云数据迁移

安全白皮书

文档版本 01

发布日期 2023-03-30

华为云计算技术有限公司

版权所有 © 华为云计算技术有限公司 2024。 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。

商标声明


和其他华为商标均为华为技术有限公司的商标。
本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目 录

 版权所有 © 华为云计算技术有限公司

1 云数据迁移安全简介

CDM 简介

云数据迁移（Cloud Data Migration，简称CDM）提供同构/异构数据源之间批量数据迁移服务，帮助客户实现数据自由流动。支持多种常用数据源，如客户自建或公有云上的文件系统，关系数据库，数据仓库，NoSQL数据库，大数据云服务，对象存储等数据源。

CDM适用于以下场景：

1. 数据上云：使用华为公有云服务时，用户可以将其历史数据或增量数据从私有云/本地数据中心/第三方公有云迁移到华为云。
2. 云服务间数据交换：用户可以在华为云的大数据服务、数据库服务、对象存储服务之间相互迁移数据。例如，可以将由MapReduce服务（MapReduce Service，简称MRS）处理的数据导入到数据仓库服务（Data Warehouse Service，简称DWS），进行交互式分析和报告统计收集。
3. 云上数据回迁到本地：用户在使用公有云计算资源对海量数据进行处理后，将结果数据回流到本地业务系统，主要是各种关系型数据库和文件系统。

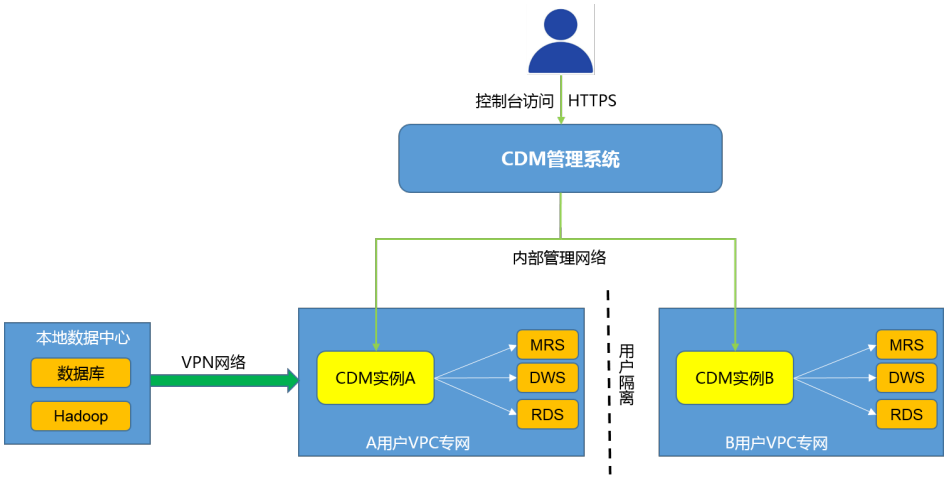
CDM 迁移原理

用户使用CDM服务时，CDM管理系统在用户VPC中发放全托管的CDM实例。此实例仅提供控制台和Rest API访问权限，用户无法通过其他接口（如SSH）访问实例。这种方式保证了CDM用户间的隔离，避免数据泄漏，同时保证VPC内不同华为云服务间数据迁移时的传输安全。用户还可以使用VPN网络将本地数据中心的数据迁移到华为云服务，具有高度的安全性。

CDM数据迁移以抽取-写入模式进行。CDM首先从源端抽取数据然后将数据写入到目的端，数据访问操作均由CDM主动发起，对于数据源（如RDS数据源）支持SSL时，会使用SSL加密传输。迁移过程要求用户提供源端和目的端数据源的用户名和密码，这些信息将存储在CDM实例的数据库中。保护这些信息对于CDM安全至关重要。

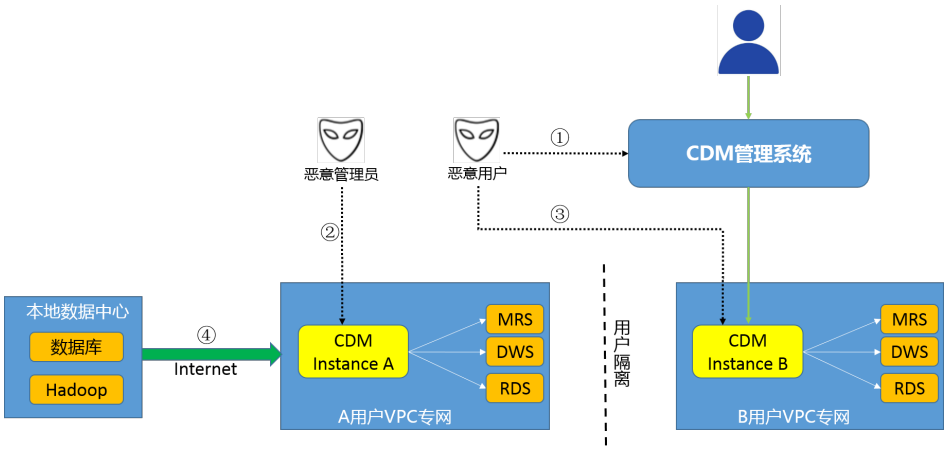
 版权所有 © 华为云计算技术有限公司

图 1-1 CDM 迁移原理



安全边界和风险规避

图 1-2 风险规避



如上图所示，CDM可能存在以下威胁：

1. 互联网威胁：恶意用户可能通过CDM控制台攻击CDM。
2. 数据中心威胁：恶意CDM管理员获取用户的数据源访问信息（用户名和密码）。
3. 恶意用户威胁：恶意用户窃取其他用户的数据。
4. 数据暴露公网：从公网迁移数据时暴露数据的威胁。

对于这些潜在的威胁，CDM提供以下机制来规避终端用户的风险：

1. 针对互联网威胁：用户不能直接通过公网登录CDM控制台。CDM提供双层安全保障机制。
   a. 华为云控制台框架要求用户访问任何控制台页面都要进行用户认证。
   b. Web应用防火墙（Web Application Firewall，简称WAF）过滤所有控制台的请求内容并停止请求攻击代码/内容。
2. 针对数据中心威胁：用户必须向CDM系统提供迁移源端和目的端的访问用户名和密码信息，才能完成数据迁移。避免CDM管理员获取此类信息并攻击用户的重要数据源对于CDM非常重要，CDM为此类信息提供三级保护机制。

版权所有 © 华为云计算技术有限公司

a. CDM在本地数据库中存储经过AES-256加密的密码，确保用户隔离。本地数据库使用用户Ruby运行，数据库仅侦听127.0.0.1，用户没有远程访问数据库的权限。

b. 用户实例发放完毕后，CDM将虚拟机的root和Ruby用户密码更改为随机密码且不会保存在任何地方，以阻止CDM管理员访问用户实例和含有密码信息的数据库。

c. CDM实例迁移以推拉模式进行，因此CDM实例在VPC上没有侦听端口，用户无法从VPC访问本地数据库或操作系统。

3. 针对恶意用户的威胁：CDM对每个用户，使用单独的虚拟机来运行各自的CDM实例，用户之间的实例是完全隔离和安全的。恶意用户无法访问其他用户的实例。

4. 针对数据暴露公网的威胁：CDM的抽取-写入模型下，即使CDM绑定了弹性IP，也不会开放端口到弹性IP，攻击者无法通过弹性IP来访问和攻击CDM。不过从公网迁移数据的方式下，由于用户数据源也会暴露在公网，存在被第三方攻击的威胁，推荐用户在数据源服务器上通过ACL或防火墙对源端进行防护，比如仅放通来自CDM绑定的弹性IP的访问请求。

 版权所有 © 华为云计算技术有限公司

2 云数据迁移安全结论

访问控制

只有华为云统一身份认证服务（Identity and Access Management，简称IAM）授权的用户才能访问CDM控制台和API。推拉模式下，CDM在VPC上没有开放侦听端口，用户无法从VPC访问实例，具有高度的安全性。

数据传输安全

CDM在用户VPC中运行，网络隔离确保数据传输的安全性。支持SSL的数据源，如RDS、SFTP等，可以使用SSL。CDM还支持公网数据源的数据上云，用户可以利用VPN和SSL技术来避免传输安全风险。

用户和网络隔离

CDM实例运行在用户独立的VPC内，VPC允许用户通过配置VPC入站IP范围，来控制连接CDM的IP地址段。CDM实例部署在用户VPC后，用户可以综合运用子网和安全组的配置，来完成CDM实例的隔离，提升CDM实例的安全性。

数据加密

用户数据源的访问信息（用户名和密码）存储在CDM实例的数据库中，并采用AES-256加密，CDM管理员无法访问。

数据删除

用户删除CDM实例时，存储在实例中的数据都会被删除，任何人都无法查看及恢复数据。

版权所有 © 华为云计算技术有限公司