阿里云

数据库
稳定性
白皮书

目录

前言

数据库稳定性是指数据库系统在长期运行过程中，能够持续保持稳定的状态，并保证数据的完整性、一致性、可靠性和可用性。

阿里云认为数据库的稳定性是数据类业务的生命线，数据库给客户带来的核心价值是确保数据业务永远在线，从而让数据价值不断放大。阿里云瑶池数据库通过自动弹性伸缩、资源实时监控和诊断、自动故障修复、数据备份和恢复、高可用容灾、数据库自治等多维度技术手段，提供稳如磐石的数据库使用体验。

数据库稳定性概述

阿里云瑶池数据库简介

阿里云拥有国内强大且丰富的云数据库产品家族——瑶池数据库。瑶池数据库涵盖关系型数据库、非关系型数据库、数据仓库、数据库生态工具四大板块，含 PolarDB、RDS、AnalyticDB、Lindorm、MongoDB、DMS 等产品家族，可以为企业数据生产和集成、实时处理、分析与发现、开发与管理提供全链路生命周期的服务。

数据库给客户带来的核心价值是确保数据业务永远在线，让数据价值不断放大。基于此，阿里云瑶池数据库将打造云原生一站式数据管理与服务作为战略，不断创造客户价值。

阿里云瑶池数据库是国内唯一连续 3 年蝉联 Gartner《全球云数据库管理系统魔力象限报告》“**领导者(LEADERS)**”的服务商，在 IDC 发布的《中国关系型数据库软件市场跟踪报告》中，连续 3 年蝉联中**国关系型数据库（公有云模式）**榜首。

数据库稳定性定义

数据库稳定性是指数据库系统在长期运行过程中，能够持续保持稳定的状态，并保证数据的完整性、一致性、可靠性和可用性。因此，本文主要从如下两个方面阐述阿里云瑶池数据库的稳定性情况：

服务的可用性：
在特定时间内数据库可提供服务的概率。

数据的可靠性：
指数据的完整性、一致性和准确性。

接下来，本文将详细阐述从服务的可用性以及数据的可靠性两个衡量角度下，阿里云瑶池数据库的稳定性处于什么样的水位，以及通过哪些核心技术来保障数据库的稳定性。

数据库可用性度量

服务可用性指标：SLA

数据库业内通常使用服务可用性 SLA 指标来度量数据库服务的可用性，该指标描述的是数据库实例处于正常状态提供数据库服务的时长占实例生命周期的比例。

以 Amazon Web Services (AWS) 为例，AWS 提供的云数据库服务 Amazon RDS 可用性高达 99.95%，而 AWS Aurora 数据库的可用性高达 99.99%。

那 99.95% 和 99.99% 这两个数字如何更直观地理解呢？基于业内对服务可用性 SLA 的定义，我们假定服务周期为一年（365 天）来折算服务可用性 SLA，即一年（365 天）中，数据库服务不可用的分钟数。

例如，四个 9（99.99%）的服务可用性 SLA 承诺，表示年服务不可用分钟数不大于 (1-99.99%)×60×24×365=52.56 分钟，99.95% 的 SLA 则表示年服务不可用分钟数不大于 (1-99.95%)×60×24×365=262.8 分钟。

阿里云云数据库也采用该业内通用的服务可用性 SLA 指标来度量各云数据库产品的可用性。基于可用性指标，阿里云数据库提供了明确的可用性 SLA(Service Level Agreement,服务等级协议),规定了服务质量和赔偿条款。

阿里云数据库服务可用性将根据服务周期，以单个实例为维度，按照如下方式计算：

服务可用性 =(（服务周期总分钟数 - 服务不可用分钟数）/ 服务周期总分钟数)x100%

其中:

服务周期
一个服务周期为一个自然月，如客户使用云数据库实例不满一个月则以当月该云数据库实例累计使用时间作为一个服务周期。

服务周期总分钟数
按照服务周期内的总天数╳24（小时）╳60（分钟）计算。

服务不可用分钟数
当某一分钟内，客户所有试图与指定的云数据库实例建立连接的连续尝试均失败，则视为该分钟内该云数据库实例服务不可用。 在一个服务周期内云数据库实例不可用分钟数之和即服务不可用分钟数。

本章节以云数据库 RDS MySQL 和云原生数据库 PolarDB MySQL 为例介绍阿里云瑶池数据库在服务可用性方面的承诺。

产品				服务可用性SLA承诺
云数据库RDS MySQL	三节点企业版三可用区/高可用版独享型			不低于99.99%
	三节点企业版非三可用区/高可用版通用型			不低于99.95%
云原生数据库PolarDB MySQL	单节点版	开启存储热备集群	PSL5存储	不低于99.95%
			PSL4存储	不低于99.90%
		关闭存储热备集群	PSL5存储	不低于99.90%
			PSL4存储	不低于99.80%
	多节点集群版	存储热备集群	PSL5存储	不低于99.99%
			PSL4存储	不低于99.96%
		关闭存储热备集群	PSL5存储	不低于99.96%
			PSL4存储	不低于99.95%

从上述 SLA 表格可以看出，阿里云云数据库 RDS MySQL 和云原生数据库 PolarDB 均可达到 99.99% 的服务可用性 SLA。

灾难恢复指标: RTO 和 RPO

灾难恢复 (Disaster Recovery, DR) 又称为容灾，指通过一组策略具和程序以使基础设施或系统在自然或人为灾难后恢复或继续运行。基于 ISO22301，容灾中有两个重要的概念:

RTO（Recovery Time Objective，目标恢复时间）
灾难发生后，从业务发生中断开始到业务恢复到目标等级所需要的时间。

RPO（Recovery Point Objective，目标恢复点）
灾难发生后，可以接受的最大数据丢失量，也以时间来度量，即丢失多长时间的数据。RPO 是反映数据丢失量的指标，体现了企业能容忍的最大数据丢失量的指标。RPO 值越小，代表企业数据丢失越少，企业损失越小。

那么阿里云数据库在 RTO 和 RPO 指标上的表现如何呢？以云原生数据库 PolarDB 为例，下表展示了在不同停机场景下 PolarDB 的 RTO 和 RPO。

停机类型		数据库高可用保障措施	RTO	RPO
计划外停机	主节点损坏	• 多节点部署 • 多可用区部署 • 热备节点 • 存储热备 • 全球数据库GDN •	秒级，5~10秒	0秒
	存储集群故障		秒级	0秒
	AZ故障		秒级	0秒 （集群需开启跨AZ强一致）
	系统写坏		秒级	0秒
	Region故障		分钟级	< 2秒
计划内停机	主备切换		秒级	0秒
	小版本升级/打补丁		秒级	0秒
	大版本升级		秒级	0秒
	主机维护		秒级	0秒
	存储迁移		秒级	0秒

从上表可以看到，PolarDB 在 Region 内的物理故障和逻辑故障的情况下，可以实现数据 0 丢失，秒级恢复业务；在 Region 整体故障情况下，支持分钟级快速拉起业务，同时丢失数据小于 2 秒。

在了解了阿里云瑶池数据库在这几个稳定性指标下的目标和承诺后，下一章将详细解读为了达成这样的目标和实现这样的承诺，阿里云瑶池数据库构筑了哪些核心技术。

阿里云瑶池数据库通用稳定性能力

随着数据量的不断增加，数据库已成为企业信息化建设的重要组成部分。而数据库的稳定性直接影响到企业的业务运行和数据安全。因此，阿里云非常重视数据库的稳定性建设，全力保证客户业务系统的稳定运行，避免数据丢失和服务中断的情况发生。

阿里云瑶池数据库的稳定性能力主要包括数据库系统的高可靠性、高容错性、自我恢复能力、数据保护能力、性能优化能力、故障排除能力等。一个稳定的数据库系统应该能够在任何情况下确保数据的安全、一致和可靠，同时满足用户对数据的需求，保证高效的数据处理和快速的数据查询。此外，数据库的稳定性也需要考虑系统的可扩展性和灵活性，以适应业务负载的变化。

下文详细介绍了阿里云瑶池数据库通用的稳定性能力，这部分能力大多是由数据库生态工具提供的。除此之外，阿里云瑶池数据库产品又有各自独特的稳定性功能。本文最后以阿里云云原生数据库 PolarDB 为例，详细介绍了阿里云数据库自身具备的特有的稳定性能力。

资源实时监控和告警

阿里云数据库自治服务 DAS 提供了丰富的数据库性能监控和告警功能，对云数据库进行集中管理、统一监控，节省用户 50% 以上的管理成本，显著降低操作故障概率，并且可以快速发现和定位数据库异常，提升数据库的稳定性。

统一监控

可查看所有环境、所有集群、所有实例的性能趋势情况和实时性能情况。

低成本：
用户无需耗费人力开发和部署采集、计算、存储程序，直接使用 DAS 即可监控数据库。

指标丰富：
支持数据库各项关键指标的采集、计算和展示。

细粒度的监控：
支持用户按需设置细粒度的监控，最小支持秒级监控，帮助用户快速发现异常。

统一告警

支持数据库告警规则的自定义、告警信息的发送。

默认告警模板：
基于阿里巴巴的数据库运维经验，为各种数据库引擎定义了默认的告警模板，用户可以直接使用。

灵活配置：
支持各种告警规则、告警模板、告警联系人、告警联系组的灵活配置，用户可以为企业内不同的使用者定义不同的告警模板。

告警提醒：
DAS 自动检查数据库实例是否配置了告警。如果某数据库实例没有配置告警，DAS 会向用户发送提醒，从而防止业务因缺少告警而受损。

智能诊断和优化（自治服务）

此外，阿里云数据库自治服务 DAS 基于机器学习和专家经验，实现了数据库自感知、自修复、自优化、自运维及自安全的云服务，保障您的数据库持续可用。

DAS 实现了无人工参与的自治场景支持，且数据库自治服务系统自身具备不断构建自学习能力，例如异常的自动标注、案例系统设置、异常模拟、量化反馈评估等，依托线上业务场景的丰富性积累，沉淀大量案例，以案例为驱动，加速自我进化，不断提升自治的有效性。

基于以上理念，DAS 拥有四大核心自治特性：

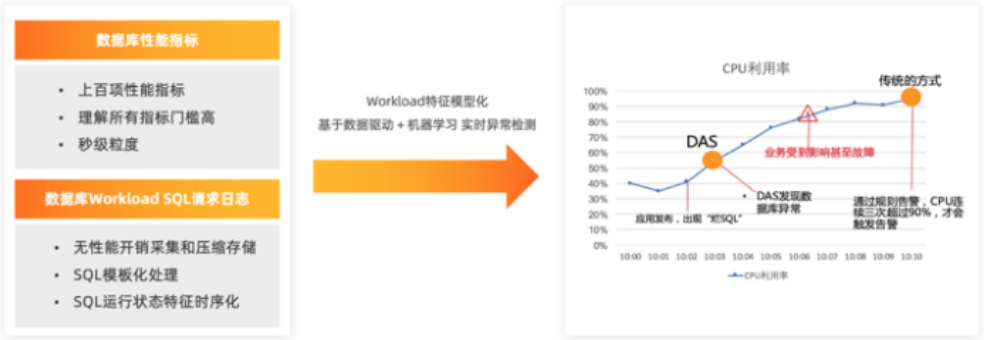
7×24 实时异常检测 **故障自愈** **自动优化** **自动弹性**

下文将从数据库稳定性的角度，以双 11 期间的实际案例作为切入点，介绍 7×24 实时异常检测、故障自愈、自动优化和自动弹性这四大自治特性。

7×24 实时异常检测

DAS 的 7x24 实时异常检测通过机器学习算法，实时对数据库的 Workload 进行异常检测，相比传统基于阈值的告警方式，能够更及时地发现数据库的异常，而不是靠故障驱动。您可以采集各种数据，比如从链路上采集数百个数据库性能指标和从链路上采集已加载 SQL 语句的查询日志，海量数据的离在线处理与存储，基于机器学习和数据库领域预测算法，实现各业务数据库实例的持续模型训练，实时模型预测和实时异常检测分析能够顺利运行。相比传统基于规则和基于阈值的方式，实时异常检测具备以下优势：

- 检测范围更广，例如不仅限监控指标，还包括 SQL、日志、锁等。
- 实现准实时的检测，大大超前传统方式发现异常。
- 基于 AI 和异常驱动的检测技术，而非故障驱动的检测。
- 具备周期性识别能力，自适应业务特征，拥有提前预测能力。



现实中常见的 workload 场景，如毛刺特征、周期性特征、趋势性特征、均值偏移特征等， 异常检测服务都能够准确自动识别，并支持多种时序特征叠加识别，识别出异常后，会触发基于根因的全局诊断分析，以及后续的异常恢复、优化自治场景。

特征	特征定义	常见场景	示例
毛刺（Spike）	短时间内突增突降	• 故障抖动，数据库活跃线程飙升 • 瞬间流量波动，如：秒杀业务场景	
周期（Seasonality）	周期规则性变化	• 业务高峰和低峰 • 工作日和周末时间段波动区别 • 周期性任务	
趋势（Trend）	按某种规则稳步地增长或下降	• 流量、CPU利用率等逐渐增长 • 库表、磁盘空间使用增长	
均值偏 (Meanshift)	某个时间点发生转折变化	• 业务大促 • 磁盘空间进行日志清理	

故障自愈

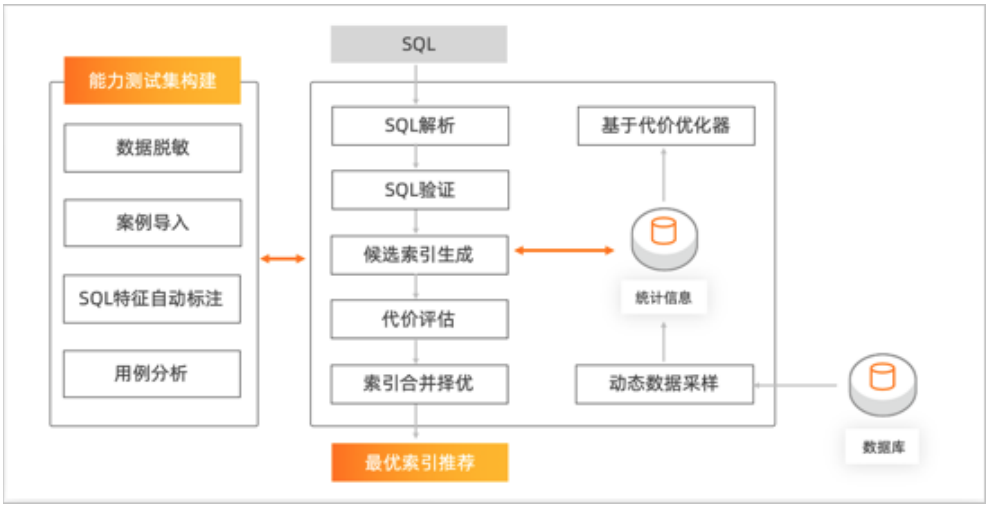
通过 7x24 实时异常检测， 数据库实例异常完成实时检测发现，DAS 自动进行根因分析，自动执行相关止损或修复操作,帮助数据库自动恢复,减少对企业业务的影响。如下图为双 11 期间自动 SQL 限流的实际案例:

某实例于 2020 年 11 月 5 日 12 点 31 分，活跃会话数和 CPU 开始骤然飙升，DAS 异常检测中心于 12 点 33 分确定此次飙升为一次数据库异常而非抖动尖刺，触发 SQL 自动限流根因诊断，12 点 34 分诊断完成，共发现两条导致该次异常的问题 SQL，发现问题 SQL 后随即发起自动限流，活跃会话数开始降低，存量已提交问题 SQL 执行结束后，活跃会话数下降至正常水位，CPU 使用率同时也恢复到正常。整个过程满足“1-5-10”异常自愈能力，即 1 分钟发现，5 分钟定位，10 分钟恢复。

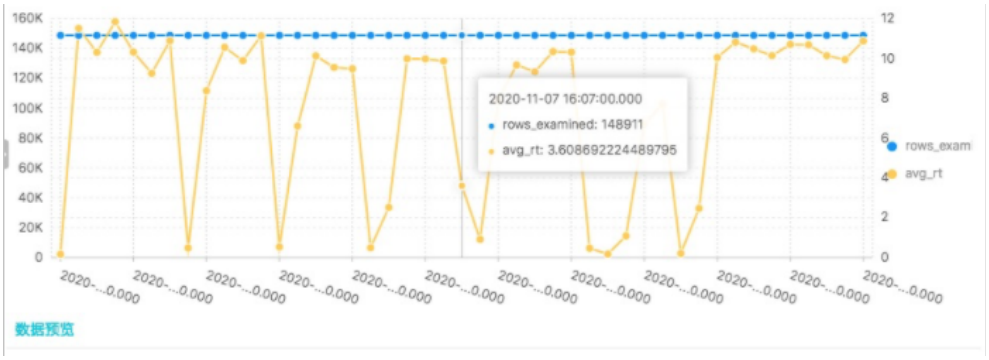
自动优化

DAS 基于全局 workload 和真实的业务场景，持续对数据库进行 SQL Review 和优化。按照经验，约 80% 的数据库性能问题可通过 SQL 优化手段解决，但 SQL 优化一直以来都是一个非常复杂的过程，需要多方面的数据库领域专家知识和经验，另外，由于 SQL 工作负载不断变化，SQL 优化还是一项非常耗时繁重的任务，这些都决定了 SQL 优化是一项高门槛，高投入且非常专业的工作。数据库自治服务 (DAS) 基于全局 workload 和真实的业务场景，持续对数据库进行 SQL Review 和优化，就像有一个不知疲倦的专业 DBA 一直在守护着您的数据库，时刻保障着数据库的可用性和稳定性。同时，DAS 的 SQL 诊断能力有与传统不同的技术特征，如：

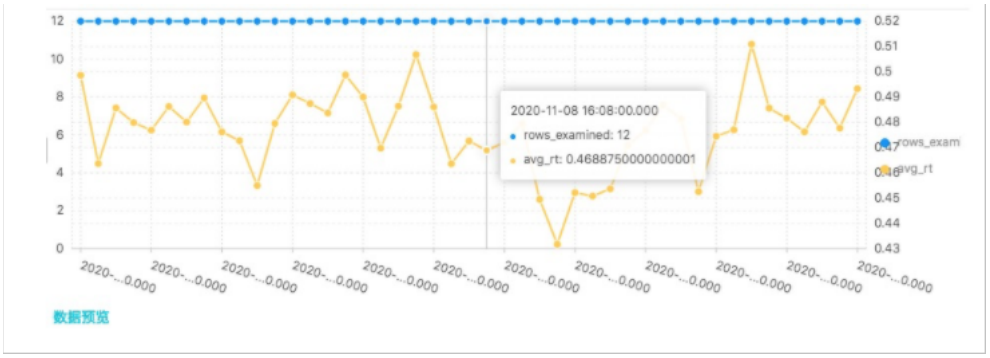
- 它采用外置式的，基于代价模型方式实现索引语句改写推荐，以及性能瓶颈问题识别和推荐，避免传统规则式的,过于机械化，推荐质量无法保证，无法量化性能提升收益等问题。
- DAS 在这些方面提供了足够覆盖度的场景：测试用例的正式特征库、在线用例的自动反馈提取、阿里巴巴多样化的应用场景。
- 基于全局的 Workload 优化，基于 Workload 特征，例如 SQL 执行频率，读写比等进行优化，最大限度地消除局部优化的片面性弊端。



下面是双 11 期间自动 SQL 优化一个实际案例：DAS 于 11 月 7 日通过负载异常检测到某实例因慢 SQL 引起的负载异常，自动触发 SQL 优化闭环，SQL 语句优化上线后，经过持续 24 小时优化效果跟踪完成优化收益评估，优化效果显著，如优化之前后的平均 RT 及扫描行数如下图所示，据统计，在优化之前，被优化 SQL 的平均扫描行数为 148889.198，平均 RT 为 505.561 毫秒。而优化之后，平均扫描行数为 12.132，大约降为优化前的万分之一，而平均 RT 降至 0.471 毫秒，也大约降低到优化前的千分之一。



自动 SQL 优化前平均 RT 及扫描行数



自动 SQL 优化后平均 RT 及扫描行数

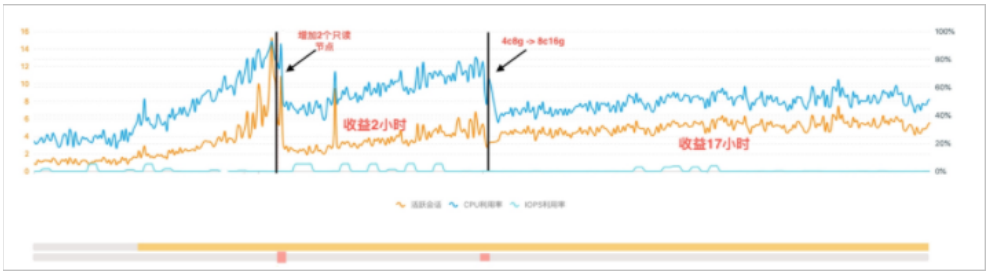
自动弹性

云上数据库提供基于计算规格的选项以及存储容量供用户选择，当用户业务 Workload 规模变化时可适当进行弹性扩缩容，但对于云原生应用而言，数据库能够根据业务 Workload 的变化自动决定最合适的规格，使用最小的资源完成业务所需的数据库容量。DAS 基于 AI 的时序列预测，能够自动对数据库的业务模型、容量水位进行计算和预测，实现及时按需自动扩缩容，即自动弹性 (Autoscaling)。

DAS 的自动弹性实现了一套完整的数据闭环，包含性能采集、决策中心、算法模型、规格建议模块、管控执行以及任务跟踪评估等。

- 性能采集负责对实例进行实时性能数据采集，涉及数据库的多项性能指标信息、规格配置信息、实例运行会话信息等。
- 决策中心模块则会根据当前性能数据、实例会话列表数据等信息进行全局判断，以基于根因的全局自治，例如可通过 SQL 限流来解决当前计算资源不足的问题则会采取限流处理。若确实为突增的业务流量，则会继续进行弹性服务流程。
- 算法模型是整个自动弹性服务的核心，负责对数据库实例的业务负载异常检测和容量规格模型推荐进行计算，解决核心的扩容时机、扩容方式、计算规格选择问题。
- 规格推荐和验证模块生成具体的推荐规范，并检查推荐的规范是否适合数据库实例的部署类型和实际运行环境。该模块还会重复检查推荐的规范是否可以在当前区域购买，这确保了推荐的规范可以在管理端顺利使用。
- 管控执行负责按照产出的规格建议进行分发执行。
- 状态跟踪最后用于衡量和跟踪规格变更前后数据库实例上的性能变化情况。

下图是双 11 期间自动 SQL 优化一个实际案例。某自治服务接入实例，用户的业务流量不断上升，PolarDB 实例的 CPU 使用率不断飙高并达到了高负载状态，此时 DAS 的 Autoscaling 算法精确的识别出了实例当前的异常状态，于是自动为实例增加了两个只读节点，实例的 CPU 使用率成功降到了较低的水位；在维持该状态两个小时后，用户实例的流量依然在不断上升，并第二次触发 DAS 的 Autoscaling，DAS Autoscaling 将实例的规格从 4 核 8GB 升级到 8 核 16GB，并平稳持续了 10 多个小时，帮助用户顺利度过了业务高峰期。



数据备份和恢复

数据是企业重要的生产资料，关键数据的丢失可能会给企业致命一击，因为数据是计算机系统存在的原因和基础。数据往往是不可再生的，一旦发生数据丢失，企业就会陷入困境：客户资料、技术文件、财务账目等客户、交易、生产数据可能被破坏得面目全非。 一般数据从生产到存储，主要经过应用、中间件、数据库、操作系统、存储或者磁盘驱动、网络、存储交换机等。其中数据库的备份设计尤为重要。

阿里云数据库本身提供基础备份的能力，而数据库备份 DBS（Database Backup）则为阿里云数据库附加了更强大更全面的备份恢复能力。

数据库备份 DBS 是阿里云提供的低成本、高可靠的云原生数据库备份平台。DBS 支持备份恢复 MySQL、SQL Server、PostgreSQL、Oracle 等近 10 种数据源，同时支持阿里云数据库、阿里云 ECS 数据库、本地数据中心、其他云厂商等环境，允许通过数据库网关 DG 私网、专线、VPN 网关等网络进行接入，有效解决备份耗时耗成本等问题。

数据库备份 DBS 提供数据全量备份、增量备份、异地备份以及数据恢复功能等，本文中重点介绍异地备份功能。

为什么需要异地备份？

作为完整数据库灾备方案，除了要有本地数据库备份外，还要在异地做数据库备份；传统方案是将备份集拷贝到本机其他盘、其他机器，这些都无法抵御地震、台风等自然灾害；如果要做到异地容灾，需要用户在其他地区自行搭建备份机房，前期投入成本很大。

阿里云瑶池数据库如何做到异地备份？

DBS 提供了异地备份（跨地域备份）的功能。阿里云瑶池数据库可以通过 DBS 将数据备份到阿里云 OSS 上，实现异地备份。异地备份可大大提高数据安全能力，您可以就近进行恢复、备份操作，降低数据库的 RPO 和 RTO，解决数据库异地灾备问题。

DBS 异地备份的优势有哪些？

传统的异地备份功能，需要通过专线或公网来打通地域间的网络，从而将数据同步至异地。而相比传统的异地备份，DBS 异地备份功能具有如下技术优势：

安全性高

- 网络安全：通过专有网络进行跨地域传输，不走公网，提高数据传输的安全性。
- 传输安全：在备份传输过程中，采用 SSL 和 AES256 加密技术，保护备份数据传输过程中的数据安全。
- 存储安全：提供 BYOK 功能，支持基于 KMS 实现备份数据加密，用户可以使用自己的 KMS 数据密钥加密备份数据。

备份传输速度稳定

DBS 所支持的异地备份功能支持自动转储（复制）备份数据，整合阿里云企业网 CEN（Cloud Enterprise Network），提供优质、高效、稳定的网络传输，帮助您轻松实现地域间数据备份与同步。

技术成本显著降低

支持灵活配置异地备份的备份策略与备份数据的生命周期，极大降低了操作成本与存储成本，例如可以配置：每天进行同地域的备份，每七天进行一次异地备份（即将数据备份至不同地域），将数据备份至异地的同时，可通过降低备份频率节省异地备份的存储成本。

转储策略灵活

DBS 的转储策略很灵活，无需和备份策略绑定。DBS 自有的调度算法可以为不同层级的存储之间提供独立的备份调度周期及数据过期保留时间（TTL/retention policy）。客户可以根据业务需要设置不同存储之间的保留策略和调度策略。

储介质多样化

DBS 转储支持多样化存储介质 / 协议，包括 OSS 存储（标准 OSS、归档 OSS、miniOSS）、NFS 协议存储（NAS）、Samba 协议存储、S3 协议存储、磁带等。

高可用容灾架构

高可用容灾是指在相隔较远的两地，建立两套或多套功能相同的 IT 系统，互相之间可以进行健康状态监视和功能切换，当一处系统因意外（如火灾、地震等）停止工作时，整个应用系统可以切换到另一处，使得该系统功能可以继续正常工作。

阿里云瑶池数据库如何做到高可用容灾？

阿里云数据传输服务 DTS（Data Transmission Service）提供了数据实时同步功能，可以保证主备系统的数据一致性，从而确保故障发生前后业务的连续性。您可以在同城 / 异地构建多个业务单元。各个业务单元之间通过 DTS 实现数据的双向实时同步，保障全局数据的一致性。当任何一个单元出现故障时，您只需将该单元的流量切换至其他单元即可，可实现业务的秒级恢复，有效地保障了服务的高可用性。以云数据库 RDS 为例，当主节点故障时，主备节点秒级完成切换，整个切换过程对应用透明；备节点故障时，RDS 会自动新建备节点以保障高可用。

DTS 支持的高可用容灾架构的优势有哪些?
阿里云瑶池数据库在高可用容灾场景下具有如下关键优势:

- 业务连续性
容灾不仅保护数据,更重要的目的在于保证业务的连续性。
- 数据实时性
在容灾系统中,两地的数据是实时一致的。
- 秒级切换
故障情况下,容灾系统的切换时间是几秒钟至几分钟。

根据灾备的等级,阿里云瑶池数据库的高可用容灾方案可分为同城容灾和异地容灾两种方式。前者的主备节点位于同一地域(同个可用区或不同可用区),后者的主备节点位于不同地域。异地容灾的高可用灾备等级比同城容灾更高,适用于对于业务连续性要求更高的行业(如电力行业),避免因整个地域长时间不可以造成的业务中断。

接下来我们详细了解一下同城容灾和异地容灾两种方案的架构和特点。

同城容灾

同城容灾,顾名思义就是指数据库集群的主备节点位于同一个地域(同个可用区或不同可用区),比如主备节点都位于杭州 A 可用区,或者主节点位于杭州 A 可用区,备节点位于杭州 B 可用区。

根据数据库的主备节点是否在同一个可用区,同城容灾可分为单可用区容灾和多可用区容灾。

单可用区容灾:
主备节点位于同一个可用区。主备节点位于两台不同的物理服务器上,可用区内的机柜、空调、电路、网络都有冗余,保障高可用性。

多可用区容灾:
主备节点位于同一地域的不同可用区,提供跨可用区的容灾能力。当一个可用区出现故障时,流量可以在短时间内切换到另一个可用区。整个切换过程对用户透明,应用代码无需变更。多可用区容灾具有比单可用区容灾更高的可用性。

同城容灾方式下,用户可选择复制加高可用、A-S (Active-Standby) 模式或 A-A (Active-Active) 模式:

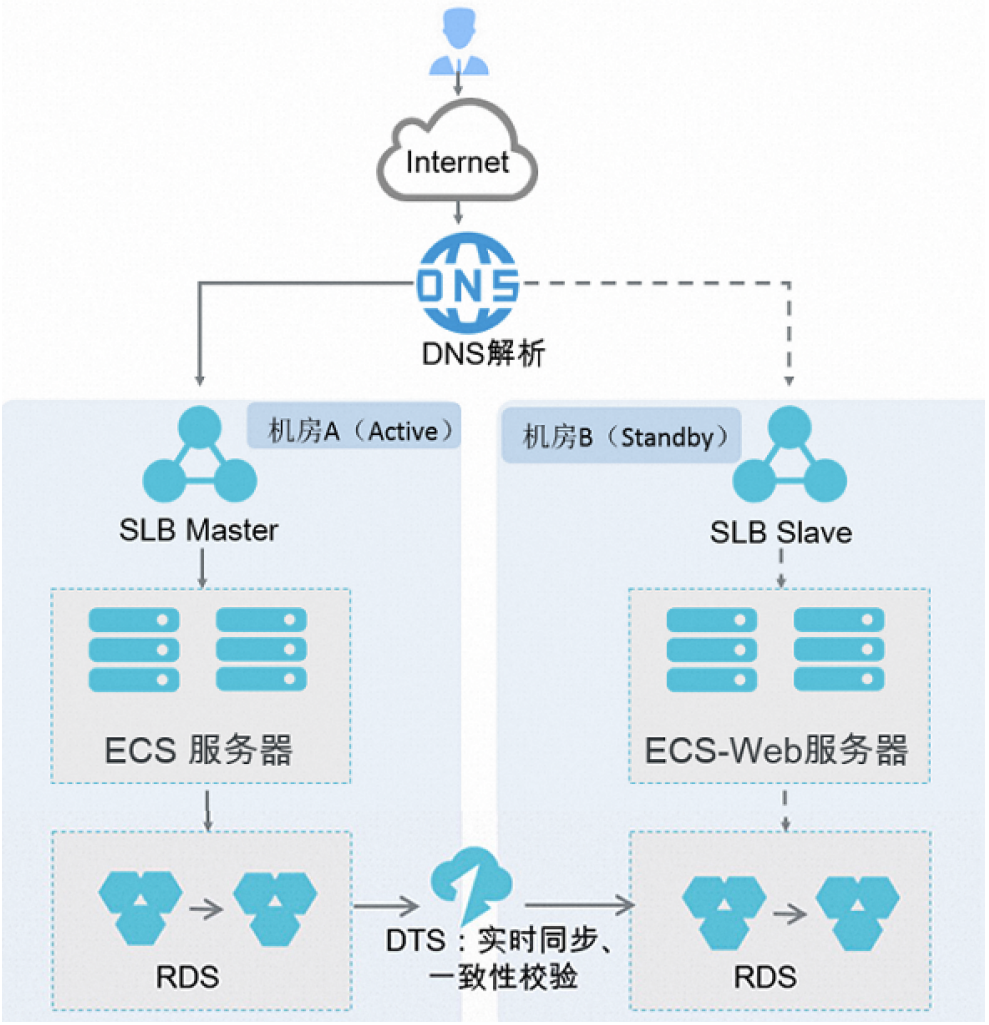
复制加高可用:
同城两机房中均部署数据库,通过复制从主用机房的数据库将数据复制备份至备机房的数据库中。当主用机房的数据库发生故障时,业务切换至备用机房的数据库。

A-S 模式:
同城两机房中部署完全一致的系统,其中一个机房 (Standby) 的资源完全用于备份,不对外提供业务。当主用机房 (Active) 发生故障时,业务切换至备用机房。

A-A 模式：

同城两机房中部署完全一致的系统，两个机房均对外提供业务，但在两个机房中均预留一部分资源作为备份。当其中一个机房发生故障时，业务会切换到另一个机房，占用另一个机房预留的资源。 如果用户资源充足且对同城容灾要求较高时，建议采用 A-S 模式；如果用户资源紧张，建议采用 A-A 模式。

我们以 A-S 模式为例对同城容灾架构进行详细介绍。以用户在阿里云的同一个 Region 购置了两套最简 IT 系统为例，数据库的 A-S 同城容灾解决方案架构示意如下：



架构说明：

- 在机房 A 和机房 B 分别部署对等的 RDS 数据库实例。
- RDS 实例之间通过 DTS 进行数据的实时同步和一致性校验，保证两个机房 RDS 数据的一致性。

备份与容灾情况：

- 如果机房 A 数据库故障或者机房整体故障，流量直接转移到机房 B 访问，暂时抛弃机房 A 的资源，等机房 A 恢复后重新配置为备选可用区。
- 如果只是应用端故障则将流量转移至机房 B 访问，同时需要对数据段做一致性校验，校验完成后，机房 B 成为主数据库，机房作为备数据库，数据同步的流向发生变化，由机房 B 同步到机房 A。

异地容灾

为了提供更高的可用性，阿里云数据库还支持跨地域的数据容灾。对于一些大型企业在业务安全性、服务可用性和数据可靠性方面既要求具备同城容灾又要求具备异地容灾时，可以采用异地容灾方案。

为什么需要异地容灾？

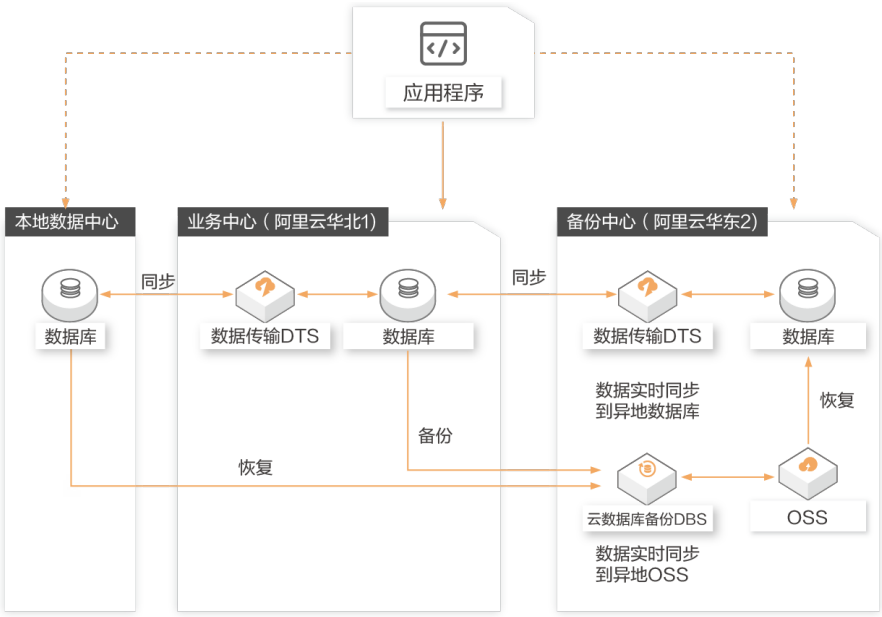
由于地区断电、断网等客观原因，产品可用性并不能达到 100%。当出现这些故障时，如果用户业务部署在单个地区，那么就会因为地区故障导致服务不可用，且不可用时间完全依赖故障恢复时间。 为了解决地区故障导致的服务不可用，提高服务可用性，可以构建异地灾备中心。当业务中心发生地区故障时，直接将业务流量切换到灾备中心，立刻恢复服务。

阿里云瑶池数据库如何做到异地容灾？

您可以通过数据传输服务 (DTS) 实现主实例和异地灾备实例之间的实时同步。主实例和灾备实例均具备主备高可用架构，当主实例所在区域发生突发性自然灾害等状况，主实例的主备节点均无法连接时，可将异地灾备实例切换为主实例，在应用端修改数据库连接地址，即可快速恢复应用的业务访问。 以云数据库 RDS 为例,用户可以将杭州地域的 RDS 实例 A 通过数据传输 DTS 服务异步复制到上海地域的 RDS 实例 B，实例 B 是一个完整独立的 RDS 实例，拥有独立的连接地址、账号和权限。

创建灾备实例后，当实例 A 所在地域发生短期不可恢复的重大故障时，用户在另外一个地域的实例 B 随时可以进行容灾切换,切换完成后,用户通过修改应用程序中的数据库连接配置,可以将应用请求转到实例 B 上，进而获得高于地域极限的数据库可用性。

如下为典型的异地灾备方案架构。通过数据库备份 DBS 和数据传输 DTS 构建备份中心。当业务中心故障时，将业务流量切换到备份中心。



当业务中心故障时，将业务流量切换到本地数据中心，或者备份中心

云原生数据库 PolarDB 稳定性最佳实践

在整体了解了阿里云数据库通用的稳定性能力和高可用容灾的稳定性架构后，下文以云原生数据库 PolarDB 为例，重点介绍阿里云瑶池数据库中交易型数据库的稳定性能力和保障。

PolarDB 简介

PolarDB 是阿里巴巴自研的新一代云原生数据库，在计算存储分离架构下，利用了软硬件结合的优势，为用户提供具备极致弹性、高性能、海量存储、安全可靠的数据库服务。100% 兼容 MySQL 和 PostgreSQL 生态，高度兼容 Oracle 语法。

PolarDB 共有三个引擎，分别为 PolarDB MySQL 版、PolarDB PostgreSQL 版、PolarDB 分布式版。经过阿里巴巴多年双十一活动的最佳实践验证，让用户既享受到开源的生态灵活性，又获得商业云原生数据库的高性能和安全等企业级能力。

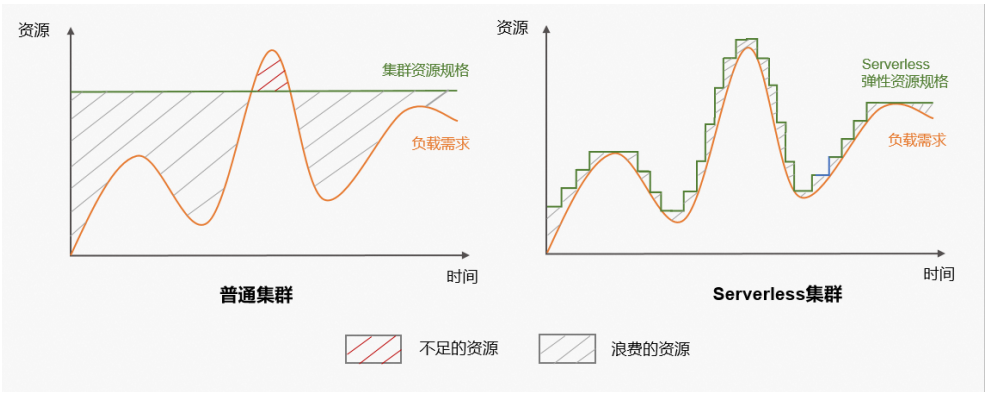
PolarDB 稳定性和高可用能力

Serverless 秒级动态弹降

云计算环境下，企业和个人通过开启云服务，即可以得到所需的软件功能、计算资源、存储空间，并按实际使用量付费。在业务量逐步上涨的过程中，用户需要不断提升计算和存储资源来满足业务需要。因此，扩展性是云数据库服务非常重要的服务指标。在创建云数据库时，客户往往需要比较保守的去配置数据库集群的资源，包括 CPU、内存、存储以及连接数等多种参数配置，以确保业务能够在波峰和波谷都能平稳运行。在这种情况下，客户购买的集群资源在业务波谷时期会被闲置，导致整体成本偏高；更重要的是，在业务压力增长阶段，集群资源又应对不足，影响业务的连续性和稳定性。

在上文中,我们详细介绍了阿里云数据库自治服务 (DAS) 为阿里云数据库提供的自动弹性能力。而除此之外，在 PolarDB 中，PolarDB Serverless 集群提供了 CPU、内存、存储、网络资源的秒级动态弹降能力，为客户提供更灵活、弹升范围更广、弹性速度更快，并且性价比也更高的扩展性方案。

在业务波动较大的场景下，普通集群和 PolarDB Serverless 集群资源使用和规格变化情况如下图：



由上图可以看到，在业务波动较大的场景下，普通集群和 Serverless 集群的集群资源在不同业务负载情况下的表现有如下显著对比：

普通集群：在波谷期浪费的资源较多，在高峰期资源不足，业务受损。

Serverless 集群：

- 由于其规格随业务需求量随时调整，总体浪费的资源很少，提升了资源利用率，降低了资源使用量。
- 在高峰期也能完全满足业务需求，保证业务不受损，提高了系统的稳定性。
- 打破固定资源付费模式，真正做到了负载与资源动态匹配的按量付费模式，可节省大量成本。
- 无需手动变配，提高了运维效率，提升了运维管理人员和研发人员的幸福感。
- 支持自动启停能力。当没有连接时，集群自动暂停，释放计算成本；当请求到来时，集群自动无感启动。
- 对高吞吐写入场景和高并发业务场景进行了设计优化，同时提供了弹性伸缩能力，适合业务数据量大、并具有典型的业务访问波峰波谷场景。

而相比 DAS 提供的自动弹性能力，Serverless 集群秒级动态弹降的特性具有如下显著优势：

高可用：多节点的架构保障了 Serverless 集群的高可用，服务等级协议 SLA 与普通集群相同，共同保证了 Serverless 集群的稳定运行。

高弹性:

- 弹升范围广: Serverless 业内自动弹升范围最广的云数据库，支持自动横向扩容，单集群支持 0~1000 核范围内的无感伸缩。
- 秒级弹升: 从容应对业务负载突增，5 秒完成探测，1 秒完成弹升；同时在业务负载下降时，集群资源阶梯性自动释放。
- 业务无感: 弹升过程业务无影响。

数据强一致: 支持高性能模式的全局一致性,在集群内实现数据强一致,数据写入后在只读节点上立即可读,性能与弱一致性基本一致。

低成本: 以计算能力（PCU）定价，真正做到按量付费，帮助客户节省成本。成本下降最高可达 80%。

高数据可靠性存储

在存储领域，通常会采用“九个 9”来作为一个存储系统数据可靠性和持久性的衡量指标。“九个 9”（Nine nines）即 99.9999999%，也被称为“可靠性的理论极限”（Theoretical Limit of Reliability）。99.9999999% 的数据可靠性意味着在一定时间内，只有很少的可能会发生故障或出现数据丢失的情况。具体来说,这个数字表示在一亿个数据操作中,只有一个可能会导致数据丢失或系统故障。这个数字非常小,说明了这个存储系统的数据可靠性非常高。

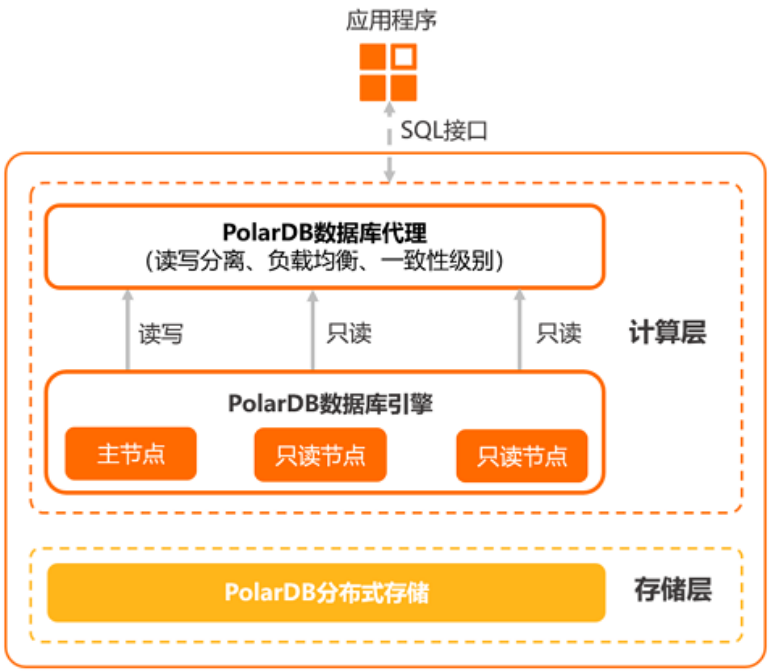
云原生数据库 PolarDB 是阿里巴巴自研的云原生 HTAP 数据库，基于云原生架构、计算存储分离、软硬件一体化设计，为用户提供具备超高弹性和性能、高可用和高可靠保障、高性价比的数据库服务。在存储方面，PolarDB 数据库服务提供了两种存储类型: PSL4（PolarStore Level 4）和 PSL5（PolarStore Level 5）。其针对数据可靠性和数据持久性的承诺如下:

存储类型	数据可靠性保证
PSL5	99.99999999%（十个9）
PSL4	99.9999999%（九个9）

从上表可以看到，PSL5 的数据可靠性更强（达到了“十个 9”），更适合于对性能和可靠性要求高，以数据库为核心系统的业务场景，如金融、电商、政务和大中型互联网业务等。而 PSL4 存储，通过阿里巴巴自研的硬件压缩盘（Smart-SSD）技术，在物理 SSD 磁盘层面压缩、解压缩存储的数据，在数据可靠性依旧达到业内顶尖的“九个 9”的基础上，实现了更低的存储价格，为客户提供了更高性价比的选择。

热备秒级切换

PolarDB 采用双活（Active-Active）的高可用集群架构。当系统发生故障时，可读写的主节点和只读节点之间会自动进行故障切换（Failover）。



在此基础上，PolarDB 支持为只读节点开启热备功能，使其成为“热备节点”，从而在主备切换的过程中实现快速切换和事务保持，并实现 5~10 秒内完成主备切换，真正达到“秒级切换”，事务不中断，应用无报错。

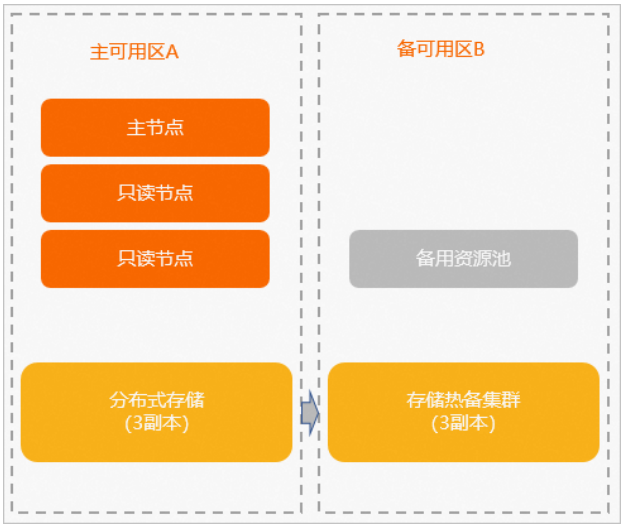
PolarDB 的热备秒切技术从故障探测、切换速度和切换体验三个方面对切换场景进行了优化，包括计划内的切换，如集群升降配和小版本升级，以及计划外的容灾切换。整合了多项技术，来解决用户的痛点问题：

- 引入全新的高可用模块 Voting Disk Service（简称 VDS），该模块基于共享存储架构，实现自治的集群节点管理，大幅降低故障检测和集群选主耗时。
- 新增支持全局预热系统的热备节点，通过对存储引擎内部的多个模块提前预热，优化升主的执行耗时。
- 结合数据库代理（PolarProxy），支持连接保持和事务保持功能。在集群升降配或小版本升级过程中，开启连接保持和事务保持功能后，系统会尽可能地保证用户的连接和事务不中断，实现基本无感知的主动运维。

跨可用区自动秒级切换

除了单个可用区内的多节点高可用架构的热备秒级能力外，PolarDB 还支持多可用区高可用架构。相比单可用区，多可用区架构具备更高的容灾能力，可以抵御机房级别的故障：当主可用区故障（如主可用区所有计算节点同时故障）时，集群会自动进行主备可用区切换，集群可秒级切换到备可用区，确保集群的可用性。

在多可用区架构中，数据分布在多个可用区内，主可用区和备可用区各保存 3 副本数据（共 6 副本数据），具有更高的 SLA 可靠性保障。计算节点暂时要求位于主可用区，备可用区的存储热备集群用于主可用区故障时进行故障切换。



存储热备集群部署在 PolarDB 集群所在地域的备可用区或者同一可用区下的不同机房，具有独立的存储能力，用于集群的热备切换。当 PolarDB 整个集群或者主可用区不可用时，存储热备集群会快速切换为主节点承担集群的读写业务。

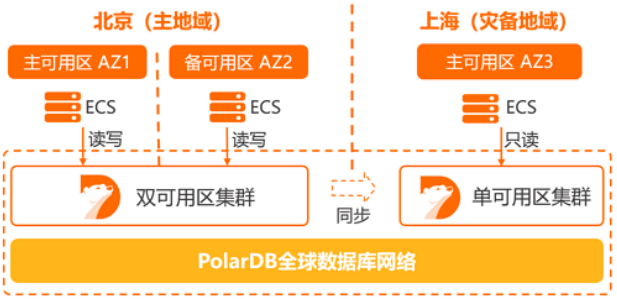
全球数据库 GDN

全球数据库网络（Global Database Network，简称 GDN）是 PolarDB 在异地容灾场景下的高可用实践，其是由分布在同一个国家内多个地域的多个 PolarDB 集群组成的网络。每个 GDN 中包含一个主 PolarDB 集群（Primary）和多个从 PolarDB 集群（Secondary），多个 PolarDB 集群之间数据保持同步。当主集群出现地域级别的故障时，可以将业务分钟级切换到从集群，从而快速恢复业务。

GDN 通常应用在异地灾备和异地多活两类场景中。下文将详细介绍 GDN 在异地灾备场景下的常见部署架构和优势。

异地灾备场景通常应用在银行、保险、证券等行业，不论业务部署在一个或多个地域，都能通过 GDN 实现异地灾备。以下以两地三中心架构为例，可以看到，一个异地灾备场景下的 GDN 由两个 PolarDB 集群组成：

- 北京的双可用区集群，覆盖北京 AZ1 和北京 AZ2。
- 上海的单可用区集群，位于上海 AZ3。



应用部署在北京，对北京 AZ1 的数据库进行本地读写。当北京的双可用区集群出现故障时，该架构下的 GDN 的异地灾备策略如下：

- 当北京 AZ1 故障时，优先切换到北京 AZ2。
- 当北京 AZ1 和 AZ2 均故障时，切换到上海 AZ3。