

华为云 确定性运维

第七期

内部发行
免费赠阅

前沿观点

大咖访谈

技术创新

业务提升

运维动态

运维效能再进阶，以AI驱动质效双跃升

HyperRouter: 从四层负载均衡设计与开发，看确定性运维实践

— 华为云爱尔兰可靠性实验室首席软件工程师 Linhua Tang

SRE深度思考：云演进新阶段，高阶服务与客户成功定义核心竞争力

— 华为云SRE部 沈幸峰 李昂 李川

以智能可观测技术，构建数据驱动的主动运维治理体系

— 信通院云大所副主任 郑立

美宜佳AI应用实践，赋能万店智慧零售新业态

— 美宜佳控股有限公司CIO 代迪





主 办	华为云计算技术有限公司SRE部
顾 问	高江海 安 宇 林华鼎
主 编	梅守祯
副主编	郑梦杰 冉冉
撰 稿	程凌 湛宁坤 代迪 窦顺阳 李昂 李川 Linhua Tang 刘军 沈幸峰 王福强 吴祥 薛黎 郑立 (按姓氏首字母排序)
编 辑	贾一夫 卢素娟 孟淼 吴顺群 吴文辉 张飞
联 系	电子专刊、索阅、投稿、建议和意见反馈请联系SRE部 邮箱: snzx1@huawei.com
地 址	东莞市松山湖高新技术产业开发区环湖路9号华为溪流背坡村
邮 编	523830
无担保声明	本资料内容仅供参考, 华为云计算技术有限公司不对本资料所有内容提供任何明示或暗示的保证, 包括但不限于适销性或者使用与某一特定目的的保证, 在法律允许的范围内, 华为云计算技术有限公司在任何情况下都不对使用本资料任何内容而产生的任何特殊的、间接的、继发性的损害进行赔偿, 也不对任何利润、数据、商誉或预期节约的损失进行赔偿。(内部发行, 免费赠阅)
版权归属	版权归华为云计算技术有限公司所有, 保留一切权利。 非经本公司书面许可, 任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部, 并不得以任何形式传播。

目录



前沿观点

- HyperRouter: 从四层负载均衡设计与开发, 看确定性运维实践 02
- SRE深度思考: 云演进新阶段, 高阶服务与客户成功定义核心竞争力 08



大咖访谈

- 以智能可观测技术, 构建数据驱动的主动运维治理体系 13
- 美宜佳AI应用实践, 赋能万店智慧零售新业态 17



技术创新

- 移动终端+ChatOps, 破局传统运维局限 22
- 构建数智运维保障体系, 解锁业务增长新动力 26



业务提升

- ECDN助企业内网降本增效、数据“秒速”畅通 31
- 识变求稳, 用定制化检查夯实系统运维根基 35

usenix SRE CON[®]

SREcon是全球顶级SRE会议，自2014年首届举办以来，已成为站点可靠性工程领域最具影响力的盛会。它汇聚全球系统运维与大规模分布式技术领域的工程师及专家，推动深度技术交流与行业创新文化的落地。

2025年10月7日，SREcon25在爱尔兰都柏林举行。华为云爱尔兰可靠性实验室首席专家Linhua Tang首次受邀在主会场发表Plenary（Keynote）演讲。

华为云亮相SREcon25主会 揭秘大规模负载均衡架构实践



Linhua Tang

华为云爱尔兰可靠性
实验室首席专家

Linhua Tang分享了华为云SRE团队构建四层负载均衡服务Hyper-Router的确定性运维实践，从中提炼出六大lessons，并深入阐释了构建高可靠、高性能、大规模服务的关键决策与逻辑。

Lesson 1

通过跨服务的协同设计可以简化系统复杂性，提升整体系统可靠性。

Lesson 3

自愈及隔离能力应该是系统设计的基本原则之一而不是作为次要特性，以充分利用系统特征实现高效、可靠的自愈及隔离机制。

Lesson 5

Cell架构和Shuffle sharding可以有效降低大规模多租系统的爆炸半径。

Lesson 2

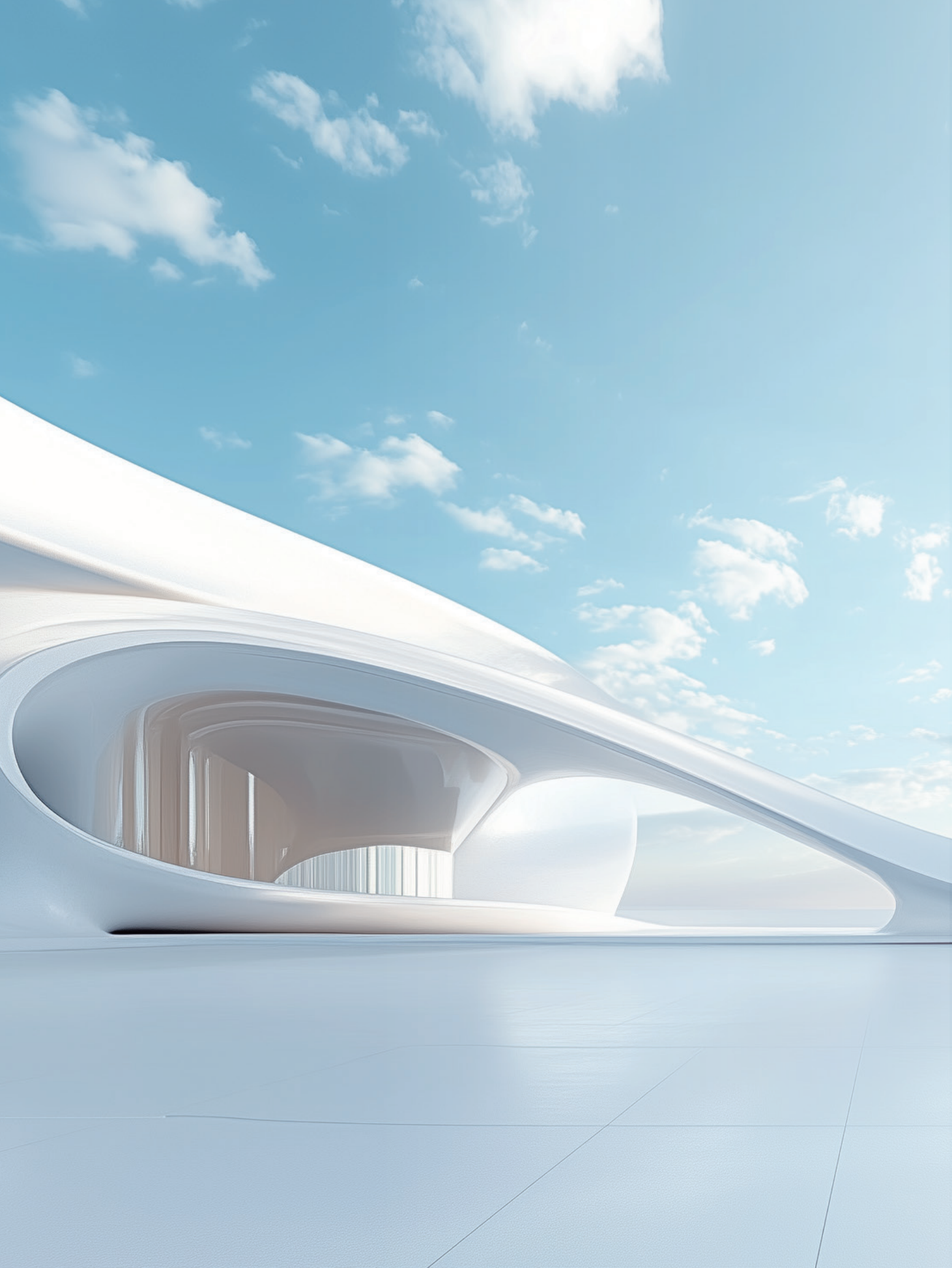
用户态内核旁路提供了最大灵活性和性能，但会牺牲硬件复用与能效；工程实践中的选择取决于性能与资源利用之间的权衡。

Lesson 4

点对点存储通信架构非常适于需要极高A、P（CAP理论中的Availability及Partition Tolerance）及最小化外部依赖的系统。

Lesson 6

CUJ是一个行之有效的以用户为中心的可观测框架和实践。



“ 序言



高江海
华为公有云业务部总裁

让运维成为智能世界变革的加速器

云计算带来的普惠技术，成为推动数字化转型的重要引擎。越来越多企业将核心业务部署到云端，华为公有云业务保持了高速增长，客户来自千行万业，使用云的能力参差不齐。如何提供安全、稳定、高质量、易用的服务，是我们面对的一个挑战，这对云底座的“质量”提出极为严苛的要求，需要采取一系列措施来预防和解决。如何有效管好云，持续为业务赋能，是企业关注的重要命题。

安全可信是前提，稳定可靠是基础，资源高效是关键，业务敏捷是核心。

“确定性运维能力”是我们运维最主要的能力，我们要从设计、开发、部署、监控、运维等环节发力进行全面质量管理发力，对故障发生、故障范围、故障恢复、故障影响做到确定性控制，实现最小化、网格化的管理，从流程、意识、质量文化、考核、工具等，帮助企业完成从传统维护团队向IT平台服务的转型，将业务高速发展带来的“不确定性”变成SLO（服务质量目标）的“确定性”，是我们不断夯实的基础能力。

2023年9月华为云正式发布维享会（确定性运维经验交流分享会），华为云希望建立这样一个平台链接全球客户，与众多行业大咖一起探讨新技术、碰撞新思想，通过专刊、白皮书和案例集等载体，进行观点和实践经验分享，探讨业务上云后的最佳实践和创新融合，相互碰撞和启发，共同构建安全可靠的“确定性”运维世界。



HyperRouter: 从四层负载均衡设计与开发，看确定性运维实践



Linhua Tang

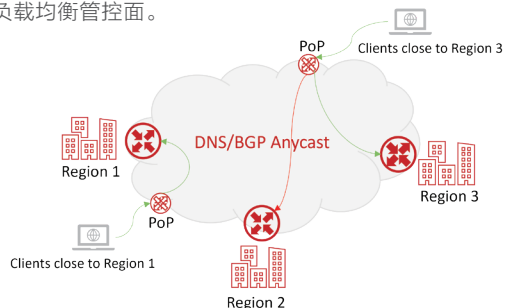
华为云爱尔兰可靠性
实验室首席专家

摘要：本文基于SREcon25 EMEA主题演讲HyperRouter: Lessons Learnt from Building an L4 Load Balancing Service整理成，在SREcon官网上可以找到演讲视频和材料。

一、确定性运维与HyperRouter的定位

在华为云的运维实践中，SRE（站点可靠性工程）团队以云服务全生命周期为核心，构建了贯穿始终的质量看护机制，最终实现现网服务高可用的可预期性。这种经过实践验证的运维理念与系统化的质量管理体系，被定义为“确定性运维”。区别于传统运维“依赖经验、事后补救”的模式，确定性运维通过覆盖全流程的质量管理体系，将抽象的可靠性目标转化为可量化、可保障的具体指标——在现网环境中，确定性运维精准实现确定性故障率、确定性恢复时长与确定性影响范围，为云服务的稳定运行提供坚实且明确的保障。

华为云全局负载均衡服务是确定性运维的重要组成部分，华为云通过全局负载均衡实现高效跨局点容灾、服务无损升级、爆炸半径可控等关键确定性运维能力。华为云全局负载均衡由多个服务组成，包括基于DNS的全局负载均衡、四层负载均衡、七层负载均衡、Service mesh以及全局负载均衡管控面。



本文旨在分享华为云全局负载均衡的重要组成部分HyperRouter在设计、开发、运维过程中的确定性运维实践。

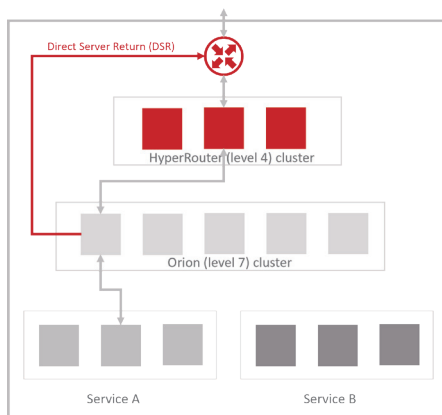
二、HyperRouter 设计与开发的六大核心实践

（一）Lesson 1: 通过跨服务的协同设计可以简化系统复杂性，提升整体系统可靠性

1. 四层与七层负载均衡协同设计

四层与七层负载均衡常以组合形式部署，七层负载均衡单独工作的场景较少。核心原因在于：应用层（七层）的有状态特性使其对单节点故障容错能力弱，如ECMP的简单哈希机制易导致大量TCP连接中断。因此，四层负载均衡的关键作用之一，便是在节点故障、扩缩容等场景下，提升七层负载均衡的可靠性。

设计四层负载均衡前，需先明确其在全局负载均衡平台中的定位，以及如何与上下游高效协同，从而简化四层架构；同时，需推动四层负载均衡尽可能实现无状态化，以获得最优横向扩展能力。基于此，HyperRouter在设计上主要实现了Google Maglev一致性负载均衡算法，并仅支持Direct Server Return（DSR）模式。



Maglev一致性负载均衡算法相比其他负载均衡算法的最大优势是最小化后端节点故障导致的TCP流中断；其次是让HyperRouter的每个数据面节点在无需任何状态同步的前提下，均能将同一TCP流的报文转发至同一个后端七层节点。为了进一步提升TCP流可靠性，HyperRouter通过Connection Tracking Table (CTT)机制实现同一个TCP流的报文总是转发到同一个后端健康节点，但CTT并非强依赖的有状态模块。

Direct Server Return (DSR) 模式的价值，在于让四层负载均衡专注于入向 (ingress) 流量转发，无需处理出向 (egress) 流量——这不仅能进一步提升四层负载均衡的横向扩展能力，还能显著降低资源开销。以简单的HTTP GET请求为例：一个`GET / HTTP/1.1`请求通常仅几十字节，单个TCP报文即可承载；但其响应数据量差异极大，小到几十字节，大到KB、MB甚至GB级。若四层负载均衡处理这类响应，会无谓消耗CPU与内存资源。通过DSR模式，七层负载均衡可直接将响应返回客户端，从而最大程度减少四层负载均衡的性能损耗 (Overhead)。

四层和七层负载均衡协同设计的实践可以应用于其他相关的服务，即通过跨服务的协同设计可以简化系统复杂性，提升整体系统可靠性。

2.HyperRouter设计目标

明确HyperRouter在全局负载均衡中的定位后，我们来聚焦其设计目标。和多数分布式系统一致，四层负载均衡服务的设计需围绕四个“简约却不简单”的核心目标展开：高性能、高可扩展、高可靠、可演进。

1)高性能：对四层负载均衡而言，高性能不仅要求以最少资源实现接近线路速率 (line-rate) 的处理能力，还需极致发挥CPU、网卡等硬件资源潜力——例如单CPU核需支撑千万级每秒数据包 (PPS) 处理量；同时，为避免对上层业务时延产生影响，每个报文的处理速度必须达到纳秒级。

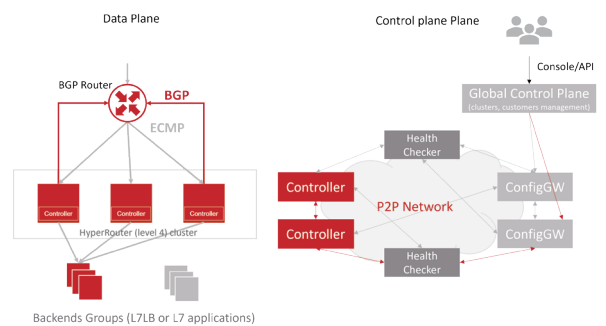
2)高可扩展：随着华为云服务数量与业务规模高速增长（如并发 TCP 连接从千万级攀升至十亿级），四层负载均衡服务必须具备极强的横向扩展能力，以适配业务体量的快速变化。

3)高可靠：无论是内部还是外部客户，均期望负载均衡提供近乎零中断的服务。因此，四层负载均衡需在各类场景下保障高可用与高可靠，包括日常软件升级、配置变更，以及硬件故障、网络故障等异常情况。

4)可演进：新场景、新需求、新硬件、新协议的出现，要求负载均衡系统能以最低成本、最高效率实现无断代的持续迭代，确保长期适配业务发展。

3.四层负载均衡架构设计

HyperRouter主要由两部分组成：数据面和管理面。



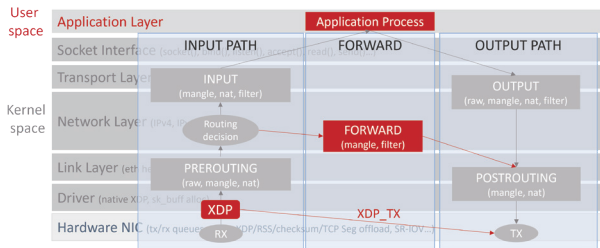
HyperRouter的架构可明确划分为数据面与管控面：数据面功能相对简洁，核心聚焦网络报文的接收、处理与转发，但对性能（如转发速率）和可靠性的要求极高；管控面则承担更复杂的管理职能，涵盖数据面集群调度、配置下发、BGP路由发布、全链路可观测、健康探测及自愈能力建设，对性能要求相对较低。

（本次分享重点并非HyperRouter的架构细节，因此不对各模块展开深入探讨）

(二) Lesson 2: 用户态内核旁路提供了最大灵活性和性能，但会牺牲硬件复用与能效；工程实践中的选择取决于性能与资源利用之间的权衡。

用户态高性能数据面

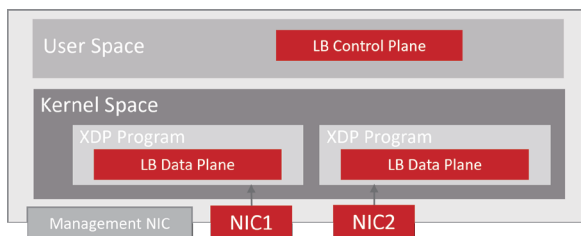
Linux内核网络栈功能完备，但在高性能场景下表现出局限性。其上下文切换、内存拷贝等通用机制导致延迟增加，在千万级PPS规模下成为性能瓶颈。



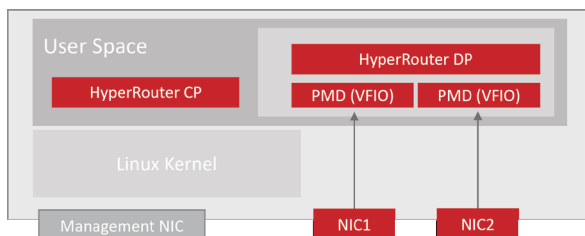
为突破性能瓶颈，需尽可能绕开Linux协议栈，Linux原生的“kernel-bypass”技术XDP（eXpress Data Path）是可行方案之一。它通过在Linux网络入口处运行eBPF程序处理原始报文，能实现高性能、低CPU开销的报文处理与转发。

但XDP存在明显局限：

- 1) eBPF程序的编程模型有约束，无法像通用程序那样灵活实现各类逻辑；
- 2) XDP要求每个网卡运行一个实例，而华为云数据面需至少支持双网卡，以适配双ToR交换机的流量处理需求，这与XDP的特性存在适配矛盾。



另一个方案就是DPDK（Data Plane Development Kit）。DPDK提供了在用户态直接接管网卡的能力，实现接近线速的转发性能，但代价是需要专用CPU与NIC，并且其关键的Poll model让CPU利用率始终保持100%。



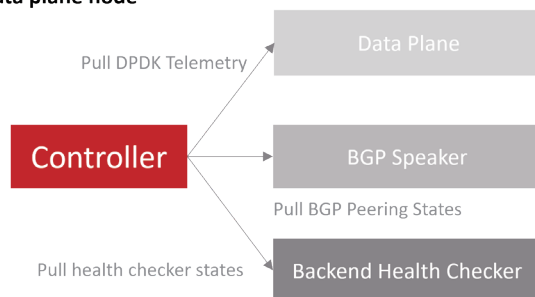
通常数据面的技术选型是一个“单向门决策”（One-way door decision）——一旦确定，切换方案的成本极高，因此需要尤为慎重。对我们而言，灵活性、可演进性与高性能等高优先级目标是我们最终选择DPDK的主要原因。

(三) Lesson 3: 自愈及隔离能力应该是系统设计的基本设计原则之一而不是作为次要特性，以充分利用系统特征实现高效、可靠的自愈及隔离机制。

自愈及隔离作为基本设计原则

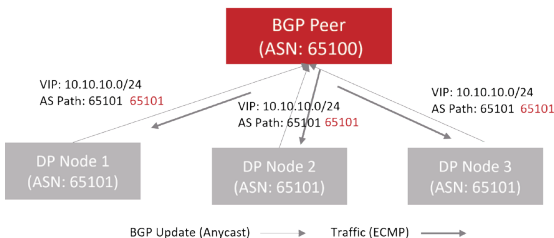
在分布式系统中，节点失效、硬件故障和网络分区等是必然事件。为此，HyperRouter从设计阶段就将自愈与隔离作为“一等公民”（First-class principle）。

Data plane node



首先是数据面节点的自我健康检查机制：Controller会持续拉取节点的核心数据（包括数据面状态与指标、BGP speaker状态、后端健康检查结果等），综合判断节点健康度。一旦检测到异常，会立即触发恢复或隔离动作，例如进程重启、配置动态调整，或直接将节点隔离。

以“转发错误（forwarding error）过高”场景为例：HyperRouter不会采用撤销BGP路由的方式转移流量，而是通过调整BGP AS Path（如追加或减少同一ASN）来改变路由权重，快速将流量导向健康节点，避免故障影响扩散。选择这一方案的核心考量是：自我探测可能因各类干扰出现误判（false positive），若直接撤销路由，可能导致报文无法正常转发，反而引发新问题。



尽管这套自愈与隔离机制并非适用于所有类型的应用，但“将自愈及隔离能力纳入设计”应作为系统设计的核心原则之一，而非事后补充的补丁方案。唯有在设计阶段就充分结合系统自身特征进行规划，才能构建出更高效、更可靠的自愈与隔离机制。

1.基于点对点架构实现系统依赖最小化

The diagram illustrates a Peer-to-Peer network topology with six peers, each represented by a red box. The peers are labeled as follows:

- Peer (Controller) - Top center
- Peer (Controller) - Middle left
- Peer (Config Gateway) - Bottom left
- Peer (Config Gateway) - Bottom center
- Peer (Health Check) - Middle right
- Peer (Health Check) - Bottom right

The connections between the peers are as follows:

- Peer (Controller) (Top center) is connected to Peer (Controller) (Middle left), Peer (Config Gateway) (Bottom left), Peer (Config Gateway) (Bottom center), Peer (Health Check) (Middle right), and Peer (Health Check) (Bottom right).
- Peer (Controller) (Middle left) is connected to Peer (Controller) (Top center) and Peer (Config Gateway) (Bottom left).
- Peer (Config Gateway) (Bottom left) is connected to Peer (Controller) (Middle left) and Peer (Config Gateway) (Bottom center).
- Peer (Config Gateway) (Bottom center) is connected to Peer (Controller) (Top center), Peer (Controller) (Middle left), Peer (Config Gateway) (Bottom left), Peer (Health Check) (Middle right), and Peer (Health Check) (Bottom right).
- Peer (Health Check) (Middle right) is connected to Peer (Controller) (Top center) and Peer (Health Check) (Bottom right).
- Peer (Health Check) (Bottom right) is connected to Peer (Config Gateway) (Bottom center) and Peer (Health Check) (Middle right).

因此，HyperRouter采用了点对点去中心化架构：各节点可独立完成状态发现与同步，即便部分区域发生网络分区，整个系统仍能维持核心功能正常运行。为此，我们从零构建了一套专为全局负载均衡设计的点对点框架，该框架由存储层、KV语义层、通信层及应用层组成。这种架构的核心价值在于，即便出现网络分区，各“孤岛”节点仍能最大限度自主做出正确决策，从根本上避免了因依赖中心化组件而可能引发的级联故障。

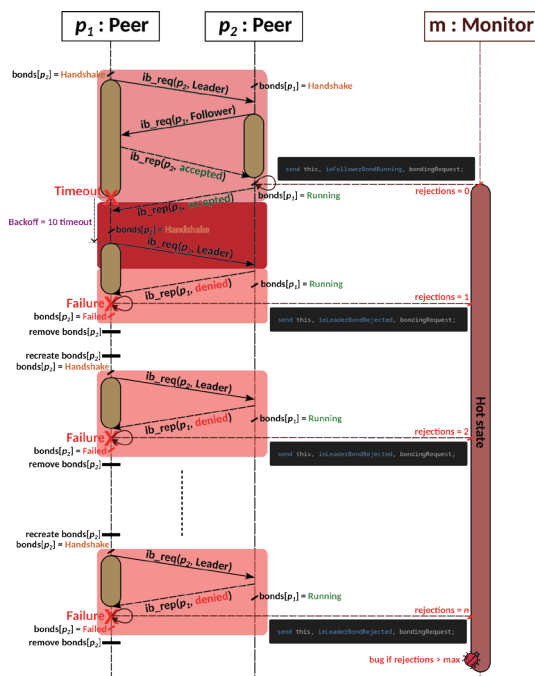
```

graph TD
    A[Application Layer  
Role specific logic, subscribe to specific messages]
    B[Communication Layer  
gRPC based bonds (connections) management, message propagations]
    C[KV-Storage Semantics  
Process message in RAM, version conflict resolving]
    D[Persistence Layer  
Local Disk, Remote Storage]
    A --- B --- C --- D
  
```

The diagram illustrates the four layers of the gRPC architecture, stacked vertically. Each layer is represented by a colored rectangular box with a title and a description.

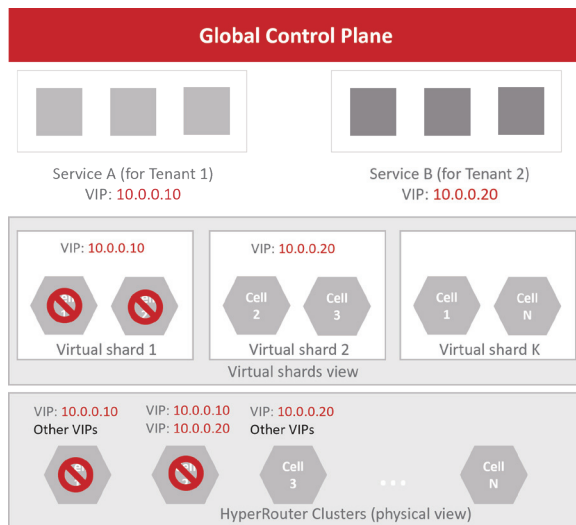
- Application Layer** (Dark Red): Role specific logic, subscribe to specific messages
- Communication Layer** (Red): gRPC based bonds (connections) management, message propagations
- KV-Storage Semantics** (Light Red): Process message in RAM, version conflict resolving
- Persistence Layer** (Pink): Local Disk, Remote Storage

为确保点对点系统的正确性，我们联合爱尔兰云可靠性实验室的形式化方法团队，采用P语言与TLA+工具对管控面协议进行了严格验证。



确定性爆炸半径

在多租户系统中，单个租户的异常若未加管控，可能扩散至整个系统并引发严重故障。为解决这一问题，HyperRouter借鉴了Cell架构与Shuffle Sharding技术。关于这两种技术的细节，我们曾在去年的SRECon会议上，结合华为云爆炸半径控制的方法与实践做过深入分享，本文仅聚焦其在四层负载均衡服务中的具体实现展开介绍。



作为负载均衡服务，HyperRouter无法依赖“Cell Router”实现Cell化，因此需由管控面直接负责Cell与虚拟分区的分配和管理。具体而言，通过将租户划分为多个Cell与虚拟分片，可将故障严格限制在较小范围内；而Shuffle Sharding的分片重叠设计，能进一步降低租户间的耦合概率，从而显著缩小故障爆炸半径。

举例：假设共有10个Cell（ $N=10$ ），若采用静态分区且每个分区包含2个Cell，其爆炸半径为1/5；但通过Shuffle Sharding机制，为每个逻辑负载均衡（LB）分配独立虚拟分区，即便该虚拟分区对应的Cell完全故障，其爆炸半径也仅为 $1/C(10,2) = 1/45$ （其中C代表组合数），故障影响范围大幅缩减。这种设计可有效提升了多租户环境下的隔离能力，是确保确定性运维的关键环节。

（六）Lesson 6: CUJ是一个行之有效的以用户为中心的观测框架和实践

CUJ: 以用户为中心的观测

在SRE的方法与实践中，可观测性是保障系统稳定的关键环节。但面对复杂系统中动辄成千上万的指标、日志与事件，如何精准切入、聚焦核心信号，进而设计合理的SLO（服务等级目标）与告警策略，成为亟待解决的问题。

传统可观测性方案多从系统自身维度（如CPU使用率、网络丢包率）出发，难以全面反映用户实际体验。为此，HyperRouter引入了“关键用户旅程（Critical User Journeys, CUJ）”方法论，以用户为核心重构可观测性体系。

CUJ由Google率先提出，目前已在业界广泛应用（例如Netflix便将其用于游戏服务的可观测性设计）。以HyperRouter管控面“配置下发”的CUJ为例，我们可清晰看到其应用逻辑：用户“新建一个负载均衡（LB）”的操作，会涉及API调用、管控面处理、BGP路由下发、数据面接入四个核心环节。通过对这一全链路的追踪与度量，我们不仅能直观评估用户感知的性能（如新建LB的耗时），还能发现传统监控未覆盖的盲区——例如DPDK内部交互延迟这类直接影响用户体验却易被忽略的问题。



Control Plane CUJ: The system receives the LB configurations from an external system (the UI which allows the user to enter them) and populates the new virtual LB routes on the BGP.

CUJ提供了一种面向用的可观测方法，使指标、SLO、监控等设计更贴近用户的实际体验。

三、HyperRouter实践核心经验总结

HyperRouter作为华为云全局负载均衡的关键组件，其设计、开发与运维过程，为云原生领域分布式系统的“确定性运维”落地沉淀了可复用的核心经验，具体可归纳为六大维度：

1.跨服务协同是简化复杂度的核心能力

四层与七层负载均衡的组合部署是常见场景，通过明确HyperRouter的定位（聚焦入向流量转发），结合Google Maglev一致性算法（实现无状态精准转发）与DSR模式（剥离出向流量处理），可实现跨服务无状态联动——既解决七层有状态服务的容错短板与ECMP重哈希断连问题，又降低四层负载均衡的资源开销，为系统整体可靠性奠定基础。

2.技术选型需理性权衡“单向门决策”

数据面框架的选择直接决定性能上限，Linux内核网络栈存在高性能瓶颈，XDP虽低耗但灵活性与双网卡适配性不足；而DPDK虽需专用CPU/NIC且CPU利用率高，却能满足“单核千万级PPS、纳秒级报文处理”的高性能需求，且适配可演进目标。这种“性能-资源-灵活性”的权衡逻辑，可迁移至其他对性能与可扩展性有要求的组件选型场景。

3.自愈与隔离必须前置为设计“一等公民”

分布式系统中故障不可避免，需从设计阶段将自愈与隔离能力纳入核心逻辑：通过Controller持续拉取节点状态（数据面指标、BGP状态等）实现健康度动态判断，用“BGP AS Path调整”替代“路由撤销”处理故障（避免误判导致的转发中断），既能快速转移流量，又能降低故障扩散风险。这种“主动探测-精准决策-最小化影响”的机制，比事后补丁式方案更高效可靠。



4.点对点架构适配高可用与分区容错需求

基于 CAP 理论，HyperRouter 管控面优先保障可用性（A）与分区容错性（P），采用点对点去中心化架构（含存储层、KV 语义层、通信层），使节点可独立完成状态发现与同步——即便出现网络分区，“孤岛”节点仍能自主决策，避免中心化组件依赖引发的级联故障；搭配 P 语言与 TLA+ 形式化验证，可提前规避状态同步无限重试等深层隐性问题。

5.Cell+Shuffle Sharding 实现确定性爆炸半径

多租户系统需严格控制故障影响范围，HyperRouter 通过管控面直接分配 Cell 与虚拟分片，结合 Shuffle Sharding 的动态分片设计，将故障爆炸半径从静态分区的 1/5（10 个 Cell、单分区 2 个 Cell）压缩至 1/45（组合数计算），有效隔离单租户异常，为多租场景下的“确定性影响范围”提供技术支撑。

6.CUJ 重构可观测性以对齐用户体验

传统系统维度监控（如 CPU、丢包率）无法精准反映用户感知，通过 CUJ 方法论（如“新建负载均衡”全链路：API 调用→管控面处理→BGP 下发→数据面接入门），可将监控指标与用户核心操作绑定，既能直观度量用户体验（如新建 LB 耗时），又能发现 DPDK 内部交互延迟等传统监控盲区，让可观测性真正服务于用户体验保障。

四、总结

本文以 HyperRouter 为实践载体，阐述了“确定性运维”理念从抽象目标到具象落地的全流程。这一实践充分证明，系统可靠性的核心并非单纯依赖冗余配置或防御性设计，而是源于结构清晰的架构规划、理性严谨的技术权衡，以及可度量的闭环反馈机制。

展望未来，随着云基础设施向更高带宽、更新协议、更智能硬件演进，HyperRouter 将持续优化性能、适配新场景，进一步拓展确定性运维的应用边界，为云服务全生命周期的稳定运行提供更坚实的技术支撑，也为行业可靠性实践提供更多可参考的落地范式。



SRE深度思考：云演进新阶段， 高阶服务与客户成功定义核心竞争力



沈幸峰
华为云SRE
运维架构师



李昂
华为云SRE
运维平台负责人



李川
华为云SRE
运维平台架构师

摘要：云计算已从提供基础设施资源，过渡到高阶服务与业务深度融合，并以卓越运维能力保障客户成功的阶段，这不仅是技术的演进，更是客户需求驱动下的产业升级。本文将分析云计算发展及客户用云模式的变化，洞察云服务核心竞争力，并探讨领先云厂商如何助力行业用好云，助力千行百业真正“用好云”，实现商业价值跃迁。

一、引言：云的演进——从资源池到价值引擎

回顾云计算的发展历程，其角色和重心经历了深刻的转变，已经从资源提供者转变为了价值共创伙伴。宛如一

条奔腾的河流，不断拓宽、加深，最终汇入价值创造的海洋。



1.1 奠基时代：基础设施的普及（Cloud 1.0）

最初，云计算着力于解决企业自建IDC成本高昂、资源闲置、扩容困难的痛点。云厂商的核心任务是提供稳定、弹性、经济的计算、存储、网络资源（IaaS）。该阶段的目标是修建信息高速公路，“让资源触手可及”，核心竞争力在于规模、成本和基础的稳定性。



1.2 深耕时代：平台服务的崛起（Cloud 2.0）

当基础设施不再是稀缺品，客户的需求开始上移。他们不再满足于简单的虚拟机，而是渴望更高效的应用开发、部署和管理能力。于是，数据库、中间件、大数据、AI等平台服务（PaaS）应运而生。云厂商开始提供更“趁手”的工具，如同在高速公路上提供高性能跑车和智能导航，帮助客户更快地构建应用、抵达目的地。服务的“丰富度”和“易用性”成为新的竞争焦点。

随着基础设施资源的普及，客户对高效应用开发、部署和管理的需求日益增长，促使数据库、中间件、大数据、AI等PaaS服务的出现。为满足应用开发与部署需求，云厂商提供工具，竞争力转向服务的丰富度与易用性。



1.3 共创时代：赋能业务成功的新篇章（Cloud 3.0 – 现在与未来）

如今，我们正迈入一个更激动人心的新阶段。随着企业核心业务系统、创新应用全面拥抱云，挑战不再仅仅是“上云”，而是如何在云上持续构建、运行和优化这些日益复杂的系统，确保它们不仅稳定可靠，更能敏捷迭代、创造价值。客户需要的不再仅仅是资源或工具，而是能够帮助他们驾驭云端复杂性、保障业务连续性、加速数字化转型的“智慧”和“能力”。随着基础设施资源的普及，客户对高效应用开发、部署和管理的需求日益增长，促使数据库、中间件、大数据、AI等PaaS服务的出现。为满足应用开发与部署需求，云厂商从提供工具、竞争力转向提高服务的丰富度与易用性。

云厂商的角色，正从“供应商”转变为“价值共创伙伴”，核心使命是赋能客户“用好云”，驱动其业务成功。

随着企业核心业务系统与创新应用全面上云，关注点从“上云”转向云上系统持续构建、运行与优化，需求集中在稳定性、可靠性、敏捷迭代与价值创造。客户需要可应对云端复杂性、保障业务连续性、推进数字化转型的能力。云厂商角色从供应商转变为价值共创伙伴，核心使命是赋能客户“用好云”，助力其业务成功。



二、大势所趋：云服务从供给侧往需求侧进化，与客户用云模式同频共振

随着数字化转型的加速推进，企业对云服务的需求正在发生深刻变革。早期的“上云”更多是基础设施层面的迁移，是一种技术替代的过程；而如今，企业的用云模式已经进入深水区，呈现出从“搬迁上云”到“生于云端”，再到“业务与云深度融合”的演进路径，客户需求也逐步从功能满足向能力构建跃迁。

2.1 从“搬迁上云”到“生于云端”

早期的“上云”多是将传统应用原封不动地搬到虚拟机上（Lift & Shift），并未充分利用云的优势。而现在，越来越多的新应用从诞生之初就基于云原生架构设计，充分利用容器、微服务、Serverless等技术，实现敏捷开发、弹性伸缩和持续交付。这种“生于云端”的模式，对云平台的服务能力和整合度提出了更高要求。



2.2 业务与云深度融合的挑战

当企业的核心交易系统、大数据分析平台、AI训练集群等关键业务负载都运行在云上时，应用架构变得极其复杂，分布式、微服务化成为常态。系统的稳定性、安全性、性能优化、成本管理等运维挑战也随之指数级增长，对确定性运维诉求强烈。任何一个环节的疏漏，都可能对业务造成严重影响。所谓的确定性运维，就是通过面向云服务全生命周期的质量看护，实现现网可预期的高可用质量结果的一种运维理念以及质量管理体系。



2.3 呼唤“用好云”的智慧与能力

面对云上业务的复杂性，企业仅仅依靠自身力量构建和维护一支庞大、高水平的运维团队既不经济，也难以跟上技术发展的速度。他们迫切需要云厂商不仅提供先进的技术，更能输出成熟的方法论、高效的工具链以及专业的支持服务，帮助他们建立起与云原生时代相匹配的运维治理体系，实现确定性运维。

企业应对云上业务复杂性，靠自身构建和维护庞大高水平运维团队，不经济且难跟上技术发展速度。企业希望云厂商除提供技术外，还能输出成熟方法论、高效工具链和专业支持服务，以助其建立适配云原生时代的运维治理体系。

客户需要的，是一种能够将云厂商自身大规模运维经验转化为客户侧能力的“赋能”。

三、核心洞察：云的核心竞争力将有高阶服务和专业服务重塑

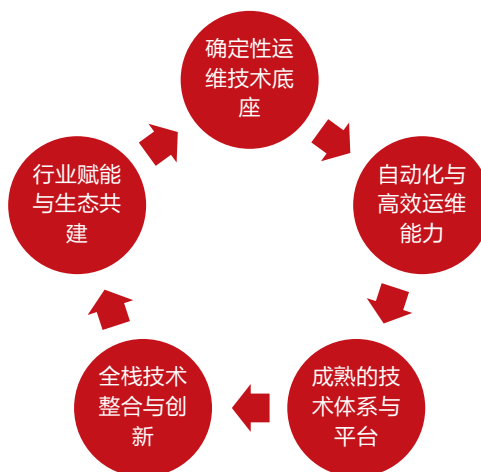
3.1 高阶服务：加速创新的“倍增器”

在数字化转型浪潮中，速度就是生命。客户需要能够快速响应市场变化、迭代业务逻辑的强大武器。数据库、AI平台、大数据服务、应用中间件等高阶云服务，正是这样的“倍增器”。它们将复杂的技术封装起来，让客户可以像搭积木一样快速构建应用，聚焦于业务创新本身，而不是底层技术的研发和维护。**云厂商提供高阶服务的能力，直接决定了其赋能客户创新的深度和广度。**

在数字化转型进程中，快速响应市场变化、迭代业务逻辑至关重要。数据库、AI平台、大数据服务、应用中间件等高阶云服务，可封装复杂技术，助力客户高效构建应用，使其专注于业务创新而非底层技术运维。云厂商的高阶服务能力，直接决定其赋能客户创新的成效。

3.2 SRE赋能：驾驭复杂性的“稳定锚”

当业务全面构建于云上，系统的可靠性、韧性成为生命线。SRE（Site Reliability Engineering，站点可靠性工程）作为一套源于互联网巨头应对大规模、高复杂度系统挑战的先进理念和实践，其核心在于通过工程化、自动化的手段保障服务稳定运行。**未来的云厂商，不仅要自身具备强大的SRE能力来保障平台稳定，更关键的是，要具备将这种能力“外溢”给客户的机制和意愿，通过提供智能化的运维工具、专业的咨询服务、甚至联合运维等方式，帮助客户构建起强大的云上运维体系，成为他们驾驭复杂性的“稳定锚”。**





3.3 两者融合：云厂商的终极价值主张

领先的高阶服务与卓越的SRE赋能，两者相辅相成，缺一不可。前者是创造价值的引擎，后者是保驾护航的基石。能够将这两者完美融合，提供从技术创新到稳定运行的全方位支撑，帮助客户在云上安心发展、持续成功的云厂商，才能真正赢得客户的长期信赖，构建起难以逾越的竞争壁垒。两者的融合是价值创造与风险控制的统一，高阶服务让客户敢于创新、快速试错，而SRE赋能则为这种创新提供了坚实的安全网，确保创新带来的不是混乱而是持续的价值输出；两者的融合是从“能用”到“好用”再到“持续领先”的进化，高阶服务解决了“能用”的问题，让复杂技术触手可及。而SRE赋能则致力于提升服务的“好用”程度——更高的可用性、更低的延迟、更强的弹性；两者的融合是真正的“价值共创”：云厂商提供的不仅仅是工具（高阶服务），更是确保这些工具能持续、高效、稳定地创造价值的能力（SRE赋能）。这种“技术创新”与“稳定运行”的双重承诺，使得云厂商从单纯的资源提供者，转变为与客户共同成长的价值共创伙伴。

这正是云服务从“资源提供”走向“价值共创”的终极体现。

四、华为云的探索与实践：与时代同频，与客户共生

作为云计算领域的积极探索者和实践者，华为云始终紧随产业发展脉搏，深刻理解客户需求的变化，致力于成为客户数字化转型道路上最值得信赖的伙伴。

4.1 技术引领：构建满足未来需求的云服务

面对客户对加速创新和深度用云的渴望，华为云持续投入，打造了一系列业界领先的高阶服务，如分布式数据库GaussDB、云容器引擎CCE Turbo、AI开发生产线ModelArts等。这些服务并非空中楼阁，而是深度结合了华为自身多年复杂业务场景的实践和对企业级需求的理解，旨在为客户提供真正强大、易用、可靠的技术底座，这是对“高阶服务成为创新倍增器”这一行业趋势的积极响应。

4.2 能力外溢：从“经验沉淀”到“客户赋能”

当企业将核心业务迁移上云，挑战才刚刚开始——复

杂的分布式架构如何设计？突发的流量洪峰如何应对？微服务间的调用链路如何保障稳定性？这些问题已远超传统运维的范畴。云厂商的使命，正是将自身在大规模业务实践中积累的SRE经验，转化为客户可感知、可落地的能力，帮助客户跨越“上云”与“用好云”之间的鸿沟。

4.2.1 场景一：架构调优——让业务与云“深度适配”

许多企业在迁移上云初期，往往简单复刻原有架构，导致资源浪费、性能瓶颈等问题。某金融客户曾因数据库读写分离设计不合理，在业务高峰期出现严重延迟。华为云SRE团队通过架构健康度评估工具OA，结合对底层链路拓扑的感知、应用SLA评估的经验，帮助客户重构读写链路，引入高可用架构调优方案，最终识别客户的SLA风险点，将系统可靠性提升10倍，并将系统吞吐量提升3倍，延迟降低80%。云厂商的价值在于：

- 提供架构评估模型，识别单点故障、资源瓶颈等隐患
- 基于行业特性（如金融交易低延迟、电商大促高并发）输出最佳实践
- 通过灰度发布、混沌工程等手段验证架构韧性

4.2.2 场景二：隐患分析——为业务戴上“预警头盔”

某企业曾因未及时清理日志导致磁盘爆满，引发核心生产线停摆。这类“小问题引发大故障”的案例屡见不鲜。华为云将内部运维中沉淀的隐患预测算法开放给客户：

- 通过算法智能分析历史故障数据，预判潜在风险（如配置错误、容量不足）
- 提供风险画像看板，量化评估业务模块的脆弱性等级
- 自动生成修复建议（如自动扩容策略、冗余链路配置）

这种“治未病”的能力，让客户运维从“救火式响应”转向“预防性治理”。

4.2.3 场景三：自助诊断——让问题定位“开箱即用”

过去，客户遇到性能抖动或异常告警时，往往需要反复提交工单、等待云厂商支持团队介入排查，效率低下。某零售客户在促销期间遭遇API接口响应延迟，通过华为云提供的自助诊断平台COC，仅需三步操作：

- **智能根因分析：**平台自动关联日志、指标、链路追踪数据，定位到某微服务线程池配置不足；
- **可视化拓扑：**展示异常服务的上下游依赖，快速确定影响范围；
- **一键修复建议：**根据历史最佳实践，推荐线程池参数调优方案，并支持自动修复。

整个过程从发现问题到解决仅耗时10分钟，而传统方式可能需要数小时甚至更久。这种能力的背后，是华为云将内部SRE日常故障排查的决策逻辑（如根因定位算法、影响面分析模型）封装为标准化工具，让客户无需深究技术细节即可自主解决问题。

4.3 伙伴同行：共赴业务成功的星辰大海

我们将进一步深化AI与云的融合，推动AI从“可用”走向“易用”和“精用”，构建更加普惠、低门槛的AI基础设施和开发平台，帮助客户在业务的每一个环节都能便捷地嵌入和利用AI能力。我们将推动云计算与物联网、边缘计算、数字孪生等技术的深度集成，打破数字世界与物理世界的边界，助力客户实现物理资产的数字化管理、生产流程的智能化升级以及全新交互体验的创造。我们将把安全与合规置于最高优先级，构建端到端、全链路的数据安全与治理体系，为客户提供一个能够安心上云、放心用数、合规创新的“数据绿洲”，确保其数据主权和商业机密得到最高级别的保护。我们将进一步提升云平台的开放性和灵活性，提供更加模块化、可编排的云服务和更高效的DevOps工具链，深化与行业伙伴的生态合作，打造更丰富的行业解决方案和应用市场。赋予客户极致的业务敏捷性，使其能够以前所未有的速度响应市场变化、进行业务创新和模式迭代。通过与生态伙伴的协同，客户可以快速获取所需能力，实现更高效的价值共创和商业成功。

华为云坚信，云厂商的成功源于客户的成功。我们追求的不是简单的买卖关系，而是与客户建立长期、共生的伙伴关系。通过持续的技术创新和服务升级，通过开放的生态合作，我们希望与千行百业的客户一起，共同探索云的无限可能，共赴业务成功的星辰大海。

五、结语：选择远见，拥抱未来

云计算的浪潮奔涌向前，其核心价值正从提供资源，转向赋能成功。以领先的高阶服务为引擎，以卓越的SRE能力为保障，深度融入客户业务，助力其在云上构建韧性、敏捷、智能的现代化应用——这已成为衡量顶尖云厂商价值的新标尺。华为云，正以其前瞻的技术布局、深厚的行业理解以及“以客户成功为中心”的服务理念，积极拥抱这一变革。选择与具备远见、能够提供“高阶服务 + SRE赋能”综合价值的云伙伴同行，无疑是企业在数字化浪潮中把握先机、赢得未来的明智之选。

随着云计算的不断发展，其核心价值已从单纯提供资源转向赋能客户成功。如今，顶尖云厂商的价值衡量标准在于能否以高阶服务为引擎，以卓越的SRE能力为保障，深度融入客户业务，助力客户在云上构建韧性、敏捷、智能的现代化应用。华为云凭借前瞻的技术布局、深厚的行业理解和“以客户成功为中心”的服务理念，积极拥抱这一变革。企业在数字化转型中，选择具备“高阶服务+SRE赋能”综合价值的云伙伴，如华为云，无疑是把握先机、赢得未来的明智之举。



以智能可观测技术构建数据驱动的主动治理体系



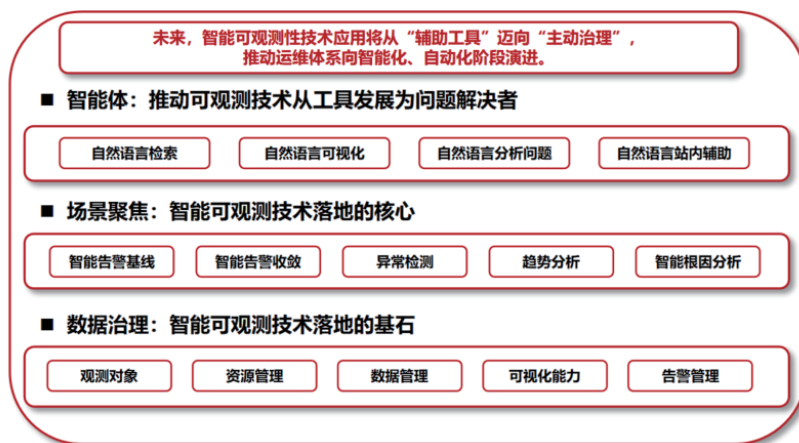
郑立
信通院云大所副主任

摘要：在云原生和AI技术发展背景下，智能可观测技术成为解决动态系统治理难题的关键。中国信通院开展可观测性标准研究，为企业提供参考。未来，该技术将向自主治理演进，在运维模式等十个维度呈现发展趋势。

一、引言

随着云原生架构的普及和AI技术的快速发展，系统的复杂性与数据规模呈指数级增长，传统运维模式已难以满足高效、精准的稳定性保障需求，在此背景下，智能可观测性技术成为解决动态系统治理难题的关键。本文基于中

国信通院在可观测性领域的标准研究，结合行业技术演进方向，探讨如何以智能可观测性技术构建数据驱动的主动治理体系。



二、背景

1.1 AI时代下，可观测技术与智能化技术加速融合

AI与可观测性技术的结合，旨在通过数据驱动实现“观测、分析、决策、优化”闭环，实现系统稳定性保障从“人工经验主导”向“算法协同驱动”的范式跃迁。其核心价值如下：

1) 复杂系统与数据规模的治理需求

微服务、容器化等技术促使系统拓扑结构日趋动态化，传统人工监控难以应对海量指标（Metrics）、日志（Logs）、追踪（Traces）等多模态数据的实时分析。通过智能化技术实现数据自动化关联与异常模式识别，可显著降低人工干预成本，提升复杂环境下的运维鲁棒性（鲁棒性是指系统在面对各种内外部干扰时，仍能保持其性能稳定的能力）。

2) 诊断能力强化与根因定位优化

基于算法与模型的自动化分析机制可初步实现问题根因的快速定位，尤其在固定场景（如服务链路中断、资源过载）中，其准确率接近专家经验水平。此类技术通过因果推理与历史数据比对，减少故障排查对个体经验的依赖，缩短平均修复时间。

3) 实现未知问题的探索和创新

针对低频异常事件与非结构化数据（如日志文本），可借助无监督学习与模式识别等智能技术，实现对系统潜在问题的主动探索与发现，如业务逻辑缺陷或安全漏洞等。这类能力突破了传统阈值告警只能识别已知异常的局限，推动系统在未知问题识别与创新性诊断方面迈出关键一步。

4) 优化资源利用与运维效率

通过时序预测与自适应算法，智能化可观测系统可动态优化资源分配策略，实现负载均衡与成本控制，同时，其预测性运维能力可将故障响应从“事后补救”转向“事前预防”，推动运维模式从被动响应升级为主动治理。

1.2 凝聚共识，中国信通院联合行业持续开展可观测性标准研究

中国信通院自2021年起牵头开展可观测性技术及标准研究，陆续发布《可观测性平台能力要求》、《根因分析平台能力要求》、《可观测性能力建设成熟度模型》、《可观测性技术发展研究报告》等多项标准和研究成果，2024年12月，中国信通院联合行业头部企业共同发布《云计算系统智能化可观测性能力成熟度模型》，并在中国通信标准化协会完成行业标准立项，这些标准为企业提供了技术建设的路径参考和能力评估框架。

三、数据治理：智能可观测技术落地的基石



在智能可观测体系的构建过程中，数据治理是实现高质量智能模型落地的坚实基础。智能模型的有效运行依赖于“数据+算法+算力”的协同驱动，其中数据作为基础要素，其质量直接决定了模型的可用性与泛化能力。在实

际应用中，数据质量面临诸多挑战，如故障数据稀缺、数据来源杂乱、数据模型不统一以及持续标注维护难度大等问题，这成为智能化可观测技术落地过程中最大的瓶颈。在此背景下，我们需要构建可观测底座能力，通过被观测对象的管理、资源管理、数据管理（包括数据采集、标准化、存储、安全与处理）、构建可视化能力、告警管理等一系列环节，实现全面、有效、持续、高质量的可观测数据的获取和有效治理，为智能化技术的应用提供坚实的基础。

四、场景聚焦：智能可观测技术落地的核心

在数据治理的基础上，智能可观测能力的落地需进一步结合实际场景开展探索与应用。当前，智能化可观测技术的建设路径呈现出明显的“任务驱动”或“场景驱动”特征。不同系统场景下的业务需求差异巨大，因此往往需要构建具备针对性的算法模型，既可以基于专家规则进行定制，也可以通过端到端的AI模型进行自动学习与推理。这种灵活建模机制使得智能可观测系统能够适配多样化的场景需求，实现从感知、诊断到优化的全过程闭环。

3.1 智能告警基线

智能告警基线可以有效解决误告警的问题，它是指针对大量观测对象的指标时域变化趋势自动生成告警基线，基线随业务峰值特点等进行动态调整的能力。在很多业务场景下，如果我们的告警阈值定的过高，就会漏掉真实的故障问题，阈值过低，就会产生大量无效告警。因此可以先通过区分业务场景、区分昼夜差异，在空间维度和时间维度设置不同的基线使得告警更加准确，也可以设置基于规则的动态基线，最后，可以使用智能化的手段生成定制化、动态化的告警基线，如从历史数据挖掘动态阈值，面向不同指标类型自适应的生成告警基线，基于历史数据天、周、月变化趋势的生成告警基线，对未超动态阈值的突变、畸变，智能产生告警等。

3.2 智能告警收敛

智能告警收敛的价值在于对真实问题告警的合并。它是指通过智能化的算法和策略对海量的告警事件进行关联分析、合并和智能抑制的能力。比如，可以在设置一些规则在同维度进行收敛，或基于识别的问题根因、基于同一根因和依赖关系进行收敛，以及利用AI算法和策略对海量的告警事件进行关联分析、合并和智能抑制。

3.3 异常检测

异常检测是根据历史可观测性数据进行数据分析和训练得到AI模型，在系统运行时，从监控数据中使用训练得到的AI模型进行实时推理发现异常，为后续的诊断提供基础。传统的异常检测模型存在依赖个人经验、配置复杂、随业务变更要不断调整、受大促和异常值影响大的痛点，融合AI能力的智能化异常检测在突增、突降等异常场景下效果好，比较适用于关键业务风险感知和变更风险的感知。

3.4 趋势分析

相对于异常检测，趋势预测是一种主动运维方法，它通过对时间维度的数据趋势分析和可能出现的故障分析来预测趋势劣化和故障发生可能性，在系统亚健康状态下能够提前发现系统隐患，主动开展健康度巡检，防患于未然。该能力适合在基础监控指标的预测、业务预测和容量规划场景落地。

3.5 智能根因分析

在异常检测或趋势预测系统发现故障或劣化趋势后，需要使用根因分析进行定界定位。智能根因分析是通过数据挖掘、数据分析，运用机器学习或大模型方法，自动或辅助识别问题根因。在传统根因分析方法的基础上，可以通过以下方式增强应用效果：

1) 使用大语言模型（LLM）赋能根因分析：通过大模型结合小模型和Agent，使用带有样例的提示词，推理得到根因/根因排序的能力；

2) 根据RAG（检索增强生成）推荐处理建议：使用RAG结合历史故障案例和故障知识图谱（知识库）给出推荐处理建议的能力；

3) 根因分析模型离线/在线更新：如大模型通过预训练完成，小模型和提示词可以离线/在线重新进行数据分析和训练完成更新。

五、推动可观测技术从工具发展为问题解决者

智能体 = 大模型 + 规划 + 记忆 + 工具使用 +

“脑”+“手”
行动智能

“脑”
认知智能

与传统的AI工具不同，智能体不仅具备决策能力，还能够执行相关任务，类似于给“大脑”配备了“手”，使其具备了更加全面的行动力。因此，智能体的出现推动了AI从单纯的工具角色升级为“自主问题解决者”，这一变革为各行业的可观测性技术提供了新的发展方向，业界当前不断地涌现出更多的落地场景和应用，如：

1) 自然语言检索：利用自然语言直接查询和检索数据库或数据仓库中的数据的技术。

2) 自然语言可视化：利用自然语言生成数据可视化的技术，旨在让用户通过简单的文字描述生成符合意图的数据图表。

3) 自然语言分析问题：通过自然语言输入自动识别、诊断并分析问题的技术。

4) 自然语言站内辅助：通过自然语言对话为用户提供实时帮助的技术。

六、未来展望

1.从AIOPS到
ChatOPs(运维模式)

6.从告诉到决策
(信任程度)

2.从小场景到全链路
(场景覆盖)

7.从静态知识到实时更新
(认知迭代)

3.从单模到融合
(模型选择)

8.从执行者到开发者
(岗位职责)

4.从被设定到自适应
(策略选择)

9.从单智能体到多体协同
(协作模式)

5.从黑盒到白盒
(思考过程)

10.从智能助手到数字
员工(角色定位)

未来，随着AI与大模型的深度渗透，智能可观测性技术应用将从“辅助工具”迈向“自主治理”，推动运维体系向更高层的智能化、自动化阶段演进，在此，我们从十个维度分析给出智能可观测性技术的发展趋势，具体如下：

1) 在运维模式方面，从AIOPS到ChatOps：演进路径为人工运维、自动化运维、平台化运维、AIOps、ChatOps。

2) 在场景覆盖方面，从小场景到全链路：智能可观测技术的应用场景将持续增加，最终覆盖故障前、中、后的全链路场景，甚至将全面覆盖IT运维的全生命周期流程。

3) 在模型选择方面，从单模到融合：机器学习算法对于数据处理有天然优势，大模型对于语义识别和逻辑推理有天然优势，两者相结合会更好地融合场景落地。

4) 在策略选择方面，从被设定到自适应：系统能够根据环境的变化和数据反馈，自动调整自身的行为、策略或参数。

5) 在思考过程方面，从黑盒到白盒：传统模型存在可解释性差的问题，思维链（CoT）可以展示思考的逻辑步骤，使得故障定位/根因分析过程可追溯、可验证。

6) 在信任程度方面，从告知到决策：当前的智能可观测技术是告诉人类问题原因，未来是知道原因后可以自主决策和执行下一步动作。

7) 在认知迭代方面，从静态知识到实时更新：当前大模型无法实现参数的实时更新，未来将和人类一样，根据不断变化的信息实现自驱的认知迭代。

8) 在岗位职责方面，从执行者到开发者：传统的运维工程师主要负责监控系统、排查故障、处理告警等日常运维任务，角色逐渐转变为智能运维系统和Agent的开发者。

9) 在协作模式方面，从单智能体到多体协同：在多个Agent系统中，多个AI Agent协同工作，通过管理者和指挥者的协调，实现故障诊断和运维活动的增强。

10) 在角色定位方面，从智能助手到数字员工：当前的智能化技术只能充当助手角色，未来逐步承担更多的人类职责。



美宜佳AI应用实践，赋能万店智慧零售新业态



代迪

美宜佳控股有限公司CIO

摘要：美宜佳作为中国知名便利店品牌，通过AI的深度应用，赋能其遍布全国的万店网络，引领智慧零售新业态。本文主要介绍美宜佳AI实践探索的思路，聚焦构建AI与数据中台，打通内外部数据，落地智慧门店、数字员工和门店订货等具体业务场景，有效提升供应链效能与门店服务体验。

一、背景与挑战

美宜佳作为中国门店数排名第一的便利店品牌（数据来源于：CCFA）的便利店品牌，始终紧跟消费升级趋势与技术发展浪潮，坚持以特许加盟为核心的发展体系，并依托该模式将加盟店网络广泛覆盖至240余座城市。在品牌发展历程中，美宜佳持续推进自我革新，先后完成三次商业模式迭代与四次信息系统升级，为规模化扩张与数字化转型奠定坚实基础。截至2025年7月，美宜佳门店数量已突破4万家。随着规模与数字化程度的同步提升，美宜佳全面加快AI建设新征程，然而，从规模化扩张到数字化深耕、再到AI化升级的进阶中，美宜佳也面临着诸多亟待解决的挑战：

数据孤岛与数据治理复杂：海量数据分散在不同平台，加盟店数据采集标准不统一，不仅造成数据质量不稳定，还导致营销管理智能化水平偏低，数据孤岛问题日益凸显；

供应链效率待提升：订货模式过度依赖人工经验，造成高频缺货与高库存积压并存问题，难以适配各区域之间的需求差异，进一步制约了整体效率；

运营成本高：随着门店规模突破万店，一方面人力成本呈现刚性上涨态势，租金支出也持续攀升；另一方面，随着规模扩大，加盟管理的隐性成本也在不断增加，这些因素共同推高了整体运营成本。



AI&数据中台架构图

二、方案实践

面对数字化时代的多重挑战，美宜佳携手华为云，将“AI中台”作为破局的关键：首先以门店高频场景为切入点，打造实时、高质、融合、标准且安全的数据基座，把数据沉淀为可共享和可复用的企业资产，其次打通全供应链数据，既提升运营效能又实现成本精细化管理，最终使AI技术成为企业持续领跑的新生产力。

AI&数据中台，整体架构主要包括AI模型资产、AI服务平台、数据中台、数据源、算力&IT平台等模块：

AI模型资产：围绕门店订货、根因定位、指标预测等高频场景，沉淀一套可迭代的算法资产：线上A/B测试持续调优，线下版本管理与效果监控闭环运行，让模型“训练-评估-升级-复用”全流程运行，在追求业务最优的同时，实现模型资产的统一沉淀与可持续治理。

场景一：智慧门店—AI实践重塑“人、货、场”协同



人工智能正深度重塑新零售，以消费者体验为中心的数据驱动业态迅速崛起，智慧门店成为新零售主战场。美宜佳第六代店型应运而生，通过构建全智能、全自助、全天候的运营体系，抢占市场先机。

智慧门店采用“AI店长+数字店员+云销售+店务员+屏幕营销”的数字运营模式，将AI能力植入到门店整体运作流程中：

数据融合：聚焦业务主航道，建立数据质量基线，贯通会员、门店、商品、人员等关键主数据全流程，完成数据资产梳理，实现数据流实时运转、按需获取以支撑ROADS体验，通过数据联接，为公司数字化运营提供支撑；

模型应用：部署多模态AI模型，通过增量训练持续优化，形成数据采集-实时决策-效果反馈的智能流转闭环；

AI服务平台：以“人、货、场”为主线，一站式承载模型开发-部署-运维全生命周期，在精准营销、柔性供应和全渠道管理等场景下实现模型快速构建、持续微调与无缝升级；

数据中台&数据源：统一接入、清洗、加工、理解并结构化原始数据，沉淀为同源、实时可视、分级分类且安全可控的数据资产与统一的数据底座；

算力&IT平台：支撑数据处理与模型运行的硬件基础设施，提供一站式算力&IT平台云服务。

美宜佳通过重塑组织模式，构建共享中台，将数据、技术和资源集中沉淀并开放复用，业务单元在充分自主的同时高效协同。依托AI底座，**锁定智慧门店、数字员工和门店订货**三类场景，把算法深度嵌入运营全链路，实现从数据到决策、从算法应用到利润增长的价值提升。

智能陈列：AI模型结合图像识别技术，自动检测商品排面占比、价格签错位等情况，实时捕捉货架状态；

客流热力：依托多源感知层关联客流波动与环境因素，量化促销活动、陈列调整对客流变化的贡献度，识别轨迹热力图，为运营优化提供科学决策依据；

数字报告：借助AI平台智能生成巡检与门店效能诊断报告，以数据可视化呈现为依据，推动业务效能升。

美宜佳的智慧门店通过“模型-数据-算法-知识-运营”的闭环体系，推动3万多家门店从机械化执行迈向有机智能体，实现对门店客流的数字化精准管控，重构门店与消费者的深度连接。

场景二：数字员工—AI Agent重构门店服务链



零售智能体

AIGC与LLM驱动零售变革，智能体成为焦点。美宜佳顺势而为，通过“感知-交互-锁定-服务”四阶流程，将客户需求实时转化为个性化的购物体验。

四阶运作流程：

环境感知层：多模态输入融合，生成动态更新的门店数字孪生体；

决策生成层：预设可执行任务指令包，零售AI数据模型训练、优化、推理等进行决策驱动；

任务执行层：通过穿戴通信设备，在商品流转、顾客服务、环境调控、资金处理等方面接收Agent指引完成任务执行；

效果评估层：量化评估矩阵，采用短期调优与长期进化双通道反馈调优。

在美宜佳零售业务场景中，AI Agent 智能体数字员工的价值突出，具体体现在四大核心优势：

1、开箱即用，部署无忧：仅需简单指令，即可快速激活并投入门店工作，实现从部署到应用的无缝衔接，大幅降低时间与成本投入；

2、行业深耕，模型沉淀：深度整合零售业垂直领域大语言模型，精准把握零售行业特性，快速沉淀并适配智慧门店数字员工场景，为业务决策提供支撑；

3、全流程自动化：将客户咨询接收、行为识别、精准服务等繁琐任务拆解为可执行的微小单元，实现全流程自动化处理，有效减少人为误差，同时提升工作效率与优秀店员经验的复制性；

4、人机协同，灵活应变：数字员工不仅具备强大的自动化能力，还能与业务人员无缝协作，支持在复杂情境下进行灵活调整，共同应对门店顾客需求复杂多变，全面提升运营管理水平，实现业务效益最大化。

作为美宜佳数字化转型的重要突破，智能体数字店员可随场随身解答问题、实现业务流程自动化。其依托垂类大语言模型与智能屏幕理解技术，能灵活、智能且精准地解读不同业务场景指令，实时感知用户需求并快速响应，大幅提升门店消费者的购买体验，构建智能交互的体验闭环。



移动终端+ChatOps，破局传统运维局限



王福强
华为云GTS云服务运维总监



吴祥
华为云GTS云服务自动化运维专家

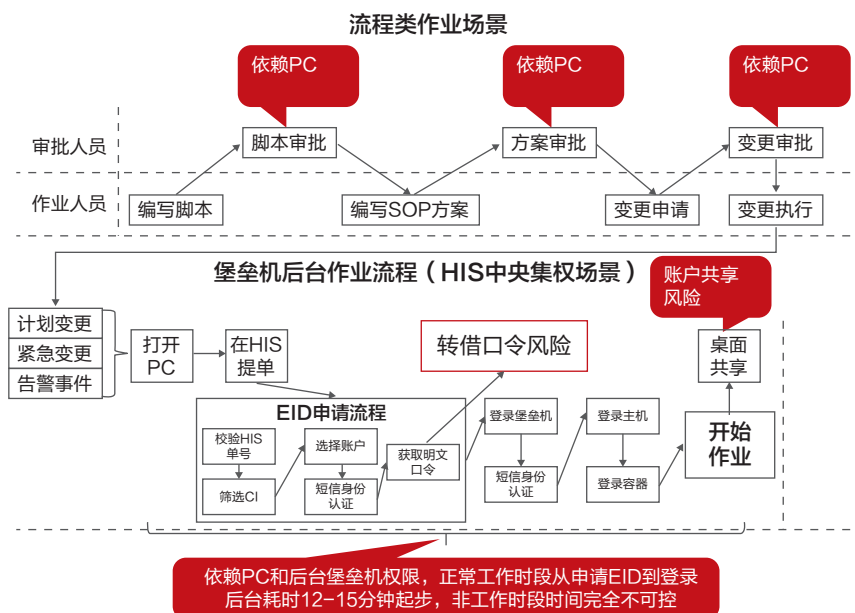
摘要：随着云计算、大数据、人工智能等技术的快速发展，企业IT系统的复杂性和规模性日益增加，这对运维团队提出了更高要求。传统的运维模式已难以满足快速响应、高效协同的需求，因此，移动运维终端和ChatOps作为一种新兴的运维协作方式应运而生。本文讲述了GTS SRE团队引入移动运维终端和ChatOps理念解决团队痛点的实践。

一、引言

随着云计算、大数据、人工智能等技术的快速发展，企业IT系统的复杂性和规模性日益增加，这对运维团队提出了更高要求。传统的运维模式已难以满足快速响应、高效协同的需求，因此，移动运维终端和ChatOps作为一种新兴的运维协作方式应运而生。本文讲述了GTS SRE团队引入移动运维终端和ChatOps理念解决团队痛点的实践。

二、业务背景

GTS SRE团队依托iDOE智能数字化运维使能平台，承担GTS数字化交付/技术作业、伙伴使能、运营运维等交付数字化平台，以及HEC上部署的外销SaaS服务的运维职责，业务部署覆盖HEC/HIS上全球30+数据中心。随着越来越多的交付工具转入团队运维，业务规模快速增长，传统运维模式逐渐暴露出响应慢、协作效率低、问题解决效率低等问题，其中团队面临的如下几个核心痛点更为突出：





1、作业严重依赖PC和后台堡垒机权限

大量作业场景严重依赖PC和后台堡垒机权限，如流程类作业（账号权限审批、工单审批/转派、告警/事件响应等）、变更类作业（升级、搬迁、采集机处理、网络诊断、主机/服务后台配置修改、紧急故障处理等）。人不在PC旁时，各项工作被阻塞；正常工作时段（人在PC旁），申请EID口令到登录主机后台耗时12-15分钟起步，非工作时段（尤其是周末，人不在PC旁）时间完全不可控，紧急故障情况下严重影响故障恢复时长。

2、作业可信风险

大量的后台手工作业场景，方案与执行不一致风险（误操作、漏操作、重复操作、操作顺序错误）、非法授权（违背最小化授权原则、非法提权、越权操作、违规转借口令等）等作业可信风险难以彻底规避。

1）流程类作业场景：如自动化脚本审批、变更方案审

批、变更单审批、事件单响应、权限角色（IDOE角色、华为云权限、堡垒机权限、FS账号权限等）审批、告警处理（确认、清除、派单等）。

2）自动化查询场景：日常告警或事件处理中高频的查

询场景，如服务/进程状态查询、日志查询、参数配置查询、版本信息查询、API/URL拨测检测、网络连通性检测、K8S集群健康检查、慢SQL查询等。

3）自动化变更场景：日常告警或事件处理中高频的快

速恢复手段（或三板斧），如主机重启、堆栈重启、应用重启、Pod重启、主备切换、日志清理、进程/会话查杀、redis缓存清理、应急预案执行等。

4）运维持续改进：运维KPI晾晒、工单晾晒/催办、

告警晾晒/催办、排班、工单信息查询、资产信息查询等。

5）安全运维场景（单列）：如安全告警处理、安全事

件处理、安全态势感知等。

三、方案实践

1、场景分析

结合GTS云服务运维业务现状分析，SRE团队识别出五大类场景急需摆脱对PC及堡垒机后台权限的依赖：

运维持续改进	流程类作业	自动化查询	自动化变更	安全运维
• 运维KPI晾晒 • 工单晾晒/催办 • 告警晾晒/催办 • 工单信息查询 • 资产信息查询 • 排班管理	• 事件单响应 • 事件单关单 • 变更单审批 • 实施单延期 • 服务请求审批 • SOP方案审批 • 问题单响应 • 原子脚本审批 • 告警处理	• 服务/进程状态 • 日志查询 • 参数配置查询 • 版本信息查询 • 拨测检测 • 网络连通性 • K8s集群检查 • 慢SQL查询	• 主机重启 • 堆栈重启 • 应用重启 • Pod重启 • 主备切换 • 日志清理 • 进程/会话查杀 • Agent管理 • 应急预案执行	• 安全告警 • 安全事件 • 安全态势 • 安全账号纳管 • 安全防护 • 版本合规

2、方案选型

经过对各类应用场景的落地成本和易用性进行评估，
最终决策按照如下策略进行方案选型：

业务场景	场景描述	iDOE接口成熟度	移动终端	ChatOps
流程类作业场景	如自动化脚本审批、变更方案审批、变更单审批、事件单响应、权限角色（iDOE角色、华为云权限、堡垒机权限、FS账号权限等）审批、告警处理（确认、清除、派单等）	成熟	适合	
自动化查询场景	日常告警或事件处理中高频的查询场景，如服务/进程状态查询、日志查询、参数配置查询、版本信息查询、API/URL拨测检测、网络连通性检测、K8S集群健康检查、慢SQL查询等	不成熟	部分适合	适合
自动化变更场景	日常告警或事件处理中高频的快速恢复手段（或三板斧），如主机重启、堆栈重启、应用重启、Pod重启、主备切换、日志清理、进程/会话查杀、redis缓存清理、应急预案执行等	不成熟	部分适合	适合
运维持续改进	运维KPI晾晒、工单晾晒/催办、告警晾晒/催办、排班、工单信息查询、资产信息查询等	成熟	部分适合	部分适合
安全运维场景	如安全告警处理、安全事件处理、安全态势感知等	成熟	适合	

其中移动运维终端能力选择基于GDE Link平台（GDE面向开发者提供的低代码移动端开发平台，用户只需低门槛的拖拉拽方式就能高效地构建移动办公解决方案）构建，而ChatOps能力则基于公司内部2023年的爆款内源Welink机器人应用小鲁班（不依赖于运行环境，不需要安装脚本，统一通过welink实现消息发送与回复；提供了一个自定义插件的框架，用户可以基于该框架自行实现小鲁班不具备的能力）构建。

3、业务设计与实现

(1) 基于GDE Link的移动终端运维

如前文所述，移动终端运维能力主要聚焦日常流程类作业（如各类工单或账号权限审批，工单转派、待办查询、告警响应等）以及安全运维场景。对于一些常见的固化的自动化操作流程，如拨测、Agent重启等也在移动运维终端提供了便捷入口。



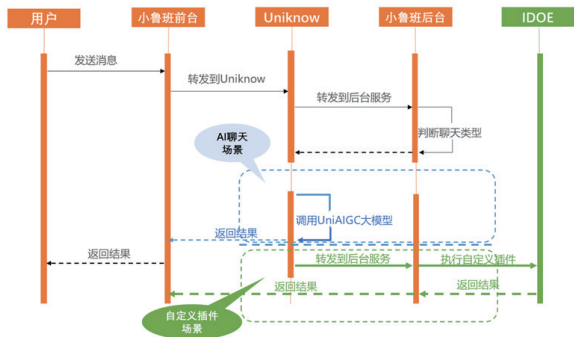
(2) 基于小鲁班的ChatOps运维

ChatOps运维能力基于小鲁班Welink机器人服务构建，考虑到现网运维面临的操作不确定风险，当前未使用小鲁班的AI能力，而是基于小鲁班的插件框架开发自定义插件调用iDOE平台的接口完成一些典型场景的运维作业，如KPI、变更、事件、告警、CMDB信息查询，以及高频的自动化查询或低风险变更操作。当前SRE团队共实现了如下2个自定义插件：

- /idoe信息查询（支持调用iDOE平台接口查询KPI、变更、事件、告警以及CMDB信息）智能化告警：提供多维度多指标告警，并将告警信息对接客户系统，当系统出现异常时，能及时自动通知相关人员，以便快速处理。
- /idoe执行指令（支持调用iDOE平台接口执行自动化查询或变更指令）

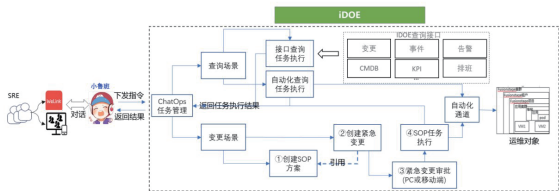
信息交互流程设计

当小鲁班接收到用户发送一条消息后，会首先筛选出此人/群/部门所拥有的功能，然后再根据正则表达式匹配判断此消息触发的是哪一个功能插件（小鲁班自带的通用插件还是用户自行开发的自定义插件），如果都不是，则说明此用户希望使用的是AI聊天，则会调用场景AI的UniAIGC服务来回答用户的问题。



用户旅程设计

对于查询类指令（无论是通过iDOE接口查询iDOE平台的运维数据，还是通过iDOE自动化通道执行查询类指令），无需通过审批即可下发执行；对于变更类指令，当前已经于iDOE变更管理系统紧急变更流程端到端打通，当用户发出指令后，会自动创建SOP变更方案和变更电子流，等待相关审批责任人通过PC或移动运维终端审批后，即可自动执行。



功能体验

用户通过Welink向小鲁班机器人服务发出口令指令后，小鲁班会对用户进行短信验证码身份认证，经过身份认证后，才可以执行idoe插件自带的指令。



安全控制

iDOE平台不仅本身提供了一些固定场景的ChatOps指令，如KPI查询、变更单查询、事件单查询、告警查询、CMDB信息查询等，还提供了一套强大的ChatOps指令封装框架，支持用户将Shell脚本、Shell指令、以及经过审

批的变更SOP方案封装成ChatOps指令，通过Welink入口执行。为了保证运维的安全性，iDOE平台进行了一系列的安全设计：

- 所有ChatOps指令必须明确标识查询或变更属性，且上线过程必须经过CAB（重大变更评审委员会）集体评审
- 所有ChatOps指令都带有用户白名单控制，限定特定的用户或群组才可以执行
- 所有ChatOps指令执行时必须经过短信验证码身份认证
- 所有ChatOps指令的执行都会有记录任务日志，对于变更类ChatOps指令的执行与现网SOP自动化通道、紧急变更流程端到端打通，保证操作标准化、自动化，变更有审批，过程可追溯

四、业务提升

随着移动运维终端以及ChatOps运维能力的逐步落地，GTS SRE团队对PC和后台堡垒机的依赖正在逐步减轻，各项日常运维工作效率明显提升。

运维能力	应用情况
移动运维终端	• 变更单和方案审批、总览类使用移动端处理量占比分别为20%、5%、10% • 告警及时处理率从94%升到99%， • 工作时段，审批时长从5分钟降至2分钟；非工作时段，事件单响应平均时长从20分钟下降到5分钟，审批时长从30分钟降至16分钟
ChatOps	自上线以来，GDE等7大业务投入使用，开发指令92+条，识别价值场景58+个，使用指令600+次，相比登录后台作业（工作时段节省3-5分钟，非工作时段节省10分钟），典型巡检场景效率提升40%，应急演练场景效率提升60% 通过ChatOps入口执行的作业全部为自动化流程操作（无需登录堡垒机后台执行繁琐人工步骤），与变更管理流程端到端打通，作业方案与作业授权、作业执行完全一致性保证，作业可信风险有效削减

五、案例总结

GTS SRE团队通过引入移动终端运维和ChatOps理念帮助团队有效地降低了对PC和后台堡垒机权限的依赖，随时随地可以接入现网运维，让“喝着咖啡做运维”不再是妄想。但是当前团队在此方面的努力仍然处于非常初级的水平，很多交互体验（如ChatOps对话）与用户的预期仍存在较大差距，还有大量的运维作业场景需要团队去思考如何通过移动终端和ChatOps去落地。随着大预言模型等AI能力在运维领域的应用实践逐步成熟，团队也会不断探索AI能力在GTS云服务运维领域的落地。

构建数智运维保障体系， 解锁业务增长新动力



刘军

华为云aPaaS与应用SRE专家

摘要：D项目通过AppStage一站式智能运维中心构建数智运维保障体系，提升业务增长动力。在云原生架构重构后，D项目面临版本快速迭代和运维工具碎片化的挑战，AppStage通过变更管理、监控、稳定性管理、资源管理、安全和治理等模块，实现高效运维。通过无人值守变更、端到端监控、混沌工程演练和自动化资源管理，D项目实现了运维效率和系统可靠性的显著提升。

一、背景介绍

2022年D项目业务转型升级，以云原生为基础技术底座进行架构重构，2023年新架构上线并逐渐承载新注册用户。D项目新架构部署在华为公有云上，集成了多种云服务组件，业务链路复杂以及碎片化的云上运维工具，给运维保障工作带来了极大的挑战。另外业务上线初期版本快速迭代，业务上线速度和现网稳定性之间存在冲突，迫切需要一套运维平台来完成数智化的转型，帮忙D项目提升运维效率和质量，提升系统的可靠性。

二、案例实践.基于AppStage的一站式智能运维中心管理

AppStage运维中心基于云原生应用运维场景打造，从传统运维的以资源管理为核心升级为以应用管理为核心，沉淀华为内部运维经验，构建多个运维能力模块，形成强大运维生态系统。运维中心提供了一站式智能化运维平台，可将已创建的资源纳管至运维中心进行统一管理，支持纳管VPC、主机、数据库、容器集群。



AppStage一站式智能运维中心



AppStage运维中心基于云原生应用运维场景打造，从传统运维的以资源管理为核心升级为以应用管理为核心，沉淀华为内部运维经验，构建多个运维能力模块，形成强大运维生态系统。运维中心提供了一站式智能化运维平台，可将已创建的资源纳管至运维中心进行统一管理，支持纳管VPC、主机、数据库、容器集群。

资源类型	相关操作
终端节点	删除终端节点：在华为云控制台选择“网络 > VPC终端节点”，进入“终端节点”页面删除，具体操作请参见删除终端节点。
VPC	取消纳管VPC：在弹性网络服务“网络规划 > VPC纳管”页面的VPC列表中，单击VPC所在行“操作”列的“取消纳管”。
OpsAgent	升级OpsAgent：在主机管理服务“云服务器管理 > 弹性云服务器”页面的主机列表中，勾选主机，单击列表上方的“升级OpsAgent”。 卸载OpsAgent：在主机管理服务“云服务器管理 > 未纳管主机”页面的主机列表中，勾选主机，单击列表上方的“卸载OpsAgent”。 卸载OpsAgent前需要先取消纳管主机。
主机	管理主机：具体操作请参见管理主机安全。 取消纳管主机：具体操作请参见取消弹性云服务器主机纳管。
数据库	管理数据库：具体操作请参见在WiseDBA中管理数据库实例。 取消纳管数据库：在数据库治理“实例管理 > 实例列表”页面的数据库实例列表，单击数据库实例所在行“操作”列的“更多 > 取消纳管”。
容器集群	删除终端节点：在华为云控制台选择“网络 > VPC终端节点”，进入“终端节点”页面删除，具体操作请参见删除终端节点。

通过将D项目系统的资源接入AppStage运维中心统一管理，可在一个portal内完成所有运维工作，包括变更、监控、故障、演练、资源管理等，完成数智化转型。

变更管理：高效可靠的自动化变革

D项目业务复杂且版本迭代频繁，月均变更数量高达100+，传统手工变更单个耗时超1小时，耗费大量人力且易出错。针对这一痛点，项目引入AppStage部署服务，采用IaC（基础设施即代码）方式实现无人值守变更及全流程自动化。



自动化评审环节，AppStage部署服务接收变更请求后，依据预设规则和策略，自动对变更内容进行风险评估与合规性审查，从源头降低风险、节省人工成本、减少人为失误。自动化执行方面，借助IaC技术将基础设施配置代码化，变更时按代码定义自动部署更新，采用灰度发布，逐步推送变更并实时监控，发现问题及时回滚，保障业务平稳变更。自动化验证则在变更完成后，利用自动化测试工具和监控系统全面评估服务状态，避免人工验证遗漏，提升效率。

自动化变更过程具备可观察性、灰度特性、影响可控性及可回退性，并融合正向检查与反向防护提升可靠性。正向检查提前识别变更风险，如版本发布前代码静态分析、单元测试，变更前目标环境健康检查等；反向保护则在变更中实时监控性能指标，异常时暂停变更并回滚，变更后持续监控确保系统正常。

监控管理：端到端的全链路洞察

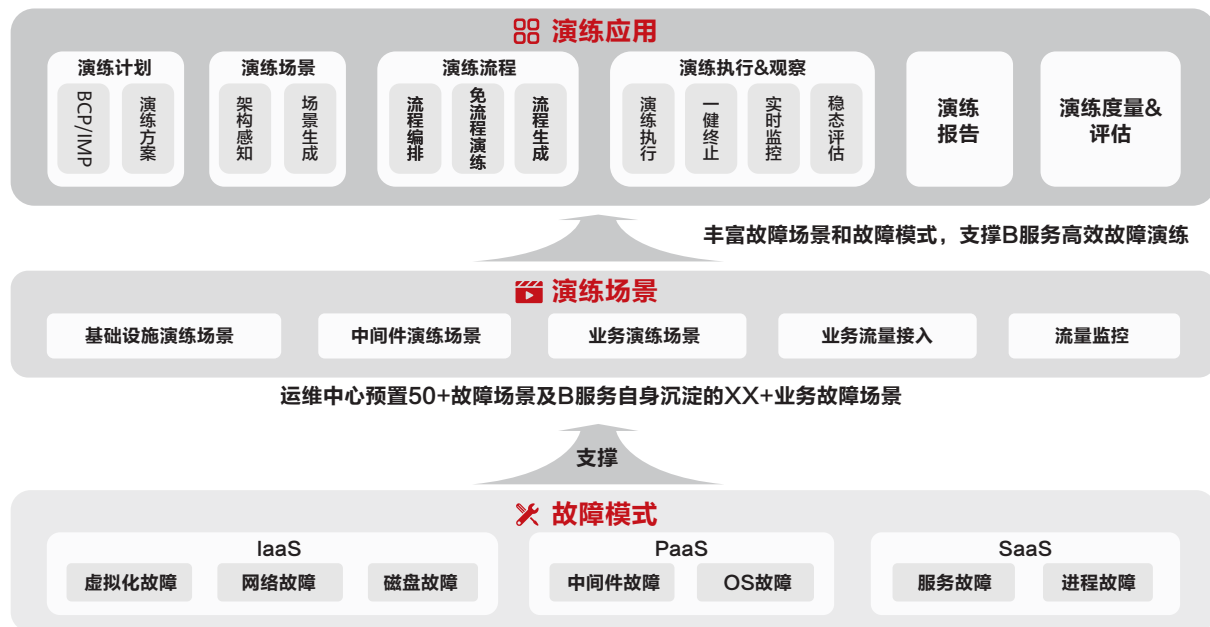
D项目集成超10个第三方组件，业务链路错综复杂，日均生成TB级日志，问题定界定位困难。为此，项目运用AppStage的可观测服务，从产品视角打造端、管、云一体化的综合观测体系，将业务全链路组件纳入可观测范围，构建端到端监控，实现现网问题的快速响应、定界与恢复。



该服务的异常检测功能采用华为基于 PB 级日志数据调优的 AI 异常检测算法，能自动检测异常并告警。借助动态阈值告警能力，AI 算法训练监控数据生成动态阈值，精准识别业务异常波动，自动触发告警，无需手动配置阈值，有效提升监控智能化水平和问题响应速度。

稳定性管理：未雨绸缪的故障演练

D 项目系统架构复杂，现网故障风险高，传统手动故障演练存在诸多弊端，如准备耗时、效率低下、人力投入大、场景覆盖不全，难以满足高频大规模系统故障演练需求。AppStage 的混沌工程演练成为破局关键，提供超 80 种故障注入方法，覆盖 IaaS - PaaS - SaaS 全层，全方位模拟故障，演练过程自动化且可控制“爆炸半径”，实现无人值守故障演练。



项目团队依据业务特点、历史故障及组网特性梳理出标准故障集，对高概率、影响大的必选演练故障场景达成 100% 端到端自动化演练能力。借助一键式故障注入，10 分钟内自动完成一次故障演练，每年演练超 50 次，全面演练可能发生的各类故障，充分验证业务可靠性，为系统稳定运行筑牢防线。

资源管理：统一管控的资源优化

D 项目系统各类资源繁杂，已全部接入 AppStage 实现统一管理，涵盖主机、容器、网络 and 数据库资源。主机管理上，实现权限管控、运维通道管控、运维过程审计及 OS 治理监测，保障主机安全可信；容器管理支持多云集群托管、资源智能调度、弹性伸缩与优化；网络管理方面，可管理 DNS、云网关、VPC、ELB 等资源，通过 IaC 定义业务访问规则，自动化部署网络隔离策略；数据库管理则具备实例管理、自动化 SQL 变更、数据合规查询及风险诊断等功能，形成云原生数据库管理全流程方案。



基于运维 CMDB 服务，围绕资源全生命周期，提供闭环资源配置管理能力，从规划、管理到治理及优化，实现资源的高效管控与利用，助力项目降本增效。

运维安全：全方位的安全保障

D项目运维涉及大量现网操作，存在人员违规、技能不达标、方案错误、流程不完善等风险，威胁系统稳定。AppStage 将安全可信要求深度融入运维流程与平台，确保合适人员在限定时间、以最小权限执行合规操作，且操作记录可追溯。



生产资源 100% 纳管，实现可视可控，包括主机和数据库特权账号密码纳管及定期修改、设置账号密码有效期、漏洞补丁管理等。作业过程 100% 可控可追溯，变更过程严格合规，生产环境操作需 3X 授权，人员权限自动化管理，操作日志安全存储查询并支持审计，数据查看提取合规管控。同时，可靠性部署下，证书、密钥、域名自动 IT 化管理，监控证书到期；版本发布软件仓合规管理，阻断不合规应用入网；安全堡垒机管理，满足可靠性可服务性规范及基线要求，全方位杜绝人为操作风险，保障运维安全。

运维治理：数据驱动的持续改进

借助 AppStage 运维座舱服务，D 项目将运维告警、事件、变更和资源等多维数据统一采集入湖，依托确定性运维经验确定关键指标，开发报表构建运维 BI，持续度量运维效率与质量，实现量化治理。变更报表呈现变更数量、失败数、手工变更数、自动化率及平均耗时；事件报表涵盖事件数、及时响应率、平均恢复时长及根因分类；告警报表显示紧急告警数、无效告警数、有效率及及时恢复率；资源报表则反映主机数、资源利用率、OS 合规率及补丁完成率等指标，为运维决策提供有力依据。

可靠性可运维基线度量						
领域	能力评审点	华为D项目				
		A服务	E服务	C服务	D服务	E服务
可靠性设计 (50%)	业务可用性、容错设计要求	9	9	8	9	8
	满足高可用集成	8	9	8	9	8
	满足双活/容灾能力	8	8	8	8	8
	满足数据可靠性、故障演练	8	8	7	8	7
可运维设计 (40%)	可运维架构设计要求	8	8	7	8	7
	满足中间件接入规范	8	6	7	9	8
	满足自动化部署、配置接入要求	7	7	4	6	5
	满足告警、AIOps接入要求	9	9	9	9	9
运维安全 (10%)	满足开放端口接入要求	2	2	2	2	2
	满足web安全接入要求	3	3	3	3	3
	满足日志安全接入要求	2	2	1	2	1
	满足安全加固接入要求	2	2	1	1	1

此外，基于可靠性/可运维能力基线，开展针对性治理，对服务可靠性、数据库、基础设施和运维安全进行持续自动化度量与优化，助力D项目业务可用性攀升至99.99%，以数字化治理推动运维效能不断提升。

通过以上各环节的紧密协同与深度整合，D项目成功构建起一套高效、智能、可靠的数智运维保障体系，在变革浪潮中稳步前行，为业务持续增长筑牢根基。

三、创新性经验

1. 无人值守变更

基于智能运维AIOps，实现无人值守变更。通过 Everything as a Code (XaC)声明，AppStage运维中心将变更的评审、执行和验证等步骤自动化，避免人因失误，减少变更过程中的步骤，变更操作时间从1小时缩短为10分钟以内，实现无人值守变更，支撑D项目月均100+的变更，极大提升了运维效率。

2. 故障生命周期管理

通过端、管、边、云的联动监控，实现故障生命周期智能化管理。AppStage运维中心通过端、管、边、云的联动监控，让指标、日志、调用链全栈可观测，打通了数据孤岛。在入口APIG接口成功率告警后，通过AI异常检测算法及AI根因诊断等能力，实现1分钟发现、5分钟定位、10分钟恢复，大大降低了故障恢复时间，提升了业务质量。

3. 无人值守的故障演练

通过混沌工程故障注入，充分验证应用可靠性。AppStage运维中心支持80多种故障注入方式，预定义了50多种故障演练场景，通过模拟各种故障、全自动化演练，主动给应用“打疫苗”，使业务对故障具备免疫能力，提升了业务的稳定性。

四、应用成效

通过接入AppStage一站式运维中心后，D项目在变更、监控、演练、安全和可靠性的成就如下：

- **变更管理：**在版本频繁迭代、业务持续高峰的情况下，通过部署服务月均变更100+，变更全流程自动化，变更效率提升90%（1小时->10分钟），未发生变更导致的现网故障；
- **监控管理：**通过将业务全链路关键组件接入可观测服务，构建全链路端到端监控，根据AIOps异常检测实现异常及时感知，告警有效率超过90%，现网故障告警覆盖率100%；
- **稳定性管理：**通过一键式单点故障注入能力，10分钟自动化完成一次故障演练，单次故障演练时间下降70%，演练场景100%覆盖历史发生概率高、影响大的故障；
- **运维安全：**生产环境资源100%纳管，运维作业过程100%可控可追溯，资源自动化管理杜绝手工操作，未发生人因导致的现网故障；

可靠性/可运维治理：识别应用集群负载不均衡、双云配置不规范、ELB健康检查未配置、主机漏洞等多个风险项，通过运营报表持续推动改进。

五、总结与提升

在实践过程中，D项目积累了诸多宝贵的创新性经验，这些经验直接推动了项目的应用成效。例如，无人值守变更的实现，极大地提升了变更管理的效率和可靠性，而故障生命周期管理和无人值守的故障演练则显著增强了系统的稳定性和安全性。

然而，在取得显著成效的同时，也引发了一些反思。D项目存量用户割接过程中的请求量可能比平常高出5-10倍，每一次都相当于是对系统的全链路压测，经常会出现

性能和容量瓶颈。这说明在资源规划方面还有提升空间，需要利用AppStage的AI能力和运维座舱服务，结合活跃用户数、请求量、基础资源使用率等多维度综合规划未来所需资源，避免频繁的扩容。

此外，虽然通过运维治理识别了系统可靠性 / 可运维的问题，并持续推动产品改进，但这并不能从根本上解决问题。为实现应用可用性的长治久安，需要将可靠性/可运维要求落入开发设计阶段，从源头上规避风险。结合开发阶段正向设计和运维阶段逆向治理，才能有效提升系统的可靠性和运维效率，为D项目的长期稳定发展提供有力保障





ECDN助企业内网降本增效、数据“秒速”畅通



程凌
华为云媒体SRE部长



湛宁坤
边缘云专业服务解决方案专家

摘要：在数字化浪潮中，企业需要构建强大的基础设施底座，满足内网海量用户即时、畅通的网络访问需求，以在激烈的市场竞争中，持续释放优质生产力。本文将介绍如何通过专业服务构建以企业内容分发网络（以下称ECDN）为核心的基础设施关键服务能力，支撑企业高质量发展。

一、引言

在数字化时代，企业对内容的分发和传输需求呈爆炸式增长。无论是在线培训视频观看、软件下载更新、门户访问还是文件访问传输，对网络访问的效率、稳定都有着较高的要求。

然而，传统的内容分发加速网络（CDN）在应对海量内容分发时，常常面临规划设计与实施挑战、带宽瓶颈、传输延迟、带宽成本高、运维服务体系缺失等问题，严重影响用户体验，制约企业运营效率提升。

企业内容分发网络（ECDN）是针对企业内部网络环境和业务需求设计的专业解决方案，能够更好地满足企业对专业设计与交付、内容安全、成本控制、体验提升及高质量运维等方面的要求。

二、业务挑战

以XX金融企业为例，该企业拥有1000+家分支机构，各分支机构通过专线跨地域访问总部。但现有的网络方案给企业带来了业务不畅和成本高企的双重压力，具体表现如下：

（一）延迟与稳定性难题

- **跨域业务痛点：**分支机构通过专线访问总部，传输受物理距离、网络拥塞影响，首包时延高，下载速率低。

- **高并发与流量突增冲击：**在线视频培训、企业内网大文件下载、办公APP下载更新等场景下，瞬时流量可达日常的X倍以上。

- **自建系统局限：**企业采用单机开源缓存时，多节点协同困难，容灾能力弱，故障恢复时间长达数小时。

（二）成本与运维效率失衡

- **成本高企：**缓存命中率不足50%，重复回源加剧带宽消耗，企业年专线费用超千万元；静态带宽扩容成本高昂，且无法应对突发流量。

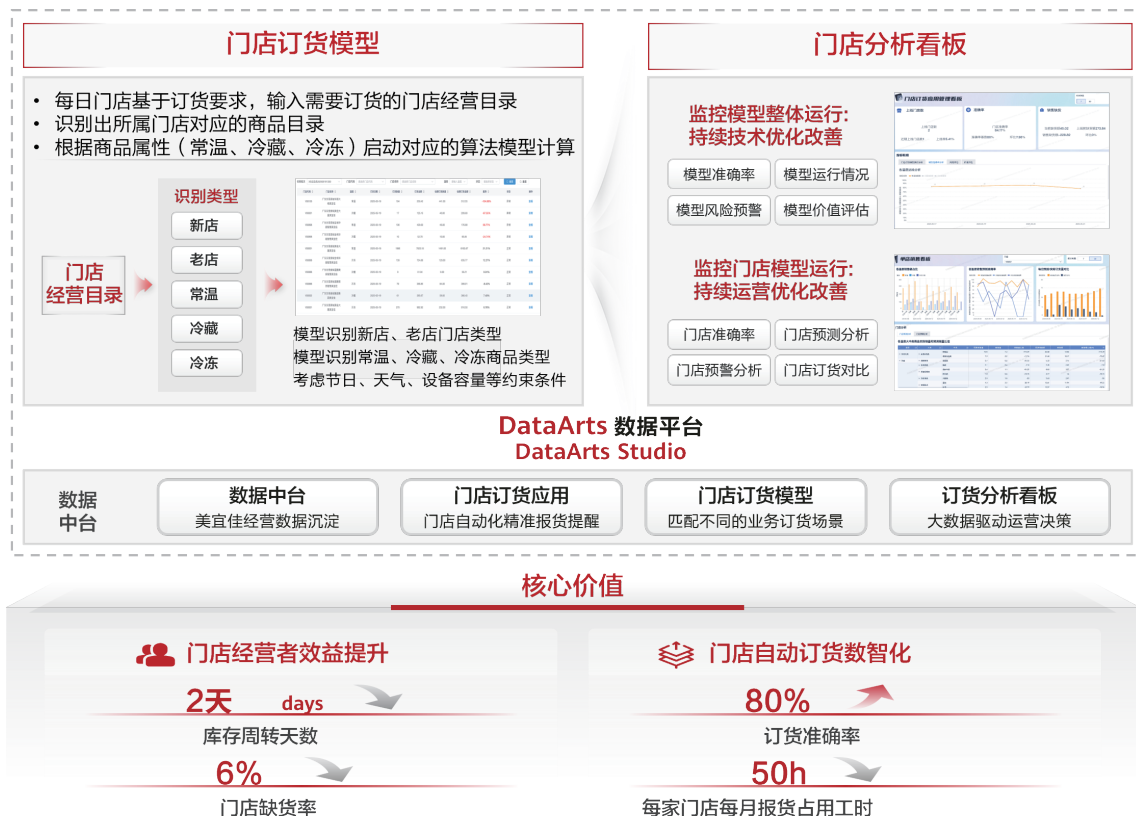
- **人力依赖：**80%的故障排查依赖人工经验，平均修复时间（MTTR）超过2小时，误操作风险高。突流情况下，人工调度响应速度慢，易错失黄金处理时间。

三、方案实践

（一）ECDN方案，架构创新破局业务挑战

区别于企业传统的采购专线网络支撑业务的模式，ECDN（Enterprise Content Delivery Network）则将**加速节点部署在局域网内**，通过一站式配置、统一管控、统一调度和缓存代理，有效优化内容传输，为网内用户提供更安全、更高效的访问体验。

场景三：门店订货—预测引擎破解库存难题



门店订货模型

为优化内部管理、升级经营模式并提升门店经营水平，美宜佳以AI算法与大模型为技术核心，结合日预计需求量、机会损失、滞销折损、促销活动及安全库存等关键因素，为门店提供商品订货建议量，既能减少人工操作时间，又能实现日销售额提升与商品报废率降低。

具体实践围绕“**技术支撑 - 场景落地 - 监控保障**”构建体系：

1.技术支撑—构建门店订货模型：整合销售预测、库存修正、节假日调优、相似店匹配、异常销售剔除、促销活动适配六大算法，形成智能订货能力。每日门店按需求输入经营目录后，系统先识别门店类型（新店 / 老店）及对应商品目录，再根据商品属性（常温 / 冷藏 / 冷冻），结合节日、天气、设备容量等约束条件启动专属算法，输出最优订购建议，最终实现库存与客流量双提升；

2.场景落地—搭建门店分析看板：从库存预警、缺货预警、模型执行、门店上线、提升效果、经营风险预警六大维度，全方位支撑门店数据运营分析。

3.监控保障—建立双重监控体系：一方面监控模型整体运行（含准确率、运行状态、风险预警、价值评估），另一方面聚焦门店模型运行（含门店准确率、预测分析、预警分析、订货对比）。

美宜佳以数据中台为底座，门店订货模型与分析看板深度协同，借助AI实时聚合多维度数据并精准分析，实现模型执行可跟踪、订货异常可溯源，实时掌握门店模型动态，推动门店运营从“经验复盘”升级为“数据沙盘”。同时，系统将态势感知能力赋予店长，支持一屏触达多维数据、明晰因果链条并匹配精准行动，让门店运营从“人为驱动”跃迁至“数字制导”，核心成效显著。

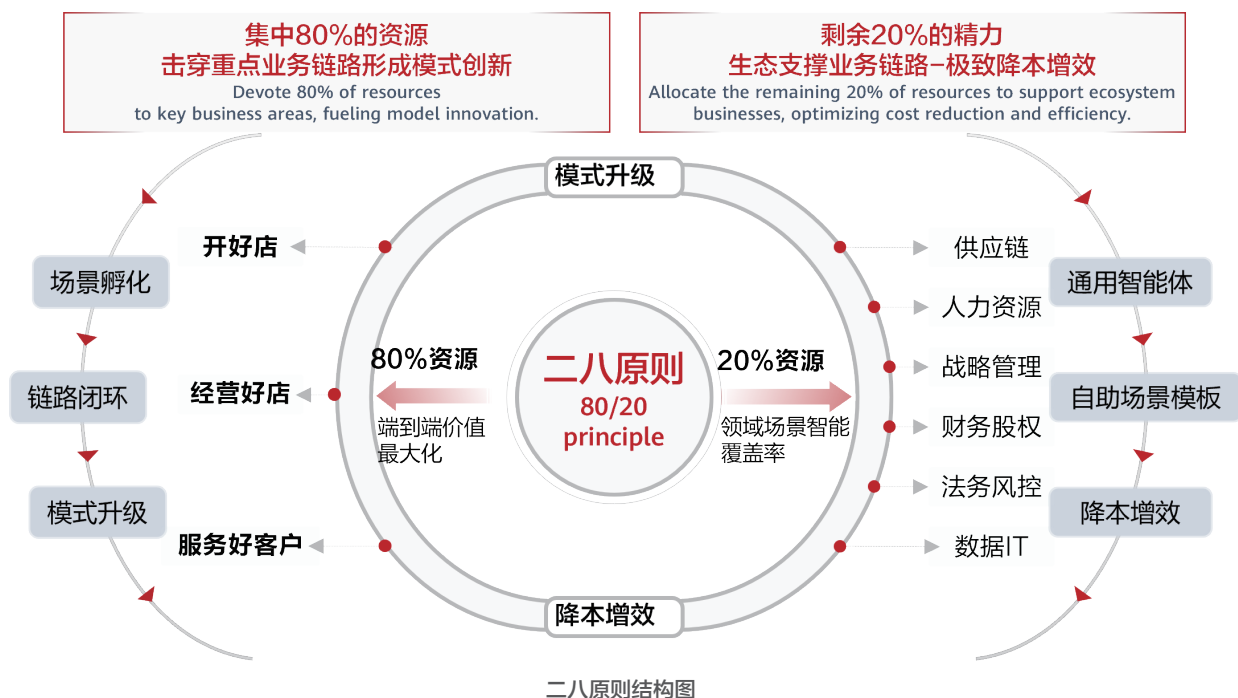
门店经营者效益提升：库存周转天数减少2天，缺货率下降6%；AI围绕经营基线稳定运行并持续优化，在鲜食版块实现商品损耗的精准控制；

门店订货数智化升级：订货准确率提升80%，单店每月报货工时减少50小时，订货效能与人工成本双优化。

三、经验分享

在AI场景实践探索中，美宜佳聚焦业务高价值场景推进均衡发展，通过科学分配资源投入实现效能最大化：一方面，集中80%资源击穿重点业务链路，以深度投入形成模式

创新突破；另一方面，将20%精力投向生态支撑业务链，夯实业务基础。两者相互协同，最终构建起“模式创新驱动增长、生态支撑助力降本”的良性循环。



AI场景的成功落地，有效驱动美宜佳整体组织与业务效能的跃升，核心实践经验可总结为五点：

1.科学规划AI投入：以二八原则聚焦生态提效关键领域，借力行业资源加速技术落地；同时深度剖析业务流程，精准识别核心挑战，据此合理分配模型、算力等资源；

2.夯实数据资产根基：构建完善的数据治理体系，保障数据的准确性、完整性与一致性；通过搭建联合数仓融合内外部数据，并严格守护数据隐私安全，为AI应用筑牢基础；

3.深耕垂直领域模型：针对零售业务全链路，明确聚焦智慧门店、数字员工、门店订货三大核心场景，开发专属垂直领域模型，提升AI应用精准度，实现业务决策与预测的高效赋能；

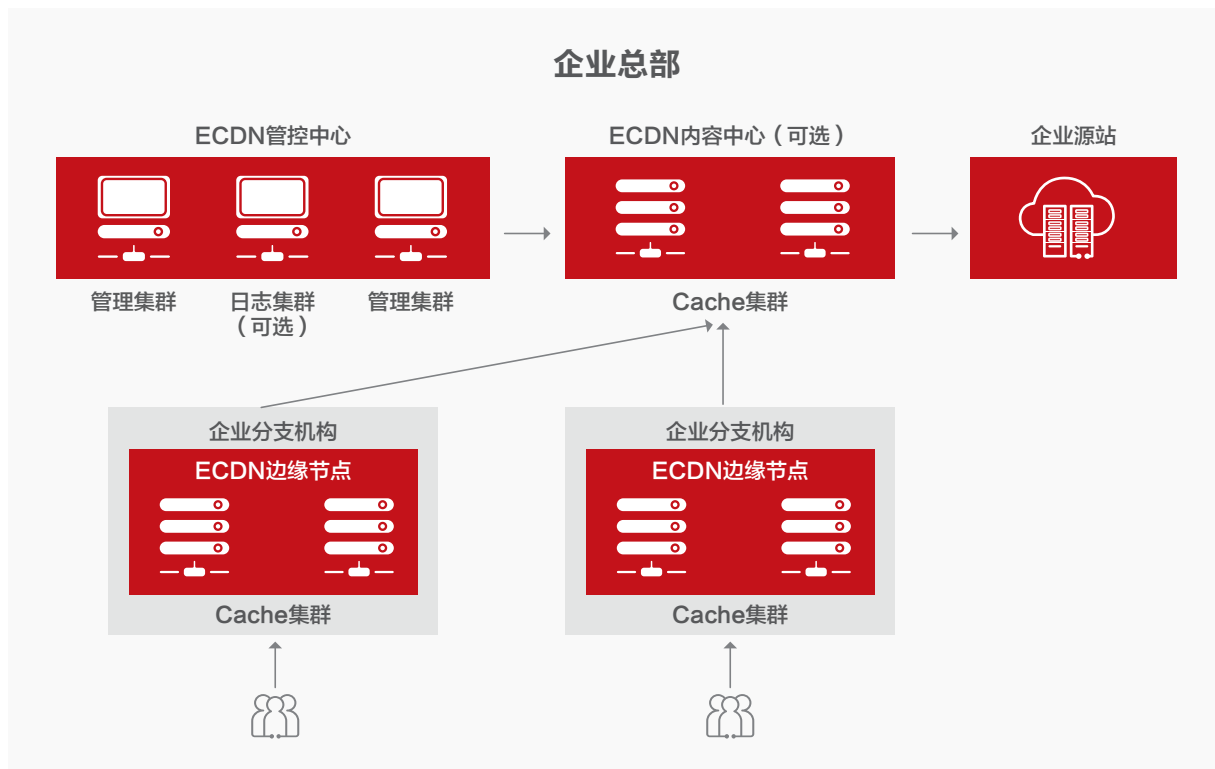
4.拓宽AI应用边界：以运营需求为导向确定投入方向，主动探索AI在创新场景中的应用价值，不设限AI能力范围，大胆将技术落地到具体业务场景中验证与迭代；

5.重视知识传承复用：通过多元方式推动知识共享与传承，尤其注重组织协同构建优秀店长经验的沉淀与复刻体系，实现多类目业务知识的系统化积累。

四、总结

美宜佳在AI领域的探索与实践表明，AI的核心价值在于成为“面向业务场景的业务能力”。它虽起源于工具，却不限于工具，更可以化身优秀的数字员工，承载的是无数优秀业务员工在实践中积累的知识经验，能让更多人具备同等专业能力，最终构建起“人+数字员工”的协同工作模式，驱动组织能力整体跃升与业务效能持续优化。

AI探索是一场持续渐进的长期实践。未来，美宜佳将在现有AI实践基础上，进一步深化AI核心竞争力构建，推动AI能力向集团多元化业务版块渗透，全方位打造企业与AI深度协同的全新发展格局。



方案架构

1. 应用场景

- **网站加速**：适用于有静态内容加速需求的内部网站、论坛、OA 系统、博客等。通过将页面资源更新缓存到近端节点，减轻源站访问并发压力，同时对网页大图进行压缩优化，实现秒级加载，缩短页面响应时间。
- **下载加速**：适用于通过 http/https 提供下载业务的网站、下载工具等，以及软件实时更新业务。将页面资源缓存到近端节点，提升下载速度，减少源站压力。
- **点播加速**：针对政企、金融等多分支机构内部培训视频等场景，将体积大的视频文件预加载到加速节点，供用户快速下载。以自主可控底层协议栈算法，保证播放流畅不卡顿及安全可靠。

2. 关键特性

- **安全加速**：控制信令和业务数据在企业内网传输，全链路采用 HTTPS 确保链路安全，防止数据在传输过程中被窃取或篡改。
- **多级容灾**：基于全局调度能力，实现跨设备、跨集群、直接回源的多级容灾，保障服务的连续性和稳定性，即使部分节点出现故障，也能确保用户正常访问。

- **访问控制**：URL 鉴权、IP 黑白名单和远端鉴权，严格控制用户访问权限，只有授权用户才能访问相应内容，提高系统安全性。
- **灵活可扩展**：对接企业权限管理、告警通知系统，根据企业现有系统进行定制化扩展，提升系统的适配性。

（二）“专优智”服务，高质量支撑方案落地

针对企业面临的具体挑战，SRE 专家以系统化解决问题的思路作为指导，通过“专优智”服务体系，面向企业提供涵盖从建设到运维的全场景专业服务，以体系化作战方式帮助企业实现内容分发网络（ECDN）的高标准规划、高效率建设、高质量服务。

1、“专”：专家团队提供端到端服务

SRE 专家分析企业现有问题后，拉起研发专家、产品专家、网络专家、交付专家、运维专家等组成专项工作组，通过专业分工、协同作战方式，为企业提供问题诊断、需求分析、节点及管控平台部署方案规划、交付实施、运维保障等端到端的服务，保证用户的体验，减少投入成本。



- **产品专家**：流量分析、优化测算，明确接入系统及应用清单，定量分析带宽压力下降比例和节省资费。
- **研发专家**：提供业务及技术咨询，方案规划设计及运维阶段的产品底层技术支持。
- **网络专家**：组网分析、规划设计网络架构方案，支撑网络运维及优化。
- **交付专家**：制定部署实施方案，系统软件部署、域名配置等。
- **运维专家**：业务接入，上线后的日常运维、重大事件保障、故障处理等全局运维服务。

整个服务过程都建立在标准化服务流程基础上，高效完成规划建设及稳健运行。

- **信息调研**：评估网络架构、业务并发量等，识别核心痛点，输出调研报告。
- **规划设计**：规划设计符合业务需求的配置方案、网络方案、部署方案。
- **部署实施**：基于项目建设规划，制定具体的实施方案（软装、配置等）。
- **验收上线**：按计划执行测试验收用例，符合验收条件后安排系统上线，承接业务。
- **运维管理**：制定系统化的运维保障方案（日常运维、应急预案、重保服务，运维智能化）。

2、“优”：“标准SOP护航方案优质高效落地



3、“智”：“智能运维体系确保业务稳健可靠

建设后，服务团队帮助客户建立起完善的智能监控与预警系统，实时监控 ECDN 系统的运行状态。一旦发现异常情况，及时发出预警信息，运维人员能够迅速响应并处理。

同时，定期对系统进行性能评估和优化，根据业务发展和用户需求的变化，不断调整节点布局和缓存策略。

以“专家支持-智能监控-智能自愈”为核心，构建三层“专+智”运维架构：

三级响应专家体系（L1-L3）：

针对客户缺少运维专家的情况，SRE专家通过搭建三层专家服务体系，建立“1+1”（远程服务+驻场服务）本远协同运维机制，帮助客户打造分层、分级运维保障服务能力。

- L1：在线服务工程师/技术服务经理，负责日常运维。
- L2：运维专家，负责突流业务等专项保障，重大问题处理。
- L3：产品研发专家，负责系统架构优化。

全场景实时智能监控与预警

- 多维度监控：实时监控节点状态、流量、用户访问情况等，帮助运维人员及时发现和解决问题。
- 智能化告警：提供多维度多指标告警，并将告警信息对接客户系统，当系统出现异常时，能及时自动通知相关人员，以便快速处理。

智能故障诊断与快速自愈

- 结合日志关联分析与网络拓扑追踪，提高异常定位效率，运维人员可根据日志快速找出问题根源，减少故障处理时间。
- 单机故障，可实现秒级切换，确保业务连续可靠。

四、服务价值

针对企业网络业务访问体验差、宽成本高、运维管理困难等突出挑战，SRE专家深入分析原因后，通过搭建“专家团队+标准流程+智能工具”服务体系，为企业量身定制ECDN方案（组网、密度、缓存、交付、运维等），在局域网内构建缓存集群，实现办公和生产系统隔离，保证数据安全。

通过定制化的方案规划设计、需求开发、集成配置及专项运维保障，用户建设的ECDN系统相比原有系统下载速率提升近300%，首包时延降低80%+，行内App更新峰值带宽占用降低70%+，带宽成本降低XXX万元；通过统一ECDN管理中心提供一站式运维，运维效率提高60%，维护成本降低40%+。

五、结语

以“专业规划交付、智能运维管理”体系化专业服务，解决企业内容分发的传统痛点，以“成本-体验-运维”三重升级，重新定义服务价值：

- **成本优化**：平均降低30%到50%的分发成本；
- **体验升级**：用户访问速度提升2-3倍，首包时延降低80%；
- **运维增强**：3级服务体系、智能运维工具，使得运维效率平均提升40%。

当企业数字化转型进入深水区，ECDN正从技术工具进化为战略级基础设施。Gartner预测，到2025年采用智能运维体系的企业将减少43%IT支出（数据来源：Gartner 2024技术成熟度报告）。

通过将智能运维、数据安全等技术深度融入内容分发网络，企业不仅能实现网络效能的量级突破，更将构建起支撑业务创新的数字化基座。这场由ECDN驱动企业内网效能提升的革命，正在重新定义企业内网服务的价值边界。





识变求稳，用定制化 检查夯实系统运维根基



薛黎

华为云深圳云业务部服务主管



窦顺阳

华为云大客户支持与保障深圳负责人

摘要：某企业在推进大数据服务上云及容器化改造时，面临系统监控复杂、故障排查困难和追求极致性能等挑战。企业通过引入华为云确定性运维服务，以定制化健康检查和架构梳理两大策略，成功识别并处理多项业务风险，实现云上系统安全、稳定、高效运行。最终，企业显著提升系统稳定性、可靠性和资源利用率，为类似企业运维实践提供参考。

一、背景

在数字化浪潮下，大数据、容器化等新兴技术正重塑游戏行业的竞争格局。某知名游戏公司，专注于手游业务，开发并运营多款爆款游戏，业务覆盖沙盒平台运营、IP文创、编程培训等多个领域，是国内头部游戏企业之一。为提升业务开发与运营效率、深度挖掘海量数据的增值价值、创新商业拓展模式，该公司对其云上核心业务进行了重大变革，涵盖大数据服务上云、容器化改造等关键内容。

- **资源容量能效低：**官网接入层网络配置未优先采用弹性伸缩方案，导致资源浪费，无法灵活应对流量波动。

单位实例风险系数高，平均10个实例积聚1.8个中高危风险

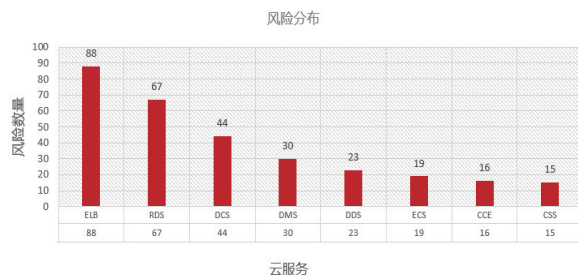


二、业务现状

1. 业务快速变化，亚健康问题频发，风险实例众多

- **新增节点挑战：**以XX精英上海Region为例，新增CCE节点后，节点池配额不足，同时存在不可调度及工作负载实例为单POD等问题，影响业务高可用。
- **配置修改问题：**XX世界底层ECS实例更换后，出现实例反亲和风险，导致资源分配不合理，影响业务性能。
- **数据安全隐患：**IAM账号未开启操作保护，IMS服务存在私有镜像，AntiDDoS流量清洗告警配置缺项，威胁业务数据安全。

如图所示：全量资源实例风险系数高，容易引发业务中断、性能下降、用户体验降低等问题。



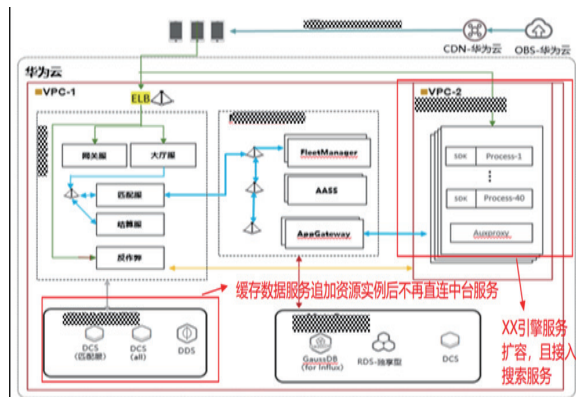
如图所示：关键核心资源累积风险高，风险项重点集中在ELB、DCS、RDS、DDS、ECS、容器和网络，数据和计算的高敏感区域存在失效隐患，价值数据及核心业务易损。

2. 架构革新，新架构可视化

现状：云上业务频繁变动，原有架构难以适应新需求。例如，业务引入大数据服务后，底层数据抓取、缓存及上层调用路径发生变化，系统业务交互及数据流向改变。

挑战：容器化改造后，新增实例和业务关联关系在架构中无法体现，增量业务开发、高效方案沟通及重大问题快速溯源困难。

需求：存量业务架构变动大，需要重新梳理，新增上云业务同已有业务端到端交互逻辑复杂，数据流转路径，上下层及同层调用关系等全量梳理；做到云上架构可视化（涵盖业务、应用、部署等层面），真正实现“一图知天下”。



3. 新服务上线，引入迁移MRR专项检查，识别云上非标配置检查盲点

现状：新服务上云后，非标配置与业务强相关，无法单纯依赖工具和基线配置标准做检查。

挑战：部分服务配置参数为非标配置，导致服务带病作业。例如：存在通用型规格资源争抢风险，影响数据库性能稳定。MCS类型的RDS实例存储空间类型为本地SSD盘，磁盘大小与规格不匹配，无法直接扩缩容，且sync_binlog参数设置为双非1，存在数据丢失风险。

DLINK_REQ_TEC_MSG_ACQ.Topic副本数为1，在节点出现故障的情况下，单副本Topic查询消息时可能会报“内部服务错误”

BDD_SEA_APPLET_TRACK_LOAD_PROD_FS_AGENT_SHADOW_SET.Topic副本数为1，在节点出现故障的情况下，单副本Topic查询消息时可能会报“内部服务错误”

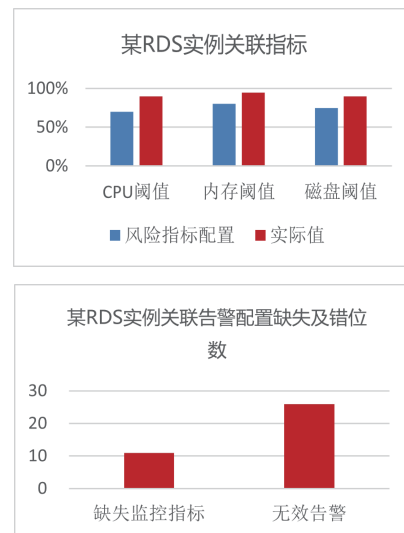
MCS类型的RDS实例存储空间类型是本地SSD盘，磁盘规格和磁盘大小强相关，如果磁盘大小和规格不匹配，无法直接扩缩容

sync_binlog, innodb_flush_logs_at_trx_commit参数设置为双非1，有数据丢失风险

4. 监控预警偏离实际业务需求

现状：监控预警基线与业务实际需求存在偏差，监控缺失或预警错配，导致风险处置错过最佳时间窗。

挑战：业务模块中，RDS for MySQL资源等待行锁指标达到规格上限时，其他资源指标（如CPU、内存）未达瓶颈，但会出现数据库连接超时、响应慢等问题。若未配置该指标，仅监控基础资源使用率，业务受损时无法及时发现并扩容。且配置指标过多导致无效告警频繁，无法有效反映业务真实情况。



三、方案实践

基于企业业务改造的“变”之特性，结合现有业务现状，推出“一变二查三稳”方案策略：

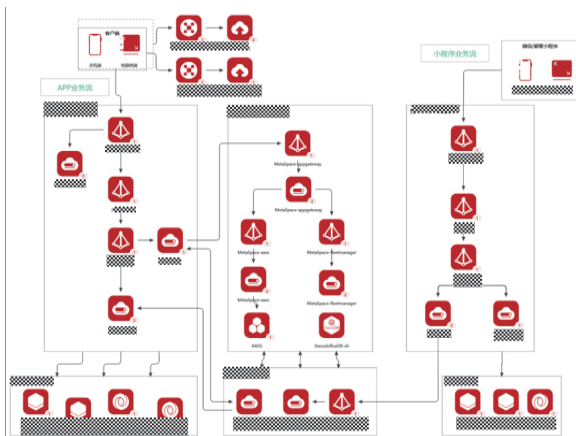
- 1. 架构梳理：**更新新旧架构，解决业务变动后的架构痛点；
- 2. 健康检查：**基于华为云可用性检查五大维度，从性能效率，可靠性，可服务性，容量优化，安全性综合识别云上风险及非标配置；
- 3. 稳定性保障：**通过上云迁移MRR、大数据利剑行动、深海巡航巡检，确保业务变动期间系统的稳定性。



1. “一变”——架构梳理服务，解决业务架构之“变”

企业架构可视化应涵盖云上系统级服务、组件关联关系及数据流转图，既能指导研发人员进行游戏内容扩展和关联业务变更设计，又能在云上业务出现问题时，通过业务关联和数据流向快速定位问题，实现快速决策与处理。

架构梳理聚焦存量业务更新和新增小游戏入口的架构变化，需要梳理游戏登录、充值等接口的交互关系、缓存数据与上层应用商城调用的逻辑关系，明确各模块关联和数据流转路径，为业务稳定运行及后续优化提供清晰蓝图



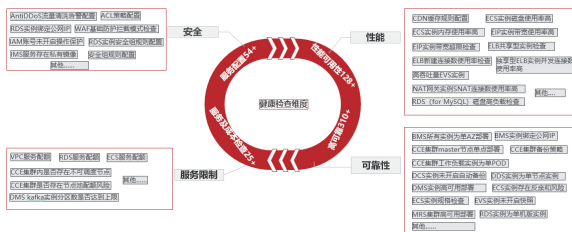
2. “二查”——健康检查服务及监控优化，识别存量业务风险

1) 健康检查服务

根据业务系统现有架构，全面梳理计算、网络、存储、中间件、数据库、安全等云服务的运行参数和指标分析，并结合用云基线最佳实践，从可靠性、安全性、性能效率、可服务性、容量瓶颈等五大维度展开健康检查。

如图所示：安全维度检查涵盖企业敏感账户、安全配置组策略、WAF安全产品防护拦截模式等配置信息；性能维度包含各云服务规则配置、宽带、并发数、磁盘负载、

CPU利用率等信息；可服务性以云服务配额、集群部署方式、节点池配额、消息中间件分区等为侧重点；可靠性偏重云上业务高可用部署方式、是否为敏感业务单节点部署、核心数据备份策略等；



本次健康检查为提高检查效率，引入专家团队及运维基线分析工具，共识别业务风险项300+，涉及业务爆炸半径、恢复时长、问题引入概率及历史问题故障率等关键指标。

如图所示：针对风险项进行优先级分析，综合考虑整改是否对业务有损、有无业务面感知等因素，并输出整改建议。例如，在保障半径方面，从跨AZ、AZ内、集群内、单点等维度进行评估；在恢复时长方面，以无业务影响、1-3min、3-5min、5-10min及10min以上等维度进行划分，综合评估整改优先级。



如图所示：整体评估整改风险情况，联合华为云后端专家，针对企业业务特性输出优化建议，包含风险处理操

作方法、业务是否有损及感知说明等，便于针对现网业务风险进行处理。

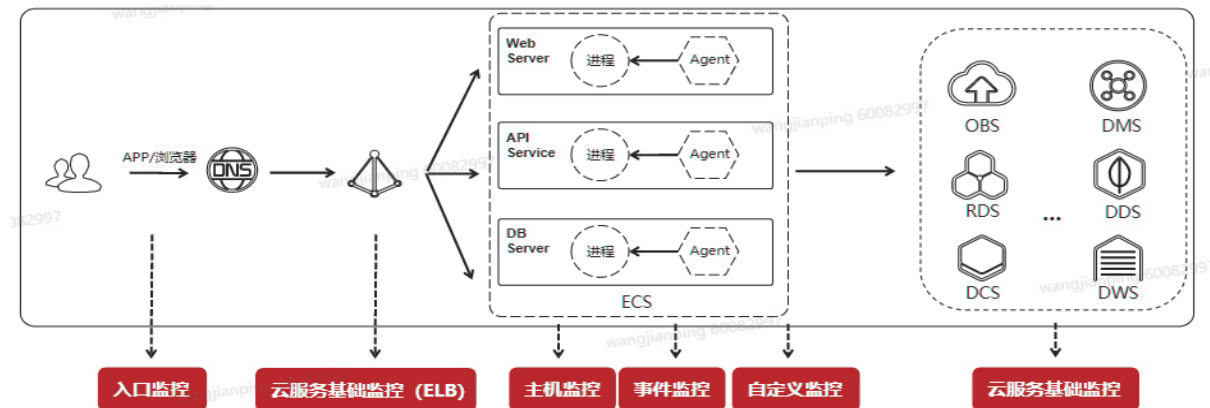
资源名称	资源ID	资源ID	资源ID	资源ID
XXX-XX服务	d058e7bXXX	集群版本过低可能会存在安全漏洞、性能瓶颈或者兼容性问题，造成数据泄露、运行缓慢或崩溃等问题	CCE前端管理界面，点击左侧的集群升级，根据提示升级集群。	集群升级时，以下几点注意事项可能会对您的业务存在影响，请您关注：XXXXXXXX
XXX-XX服务	4bdc83XXX	DDS集群、副本集实例的所有节点在同一个可用区，可能存在AZ级故障无法切换恢复风险	DDS迁移可用区，操作指导：XXXXXX	可用区迁移所需时间和数据量大小有关，并且会伴随60秒以内的闪断，请在业务低峰期执行迁移可用区操作

2) 监控优化提升服务

监控告警优化服务主要针对云服务的告警、资源负载和资源健康情况进行检查。从业务架构、业务场景、历史故障进行分析，结合华为云监控配置成熟度经验，提供监控告警优化建议；

本次监控告警优化，从企业历史用云数据，业务端到端各项指标流量顶峰，低峰，流量波动等多维度展开分析。

如图所示：本次监控优化提升方案从高向下，进行端到端梳理企业监控需求，包含游戏流量入口网络监控、负载均衡指标监控、各项业务应用承接的计算和读写指标监控及基础云服务核心业务指标监控，真正做到云上业务运行态可观测，预警可检测，风险可监控。



1. 三稳”——为增量业务保驾护航

1) 深海巡航服务——现网业务定时定期深度巡检，稳现网业务稳定

作为华为云运维服务重要一环，由华为云支持服务开展的一项巡检服务，主要包括对现网事件单每日分析、重点问题跟踪、关键问题清零等，在本次服务中华为云支持服务主要保障企业大数据上云及容器化改造过程的业务稳定性。

2) 上云迁移MRR服务——高阶服务迁移端到端守护，稳迁移过程可靠

上云迁移MRR是健康检查服务针对上云迁移场景中，各项高阶服务进行全量健康度体检的能力补充。涵盖迁移整个过程、包括迁移前方案设计、转维前交付实施、转维后稳定运营等。

如图所示：上云迁移在整个迁移过程中，从业务实际出发，通过业务架构分析及需求，分别从增量业务的可靠性部署、安全要求、可服务性等维护展开分析、输出优化整改意见。

领域	云服务	风险分类	风险分类	优化建议	风险详解
大数据	MRS	可靠性	中风险	在console页面选择补丁管理，安装最新的补丁；	野补丁代码回溯难度大，容易出现不兼容问题导致业务受损
大数据	MRS	可靠性	中风险	集群创建时，请选择多AZ，将集群的节点分布在不同的AZ中。如果集群创建时没有做AZ可选，请联系华为侧技术支持人员	单AZ故障影响租户业务
大数据	MRS	可靠性	中风险	集群master/slave节点分在同一个虚拟机反亲和组，可在ECS的节点组中查看该集群的存储/计算节点是否在同一个节点组中。Worker节点至少有3个节点分在同一个虚拟机反亲和组。	单台物理机故障影响客户业务
大数据	MRS	可靠性	中风险	告警治理，MRS的告警处理指导如下： https://support.huaweicloud.com/usermanual-mrs/ALM-12001.html ，进入告警页面后搜索告警ID即可	告警存在隐患，影响客户业务

3) 利剑行动服务——增量业务上云预期稳，稳云资源实例可服务性有保障

利剑行动聚焦高阶服务风险识别和消减，事件清淤等，将现网业务中包含高阶大数据服务纳入行动专项名单，组织华为云专家针对企业重点事件进行分析，针对大数据服务运维需求进行方案设计以及现网风险巡检，提升企业新增业务稳定性。

四、全方位提升

通过本次定制化健康检查服务交付，完成XX世界上海、广州等多Region云服务资源深度健康检查，针对XX精英全量云资源实例部署及业务架构进行从0到1的全量梳理，使得企业用云风险消减300+，主要体现在以下方面：

1) 可靠性提升：通过梳理企业多Region及跨AZ业务部署，输出7个云服务共25项架构可优化项，完成优化闭环，进一步提升架构的合理性与稳定性。

2) 性能优化提升：结合业务历史数据和业务增长趋势、业务现状及预期峰值，进行容量及资源分配策略优化，识别出6项容量扩展风险，并完成整改闭环，确保资源合理利用与业务平稳发展。

3) 监控告警提升：针对业务核心资源，全面梳理现有监控告警规则、识别冗余告警和误报问题、优化告警策略、降低监控噪音，完成3个重点资源（1个ELB后端异常主机梳理、2台ECS的CPU资源使用率频繁超阈值）告警的优化和降噪，监控有效性与精准性显著提升。

4)可服务性提升：结合历史业务问题进行问题根源分析、设计解决方案、闭环历史全量问题单。针对业务涉及DDS、ECS、ELB等云服务，识别N+业务问题，实现全量隐患消减。

5)安全提升：梳理云上白名单场景，完成虚拟网络服务3个、CSS服务1个、MRS服务1个、EVS服务1个、数据库服务6个、DCS服务1个、CCE服务1个，提升运维流程的安全性及高效性。

五、关键经验分享

1.通过最佳用云实践对存量业务进行风险分析，针对盲点风险补齐运营短板，真正做到**早识别、早处置，防患于未然**。

2.增量业务需提前分析业务运营痛点，尽早介入，识变求稳，可有力支撑新业务稳定运营，抢占市场先机。

3.架构决定基础，在业务优化过程中，架构的梳理尤为关键。只有梳理业务和部署架构后，才能更深层次地针对业务进行定制化检查，实现精准运维。



六、案例总结

在数字化、新业态的变革浪潮下，云上业务不断求变求进。对于运维人员而言，积极拥抱变化，并在变中求稳，是其核心目标。本文的案例实践表明，定制化的分析和检查是提升业务系统稳定性与资源健康度透明化的关键一步。通过实施定制化健康检查与架构梳理，企业能够有效应对业务变革带来的运维挑战，夯实系统运维根基，为业务的长远发展提供坚实保障。





维享会精选活动回顾

2025年，维享会举行多次沙龙活动，汇聚行业大咖共同探讨AI技术的创新应用，围绕华为云确定性运维实践，进行思想碰撞和实践探索，通过确定性运维体系及解决方案，推动企业云上的业务创新，助力企业运维变革，帮助企业重构“管云”模式，护航千行万业行稳致远。

西班牙巴塞罗那

MWC 2025 华为云峰会及
系列专业服务发布

3月3日-6日（同期3月2日华为云峰会）

重庆

2025 华为云中
企出海全球峰会

2025年3月24日

天津（南开大学）

华为云确定性运维
公开课走进南开大学

2025年4月14日

2025年10月7日

2025年9月19日

2025年8月28日 - 29日

爱尔兰都柏林

华为云亮相SREcon25
揭秘大规模负载均衡架构实践

上海

华为全联接大会2025

哥伦比亚

拉美华为云 SPARK 峰会



“

2025年3月3日至6日，2025年世界移动通信大会（MWC 25）在西班牙巴塞罗那举办。在巴展期间，华为云面向全球发布CloudDC确定性运维、盘古大模型专业服务、AI全场景专业服务等一系列专业服务，分享最佳解决方案和创新实践。

”

— MWC 2025 华为云峰会及系列专业服务发布



2025年世界移动通信大会

“

2025年3月24日，2025华为云中企出海全球峰会在重庆开幕，围绕当前企业出海的新趋势与机遇进行深入探讨。在峰会上，以“智利全国范围的大规模停电保障事件”为案例，分享华为云持续重保27小时，以“基础设施抗压、快速应急响应、全链路持续保障”三大能力，保障全球业务0中断，强调了稳定可靠作为华为云的质量守护。

”

— 2025 华为云中企出海全球峰会



2025 华为云中企出海全球峰会

“

2025年4月14日，华为云确定性运维公开课走进南开大学，围绕“数字化业务稳定性的未来”主题，分享 SRE 理念与确定性运维体系实践经验，课程结合确定性运维体系建设、高可用架构设计和故障快恢等前沿技术，探讨企业如何通过智能运维应对数字化转型中的挑战，展望了未来运维与AI、大数据融合的新趋势，与南开大学学生一起交流 SRE 职业要求以及当前行业现状，为学生未来就业提供宝贵的参考经验。

”

— 华为云确定性运维公开课走进南开大学



华为云确定性运维公开课走进南开大学

“

2025年4月26日，“春生万物 云启新程”游戏运维精英社走进华为活动在上海青浦·华为练秋湖研发中心成功举办。此次活动汇聚了游戏行业众多精英，围绕AI与游戏产业融合、游戏运维创新实践展开深度探讨，为行业发展提供新思路，华为云 SRE 专家受邀现场分享确定性运维理念及实践经验。

”

— “春生万物 云启新程” 游戏运维精英社走进华为



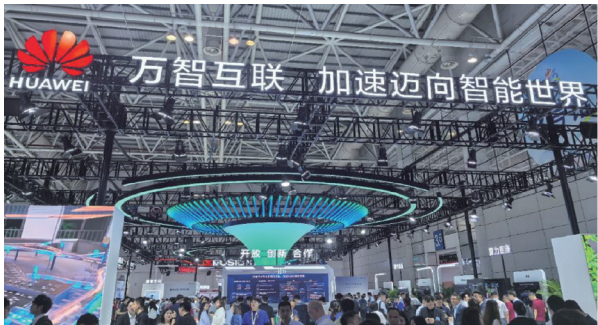
中国上海（华为练秋湖研发中心）

“

2025年5月，第八届数字中国建设峰会在福建福州举行，华为云 AI CITY “1234MNX” 架构首次公开，分享 AI 运维的确定性运维体系，通过故障快恢、风险识别与消除、运维合规等能力，保障城市云上系统安全、稳定、可靠，赋能城市高质量发展。

”

— 第八届数字中国建设峰会



第八届数字中国建设峰会

“

2025年5月16日，华为云维享会·新加坡2025会议召开，经过近两年的发展，维享会已发展成国内最大的运维专家联盟，作为迈向全球化进程的重要里程碑，首站选址亚太数字经济枢纽新加坡，吸引来自全球的40余位行业领袖、技术专家齐聚一堂，围绕AI时代下云服务韧性建设与运维实践等议题展开深度探讨，赋能企业业务稳定运行，加速企业创新，实现云上质量与运营效率的双跃升。

”

— 华为云维享会·新加坡 2025 会议



华为云维享会·新加坡 2025会议

“

在华为中国行 2025·河北沧州新质生产力峰会上，沧州与华为回顾八年合作历程，双方共建的沧州华为云计算中心采用“一云双平台”架构，依托确定性运维体系保障云上系统安全、稳定、可靠，已承载全市 54 个政府部门、160 余项业务系统，赋能民生、政务、产业领域，推动沧州实现从传统治理向智能治理的跨越。

”

— 华为中国行 2025·河北沧州新质生产力峰会



华为中国行 2025·河北 沧州新质生产力峰会

“

2025年6月21日，HDC2025华为云专业服务高峰论坛召开，本次论坛汇聚了近300位专家、客户、合作伙伴及开发者，围绕“全栈 AI 使能服务，释放企业澎湃生产力”主题展开探讨。会上，华为云现场首次发布 WISE 4A能力模型，为企业提供一套AI成熟度评估标准，介绍了华为云CloudMatrix384复杂系统下的稳定性保障能力，邀请美宜佳和中软国际客户分享 AI 应用实践经验。同时现场发布 AI 专业服务解决方案，结合“沃土云创”计划，高效赋能企业与开发者，加速行业AI应用落地。

”

— HDC 2025 华为云专业服务高峰论坛



HDC 2025 华为云 专业服务高峰论坛

“

2025年8月28日 - 29 日，拉美华为云 SPARK 峰会在哥伦比亚召开，本届大会以“以 AI 加速，以云拓展”为主题，华为云拉美首席信息官受邀分享《从混乱到控制：华为云确定性运维之道》，与现场客户伙伴一起交流云上运维创新技术与经验，与会者探讨云技术如何帮助企业降低成本、提高效率和促进业务创新，共赢云上行业智能化未来。

”

一 拉美华为云 SPARK 峰会



拉美华为云 SPARK 峰会

“

在华为全联接大会2025上，广州交易集团有限公司总经济师张石波在“华为云服务分论坛：质赢未来·云智共生”上分享了《IT数字化战略转型：面向用户体验的运维监控一体化实践》，展示了广州交易集团携手华为云服务，依托华为在数字化领域的领先技术实力，围绕“运维支撑转型”为核心逻辑构建的运维管理体系。

”

一 华为全联接大会2025



华为全联接大会2025

“

2025年10月7日，SREcon25在爱尔兰都柏林举行，华为云爱尔兰可靠性实验室首席专家首次受邀在主会场做 Plenary (Keynote) 报告，结合华为云SRE团队构建四层负载均衡服务HyperRouter的经验，基于确定性运维理念，分享如何通过跨服务协同架构设计、用户态高性能架构、服务自愈、Cell化部署和用户为中心的可观测等，实现云服务全生命周期稳定运行。此次分享展现了华为云在SRE领域的前沿实践，获得了现场专家的广泛认可。

”

一 华为云亮相SREcon25 揭秘大规模负载均衡架构实践



华为云亮相SREcon25，揭秘大规模负载均衡架构实践

**依托确定性运维体系
华为云重大事件结果质量达成业界最佳
华为云0重大事故804天**

截止2025年10月15日

华为云全栈AI使能专业服务，助力释放企业澎湃生产力

华为云首次提出的WISE 4A模型，通过建立AI成熟度评估标准，为企业提供明确的AI落地路径，帮助企业实现AI时代竞争力的新跃升。



适配度：指标为MAI（ModelAdaptabilityIndex），从知识能力、专业任务、安全合规维度，评估大模型与行业领域/应用场景适配程度。

亲和度：指标为CAI（ClusterAffinityIndex），从吞吐、时延、精度维度，度量大模型资源需求与智算平台资源供给拟合度。



集成与运营服务

开发支持 Agent开发支持 场景化方案优化

集成与优化服务

数据集 模型增训&微调 模型评测

集成与运维服务

迁移适配 AI平台技术支持 训推保障与优化

能力认证培训服务

HCCDP-AI HCCDA-AI

3套工具集
DataArts, ModelArts, Versatile

AI成熟度指标体系



MAI

能力 | 任务 | 安全



CAI

吞吐 | 时延 | 精度



SFU

单卡有效算力 | 线性加速比 | 资源池可用度



ATD

通过AI认证人数占比

可用度：指标为SFU（System Flops Utilization），从单卡有效算力、线性加速比、资源池可用情况，度量相关能力。

AI人才浓度：指标为ATD（AI-Talent Density），以“通过AI认证人数占比”为依据，度量AI人才密集程度。

安全可靠 | 智能运营
资源高效 | 业务敏捷

华为云



确定性运维官网



维享会官网