

Kubernetes 集群容器运行日志备份指南

1、背景

K8s 集群或者集群内的组件（容器）发生故障时，获取一份容器运行日志的完整快照，便于事后分析、定位故障原因。收集容器运行日志后，就可以迅速恢复现场，提高无故障时间，这对有服务等级要求的系统来说，尤为重要。

因为实际生产需要，作者编写批处理脚本文件，用于收集 K8s 集群某个命名空间下的所有 Pod/容器的运行日志，打包成压缩文件。因为日志文件包含很多重复内容，压缩比很高，经验值大约在 50~100 : 1 范围内。日志文件压缩之后，因为文件体积变小，便于下载、上传或者传递日志文件包。

批处理采集日志文件时，以命名空间作为范围边界，是因为发生故障时同一命名空间内的容器关联性高，应当一次性采集，便于时候做关联性分析。而如果以整个集群作为范围边界，恐怕日志文件太多，而且采集不相关的日志文件没有必要。

如果需要采集两个或以上的命名空间的容器日志文件，可以分别运行脚本两次或多次，只要带上相应命名空间作为参数就可以了。

2、下载采集备份脚本

2.1 准备工作

备份脚本支持 Linux/Unix 操作系统，先要确认主机已经安装了 bash、git 等工具，再确认已经安装 kubectl 命令，并且 kubectl 命令可以连接到目标 K8s 集群。

以 CentOS 操作系统为例，安装工具软件的命令如下：

```
yum install -y bash  
yum install -y git  
yum install -y kubectl
```

2.2 下载安装备份脚本

创建备份工具的根目录：

```
mkdir -p /data  
cd /data
```

从源码仓库 Github 克隆项目：

```
git clone https://github.com/solomonxu/k8s-backup-restore.git
```

修改文件属性，授予 Shell 脚本文件执行权限：

```
cd /data/k8s-backup-restore  
chmod -R a+x bin
```

3、采集容器日志文件

容器日志采集是以命名空间为范围边界的，采集 Pod/容器的运行日志。与实时采集日志只采集最新的日志条目不同，本脚本采集容器从启动到当前时刻或者结束时的所有日志条目，相当于容器全生命周期的日志快照。

3.1 采集容器运行日志文件

采集容器运行日志，输出到日志文件，并打成压缩包文件：

```
cd /data/k8s-backup-restore/bin  
./k8s_dump_logs.sh [命名空间]
```

如果传入命名空间参数，将会采集（导出）该命名空间下所有 Pod/容器的运行日志。

对于包含多个容器的 Pod，会自动识别 Pod 内的所有容器，分别导出各容器的运行日志。

如果没有传入命名空间参数，默认的命名空间是 `wisecloud-controller`。

脚本运行后，打包压缩后的日志文件包位于目录：`/data/k8s-backup-restore/data/logfile`。

脚本运行时输出日志信息，详细记录导出了哪些 Pod/容器的运行日志，日志文件分别位于什么目录位置，压缩包文件在什么位置。

3.2 容器运行日志利用

下载日志压缩包以后，解压缩就可以用文本阅读器或者 `vi` 命令查看文件内容。

解压缩日志包文件：

```
tar -xvzf xxyy.tar.gz
```

日志文件的命名格式是：

```
<podname>_<container>.log
```