КОСМИЧЕСКИЕ ИССЛЕДОВАНИЯ, 2012, том 50, № 4, с. 346–350

КРАТКИЕ СООБЩЕНИЯ

УДК 002.6

ОРГАНИЗАЦИЯ ЦЕНТРА ОБРАБОТКИ НАУЧНОЙ ИНФОРМАЦИИ ДЛЯ РАДИОИНТЕРФЕРОМЕТРИЧЕСКИХ ПРОЕКТОВ

© 2012 г. М. В. Шацкая[1], А. А. Андрианов[1], И. А. Гирин[1], Е. А. Исаев[2], В. И. Костенко[1], С. Ф. Лихачев[1], А. С. Пимаков[1], С. И. Селиверстов[1], Н. А. Федоров[1]

[1] Астрокосмический центр ФИАН, г. Москва

[2] Пущинская радиоастрономическая обсерватория АКЦ ФИАН

mshatsk@asc.rssi.ru

Поступила в редакцию 18.02.2011 г.

В настоящее время в связи с развитием наблюдательной астрономической техники и значительными достижениями в регистрирующих технологиях астрономия столкнулась с лавинообразным увеличением количества наблюдательных данных. Эти данные полученные в различных диапазонах длин волн, от гамма и рентгеновского до радиодиапазона, являются массивами огромных размеров. Отсюда возникает проблема передачи и хранения такого количества информации. Еще одна проблема внеатмосферных астрономических исследований – это колоссальные объемы вычислительной работы, сопутствующие математической обработке и анализу наблюдений [1, 2]. Такие задачи решаются при помощи вычислительной техники, которая предоставляет возможности для создания мощных систем хранения и обработки информации.

Вот почему составной частью успеха астрономических исследований является вычислительная техника.

Центр Обработки Научной информации (ЦОНИ). ЦОНИ – отказоустойчивая комплексная централизованная система, взаимосвязанных программных и аппаратных компонент, организационных процедур, предназначенная для надежного хранения и обработки информации, предоставления сервисов, приложений, обладающая высокой степенью виртуализации своих ресурсов.

Основные задачи выполняемые ЦОНИ: эффективное осуществление сбора и хранения данных в специализированном информационном хранилище в течение заданного периода времени с заданной надежностью; предоставление пользователям прикладных сервисов; обработка данных на высокопроизводительном вычислительном комплексе.

Вычислительный комплекс. Структурная схема вычислительного комплекса, созданного в Астрокосмическом Центре ФИАН, представлена на рис. 1. Это: управляющий узел; вычислительный кластер; хранилище данных суммарным объемом 80 ТБ; система резервного копирования на магнитных лентах на 32 ТБ; резервная система хранения данных 24 ТБ в Пущинской радиоастрономической обсерватории АКЦ ФИАН в г. Пущино; WEB и FTP – сервер; сети управления и передачи данных.

Внешний вид одной из стоек этого комплекса изображен на рис. 2. Стойка была спроектирована специалистами IBM и собрана специалистами компании KraftWay по заказу Астрокосмического центра. Остановимся подробнее на каждом компоненте вычислительного комплекса.

Кластер (являющийся основным компонентом комплекса) – это группа компьютеров, объединенная высокоскоростными каналами связи и предоставляющая с точки зрения пользователя единый аппаратный ресурс. Кластеры в зависимости от назначения подразделяются на: отказоустойчивые, кластеры сбалансированной нагрузки, вычислительные кластеры и др. У нас вычислительный кластер, используемый в научных исследованиях.

Для вычислительных кластеров существенными показателями являются высокая производительность процессора в операциях над числами с плавающей точкой (flops), низкая латентность (задержка) объединяющей сети и, менее существенными, скорость операций ввода-вывода, которая в большей степени важна для баз данных и web-сервисов. Вычислительные кластеры позволяют уменьшить время расчетов, по сравнению с одиночным компьютером, разбивая задание на параллельно выполняющиеся ветки, которые обмениваются данными по связывающей сети.

Представляемый в данной статье вычислительный кластер включает в себя один управляющий и 5 вычислительных серверов смонтированных в общую стойку. Сервера соединены друг с

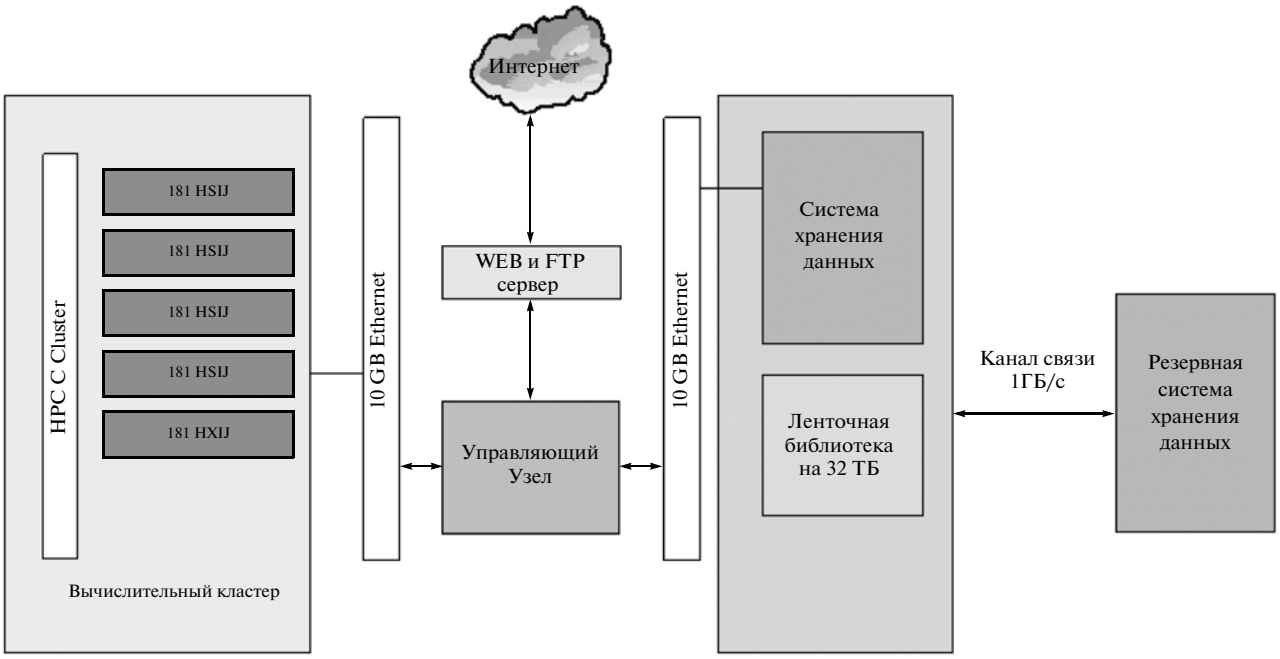

Рис. 1

другом двумя сетями. Одна из сетей является частной (в литературе чаще всего упоминается как Private) для обмена MPI трафиком с пропускной способностью 10 Гбит/с. Это реализовано за счет использования специально оптимизированного для агрегации серверов центра обработки данных высокопроизводительного коммутатора Cisco. Вторая сеть, использующая интерфейс Gigabit Ethernet является публичной (Enterprise). Она используется для управления серверами, ленточной библиотекой, хранилищем данных, а также для доступа с управляющего узла через удаленный рабочий стол к серверам кластера.

С “внешним миром” кластер общается через управляемый свитч. Он также играет роль фаервола (Firewall), ограничивая круг лиц, имеющих доступ к кластеру, и обеспечивая высокую производительность, сочетая в себе гибкость увеличения числа пользователей в сети и расширенные функции безопасности. В качестве операционной системы была выбрана Microsoft HPC server 2008 [3], являющаяся новым поколением высокопроизводительных вычислительных систем. Полностью 64-разрядная архитектура и усовершенствованные протоколы обмена информацией вместе с удобными инструментами управления и мониторинга, механизмом создания сценариев командной строки дают широкие перспективы использования кластерной платформы Microsoft HPC Server 2008. Производительность созданного вычислительного кластера по Linpack получилась равной 430 Гфлоп/с. Предусмотрено дальнейшее масштабирование комплекса до 1 Тфлоп/с.

Хранилище. Система хранения состоит из основной и резервной системы хранения данных на HDD на 104 ТБ, плюс ленточная библиотека (система резервного копирования) на 32 ТБ. Первым шагом на пути обеспечения высокой готовности (в иностранной литературе упоминается как high available) является защита наиболее важной части системы, а именно – данных. В нашем случае система хранения данных состоит из дискового массива с двумя RAID-контроллерами на базе 600Mhz RISC-процессоров, кэш-памятью 1 Гбайт, а также трех JBOD (Just Banch Of Discs) внешних дисковых массивов.

Соединение JBOD массивов между собой и контроллером массива, а также между контроллером и головным сервером осуществляется по счетверенному SAS 4x интерфейсу посредством двух широких SAS 4x портов на каждом RAID-контроллере. Такое соединение системы хранения с головным сервером позволяет обеспечить пропускную способность хост-канала 10 Гбит/с.

Контроллер дискового массива и каждая полка состоят из 16 отсеков для дисков SAS/SATAII со скоростью до 300 Мбайт/с и возможностью горячей замены. Суммарный объем дискового пространства 80 Тбайт.

Рис. 2

Отказоустойчивость системы хранения данных достигается путем использования двух резервируемых блоков питания с возможностью горячей замены и 3 вентиляторов охлаждения, также с возможностью горячей замены.

Надежность хранения достигается использованием RAID6, похожим на RAID5, но имеющим более высокую степень надежности – под контрольные суммы выделяется емкость 2-х дисков, рассчитываются 2-е суммы по разным алгоритмам. Данная технология предполагает использование наборов дисков, доступных пользователям как один логический диск. На случай неисправностей, дисковый массив содержит дополнительную емкость, обеспечивающую возможность восстановления данных.

Управление системой хранения данных осуществляется с помощью встроенного ПО RAID контроллера или через графический интерфейс ПО SANWatch. Графический интерфейс SANWatch обеспечивает все функции, необходимые для настройки, конфигурирования, администрирования и мониторинга RAID массивов, вне зависимости от их количества и удаленности. SANWatch также обеспечивает управляющий интерфейс для EonPath и Snapshot, так что управление всеми функциями защиты данных и доступа становится возможным через один интерфейс.

Ленточная библиотека Tandberg StorageLibrary T40 обеспечивающая *резервное копирование* до 32 Тбайт данных “на лету”. Конструкция библиотеки обеспечивает возможность горячей замены не только отдельных ленточных картриджей (емкостью 0.8 ТБ), но и загрузку/выгрузку магазинов с картриджами. Для организации автоматизированного учета картриджей в библиотеку встроен считыватель штрих кодов. Настройка, управление и мониторинг состояния библиотеки может осуществляться как на лицевой панели устройства, так и с помощью специального программного обеспечения Symantec Backup Exec. Предназначением ленточной библиотеки также Tandberg StorageLibrary T40 в ЦОНИ является длительное хранение полученных необработанных данных, а также редко запрашиваемых данных после обработки.

В случае выхода из строя какого-либо компонента вычислительного комплекса или при обрыве связи с ним для исключения потери данных в

Пущино организована *резервная система* хранения данных на 24 ТБ.

Для оперативного обмена данными с резервным узлом хранения информации создан прямой независимый канал связи емкостью 1 Гб/с между ПРАО ФИАН в г. Пущино и АКЦ (Москва). На первом этапе был создан канал связи в пределах г. Пущино. На втором этапе для соединения Пущинского участка канала связи с Москвой было заключено соглашение с крупнейшим провайдером Московской области ООО "СТЭК" о предоставлении линии связи СТЭК (Пущино) – М9 для целей научных исследований. Затем был организован канал между московским центром коммутации М9 и ЦОНИ (АКЦ).

Предполагается, что информация в Центр Обработки может поступать через сеть интернета. Из мест, где нет высокоскоростных каналов связи, возможна доставка информации на жестких дисках.

Следующий компонент нашего комплекса WEB и FTP *сервер*, который обеспечивает непосредственный контакт с пользователем, т.е. предоставление прикладных сервисов.

Копирование данных, мониторинг их обработки и удаленная работа с данными организованна посредством WEB и FTP сервера, который является своего рода шлюзом между кластером и внешней сетью, что также обеспечивает дополнительную защиту кластера от вирусов и атак извне.

Кондиционирование, надежность, видеонаблюдение. Современное высокотехнологичное вычислительное и телекоммуникационное оборудование чувствительно к самым незначительным изменениям окружающей среды. Обязательным условием обеспечения его нормальной работоспособности является поддержание строго определенных температурных режимов и уровня влажности, поэтому в комнате установлена система кондиционирования и вентиляции, подключенные по схеме N + 1. В случае выхода одного кондиционера, автоматически включается другой. К тому же у них реализована функция удаленного мониторинга микроклимата в помещении, что позволяет оперативно реагировать на изменение температуры или отказ модулей.

Для серверов организована система бесперебойного электропитания на основе APC, которая исключает потерю данных в случае пропадания или скачков электроэнергии, обрыва одной или нескольких жил питающих кабелей.

Не меньшую важность для обеспечения функционирования ЦОНИ имеют системы мониторинга состояния оборудования и видеоконтроля помещения. Система мониторинга позволяет осуществлять контроль состояния инженерных систем в режиме реального времени, оперативное управление оборудованием, разграничение доступа к информации. При этом все функции по управлению инженерными системами доступны по средствам пользовательского интерфейса. Для видеоконтроля в помещение установлена система видеонаблюдения, которая позволяет визуально контролировать состояние оборудования и действие людей в помещении Центра.

Применение комплекса. Описанный в данной статье вычислительный комплекс может быть использован для космических исследований, например, для проекта Радиоастрон.

Радиоастрон – международный проект, разрабатываемый в Астрокосмическом центре Физического института им. П.Н. Лебедева, Москва, Россия. Цель проекта - проведение научных радиоастрономических наблюдений с помощью радиотелескопа, смонтированного на космическом аппарате *Спектр-Р*, который создается в НПО им. Лавочкина.

Проект предусматривает запуск космического 10-метрового радиотелескопа на высокоапогейную орбиту спутника Земли (http://www.asc.rssi.ru/radioastron/rus/index.html). Цель проекта состоит в том, чтобы создать совместно с глобальной наземной сетью радиотелескопов единую систему наземно-космического интерферометра для получения изображений, координат и угловых перемещений различных объектов Вселенной с исключительно высоким угловым разрешением (порядка 10^{-6} угловых секунд дуги).

Планируется, что с борта космического корабля по каналу ВИРК (высокоинформативный радиоканал) информация, будет поступать на наземную станцию слежения со скоростью 256 Мб/с. В зависимости от времени сеанса размер файла одного наблюдения может достигать 1.3 ТБ. Кроме потока научной информации, необходимого для построения карт распределения яркости радиоисточников, предусмотрена передача телеметрической информации о системе ориентации, работе радиоастрономических приемников (4 длины волны от 1.35 см до 91 см в двух поляризациях), работе служебных систем, температурном режиме и др. Планируется, что эта информация будет отображаться на мониторах коллективного и индивидуального пользования в режиме реального времени.

Созданный комплекс может быть использован для надежного хранения и корреляционнной обработки научных данных, получаемых с космического и наземных радиотелескопов.

Результаты и перспективы. Итогом проделанной работы стало создание мощного вычислитель-

ного комплекса, состоящего из вычислительного кластера с производительностью 430 ГФлоп/с, основной и резервной систем хранения данных на жестких дисках на 104 ТБ, системы резервного копирования на магнитных лентах на 32 ТБ, канала связи емкостью 1 ГБит/с, соединяющего основную систему хранения данных с резервной. Созданный ЦОНИ предполагается использовать для хранения и обработки массивов данных астрономических наблюдений.

Планируемое развитие ЦОНИ предусматривает наращивание объема хранилища до 200–230 ТБ и вычислительной мощности кластера до 1 Тфлоп/с, а также совершенствование программного обеспечения для увеличения функциональных возможностей центра обработки.

СПИСОК ЛИТЕРАТУРЫ

1. *Есепкина Н.А., Корольков Д.В., Парийский Ю.Н.* Радиотелескопы и радиометры. М.: Наука, 1973.
2. *Томпсон А.Р., Моран Д.М., Свенсон Д.У.* Интерферометрия и синтез в радиоастрономии. М.: Физматлит, 2003.
3. *Рэнд Моримото, Майкл Ноэл, Омар Драуби, Росс Мистри, Ерис Амарис.* Microsoft Windows Server 2008. Полное руководство / Пер. с англ. М.: ООО “И.Д. Вильямс”. 2009.