

Современные кластерные технологии в нашей компании

Аляткин Д.Ф., ОАО "Саратовнефтегеофизика", г. Саратов

Современные кластерные технологии позволяют эффективно использовать вычислительные мощности и ресурсы распределенных систем. В нашем случае это позволяет эффективно использовать ресурсы серверов и дисков. В частности, мы используем кластерные технологии для обработки данных с использованием сложных алгоритмов. Для успешной работы кластера необходимо иметь достаточные ресурсы. В нашей компании, чтобы эффективно использовать ресурсы, необходимо иметь достаточные ресурсы. С учетом того, что прогресс в области информационных технологий движется быстро, эти ресурсы со временем становятся все более и более ограниченными. Поэтому необходимо постоянно расширять и совершенствовать систему, чтобы соответствовать требованиям времени.

Большой недостаток в нашей компании стал дефицит как вычислительных мощностей, так и дискового пространства. Поэтому была необходима модернизация аппаратной части. Раньше в геофизике использовались в основном машины и дорогие RISC-системы. В нашем центре к тому времени флагманом был сервер Sun Enterprise 3500. Он вместе с еще несколькими серверами составлял вычислительное ядро нашей системы. Но у RISC-систем, которые требуют значительных инвестиций и делают компанию зависимой от выбранной архитектуры RISC, уже была достойная альтернатива - вычислительные кластеры, работающие на операционной системе Linux. Кластер - это несколько компьютеров, объединенных между собой и используемых для решения одной общей задачи. Каждый узел кластера обрабатывает свою часть общей задачи. Такие кластеры строятся на базе стандартной архитектуры Intel и открытой операционной системы Linux, поэтому позволяют компании быть независимой от какого-то одного поставщика. Кроме того, производительность такого кластера может быть легко увеличена в случае необходимости путем подключения к нему новых дополнительных серверов. Выбор был сделан в пользу кластерных технологий.

А нас в стране уже есть несколько производителей, которые предлагают кластерные системы. Рассмотрев несколько предложений, мы выбрали компанию R-Style.

Так как дефицит дискового пространства ощущался острее, вначале был развернут один из узлов и дисковый массив R-Style[®] MARSHALL[®] StoreARRAY емкостью 2TB. Чуть позже были развернуты остальные узлы. В итоге построен кластер, работающий под управлением операционной системы Linux, состоящий из 8 узлов на базе процессоров Intel Xeon (два процессора с частотой 3.066 ГГц, 2 GB оперативной памяти, два гигабитных интерфейса). Вся система представляет собой сервер R-Style MARSHALL NP 243г с подключенной к нему дисковой стойкой и семь серверов R-Style MARSHALL NP 244г. Все узлы связаны между собой по гигабитному Ethernet. Связан кластер и с нашим сервером Sun Enterprise 3500. Правда для него существует только оптический гигабитный интерфейс, поэтому в сервер был установлена карта Sbus X1140A, а далее через медиа-конвертор AlliedTelesin он соединяется с новой кластерной системой по гигабитному Ethernet.

На кластере также было развернуто приобретен-

ное программное обеспечение GeoCluster, новая версия которого предоставляет возможность GeoVector[®] и GeoCluster[®] для обработки данных. Sun Enterprise 3500. В результате кластер стал мощным инструментом для обработки данных, который позволяет эффективно использовать ресурсы, как вычислительные, так и дисковые. Кроме того, кластер позволяет эффективно использовать ресурсы, как вычислительные, так и дисковые. В будущем планируется дальнейшее расширение конфигурации и перенос работы программы кластерной операционной системы Linux.

Рост производительности системы позволяет не только строить кластеры и постоянно наращивать их вычислительную мощность, но и организовывать рабочие места, на которых можно выполнять работу, ранее выполняемую только на серверах. Примером может служить графическая станция R-Style Marshall NP 246, оснащенная двумя процессорами Intel Xeon с частотой 3.066 ГГц, двумя гигабайтами оперативной памяти и профессиональной высокопроизводительной графической картой QuadroFX 700. На данной станции проводились тестовые испытания программного обеспечения Epos3 SE компании Paradigm. Например, миграция 3D (в пакете GeoDepth) по одной из съемок 3D, отработанных нашей компанией, выполняется в 4.4 раза быстрее, чем на сервере Sun Enterprise 3500. А на кластере, состоящем из восьми узлов, близких по конфигурации к данной станции, мы получим еще более существенный прирост производительности и сможем обрабатывать гораздо большие объемы 3D, используя более современные алгоритмы миграции, такие как миграция на основе волнового уравнения.

В связи со значительным ростом дискового пространства доступного пользователям и прогнозируемым ростом объемов данных, связанным с ростом наших вычислительных мощностей, сразу же было предусмотрено новое средство резервного копирования и архивирования материалов. Был приобретен накопитель HP StorageWorks Ultrium 460 External. Картриджи для устройств имеют емкость 400 GB, и они предназначены в основном для обслуживания данных дискового массива емкостью 2TB. Данный накопитель дополняет и расширяет возможности нашей компании по резервному копированию и архивированию. Для этих целей уже используются магнитооптика, DVD, Exabyte и HP SuperStore DL180.

Проведенная модернизация существенно увеличила вычислительные мощности компании. Это открывает много новых возможностей. Прежде всего, компания теперь меньше зависит от RISC-систем и может дальше строить и развивать свою вычислительную систему уже на базе стандартной архитектуры Intel и открытой операционной системы Linux. Если возникнет необходимость, вычислительный центр может еще больше увеличить свои вычислительные мощности путем увеличения числа узлов в кластере. В планах компании перевод остального программного обеспечения на Linux для работы на кластере, прежде всего приобретение Epos3 SE. Это позволит более полно использовать возможности кластера и применить современные, требовательные к вычислительным ресурсам, алгоритмы.

ОАО "САРАТОВНЕФТЕГЕОФИЗИКА" ТЕХНОЛОГИИ ОБРАБОТКИ