

ОПТИМИЗАЦИЯ АРХИТЕКТУРЫ ВЫЧИСЛИТЕЛЬНЫХ КОМПЛЕКСОВ НА ЭТАПАХ ПРОЕКТИРОВАНИЯ И ЭКСПЛУАТАЦИИ С УЧЕТОМ СПЕЦИФИКИ ПРИКЛАДНЫХ ЗАДАЧ

В. В. Банкрутенко, М. А. Карнов, Л. А. Копылов

АО «Опытное Конструкторское Бюро Машиностроения им. И. И. Африкантова»,
г. Нижний Новгород

АО «ОКБМ Африкантов» – крупный научно-производственный центр Государственной корпорации «Росатом» с полным циклом разработки, изготовления, поставки надежных, безопасных и экономичных реакторных установок стационарного и транспортного типа, оборудования и систем РУ и АЭС и их сервисного обслуживания. Одним из этапов жизненного цикла разрабатываемого на предприятии изделия является расчетно-техническое обоснование (РТО) проекта. В рамках расчетно-технического обоснования необходимо решать различные инженерные задачи, связанные с расчетом, анализом и стимуляцией физических процессов.

Инструментом решения таких задач являются САЕ-системы. В основе работы САЕ-систем лежат различные численные методы: метод конечных элементов, метод конечных разностей, метод конечных объемов и др.

Примером современной САЕ-системы является система ANSYS (рис. 1), которая представляет собой набор программных пакетов конечно-элементного анализа и применяется при решении задач в сферах механики конструкций, механики жидкости и газа, теплообмена, электродинамики, акустики.

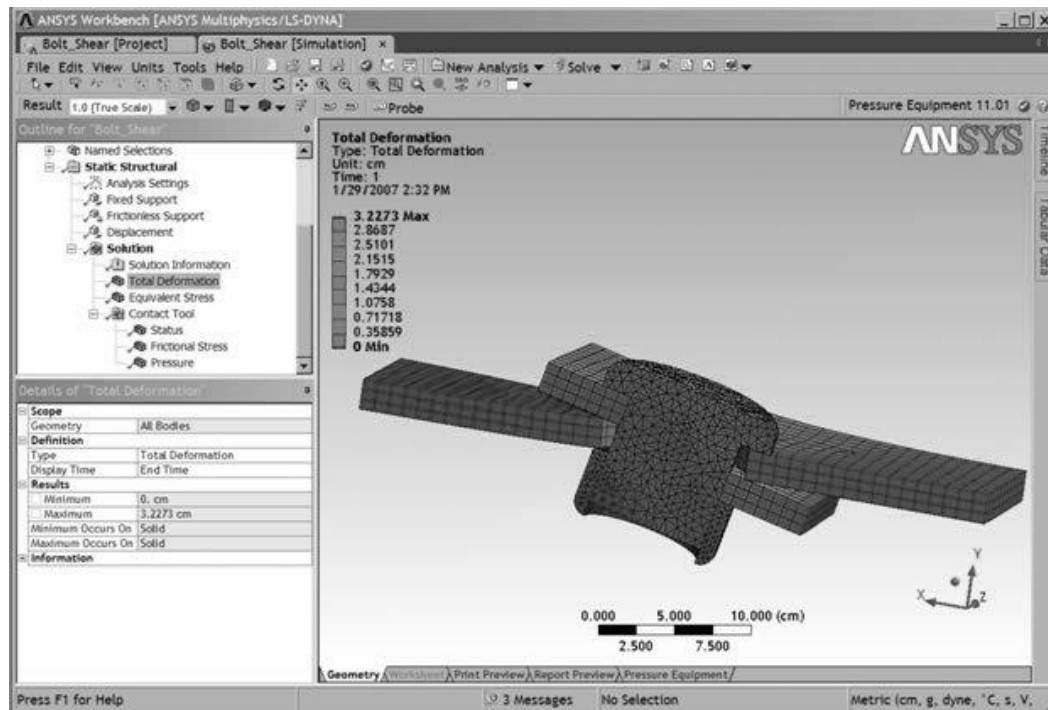


Рис. 1. Пример интерфейса системы Ansys

Типичный процесс работы пользователя с САЕ-системой состоит из следующих этапов:

1) подготовка исходных данных для расчета с использованием графических инструментов препроцессинга САЕ-системы;

2) постановка подготовленной задачи на расчет в виде запуска расчетных модулей САЕ-системы;

3) анализ полученных результатов расчета инструментами постпроцессинга САЕ-системы.

Результаты одного расчета могут быть использованы в качестве исходных данных для другого, например результаты тепловых расчетов могут использоваться для формирования исходных данных прочностных расчетов.

Специфическими особенностями работы с САЕ-системами в условиях большого разнообразия решаемых задач являются:

- большой объем генерируемых, обрабатываемых и хранимых данных (размер файлов результатов в рамках одной задачи может достигать десятков и сотен гигабайт);

- длительное время расчетов с использованием значительных вычислительных ресурсов (могут быть задействованы десятки вычислительных ядер процессоров и сотни гигабайт оперативной памяти);

- для визуализации графической информации (исходных данных для расчета, результатов расчета) могут потребоваться мощные графические ускорители;

- зачастую рабочие станции пользователей не обладают требуемыми ресурсами.

Для успешного применения САЕ-систем при решении больших задач в условиях многопользовательского режима работы используются высокопроизводительные вычислительные кластеры (HPC кластер). Это группа серверов (рис. 2), функционирующая в круглосуточном режиме, на которой производится запуск ресурсоемких расчетных задач пользователей, обеспечивается хранение больших объемов данных, предоставляются ресурсы мощных графических ускорителей для визуализации графической информации.



Рис. 2. Внешний вид вычислительного кластера

При проектировании архитектуры вычислительного кластера необходимо учитывать специфику решаемых прикладных задач, особенности используемого прикладного программного обеспечения и уделить особое внимание характеристикам следующих функциональных подсистем:

- подсистема хранения данных – например, дисковые массивы, обеспечивающие отказоустойчивость и быстрый доступ к данным, обоснованность применения твердотельных дисков (SSD), организация многоуровневого хранения, архивного хранения и т. п.;

- вычислительное поле (или несколько полей), состоящее из узлов, предназначенных для запуска расчетных задач в пакетном режиме. Узлы поля могут различаться по характеристикам, в за-

висимости от специфики решаемых задач. Одни задачи могут требовать много параллельных потоков исполнения и запускаются на множестве узлов, а другие плохо приспособлены к высокой параллельности и требуют большого объема памяти на одном узле. Например, для Ansys пакеты вычислительной гидродинамики (CFX) выигрывают от параллельного запуска на множестве потоков, а пакеты механики сплошных сред (Mechanical) не получают от многопоточного запуска преимуществ и, как правило, требуют много оперативной памяти и выполняются в пределах одного вычислительного узла;

- сетевая инфраструктура с высокой пропускной способностью и низкой задержкой. Наиболее широкое применение в настоящее время нашла технология InfiniBand с пропускной способностью десятки гигабит в секунду. В феврале 2016 года анонсировано решение Oracle Enhanced Data Rate (EDR) InfiniBand Fabric, позволяющее создавать конвергентную сетевую инфраструктуру с пропускной способностью 100 Гбит/с;

- инструментальные (фронт-энд) серверы, предназначенные для работы с исходными данными и результатами расчетов, преимущественно в режиме наглядной графической визуализации. Эти серверы оснащаются производительными многоядерными центральными процессорами, большим объемом памяти и мощной графической подсистемой. Они обеспечивают одновременную работу нескольких пользователей с графическими интерфейсами CAE-систем;

- управляющие инфраструктурные серверы, обеспечивающие управление работой кластера;
- каналы связи для организации работы территориально распределенных пользователей;
- системы разграничения доступа и защиты информации.

При эксплуатации вычислительных кластеров сотрудникам ИТ служб необходимо решать ряд сложных задач, связанных с поддержкой инфраструктуры.

Для управления большим количеством серверов важно предусмотреть соответствующие средства для автоматизации развертывания и обновления системного и прикладного программного обеспечения. Существуют комплексные инструменты для решения задач развертывания и поддержки серверной инфраструктуры. Одним из таких инструментов является система XCAT (Extreme Cloud Administration Toolkit), обеспечивающая развертывание, массовое конфигурирование и обновление программного обеспечения серверов.

Функционирование вычислительных кластеров изначально предполагает многопользовательский режим работы на доступных пользователю серверах. Для крупных систем необходимостью становится использование средства централизованного управления учетными записями пользователей. К таким средствам относятся разнообразные службы каталогов, такие как LDAP [1] и Active Directory (в составе семейства операционных систем MS Windows Server). С помощью этих инструментов можно добавлять, изменять, блокировать и удалять учетные записи пользователей в единой для всех серверов базе данных.

Для работающей в круглосуточном режиме инфраструктуры необходимо развернуть средства мониторинга состояния аппаратного и программного обеспечения. Система мониторинга должна решать следующие задачи:

- своевременное оповещение специалистов ИТ служб (инженеров и администраторов), с целью опережающей реакции на возникающие проблемы;
- накопление статистической информации для анализа работы серверной инфраструктуры в динамике и принятии решений на основе исторических данных;
- наличие единой панели обзора всей инфраструктуры для быстрого принятия решения и поиска корневой причины возникающих проблем.

Есть множество инструментов для создания служб мониторинга, как коммерческих (проприетарных), так и с открытым исходным кодом. Задачи по оповещению и обзору всей текущей ситуации решаются, например, такими системами как Nagios и Zabbix. Сбор статистики может осуществляться средствами массового сбора и накопления метрик, например, Ganglia [2], Monit, Cacti.

Система с открытым исходным кодом Nagios (рис. 3) позволяет гибко конфигурировать параметры проверок, настраивать различные средства оповещения (электронная почта, мгновенные сообщения, SMS и др.), развертывать распределенный и отказоустойчивый мониторинг.

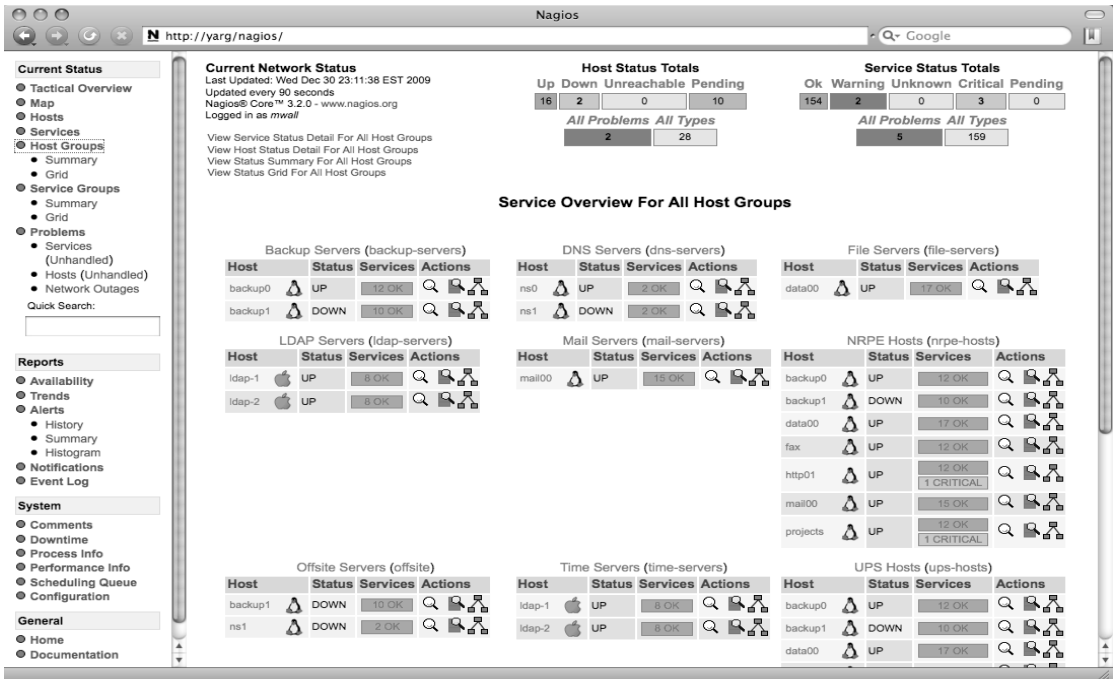


Рис. 3. Интерфейс системы Nagios

Система с открытым исходным кодом Ganglia (рис. 4) позволяет в непрерывном режиме собирать множество метрик работы серверов, и расширяема модулями для сбора дополнительных метрик. Статистика сохраняется в единой базе данных и веб-интерфейс позволяет отображать статистические данные в виде графиков в разных масштабах времени.

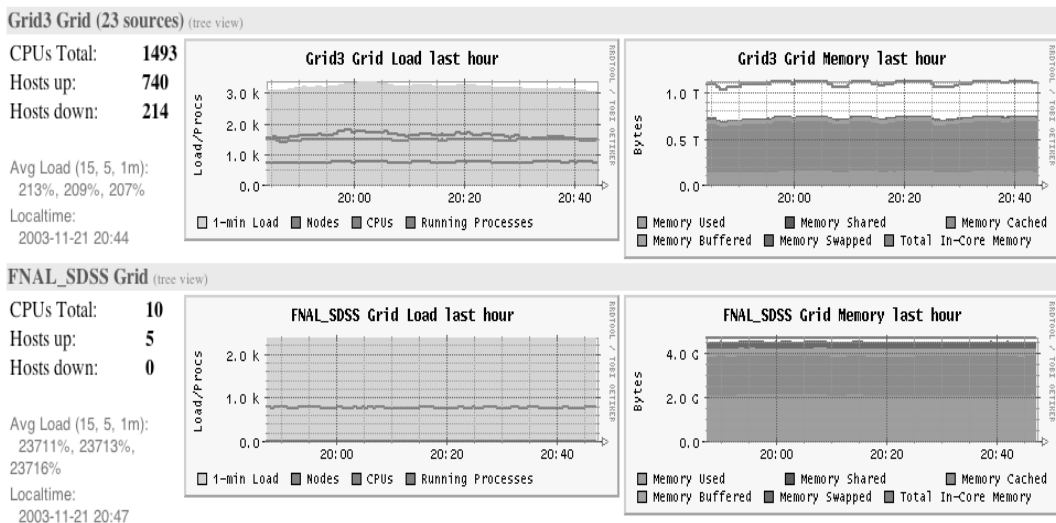


Рис. 4. Интерфейс системы Ganglia

Важной проблемой в условиях многопользовательского режима на кластерах является устранение конкуренции пользователей за вычислительные ресурсы. На современных вычислительных кластерах эта проблема решается с помощью специальных систем – менеджеров распределенных ресурсов. Основная функция таких систем – распределение вычислительных задач среди доступных вычислительных ресурсов. Примерами таких систем являются система с открытым исходным кодом Torque, система с открытым исходным кодом SLURM, HPC Job Manager (входит в состав Microsoft HPC Pack). В подобных системах пользователи не напрямую запускают вычислительные

программы, а ставят запуск программ в очередь менеджера вычислительных ресурсов. Планировщик менеджера вычислительных ресурсов определяет доступные для запуска ресурсы на кластере и производит отложенный запуск задач. Пользователь может отслеживать статус своих задач на кластере и видеть распределение вычислительных ресурсов.

Работа пользователя с вычислительным кластером требует от него специфических навыков взаимодействия с удаленной системой, таких как:

- работа в режиме командной строки при постановке и мониторинге состояния задач в случае отсутствия интерфейса графической оболочки у менеджера ресурсов кластера;
- использование инструментов препроцессинга и постпроцессинга САЕ-систем на удаленном рабочем столе инструментального сервера;
- работа с файлами (файлы исходных данных, результатов расчета) на хранилищах данных вычислительных кластеров.

Одной из задач сотрудников ИТ служб, занимающихся сопровождением вычислительных кластеров, является обучение пользователей вышеперечисленным навыкам, которое обеспечивает не просто возможность работы с принципиально отличным от привычного им способа взаимодействия с прикладными системами на персональных компьютерах, но и дает им понимание особенностей такой работы, что позволяет максимально эффективно использовать все технические возможности и особенности конкретного оборудования используемых вычислительных кластеров.

В качестве вывода необходимо отметить, что на сегодняшний день обеспечение РТО проектов является комплексом задач, требующим от сотрудников ИТ служб, занимающихся сопровождением вычислительных кластеров, как глубоких знаний выбранной в качестве платформы операционной системы, так и квалификации в области проектирования и сопровождения сложной серверной инфраструктуры [3] и понимания специфики работы САЕ-систем.

Литература

1. Butcher M. Mastering OpenLDAP. Packt Publishing, 2007.
2. Massie M., Li B., Nicholes B., Vuksan V. Monitoring with Ganglia. O'Reilly Media, 2012.
3. Эви Н., Гарт С., Трент Х., Бэн У. Unix и Linux: руководство системного администратора. М.: ООО «И. Д. Вильямс», 2012.

СОЗДАНИЕ ПАРАЛЛЕЛЬНОЙ ВЕРСИИ ПРОГРАММЫ ДЛЯ МОДЕЛИРОВАНИЯ МНОГОКОМПОНЕНТНОЙ ФИЛЬТРАЦИИ ПРИ РАЗРАБОТКЕ МЕСТОРОЖДЕНИЙ НЕФТИ И ГАЗА С ИСПОЛЬЗОВАНИЕМ DVM-СИСТЕМЫ

В. А. Бахтин, А. В. Королев¹, Н. В. Поддержюгина

Институт прикладной математики им. М. В. Келдыша РАН, г. Москва

¹Федеральный научный центр НИИ системных исследований РАН, г. Москва

Введение

DVM-система, созданная в ИПМ им. М. В. Келдыша РАН, предназначена для разработки параллельных программ научно-технических расчетов на языках C-DVMH и Fortran-DVMH. Эти языки используют единую модель параллельного программирования (DVMH-модель) и являются рас-