

UniCloud BDP 大数据平台

高危操作手册

紫光云技术有限公司
www.unicloud.com

资料版本：5W100-20211130
产品版本：UniCloud BDP (E6101)

© 紫光云技术有限公司 2021 版权所有，保留一切权利。

未经本公司书面许可，任何单位和个人不得擅自摘抄、复制本书内容的部分或全部，并不得以任何形式传播。

对于本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。

由于产品版本升级或其他原因，本手册内容有可能变更。紫光云保留在没有任何通知或者提示的情况下对本手册的内容进行修改的权利。本手册仅作为使用指导，紫光云尽全力在本手册中提供准确的信息，但是紫光云并不确保手册内容完全没有错误，本手册中的所有陈述、信息和建议也不构成任何明示或暗示的担保。

前言

本手册介绍了 UniCloud BDP 大数据平台在使用和维护过程中，所涉及到的可能会导致功能异常、数据丢失、业务中断或异常的操作。在实际使用过程中，要求运维人员请先了解各操作可能带来的风险，再按照操作建议进行操作，以防止影响业务的正常运行。

前言部分包含如下内容：

- [读者对象](#)
- [本书约定](#)
- [资料意见反馈](#)

读者对象

本手册主要适用于如下工程师：

- 网络规划人员
- 现场技术支持与维护人员
- 负责网络配置和维护的网络管理员

本书约定

1. 图形界面格式约定

格 式	意 义
<>	带尖括号“<>”表示按钮名，如“单击<确定>按钮”。
[]	带方括号“[]”表示窗口名、菜单名和数据表，如“弹出[新建用户]窗口”。
/	多级菜单用“/”隔开。如[文件/新建/文件夹]多级菜单表示[文件]菜单下的[新建]子菜单下的[文件夹]菜单项。

2. 各类标志

本书还采用各种醒目标志来表示在操作过程中应该特别注意的地方，这些标志的意义如下：

 警告	该标志后的注释需给予格外关注，不当的操作可能会对人身造成伤害。
 注意	提醒操作中应注意的事项，不当的操作可能会导致数据丢失或者设备损坏。
 提示	为确保设备配置成功或者正常工作而需要特别关注的操作或信息。
 说明	对操作内容的描述进行必要的补充和说明。
 窍门	配置、操作、或使用设备的技巧、小窍门。

3. 端口编号示例约定

本手册中出现的端口编号仅作示例，并不代表设备上实际具有此编号的端口，实际使用中请以设备上存在的端口编号为准。

资料意见反馈

如果您在使用过程中发现产品资料的任何问题，可以通过以下方式反馈：

E-mail: unicloud-ts@unicloud.com

感谢您的反馈，让我们做得更好！

目 录

1 说明	1
2 大数据平台 Web 类高危操作介绍	1
2.1 简介	1
2.2 集群管理高危操作介绍	1
2.2.1 删除集群【严重】	1
2.3 集群主机高危操作介绍	2
2.3.1 主机扩容【警告】	2
2.3.2 主机缩容（删除主机）【警告】	2
2.4 集群机架高危操作介绍	3
2.4.1 放入主机【警告】	3
2.5 告警管理高危操作介绍	3
2.5.1 删除告警阈值【一般】	3
2.5.2 删除告警通知【一般】	3
2.6 安全管理高危操作介绍	4
2.6.1 开启 Kerberos【严重】	4
2.6.2 删除用户【警告】	4
2.6.3 编辑（删除）用户授权【警告】	4
2.6.4 重置密码【警告】	5
2.6.5 编辑角色【警告】	5
2.6.6 删除密钥【警告】	5
2.6.7 编辑密钥授权【警告】	5
2.6.8 删除密钥授权【警告】	6
2.7 租户管理高危操作介绍	6
2.7.1 租户编辑（删除）用户【警告】	6
2.7.2 租户资源缩容【警告】	6
2.7.3 删除租户【严重】	7
2.8 备份恢复高危操作介绍	7
2.8.1 删除周期性同步任务【警告】	7
2.9 组件管理（公共类）高危操作介绍	7
2.9.1 添加业务组件【一般】	8
2.9.2 停止组件【警告】	8
2.9.3 重启组件【一般】	8

2.9.4 删除组件【警告】	9
2.9.5 修改组件配置【警告】	9
2.9.6 删除组件配置【警告】	10
2.9.7 将组件 Client 安装在大数据平台或大数据集群的节点上【严重】	10
2.9.8 添加进程【一般】	11
2.9.9 停止进程【警告】	11
2.9.10 重启进程【一般】	11
2.9.11 删除进程【警告】	12
2.10 组件管理（关键操作类）高危操作介绍	12
2.10.1 【ZooKeeper】修改 ZooKeeper 数据目录【警告】	12
2.10.2 【Zookeeper】修改依赖 ZooKeeper 的组件的 znode 节点【警告】	13
2.10.3 【Zookeeper】修改 Zookeeper 实例的 JVM 参数【警告】	13
2.10.4 【HDFS】删除 DataNode【警告】	14
2.10.5 【HDFS】修改 HDFS 相关数据目录【警告】	14
2.10.6 【HDFS】数据均衡【警告】	15
2.10.7 【HDFS】磁盘均衡【警告】	15
2.10.8 【YARN】删除 NodeManager【警告】	15
2.10.9 【YARN】修改 YARN 调度模式【警告】	16
2.10.10 【HBase】修改 HBase 加密参数【警告】	16
2.10.11 【HBase】修改 HDFS 上 HBase 存储目录【警告】	16
2.10.12 【Hive】修改 Hive 元数据库相关配置【警告】	17
2.10.13 【DLH】修改 DLHServer 元数据库相关配置【警告】	17
2.10.14 【Impala】重启 Impala daemon 进程【一般】	18
2.10.15 【Redis】修改 Redis 数据目录【警告】	18
2.10.16 【Oozie】修改 Oozie 关键配置【警告】	18
2.10.17 【Flume】修改 Flume 关键配置【警告】	19
2.10.18 【Storm】删除 Supervisor【警告】	19
2.10.19 【Storm】修改 Storm 关键配置【警告】	19
2.10.20 【Flink】修改 state 配置【警告】	20
2.10.21 【Kafka】修改 Kafka 数据目录【警告】	20
2.10.22 【Kafka】修改 Kafka 权限相关配置【警告】	20
2.10.23 【Kafka】修改 Kafka 监听配置【警告】	21
2.10.24 【Kafka】删除 Kafka Broker【警告】	21
2.10.25 【Elasticsearch】修改 Elasticsearch 数据目录【警告】	21
2.10.26 【Solr】修改 Solr 数据目录【警告】	22
2.11 系统管理 Web 类高危操作介绍	22

2.11.1 禁用系统用户【警告】	22
2.11.2 修改系统用户密码【一般】	23
2.11.3 修改系统用户授权【警告】	23
2.11.4 删除系统用户【严重】	23
2.11.5 修改 SSO 配置【严重】	23
2.11.6 断开 License Server 连接【警告】	24
2.11.7 系统升级【严重】	24
3 数据工厂 Web 类高危操作介绍	24
3.1 简介	24
3.2 集群管理高危操作介绍	25
3.2.1 删除大数据集群用户对数据工厂的影响【警告】	25
3.2.2 删除大数据集群对数据工厂的影响【严重】	25
3.3 数据探索高危操作介绍	25
3.3.1 删除 HDFS 文件【一般】	25
3.3.2 移动 HDFS 文件【一般】	26
3.3.3 重命名 HDFS 文件【一般】	26
3.3.4 删除库表【严重】	26
3.3.5 删除运行环境【一般】	26
3.3.6 编辑运行环境【一般】	27
3.3.7 重启运行环境【一般】	27
3.3.8 删除笔记【一般】	27
3.4 数据开发高危操作介绍	27
3.4.1 删除工作流【警告】	27
3.5 连接管理高危操作介绍	28
3.5.1 删除数据源【严重】	28
3.5.2 删除主机连接【警告】	28
4 后台类高危操作介绍	28
4.1 简介	28
4.2 命令行高危操作介绍	29
4.3 大数据平台（后台类）高危操作介绍	31
4.3.1 修改 NTP 服务配置【严重】	31
4.3.2 修改 YUM 源配置【警告】	31
4.3.3 root 配置 SSH 免密【严重】	31
4.3.4 后台修改主机 root 密码【严重】	32
4.3.5 修改主机 IP【严重】	32
4.3.6 修改主机名【严重】	32

4.3.7	对 HDFS 文件管理进行编辑【严重】	32
4.3.8	对大数据集群的元数据库进行编辑【严重】	33
4.3.9	删除大数据平台管理系统的内置 MySQL 数据库【严重】	33
4.3.10	修改主机防火墙配置【警告】	33
4.3.11	修改大数据集群中组件或进程相关端口号【警告】	34
4.3.12	未经适配的第三方软件与大数据集群混合部署【严重】	34
4.3.13	执行安装管理系统操作前修改配置文件 default.json【严重】	34
4.3.14	卸载大数据平台管理系统【严重】	34
4.3.15	安装管理系统操作后，单独安装数据工厂时修改配置文件 default.json【严重】	35
4.3.16	单独卸载数据工厂【严重】	35
4.3.17	执行大数据平台管理系统版本升级前修改配置文件 upgrade.json【严重】	36
4.4	组件管理（后台类）高危操作介绍	36
4.4.1	【组件公共】未开启 Kerberos 的环境，使用组件超级用户执行组件操作【严重】	36
4.4.2	【HDFS】删除 HDFS 数据目录【严重】	37
4.4.3	【HDFS】删除 HDFS 中存储的数据【严重】	37
4.4.4	【ZooKeeper】业务代码连接 Zookeeper【严重】	37
4.4.5	【Elasticsearch】删除 Elasticsearch 数据目录【严重】	38
4.4.6	【Elasticsearch】删除 Elasticsearch 集群索引【严重】	38
4.4.7	【Solr】删除 Solr 数据目录【严重】	38
4.4.8	【Solr】删除 Solr 集群索引【严重】	39
4.4.9	【Kafka】删除 Kafka 数据目录【严重】	39
4.4.10	【Redis】删除 Redis 数据目录【严重】	39
4.4.11	【DLH】DLH 库表删除【严重】	40
4.4.12	【DLH】DLH 表清空【严重】	40
4.4.13	【DLH】删除 DLH 库表在 HDFS 上的存储路径【严重】	40
4.4.14	【DLH】删除 DLH 元数据库表【严重】	41
4.4.15	【HBase】修改或删除 Phoenix 系统表或系统表数据（SYSTEM.*）【严重】	41
4.4.16	【HBase】修改或删除 HBase 元数据表【严重】	41
4.4.17	【HBase】删除 HBase 在 HDFS 上的存储路径【严重】	41
4.4.18	【Hive】Hive 库表删除【严重】	42
4.4.19	【Hive】Hive 表清空【严重】	42
4.4.20	【Hive】删除 Hive 库表在 HDFS 上的存储路径【严重】	42
4.4.21	【Hive】删除 Hive 元数据库表【严重】	42
4.4.22	【Impala】删除 Impala 元数据库表【严重】	43
4.4.23	【Impala】修改 Impala 元数据库表数据【严重】	43
4.4.24	【Impala】删除 Impala 在 HDFS 上的存储路径【严重】	43

4.4.25	【Spark】Spark 库表删除【严重】	44
4.4.26	【Spark】Spark 表清空【严重】	44
4.4.27	【Spark】删除 Spark 库表在 HDFS 上存储路径【严重】	44
4.4.28	【Spark】删除 Spark 元数据库表【严重】	44

1 说明

本文档介绍的所有高危操作，根据危险程度共分为严重、警告、一般三种级别。在实际使用过程中，请先了解各操作可能带来的风险，再按照操作建议进行操作，以防止影响业务的正常运行。

- 严重：功能全部丧失、数据丢失或者业务中断
- 警告：大部分功能丧失或者影响业务
- 一般：部分功能不可用或者不影响业务

2 大数据平台 Web 类高危操作介绍

2.1 简介

高危险的 Web 类操作只能由有资质、且经过培训的维护人员执行。如果操作不当，可能会导致集群异常、组件异常、业务中断、业务运行异常、重要文件被删除、所有配置被清除、用户无法登录、用户下线等现象发生。

Web 类高危操作包括但不限于本手册中介绍的内容，在进行 Web 类高危操作之前，请先了解可能带来的风险再进行操作。

2.2 集群管理高危操作介绍

2.2.1 删除集群【严重】

1. 操作事项

在[集群管理/集群列表]页面执行删除集群的操作，集群删除的相关信息确认并选择完成后，还需根据弹窗中的提示进行相关操作（比如：输入集群名和集群 root 密码进行二次确认），然后即可删除集群。

2. 可能带来的风险

- 删除集群操作一旦执行，不可回退或暂停。
- 执行删除集群操作时可选择是否删除数据，分为两种情况：
 - 对于在线主机，执行删除集群操作时可选择是否删除数据。若勾选删除数据，则同时删除节点上的组件和数据；若不勾选删除数据，则此时仅删除组件，该节点若想被其他大数据集群重新使用，可后续在[集群管理/主机管理/主机资源]的主机列表处手动删除对应节点上的数据。
 - 对于离线主机，执行删除集群操作时不能删除离线节点上的组件及数据（需等待集群删除操作完成且主机恢复正常后，在[集群管理/主机管理/主机资源]的主机列表处手动删除离线节点上的组件和数据，然后该节点才能被其他大数据集群重新使用）。
- 若某租户集群中已创建租户，则删除该租户集群时，对应的租户会一起被删除。

3. 操作建议

此操作风险极高，请谨慎执行。若此操作不可避免，请确认当前环境下集群是否可以删除（集群不再使用、集群无业务正在运行、集群内重要的数据已备份），待完成充分确认和数据备份后再执行删除操作。

2.3 集群主机高危操作介绍

2.3.1 主机扩容【警告】

1. 操作事项

在[集群管理/集群列表]页面选择待扩容的集群进入集群详情页面，在[主机]页签或集群操作下拉框中选择执行主机扩容的操作。

2. 可能带来的风险

- 主机扩容操作一旦执行，不可回退或暂停。
- 集群执行主机扩容操作之后，需要手动重启相关受影响的组件以使配置生效，但是重启操作可能会导致正在运行的业务中断。
- 集群执行主机扩容操作之后，部分组件短时间内可能会进行数据重分布或集群内部任务重新下发等操作，可能对当前业务有短暂的性能影响。

3. 操作建议

集群执行主机扩容操作之前建议评估业务可中断时间窗口，在有效窗口期内进行。

2.3.2 主机缩容（删除主机）【警告】

1. 操作事项

在[集群管理/集群列表]页面选择待缩容的集群进入集群详情页面，在[主机]页签或集群操作下拉框中选择执行删除主机的操作。

2. 可能带来的风险

- 删除主机操作一旦执行，不可回退或暂停。
- 集群执行删除主机操作之后，需要手动重启相关受影响的组件以使配置生效，但是重启操作可能会导致正在运行的业务中断。
- 集群执行删除主机操作之后，部分组件短时间内可能会进行数据重分布或集群内部任务重新下发等操作，可能会导致正在运行的业务响应变慢或者不稳定。
- 若主机节点上安装了 **master** 类型的进程，执行删除主机操作会导致集群组件运行异常。
- 删除数据节点可能会导致部分组件的数据丢失，删除之前需确认数据节点的数量是否大于等于数据副本数。
- 对于单机模式部署的组件（如 **Redis**），执行删除对应的主机会导致组件被删除，进而导致数据丢失，对应业务场景中断。

3. 操作建议

- 集群执行删除主机操作之前，请务必检查该主机上已部署的组件进程。评估待删除节点上的所有数据，请根据实际情况判断是否需要对该节点的数据进行迁移或备份。

- 对于安装 **master** 类型进程的主机节点，禁止执行删除主机操作，否则会导致集群组件运行异常。
- 集群执行删除主机操作之前建议评估业务可中断时间窗口，在有效窗口期内进行。
- 集群执行删除主机操作之前，需要评估主机扩容之后的集群是否能负载业务需求，否则会导致集群运行压力过大而不稳定。

2.4 集群机架高危操作介绍

2.4.1 放入主机【警告】

1. 操作事项

在[集群管理/集群列表]页面单击集群名进入集群详情页面，在[主机]页签下的机架列表，对某机架执行放入主机操作。

2. 可能带来的风险

若不按实际情况配置主机的机架信息，可能会导致数据存储不均衡。

3. 操作建议

请谨慎执行本操作。执行此操作前，建议确认主机实际所属机架信息，然后按照实际情况进行配置。

2.5 告警管理高危操作介绍

2.5.1 删除告警阈值【一般】

1. 操作事项

在[集群管理/告警管理/阈值设置]页面，执行删除告警阈值的操作。

2. 可能带来的风险

告警阈值删除后，再到达阈值时将不会再产生告警。

3. 操作建议

请谨慎执行本操作。执行此操作前，请确认此告警不再需要。

2.5.2 删除告警通知【一般】

1. 操作事项

在[集群管理/告警管理/告警通知]页面，执行删除告警通知的操作。

2. 可能带来的风险

告警通知删除后，再产生的告警将不会通过邮件进行通知。

3. 操作建议

请谨慎执行本操作。执行此操作前，请确认不再需要通过邮件进行告警通知。

2.6 安全管理高危操作介绍

2.6.1 开启 Kerberos 【严重】

1. 操作事项

在[集群管理/集群列表]页面选择目标集群进入集群详情页面，在集群操作下拉框中选择执行开启 Kerberos 的操作。

2. 可能带来的风险

- 集群开启 Kerberos 的操作一旦执行，不可回退或暂停，且 Kerberos 一旦开启后无法关闭。
- 执行开启 Kerberos 操作时，会停止当前集群中的所有组件，所有业务中断。开启 Kerberos 操作完成之后，需要手动启动所有组件以使配置生效，且当前所有业务均无法直接恢复（需要重新适配）。
- 集群开启 Kerberos 之后，用户需进行 kerberos 认证后才能进行集群操作。

3. 操作建议

请谨慎执行本操作。集群开启 Kerberos 之后，大数据平台组件的使用需增加用户认证步骤，且平台上层应用均需增加 Kerberos 校验功能，请根据项目实际需求进行充分评估后再决定是否开启 Kerberos。

2.6.2 删除用户 【警告】

1. 操作事项

在[权限管理/用户管理]页面执行删除用户的操作，并在弹窗中输入该用户的密码，然后进行确定即可删除该用户。

2. 可能带来的风险

- 该操作会导致集群内该用户的信息丢失，导致无法通过该用户连接集群。
- 如果正在使用某用户运行业务，删除该用户后，业务会中断。

3. 操作建议

请谨慎执行本操作。建议删除用户前评估是否影响自身业务，且操作前请确认选择的用户是否正确。

2.6.3 编辑（删除）用户授权 【警告】

1. 操作事项

在[集群权限/用户管理]页面执行修改用户授权的操作。

2. 可能带来的风险

若执行该操作时删除或修改了一些用户角色授权，则会导致该用户相关权限改变，从而导致对应的用户业务应用无权限访问集群数据，进而导致业务中断。

3. 操作建议

请谨慎执行本操作。建议在编辑或删除用户授权前评估权限变更是否影响自身业务，且操作前请确认选择用户以及更新的用户授权是否正确。重置密码后需要重新下载并更新业务端认证文件。

2.6.4 重置密码【警告】

1. 操作事项

在[集群权限/用户管理]页面执行重置密码的操作。

2. 可能带来的风险

执行该操作用户密码会被重置，可能导致一些使用了该用户的应用认证失败，进而导致业务中断。

3. 操作建议

请谨慎执行本操作。建议在重置用户密码前评估密码变更是否影响自身业务，且操作前请确认选择的用户是否正确。

2.6.5 编辑角色【警告】

1. 操作事项

在[集群权限/角色管理]页面执行编辑角色的操作。

2. 可能带来的风险

若执行该操作时删除或修改了一些组件权限，则会导致该角色已关联的用户权限改变，从而导致对应的用户业务应用无权限访问集群数据，进而导致业务中断。

3. 操作建议

请谨慎执行本操作。建议编辑角色前评估权限变更是否影响自身业务，且操作前请确认选择的角色是否正确。

2.6.6 删除密钥【警告】

1. 操作事项

在[集群权限/密钥管理]页面执行删除密钥的操作。

2. 可能带来的风险

该操作会导致加密区数据无法读写。

3. 操作建议

请谨慎执行本操作。建议操作前评估删除密钥是否影响自身业务，且操作前请确认选择的密钥是否正确。

2.6.7 编辑密钥授权【警告】

1. 操作事项

在[集群权限/密钥管理]页面执行编辑密钥授权的操作。

2. 可能带来的风险

该操作会导致对应加密区数据访问权限改变。

3. 操作建议

请谨慎执行本操作。建议操作前评估密钥授权变更是否影响自身业务，且操作前请确认选择的密钥授权是否正确。

2.6.8 删除密钥授权【警告】

1. 操作事项

在[集群权限/密钥管理]页面执行删除密钥授权的操作。

2. 可能带来的风险

该操作会导致对应加密区数据无法读写。

3. 操作建议

请谨慎执行本操作。建议操作前评估密钥授权删除是否影响自身业务，且操作前请确认选择的密钥授权是否正确。

2.7 租户管理高危操作介绍

2.7.1 租户编辑（删除）用户【警告】

1. 操作事项

在[集群权限/租户管理]页面执行编辑（删除）租户用户的操作，或者在[集群权限/用户管理]页面执行编辑（删除）租户用户的操作。

2. 可能带来的风险

删除租户用户后会导致该用户无权限访问租户内的组件资源数据，若此用户正在使用租户内的组件资源，将导致该用户的业务发生中断等异常。

3. 操作建议

请谨慎执行本操作。建议操作前评估此用户不再需要使用租户内的组件资源，且操作前请确认选择的租户用户是否选择正确。

2.7.2 租户资源缩容【警告】

1. 操作事项

在[权限管理/租户管理]页面选择待缩容的租户进入租户详情页面，在[资源列表]页签执行租户资源缩容的操作。

2. 可能带来的风险

- 租户资源缩容将导致组件资源容量减少。
- 若修改后的 YARN 组件资源容量小于现有业务已占用的资源容量（现有业务已占用的租户资源容量，在租户详情页可查看），将导致相关任务执行阻塞。

3. 操作建议

请谨慎执行本操作。建议操作前请务必根据业务场景的实际需求情况对租户资源容量进行合理规划，租户资源执行缩容操作后须确保缩容后的容量不小于现有业务需要的容量。

【说明】当前版本中，Kafka 租户资源禁止执行缩容操作；HDFS、Hive 和 HBase 租户资源在执行缩容操作时，会判断不允许小于已使用的资源值的大小；YARN 租户资源的已使用值是动态的，所以允许出现超额情况，未加限制。

2.7.3 删除租户【严重】

1. 操作事项

在[集群权限/租户管理]页面执行删除租户的操作，并输入租户名称和主用户密码后，然后进行确定后即可删除租户；或者在租户详情页面执行删除某组件的租户资源的操作。

2. 可能带来的风险

- 删除租户的操作一旦执行，不可回退或暂停。
- 删除租户的操作或删除组件的租户资源的操作均会导致该租户中的数据和组件资源数据一并被删除，该租户中的业务数据被全部清空，业务中断。

3. 操作建议

此操作风险极高，请谨慎执行。若此操作不可避免，请确认当前环境下租户或租户内组件资源是否可以删除（租户不再被业务使用、租户内重要的数据已备份），待完成充分确认和数据备份后再执行删除操作。

2.8 备份恢复高危操作介绍

2.8.1 删除周期性同步任务【警告】

1. 操作事项

当周期性同步任务处于未启动状态时，在[集群管理/备份恢复]页面对其执行删除操作。

2. 可能带来的风险

删除周期性同步任务后，源集群中新增的数据将不能在周期时间内同步到目的集群。此时，若源集群出现故障，则会因数据没有备份到目的集群而造成数据丢失。

3. 操作建议

请谨慎执行本操作。执行此操作前，请确认该同步任务所涉及的数据已经不需要备份到目的集群。

2.9 组件管理（公共类）高危操作介绍



大数据平台中包含多个大数据组件，本章节仅介绍公共类的组件管理相关高危操作。关于各个组件的配置相关的重点高危操作详情请参见 [2.10 组件管理（关键操作类）高危操作介绍](#) 章节。

2.9.1 添加业务组件【一般】

1. 操作事项

在[集群管理/集群列表]页面选择目标集群进入集群详情页面，在[组件]页签或集群操作下拉框中选择执行添加业务组件的操作。

2. 可能带来的风险

- 执行添加业务组件操作之后，需要手动重启相关受影响的组件以使该组件加入集群生效，但是重启操作可能会导致正在运行的业务中断。
- 添加业务组件操作执行不当时，可能会造成集群异常，进而影响该集群上正在运行的业务。

3. 操作建议

请谨慎执行本操作。建议在集群中无业务运行时，再依据实际业务需求执行添加业务组件的操作。

2.9.2 停止组件【警告】

1. 操作事项

在[集群管理/集群列表]页面选择目标集群进入集群详情页面，在[组件]页签、集群操作下拉框中或组件详情页面的组件操作下拉框中选择执行停止组件的操作，或者在[集群管理/组件管理]页面组件操作下拉框中选择执行停止组件的操作。

2. 可能带来的风险

- 执行停止组件操作，若集群中有业务正在进行读写操作，则可能产生坏文件或者数据丢失。
- 执行停止组件操作之后，将中止该组件及与其有依赖关系的其他组件上的业务访问。若该组件上有任务正在运行，业务将中断。
- 执行停止组件操作之后，与其有依赖关系的其他组件可能也会受到影响（比如：**Spark** 依赖 **Hive**，当 **Hive** 停止后 **Spark** 部分功能会异常）。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- **ZooKeeper** 等基础组件禁止单独停止。
- 建议在集群中无业务运行时，再依据实际业务需求执行停止组件的操作。
- 若集群中有业务正在运行，首先需确认待停止组件是否影响业务运行，然后在业务中断时间窗口内执行停止组件的操作。

2.9.3 重启组件【一般】

1. 操作事项

在[集群管理/集群列表]页面选择目标集群进入集群详情页面，在[组件]页签、集群操作下拉框中或组件详情页面的组件操作下拉框中选择执行重启组件的操作，或者在[集群管理/组件管理]页面组件操作下拉框中选择执行重启组件的操作。

2. 可能带来的风险

- 执行重启组件操作，若集群中有业务正在进行读写操作，则可能产生坏文件或者数据丢失。

- 执行重启组件操作的过程中，会中止该组件及与其有依赖关系的其他组件上的业务访问。若该组件上有任务正在运行，业务将短暂中断。
- 重启组件操作执行不当时，可能会造成该组件长时间无法启动，影响该组件关联的业务运行。若组件长时间无法启动，请尽快联系紫光云技术支持。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- ZooKeeper 等基础组件禁止单独重启。
- 建议在集群中无业务运行时，再依据实际业务需求执行重启组件的操作。
- 若集群中有业务正在运行，首先需确认待重启组件是否影响业务运行，然后在业务中断时间窗口内执行重启组件的操作。

2.9.4 删除组件【警告】

1. 操作事项

在[集群管理/集群列表]页面选择目标集群进入集群详情页面，在[组件]页签或组件详情页面的组件操作下拉框中选择执行删除组件的操作，或者在[集群管理/组件管理]页面的组件操作下拉框中选择执行删除组件的操作。

2. 可能带来的风险

执行删除组件操作会删除该组件的所有业务，造成不可逆的业务中断。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请确认当前环境下该组件是否可以删除（组件不再使用、组件无业务正在运行、组件中重要的数据已备份），待完成充分确认和数据备份后再执行删除操作。

2.9.5 修改组件配置【警告】

1. 操作事项

在[集群管理/集群列表]页面选择目标集群进入集群详情页面，在[组件]页签选择目标组件进入组件详情页面，在[配置]页签执行修改组件配置的操作，或者在[集群管理/组件管理]页面的[配置]页签执行修改组件配置的操作。组件配置修改完成并保存后，需要重启相关受影响的组件。

2. 可能带来的风险

- 组件执行修改配置操作之后，需要手动重启相关受影响的组件以使配置生效，但是重启操作可能会导致业务中断。
- 修改组件的重要配置可能会导致组件异常，组件可能无法启动或数据丢失，影响业务运行。
- 若修改 JVM 信息，则可能会影响组件的稳定性或性能。
- 组件配置修改不当时，可能会造成该组件无法启动，影响该组件关联的业务运行。若组件长时间无法启动，请尽快联系紫光云技术支持。
- 某些组件的部分参数配置项修改后，可能不满足集群运行和使用需求。例如：若使用 Log 参数值配置不当会影响磁盘空间等。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 该操作需要组件专业人员执行，要求对组件配置信息十分了解。
- 建议在集群中无业务运行时，再依据实际业务需求执行修改组件配置的操作（且务必保持修改后的配置正确且有效）。
- 若集群中有业务正在运行，组件配置不可修改，此时若必须修改则要求在业务中断时间窗口内执行。
- 即使组件数据不再需要或组件配置修改符合业务需求，也不建议对组件执行修改元数据、数据目录及其文件权限等操作。

2.9.6 删除组件配置【警告】

1. 操作事项

在[集群管理/集群列表]页面选择目标集群进入集群详情页面，在[组件]页签选择目标组件进入组件详情页面，在[配置]页签执行删除组件配置的操作，或者在[集群管理/组件管理]页面的[配置]页签执行删除组件配置的操作。组件配置删除完成并保存后，需要重启相关受影响的组件。

2. 可能带来的风险

- 组件执行删除配置操作之后，需要手动重启相关受影响的组件以使配置生效，但是重启操作可能会导致业务中断。
- 删除组件的重要配置将导致组件异常，组件可能无法启动或数据丢失，影响业务运行。
- 组件配置删除不当时，可能会造成该组件无法启动，影响该组件关联的业务运行。若组件长时间无法启动，请尽快联系紫光云技术支持。
- 某些组件的部分参数配置项删除后，将使用原始默认值，不满足集群运行和使用需求。例如：若使用 JVM 参数默认值会影响集群性能，若使用 Log 参数默认值会影响磁盘空间等。

3. 操作建议

请谨慎执行本操作。该操作需要组件专业人员执行，要求对组件配置信息十分了解。若集群中有业务正在运行，组件配置禁止删除。

2.9.7 将组件 Client 安装在大数据平台或大数据集群的节点上【严重】

1. 操作事项

将下载的组件 Client，安装在大数据平台管理节点或者大数据集群的各个节点上。

2. 可能带来的风险

- 可能会影响大数据平台的内部服务。
- 可能与大数据集群中已安装的组件产生冲突。

3. 操作建议

此操作风险较高，禁止执行。

建议将下载的组件 Client，安装至与大数据平台及大数据集群均无关的主机节点上，但要求安装 Client 的主机与大数据平台和大数据集群可以网络互连。

2.9.8 添加进程【一般】

1. 操作事项

在[集群管理/集群列表/组件]页签选择目标组件进入组件详情页面，在组件操作下拉框中选择执行添加进程的操作，或者在[集群管理/组件管理]页面的组件操作下拉框中选择执行添加进程的操作。

2. 可能带来的风险

- 执行添加进程操作之后，需要手动重启相关受影响的组件以使该进程加入集群生效，但是重启操作可能会导致正在运行的业务中断。
- 添加进程操作执行不当时，可能会造成集群异常，进而影响该集群上支撑的业务。

3. 操作建议

请谨慎执行本操作。建议在集群中无业务运行时，再依据实际业务需求执行添加进程的操作。

2.9.9 停止进程【警告】

1. 操作事项

在[集群管理/集群列表/组件]页签选择目标组件进入组件详情页面，在[部署拓扑]页签执行停止进程的操作，或者在[集群管理/组件管理]页面的[部署拓扑]页签执行停止进程的操作。

2. 可能带来的风险

- 执行停止进程操作之后，将中止该进程及与其有依赖关系的其他组件上的业务访问。
- 对于组件的相同进程，若均进行停止操作，则无法对外提供服务。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 建议在集群中无业务运行时，再依据实际业务需求执行停止进程的操作。
- 若集群中有业务正在运行，首先需确认待停止进程是否影响业务运行，然后在业务中断时间窗口内执行停止进程的操作。
- 对于部分组件，需关注停止进程后，该组件数据节点的数量是否大于等于数据副本数，以防止出现副本丢失导致组件异常。

2.9.10 重启进程【一般】

1. 操作事项

在[集群管理/集群列表/组件]页签选择目标组件进入组件详情页面，在[部署拓扑]页签执行重启进程的操作，或者在[集群管理/组件管理]页面的[部署拓扑]页签执行重启进程的操作。

2. 可能带来的风险

- 执行重启进程操作的过程中，会中断该进程及与其有依赖关系的其他组件上的业务访问。
- 重启进程操作执行不当时，可能会造成该组件长时间无法启动，影响该组件关联的业务运行。若组件长时间无法启动，请尽快联系紫光云技术支持。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- ZooKeeper 等基础组件的进程禁止单独重启。
- 建议在集群中无业务运行时，再依据实际业务需求执行重启进程的操作。
- 若集群中有业务正在运行，首先需确认待重启进程是否影响业务运行，然后在业务中断时间窗口内执行重启进程的操作。

2.9.11 删除进程【警告】

1. 操作事项

在[集群管理/集群列表/组件]页签选择目标组件进入组件详情页面，在[部署拓扑]页签执行重启进程的操作，或者在[集群管理/组件管理]页面的[部署拓扑]页签执行重启进程的操作。

2. 可能带来的风险

- 执行删除进程操作会删除该进程上的所有业务，造成不可逆的业务中断。
- 部分组件执行删除进程，会导致数据重新分布，造成一定的性能损耗。
- 执行删除进程操作之后，需要手动重启相关受影响的组件，但是重启操作可能会导致正在运行的业务中断。
- 删除进程操作执行不当时，可能会造成该组件异常，影响该组件关联的业务运行。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 建议在集群中无业务运行时，再依据实际业务需求执行删除进程的操作。
- 若集群中有业务正在运行，首先需确认待删除进程是否影响业务运行，然后在业务中断时间窗口内执行删除进程的操作。
- 待删除进程中若有重要的数据，则需要进行备份。
- 对于部分组件，需关注删除进程后，该组件数据节点的数量是否大于等于数据副本数，以防止出现副本丢失导致组件异常。

2.10 组件管理（关键操作类）高危操作介绍

2.10.1 【ZooKeeper】修改 ZooKeeper 数据目录【警告】

1. 操作事项

在 ZooKeeper 组件详情页面，在[配置]页签执行修改 ZooKeeper 数据目录的操作（即修改配置项 zoo.cfg 中 dataDir 的属性值）。

2. 可能带来的风险

- HDFS/YARN/HBase/Hive 等众多组件均依赖 ZooKeeper，会在 ZooKeeper 中保存元数据信息，所以修改 ZooKeeper 相关的数据目录将会影响到相关组件的正常运行。
- 修改 ZooKeeper 数据目录可能会导致之前数据的丢失，从而影响依赖 Zookeeper 的组件的正常运行。
- 修改 ZooKeeper 数据目录之后，需要重启 Zookeeper 以使配置修改生效，此时依赖 Zookeeper 的组件也需要重启，影响其正常运行。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 操作前请务必提前对数据进行备份。
- 建议修改后的数据目录仍在数据盘。对于新增的数据目录，需确保 Zookeeper 拥有读/写权限，且必须保证该目录为空。
- 执行本操作前，要求关闭业务系统对 Zookeeper 集群的访问。

2.10.2 【Zookeeper】修改依赖 ZooKeeper 的组件的 znode 节点【警告】

1. 操作事项

修改依赖 ZooKeeper 的组件的 znode 节点相关配置。例如：HBase 依赖 ZooKeeper，在 HBase 组件详情页面，在[配置]页签执行修改 znode 节点相关配置（即修改配置项 hbase-site 中 zookeeper.znode.parent 的属性值）。

2. 可能带来的风险

HDFS/YARN/HBase/Hive/Solr 等依赖 ZooKeeper 的组件在 ZooKeeper 中都有自己的 znode 节点来存储自己的元数据，所以，若修改组件配置中的该属性则可能会导致元数据的丢失，进而导致大数据组件运行异常，影响业务的正常运行。

3. 操作建议

请谨慎执行本操作。建议操作之前，请注意：

- 请确认 HDFS/HBase/Solr 等依赖 ZooKeeper 的组件是否可以变更 znode 节点路径，同时做好 Zookeeper 相关数据备份。
- 请需要确认组件停止是否不影响业务运行。

2.10.3 【Zookeeper】修改 Zookeeper 实例的 JVM 参数【警告】

1. 操作事项

在 ZooKeeper 组件详情页面，在[配置]页签执行修改 Zookeeper 实例的 JVM 参数的操作（即修改配置项 zookeeper-env 中 content/zk_server_heapsize 的属性值）。

2. 可能带来的风险

- 组件执行修改配置操作之后，需要手动重启相关受影响的组件以使配置生效，但是重启操作可能会导致业务中断。
- 组件配置修改不当时，可能会造成该组件及与其有依赖关系的其他组件无法启动，影响该组件关联的业务运行。若组件长时间无法启动，请尽快联系紫光云技术支持。

3. 操作建议

请谨慎执行本操作。该操作需要组件专业人员执行，要求对组件配置信息十分了解。执行本操作前，请注意：

- 要求在集群中无业务运行。
- 依据实际业务需求修改 Zookeeper 实例的 JVM 参数时务必要保证修改后的配置正确且有效。另外，修改组件 GC、HEAPSIZE 等 JVM 参数会影响组件运行性能，同时需要根据集群节点的内存、CPU 以及当前节点系统负载去设置合理的值。

2.10.4 【HDFS】删除 DataNode 【警告】

1. 操作事项

在 HDFS 组件详情页面，在[部署拓扑]页签执行删除 DataNode 进程的操作。

2. 可能带来的风险

DataNode 管理本节点的数据，若删除 DataNode 会将降低 HDFS 组件的存储容量并增大数据丢失的风险。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 一般情况下，不建议对 HDFS 执行缩容操作。
【说明】当集群中只有 3 个 DataNode 进程时，不允许执行删除操作。
- 确认当前环境下该进程是否可以删除（该 DataNode 节点上无业务正在运行、删除 DataNode 后的数据节点可以满足数据存储需求）。
- 删除 DataNode 后，HDFS 为保证所有数据块的副本数达到配置的副本数（默认值为 3），会自行对部分数据块进行复制，以确保数据可靠。
- 当需要删除多个 DataNode 时，为保证数据不丢失，建议依次删除 DataNode（即删除一个 DataNode，请等待一段时间后，尽量保证在 HDFS 数据同步完成之后，再删除下一个 DataNode）。

2.10.5 【HDFS】修改 HDFS 相关数据目录 【警告】

1. 操作事项

在 HDFS 组件详情页面，在[配置]页签执行修改 HDFS 相关数据目录的操作。

2. 可能带来的风险

HDFS 包括多个数据目录配置项，各个目录配置信息以及可能导致的风险如[表 2-1](#)所示。

表2-1 HDFS 数据目录

数据目录配置项	存储数据	修改可能导致的风险
dfs.namenode.name.dir	NameNode 元数据信息的目录	元数据记录了 HDFS 数据存储的实际地址，删除元数据信息的目录可能会导致数据全部丢失
dfs.datanode.data.dir	HDFS 实际存储数据的目录	变更或删除数据目录会导致数据丢失
dfs.namenode.checkpoint.dir	HDFS 未开启 HA 时，SecondaryNameNode 存储镜像文件的目录	存储元数据的快照数据，删除可能导致元数据无法恢复
dfs.journalnode.edits.dir	HDFS 开启 HA 时，Journalnode 存储操作日志信息目录	用于同步主备 NameNode 之间的元数据信息，修改可能会导致部分最新的元数据丢失或导致组件异常

3. 操作建议

请谨慎执行本操作。一般情况下，集群正常且已保存数据的情况下不建议修改 HDFS 数据目录的相关配置。若遇到磁盘损坏等问题，请尽快联系紫光云技术支持。

2.10.6 【HDFS】数据均衡【警告】

1. 操作事项

数据均衡指在 HDFS 的所有 DataNode 节点之间进行数据均衡，此时需要跨数据节点进行数据迁移。在 HDFS 组件详情页面，在组件操作下拉框中选择执行数据均衡的操作。数据均衡操作的运行时间取决于集群大小和数据的不平衡程度。

2. 可能带来的风险

- 数据均衡在执行过程中会消耗集群资源，影响集群的性能。
- 在数据量较大的情况下，数据均衡操作需要较长的时间才能执行完成，会对集群性能造成影响。

3. 操作建议

建议在集群中无业务运行时，执行数据均衡的操作。在执行数据均衡操作时，请选择合适的均衡阈值，均衡阈值配置越小则均衡器需要执行的工作越多，耗费的时间就更长。

2.10.7 【HDFS】磁盘均衡【警告】

1. 操作事项

磁盘均衡指在 HDFS 内各个 DataNode 节点上的不同磁盘之间进行数据均衡，此时不需要跨数据节点，仅需要在同一个 DataNode 节点上的不同磁盘之间进行数据迁移。在 HDFS 组件详情页面，在[部署拓扑]页签执行磁盘均衡的操作。

2. 可能带来的风险

- 磁盘均衡在执行过程中会消耗集群资源（比如：磁盘 IO、网络带宽），影响集群的性能。
- 在数据量较大的情况下，磁盘均衡操作需要较长的时间才能执行完成，且需要长时间读写磁盘，会对集群性能造成影响。

3. 操作建议

建议在集群中无业务运行时，执行磁盘均衡的操作。执行磁盘均衡操作前，需提前做好合理规划（选择合适的均衡容忍百分比和可使用的最大网络带宽），避免磁盘均衡操作运行太长时间，使用过多的集群资源，从而影响集群上的其他业务。

2.10.8 【YARN】删除 NodeManager【警告】

1. 操作事项

在 YARN 组件详情页面，在[部署拓扑]页签执行删除 NodeManager 进程的操作。

2. 可能带来的风险

NodeManager 负责本节点资源、任务的调度与管理，删除 NodeManager 可能导致正在运行的任务失败，并造成集群可用资源的减少。

3. 操作建议

请谨慎执行本操作。一般情况下，要求在集群部署前就对 **NodeManager** 数量做好合理规划。若此操作不可避免，请确认当前环境下该进程是否可以删除（该 **NodeManager** 节点上无业务正在运行）。

2.10.9 【YARN】修改 YARN 调度模式【警告】

1. 操作事项

在 **YARN** 组件详情页面，在[配置]页签执行修改 **YARN** 调度模式的操作（即修改配置项 `yarn.resourcemanager.scheduler.class` 的值）。

2. 可能带来的风险

- 当前版本中，租户集群调度模式仅支持公平调度，若将其改为其他调度模式，则会导致租户计算资源无法正常使用，且可能导致集群上 **YARN** 组件无法正常使用，从而造成业务中断。
- 当前版本中，独立集群调度模式默认为公平调度，可根据实际业务需求修改。在数据量很大的情况下，容量调度模式可能会出现 **YARN** 卡顿。

3. 操作建议

请谨慎执行本操作。当前版本中，租户集群禁止修改 **YARN** 调度模式。若集群根据实际业务需求需要修改 **YARN** 调度模式，请咨询紫光云技术支持。

2.10.10 【HBase】修改 HBase 加密参数【警告】

1. 操作事项

在 **HBase** 组件详情页面，在[配置]页签执行修改 **HBase** 加密参数的操作（即修改自定义配置 `hbase-site` 中的“加密属性”）。

2. 可能带来的风险

HBase 加密类配置参数若修改不当将导致组件无法启动，数据无法正常访问（`*.encryption.*`、`keyprovider.*`）。

3. 操作建议

请谨慎执行本操作。该操作需要组件专业人员执行，要求对组件配置信息十分了解。依据实际业务需求修改 **HBase** 加密参数时，请注意：

- 务必要保证修改后的配置正确且有效。
- 组件执行修改配置操作之后，需要手动重启相关受影响的组件以使配置生效，所以执行此操作时还需确认集群是否可以重启，另外重启集群之前建议关闭业务系统对 **HBase** 的读/写操作。

2.10.11 【HBase】修改 HDFS 上 HBase 存储目录【警告】

1. 操作事项

在 **HBase** 组件详情页面，在[配置]页签执行修改 **HBase** 在 **HDFS** 上的存储目录的操作（即修改配置项 `hbase-site` 中 `hbase.rootdir` 的属性值）。

2. 可能带来的风险

- 执行修改配置操作之后，需要手动重启相关受影响的组件以使配置生效，重启操作会导致业务系统暂时无法访问 HBase 集群。
- 修改 HBase 在 HDFS 上的存储目录可能导致业务数据丢失，HBase 组件无法启动。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 操作前请确认数据是否需要，如果需要请对数据进行备份。
- 建议修改后的数据目录仍在数据盘，变更的数据目录需确保 HBase 拥有读/写权限，且必须保证该目录为空。
- 执行此操作前，建议停止业务系统对 HBase 集群的读写操作。

2.10.12 【Hive】修改 Hive 元数据库相关配置【警告】

1. 操作事项

在 Hive 组件详情页面，在[配置]页签执行修改 Hive 元数据库相关配置的操作（即修改配置项 `javax.jdo.option.ConnectionDriverName`、`javax.jdo.option.ConnectionURL`、`javax.jdo.option.ConnectionUserName`、`ambari.hive.db.schema.name`、`hive_database_type` 的值）。

2. 可能带来的风险

- 若修改后的 Hive 元数据库参数有误，则会由于 Hive 连接元数据库失败造成 Hive 不可用。
- 若修改后的 Hive 元数据库参数正确，且该元数据库无其他集群的 Hive 正在使用，也会由于 Hive 切换元数据库造成原来的元数据在当前 Hive 中不可见。
- 若修改后的 Hive 元数据库参数正确，但是该元数据库已有其他集群的 Hive 正在使用，则会造成原来的元数据在当前 Hive 中不可见，还可能会造成同时使用该元数据库的 Hive 元数据混乱。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请务必确保修改后的元数据库配置当前 Hive 可以连接，并且对应的元数据库无其他集群的 Hive 正在使用。

2.10.13 【DLH】修改 DLHServer 元数据库相关配置【警告】

1. 操作事项

在 DLH 组件详情页面，在[配置]页签执行修改 DLHServer 元数据库相关配置的操作（即修改配置项 `javax.jdo.option.ConnectionDriverName`、`javax.jdo.option.ConnectionURL`、`javax.jdo.option.ConnectionUserName`、`ambari.hive.db.schema.name`、`hive_database_type` 的值）。

2. 可能带来的风险

- 若修改后的 DLH 元数据库参数有误，则会由于 DLH 连接元数据库失败造成 DLH 不可用。
- 若修改后的 DLH 元数据库参数正确，且该元数据库无其他集群的 DLH 正在使用，也会由于 DLH 切换元数据库造成原来的元数据在当前 DLH 中不可见。

- 若修改后的 DLH 元数据库参数正确，但是该元数据库已有其他集群的 DLH 正在使用，则会造成原来的元数据在当前 DLH 中不可见，同时还可能会造成同时使用该元数据库的 DLH 元数据混乱。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请务必确保修改后的元数据库配置当前 DLH 可以连接，并且对应的元数据库无其他集群的 DLH 正在使用。

2.10.14 【Impala】重启 Impala daemon 进程【一般】

1. 操作事项

在 Impala 组件详情页面，在[部署拓扑]页签执行重启 Impala daemon 进程的操作。

2. 可能带来的风险

由于 Impala 查询记录保存在 Impala daemon 缓存中，若重启该进程 Impala 查询记录相关信息将会丢失，且不可恢复。

3. 操作建议

请谨慎执行本操作。若依据实际业务需求，需要重启 Impala daemon 进程，请确认 Impala 之前的查询记录不再需要保留。

2.10.15 【Redis】修改 Redis 数据目录【警告】

1. 操作事项

在 Redis 组件详情页面，在[配置]页签执行修改 Redis 数据目录的操作（即修改配置项 `redis.dir` 的值）。

2. 可能带来的风险

该操作可能会导致数据丢失，Redis 启动失败，无法对外提供服务，进而对应业务中断。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 操作前请确认之前的数据是否需要，如果需要请对数据进行备份。
- 执行此操作前，建议停止业务系统对 Redis 集群的读写操作。
- 建议修改后的数据目录仍在数据盘，变更的数据目录需确保 Redis 拥有读/写权限，且必须保证该目录为空。

2.10.16 【Oozie】修改 Oozie 关键配置【警告】

1. 操作事项

在 Oozie 组件详情页面，在[配置]页签执行修改 Oozie 关键配置的操作。

2. 可能带来的风险

- `oozie-env` 相关配置项定义了 Oozie 组件启动过程中的环境变量、运行时用户及数据的保存目录等配置，若修改这些配置项的值可能会影响 Oozie 的稳定性或性能。

- 若修改 `oozie_user`、`oozie_tmp_dir`、`oozie_data_dir` 等关键配置项的值，可能导致 Oozie 无法正常使用。

3. 操作建议

请谨慎执行本操作。该操作需要组件专业人员执行，要求对组件配置信息十分了解。

2.10.17 【Flume】修改 Flume 关键配置【警告】

1. 操作事项

在 Flume 组件详情页面，在[配置]页签执行修改 Flume 关键配置的操作。

2. 可能带来的风险

- 若修改 `flume_user`、`flume_run_dir` 等关键配置项的值，可能导致 Flume 无法正常使用。
- 若修改 content 配置中 flume 的 JVM 信息，则可能会影响 Flume 的稳定性或性能。

3. 操作建议

请谨慎执行本操作。该操作需要组件专业人员执行，要求对组件配置信息十分了解。另外，`flume_user`、`flume_run_dir` 配置为服务内部使用，不建议修改。

2.10.18 【Storm】删除 Supervisor【警告】

1. 操作事项

在 Storm 组件详情页面，在[部署拓扑]页签执行删除 Supervisor 进程的操作。

2. 可能带来的风险

Supervisor 为 Storm 任务的计算节点，删除 Supervisor 会将降低 Storm 集群的计算资源，甚至可能会导致正在运行的任务中断。

3. 操作建议

请谨慎执行本操作。一般情况下，要求在集群部署前就评估流计算业务对计算资源的消耗，对 Supervisor 数量做好合理规划。集群在使用过程中，不建议对 Storm 执行缩容操作。

2.10.19 【Storm】修改 Storm 关键配置【警告】

1. 操作事项

在 Storm 组件详情页面，在[配置]页签执行修改 Storm 关键配置的操作。

2. 可能带来的风险

- 若修改 `storm_user`、`storm.local.dir`、`ranger` 相关配置等关键配置项的值，可能导致 Storm 无法正常使用或中间数据丢失。
- 若修改 Ranger 相关配置，则可能导致 Storm 组件权限功能无法实现，从而导致 Storm 任务异常中断或数据不一致等异常出现。

3. 操作建议

请谨慎执行本操作。该操作需要组件专业人员执行，要求对组件配置信息十分了解。

2.10.20 【Flink】修改 state 配置【警告】

1. 操作事项

在 Flink 组件详情页面，在[配置]页签执行修改 state 配置的操作（即修改 flink-conf 中 state 相关配置项 state.backend、state.backend.incremental、state.checkpoints.dir、state.savepoints.dir 的值）。

2. 可能带来的风险

Flink 的 state 相关配置定义了开启 checkpoint 机制的 Flink 作业使用的状态后端类型、状态储存路径等，所以若修改相关配置会造成已提交的开启了 checkpoint 机制的作业重启后，无法保证 exactly-once 语义，作业处理流程可能会出现丢失数据或者重复处理数据的现象。

3. 操作建议

请谨慎执行本操作。操作前请确认不存在正在运行的开启 checkpoint 机制的作业。

2.10.21 【Kafka】修改 Kafka 数据目录【警告】

1. 操作事项

在 Kafka 组件详情页面，在[配置]页签执行修改 Kafka 数据目录的操作（即修改配置项 log.dirs 的值）。

2. 可能带来的风险

配置项 log.dirs 的值定义了 Kafka 存储数据使用的数据目录，该目录除了包括原始数据外，还包括一些服务自身生成使用的索引数据和配置数据。所以：

- 若该数据目录内容被修改，可能会造成 Kafka 因无法识别已修改内容引发进程故障，导致进程无法启动。
- 若新添加的数据目录有其他非 Kafka 生成的文件，同样存在上述隐患。
- 该操作可能会导致数据丢失，Kafka 启动失败，无法对外提供服务，进而对应业务中断。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 操作前请确认之前的数据是否需要，如果需要请对数据进行备份。
- 执行此操作前，建议停止业务系统对 Kafka 集群的读写操作。
- 建议修改后的数据目录仍在数据盘。对于新增的数据目录，需确保 Kafka 拥有读/写权限，且必须保证该目录为空。

2.10.22 【Kafka】修改 Kafka 权限相关配置【警告】

1. 操作事项

在 Kafka 组件详情页面，在[配置]页签执行修改 Kafka 权限相关配置的操作（即修改配置项 sasl.enabled.mechanisms、sasl.mechanism.inter.broker.protocol、security.inter.broker.protocol sasl.kerberos.principal.to.local.rules、kafka_jaas_conf 的值）。

2. 可能带来的风险

权限相关配置定义了 Kafka 使用的认证机制，且多个配置项协同工作，若修改不当可能导致 Kafka 故障无法启动、业务请求无法通过权限认证等，造成对应业务中断。

3. 操作建议

请谨慎执行本操作。一般情况下，Kafka 权限相关配置不建议修改。若根据业务开发需要必须修改权限相关配置，请在充分理解各配置项作用或咨询紫光云技术支持后再进行修改。

2.10.23 【Kafka】修改 Kafka 监听配置【警告】

1. 操作事项

在 Kafka 组件详情页面，在[配置]页签执行修改或删除 Kafka 监听配置的操作（即修改配置项 listeners、advertised.listeners 的值）。

2. 可能带来的风险

配置项 listeners、advertised.listeners 的值定义了 Kafka 监听信息和对外发布到 ZooKeeper 对外提供服务的信息，若修改不当可能导致 Kafka 故障无法启动、业务无法连接到 Kafka。

3. 操作建议

请谨慎执行本操作。若根据业务开发需要必须修改监听相关配置，请在充分理解各配置项作用或咨询紫光云技术支持后再进行修改。

2.10.24 【Kafka】删除 Kafka Broker【警告】

1. 操作事项

在 Kafka 组件详情页面，在[部署拓扑]页签执行删除 Kafka Broker 进程的操作。

2. 可能带来的风险

删除 Kafka Broker 进程可能会造成该节点上 Topic 副本数据丢失。另外，若该 Topic 为单副本，则删除此 Kafka Broker 进程后会导致该 Topic 数据不可用，对应业务中断。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 操作前请确保剩余 Kafka Broker 进程数大于所有 Topic 的副本数的最大值。
- 或者根据 Topic 分布情况提前将待删除 Kafka Broker 节点上的数据迁移到其他不被删除 Kafka Broker 进程的节点上，然后再执行删除操作。

【说明】数据迁移请使用 Kafka 内置脚本工具 kafka-reassign-partitions.sh 完成。

2.10.25 【Elasticsearch】修改 Elasticsearch 数据目录【警告】

1. 操作事项

在 Elasticsearch 组件详情页面，在[配置]页签执行修改 Elasticsearch 数据目录的操作（即修改配置项 elasticsearch-config 中 path.data 的属性值）。

2. 可能带来的风险

- 修改 Elasticsearch 数据存储目录（新增数据目录/变更数据目录）可能导致 Elasticsearch 启动失败，无法对外提供服务，进而对应业务中断。
- 删除 Elasticsearch 数据存储目录可能会导致数据丢失，启动/加载索引失败，进而对应业务中断。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 操作前请确认之前的数据是否需要，如果需要请对数据进行备份。
- 执行此操作前，建议停止业务系统对 Elasticsearch 集群的读写操作。
- 建议修改后的数据目录仍在数据盘。对于新增的数据目录，需确保该目录仅供 Elasticsearch 组件使用并且未创建，配置修改之后 Elasticsearch 在重启时会自动新创建该目录并赋予其正确的权限。

2.10.26 【Solr】修改 Solr 数据目录【警告】

1. 操作事项

在 Solr 组件详情页面，在[配置]页签执行修改 Solr 数据目录的操作（即修改配置项 Solr-config-env 中 solr.data.home 的属性值）。

2. 可能带来的风险

该操作可能会导致数据丢失，Solr 启动失败，无法对外提供服务，进而对应业务中断。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 操作前请确认之前的数据是否需要，如果需要请对数据进行备份。
- 执行此操作前，建议停止业务系统对 Solr 集群的读写操作。
- 建议修改后的数据目录仍在数据盘，变更的数据目录需确保 Solr 拥有读/写权限，且必须保证该目录为空。

2.11 系统管理Web类高危操作介绍

2.11.1 禁用系统用户【警告】

1. 操作事项

通过系统超级管理员或安全管理员访问大数据平台管理系统时，在[系统/系统用户]页面对目标用户执行<禁用>操作。

2. 可能带来的风险

- 被禁用的用户将无法登录大数据平台管理系统，进而无法对大数据集群执行管理操作。
- 若被禁用的用户为安全管理员或系统超级管理员，将无法完成流程审批。
- 若禁用的用户为系统超级管理员，使用其他用户登录大数据平台管理系统时，将无法获取所有的管理操作权限，各用户拥有的操作权限由该用户绑定的角色决定。

3. 操作建议

- 系统超级管理员 admin 默认具有大数据平台管理系统的所有管理操作权限。在实际使用过程中，建议避免仅使用一个系统超级管理员 admin 来执行大数据平台的所有管理操作。若计划实现用户分权（比如：三权分立），可通过新建多个用户并分别绑定“安全管理员、运维管理员、审计管理员、普通用户”角色，实现用户权限的差异，然后使用这些用户可分别执行大数据平台的管理、运维、审计、使用等操作。

- 在实际使用过程中，若大数据平台使用了用户分权（比如：三权分立），则建议禁用系统超级管理员 **admin**。

2.11.2 修改系统用户密码【一般】

1. 操作事项

通过系统超级管理员或安全管理员访问大数据平台管理系统时，在[系统/系统用户]页面对目标用户执行<修改密码>操作。

2. 可能带来的风险

被修改密码的用户需要使用修改后的新密码进行登录，否则将无法登录到大数据平台管理系统。

3. 操作建议

请谨慎执行本操作。系统超级管理员或安全管理员修改其他用户的登录密码后，需要及时通知到对应用户。

2.11.3 修改系统用户授权【警告】

1. 操作事项

通过系统超级管理员或安全管理员访问大数据平台管理系统时，在[系统/系统用户]页面对目标用户执行<修改授权>操作。

2. 可能带来的风险

用户授权变更后，该用户原拥有的界面操作权限将会变更为新绑定角色对应的权限。

3. 操作建议

请谨慎执行本操作。修改用户授权操作会影响用户的界面操作权限，新用户第一次被修改授权后，一经投入使用建议不再随意修改授权。

2.11.4 删除系统用户【严重】

1. 操作事项

通过系统超级管理员或安全管理员访问大数据平台管理系统时，在[系统/系统用户]页面对目标用户执行<删除>操作。

2. 可能带来的风险

系统用户被删除后，该用户将不能登录管理系统，相关数据也被清除且不能恢复。

3. 操作建议

请谨慎执行本操作。因删除系统用户风险较高，强烈建议以禁用用户的方式代替删除用户操作。

2.11.5 修改 SSO 配置【严重】

1. 操作事项

在[系统/系统配置/安全设置/SSO 认证]页面，新增开启 SSO 认证进行相关配置或者修改 SSO 认证时进行配置修改的操作。

2. 可能带来的风险

SSO 认证若配置错误，则可能导致管理系统无法登录、用户信息无法同步。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请务必确保 SSO 配置修改正确。

2.11.6 断开 License Server 连接【警告】

1. 操作事项

系统已配置并连接到 License Server 获取授权后，在[系统/软件授权]页面单击<断开连接>按钮并确认，断开 License Server 连接。

2. 可能带来的风险

产品已获得的授权会被 License Server 回收，此时进入集群管理与数据工厂页面时会被强制跳转到软件授权页面，无法执行相关管理操作，但集群业务无影响。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请确认在管理系统不再需要执行相关管理操作。

2.11.7 系统升级【严重】

1. 操作事项

在[系统/系统升级]页面执行补丁包升级的操作。

2. 可能带来的风险

- 系统升级过程中，系统管理、集群管理、数据工厂服务可能不可用。
- 若升级失败，则可能会导致升级后管理系统和集群管理服务无法使用或数据工厂功能异常。

3. 操作建议

请谨慎执行本操作。若此操作不可避免，请注意：

- 执行系统升级操作前，请对系统和集群的关键数据进行备份。
- 执行升级操作前，请务必确保系统运行正常，且系统无业务请求。建议评估业务可中断时间窗口，在有效窗口期内进行。
- 系统升级过程中，系统和集群的服务均不可用。若系统升级失败，则可能会导致系统和集群使用异常，所以升级前请咨询紫光云技术支持。
- 不支持直接跨版本补丁包升级，只支持相邻版本之间升级。若版本间存在多个补丁包，升级前请确认补丁包版本的连续性，且补丁包均已正确上传指定路径。

3 数据工厂 Web 类高危操作介绍

3.1 简介

高风险的 Web 类操作只能由有资质、且经过培训的维护人员执行。如果操作不当，可能会导致业务中断、业务运行异常、重要文件被删除、所有配置被清除、用户无法登录、用户下线等现象发生。

Web 类高危操作包括但不限于本手册中介绍的内容，在进行 Web 类高危操作之前，请先了解可能带来的风险再进行操作。

3.2 集群管理高危操作介绍

3.2.1 删除大数据集群用户对数据工厂的影响【警告】

1. 操作事项

在[集群权限/用户管理]页面执行删除大数据集群用户的操作。

2. 可能带来的风险

该操作若涉及数据工厂数据源依赖的集群用户，则会导致数据工厂数据源不可用，从而导致依赖该数据源的库表管理、查询编辑、Notebook、工作流等模块的相关操作运行失败报错。若要恢复上述模块正常运行，需重新选择存在的用户进行数据源重连或选择其他集群进行数据源新建。

3. 操作建议

此操作风险高，请谨慎执行。若此操作不可避免，请确认当前环境下所依赖的集群是否使用该用户，待完成充分确认后再执行删除大数据集群用户的操作。

3.2.2 删除大数据集群对数据工厂的影响【严重】

1. 操作事项

在[集群管理/集群列表]页面执行删除集群的操作。

2. 可能带来的风险

该操作若涉及数据工厂数据源依赖的集群，则会一并删除数据工厂内使用该集群创建的数据源，从而导致依赖该数据源的库表管理、查询编辑、Notebook、工作流等模块的相关操作运行失败报错且无法恢复。

3. 操作建议

此操作风险极高，请谨慎执行。若此操作不可避免，请确认当前环境下集群是否可以删除（集群不再使用、集群无业务正在运行、集群内重要的数据已备份），且集群没有被数据工厂中任何数据源使用，待完成充分确认和数据备份后再执行删除操作。

3.3 数据探索高危操作介绍

3.3.1 删除 HDFS 文件【一般】

1. 操作事项

在[数据工厂/数据探索/文件管理]页面执行删除 HDFS 文件的操作。

2. 可能带来的风险

文件管理模块是以可视化方式，为用户提供管理 HDFS 文件系统功能，在 Web 页面上对 HDFS 文件的删除会同步删除 HDFS 组件文件，该操作可能会导致重要数据丢失或使用该文件的业务不可用。

3. 操作建议

请谨慎执行本操作。在对文件进行删除时，请先确保无重要业务使用该文件，如有重要数据请及时做好备份。

3.3.2 移动 HDFS 文件【一般】

1. 操作事项

在[数据工厂/数据探索/文件管理]页面执行移动 HDFS 文件的操作。

2. 可能带来的风险

文件管理模块是以可视化方式，为用户提供管理 HDFS 文件系统功能，在 Web 页面上对 HDFS 文件的移动会同步移动 HDFS 组件文件，该操作可能会导致重要数据丢失或使用该文件的业务不可用。

3. 操作建议

请谨慎执行本操作。在对文件进行移动时，请先确保无重要业务使用该文件。

3.3.3 重命名 HDFS 文件【一般】

1. 操作事项

在[数据工厂/数据探索/文件管理]页面执行重命名 HDFS 文件的操作。

2. 可能带来的风险

文件管理模块是以可视化方式，为用户提供管理 HDFS 文件系统功能，在 Web 页面上对 HDFS 文件的重命名会同步重命名 HDFS 组件文件，该操作可能会导致重要数据丢失或使用该文件的业务不可用。

3. 操作建议

请谨慎执行本操作。在对文件进行重命名时，请先确保无重要业务使用该文件。

3.3.4 删除库表【严重】

1. 操作事项

在[数据工厂/数据探索/库表管理]页面执行删除数据库（数据表或视图）的操作。

2. 可能带来的风险

库表管理模块为 Hive、Impala、SparkSQL、Phoenix 等数据源提供库表管理功能，删除操作会删除 Hive、Impala、SparkSQL、Phoenix 等数据源存储的真实数据文件，导致数据不可修复。

3. 操作建议

请谨慎执行本操作。如需删除库表，请先确认是否为业务使用数据，并对重要数据及时做好备份。

3.3.5 删除运行环境【一般】

1. 操作事项

在[数据工厂/数据探索/Notebook]页面执行删除运行环境的操作。

2. 可能带来的风险

通过 Notebook 直接和 HBase、Spark 进行交互。删除 Notebook 中运行环境可能会导致使用该运行环境的笔记不可用。

3. 操作建议

请谨慎执行本操作。如需删除运行环境，请先确保无重要业务笔记使用该运行环境。

3.3.6 编辑运行环境【一般】

1. 操作事项

在[数据工厂/数据探索/Notebook]页面执行修改运行环境分组的操作。

2. 可能带来的风险

修改 Notebook 运行环境分组可能会导致使用该运行环境的笔记不可用。

3. 操作建议

请谨慎执行本操作。如需修改运行环境分组，请先确保无重要业务笔记使用该运行环境。

3.3.7 重启运行环境【一般】

1. 操作事项

在[数据工厂/数据探索/Notebook]页面执行重启运行环境的操作。

2. 可能带来的风险

重启 Notebook 运行环境可能会导致使用该运行环境且正在运行的笔记中断报错。

3. 操作建议

请谨慎执行本操作。如需重启运行环境，请先确保无正在运行的重要业务笔记使用该运行环境。

3.3.8 删除笔记【一般】

1. 操作事项

在[数据工厂/数据探索/Notebook]页面执行删除笔记的操作。

2. 可能带来的风险

删除 Notebook 笔记不可恢复，删除后可能会导致重要业务笔记丢失。

3. 操作建议

请谨慎执行本操作。如需删除笔记，请先确保该笔记非重要业务笔记。

3.4 数据开发高危操作介绍

3.4.1 删除工作流【警告】

1. 操作事项

在[数据工厂/数据开发/工作流设计]页面执行删除工作流的操作。

2. 可能带来的风险

workflows 设计是按照时间和依赖关系进行任务调度管理，删除 workflow 可能导致正在调度的任务中断且不可恢复。

3. 操作建议

请谨慎执行本操作。如需删除 workflow，请先确保该 workflow 无重要业务且任务处于停止状态。

3.5 连接管理高危操作介绍

3.5.1 删除数据源【严重】

1. 操作事项

在[数据工厂/连接管理/数据源连接]页面执行删除数据源的操作。

2. 可能带来的风险

数据源供数据探索服务使用，删除数据源可能导致数据探索模块的相关文件管理不可用、相关运行环境及使用该运行环境的笔记不可用、数据开发模块相关工作流不可用。

3. 操作建议

请谨慎执行本操作。如需删除数据源，请先确保暂无其他重要业务使用该数据源。

3.5.2 删除主机连接【警告】

1. 操作事项

在[数据工厂/连接管理/主机连接]页面执行删除主机连接的操作。

2. 可能带来的风险

主机连接实现对外部源的有效管理，供数据开发使用。删除主机连接可能导致使用该主机的工作流不可用。

3. 操作建议

请谨慎执行本操作。如需删除主机连接，请先确保该主机无重要 workflow 业务使用。

4 后台类高危操作介绍

4.1 简介

高危险的后台类操作只能由有资质、且经过培训的维护人员执行。如果对此类命令操作不当，可能会导致集群异常、组件异常、业务中断、业务运行异常、重要文件被删除、所有配置被清除、用户无法登录、用户下线等现象发生。

后台类高危操作包括但不限于本手册中介绍的内容，在对高危命令进行操作之前，请先了解可能带来的风险再进行操作。

4.2 命令行高危操作介绍

功能分类	操作项	操作举例	可能带来的风险	操作建议
集群管理	修改主机名	<ul style="list-style-type: none"> 通过 vi、vim 命令修改/etc/hosts 文件 通过 hostnamectl 命令修改主机名 	执行此操作将修改集群内主机名的名称，此操作将导致含有原主机名的配置文件失效，从而导致主机业务不可用	请勿执行本操作
系统管理	修改系统时间	date -s	执行此操作将修改主机节点的系统时间，若集群中各节点的时间不同步，将导致集群业务异常	请谨慎执行本操作，操作前请务必评估所有可能产生的风险
系统管理	节点重启/关机	<ul style="list-style-type: none"> reboot shutdown poweroff 	执行此操作将使主机节点重启或关机，将导致该节点不可用，节点上的业务将中断	请谨慎执行本操作，操作前请务必评估所有可能产生的风险
系统管理	结束任务进程	<ul style="list-style-type: none"> kill killall 	执行此操作将结束执行中的任务进程。若误操作，将导致对应的功能或业务异常	请谨慎执行本操作，操作前请务必评估所有可能产生的风险
系统管理	停止crond守护进程	<ul style="list-style-type: none"> service crond stop systemctl stop crond.service 	执行此操作将停止系统的守护进程crond，crond是Linux用来定期执行程序命令，停止该进程会造成操作系统异常，从而导致集群业务异常	请勿执行本操作
系统管理	远程下载脚本	wget	执行此操作可以远程下载脚本，因来源未知的脚本的安全性无法保证，存在风险，所以运行此类脚本可能导致无法预知的问题	请谨慎执行本操作，操作前请明确脚本来源，务必确保脚本安全可靠无风险
系统管理	运行来源未知的脚本	无	因来源未知的脚本的安全性无法保证，存在风险，且脚本代码中可能隐藏高危操作指令，所以运行此类脚本可能导致无法预知的问题	请谨慎执行本操作，操作前请明确脚本来源，务必确保脚本安全可靠，并评估运行脚本可能产生的风险
磁盘管理	卸载分区磁盘	umount	执行此操作可以卸载分区磁盘。若误操作，可能破坏该磁盘设备上的文件系统和已有数据，从而导致相关业务异常或数据丢失	请勿执行本操作

功能分类	操作项	操作举例	可能带来的风险	操作建议
网络配置	修改交换机配置	lacp period short	执行此操作可以修改交换机配置。若误操作，将导致集群的网络异常，从而导致集群业务异常	请谨慎执行本操作，操作前请提前做好集群组网规划，并评估所有可能产生的风险
网络配置	修改节点的网络配置	<ul style="list-style-type: none"> • ifdown • ifup • route 	执行此操作可以修改节点的网络配置。若误操作，将导致节点的网络异常，从而导致集群业务异常	请谨慎执行本操作，操作前请提前做好集群组网规划，并评估所有可能产生的风险
文件管理	修改文件或目录的所属群组	chgrp	执行此操作可以修改文件或目录的所属群组。若误操作，可能导致集群业务异常	请谨慎执行本操作，操作前请评估所有可能产生的风险
文件管理	修改文件或目录的拥有者	chown	执行此操作可以修改文件或目录的拥有者。若误操作，可能导致集群业务异常	请谨慎执行本操作，操作前请评估所有可能产生的风险
文件管理	修改文件或目录的权限	chmod	执行此操作可以修改文件或目录的权限。若误操作，可能导致集群业务异常	请谨慎执行本操作，操作前请评估所有可能产生的风险
文件管理	递归删除目录	rm -rf	执行此操作将删除目录中的子目录和文件，并忽略提示信息。若误操作，可能导致重要文件或数据被删除	请谨慎执行本操作，操作前请务必确认该目录及其所有子目录和文件均不再需要
文件管理	强制拷贝或移动文件	<ul style="list-style-type: none"> • cp -rf • mv -f 	执行此操作将强制拷贝或移动路径下的文件至新路径，并忽略提示信息。若误操作，可能导致路径下原有文件被覆盖	请谨慎执行本操作，操作前请务必确认待移动的文件与新路径下的原有文件不存在重名
文件管理	操作设备文件描述符	<ul style="list-style-type: none"> • echo • dd 	执行此操作可以修改硬盘设备映射到/dev/目录下的文件。若误操作，可能破坏磁盘设备上的文件系统和已有数据，从而导致相关业务异常或数据丢失	请勿执行本操作
文件管理	修改文件	通过vi、vim命令修改文件	执行此操作可以修改文件内容。若误操作，可能导致文件内容错误，从而导致无法预知的问题	请谨慎执行本操作，操作前请评估所有可能产生的风险
用户管理	删除操作系统用户	userdel	执行此操作可以删除操作系统内部用户。若误操作，将导致该	请勿执行本操作

功能分类	操作项	操作举例	可能带来的风险	操作建议
			用户的服务异常	

4.3 大数据平台（后台类）高危操作介绍

4.3.1 修改 NTP 服务配置【严重】

1. 操作事项

后台登录集群内的某主机节点，在/etc/chrony.conf 下修改 NTP 服务配置。

2. 可能带来的风险

- 如果修改为不配置 NTP 同步，大数据集群中各主机节点之间将不再进行时间同步，这样可能导致集群内各主机之间的时间差越来越大，从而导致业务中断。
- 如果变更 NTP 服务器，则可能因为新的 NTP 服务器与大数据集群时间差异较大，导致集群系统时间跳变，从而导致集群异常，影响用户的大数据业务。

3. 操作建议

请谨慎执行本操作。操作前请关注如下事项：

- 建议为大数据集群配置 NTP 同步。
- 变更 NTP 服务器时，请确保大数据集群与更换后 NTP 服务器的时间差尽量小。

4.3.2 修改 YUM 源配置【警告】

1. 操作事项

后台登录集群内的某主机节点，在/etc/yum.repos.d/下修改 YUM 源配置。

2. 可能带来的风险

若 YUM 源配置修改不当，则可能导致集群 YUM 源不可用，进而引起集群扩容、安装组件等异常。

3. 操作建议

请谨慎执行本操作。如非必要情况，不建议对 YUM 源配置进行修改。若此操作不可避免，要求对 YUM 源只能添加不能删改，且需确认添加内容与原内容无冲突 RPM。

4.3.3 root 配置 SSH 免密【严重】

1. 操作事项

后台登录集群内的某主机节点，为 root 用户配置 SSH 免密访问。

2. 可能带来的风险

若集群内主机 root 用户配置免密，则分布式集群内部访问及操作不会再鉴别密码，易产生误删数据或文件的风险。

3. 操作建议

此操作风险极高，禁止执行。

4.3.4 后台修改主机 root 密码【严重】

1. 操作事项

后台登录集群内的某主机节点，通过 `passwd` 命令直接修改 `root` 用户的密码。

2. 可能带来的风险

在后台私自修改集群内主机节点 `root` 用户的密码会导致集群功能异常，若修改后的密码丢失将导致集群无法访问和使用。

3. 操作建议

此操作禁止执行。严禁在后台修改主机 `root` 用户的密码，当集群部署完成后可以在 `Web` 管理界面的执行修改 `root` 密码的操作，且修改后的密码需妥善保管。

4.3.5 修改主机 IP【严重】

1. 操作事项

在后台修改主机 `IP` 地址，包括修改集群内主机的 `IP` 地址或管理节点的 `IP` 地址。

2. 可能带来的风险

- 修改集群内主机的 `IP` 地址，将导致集群不可用，业务中断。
- 修改管理节点的 `IP` 地址，可能导致管理节点无法登录，无法进行集群管理。

另外，集群内主机的 `IP` 地址或管理节点的 `IP` 地址一旦被修改，即使再还原也可能由于不可控因素导致集群或者管理节点无法恢复。

3. 操作建议

此操作禁止执行。

4.3.6 修改主机名【严重】

1. 操作事项

后台登录集群内的某主机节点，通过修改 `/etc/hosts` 文件或使用 `hostname` 命令修改集群主机名。

2. 可能带来的风险

执行此操作将修改集群内主机名的名称，导致含有原主机名的配置文件失效，从而导致主机业务不可用。

3. 操作建议

此操作禁止执行。

4.3.7 对 HDFS 文件管理进行编辑【严重】

1. 操作事项

后台登录集群内的某主机节点，通过 `hdfs` 命令修改或删除 `HDFS` 相关文件、目录及其权限。

2. 可能带来的风险

HDFS 文件存储了 HBase、Hive、Spark、YARN 等组件的元数据、中间数据和业务数据。在后台通过 `hdfs` 命令修改或删除 HDFS 相关文件、目录及其权限，若这些数据被误操作，将导致组件异常、数据丢失或业务中断等故障。

3. 操作建议

请谨慎执行本操作，操作前请务必评估所有可能产生的风险。

4. 操作步骤

务必谨慎执行的 `hdfs` 文件管理命令包括：`hdfs dfs -rm -f`、`hdfs dfs -chmod`、`hdfs dfs -chgrp`、`hdfs dfs -chown` 等。

4.3.8 对大数据集群的元数据库进行编辑【严重】

1. 操作事项

后台登录集群内的某主机节点，通过命令行或使用第三方数据库客户端软件连接大数据集群的元数据库 MySQL，对数据库中的表执行删除或修改等操作。

2. 可能带来的风险

在大数据集群的元数据库 MySQL 的文件中存储了 HBase、Hive、Spark、YARN 等组件的元数据、中间数据和业务数据，若这些数据被误操作，将导致集群及组件异常、数据丢失或业务中断等故障。

3. 操作建议

请谨慎执行本操作，操作前请务必评估所有可能产生的风险。

4. 操作步骤

务必谨慎执行的数据库命令包括：`drop database`、`delete`、`truncate table`、`update` 等。

4.3.9 删除大数据平台管理系统的内置 MySQL 数据库【严重】

1. 操作事项

后台登录集群内的某主机节点，通过命令行或使用第三方数据库客户端软件连接大数据平台管理系统的内置 MySQL 数据库，对数据库执行删除操作。

2. 可能带来的风险

该操作会导致大数据平台管理系统不可用，造成的影响不可恢复。

3. 操作建议

此操作禁止执行。

4.3.10 修改主机防火墙配置【警告】

1. 操作事项

后台登录集群内的某主机节点，修改主机的防火墙配置。

2. 可能带来的风险

该操作可能会导致集群或者节点异常，影响集群的正常使用。

3. 操作建议

请谨慎执行本操作。

4.3.11 修改大数据集群中组件或进程相关端口号【警告】

1. 操作事项

后台登录集群内的某主机节点，通过命令修改大数据平台中组件或进程相关端口号。

2. 可能带来的风险

该操作可能会导致组件或进程异常，无法访问。

3. 操作建议

请谨慎执行本操作。该操作需要组件专业人员执行，要求对组件端口配置信息十分了解。

4.3.12 未经适配的第三方软件与大数据集群混合部署【严重】

1. 操作事项

在部署大数据集群的主机节点上，部署未经适配的第三方软件。

2. 可能带来的风险

未经适配的第三方软件与大数据集群混合部署，可能导致与大数据集群进程或端口冲突，出现资源抢占等异常，导致集群不可用。

3. 操作建议

此操作禁止执行。

4.3.13 执行安装管理系统操作前修改配置文件 default.json【严重】

1. 操作事项

执行安装管理系统操作前，通过后台进入安装目录，修改安装脚本配置文件 default.json。

2. 可能带来的风险

default.json 文件中包含多个配置项，若修改不当则可能导致端口冲突、数据库初始化失败、服务部署启动失败等故障。

3. 操作建议

此操作风险极高，请谨慎执行。若此操作不可避免，请务必先检查端口是否占用、命名是否规范等。如果遇到安装失败，请执行卸载操作后，然后重新修改配置后再进行安装。

4.3.14 卸载大数据平台管理系统【严重】

1. 操作事项

后台登录安装大数据平台管理系统的管理节点，进入安装目录下执行 `sh uninstall.sh` 卸载脚本，卸载大数据平台（含数据工厂）。

2. 可能带来的风险

- 若大数据平台管理系统服务异常，对于非融合部署，该操作仅会卸载大数据平台管理系统，而大数据平台会完全失去 Web 管理和监控运维能力，但是集群运行不受影响，用户业务也不受影响。
- 若大数据平台管理系统服务异常，对于融合部署，该操作不仅会卸载大数据平台管理系统，还会导致大数据集群不可用，用户业务异常。
- 若大数据平台管理系统服务运行正常，仅当大数据平台中删除所有集群后，该操作才可执行，此时大数据平台管理系统会被卸载，管理系统上相关数据被删除。

3. 操作建议

此操作风险极高，请勿随意执行。若此操作不可避免，请先卸载大数据平台内的所有大数据集群，并且确保大数据平台管理系统服务正常后再执行卸载操作。

4.3.15 安装管理系统操作后，单独安装数据工厂时修改配置文件 default.json 【严重】

1. 操作事项

安装大数据平台管理系统操作后，单独安装数据工厂时，通过后台进入安装目录，修改安装脚本配置文件 default.json。

2. 可能带来的风险

default.json 文件中包含多个配置项，单独安装数据工厂时仅需要将 df.install 的值由 false 改为 true。若修改了其他配置项，则可能导致故障，比如：

- 若修改 Mysql 的相关配置，可能导致数据工厂数据库连接异常。
- 若修改 Kong、Keystone、Haproxy 的相关配置，可能引起数据工厂的请求访问不通。
- 若修改 Webserver 的相关配置，可能引起管理页面访问异常。
- 若修改 Ambari、LDAP 的相关配置，可能导致数据源新建异常。

3. 操作建议

请谨慎执行本操作，此种情况下仅需要将 df.install 的值由 false 改为 true，不要随意修改其他配置项的值。

4.3.16 单独卸载数据工厂 【严重】

1. 操作事项

后台登录大数据平台的管理节点，进入安装目录下执行 sh uninstall_df.sh 卸载脚本，单独卸载数据工厂。

2. 可能带来的风险

该操作会导致数据工厂不可用，数据工厂内数据源、Notebook 运行环境、笔记和工作流数据均会丢失。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请确认不需要再使用数据工厂，且数据工厂相关的数据均已不再使用或已完成备份。

4.3.17 执行大数据平台管理系统版本升级前修改配置文件 upgrade.json 【严重】

1. 操作事项

后台登录大数据平台的管理节点，进入安装目录，修改配置文件 `upgrade.json` 中支持的版本号或对应的大数据平台管理系统的虚 IP 配置，然后执行 `sh upgrade.sh` 版本升级脚本。

2. 可能带来的风险

- 系统版本升级过程中，大数据平台管理系统服务不可用，页面将无法访问；升级完成后，若升级成功，则对应的大数据平台管理系统将恢复使用。
- 若对不支持的版本进行升级，则会升级失败，同时可能导致升级失败后整个大数据平台管理系统服务不可用且不可恢复。
- 若对要升级的大数据平台管理系统的访问虚 IP 配置错误，则可能导致升级失败或虚 IP 对应的大数据平台管理系统服务不可用且不可恢复。

3. 操作建议

此操作风险极高，请勿随意执行。若此操作不可避免，请注意：

- 务必确认已安装的版本与升级所支持的版本是否匹配。若版本匹配，即已安装版本支持升级，请认真核对已安装大数据平台管理系统访问虚 IP 并进行配置。
- 严禁修改配置文件 `upgrade.json` 中支持的版本号配置。

4.4 组件管理（后台类）高危操作介绍

4.4.1 【组件公共】未开启 Kerberos 的环境，使用组件超级用户执行组件操作 【严重】

1. 操作事项

由于 `hdfs` 用户是 HDFS 组件的超级用户，因此该用户可以对 HDFS 组件执行任意操作（比如：对文件进行增/删）。

未开启 Kerberos 的集群环境中，在集群内直接使用 `hdfs` 用户操作 HDFS 组件；或者在 HDFS Client 客户端，在后台通过 `useradd hdfs` 命令添加了 `hdfs` 用户，然后使用该用户操作 HDFS 组件。

2. 可能带来的风险

- 若使用 `hdfs` 用户进行异常操作，则会造成 HDFS 组件异常或数据不安全等问题。
【严重】同样的，使用 `hbase`、`yarn`、`hive` 等组件超级用户对对应的组件进行异常操作，也会造成组件异常等问题。

3. 操作建议

由于未开启 Kerberos 环境中，无法检验用户身份，建议不要在后台使用组件超级用户（`hdfs`、`oozie`、`mapred`、`yarn`、`hive`、`spark`、`hbase`、`atlas`、`kafka`、`elasticsearch`、`solr`）执行组件操作。

4.4.2 【HDFS】删除 HDFS 数据目录【严重】

1. 操作事项

在后台删除 HDFS 数据目录，其中数据目录为“Web 管理界面的大数据集群中，在 HDFS 组件详情页面的 [配置] 页签 `hdfs-site` 的 `dfs.namenode.name.dir`、`dfs.datanode.data.dir`、`dfs.namenode.checkpoint.dir` 和 `dfs.journalnode.edits.dir` 配置项中定义的数据目录”

2. 可能带来的风险

HDFS 包括多个数据目录配置项，各个目录配置信息以及可能导致的风险如[表 4-1](#) 所示。

表4-1 HDFS 数据目录

数据目录配置项	存储数据	修改可能导致的风险
<code>dfs.namenode.name.dir</code>	NameNode元数据信息的目录	元数据记录了HDFS数据存储的实际地址，删除元数据信息的目录可能会导致数据全部丢失
<code>dfs.datanode.data.dir</code>	HDFS实际存储数据的目录	删除数据目录会导致数据丢失
<code>dfs.namenode.checkpoint.dir</code>	HDFS未开启HA时，SecondaryNameNode存储镜像文件的目录	存储元数据的快照数据，删除可能导致元数据无法恢复
<code>dfs.journalnode.edits.dir</code>	HDFS开启HA时，Journalnode存储操作日志信息目录	用于同步主备NameNode之间的元数据信息，删除可能会导致部分最新的元数据丢失或导致组件异常

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.3 【HDFS】删除 HDFS 中存储的数据【严重】

1. 操作事项

在后台通过命令（如：`hadoop fs -rm -r /path`）删除 HDFS 中存储数据的目录。

2. 可能带来的风险

这些目录可能是用户或相关组件（如：HBase、Hive、DLH 等）用于存储数据的路径，删除会导致用户数据丢失或组件使用异常。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请在删除前务必仔细确认目录中存储数据的用途，减少误操作。

4.4.4 【ZooKeeper】业务代码连接 Zookeeper【严重】

1. 操作事项

在后台，把不涉及大数据集群中组件的客户业务代码，连接到大数据集群上的 Zookeeper。

2. 可能带来的风险

客户业务代码如果存在 bug，则可能出现 ZooKeeper 连接泄露，导致 ZooKeeper 的连接数不足，从而影响大数据集群中其他依赖 ZooKeeper 的大数据组件不可用。

3. 操作建议

此操作风险极高，禁止执行。大数据集群中的 ZooKeeper 仅供集群内部组件使用，外部业务禁止使用该 Zookeeper。

4.4.5 【Elasticsearch】删除 Elasticsearch 数据目录【严重】

1. 操作事项

在后台删除 Elasticsearch 数据目录，其中数据目录为“Web 管理界面的大数据集群中，在 Elasticsearch 组件详情页面的[配置]页签配置项 elasticsearch-config 中 path.data 的属性值定义的数据目录”。

2. 可能带来的风险

该操作会导致数据丢失，Elasticsearch 无法启动，无法对外提供服务，进而导致业务中断。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 操作前请确认之前的数据是否需要，如果需要请对数据进行备份。
- 执行此操作前，建议停止业务系统对 Elasticsearch 集群的读写操作。

4.4.6 【Elasticsearch】删除 Elasticsearch 集群索引【严重】

1. 操作事项

在后台通过 API 删除 Elasticsearch 集群索引。

2. 可能带来的风险

Elasticsearch 索引数据不可轻易删除，否则会导致该 Index 相关业务出现异常。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 删除之前务必确认待删除索引是否还需要，若需要请进行索引数据备份。
- 执行此操作前，请停止对该索引的读/写操作，并且确保业务可中断及修改值正确有效。
- 如果索引数据量较大，删除过程会比较久，执行此操作前建议添加删除索引 API 的超时时间设置或者重试。

4.4.7 【Solr】删除 Solr 数据目录【严重】

1. 操作事项

在后台删除 Solr 数据目录，其中数据目录为“Web 管理界面的大数据集群中，在 Solr 组件详情页面的[配置]页签配置项 Solr-config-env 中 solr.data.home 的属性值定义的数据目录”。

2. 可能带来的风险

该操作会导致数据丢失，Solr 无法启动，无法对外提供服务，进而导致业务中断。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 操作前请确认之前的数据是否需要，如果需要请对数据进行备份。
- 执行此操作前，建议停止业务系统对 Solr 集群的读写操作。

4.4.8 【Solr】删除 Solr 集群索引【严重】

1. 操作事项

在后台通过 API 删除 Solr 集群索引。

2. 可能带来的风险

Solr 索引数据不可轻易删除，否则会导致该 Collection 相关业务出现异常。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 删除之前务必确认待删除索引是否还需要，若需要请进行索引数据备份。
- 执行此操作前，请停止对该索引的读/写操作，并且确保业务可中断及修改值正确有效。
- 如果索引数据量较大，删除过程会比较久，执行此操作前建议添加删除索引 API 的超时时间设置或者重试。

4.4.9 【Kafka】删除 Kafka 数据目录【严重】

1. 操作事项

在后台删除 Kafka 数据目录，其中数据目录为“Web 管理界面的大数据集群中，在 Kafka 组件详情页面的[配置]页签 kafka-broker 的 log.dirs 配置项中定义的数据目录”。

2. 可能带来的风险

- 删除 Kafka 正在使用的数据目录会造成 Topic 数据丢失，对应业务将无法读取已删除数据，业务中断。
- 如果 Topic 为单副本，则数据丢失后无法恢复。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 不建议手动删除 Kafka 数据，极易造成数据丢失或 Kafka 组件无法使用。
- 如需进行 Kafka 历史数据进行清理操作，建议通过缩短 Kafka 数据保留周期的配置来完成，详情请咨询紫光云技术支持。

4.4.10 【Redis】删除 Redis 数据目录【严重】

1. 操作事项

在后台删除 Redis 数据目录，其中数据目录为“Web 管理界面的大数据集群中，在 Redis 组件详情页面的[配置]页签 redis-site 的 redis.dir 配置项中定义的数据目录”。

2. 可能带来的风险

删除 Redis 正在使用的数据目录会造成数据丢失，且数据无法恢复，对应业务将无法读取已删除数据，业务中断。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 不建议手动删除 Redis 数据，极易造成数据丢失或 Redis 组件无法使用。

4.4.11 【DLH】DLH 库表删除【严重】

1. 操作事项

在后台通过 DLH SQL 命令行执行库表删除的操作。

2. 可能带来的风险

- 删除库操作会将库以及该库下所有表的元数据删除，该库对应存储介质上真实文件目录也会被删除，造成整库生产数据丢失。
- 单独删除管理表会删除该表元数据以及真实数据文件，造成该表生产数据丢失。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请务必确保库/表数据已不再需要，待完成充分确认后再执行删除操作。

【说明】针对管理表操作，如果有删除表保留真实文件的需求，可以使用外部表，删除外部表只会删除对应的元数据而不会删除真实数据文件。

4.4.12 【DLH】DLH 表清空【严重】

1. 操作事项

在后台通过 DLH SQL 命令行执行表的 truncate 清空的操作。

2. 可能带来的风险

表清空操作会对表数据文件进行清除，执行该操作后数据不可恢复，将造成生产数据丢失。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请务必确保表数据已不再需要，待完成充分确认后再执行清空操作。

4.4.13 【DLH】删除 DLH 库表在 HDFS 上的存储路径【严重】

1. 操作事项

在后台直接通过 HDFS 命令删除 DLH 库表在 HDFS 上对应的存储路径。

2. 可能带来的风险

该操作会导致生产数据丢失，且无法恢复。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.14 【DLH】删除 DLH 元数据库表【严重】

1. 操作事项

在后台通过 MySQL 客户端执行 DLH 元数据库表删除的操作。

2. 可能带来的风险

该操作会导致 DLH 默认数据源的库表元数据丢失。

3. 操作建议

此操作风险极高，如无必要，请不要执行此操作。

4.4.15 【HBase】修改或删除 Phoenix 系统表或系统表数据 (SYSTEM.*)【严重】

1. 操作事项

在后台通过 Phoenix 客户端登录集群内主机，修改或删除 Phoenix 系统表或系统表数据 (SYSTEM.*)。

2. 可能带来的风险

修改或删除 Phoenix 系统表将导致元数据不可恢复，无法识别库表数据，无法正常进行业务操作。

3. 操作建议

此操作禁止执行。

4.4.16 【HBase】修改或删除 HBase 元数据表【严重】

1. 操作事项

在后台通过 HBase 客户端登录集群内主机，修改或删除 HBase 元数据表。

2. 可能带来的风险

修改或删除 HBase 元数据表将导致数据不可恢复，无法识别库表数据，无法正常进行业务操作。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.17 【HBase】删除 HBase 在 HDFS 上的存储路径【严重】

1. 操作事项

在后台直接通过 HDFS 命令删除 HBase 在 HDFS 上对应的存储路径。

2. 可能带来的风险

该操作会导致 HBase 业务数据丢失，HBase 组件启动失败，从而导致业务中断。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 操作前请确认数据是否需要，如果需要请对数据进行备份。
- 执行此操作前，建议停止业务系统对 HBase 集群的读写操作。

- 若 HDFS 开启了回收站功能，可以在设置的保留时间段之内进行误删恢复，但若超过保留时间则误删数据无法恢复。

4.4.18 【Hive】Hive 库表删除【严重】

1. 操作事项

在后台直接通过 Hive SQL 命令行执行库表删除的操作。

2. 可能带来的风险

- 删除库操作会将库以及该库下所有表的元数据删除，该库对应存储介质上真实文件目录也会被删除，造成整库生产数据丢失。
- 单独删除管理表会删除该表元数据以及真实数据文件，造成该表生产数据丢失。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请务必确保库/表数据已不再需要，待完成充分确认后再执行删除操作。

【说明】针对管理表操作，如果有删除表保留真实文件的需求，可以使用外部表，删除外部表只会删除对应的元数据而不会删除真实数据文件。

4.4.19 【Hive】Hive 表清空【严重】

1. 操作事项

在后台通过 Hive SQL 命令行执行表的 truncate 清空的操作。

2. 可能带来的风险

表清空操作会对表数据文件进行清除，执行该操作后数据不可恢复，将造成生产数据丢失。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请务必确保表数据已不再需要，待完成充分确认后再执行清空操作。

4.4.20 【Hive】删除 Hive 库表在 HDFS 上的存储路径【严重】

1. 操作事项

在后台直接通过 HDFS 命令删除 Hive 库表在 HDFS 上对应的存储路径。

2. 可能带来的风险

该操作会导致生产数据丢失，且无法恢复。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.21 【Hive】删除 Hive 元数据库表【严重】

1. 操作事项

在后台通过 MySQL 客户端直接删除 Hive 使用的元数据库表。

2. 可能带来的风险

该操作会导致 Hive 组件现有的库表元数据丢失，Hive 不可用。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.22 【Impala】删除 Impala 元数据库表【严重】

1. 操作事项

在后台通过 MySQL 客户端直接删除 Impala 元数据库表，Impala 和 Hive 共用一个数据库。

2. 可能带来的风险

- 该操作会导致元数据丢失，Impala 组件不可用，业务数据无法访问。
- 该操作删除的 Impala 元数据库表，一旦删除则无法恢复。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.23 【Impala】修改 Impala 元数据库表数据【严重】

1. 操作事项

在后台通过 MySQL 客户端修改 Impala 元数据库表数据。

2. 可能带来的风险

- 该操作会导致元数据丢失，Impala 组件不可用，业务数据无法访问。
- 该操作修改的 Impala 元数据库表数据，一旦修改则无法恢复。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.24 【Impala】删除 Impala 在 HDFS 上的存储路径【严重】

1. 操作事项

在后台直接通过 HDFS 命令删除 Impala 在 HDFS 上对应的存储路径。

【说明】Impala 的数据目录包括两种情况：

- 如果为内部表，Impala 数据目录和 Hive 共用，其中数据目录为“Web 管理界面的大数据集群中，在 Hive 组件详情页面的[配置]页签配置项 hive-site 的 hive.metastore.warehouse.dir 属性值定义的数据目录”。
- 如果为外部表，则数据目录为创建表时通过[LOCATION 'hdfs_path']指定的目录。

2. 可能带来的风险

该操作会导致数据丢失，对应业务将无法读取已删除数据，从而导致业务中断。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请注意：

- 操作前请确认数据是否需要，如果需要请对数据进行备份。

- 执行此操作前，建议停止业务系统对 Impala 集群的读写操作。
- 若 HDFS 开启了回收站功能，可以在设置的保留时间段之内进行误删恢复，但若超过保留时间则误删数据无法恢复。

4.4.25 【Spark】Spark 库表删除【严重】

1. 操作事项

在后台通过 Spark SQL 命令行执行库表删除的操作。

2. 可能带来的风险

- 删除库操作会将库以及该库下所有表的元数据删除，该库对应存储介质上真实文件目录也会被删除，造成整库生产数据丢失。
- 单独删除管理表会删除该表元数据以及真实数据文件，造成该表生产数据丢失。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请务必确保库/表数据已不再需要，待完成充分确认后再执行删除操作。

4.4.26 【Spark】Spark 表清空【严重】

1. 操作事项

在后台通过 Spark SQL 命令行执行表的 truncate 清空的操作。

2. 可能带来的风险

表清空操作会对表数据文件进行清除，执行该操作后数据不可恢复，将造成生产数据丢失。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。若此操作不可避免，请务必确保表数据已不再需要，待完成充分确认后再执行清空操作。

4.4.27 【Spark】删除 Spark 库表在 HDFS 上存储路径【严重】

1. 操作事项

在后台直接通过 HDFS 命令删除 Spark 库表在 HDFS 上对应的存储路径。

2. 可能带来的风险

该操作会导致生产数据丢失，且无法恢复。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。

4.4.28 【Spark】删除 Spark 元数据库表【严重】

1. 操作事项

在后台通过 MySQL 客户端直接删除 Spark 使用的元数据库表。

2. 可能带来的风险

该操作会导致 Spark 组件现有的库表元数据丢失，Spark 将不可用。

3. 操作建议

此操作风险极高，请勿随意执行，执行前请咨询紫光云技术支持。