

# Организация вычислительной инфраструктуры для решения задач инженерного анализа

В.В. Банкрутенко, М.А. Карпов, Л.А. Копылов, Д.Е. Мохин

**Аннотация.** В статье представлена задача построения высокопроизводительных вычислительных кластеров для работы с современными CAE-системами. Приведены основные проблемы, стоящие перед IT-службами при решении данной задачи, а также способы их решения. Рассмотрены принципы проектирования HPC-кластеров с учетом специфики решаемых задач. Показаны основные подсистемы вычислительного кластера. Дан обзор инструментов, связанных с автоматизацией работы, поддержкой и обслуживанием сложной серверной инфраструктуры.

**Ключевые слова:** CAE, HPC-кластер, системное администрирование, мониторинг, менеджмент распределенных ресурсов.

## Введение и постановка задачи

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 симуляцией физических процессов.

Инструментом решения таких задач являются CAE-системы. В основе работы CAE-систем лежат различные численные методы: метод конечных элементов, метод конечных разностей, метод конечных объемов и др.[1].

Примером современной CAE-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.

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

1. подготовка исходных данных для расчета с использованием графических инструментов препроцессинга CAE-системы;

2. постановка подготовленной задачи на расчет в виде запуска расчетных модулей CAE-системы;

3. анализ полученных результатов расчета инструментами постпроцессинга CAE-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.

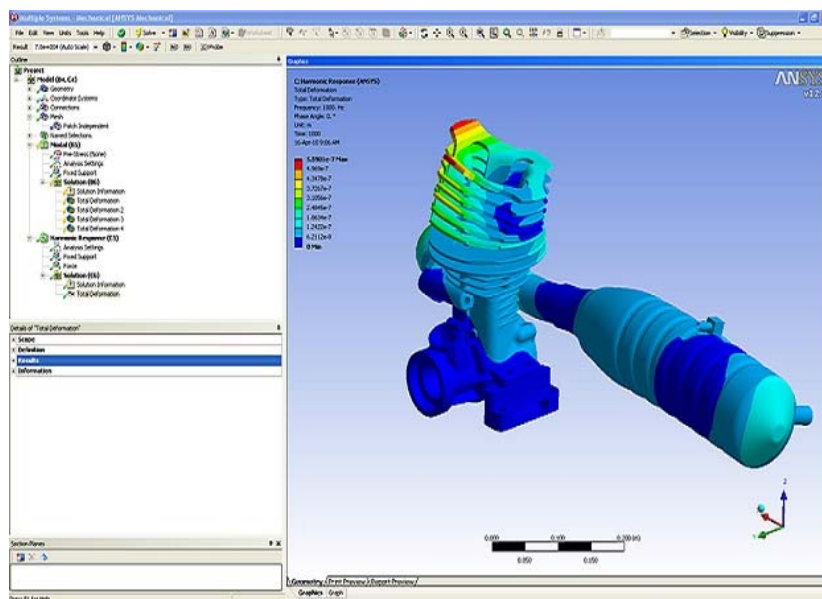


Рис. 1. Пример интерфейса системы ANSYS

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.

Перед IT-службами предприятия стоит задача проектирования, развертывания и обслуживания вычислительно инфраструктуры, которая бы удовлетворяла вышеприведенной специфике САЕ-систем.

## Решение задачи

Для успешного применения САЕ-систем при решении больших задач в условиях многопользовательского режима работы используются

высокопроизводительные вычислительные кластеры (HPC cluster). 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.

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 зависимости 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 вычислительных 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 [2] с пропускной способностью десятки гигабит в секунду. В феврале 2016 года анонсировано решение Oracle Enhanced Data Rate (EDR) InfiniBand Fabric [3], 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 САЕ-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.

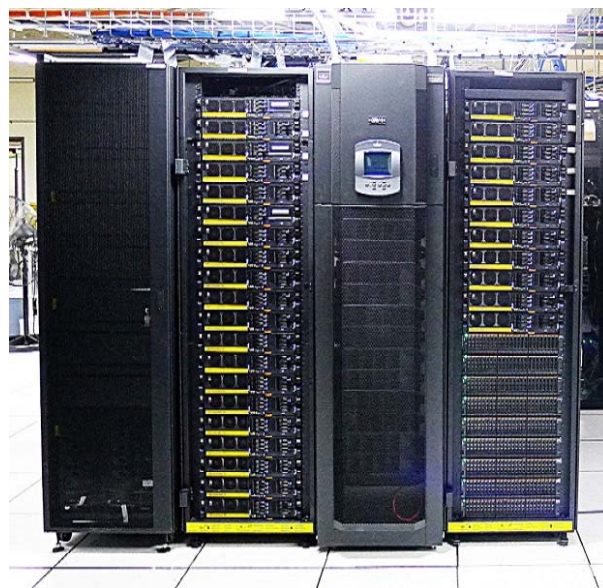


Рис. 2. Внешний вид вычислительного кластера

Одним из таких инструментов является система XCAT (Extreme Cloud Administration Toolkit) [4], обеспечивающая развертывание, массовое конфигурирование и обновление программного обеспечения серверов.

Функционирование вычислительных кластеров изначально предполагает многопользовательский режим работы на доступных пользователям серверах. Для крупных систем необходимостью становится использование средства централизованного управления учетными записями пользователей. К таким средствам относятся разнообразные службы каталогов, такие как LDAP (реализация – OpenLDAP [5]) 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, Monit, Cacti.

Система с открытым исходным кодом Nagios [6] (Рис. 3) позволяет гибко конфигурировать параметры проверок, настраивать различные средства оповещения (электронная почта, мгновенные сообщения, SMS и др.), развертывать распределенный и отказоустойчивый мониторинг.

Система с открытым исходным кодом Ganglia [7] (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.

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,

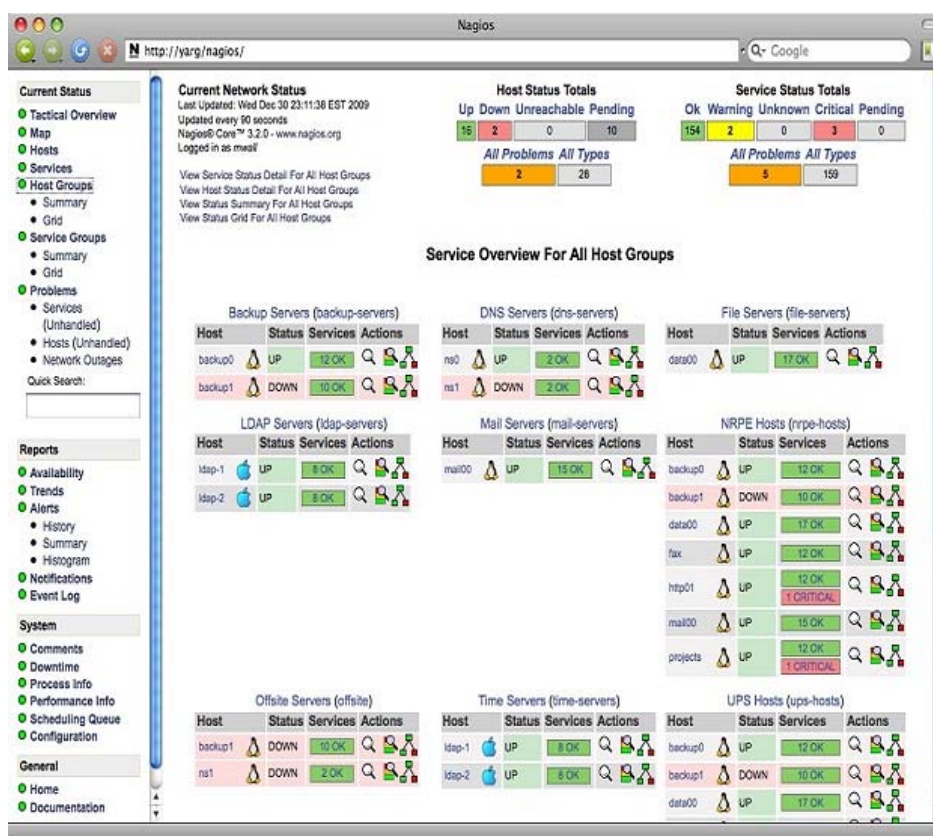


Рис. 3. Интерфейс системы Nagios



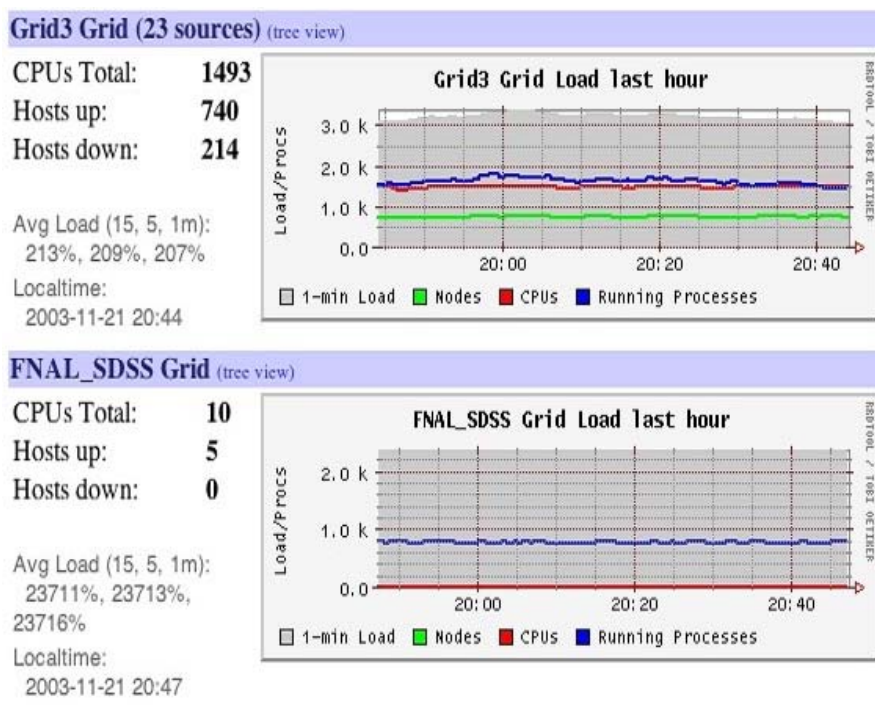


Рис. 4. Интерфейс системы Ganglia

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 спосо-

ба 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.

### Заключение

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 и понимания специфики работы САЕ-систем.

### Литература

1. Computer-aided engineering [Электронный ресурс] [[https://ru.wikipedia.org/wiki/Computer-aided\\_engineering](https://ru.wikipedia.org/wiki/Computer-aided_engineering)]. Ссылка проверена 31.03.2017.
2. Infiniband [Электронный ресурс] [<https://ru.wikipedia.org/wiki/InfiniBand>]. Ссылка проверена 31.03.2017.

3. Oracle EDR Infiniband Fabric [Электронный ресурс] [https://www.oracle.com/ru/corporate/pressrelease/edr-infiniband-fabric-20160222.html]. Ссылка проверена 31.03.2017.
4. XCAT [Электронный ресурс] [https://xcat.org/]. Ссылка проверена 31.03.2017.
5. OpenLDAP [Электронный ресурс] [http://www.openldap.org/]. Проверено 31.03.2016.
6. Nagios [Электронный ресурс] [https://www.nagios.org/]. Ссылка проверена 31.03.2017.
7. Ganglia [Электронный ресурс] [http://ganglia.info/]. Ссылка проверена 31.03.2017.
8. Matt Butcher. Mastering OpenLDAP. 2007 Packt Publishing – 482 с.
9. Matt Massie, Bernard Li, Brad Nicholes, and Vladimir Vuksan. Monitoring with Ganglia. O'Reilly Media, 2012 – 254 с.
10. Немет Эви, Снайдер Гарт, Хейн Трент, Уэйли Бэн. Unix и Linux: руководство системного администратора. “И.Д. Вильямс”, 2012 – 1312 с.

**Банкрутенко Владимир Викторович.** Главный специалист АО "ОКБМ Африкантов". Окончил ГГУ им. Лобачевского в 1972 году. Кандидат технических наук. Количество печатных работ: 39. Область научных интересов: информационные технологии. E-mail: bankrutenko@okbm.nnov.ru

**Карпов Михаил Александрович.** Инженер по ИТ АО "ОКБМ Африкантов". Окончил НГТУ им. Алексея в 2010 году. Область научных интересов: информационные технологии. E-mail: karpov\_ma@okbm.nnov.ru

**Копылов Леонид Александрович.** Начальник бюро АО "ОКБМ Африкантов". Окончил ННГУ им. Лобачевского в 1997 году. Область научных интересов: информационные технологии. Количество печатных работ: 4. E-mail: kopylov@okbm.nnov.ru

**Мохин Денис Евгеньевич.** Ведущий инженер АО "ОКБМ Африкантов". Окончил ННГУ им. Лобачевского в 2004 году. Количество печатных работ: 1. Область научных интересов: информационные технологии. E-mail: mohin@okbm.nnov.ru

## Organization of computational infrastructure for computer-aided engineering analysis

V.V. Bankrutenko, M.A. Karpov, L.A. Kopylov, D.E. Mokhin

**Abstract.** The paper presents the problem of high-performance computer clustering to operate the current CAE systems. The main issues being the matter of IT services when solving this problem as well as the methods of their solution are given. The principles of HPC clusters designing in view of the specific character of the problems to be solved are considered. The main subsystems of the compute cluster are shown. The overview of the tools for complex server infrastructure operation computerization, support, and maintenance is given.

**Keywords:** CAE, HPC-cluster, system administration, monitoring, distributed resource management.

### References

1. Computer-aided engineering [Online resource]. [https://ru.wikipedia.org/wiki/Computer-aided\_engineering]. Link checked 03.03.2017.
2. Infiniband [Online resource] [https://ru.wikipedia.org/wiki/InfiniBand]. Link checked 03.03.2017.
3. Oracle EDR Infiniband Fabric [Online resource] [https://www.oracle.com/ru/corporate/pressrelease/edr-infiniband-fabric-20160222.html]. Link checked 03.03.2017.
4. XCAT [Online resource] [https://xcat.org/]. Link checked 03.03.2017.
5. OpenLDAP [Online resource] [http://www.openldap.org/]. Link checked 03.03.2017.
6. Nagios [Online resource] [https://www.nagios.org/]. Link checked 03.03.2017.
7. Ganglia [Online resource] [http://ganglia.info/]. Link checked 03.03.2017.
8. Matt Butcher. Mastering OpenLDAP. 2007 Packt Publishing – 482 s.
9. Matt Massie, Bernard Li, Brad Nicholes, and Vladimir Vuksan. Monitoring with Ganglia. O'Reilly Media, 2012 – 254 s.
10. Evy Nemeth, Garth Snyder, Trent R. Hein, Ben Whaley. Unix and Linux system administration handbook. “I.D. Williams”, 2012 – 1312 s.

**Bankrutenko Vladimir Victorovich:** Place of work: JSC "OKBM Afrikantov", department of IT development and implementation  
Position: Chief Specialist Science degree: Candidate of Technical Sciences Institution: GSU them. Lobachevsky, 1972 Number of publications: 39 Research interests: IT. E-mail: bankrutenko@okbm.nnov.ru

**Karpov Mikhail Aleksandrovich:** Place of work: JSC "OKBM Afrikantov", department of IT development and implementation  
Position: IT Engineer Institution: NGTU them. Alekseeva, 2010 Research interests: IT. E-mail: karpov\_ma@okbm.nnov.ru

**Kopylov Leonid Aleksandrovich:** Place of work: JSC "OKBM Afrikantov", department of IT development and implementation  
Position: Head Office Science degree: Master of mathematics Institution: UNN them. Lobachevsky, 1997 Number of publications: 4  
Research interests: IT. E-mail: kopylov@okbm.nnov.ru

**Mokhin Denis Evgenievich:** Place of work: JSC "OKBM Afrikantov", department of IT development and implementation  
Position: Leading Engineer IT Institution: UNN them. Lobachevsky, 2004 Number of publications: 1 Research interests: IT  
E-mail: mohin@okbm.nnov.ru