

# 使用 AWS 执行备份和恢复的方法

2016 年6月



© 2016, Amazon Web Services, Inc. 或其附属公司。保留所有权利。

## 声明

本文档仅用于参考。本文档代表截至其发行之日的 **AWS** 的最新产品服务和实践，如有变更，恕不另行通知。客户负责对此文件的信息以及对 **AWS** 的产品或服务的任何使用进行自我独立的评估，每项产品或服务均按“原样”提供，且不提供任何类型的保证，不管是明示还是暗示。本文档不形成 **AWS**、其附属公司、供应商或许可方的任何保证、表示、合同承诺、条件或担保。**AWS** 对其客户承担的责任和义务受 **AWS** 协议制约，本文档不是 **AWS** 与客户之间的协议的一部分，也不构成对该协议的修改。

# 目录

摘要	4
简介	4
为何将 <b>AWS</b> 用作数据保护平台？	4
用于数据保护的 <b>AWS</b> 存储服务	5
<b>Amazon S3</b>	5
<b>Amazon Glacier</b>	6
<b>AWS Storage Gateway</b>	6
<b>AWS Transfer Services</b>	6
设计备份和恢复解决方案	7
云原生基础设施	8
基于 <b>EBS</b> 快照的保护	9
数据库备份方法	14
本地到 <b>AWS</b> 基础设施	17
混合环境	20
将基于 <b>AWS</b> 的应用程序备份到数据中心	21
将备份管理迁移到云以实现可用性	21
混合方案示例	22
使用 <b>AWS</b> 对数据进行存档	23
在 <b>AWS</b> 中保护备份数据	24
结论	24
撰稿人	25
文档修订	25

## 摘要

本文的目标受众是企业解决方案架构师、备份架构师和负责保护企业 IT 环境中数据的 IT 管理员。本文介绍了可使用 AWS 实施来增强或替换备份和恢复解决方案的生产工作负载和体系结构。利用这些方法，可以降低成本并提高可扩展性和持久性，从而满足恢复时间目标 (RTO)、恢复点目标 (RPO) 和合规性要求。

## 简介

随着企业数据的增长速度加快，保护数据的任务变得更具挑战性。有关备份方法的持久性和可扩展性的问题已变得司空见惯，其中包括以下问题：云如何帮助满足我的备份和存档需求？

本文讲述了大量备份架构（云原生应用程序、混合环境和本地环境）和关联的 AWS 服务，它们可用于构建可扩展的可靠数据保护解决方案。

## 为何将 AWS 用作数据保护平台？

Amazon Web Services (AWS) 是一款安全、高性能、灵活、高性价比且易于使用的云计算平台。AWS 负责处理各种各样繁重的工作，并提供可用于构建可扩展的备份和恢复解决方案的工具和资源。

将 AWS 纳入到您的数据保护战略中，可以获得许多优势：

- **持久性：** [Amazon Simple Storage Service \(Amazon S3\)](#) 和 [Amazon Glacier](#) 旨在为其中存储的对象提供 99.999999999%（11 个 9）的持久性。这两个平台都提供了用于备份数据的可靠场所。
- **安全性：** AWS 提供了用于对传输中数据和静态数据进行访问控制和加密的大量选项。
- **全球基础设施：** 在全球范围都可以享有 AWS 服务，您可以在满足合规性要求的区域中备份和存储数据。

- **合规性：** AWS 基础设施已获得符合多项标准的认证，例如服务性机构控制体系 (SOC)、《鉴证业务准则公告第 16 号》(SSAE 16)、国际标准化组织 (ISO) 27001、支付卡行业数据安全标准 (PCI DSS)、健康保险可携与责任法 (HIPPA)、[SEC<sup>1</sup>](#)、联邦风险与授权管理项目 (FedRAMP)，您可以轻松地让备份解决方案融入您现有的合规性体系中。
- **可扩展性：** 使用 AWS，您无需担心容量。可根据您的需求变化来增加或减少使用量，而不会产生管理开销。
- **更低的 TCO：** AWS 的运营规模降低了服务成本，并且有助于降低存储的总拥有成本 (TCO)。AWS 通过降低价格的方式让客户享受到成本节约带来的实惠。
- **按需支付定价：** 在需要 AWS 服务时才购买，并且可以只在计划使用这些服务的时段内购买。AWS 定价不需要预付费用，不会产生服务终止罚金，也不需要签订长期合同。

## 用于数据保护的 AWS 存储服务

Amazon S3 和 Amazon Glacier 是用于备份和存档的理想服务。这二者都是持久的、低成本的存储平台。它们都提供无限容量，并且在备份数据集增长时无需进行卷或介质管理。按实际用量付费模式和较低的每月每 GB 成本，使得这些服务非常适合数据保护使用情形。

### Amazon S3

Amazon S3 提供了高度安全的、可扩展的对象存储。

可以使用 Amazon S3 随时从任何位置在 Web 上存储和检索任意数量的数据。Amazon S3 将数据作为对象存储在称为“存储桶”的资源中。AWS Storage Gateway 和许多第三方备份解决方案可以为您代为管理 Amazon S3 对象。您可以在存储桶中存储任意所需数量的对象，并且可以在存储桶中写入、读取和删除对象。单个对象的最大大小为 5 TB。

---

<sup>1</sup> <https://aws.amazon.com/about-aws/whats-new/2015/09/amazon-glacier-receives-third-party-compliance-assessment-for-sec-rule-17a-4f-from-cohasset-associates-inc/>

Amazon S3 提供了一系列存储分类，适合不同的使用情形。包括：

- **Amazon S3 标准**，适用于频繁访问的数据的通用存储。
- **Amazon S3 标准 — 不经常访问**，适用于长时间保存但不常访问的数据。
- **Amazon Glacier**，适用于长期存档。

Amazon S3 还提供了生命周期策略，您可以配置这些策略以便在整个数据生命周期内管理数据。设置策略后，无需更改您的应用程序，即可将数据迁移到相应分类的存储。有关更多信息，请参阅 [S3 存储分类](#)。

## Amazon Glacier

Amazon Glacier 是一项成本极低的云存档存储服务，可为数据存档和联机备份提供安全持久的存储。为了保持低成本，Amazon Glacier 针对不经常访问的数据（可以接受若干小时的检索时间的数据）进行了优化。借助 Amazon Glacier，您能够可靠地存储大量或少量数据，费用低至每月每 GB 0.007 美元，与本地解决方案相比，大大节省了成本。Amazon Glacier 非常适合存储具有长期或无限期保留要求的备份数据以及长期数据存档。有关更多信息，请参阅 [Amazon Glacier](#)。

## AWS Storage Gateway

AWS Storage Gateway 可将本地软件设施与基于云的存储连接起来，在本地 IT 环境与 AWS 存储基础设施之间提供无缝、高度安全的集成。有关更多信息，请参阅 [AWS Storage Gateway](#)。

## AWS Transfer Services

除了第三方网关和连接器之外，您还可以使用 AWS 选项（如 AWS Direct Connect、AWS Snowball、AWS Storage Gateway 和 Amazon S3 Transfer Acceleration）来快速传输您的数据。有关更多信息，请参阅 [云数据迁移](#)。

## 设计备份和恢复解决方案

在制定用于备份和还原数据的全面性策略时，您必须先确定可能发生的失败或灾难情况及其可能对业务造成的影响。在某些行业，您必须考虑数据安全、隐私和记录保留方面的法规要求。

您在实施备份过程时，应该能够提供适当的精细度级别来满足业务的 **RTO** 和 **RPO**，包括：

- 文件级恢复
- 卷级恢复
- 应用程序级恢复（例如，数据库）
- 映像级恢复

以下部分介绍基于基础设施组织结构的备份、恢复和存档方法。IT 基础设施可大致分为以下类别：云原生、本地和混合。

## 云原生基础设施

此方案是指完全存在于 AWS 中的工作负载环境。如下图所示，它包括 Web 服务器、应用程序服务器、监控服务器、数据库和 Active Directory。

如果您在 AWS 上运行您所有的服务，则可利用许多内置功能来满足数据保护和恢复需求。

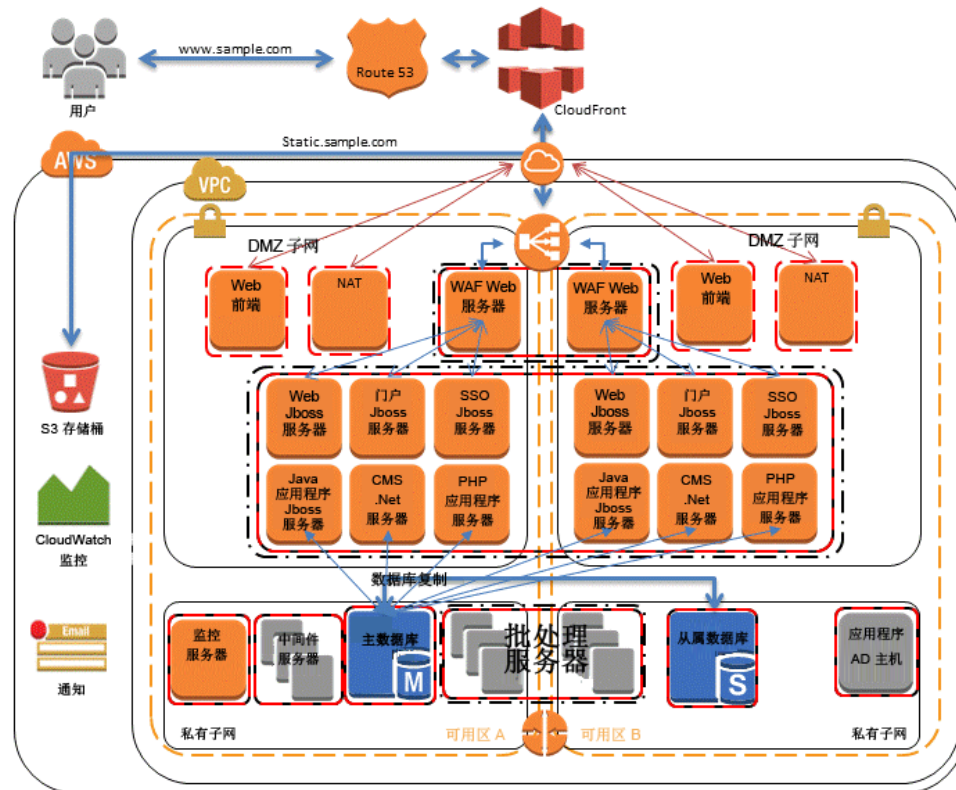


图 1: AWS 原生方案

## 基于 EBS 快照的保护

当服务在 [Amazon Elastic Compute Cloud](#)<sup>2</sup> (Amazon EC2) 中运行时，计算实例可使用 Amazon Elastic Block Store (Amazon EBS) 卷来存储和访问主要数据。可将此数据块存储用于结构化数据（例如数据库）或非结构化数据（例如卷上的文件系统中的文件）。

借助 Amazon EBS，您能够创建任何 Amazon EBS 卷的快照（备份）。它会制作卷的副本并将该副本放入 Amazon S3，其中该副本将以冗余方式存储在多个可用区中。第一个快照是卷的完整副本；后续的快照操作只会存储增量的数据块级更改。

这种方法可以快速且可靠地还原完整卷数据。如果您只需部分还原，则可以将卷附加到正在运行的实例（采用不同设备名称），安装该实例，然后使用操作系统复制命令将数据从备份卷复制到生产卷。

也可以使用控制台中提供的 Amazon EBS 快照复制功能或者从命令行，在 AWS 区域之间复制 Amazon EBS 快照，如 [Amazon Elastic Cloud Compute 用户指南](#) 中所述。<sup>3</sup> 您可以使用此功能将备份存储到其他区域，而不必管理底层复制技术。

---

<sup>2</sup> <http://aws.amazon.com/ec2/>

<sup>3</sup> <http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-copy-snapshot.html>

## 创建 EBS 快照

在创建快照时，您可以直接将数据存储到基于磁盘的持久存储来保护数据。可以使用 AWS 管理控制台、命令行界面 (CLI) 或 API 来创建 Amazon EBS 快照。

在 Amazon EC2 控制台的 **Elastic Block Store Volumes** 页上，从 **Actions** 菜单中选择 **Create Snapshot**。在 **Create Snapshot** 对话框中，选择 **Create** 以创建将存储在 Amazon S3 中的快照。

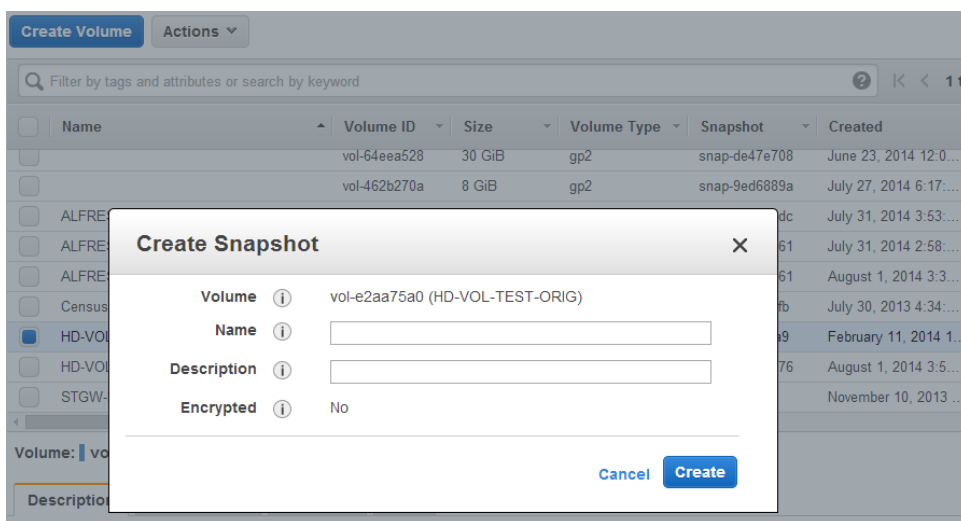


图 2: 使用 EC2 控制台创建快照

要使用 CLI 命令创建快照，请运行以下命令：

```
➤ aws ec2 create-snapshot
```

您可以计划定期运行 `aws ec2 create-snapshot` 命令来备份 EBS 数据。Amazon S3 的定价十分实惠，使您能够保留许多代的数据。而且，由于快照基于数据块，因此在创建初始快照后，只会针对已更改的数据消耗空间。

## 从 EBS 快照还原

要从快照还原数据，您可以使用 **AWS** 管理控制台、**CLI** 或 **API** 从现有快照创建卷。

例如，执行以下步骤可将卷还原到以前的时间点备份：

1. 使用以下命令从备份快照创建卷：

```
➤ aws ec2 create-volume --region us-west-1b --snapshot-id mysnapshot-id
```

2. 在 **Amazon EC2** 实例上，卸载现有卷。

在 **Linux** 中，使用 `umount`。在 **Windows** 中，使用 **Logical Volume Manager (LVM)**。

3. 使用以下命令从实例分离现有卷：

```
➤ aws ec2 detach-volume --volume-id oldvolume-id --instance-id myec2instance-id
```

4. 使用以下命令附加已从快照创建的卷：

```
➤ aws ec2 attach-volume --volume-id newvolume-id --instance-id myec2instance-id --device /dev/sdf
```

5. 在正在运行的实例上重新安装卷。

## 创建一致备份或热备份

执行备份时，最好是让系统处于不执行任何 I/O 的状态。在理想情况下，计算机不接受流量，但是随着 24/7 IT 运营成为常态，这种理想状态越来越少见。

因此，您必须停用文件系统或数据库，才能创建干净备份。具体操作方式取决于数据库或文件系统。

针对数据库的过程如下：

- 如果可能，将数据库设为热备份模式。
- 运行 **Amazon EBS** 快照命令。
- 使数据库退出热备份模式；如果正在使用只读副本，则终止只读副本实例。

针对文件系统的过程类似，但取决于操作系统或文件系统的功能。例如，对于 **XFS** 文件系统来说，可以刷新其数据来获得一致备份。有关更多信息，请参阅 [xfs freeze](#)。<sup>4</sup>

如果文件系统不支持冻结功能，则您应卸载文件系统，发出快照命令，然后重新安装文件系统。或者，您可以使用支持 I/O 冻结的逻辑卷管理器来实现此过程。

快照过程在后台继续进行，创建快照操作的执行速度非常快，并且捕获的是时间点，因此您备份的卷只需卸载几秒钟时间即可。由于备份占用时间尽可能短，因此中断时间是可预测的，并且可安排好时间。

---

<sup>4</sup> <https://access.redhat.com/documentation/en-US/Red Hat Enterprise Linux/6/html/Storage Administration Guide/xfsfreeze.html>

## 执行多卷备份

在某些情况下，您可以使用逻辑卷管理器跨多个 Amazon EBS 卷使数据条带化，以此提高潜在吞吐量。使用逻辑卷管理器（例如 `mdadm` 或 `LVM`）时，需要从卷管理器层（而不是底层 EBS 卷）执行备份，这十分重要。这可确保所有元数据都一致，并且子组件卷条理清晰。

可通过多种方式达到此目的。例如，您可以使用 [alestic.com](https://github.com/alestic/ec2-consistent-snapshot)<sup>5</sup> 创建的脚本。内存缓冲区应刷入到磁盘中；应停止文件系统到磁盘的 I/O；并且应同时为构成 RAID 集的所有卷启动快照。在启动卷的快照（通常是另外的 1 个或 2 个）后，文件系统可以继续其操作。应该对快照进行标记，以便能够在还原期间集中管理这些快照。

您也可以从逻辑卷管理器或文件系统级别执行这些备份。在这些情况下，通过使用传统备份代理，您能够通过网络来备份数据。互联网和 [AWS Marketplace](https://aws.amazon.com/marketplace/) 中提供了大量基于代理的备份解决方案。<sup>6</sup> 请记住，基于代理的备份软件需要一致的服务器名称和 IP 地址。因此，将这些工具与 Amazon [Virtual Private Cloud \(VPC\)](https://aws.amazon.com/vpc/)<sup>7</sup> 中部署的实例结合使用是确保可靠性的最佳方法。

另外一种方法是创建主系统卷的副本（存在于单个大型卷上）。这样可简化备份过程，因为只需备份一个大型卷，并且备份不在主系统上进行。但是，您应先确定单个卷能否在备份期间充分地执行，以及最大卷大小是否适合应用程序。

---

<sup>5</sup> <https://github.com/alestic/ec2-consistent-snapshot>

<sup>6</sup> <https://aws.amazon.com/marketplace/>

<sup>7</sup> <http://aws.amazon.com/vpc/>

## 数据库备份方法

AWS 提供了许多针对数据库的选项。您可以在 EC2 实例上运行自己的数据库，或者使用由 [Amazon Relational Database Service](#)<sup>8</sup>(Amazon RDS) 提供的托管服务数据库选项之一。如果您在 EC2 实例上运行您自己的数据库，则可使用本机工具（例如，[MySQL](#)<sup>9</sup>、[Oracle](#)<sup>10</sup>、[MSSQL](#)<sup>11</sup>、[PostgreSQL](#)<sup>12</sup>）将数据备份到文件，或使用“[基于 EBS 快照的保护](#)”中介绍的方法之一来创建包含数据的卷的快照。

### 使用数据库副本备份

对于基于 Amazon EBS 卷的 RAID 集构建的数据库，可以通过创建数据库的只读副本来消除主数据库上的备份工作。这是在单独的 Amazon EC2 实例上运行的数据库的最新副本。可以使用与源类似的多个磁盘创建副本数据库实例，也可以将数据整合到单个 EBS 卷中。随后，您可以使用“[基于 EBS 快照的保护](#)”中介绍的过程之一来为 EBS 卷拍摄快照。通常，此方法用于需要全天候运行的大型数据库。在这种情况下，所需的备份时段会过长，而生产数据库是不可能关闭如此长时间的。

---

<sup>8</sup> <https://aws.amazon.com/rds/>

<sup>9</sup> <http://dev.mysql.com/doc/refman/5.7/en/backup-and-recovery.html>

<sup>10</sup> [http://docs.oracle.com/cd/E11882\\_01/backup.112/e10642/rcmbckba.htm#BRADV8003](http://docs.oracle.com/cd/E11882_01/backup.112/e10642/rcmbckba.htm#BRADV8003)

<sup>11</sup> <http://msdn.microsoft.com/en-us/library/ms187510.aspx>

<sup>12</sup> <http://www.postgresql.org/docs/9.3/static/backup.html>

## 将 Amazon RDS 用于备份

Amazon RDS 包含用于自动执行数据库备份的功能。Amazon RDS 可创建数据库实例的存储卷快照，并备份整个数据库实例而不只是备份单独的数据库。

Amazon RDS 提供两种不同的方法来备份和还原数据库实例：

- **自动备份**可实现数据库实例的时间点恢复。默认情况下，在创建新的数据库实例时，将启用自动备份。在您创建数据库实例时，Amazon RDS 将在您定义的时段内对数据执行完整的每日备份。可以为自动备份配置保留期，最长为 35 天。Amazon RDS 利用这些定期数据备份以及事务日志，使您能够将数据库实例还原到保留期内的任一时间点，这个保留期最长为 LatestRestorableTime 指定的时间（通常为最后 5 分钟）。要查找数据库实例的最新可还原时间，可使用 DescribeDBInstances API 调用或查看 [AWS 管理控制台](#) [AA1] 中该数据库的 **Description** 选项卡。

启动时间点恢复时，会将事务日志应用到最合适的每日备份，以便将数据库实例还原到您要求的时间。

- **数据库快照**是用户启动的备份，可让您按所需的频率备份数据库实例的已知状态，然后您可以随时还原到该状态。您可以使用 [AWS 管理控制台](#) 或 CreateDBSnapshot API 调用来创建数据库快照。这些快照的保留期没有限制。它们将一直保存，直至您使用控制台或 DeleteDBSnapshot API 调用来明确将其删除。

在将数据库还原到某个时间点或者从数据库快照还原数据库时，将创建具有新终端节点的新数据库实例。这样一来，您便可以从特定数据库快照或时间点创建多个数据库实例。

您可以使用 [AWS 管理控制台](#) 或 DeleteDBInstance 调用来删除旧的数据库实例。

## 使用 AMI 备份 EC2 实例

AWS 将系统映像存储在称为 Amazon 系统映像 (AMI) 的位置。这些映像包含启动实例所需的根卷的模板。您可以使用 AWS 管理控制台或 `aws ec2 create-image` CLI 命令，将根卷备份为 AMI。

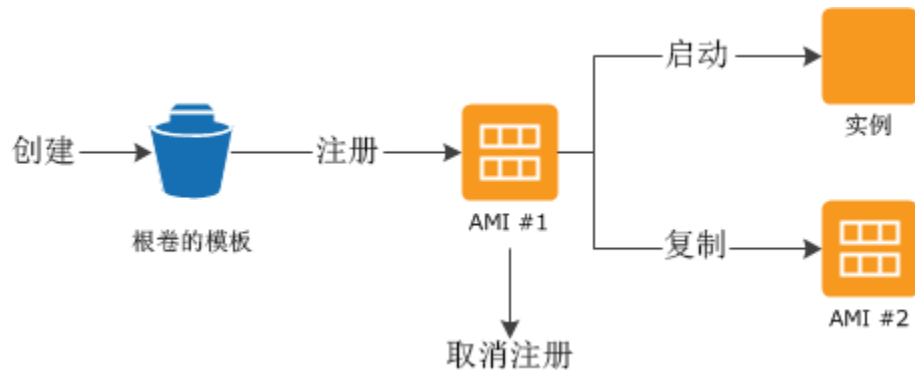
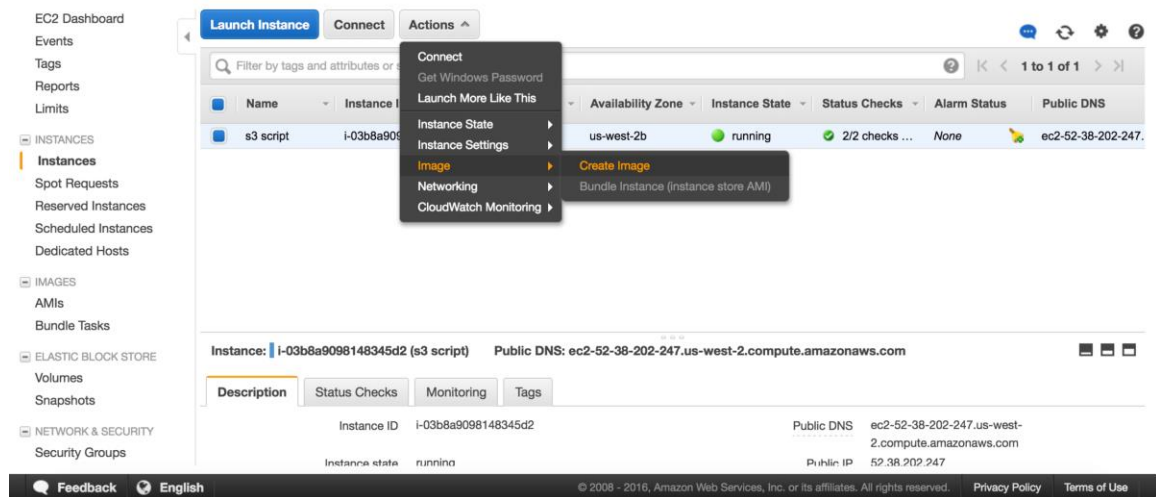


图 3: 使用 AMI 来备份和启动实例

在您注册 AMI 时，它会使用 Amazon EBS 快照存储在您的账户中。这些快照位于 Amazon S3 中，能够长时间保存。



[PAL2]

图 4: 使用 EC2 控制台创建系统映像

在创建 Amazon EC2 实例的 AMI 后，您可以使用此 AMI 重新创建实例或启动实例的多个副本。您还可以将 AMI 从一个区域复制到另一个区域，以便用于应用程序迁移或灾难恢复。

## 本地到 AWS 基础设施

此方案是指在云中没有任何组件的工作负载环境。所有资源（包括 Web 服务器、应用程序服务器、监控服务器、数据库和 Active Directory 等）都位于客户数据中心，或者通过主机托管的方式托管。

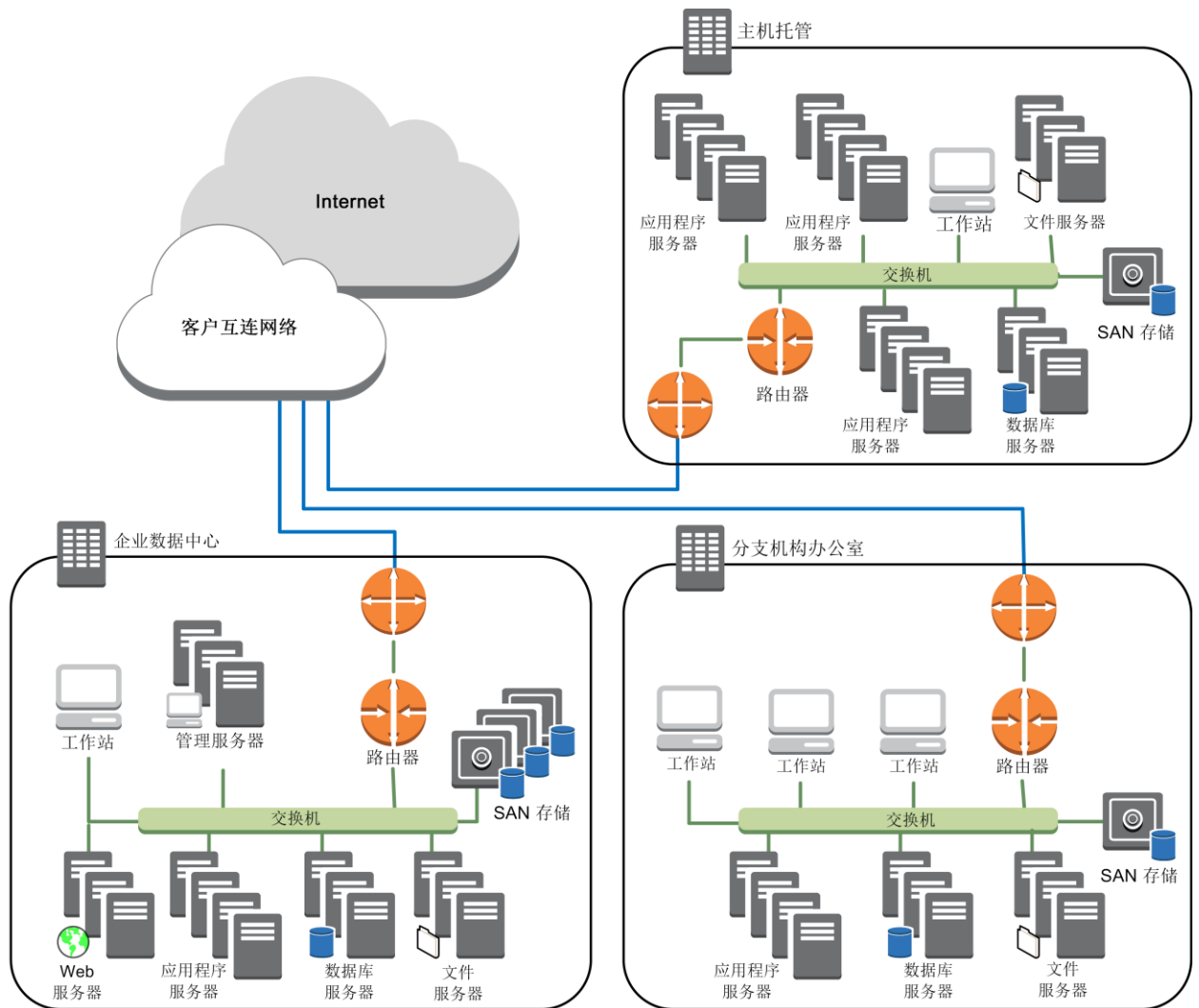


图 5: 本地环境

通过在此方案中使用 AWS 存储服务，您能够专注于备份和存档任务。在完成备份任务时，您无需担心存储扩展或基础设施容量。

Amazon S3 和 Amazon Glacier 提供原生 API，并且可通过 Internet 使用。这可使备份软件供应商直接将其应用程序与 AWS 存储解决方案集成，如下图所示。

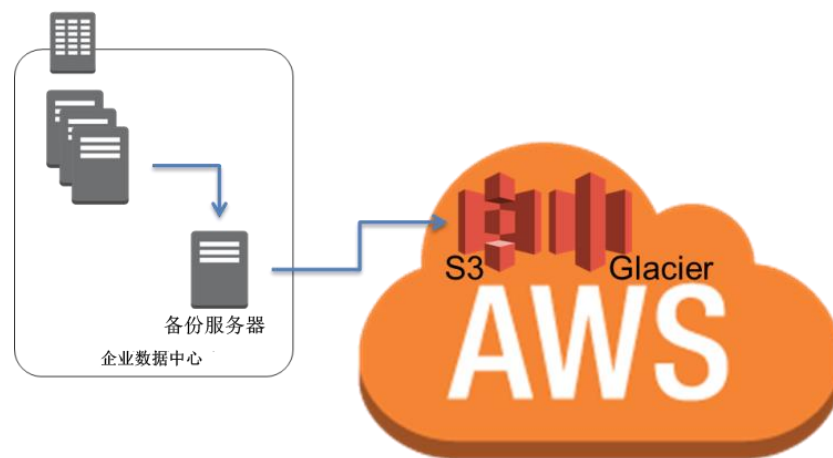


图 6: 针对 Amazon S3 或 Amazon Glacier 的备份连接器

在此方案中，备份和存档软件直接通过 API 连接 AWS。由于备份软件能够识别 AWS，因此它将本地服务器中的数据直接备份到 Amazon S3 或 Amazon Glacier。

如果现有备份软件并不原生支持 AWS 云，则可使用 AWS Storage Gateway 产品。[AWS Storage Gateway](http://aws.amazon.com/storagegateway/)<sup>13</sup> 是一个虚拟设备，它在您的数据中心与 AWS 存储基础设施之间提供无缝且安全的集成。利用此服务，您可以安全地将数据存储到 AWS 云，即可享用可扩展且经济实惠的存储。Storage Gateway 支持适用于现有应用程序的行业标准存储协议，并将所有加密的数据安全地存储在 Amazon S3 或 Amazon Glacier 中。

<sup>13</sup> <http://aws.amazon.com/storagegateway/>

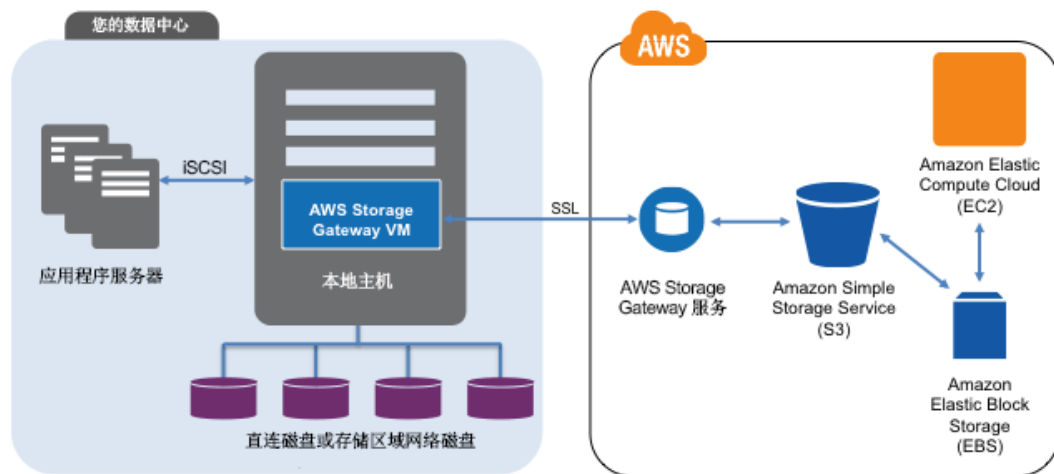


图 7：将本地存储连接到 AWS 存储

AWS Storage Gateway 支持以下配置：

- **卷网关：**卷网关提供了基于云的存储卷，可以从本地应用程序服务器将该存储卷作为 Internet 小型计算机系统接口 (iSCSI) 设备安装。该网关支持以下卷配置：
  - **网关缓存卷：**您可以将主数据存储在 Amazon S3 中，并将您经常访问的数据保留在本地。网关缓存卷可大大节省主存储成本，尽可能地减少扩展本地存储的需求，并保留对经常访问的数据的低延迟访问。
  - **网关存储卷：**如果需要对整个数据集进行低延迟访问，则可以配置本地数据网关以便在本地存储主数据，然后以异步方式将此数据的时间点快照备份到 Amazon S3。网关存储卷提供了持久且价格低廉的场外备份，您可以在本地恢复或从 Amazon EC2 恢复这些备份。
- **网关 — 虚拟磁带库（网关 VTL）：**利用网关 VTL，您可以获得无限制的虚拟磁带集合。每个虚拟磁带可存储在由 Amazon S3 支持的虚拟磁带库中或由 Amazon Glacier 支持的虚拟磁带架中。虚拟磁带库提供了一个行业标准 iSCSI 接口，您的备份应用程序可通过此接口来联机访问虚拟磁带。当不再需要立即访问或频繁访问虚拟磁带中保存的数据时，您可以使用备份应用程序将虚拟磁带从其虚拟磁带库移至虚拟磁带架，从而进一步降低存储成本。

这些网关可充当提供标准 iSCSI 设备的即插即用设备，可集成到备份或存档框架中。可以将 iSCSI 磁盘设备用作备份软件或网关 VTL 的存储池，来将基于磁带的备份或存档直接卸载到 Amazon S3 或 Amazon Glacier。

通过这种方法，您的备份和存档将自动在场外进行（以提供合规性）并存储在持久性介质中，从而消除场外磁带管理带来的复杂性和安全风险。

## 混合环境

就这一点讨论的两种基础设施部署（云原生和本地）可以结合成为一个混合方案。在这个方案中，工作负载环境具有本地组件和 AWS 基础设施组件。资源（包括 Web 服务器、应用程序服务器、监控服务器、数据库和 Active Directory 等）位于客户数据中心内或 AWS 中。AWS 云中运行的应用程序是与本地运行的应用程序相连的。

这正日益成为一种面向企业工作负载的常见方案。许多企业拥有自己的数据中心，并利用 AWS 来扩容。通常，这些客户数据中心通过高容量网络链路来连接到 AWS 网络。例如，利用 [AWS Direct Connect](#)<sup>14</sup>，您可以建立从本地到 AWS 的专用私有连接。<sup>[PAL3]</sup>这可在将数据上传到云时提供带宽和一致延迟，从而实现数据保护以及面向混合工作负载的一致性能和延迟。

---

<sup>14</sup> <http://aws.amazon.com/directconnect/>

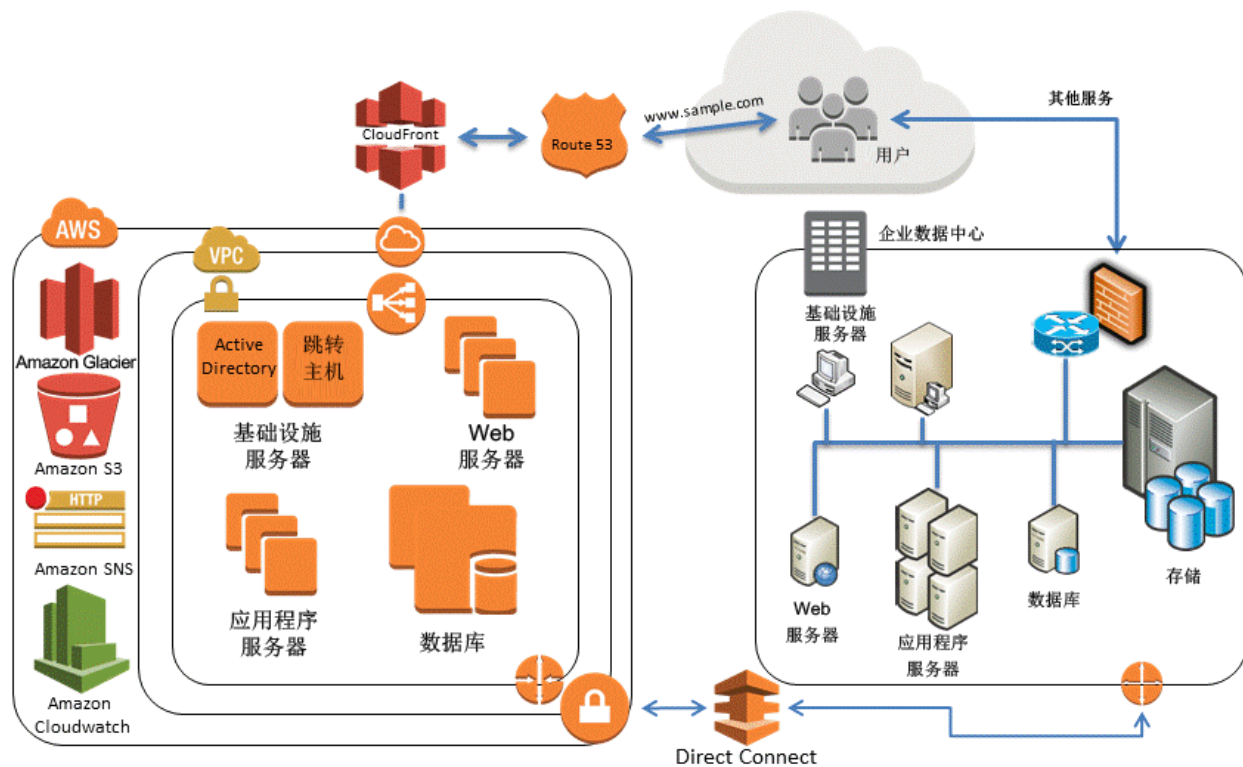


图 8：混合基础设施方案

设计精良的数据保护解决方案通常组合使用云原生和本地解决方案中介绍的方法。

## 将基于 AWS 的应用程序备份到数据中心

如果您已具有可备份本地服务器数据的框架，则可通过 VPN 连接或 AWS Direct Connect 轻松将它扩展到 AWS 资源。您可以在 Amazon EC2 实例上安装备份代理，并依据数据保护策略对这些实例进行备份。

## 将备份管理迁移到云以实现可用性

根据您的备份体系结构，对于所要保护的服务，您可能拥有一个主备份服务器以及一个或多个本地介质或存储服务器。在这种情况下，您可能希望将主备份服务器移到 Amazon EC2 实例，以防出现本地灾难，并且实现高度可用的备份基础设施。

要管理备份数据流，您可能还需要在 Amazon EC2 实例上创建一个或多个介质服务器。在 Amazon EC2 实例上创建介质服务器后，将节省您的 Internet 传输成本，并在备份到 S3 或 Amazon Glacier 时提高总体备份和恢复性能。

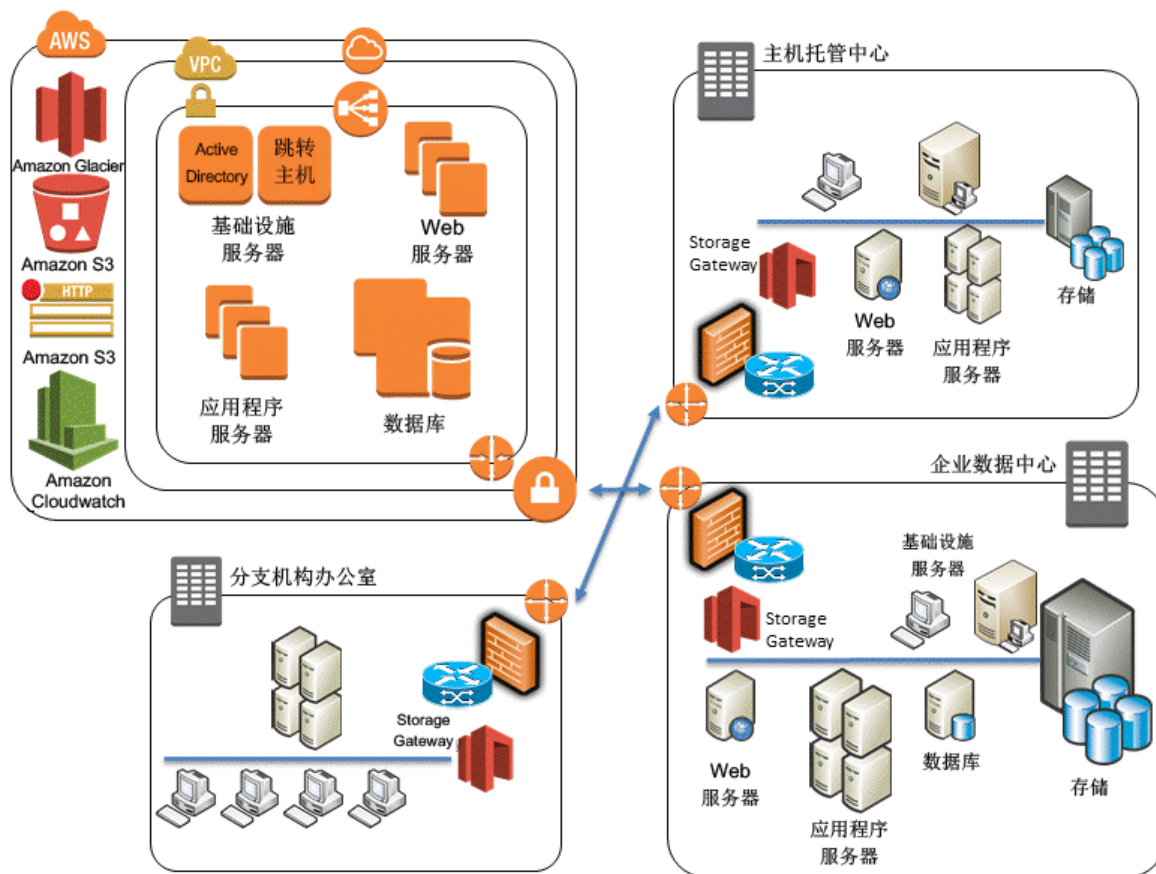


图 9：在混合方案中使用网关

## 混合方案示例

假定您正在管理一个用于备份 Amazon EC2 实例、独立服务器、虚拟机和数据库的环境。这个环境具有 1,000 个服务器，并且您需要备份操作系统、文件数据、虚拟机映像和数据库。需要备份 20 个数据库（MySQL、Microsoft SQL Server 和 Oracle 的组合）。

您的备份软件具有可备份操作系统、虚拟机映像、数据卷、SQL Server 数据库和 Oracle 数据库（使用 RMAN）的代理。对于备份软件不具有其代理的应用程序（如 MySQL），您可能使用 `mysqldump` 客户端实用程序来创建数据库转储文件并将其保存到磁盘，以便标准备份代理能够保护数据。

为了保护此环境，您的第三方备份软件最有可能具有一个全局目录服务器或主服务器（可控制备份、存档和还原活动）以及多个介质服务器（这些服务器连接到基于磁盘的存储、线性磁带开放 (LTO) 磁带驱动器和 AWS 存储服务）。

通过 AWS 存储服务扩展备份解决方案的最简单方式是：利用备份供应商提供的对 Amazon S3 或 Amazon Glacier 的支持。我们建议您与供应商协作，了解其集成和连接器选项。有关与 AWS 开展合作的备份软件供应商的列表，请参阅我们的[合作伙伴目录](#)<sup>15</sup>。

如果您的现有备份软件本身不支持将云存储用于备份或存档，则可在备份软件与 Amazon S3 或 Amazon Glacier 之间使用存储网关设备（如网桥）。

有许多第三方网关解决方案可供使用。您也可以使用 AWS Storage Gateway 虚拟设备来解决这个问题，因为它使用常见技术，例如基于 iSCSI 的卷和虚拟磁带库 (VTL)。此配置需要使用受支持的管理程序 (VMware 或 Microsoft Hyper-V) 和本地存储来托管设备。

## 使用 AWS 对数据进行存档

当您需要出于合规性或企业原因保留数据时，可对数据进行存档。与备份不同，存档将保留所有数据副本直至保留策略过期；而执行备份通常是为了将生产数据副本保留一段较短的时间，以便在数据损坏或数据丢失的情况下进行恢复。

设计优异的存档应该符合以下标准：

- 可确保长期完整性的数据持久性
- 数据安全
- 易于恢复
- 成本低廉

另外，有的法规或合规性要求使用不可变的数据存储。

Amazon Glacier 提供了低成本的存档、静态数据的本机加密、99.999999999%（11 个 9）的持久性以及无限容量。

---

<sup>15</sup> <http://www.aws-partner-directory.com/PartnerDirectory/PartnerSearch?type=ISV>

对于需要快速检索数据的使用情形来说，“Amazon S3 标准 - 不经常访问”是非常不错的选择。而对于不常访问数据且可接受数小时检索时间的使用情形来说，Amazon Glacier 是良好的选择。

可通过 S3 中的生命周期规则或 Amazon Glacier API 将对象分层到 Amazon Glacier 中。利用 Amazon Glacier 文件库锁定功能，您可以轻松利用文件库锁定策略对单独的 Amazon Glacier 文件库进行部署并实施合规性控制。您可以在一个文件库锁定策略中指定类似“一次写入，多次读取”(WORM) 这样的控制，并且可以锁定该策略以防将来被人编辑。有关更多信息，请参阅 [Amazon Glacier](#)。

## 在 AWS 中保护备份数据

数据安全是一个常见问题。AWS 非常重视安全。这是我们推出的每项服务的基础。存储服务（如 Amazon S3）提供了用于对静态数据和传输中的数据进行访问控制和加密的强大功能。所有 Amazon S3 和 Amazon Glacier API 终端节点都支持对传输中的数据进行 SSL 加密。默认情况下，Amazon Glacier 会对所有静态数据进行加密。利用 Amazon S3，客户可通过以下方式对静态对象选择服务器端加密：让 AWS 管理加密密钥，在上传对象时提供自己的密钥或对加密密钥使用 AWS Key Management Service (AWS KMS)<sup>16</sup>集成。或者，客户可以先对自己的数据进行加密，然后再将数据上传到 AWS。有关更多信息，请参阅 [Amazon Web Services: 安全流程概述](#)。

## 结论

Gartner 已将 AWS 评定为公共云存储服务领域的领导者<sup>17</sup>。AWS 已做好充分准备，能够帮助组织将其工作负载转移到基于云的平台，即下一代备份。AWS 提供了经济高效的、可扩展的解决方案，来帮助组织平衡他们在备份和存档方面的要求。这些服务能够与您当前使用的技术很好地集成。

---

<sup>16</sup> <http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingKMSEncryption.html>

<sup>17</sup> <http://www.gartner.com/technology/reprints.do?id=1-1WWKTQ3&ct=140709&st=sb>

## 撰稿人

本文的撰稿人包括：

- Pawan Agnihotri, Amazon Web Services 解决方案架构师
- Lee Kear, Amazon Web Services 解决方案架构师
- Peter Levett, Amazon Web Services 解决方案架构师

## 文档修订

更新时间：2016 年 5 月