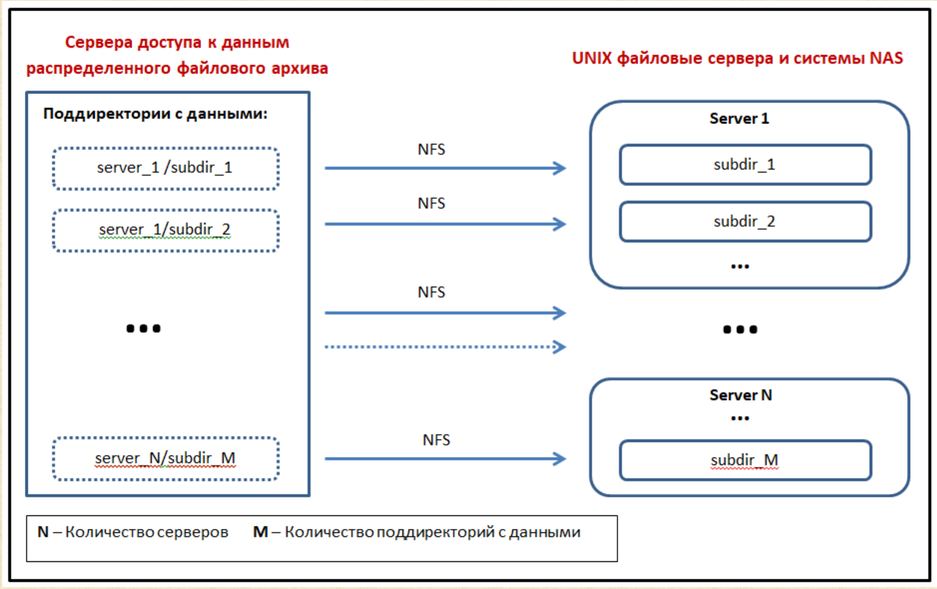
# Поддержка функционирования сверхбольших файловых хранилищ спутниковых данных ЦКП «ИКИ-Мониторинг»

***Прошин А.А., Бурцев М.А., Радченко М.В.***

В настоящее время архивы ЦКП «ИКИ-Мониторинг» содержат более 2,5 петабайт спутниковых данных, находящихся в режиме непосредственного доступа. При этом ежедневно в архивы поступает более 2 терабайт новых данных. Основная часть этих данных располагается непосредственно в ИКИ РАН, для чего в настоящее время используется 25 серверов под управлением операционной системы UNIX и 4 специализированных системы хранения данных NAS. Реализация распределенных файловых архивов базируется на использовании протокола NFS.

Для надежной и эффективной работы рассматриваемого программно-аппаратного комплекса необходима автоматизация и оптимизация основных процедур по работе с ним, а также использование специальных инструментов для контроля за его работой. Для решения этих задач специалистами отдела «Технологии спутникового мониторинга» ИКИ РАН был разработан целый ряд программных решений, представленных ниже.

*Схема построения распределенного файлового архива*




## Программа автоматического управления NFS-соединениями

Программа предназначена для поддержания в актуальном состоянии большого числа различных NFS-соединений, набор которых существенно отличается для различных групп серверов и конфигурируется централизованно. Позволяет автоматически отключать проблемные соединения и восстанавливать их после устранения сбоев с оборудованием или сетью, что является крайне важным для корректной работы серверов, обеспечивающих доступ к данным архивов.

```
# 193.232.9.115:/data/HRSATDB_FDB_STORE on /mnt/epsilonstore_hrsatdb_fdb_store
> Check mount exists .. Ok
> Check mount directory .. Ok
> Check web access .. Ok
# already mounted

# 193.232.9.223:/data/HRSATDB_FDB_STO
> Check mount exists .. Ok
> Check mount directory .. Ok
> Check web access .. Failed
> Call umount .. # unmounted

# 193.232.9.150:/sdata/HRSATDB_FDB_ST
> Check mount exists .. Ok
> Check mount directory .. Ok
> Check web access .. Ok
# already mounted

# 193.232.9.150:/data1/HRSATDB_FDB_ST
> Check mount exists .. Ok
> Check mount directory .. Ok
> Check web access .. Ok
# already mounted
```

```
/mnt/lambdastore_hrsatdb_fdb_store       unmounted
/mnt/mrsatdb_fdb_store                   already mounted
/mnt/mstore_fdb_store                    already mounted
/mnt/mustore_radarsat                    already mounted
/mnt/new_wkstore                         already mounted
/mnt/nustore_hrsatdb_fdb_store           already mounted
/mnt/nustore_mrsatdb_fdb_store           already mounted
/mnt/nustore_smisdata_fdb_store          already mounted
/mnt/nustore_zetastore_fdb_store         already mounted
/mnt/rdxps9                              already mounted
/mnt/rstore_a                            already mounted
/mnt/rstore_b                            already mounted
/mnt/smisdata_fdb_store                  already mounted
/mnt/smisvega_fdb_store                  already mounted
/mnt/thetastore_hrsatdb_fdb_store        already mounted
/mnt/thetastore_mrsatdb_fdb_store        already mounted
/mnt/zetastore_fdb_store                 already mounted

already mounted : 52
      unmounted : 1
          total : 53

Some of mounts are invalid
```

## Программа автоматической диспетчеризации входных потоков данных

Программа позволяет отслеживать степень заполнения дисковых пространств серверов хранения и автоматически переключать поступление данных на другой сервер, располагающий свободным дисковым пространством. Приоритет при выборе сервера задается в конфигурации, заданной отдельно для каждого из архивов, что позволяет достаточно гибко распределять новые данные по различным серверам. Также реализована дополнительная проверка доступности текущего сервера, необходимая для корректной реакции на различные сбойные ситуации.

```
*** Configuration: ***
Space for next subdir   : 2Tb
Storage data subdir     : hrsat/fragments/qlook
Local data dir          : /data/fdb_store/unisat_hrsat/fragment_products/product
Preferable mount dirs   : /mnt/rdxps9, /mnt/duostorec_hrsatdb_fdb_store_data2, /mnt/kilostore_hrsatdb_fd
Required next
subdirs count           : 1
Critical free space
for current subdir      : 1024Gb

### Analize info ###
# Suitable mount dirs:
/mnt/duostorea_hrsatdb_fdb_store_data1   - 2Tb
/mnt/duostoreb_hrsatdb_fdb_store_data1   - 2Tb
/mnt/duostorea_hrsatdb_fdb_store_data2   - 2.4Tb
/mnt/epsilonstore_hrsatdb_fdb_store      - 2.5Tb
/mnt/rstore_b                            - 4.0Tb
/mnt/exoplanet                           - 6Tb
/mnt/nustore_hrsatdb_fdb_store           - 8Tb
/mnt/kilostore_hrsatdb_fdb_store         - 11Tb
/mnt/duostorec_hrsatdb_fdb_store_data2   - 35Tb
/mnt/rdxps9                              - 398Tb
# Next mount dir        : /mnt/rdxps9
# Max ID                : 186077692
# Current subdir        : 01860 /mnt/rdxps9 (398Tb)
# Last subdir           : 01861 /mnt/rdxps9 (398Tb)
# Next subdirs count    : 1
# Data upload link      : /data/fdb_store/current_storage => /mnt/rdxps9
# Data archive link     : /data/unisat/current_storage => /mnt/rdxps9
```

## Блок мониторинга состояния распределенных файловых архивов

На каждом из серверов хранения периодически запускаются процедуры для контроля за состоянием дисковых массивов, получения различной информации о заполнения дискового пространства, тестирования скорости чтения и записи и др. Для сбора и анализа этой информации на одном из наших серверов реализована специальная база данных, в которой также содержится необходимая справочная информация об используемых серверах хранения данных, включая информацию о том, каким архивам данных соответствуют те или иные директории серверов хранения. Это позволяет получить интегральные характеристики, описывающие текущее состояние распределенного файлового хранилища, а также прогнозировать постоянно растущие потребности в расширении дискового пространства.

**Информация о состоянии томов хранения данных, размещенных на серверах отдела 56**

| № | Сервер | Диск | Размер (Тб) | Свободно (Тб) | Занято (Тб) | Занято (%) | Тип ФС | Скорость чтения (Мб/с) | Скорость записи (Мб/с) | Дата измерения GMT | Состояние RAID массива | Проверка перезаполнения | Информация об архивах: Архив | Директория | Размер (Тб) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | alphastore | /data | 55.6 | 8.6 | 47.0 | 84% | ZFS | 651 | 207 | 2019-11-07 | ● | ● | modis_products | /data/smisstore | 37.0 |
| | | | | | | | | | | | | | Прочие данные | - | 10.0 |
| 2. | betastore | /data | 56.0 | 1.4 | 50.1 | 89% | UFS | 291 | 361 | 2019-11-07 | ● | ● | unisat_hrsat | /data/HRSATDB_FDB_STORE/hrsat | 13.7 |
| | | | | | | | | | | | | | unisat_mrsat | /data/MRSATDB_FDB_STORE/mrsat | 6.0 |
| | | | | | | | | | | | | | unisat_radarsat | /data/UNISAT_RADARSAT_FDB_STORE/unisat_radarsat | 26.6 |
| | | | | | | | | | | | | | Прочие данные | | 3.8 |
| 3. | d902 | /data | 15.1 | 0.6 | 14.5 | 96% | ZFS | 996 | 193 | 2019-11-08 | ● | ● | unisat_mrsat | /data/MRSATDB_FDB_STORE/mrsat | 12.1 |
| | | | | | | | | | | | | | Прочие данные | - | 2.4 |
| 4. | d902 | /mnt/exoplanet | 295.0 | 6.6 | 288.4 | 97% | NFS | 101 | 106 | 2019-11-08 | ● | ● | unisat_comp | /mnt/exoplanet/unisat_comp | 0.2 |
| | | | | | | | | | | | | | unisat_hrsat | /mnt/exoplanet/hrsat | 147.8 |
| | | | | | | | | | | | | | unisat_mrsat | /mnt/exoplanet/mrsat | 23.6 |
| 5. | d902 | /mnt/new_wkstore | 50.0 | 0.5 | 49.5 | 99% | NFS | 104 | 73 | 2019-11-08 | | | | | |
| 6. | d902 | /mnt/rstore_a | 202.6 | 27.4 | 166.4 | 82% | NFS | 110 | 97 | 2019-11-08 | | | | | |
| 7. | d902 | /mnt/rstore_b | 217.2 | 4.0 | 203.8 | 93% | NFS | 104 | 94 | 2019-11-08 | | | | | |
| 8. | deltastore | /data | 53.6 | 0.8 | 48.5 | 90% | UFS | 134 | 303 | 2019-11-07 | | | | | |
| 9. | deltastore | /sdata | 2.2 | 0.4 | 1.6 | 73% | UFS | 142 | 106 | 2019-11-07 | | | | | |
| 10. | duostorea | /data1 | 76.0 | 2.2 | 67.8 | 89% | UFS | 349 | 601 | 2019-11-07 | | | | | |
| 11. | duostorea | /data2 | 74.9 | 2.4 | 70.2 | 93% | UFS | 332 | 241 | 2019-11-07 | | | | | |
| 12. | duostorea | /sdata | 71.9 | 0.9 | 65.2 | 90% | UFS | 714 | 929 | 2019-11-07 | | | | | |
| 13. | duostoreb | /data1 | 74.1 | 2.2 | 65.9 | 89% | UFS | 400 | 477 | 2019-11-08 | | | | | |
| 14. | duostoreb | /data2 | 75.4 | 1.6 | 71.5 | 94% | UFS | 452 | 923 | 2019-11-08 | | | | | |
| 15. | duostoreb | /data3 | 70.4 | 1.0 | 67.3 | 95% | UFS | 399 | 357 | 2019-11-08 | | | | | |
| 16. | duostorec | /data1 | 74.1 | 1.8 | 68.6 | 92% | UFS | 283 | 232 | 2019-11-08 | | | | | |
| 17. | duostorec | /data2 | 76.0 | 35.3 | 39.2 | 51% | UFS | 546 | 1604 | 2019-11-08 | | | | | |

**Статистика по дисковому пространству, занимаемому архивами отдела 56**

Выбор даты: 2019-11-08

| Архив | Объем (Тб) | Приращение объема за 10 дней* (Тб) | Информация о томах хранения: Сервер | Раздел | Директория | Размер (Тб) | Дата измерения |
|---|---|---|---|---|---|---|---|
| modis_products | 56.6 | 0.1 | alphastore | /data | /data/smisstore | 37.0 | 2019-11-07 |
| | | | iotastore | /data | /data/modis_products/THDD | 19.7 | 2019-11-07 |
| unisat_comp | 27.6 | 0.0 | d902 | /mnt/exoplanet | /mnt/exoplanet/unisat_comp | 0.2 | 2019-11-08 |
| | | | d902 | /mnt/rstore_a | /mnt/rstore_a/unisat_comp | 2.2 | 2019-11-08 |
| | | | epsilonstore | /data | /data/ZETASTORE_FDB_STORE/unisat_comp | 7.8 | 2019-11-08 |
| | | | zetastore | /data | /data/fdb_store/regions_products | 17.4 | 2019-11-08 |
| unisat_hrsat | 1639.0 | 3.7 | betastore | /data | /data/HRSATDB_FDB_STORE/hrsat | 13.7 | 2019-11-07 |
| | | | d902 | /mnt/exoplanet | /mnt/exoplanet/hrsat | 147.8 | 2019-11-08 |
| | | | d902 | /mnt/new_wkstore | /mnt/new_wkstore/HRSATDB_FDB_STORE | 22.3 | 2019-11-08 |
| | | | d902 | /mnt/rstore_b | /mnt/rstore_b/hrsat | 200.9 | 2019-11-08 |
| | | | deltastore | /data | /data/HRSATDB_FDB_STORE/hrsat | 17.2 | 2019-11-07 |
| | | | duostorea | /data1 | /data1/HRSATDB_FDB_STORE | 67.8 | 2019-11-07 |
| | | | duostorea | /data2 | /data2/HRSATDB_FDB_STORE | 68.5 | 2019-11-07 |
| | | | duostorea | /sdata | /sdata/HRSATDB_FDB_STORE | 21.9 | 2019-11-07 |
| | | | duostoreb | /data1 | /data1/HRSATDB_FDB_STORE | 65.9 | 2019-11-08 |
| | | | duostoreb | /data2 | /data2/HRSATDB_FDB_STORE | 71.5 | 2019-11-08 |
| | | | duostoreb | /data3 | /data3/HRSATDB_FDB_STORE | 26.2 | 2019-11-08 |
| | | | duostorec | /data1 | /data1/HRSATDB_FDB_STORE | 68.6 | 2019-11-08 |
| | | | duostorec | /data2 | /data2/HRSATDB_FDB_STORE | 38.9 | 2019-11-08 |
| | | | epsilonstore | /data | /data/HRSATDB_FDB_STORE/hrsat | 6.3 | 2019-11-08 |
| | | | etastore | /data | /data/HRSATDB_FDB_STORE/hrsat | 35.9 | 2019-11-07 |
| | | | firemaps | /data | /data/HRSATDB_FDB_STORE/hrsat | 49.7 | 2019-11-07 |
| | | | fktweb | /mnt/nas | /mnt/nas/fdb_store | 317.9 | 2019-11-08 |
| | | | fktweb | /mnt/nas2 | /mnt/nas2/HRSATDB_FDB_STORE | 118.4 | 2019-11-08 |
| | | | gammastore | /data | /data/HRSATDB_FDB_STORE/hrsat | 34.1 | 2019-11-07 |
| | | | hstore | /data | /data/fdb_store/hrsat | 70.6 | 2019-11-07 |
| | | | kappastore | /data | /data/HRSATDB_FDB_STORE/hrsat | 51.7 | 2019-11-07 |
| | | | kilostore | /data | /data/HRSATDB_FDB_STORE/hrsat | 56.7 | 2019-11-07 |
| | | | kilostore | /sdata | /sdata/HRSATDB_FDB_STORE/hrsat | 13.6 | 2019-11-07 |
| | | | nustore | /data | /data/HRSATDB_FDB_STORE/hrsat | 37.1 | 2019-11-07 |
| | | | thetastore | /data2 | /data2/HRSATDB_FDB_STORE/hrsat | 15.6 | 2019-11-07 |
| unisat_mrsat | 158.1 | 0.6 | betastore | /data | /data/MRSATDB_FDB_STORE/mrsat | 6.0 | 2019-11-07 |
| | | | d902 | /data | /data/MRSATDB_FDB_STORE/mrsat | 12.1 | 2019-11-08 |
| | | | d902 | /mnt/exoplanet | /mnt/exoplanet/mrsat | 23.5 | 2019-11-08 |
| | | | d902 | /mnt/new_wkstore | /mnt/new_wkstore/MRSATDB_FDB_STORE | 3.0 | 2019-11-08 |
| | | | d902 | /mnt/rstore_a | /mnt/rstore_a/mrsat | 28.0 | 2019-11-08 |
| | | | deltastore | /data | /data/MRSATDB_FDB_STORE/mrsat | 14.9 | 2019-11-07 |
| | | | deltastore | /sdata | /sdata/MRSATDB_FDB_STORE/mrsat | 1.5 | 2019-11-07 |
| | | | epsilonstore | /data | /data/MRSATDB_FDB_STORE/mrsat | 10.1 | 2019-11-08 |
| | | | gammastore | /data | /data/MRSATDB_FDB_STORE/mrsat | 13.6 | 2019-11-07 |

## Поддержка бесперебойной работы серверов хранения

На каждом из серверов функционирует ПО PMS (Process Monitoring System), разработанное в ИКИ РАН. Для своевременного выявления сбоев и неполадок в работе серверов, оперативного оповещения о них персонала и сопровождения процесса их устранения используется подсистема ведения сбоев, входящая в состав разработанной в ИКИ РАН Системы Документирования и Контроля Проектов (СДКП).

## Используемые программно-аппаратные решения

Основная часть серверов хранения данных функционирует под ОС FreeBSD. При этом для хранения данных используются дисковые массивы RAID 6. Как правило, сервера поддерживают удаленное управление (IPMI и т.п.). Для хранения файлов данных используются также специализированные системы хранения, в первую очередь, программно-определяемые СХД на базе ПО RAIDIX.

Институт космических исследований (ИКИ РАН), Отдел технологий спутникового мониторинга
117810 Москва, Профсоюзная ул. 84./32, тел.:(495) 333-53-13, факс:(495)9133040, e-mail: info@smis.iki.rssi.ru