledge Q&A, Pangu's intent recognition, OCR, coupled with curated medical knowledge bases, help deliver high-quality answers to patient questions and ensure high accuracy and efficiency in interpreting patients' lab test results. In imaging report structuring, a standardized labeling process and a high-performance model ensure high accuracy in image

recognition. In Al-assisted diagnostics, hierarchical data management and optimized processing of overlong tokens ensure efficient computation of complex clinical course information. These solutions improve medical services by helping doctors make more accurate diagnosis, cutting labor costs, and improving structured medical data management.

RPA for Al-powered Government Services



Abstract

This case describes how a development area uses Huawei Cloud robotic process automation (RPA) to offer Al-powered government services. With RPA, government services become smarter and more user-centric.

Background

Industries are harnessing the power of AI to go digital in real-world scenarios and stay relevant in the fastgrowing digital economy. One of the Al technologies used for this is robotic process automation (RPA), which simulates manual labor to efficiently complete a large number of repetitive tasks with fixed rules. Easy to operate and widely adopted, RPA has become an indispensable tool for the digital

transformation of government services.

The administrative service center of a development area leverages the digital robot integrated with Huawei RPA to handle requests in just a few clicks. Intelligent information filling and review streamlines the government service process, and

digital service representatives facilitate civil services, allowing the city to better serve both people and businesses. This project was rated by the province National Development and Reform Commission as one of the best practices of reforming the administrative approval system. It has been widely adopted nationwide and won high recognition.



Challenges

One of the key measures for modernizing government services is to provide a unified portal for all services. In particular, the one-stop enterprise registration service is crucial for creating a business-friendly environment. However, the initiative faces many challenges:



More user-friendly request handling



Understaffing



Ever-evolving government services

A prospective entrepreneur had to go back and forth between five service windows, with piles of forms spilling out of their hands. After completing all these, it still took five days to obtain approval.

There were so many manual operations, such as data filling that took 1 hour. The understaffed service center was unable to complete service handling in time.

Subjective personal experience, without the backing of data, simply is not enough to consistently deliver innovative solutions.

Solution

To address these challenges, the development area pioneers intelligent administrative approval with digital robots. These Al service representatives in the administrative service center

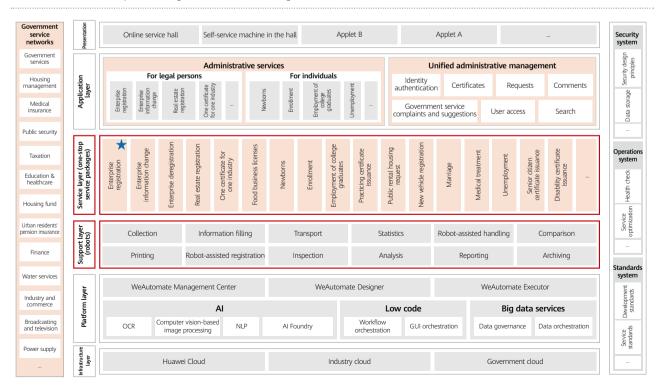
reshape the digital service model by completing each request in just a few clicks.

Here is the scenario-specific solution:

Digital robot solution for government services

The architecture of the RPA-powered digital robot solution for government

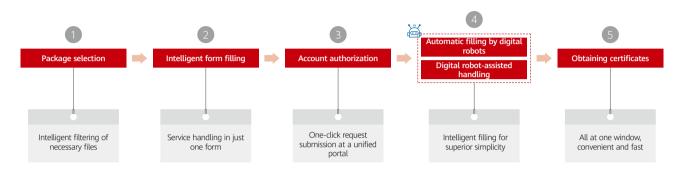
Architecture of the RPA-powered digital robot solution for government services



services consists of the application layer (automated execution of government applications), service layer (government service windows), support layer (assurance of robot development, deployment, and operating management), platform layer (integration and collaboration of data and resources), and infrastructure layer (operating environment on hardware).

The following describes the intelligent process of RPA-powered digital robots for the one-stop enterprise registration service at the service layer:

Intelligent process of the one-stop enterprise registration service



The user-to-user service process has been reconstructed to tackle the most prominent challenges, with digital robots automating service handling from end to end:

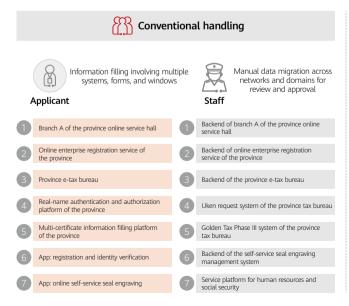
- ① Intelligent identification: Simplification and redundant information removal mean that just one form needs completing.
- ② Intelligent collection: Information collection from various devices, such as card recognizers and high-speed document scanners,

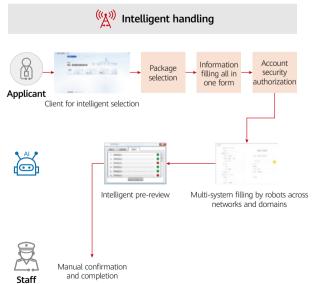
- significantly boosts efficiency.
- ③ Intelligent filling: Applicants can scan the QR code to authorize digital robots to use their accounts securely to complete requests on each private network.
- ① Intelligent assistance: Digital robots acquire key pre-review information to help the service personnel quickly extract information, taking the efficiency of pre-review to new levels.
- ⑤ Intelligent upload: After completing package services,

digital robots accurately migrate service information from the private network to the online customer service system in real time

RPA-powered robots outperform the conventional enterprise registration service in the following aspects:

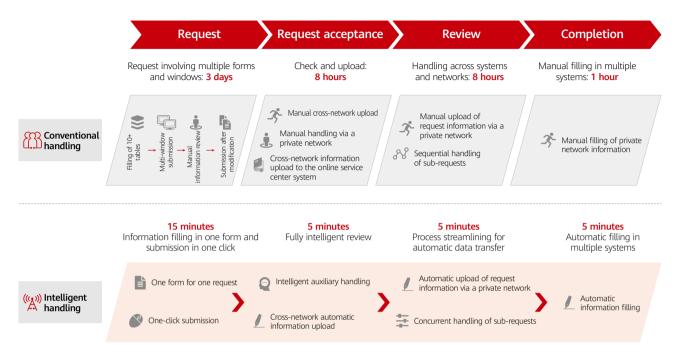
 Experience: RPA is more usercentric, improving the enterprise registration experience for startups.



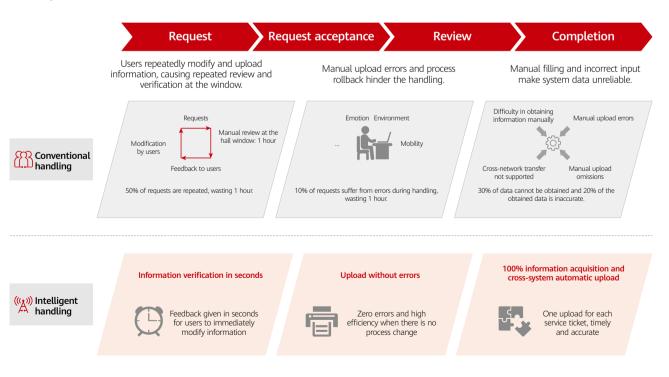




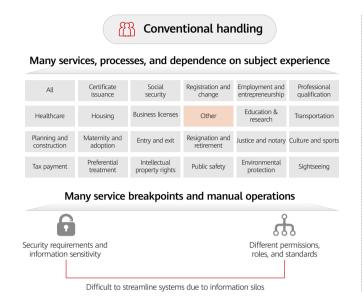
2) Efficiency: The request handling cycle is shortened from five days to 30 minutes, significantly improving experience of applicants.



3) Accuracy: RPA improves the accuracy of process operations and reduces wasted time due to repeated manual misoperations.



4) Replication: This innovative attempt in the unified portal for all services can be replicated to other areas.





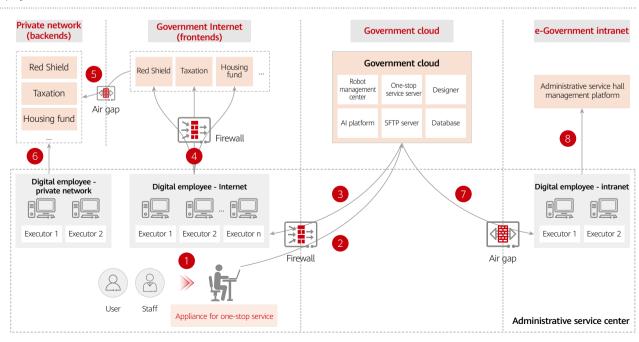
2. RPA-powered digital robot deployment

The resources of RPA-powered digital robots are deployed on

the e-Government Internet, e-Government cloud, e-Government intranet, and private network. The robot management center, onestop service server, AI platform, SFTP server, and database are deployed on a municipal e-Government cloud. See the deployment scheme:

service handling.

Deployment scheme



- ① In the Internet area, components for filling in information on the Internet are deployed, including digital employees who fill in information. After a user arrives at the service hall, they can fill in the service request information through the one-stop service.
- ② The data filled in is uploaded to the database of the e-Government cloud through the firewall.
- ③ Digital employees in the Internet area obtain data from the e-Government cloud through the firewall.
- ④ Digital employees automatically fill in information at each service

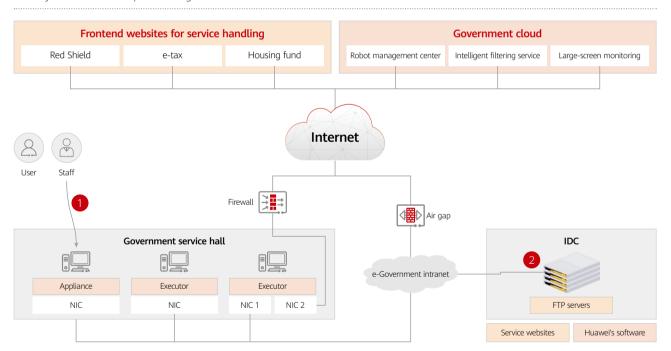
- frontend through the firewall, such as the Red Shield (an industrial and commercial information portal), taxation, and housing fund.
- ⑤ The service frontend and backend automatically synchronize data through the air gap.
- ⑥ Digital employees on the private network assist in pre-handling and pre-review.
- Digital employees on the e-Government intranet obtain data from the e-Government cloud through the air gap.
- ® Digital employees on the e-Government intranet automatically upload data to the

backend of the administrative service hall.

Security solution for RPApowered digital robots

The RPA-powered digital robot supports localized full-stack technologies and products, and is equipped with a mature trustworthiness and security management system as well as intellectual property rights. It can be installed and run on localized systems to ensure the secure operations of the RPA system.

Security solution for RPA-powered digital robots



1) Security management solution for user accounts: The system does not obtain end users' account information. Instead, it pushes QR codes to users. Users can scan the QR codes to log in. The account information of service personnel is encrypted using AES256 to prevent

account information leakage through centralized management. In addition, the robot logs all operations to make the entire process transparent and traceable.

2) Secure data storage solution:Users' service data and personal

data is encrypted and stored on SFTP file servers. For each request, the user applies for a key from the e-Government cloud server to ensure that the user's data in each request is encrypted using a unique key.

Service Improvement



10+ forms → 1 form



5+ systems → In a few clicks



7 employees → 1 robot

- 1. Faster service handling: Each request can be handled in just a few clicks, using just one table. Intelligent filling by digital robots streamlines the service process. Each request is completed in 30 minutes on average, saving both time and money for citizens.
- 2. Improved cross-network and cross-field efficiency: Intelligent handling of major services and concurrent handling of auxiliary services enable automatic filling.
- **3. Accumulated digital assets:** Knowledge acquired from services is accumulated to continuously

improve government service capabilities.

4. Smarter government services: Al equips RPA with better learning and informed decision-making to easily handle more complex and changeable government services.

Key Success Factors

The trustworthy technology stack can be deployed for tools, enablement, and support systems. Compatible with multiple trustworthy application products, RPA safeguards the long-term operating of services.

- 1. Drag-and-drop tools for easier development: Low-code/No-code development makes life easier for developers. It takes only five days to adapt to a single process.
- 2. Quick start with a deployment quide: With no need for repeated
- surveys, staff can familiarize themselves with the entire process within just half a day.
- **3. Full-stack trustworthiness:** High compatibility avoids the need for on-site adaptation.

Case Summary

The RPA-powered digital robot solution perfectly fits into an increasingly digital world and echoes the Al-favored national policies that aim to make government services smarter and more efficient. The administrative service center of a development area pioneers this

initiative with RPA-powered digital robots. The one-time handling of each request greatly improves the government approval process.

Fueled by AI innovation, the RPApowered digital robot solution can be adopted nationwide. With process streamlining and even reconstruction, the solution can be maturely applied to cross-field and general-purpose work. Less repetition means higher efficiency in the government approval service. Smarter government services will make service handling easier and people happier.

