

Неравномерное назначение кэш дисков

Иногда в силу работы встроенного механизма распределения нагрузки журналы на SSD дисках в системе хранения Р-Хранилище создаются не равномерно или вообще не попадают на SSD а находят место на том-же диске.

Симптомы

Для простой проверки распределения журналов по дискам необходимо выполнить команду:

```
ls -alh /vstorage/*/*journal
```

Пример вывода на снимке экрана ниже:

```
mchus@virt02:~$ ls -alh /vstorage/*/*journal
-rw-r----- 1 vstorage vstorage 22 Nov 29 21:23 /vstorage/mds/journal

/vstorage/0be4ee94/journal:
total 12K
drwxr-xr-x 3 root root 4.0K Sep 23 15:24 .
drwxr-xr-x 4 root root 4.0K Sep 23 15:23 ..
lrwxrwxrwx 1 root root 47 Sep 23 15:24 journal-cs-1027 -> journal-cs-e6974ec7-1e3b-4654-838e-ef62d9fc5ce6
drwx----- 3 vstorage vstorage 4.0K Sep 23 15:23 journal-cs-e6974ec7-1e3b-4654-838e-ef62d9fc5ce6

/vstorage/6bad6b71/journal:
total 12K
drwxr-xr-x 3 root root 4.0K Sep 23 15:24 .
drwxr-xr-x 4 root root 4.0K Sep 23 15:23 ..
lrwxrwxrwx 1 root root 47 Sep 23 15:24 journal-cs-1028 -> journal-cs-45d74330-ee31-4a30-b747-d6a803aeba83
drwx----- 3 vstorage vstorage 4.0K Sep 23 15:24 journal-cs-45d74330-ee31-4a30-b747-d6a803aeba83

/vstorage/71c62bd7/journal:
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 23 15:23 .
drwxr-xr-x 5 root root 4.0K Sep 23 15:23 ..

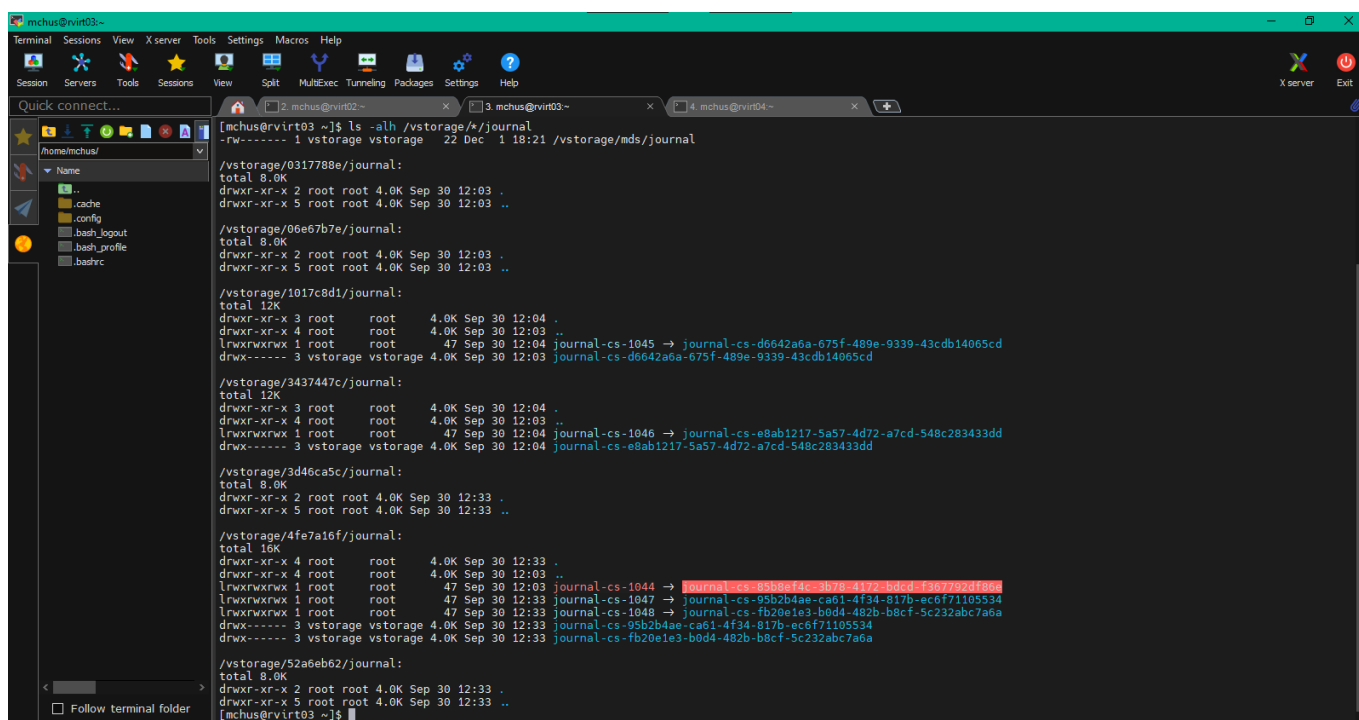
/vstorage/ac1c83db/journal:
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 23 15:23 .
drwxr-xr-x 5 root root 4.0K Sep 23 15:23 ..

/vstorage/b2a4768b/journal:
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 23 15:23 .
drwxr-xr-x 5 root root 4.0K Sep 23 15:23 ..

/vstorage/d33dcd56/journal:
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 23 15:23 .
drwxr-xr-x 5 root root 4.0K Sep 23 15:23 ..

/vstorage/defb98ad/journal:
total 16K
drwxr-xr-x 4 root root 4.0K Sep 23 15:23 .
drwxr-xr-x 4 root root 4.0K Sep 23 15:23 ..
lrwxrwxrwx 1 root root 47 Sep 23 15:23 journal-cs-1025 -> journal-cs-6c959ffd-aed0-4e36-8ffe-3fcc3049e29d
lrwxrwxrwx 1 root root 47 Sep 23 15:23 journal-cs-1026 -> journal-cs-460c2835-601c-403b-9f15-a0d9126f5b76
drwx----- 3 vstorage vstorage 4.0K Sep 23 15:23 journal-cs-460c2835-601c-403b-9f15-a0d9126f5b76
drwx----- 3 vstorage vstorage 4.0K Sep 23 15:23 journal-cs-6c959ffd-aed0-4e36-8ffe-3fcc3049e29d
[mchus@virt02 ~]$
```

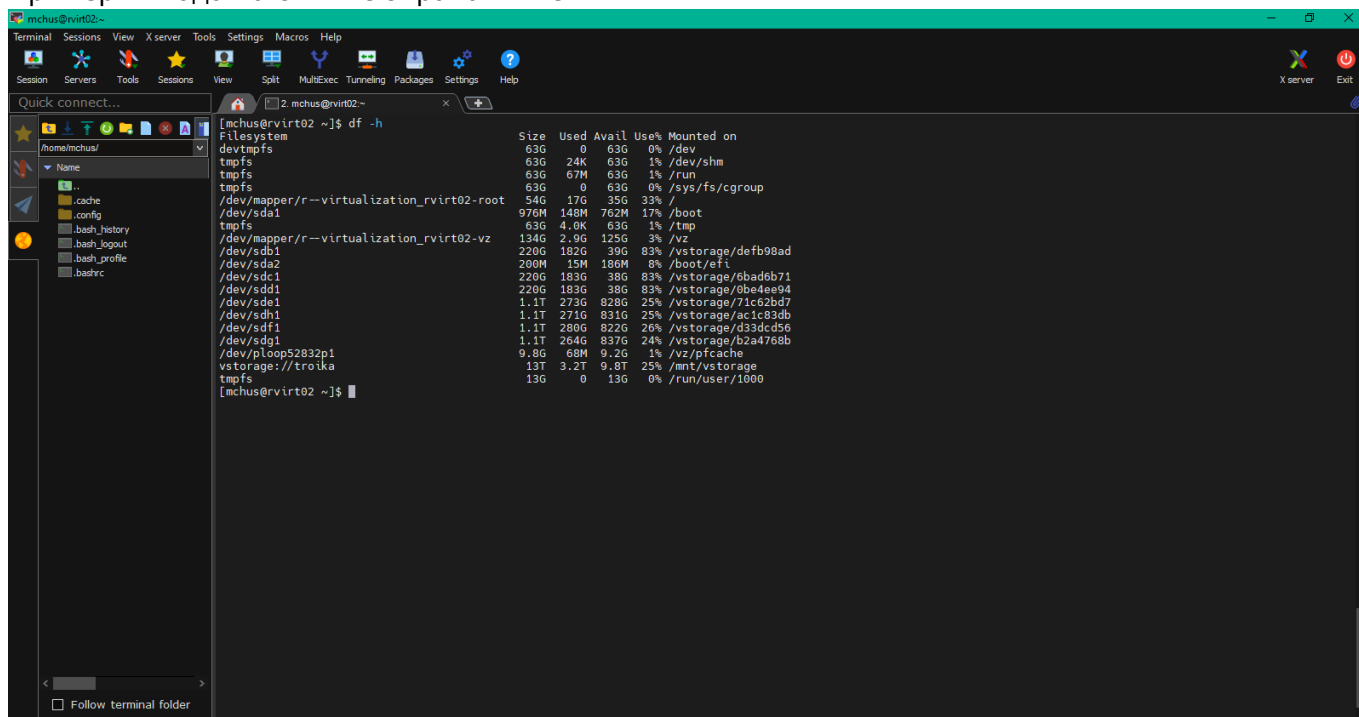
Также на диске могут быть обнаружены “мертвые” ссылки на чанк серверы которые уже удалены, пример такого сервера на снимке ниже:



Тип диска проще всего определить по его объему командой

df -h

Пример вывода на снимке экрана ниже:



В данном случае расположение журналов следующее:

- на диске /dev/sdb1 расположены два журнала для чанк сервисов №1025 и №1026;
- на диске /dev/sdd1 расположен журнал для чанк сервиса 1027;
- на диске /dev/sdc1 расположен журнал для чанк сервиса 1028;

Такое распределение является нормальным, так как сервер оснащен тремя кэширующими

дисками и четырьмя дисками для хранения.

Проверить соответствие чанк серверов физическим дискам необходимо командой:

```
vstorage list-services -c %CLUSTER_NAME%
```

Где:

- %CLUSTER_NAME% - имя кластера Р-Хранилище.

Причины

- Изменение конфигурации Р-Хранилища после назначения кэш-дисков;
- Добавление дисков хранения после назначения кэш-дисков;
- Слишком маленькое количество SSD дисков на систему;
- Слишком маленький суммарный объем SSD дисков по отношению к объему HDD дисков;

Восстановление

Вариант с изменением размера журнала

1. Рассчитать количество журналов которые должны быть размещены на каждом SSD диске по формуле $\%КОЛ-ВО\ CS\ сервисов\% / \%КОЛ-ВО\ SSD\ для\ кэширования\%$;
2. Рассчитать оптимальный объем журнала для каждого cs сервиса по формуле: $\%объем\ SSD\% * 0.8 / \%Кол-во\ HDD\ на\ каждый\ SSD\%$;
3. Проверить сколько журналов расположено на каждом SSD диске с ролью “Кэш” можно командой `ls -al /vstorage/*/journal`.
4. Необходимо удалить все cs сервисы, журналы которых ошибочно оказались на этих же дисках а не на SSD;
5. Далее на каждом SSD диске необходимо оставить количество журналов, не превышающее расчетное. Для этого необходимо:
 1. Получить список журналов SSD диска командой `ls -ls /vstorage/*/journal`;
 2. Получить соотношение идентификатора cs сервиса и локального диска командой `vstorage list-services -c %CLUSTER_NAME%`;
 3. Удалить из консоли веб-управления cs сервисы для достижения необходимого количества журналов на каждом SSD диске;
6. Очистить SSD диски от следующих ненужных файлов:
 1. неиспользуемые журналы, к которым нет символических ссылок¹⁾;
 2. символические ссылки, которые никуда не ведут²⁾;
7. Изменить размер журнала всех созданных сервисов cs командой:

```
for f in /vstorage/*/cs;
do
    echo f$;
    vstorage -c %cluster_name% configure-cs -r $f -s %размер журнала%;
    sleep 1m;
done
```

8. На свободных HDD дисках создать новые cs сервисы по одному;
9. Проверить равномерность распределения файлов журнала и при необходимости:
 1. Удалить cs, журнал которого назначен не на тот SSD;
 2. Для выравнивания занятого пространства запустить скрипт изменения размера журнала;
10. После достижения одинакового количества файлов журналов на всех SSD дисках при всех созданных cs сервисах необходимо:
 1. Проверить равномерность расхода дисковой емкости SSD, командой `df -h`
 2. В случае если занятая дисковая емкость на SSD дисках распределена не равномерно запустить скрипт изменения размера журнала;

Вариант с пересозданием файла журнала из CLI

1. Удалить старый журнал созданный не в том месте, командой: ```vstorage -с имякластера configure-cs -r /vstorage/IDчанкСервисаИЛИIDдискаSSDcКЭШ/journal/имяЖурнала -d```
2. Рассчитать размер журнала по формуле: $\text{Полезная емкость SSD (372ГБ)} * 0.8 (80\%) = 297\text{ГБ} / 4(\text{HDD на один SSD диск}) = 74\text{ГБ}$
3. Создать журнал вручную, командой: `vstorage -с имя кластера configure-cs -r /vstorage/IDчанкСервиса/cs -a /vstorage/IDдискаSSDcКЭШ/journal/ИмяФайлажурналаСчанкаСервиса -s 75776`, где 75776- размер журнала в МБ.

1)

```
/vstorage/0aee28e0/journal:
total 24
drwxr-xr-x 6 root root 4096 Dec 2 13:36 .
drwxr-xr-x 4 root root 4096 Dec 2 13:33 ..
lrwxrwxrwx 1 root root 47 Dec 2 13:35 journal-cs-1332 -> journal-cs-550170b2-87fc-4880-98e4-8a38c0de3b9d
lrwxrwxrwx 1 root root 47 Dec 2 13:35 journal-cs-1333 -> journal-cs-d2209440-2772-457a-955f-fde8f77cd423
lrwxrwxrwx 1 root root 47 Dec 2 13:36 journal-cs-1335 -> journal-cs-a7f8461e-931d-45e4-b71e-1753c437800e
drwx----- 3 vstorage vstorage 4096 Dec 2 13:35 journal-cs-550170b2-87fc-4880-98e4-8a38c0de3b9d
drwx----- 3 vstorage vstorage 4096 Dec 2 13:36 journal-cs-a7f8461e-931d-45e4-b71e-1753c437800e
drwx----- 3 vstorage vstorage 4096 Dec 2 13:35 journal-cs-d2209440-2772-457a-955f-fde8f77cd423
drwx----- 3 vstorage vstorage 4096 Dec 2 13:36 journal-cs-d729fa5b-34da-4927-b96d-b8e09e776c10
```

2)

```
drwxr-xr-x 4 root root 4096 Dec 2 13:58 .
drwxr-xr-x 4 root root 4096 Dec 2 13:33 ..
lrwxrwxrwx 1 root root 47 Dec 2 13:34 journal-cs-1329 -> journal-cs-53d85cb1-585d-4884-b68a-9b9a392bef9f
lrwxrwxrwx 1 root root 47 Dec 2 13:34 journal-cs-1330 -> journal-cs-e03bbffe-5294-4817-b191-332784e72607
lrwxrwxrwx 1 root root 47 Dec 2 13:35 journal-cs-1331 -> journal-cs-c7767fb4-2eab-4bec-8fc4-21592ab29195
drwx----- 3 vstorage vstorage 4096 Dec 2 13:34 journal-cs-53d85cb1-585d-4884-b68a-9b9a392bef9f
drwx----- 3 vstorage vstorage 4096 Dec 2 13:34 journal-cs-c7767fb4-2eab-4bec-8fc4-21592ab29195
```

From:
<https://micronode.ru/> - **micronode.ru**

Permanent link:
<https://micronode.ru/wiki/trouble/rosplatforma/uneven-cache-distribution>

Last update: **2022/12/02 21:03**

