数据安全中心

常见问题

文档版本	20
发布日期	2025-07-07

华为云计算技术有限公司

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。

华为云计算技术有限公司

地址： 贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址： https://www.huaweicloud.com/

目 录

1 产品咨询类

1.1 数据安全中心是否会保存您的数据和文件?

数据安全中心（DSC）不会保存您的数据和文件，在您授权访问数据源后，DSC会对数据进行识别、脱敏、或添加水印等操作。

数据识别的结果将展示在DSC的控制台，如何查看敏感数据识别结果，请参见**敏感数据识别结果**。

1.2 DSC 支持识别的非结构化文件类型?

数据安全中心（DSC）支持解析的非结构化文件类型如**表1-1**、**表1-2**和**表1-3**。

表 1-1 文本文档代码类

序号	文件类型	序号	文件类型
1	Access数据库文件	74	Pdf文档
2	Arff文件	75	Perl源代码
3	Asp文件	76	Pgp文件
4	Atom文件	77	Php源代码
5	Bat文件	78	Pkcs7数字证书文件
6	Bcpl源代码	79	Plist文件
7	Bib文件	80	Postgres数据库文件
8	C#源代码	81	Postscript文档
9	C/C+源代码	82	Powerpoint文档
10	Cad Sldworks文件	83	Properties文件
11	Cad文档	84	Publisher文件

序号	文件类型	序号	文件类型
12	Cbor文件	85	Python源代码
13	Cfg文件	86	Quattro-Pro电子表格
14	Chm文件	87	Redis数据库文件
15	Com可执行文件	88	Rss文件
16	Css文件	89	Rtf文档
17	Datax配置文件	90	Ruby源代码
18	Dbf文件	91	R源代码
19	Dif文件	92	Sas7Bdat文件
20	Dita文件	93	Sas文件
21	Djvu文档	94	Scala源代码
22	Dos可执行文件	95	Shell脚本
23	D源代码	96	Sqlite3数据库文件
24	Elf可执行文件	97	Sqlserver数据库文件
25	Epub电子书	98	Sql源代码
26	Excel文档	99	Ssh公钥
27	Fdf文档	100	Ssh配置文件
28	Fictionbook Xml文件	101	Ssh私钥
29	Ftp会话文件	102	Staroffice文档
30	Gnuccash财务xml文件	103	Swift源代码
31	Go源代码	104	Tab文件
32	Groovy源代码	105	Tcl源代码
33	Hdr文件	106	Text文件
34	Hocon文件	107	Tff文件
35	Html文件	108	Tnef文件
36	Htm文件	109	Tomcat Application配置文件
37	Hwp文件	110	Tomcat Users配置文件
38	Ibooks文件	111	Tomcat配置文件
39	Iis配置文件	112	Toml文件
40	Ini 文件	113	Tsd文件

序号	文件类型	序号	文件类型
41	Isa-Tab文件	114	Tsv文件
42	Iwork文档	115	Vcs文件
43	Java Jce Keystore文件	116	Visio文档
44	Java Keystore文件	117	Visualbasic源代码
45	Javascript源代码	118	Vrml虚拟现实建模语言代码
46	Java源代码	119	Webarchive文件
47	Json文件	120	Weblogic配置文件
48	Jsp源代码	121	Webvtt文件
49	Latex源代码	122	Windowsinf文件
50	Log日志文件	123	Windows帮助全文搜索引
51	Lua源代码	124	Windows预编译文件
52	Mariadb数据库文件	125	Wordperfect文档
53	Markdown文档	126	Word文档
54	Matlab源代码	127	Wpd文档
55	Mbox文件	128	Wps文档
56	Mhtml文件	129	Xdp文件
57	Microsoft Reader文档	130	Xfdf文件
58	Mongodb数据库文件	131	Xhtml文件
59	Mrs配置文件	132	Xlf文件
60	Msworks文档	133	Xliff文件
61	Mysql数据库文件	134	Xlr文件
62	Netcdf文件	135	Xlz文件
63	Objective-C源代码	136	Xml Sitemap文件
64	Obs配置文件	137	Xml文件
65	Office文档	138	Xmp文件
66	Onenote文件	139	Xps文档
67	Opendocument文件	140	Xpt文件
68	Openvpn配置文件	141	Yaml文件
69	Oracle数据库文件	142	常见数字证书文件

序号	文件类型	序号	文件类型
70	Outlook文件	143	空文件
71	Pascal源代码	144	配置文件windows Initialization
72	Pbm文件	145	其他普通未加密文本文件
73	Pcx文件	146	邮件文档

表 1-2 压缩和二进制类

序号	类型说明	序号	类型说明
1	7Zip文件	26	Lha压缩文件
2	Apk安卓程序	27	Lz4压缩文件
3	Arj文件	28	Lzma压缩文件
4	Ar文件	29	Mat文件
5	Bgp文件	30	Netcdf文件
6	Brotli压缩文件	31	Object文件
7	Bzip2压缩文件	32	Pack200压缩文件
8	Bzip压缩文件	33	Rar压缩文件
9	Cabinet压缩文件	34	Sharelib文件
10	Coredump文件	35	Snappy压缩文件
11	Cpio压缩文件	36	Tar压缩文件
12	Deflate64压缩文件	37	Tcpdump捕获文件
13	Dmg文件	38	Tika-Unix-Dump文件
14	Elf可执行文件	39	Unix压缩文件
15	Gdal文件	40	Xcompress压缩文件
16	Grb文件	41	Xlz压缩文件
17	Grib2文件	42	Xpi Firefox插件安装包
18	Grib文件	43	Xz压缩文件
19	Gzip文件	44	Zip压缩文件
20	Hdf文件	45	Zlib压缩文件
21	He5文件	46	Zstd压缩文件

序号	类型说明	序号	类型说明
22	Iso-19139地理信息文件	47	Zstd字典文件
23	Iso压缩文件	48	Z压缩文件
24	Jar文件	49	可执行文件
25	Java Class文件	50	普通压缩文件

表 1-3 图片类

序号	类型说明	序号	类型说明
1	BMP文件	4	JFIF文件
2	PNM文件	5	JPEG文件
3	PNG文件	6	TIFF文件

2 区域与可用区

2.1 什么是区域和可用区?

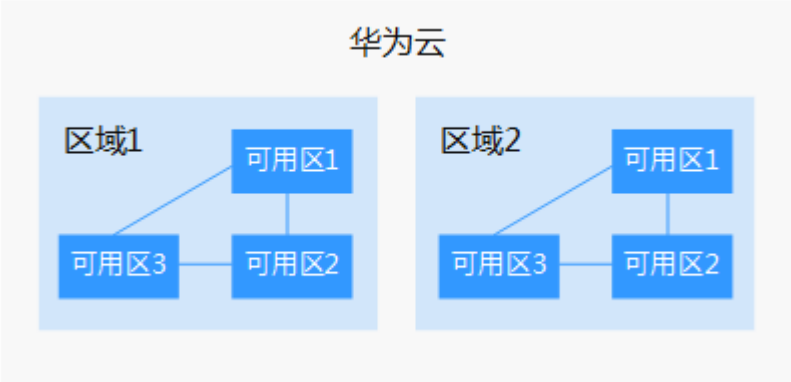
什么是区域、可用区?

我们用区域和可用区来描述数据中心的位置，您可以在特定的区域、可用区创建资源。

- 区域（Region）：从地理位置和网络时延维度划分，同一个Region内共享弹性计算、块存储、对象存储、VPC网络、弹性公网IP、镜像等公共服务。Region分为通用Region和专属Region，通用Region指面向公共租户提供通用云服务的Region；专属Region指只承载同一类业务或只面向特定租户提供业务服务的专用Region。
- 可用区（AZ，Availability Zone）：一个AZ是一个或多个物理数据中心的集合，有独立的风火水电，AZ内逻辑上再将计算、网络、存储等资源划分成多个集群。一个Region中的多个AZ间通过高速光纤相连，以满足用户跨AZ构建高可用性系统的需求。

图2-1阐明了区域和可用区之间的关系。

图 2-1 区域和可用区



目前，华为云已在全球多个地域开放云服务，您可以根据需求选择适合自己的区域和可用区。

如何选择区域?

选择区域时，您需要考虑以下几个因素：

- 地理位置

  一般情况下，建议就近选择靠近您或者您的目标用户的区域，这样可以减少网络时延，提高访问速度。

  - 在除中国大陆以外的亚太地区有业务的用户，可以选择“中国-香港”、“亚太-曼谷”或“亚太-新加坡”区域。
  - 在非洲地区有业务的用户，可以选择“非洲-约翰内斯堡”区域。
  - 在拉丁美洲地区有业务的用户，可以选择“拉美-圣地亚哥”区域。

- 资源的价格

  不同区域的资源价格可能有差异，请参见**华为云服务价格详情**。

如何选择可用区?

是否将资源放在同一可用区内，主要取决于您对容灾能力和网络时延的要求。

- 如果您的应用需要较高的容灾能力，建议您将资源部署在同一区域的不同可用区内。
- 如果您的应用要求实例之间的网络延时较低，则建议您将资源创建在同一可用区内。

区域和终端节点

当您通过API使用资源时，您必须指定其区域终端节点。有关华为云的区域和终端节点的更多信息，请参阅**地区和终端节点**。

2.2 数据安全中心可以跨区域使用吗?

数据安全中心不支持跨区域使用，即在本Region下购买的DSC版本，只能在该Region下使用。

3 资产授权类

3.1 开通云资源授权后，获得了授权资产服务的哪些权限?

开通云资源授权后，可以访问OBS桶、数据库、大数据以及资产地图，获得了授权资产服务的权限如表3-1所示。

表 3-1 对应授权项服务创建的委托

资产模块	服务策略	作用范围	备注
OBS	OBS Administrator	全局	用于配置OBS日志，获取OBS对象列表，下载OBS对象，用于获取OBS服务投递日志等
	EVS ReadOnlyAccess	区域	用于获取云硬盘列表
数据库	ECS ReadOnlyAccess	区域	用于获取自建数据库ECS列表
	RDS ReadOnlyAccess	区域	用于获取RDS数据库列表及数据库列表相关信息
	DWS ReadOnlyAccess	区域	用于获取DWS列表
	VPC FullAccess	区域	用于打通网络，VPC的端口创建，安全组规则创建等
	KMS CMKFullAccess	区域	用于使用KMS加密脱敏的场景

资产模块	服务策略	作用范围	备注
	GaussDB ReadOnlyAccess	区域	用于获取GaussDB列表
大数据	ECS ReadOnlyAccess	区域	用于获取自建大数据ECS列表
	CSS ReadOnlyAccess	区域	用于获取CSS数据集群列表及数据索引等相关信息
	DLI Service User	区域	用于获取DLI队列及数据库
	VPC FullAccess	区域	用于打通网络，VPC的端口创建，安全组规则创建等
	KMS CMKFullAccess	区域	用于使用KMS加密脱敏的场景
MRS	MRS CommonOperations	区域	用于集群查询、任务创建等
资产地图	Tenant Guest	区域	用于获取用户涉及数据存储处理等相关云服务的列表等
	OBS Administrator	全局	用于配置OBS日志，获取OBS对象列表，下载OBS对象等
	EVS ReadOnlyAccess	区域	用于云硬盘列表获取
	OBS Administrator	全局	用于OBS服务投递日志
LTS	LTS ReadOnlyAccess	区域	用于读取LTS日志组/日志流

3.2 如何排查数据库资产连通性失败?

数据库添加完成后，该数据库的“连通性”为“检查中”，此时，DSC会测试数据库的连通性，如果数据库的“连通性”为“失败”，请按照以下步骤进行排查：

步骤1 检查添加资产的IP、账号、密码、数据库名是否正确。

- 不正确，修改添加资产的IP、账号、密码、数据库名。
- 正确，执行2。

步骤2 检查您资产安全组的出方向是否全部放开。

- 没有全部放开，需要添加出方向规则，安全组的出方向全部放开后再编辑数据库重新添加，如果仍失败，执行**3**。
- 已全部放开，执行**3**。

步骤3 检查数据库对应IP子网的可用IP数是否为0。

由于DSC服务需要对数据库进行网络打通，至少需要一个可用IP数。如果数据库对应IP子网的可用IP数为0，则需要在对应数据库服务中添加可用IP。

----结束

4 数据识别和数据脱敏

4.1 DSC 能够识别哪些数据源对象?

DSC能通过内置规则和自定义规则从OBS、RDS、Elasticsearch、DWS、DLI的海量数据中分析并识别出敏感对象。

DSC支持的数据源如表4-1所示。

表 4-1 支持的数据源

数据源	具体的数据类型	扫描限制
RDS（关系型数据库）	MySQL、SqlServer、PostgreSQL类型。	采样扫描前500行数据。扫描指标QPS为300次/秒。
CSS（云搜索服务）	大数据资产	--
OBS（对象存储服务）	支持200+文件类型。	大于200MB以上的文件不会对其进行扫描；同时如果OBS桶的文件进行了加密，则无法对其扫描。
DWS（数据仓库服务）	--	--
ECS（弹性云服务器）	搭建的Mysql、TDSQL、SqlServer、PostgreSQL、Oracle数据库及ElasticSearch实例。	--
DLI（数据湖探索）	大数据资产	--
LTS（云日志服务）	日志流	--

4.2 DSC 的扫描时长和脱敏时长?

扫描时长

DSC服务扫描的时长将由您所扫描数据源的数据量、扫描规则数、扫描模式决定，表4-2中提供的扫描时长仅作参考。

表 4-2 扫描时长

数据源	数据量	扫描模式	扫描时长
RDS（关系型数据库）	1000张表	快速扫描	5分钟
CSS（云搜索服务）	1000Wdoc	快速扫描	15分钟
OBS（对象存储服务）	100M	快速扫描	1分钟
OBS（对象存储服务）	100M	全量扫描	15分钟

脱敏时长

DSC通过内置和自定义脱敏算法，实现对RDS、ES、MRS、Hive数据进行脱敏，一般情况下，脱敏时长如表4-3所示。

表 4-3 脱敏时长

数据源	数据量	脱敏时长
RDS（关系型数据库）	1000W行	40分钟
Elasticsearch实例	1000Wdoc	40分钟
MRS_HIVE	1000W行	40分钟

4.3 DSC 支持的内置识别规则有哪些?

数据安全中心根据行业敏感信息内置了包含敏感图片信息、个人敏感信息、企业敏感信息等七类规则，具体内置规则如表4-4所示。

表 4-4 内置规则

敏感数据分类	类型
个人敏感图片信息	● 驾照图片（中国内地） ● 银行卡图片（中国内地） ● 身份证图片（中国内地） ● 机动车登记证书图片（中国内地） ● 护照图片（中国内地） ● 车险保单图片（中国内地） ● 机动车行驶证图片（中国内地）

敏感数据分类	类型
个人敏感信息	• 身份证号（中国内地） • 护照号（中国内地） • 驾照号（中国内地） • 港澳通行证 • 车牌号（中国内地） • 军官证号 • 美国社会保险号码SSN • ITIN • 社保相关信息 • 车辆识别代码 • 姓名（简体中文） • 姓名（英文） • 国籍 • 性别 • 民族 • 生日 • 出生地 • 教育程度 • 工作单位 • 工作行业 • 电话号码（中国内地） • 手机号码（中国内地） • 邮箱 • 微信号 • QQ号 • 婚姻状况 • 家庭成员关系 • 宗教信仰 • 银行卡号 • 信用卡号 • 万事达信用卡 • VISA信用卡 • 信用卡安全码

敏感数据分类	类型
企业敏感信息	• 工商注册号 • 统一社会信用代码 • 纳税人识别号(税号) • 组织机构代码 • 企业类型 • 经营状态 • 企业交付信息 • 企业需求信息
设备敏感信息	• 唯一设备识别码IMEI • 移动设备识别码MEID • MAC地址 • SIM卡IMSI信息 • IPv4地址 • IPv6地址 • Linux-Passwd文件 • Linux-Shadow文件
密钥敏感信息	• SSL Certificate • Secret_Access_Key • AWS_ACCESS_KEY • AWS_SECRET_KEY • Facebook_SECRET • IAM账号密码 • GitHub_KEY • DSA私钥 • EC私钥 • 加密私钥 • RSA私钥
位置敏感信息	• GPS信息 • 精确地址（中国） • 省份（中国内地） • 邮政编码（中国内地） • 城市（中国内地） • 直辖市（中国） • 地址（中国内地）

敏感数据分类	类型
系统敏感信息	• URL链接 • LDAP • OS类型
通用敏感信息	日期

4.4 DSC 支持的内置识别模板包含哪些识别规则?

数据安全中心的识别规则模板是根据不同行业规范、针对性定制的敏感数据分类分级。通过识别规则模板可以使敏感数据自动符合合规要求。具体可识别的模板如表4-5所示。

同时支持自定义分级分类模板，最多支持20个识别模板。

内置华为云数据安全分类分级模板

表 4-5 内置分类分级模板

一级分类	二级分类	敏感等级	内置规则
个人信息	权威社会标识	L3	身份证号（中国内地）
		L3	护照号（中国内地）
		L3	驾照号（中国内地）
		L3	港澳通行证
		L2	车牌号（中国内地）
		L3	军官证号
		L3	美国社会保险号码SSN
		L3	ITIN
		L3	社保相关信息
		L2	车辆识别代码
	个人一般信息	L1	姓名（简体中文）
		L1	姓名（英文）
		L2	国籍
		L2	性别
		L2	民族
		L2	生日

一级分类	二级分类	敏感等级	内置规则
		L2	出生地
		L2	教育程度
		L2	工作单位
		L2	工作行业
		L2	电话号码（中国内地）
		L3	手机号码（中国内地）
		L3	邮箱
		L2	微信号
		L2	QQ号
	个人私密信息	L4	婚姻状况
		L4	家庭成员关系
		L4	宗教信仰
	实名认证证明	L4	驾照图片（中国内地）
		L4	银行卡图片（中国内地）
		L4	身份证图片（中国内地）
		L4	机动车登记证书图片（中国内地）
		L4	护照图片（中国内地）
		L4	车险保单图片（中国内地）
		L4	机动车行驶证图片（中国内地）
	银行账号信息	L3	银行卡号
		L3	信用卡号
		L3	万事达信用卡
		L3	VISA信用卡
		L4	信用卡安全码
企业信息	企业标识信息	L1	工商注册号
		L1	统一社会信用代码
		L1	纳税人识别号(税号)
		L1	组织机构代码
		L1	营业执照图片

一级分类	二级分类	敏感等级	内置规则
	公开披露信息	L1	企业类型
		L1	经营状态
	企业内部信息	L2	企业交付信息
		L2	企业需求信息
设备信息	终端标识信息	L2	唯一设备识别码IMEI
		L2	移动设备识别码MEID
		L2	MAC地址
		L2	SIM卡IMSI信息
	IP地址信息	L2	IPv4地址
		L2	IPv6地址
	终端配置信息	L3	Linux-Passwd文件
		L3	Linux-Shadow文件
通用信息	时间信息	L1	日期
		L1	时间
	位置信息	L4	GPS信息
		L4	精确地址（中国）
		L2	省份（中国内地）
		L2	邮政编码（中国内地）
		L2	城市（中国内地）
		L2	直辖市（中国）
		L3	地址（中国内地）
	密钥凭证信息	L3	SSL Certificate
		L4	Secret_Access_Key
		L3	AWS_ACCESS_KEY
		L4	AWS_SECRET_KEY
		L4	Facebook_SECRET
		L4	IAM账号密码
		L4	GitHub_KEY
		L4	DSA私钥

一级分类	二级分类	敏感等级	内置规则
		L4	EC私钥
		L4	加密私钥
		L4	RSA私钥
	系统网络信息	L2	URL链接
		L2	LDAP
		L1	OS类型

4.5 数据脱敏是否对原始数据有影响?

没有影响。数据脱敏功能只会对数据进行读取，脱敏后保存到您选择的目标位置，不会对源数据进行改动。

⚠ 注意

- 如果需要填写已有的数据表，请勿选择业务数据表，以免影响业务。
- 目标数据表请勿选择原数据表，以免覆盖原始数据。

4.6 DSC 对可识别和脱敏的数据的字符集是否有要求?

DSC对可识别和脱敏数据库编码格式没有任何要求。

对于MRS类型数据源基于UDF脱敏的数据源仅支持UTF-8。

DSC可以识别的数据源对象：**DSC能够识别哪些数据源对象?** 。

DSC支持识别的敏感数据类型：**查看内置规则**。

4.7 为什么创建数据库脱敏任务时，无法找到已有的数据库实例中的表?

如果您在创建数据库脱敏任务时，选择已有的数据库实例，却无法找到对应实例里的表，一般是授权错误导致的。

背景

如果只对RDS数据库配置了“只读权限”，则只支持对数据库进行敏感数据识别，不支持数据脱敏。

原因

对该数据库实例只授权了“只读”权限，授权为“只读”权限无法进行数据脱敏。

解决办法

步骤1 先删除该数据库实例。

步骤2 再参照**添加数据库资产并授权**章节对数据库实例重新授权为“读写权限”。

说明

对数据库实例进行授权时，如果配额不够，请参照**升级版本和规格**购买数据库扩展包，1个数据库扩展包含1个可添加数据库（支持RDS、DWS、ECS自建数据库、DLI、CSS、ECS自建大数据等）资产。

----结束

5 数据水印类

5.1 数据水印功能会不会修改源数据?

华为云数据安全中心服务的数据水印功能不会修改源数据。

使用数据水印功能时，DSC通过调用OBS桶数据或者本地文件，将水印信息嵌入到文件后生成新的文档，该文档会下载到您指定的本地路径，所以对源数据不会有任何影响。更详细的了解，请参考**数据水印**章节。

5.2 文档损坏后，是否可以提取出水印?

DSC提供的数字水印能力具有高鲁棒性，即水印在传输或使用过程中不易被磨灭掉，数据载体即使经过被改动或受到攻击损坏后，依然有很大概率提取出水印。

- 添加水印后的文档被删除了几页后，仍然可以提取出水印。
- 添加水印后的图片被旋转、剪裁、缩放、修图等形变后，根据形变大小决定，形变较小则可以提取。

5.3 对待注入水印的源数据有什么要求?

由于注入水印的原理是将水印原子信息嵌入到不同特征的数据中去，因此源数据特征越多，越能嵌入完整的水印信息、提高提取成功率，并且即使缺失部分数据也不影响水印提取。所以对需要注入水印的数据有如下要求：

- 待注入水印的源数据需要大于等于1000行。

  小于1000行的源数据有可能因为特征不够导致提取水印失败。
- 尽量选取数据取值比较多样的列注入水印，如果该列的值是可枚举穷尽的，则有可能因为特征不够导致提取失败。

  常见的适合嵌入水印的列如地址、姓名、UUID、金额、总数等。

6 数据审计

6.1 DSC 可以检测哪些类型的异常事件?

数据安全中心服务当前仅支持对OBS桶数据进行异常检测。

DSC根据敏感数据规则对OBS桶进行识别，根据识别的敏感数据进行监控，监控到敏感数据的异常事件相关操作后，会将监控结果展示在异常事件处理页面中，用户可根据需要对异常事件进行处理。DSC支持检测的异常类型和异常内容如表6-1所示。

表 6-1 DSC 异常检测

异常类型	异常内容
数据访问异常	• 敏感文件的越权操作。 • 敏感文件的下载操作。
数据操作异常	• 敏感文件的更新操作。 • 敏感文件的文件内容追加操作。 • 敏感文件的删除操作。 • 敏感文件的复制操作。
数据管理异常	• 添加桶时，检测到桶为公共读或公共读写桶。 • 添加桶时，检测到私有桶对匿名用户或注册用户组开通了访问/ACL访问权限。 • 含有敏感文件的桶出现桶策略更改、删除操作。 • 含有敏感文件的桶出现桶ACL更改、删除操作。 • 含有敏感文件的桶出现跨区域复制配置的更改、删除操作。 • 敏感文件的对象出现ACL更改、删除操作。

6.2 如何对 DSC 的操作记录进行审计?

DSC的所有操作都会通过API形式记录在云审计服务（Cloud Trace Service，CTS）中。

开通云审计服务后，您可以在云审计服务中查看有关DSC的所有操作记录，供安全审查使用。关于如何查看审计日志，请参见**查看DSC的云审计日志**。