



标记监控、报告和成本优化时使用最佳实践

AWS 规范性指导



AWS 规范性指导: 标记监控、报告和成本优化时使用最佳实践

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商标和商业外观不得用于任何非 Amazon 的商品或服务，也不得以任何可能引起客户混淆、贬低或诋毁 Amazon 的方式使用。所有非 Amazon 拥有的其他商标均为各自所有者的财产，这些所有者可能附属于 Amazon、与 Amazon 有关联或由 Amazon 赞助，也可能不是如此。

Table of Contents

简介 1

概述 1

先决条件 3

目标 4

实现更详细的成本跟踪 4

生成更准确的报告 4

加快清理速度 5

高效管理权限 5

最佳实践 6

标记本体 6

标签治理 7

要避免的实践 8

不一致的标记 8

标签中的错误和敏感数据 9

常见问题解答 11

是否能在部署任何基础设施时强制创建标签？ 11

标记基础设施是否足够？ 11

有哪些后续步骤？ 11

..... 12

文档历史记录 13

..... xiv

标记监控、报告和成本优化时使用最佳实践

Apoorva Patrikar , Amazon Web Services (AWS)

2023 年 10 月 ([文档历史记录](#))

标记是指以键值对的形式将元数据分配给 Amazon Web Services (AWS) 上部署的资源的过程。标记对于管理和识别资源、监控和内部管理非常重要。标签提供有关文件、代码和基础设施的其他上下文，可以帮助您构建用于监控和流程优化的控制面板。

本指南讨论了标记应用程序所有组件的最佳实践和本体。该示例侧重于数据管线。但是，您可以将相同的概念扩展到任何领域或应用程序。

概述

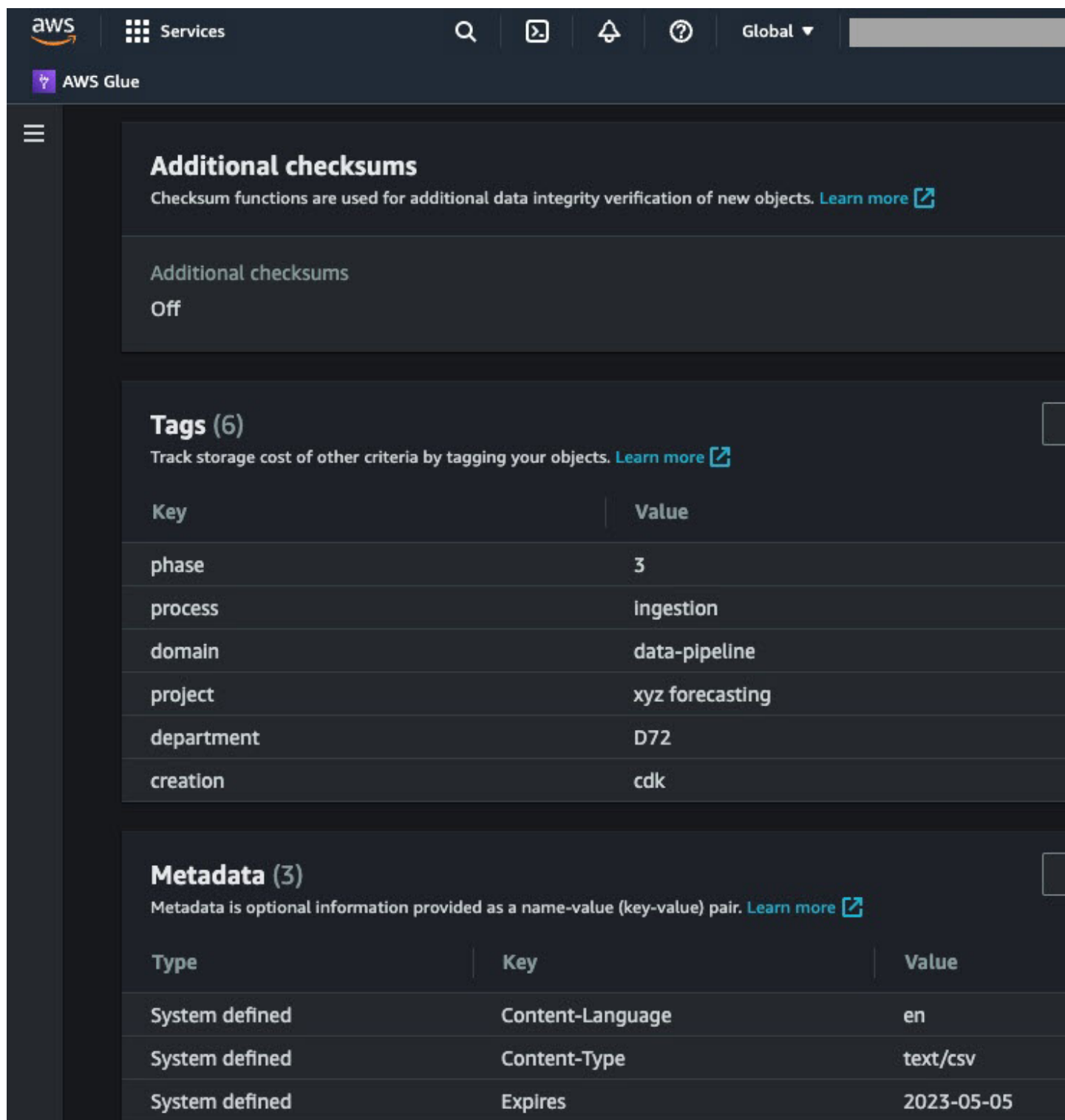
标签是有关数据文件或资源的数据。它们提供其他业务或技术信息。例如，数据的标签可以提供数据归属、领域、部门和类型等信息。资源的标签提供有关整个过程的信息。例如，标签可提供以下信息：

- 项目名称
- 运行类型
- 环境，例如 dev、int 或 prod
- 运营的数据层，例如原始、可信或优化

标签是键值对，可用于报告当前状态或了解每个粒度级别的成本、性能和成功率。标签可以按不同的属性进行分类，包括以下属性：

- 用途
- 所有者
- 环境
- Project
- 阶段

下图显示在 Amazon Simple Storage Service (Amazon S3) 上实施的对象标签。



The screenshot displays the AWS Glue console interface. At the top, the AWS logo and 'Services' menu are visible. The main header shows 'AWS Glue'. A sidebar on the left contains a hamburger menu icon. The main content area is divided into three sections:

- Additional checksums**: A section with a description and a 'Learn more' link. Below it, a toggle switch for 'Additional checksums' is set to 'Off'.
- Tags (6)**: A section with a description and a 'Learn more' link. Below it is a table of tags.
- Metadata (3)**: A section with a description and a 'Learn more' link. Below it is a table of metadata.

Key	Value
phase	3
process	Ingestion
domain	data-pipeline
project	xyz forecasting
department	D72
creation	cdk

Type	Key	Value
System defined	Content-Language	en
System defined	Content-Type	text/csv
System defined	Expires	2023-05-05

本指南介绍在将标签应用到所有组件（包括以下各项）之后，标记如何帮助您了解应用程序的指标：

- 输入文件
- 数据库

- 处理作业
- 输出文件

先决条件

在实施标记时，请注意以下要求：

- 将标签用于报告目的时，必须激活标签。有关更多信息和说明，请参阅 [AWS Billing 文档](#)。将标签用于其他目的（例如清理、自动化或检索属于特定类别服务的完整列表）时，为所有涉及组件分配标签可能就足够了。
- 在指定推荐或强制标签之前，请查看 [AWS 文档](#)，了解有关限制和要求的消息。

目标

使用标记实现以下业务成果。

实现更详细的成本跟踪

在成本管理控制面板中，您可以根据资源类型、AWS 账户、AWS 区域和其他因素对成本进行细分，但不包括业务背景。例如，控制面板不会告知您哪个部门欠了多少钱，也不会告知您数据管线的成本是否高于机器学习管线。这些问题的答案只能通过以标签形式添加业务背景来获得。然后，您可以使用标签筛选成本管理控制面板。

标记跟踪费用

要将标签用于计费，[请选择成本分配标签](#)，该标签将在 24 小时内激活。然后，您可以选择这些标签作为图表的筛选条件，也可以下载报告。有关更多信息以及示例报告，请参阅 [AWS Billing 文档](#)。

生成更准确的报告

高层管理人员和审计组织通常需要报告使用情况、成本和资源。这些报告需要深入了解技术应用程序中的每个组件如何映射到业务使用案例或部门。标记可以提供这些知识。

实施标记后，您可以使用 AWS Cost Explorer 中的筛选条件来区分不同的值。这为您提供了更高效、更准确的报告或监控选项。

- 部门或计划：部门名称（例如，D72、IT 或运营）

如果是整合账单或多个部门使用一个 AWS 账户，则可以筛选该值以准确找出每个部门的成本。

- 项目：项目名称（例如，销售预测、招聘预测）

如果是整合账单或多个项目使用一个 AWS 账户，则可以筛选该值以准确找出每个项目的成本。

- 阶段：阶段或状态名称（例如，开发、测试）

项目包含不同的阶段，例如开发、测试、错误修复和报告。通过添加此标签，您可以筛选项目每个阶段的成本或性能。这有助于更准确地规划项目的下一次迭代。

- 流程：流程名称（例如，数据摄取、数据清理/处理、机器学习、建模）

使用此信息，您可以区分每个流程的成本、时间或性能。例如，如果您发现数据清理的运行时间或成本非常高，则可以优化数据来源以降低成本并提升性能。

加快清理速度

需要每年进行一次清理，以维持成本，并确保移除不再需要用于测试、概念验证（POC）和开发的资源。如果没有正确的标签，就很难区分是生产环境中使用的资源还是最新开发版本中使用的资源。

在应用程序中标记所有生产组件后，您可以根据组件是否有标签来决定删除还是保留该组件。例如，一个长时间未运行且零标签的 AWS Lambda 函数可能是团队成员出于开发或快速测试而创建的。如果标记使用基础设施即代码实现完全自动化，则生产环境中的所有组件都将有一个与之关联的标签，例如 `key:environment` 和 `value:prod`。

高效管理权限

如果不进行标记，则在应用程序中为 AWS 服务分配 AWS Identity and Access Management（IAM）角色需要为每个服务和实例单独创建角色。这可能非常耗时，因为整个应用程序中可能有多项服务需要相同的 IAM 操作。

通过基于 [AWS Resource Groups](#)（或基于 AWS CloudFormation 堆栈）创建标签，可以更高效地分配权限。您可以将 IAM 权限分配给更广泛的组以允许或限制访问权限。在允许对一组资源执行特定操作的 IAM 策略的以下摘录中，您将简化自己的 IAM 策略。

```
"Action": [
    "resource-groups:List*",
    "resource-groups:action2",
    "resource-explorer:action3"
]
```

最佳实践

使用标签时，了解当前或潜在的公司和使用案例结构非常重要。使用这些信息，您可以选择合适的标签。例如，如果您正在为售前部门构建数据湖，并且知道有计划扩展数据湖以包括来自售后部门的数据，那么使用标签 `department` 将有助于分别确定每个部门的成本和绩效。可以更准确地识别代码或数据的规划、成本分配和优化。如果没有 `department` 标签，售前数据需要 15 分钟进行数据建模，而售后数据需要 45 分钟，则开发人员必须花更多的时间进行根本原因分析。有了 `department` 标签，开发人员就可以准确了解该从哪里入手。

标记本体

业务和技术在识别合适的标签方面共同发挥着重要作用。从业务角度来看，公司和项目将始终遵循一定的结构。例如，在欧洲、中东和非洲区域，HR 部门可能有一个预测招聘需求的项目。在这种情况下，包括来自现有结构的元数据对于报告、监控、清理和推出非常重要。同时，技术部门明白，该项目将需要以下内容：

- 通过由数据摄取、清理和处理组成的数据管线收集数据的各个阶段
- 用于进行预测数据建模的 ML 团队
- 用于代码编排的 DevOps 管线，覆盖开发、测试和生产环境

所有斜体关键字均为业务和技术组结构，这些结构务必与应用程序的组件关联。以下是典型的标记本体示例。使用该示例，下表显示了标签对应的键值对。

键	值
<code>department</code>	<code>human resources</code>
<code>region</code>	<code>EMEA</code>
<code>project</code>	<code>hiring forecast</code>
<code>phase</code>	<code>3</code>
<code>process</code>	<code>data ingestion data cleaning、processing 、modeling、或 sales forecasting</code>

domain	machine learning 或者 data pipeline
creation	cdk x framework 、 ingest pipeline、 或 manual - empty
status	development testing、 production access、 reporting 、 或 onboarding

标签治理

设置治理机制有助于使标记在所有 AWS 资源上保持一致且可编程：

- 被动治理是指查找未正确标记的资源。您可以使用[资源组标记 API](#)、[AWS Config 规则](#)和自定义脚本等工具。
- 主动治理是指不允许用户创建任何未标记的资源。您可以使用 [AWS CloudFormation](#)、[AWS Service Catalog](#)、AWS Organizations 中的[标签策略](#)或 IAM 资源级权限等工具。

要避免的标记实践

虽然在 AWS 上标记对象或基础设施时需要实施一些实践，但也需要避免一些实践。

不一致的标记

如[目标](#)一节所述，如果没有标记，就无法实现高水平的自动化、清理或监控。同样，如果标签不完整或不一致，自动化或监控所需的信息就不完整，从而导致结果不可靠。

想象一个场景：您使用标记策略来计算所有项目的总成本。该策略从概念验证阶段 (PoC) 开始，到生产阶段结束。考虑以下场景，将标签应用到销售预测项目 P1、D1 和 Pr1 示例以及售后维护项目 P2、D2 和 Pr2 示例的数据和资源。

销售预测

示例 P1：PoC 项目（缺少领域和时间戳）。

```
env: "poc"
project: "sales forecasting"
```

示例 D1：开发阶段（缺少领域）。

```
env: "dev"
project: "sales forecasting"
timestamp: 20210505T12:34:55
```

示例 Pr1：生产阶段（所有值都存在）。

```
env: "prod"
project: "sales forecasting"
domain: "machine learning"
timestamp: 20210505T12:34:55
```

对于销售预测项目：

- 示例 P1 没有提及对象来自哪个领域或时间戳。
- 示例 D1 也没有提及项目的领域。
- 示例 Pr1 包含所有所需的数据。

由于未定义领域，示例 P1 和 D1 将导致规划的报告或估算不正确。

售后维护

示例 P2：PoC 项目（缺少所有标签）。

示例 D2：开发阶段（缺少项目）。

```
env: "dev"
domain: "machine learning"
timestamp: 20210505T12:34:55
```

示例 Pr2：生产阶段（所有值都存在）。

```
env: "prod"
project: "post sales maintenance"
domain: "machine learning"
timestamp: 20210505T12:34:55
```

对于售后维护项目：

- 示例 P2 没有任何信息，因此无法对其进行跟踪。
- 示例 D2 没有提及项目名称，因此无法对其进行跟踪。
- 示例 Pr2 包含所有所需的数据。

示例 P2 和 D2 将由于缺少标签或标签不一致而导致报告不正确、计划不足或报告不足。

因此，一致地实施标签策略非常重要。

标签中的错误和敏感数据

如果使用不正确、敏感或私密的信息，则标记可能适得其反。错误的标签可能会产生误导性的结果。使用包含个人身份信息（PII）等敏感数据的标签，可能会使客户和员工面临安全风险。

标签中含有错误信息

想象一个场景：您使用标记策略来计算每个领域或每个部门的总成本。您刚刚完成数据摄取阶段，正在向机器学习迈进。以下示例包括从项目上一阶段复制的自定义标签。

```
env: "development"
```

```
project: "sales prediction"
domain: "data ingestion"
timestamp: 20210505T12:34:55
```

领域被错误地标记为上一个项目阶段的 data ingestion，而不是正确的领域（即 machine learning）。现在，data ingestion 领域的报告将显示更高的成本、时间范围和资源分配。machine learning 领域的这些报告将显示较低值。这将导致规划、预算分配和截止日期估算不正确。

拥有正确的标签对于职能系统至关重要。

标签中含有敏感信息

AWS 提供多种用于识别对象中 PII 的工具。这些工具包括 [Amazon Macie](#) 和 [AWS Glue 敏感数据检测](#)，用于找出可用于识别个人的数据。但是，切勿在标签中使用 PII 或敏感数据。

考虑以下 Amazon S3 中已脱敏或匿名化 PII 的文件为例。

```
{
  firstName: "67A1790DCA55B8803AD024EE28F616A2",
  lastName: "DRG54654DFHJGDYYRD",
  age: 21,
  city : "Frankfurt",
  probability_of_purchase: 48.858093,
  veggieName: "broccoli",
  creditcard: false
}
```

您可以看到客户的名字和姓氏已被哈希处理。但是，在本例中，该记录具有以下自定义标签。

```
owner: "Company XYZ"
about: "John Doe"
contact: "johnthegreat@email.com"
timestamp: 20210505T12:34:55
```

在这种情况下，尽管文件本身不包含 PII，但标签确实包含敏感信息。这会增加信息泄露的可能性，因为当您共享或传输文件或对象时，也会共享或传输其元数据。这也适用于数据库、表、作业和函数等其他 AWS 资源。

因此，避免在标签中使用私有信息极其重要。同样的概念也适用于关键或非公开信息。

常见问题解答

以下是有关标记的常见问题。AWS

是否能在部署任何基础设施时强制创建标签？

是的，您可以使用[服务控制策略](#)强制创建标签。

标记基础设施是否足够？

通常，仅标记基础设施是不够的。要在控制面板中获得准确的数字，必须标记在 AWS 上存储或处理的所有资源 and 数据文件。

有哪些后续步骤？

重新评估您的项目。将标签应用到所有数据和基础设施，然后设置控制面板。即使目前不需要此类控制面板，现在标记资源也将在您需要了解分项成本时大大加快控制面板的设置速度。

资源

参考

- [Implement AWS resource tagging strategy using AWS Tag Policies and service control policies](#) (博客文章)
- [用户定义的成本分配标签](#)
- [对AWS资源加标签](#)

工具

- [服务控制策略 \(SCP\)](#)

AWS 合作伙伴

- [AWS Data and Analytics Competency Partners](#)

文档历史记录

下表介绍了本指南的一些重要更改。如果您希望收到有关未来更新的通知，可以订阅 [RSS 源](#)。

变更	说明	日期
初次发布	—	2023 年 10 月 12 日

本文属于机器翻译版本。若本译文内容与英语原文存在差异，则一律以英文原文为准。