

# **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@rvirt02:~
```

Terminal Sessions View X server Tools Settings Macros Help

Session Servers Tools Sessions View Split MultiExec Tunneling Packages Settings Help

X server Exit

Quick connect...

2. mchus@rvirt02:~

```
..abrt-cli status' timed out
[mchus@rvirt02 ~]$ ls -lah /vstorage/*journal
-rw----- 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 16K
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-a430-b747-d6a803aeba83
drwx----- 3 vstorage vstorage 4.0K Sep 23 15:24 journal-cs-45d74330-ee31-a430-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/b2a4760b/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 ..

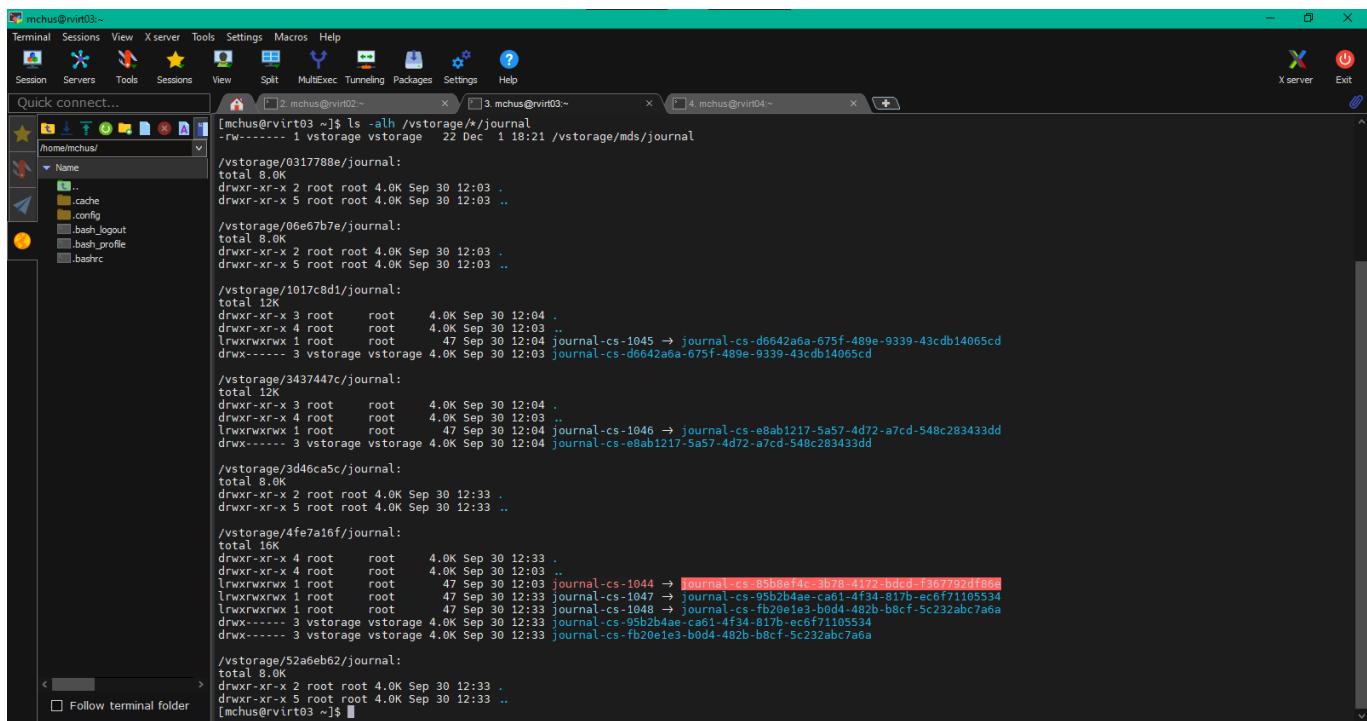
/vstorage/defb9ead/journal:
total 16K
drwxr-xr-x 4 root      root      4.0K Sep 23 15:23 .
drwxr-xr-x 4 root      root      4.0K Sep 23 15:23 ..
lrwxrwxrwx 1 root      root      47 Sep 23 15:23 journal-cs-1026 → journal-cs-6c959ffd-aed0-4e36-8ffe-3fcc3049e29d
drwx----- 3 vstorage vstorage 4.0K Sep 23 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@rvirt02 ~]$
```

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

Last update:

2022/12/02 wiki:trouble:rosplatforma:uneven-cache-distribution https://micronode.ru/wiki/trouble/rosplatforma/uneven-cache-distribution  
21:03



```
[mchus@rvirt03 ~]$ ls -lah /vstorage/*/journal
-rw----- 1 vstorage vstorage 22 Dec 1 18:21 /vstorage/mds/journal
[mchus@rvirt03 ~]$ ls -lah /vstorage/*/journal
-rw----- 1 vstorage vstorage 22 Dec 1 18:21 /vstorage/mds/journal
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 30 12:03 .
drwxr-xr-x 5 root root 4.0K Sep 30 12:03 ..

/vstorage/06e67b7e/journal:
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 30 12:03 .
drwxr-xr-x 5 root root 4.0K Sep 30 12:03 ..
drwxr-xr-x 3 vstorage vstorage 4.0K Sep 30 12:03 journal.cs-d6642a6a-675f-489e-9339-43cdb14065cd

/vstorage/1017c8d1/journal:
total 12K
drwxr-xr-x 3 root root 4.0K Sep 30 12:04 .
drwxr-xr-x 4 root root 4.0K Sep 30 12:03 ..
lwxrwxrwx 1 root root 47 Sep 30 12:04 journal.cs-1045 → journal.cs-d6642a6a-675f-489e-9339-43cdb14065cd
drwxr-xr-x 3 vstorage vstorage 4.0K Sep 30 12:04 journal.cs-e8ab1217-5a57-4d72-a7cd-548c283433dd

/vstorage/3d46ca5c/journal:
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 30 12:33 .
drwxr-xr-x 5 root root 4.0K Sep 30 12:33 ..

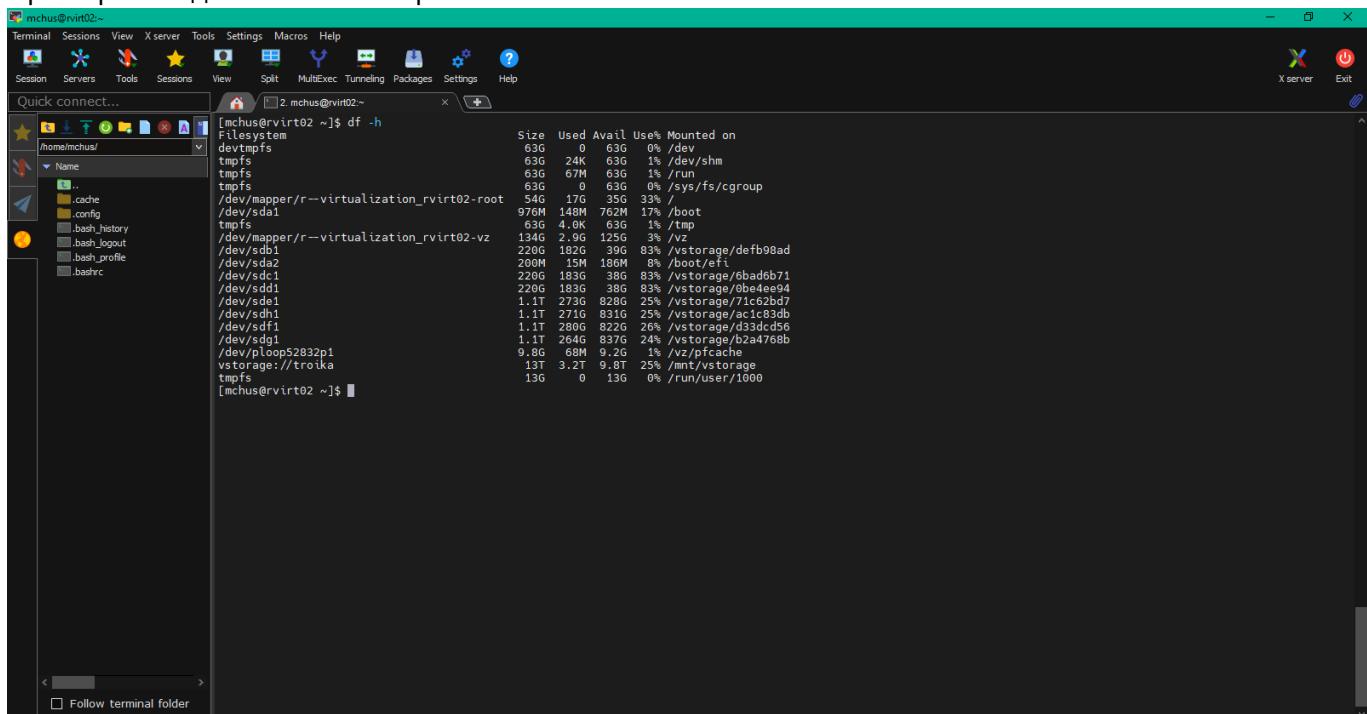
/vstorage/4fe7a16f/journal:
total 16K
drwxr-xr-x 4 root root 4.0K Sep 30 12:33 .
drwxr-xr-x 4 root root 4.0K Sep 30 12:03 ..
lwxrwxrwx 1 root root 47 Sep 30 12:03 journal.cs-1044 → journal.cs-95b2b4ae-ca61-4f34-817b-ec6f71105534
lwxrwxrwx 1 root root 47 Sep 30 12:33 journal.cs-1048 → journal.cs-fb20e1e3-b0d4-482b-b8cf-5c232abc7a6a
drwxr-xr-x 3 vstorage vstorage 4.0K Sep 30 12:33 journal.cs-95b2b4ae-ca61-4f34-817b-ec6f71105534
drwxr-xr-x 3 vstorage vstorage 4.0K Sep 30 12:33 journal.cs-fb20e1e3-b0d4-482b-b8cf-5c232abc7a6a

/vstorage/52a6eb62/journal:
total 8.0K
drwxr-xr-x 2 root root 4.0K Sep 30 12:33 .
drwxr-xr-x 5 root root 4.0K Sep 30 12:33 ..
```

Тип диска проще всего определить по его объему командой

```
df -h
```

Пример вывода на снимке экрана ниже:



```
[mchus@rvirt02 ~]$ df -h
Filesystem      Size  Used Avail Use% Mounted on
/devtmpfs        63G   0    63G  0% /dev
tmpfs           63G  24K  63G  1% /dev/shm
tmpfs           63G  678M 63G  1% /run
tmpfs           63G   0    63G  0% /sys/fs/cgroup
/dev/mapper/r-virtualization_rvirt02-root  54G  17G  35G  33% /
/dev/sda1       976M 148M 762M 17% /boot
tmpfs           63G  4,0K  63G  1% /tmp
/dev/mapper/r-virtualization_rvirt02-vz  134G  2,9G 125G  3% /vz
/dev/sdb1       220G 182G 39G  83% /vstorage/defb98ad
/dev/sda2       200M 15M 186M  8% /boot/efi
/dev/sdc1       220G 183G 38G  83% /vstorage/6bad6b71
/dev/sdd1       220G 183G 38G  83% /vstorage/6be4ee94
/dev/sde1       1.1T 2,1T 890G 25% /vstorage/71c62bd7
/dev/sdh1       1.1T 2,71G 831G 25% /vstorage/4c1c83b0
/dev/sdf1       1.1T 2,80G 822G 26% /vstorage/d33dc556
/dev/sdg1       1.1T 2,64G 837G 24% /vstorage/b2a47e8b
/dev/ploop52832p1  9,8G  68M 9,2G  1% /vz/pfcache
vstorage://troika 13T 3,2T 9,8T 25% /mnt/vstorage
tmpfs           13G   0    13G  0% /run/user/1000
[mchus@rvirt02 ~]$
```

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 символьических ссылок <sup>1)</sup>;
  2. символические ссылки, которые никуда не ведут <sup>2)</sup>;
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 -c имя кластера configure-cs -r /vstorage/IDчанкСервиса ИЛИ IDдиска SSD c КЭШ/journal/имяЖурнала -d`````
2. Рассчитать размер журнала по формуле: Полезная емкость SSD (372ГБ) \* 0.8 (80%) = 297ГБ / 4(HDD на один SSD диск)=74ГБ
3. Создать журнал вручную, командой: `vstorage -c имя кластера configure-cs -r /vstorage/IDчанкСервиса/cs -a /vstorage/IDдиска SSD c КЭШ/journal/ИмяФайлаЖурнала СчанкаСервиса -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 ..  
lrvwxrwxrwx 1 root      root     47 Dec  2 13:35 journal-cs-1332 -> journal-cs-550170b2-87fc-4880-98e4-8a38c0de3b9d  
lrvwxrwxrwx 1 root      root     47 Dec  2 13:35 journal-cs-1333 -> journal-cs-d2209440-2772-457a-955f-fde8f77cd423  
lrvwxrwxrwx 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-d729fe5b-34da-4927-b96d-b8e09e77be10
```

2)

```
drwxr-xr-x 4 root      root    4096 Dec  2 13:58 .  
drwxr-xr-x 4 root      root    4096 Dec  2 13:33 ..  
lrvwxrwxrwx 1 root      root     47 Dec  2 13:34 journal-cs-1329 -> journal-cs-53d85cb1-585d-4884-b68a-9b9a392bef9f  
lrvwxrwxrwx 1 root      root     47 Dec  2 13:34 journal-cs-1330 -> journal-cs-e03bbffe-5294-4817-b191-332784e72607  
lrvwxrwxrwx 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**