

Разработка технологических решений по созданию и использованию специализированных аппаратно-программных комплексов на базе грид-инфраструктур суперкомпьютерных ресурсов

М.Е. Кулешова, Н.И. Парамонова, Н.Н. Парамонов, О.П. Чиж

Объединенный институт проблем информатики НАН Беларуси

Приводятся результаты практического использования аппаратно-программных комплексов на базе распределенной грид-инфраструктуры суперкомпьютерных ресурсов (АПК-ГРИД) для решения различных ресурсоемких прикладных задач. В качестве инструментальной среды для проведения исследований использовался экспериментальный образец центра обработки данных (ЦОД) в составе национального ГРИД-сегмента на базе кластера «СКИФ-GPU».

1. Введение

Распределенные грид-инфраструктуры широко используются для хранения и обработки данных при реализации различных проектов. Современные распределенные вычислительные системы, как правило, имеют три уровня базового программного обеспечения (ПО): системное ПО; прикладное ПО; ПО промежуточного уровня (ПО ПУ). При выборе компонентов создаваемых АПК-ГРИД и проведении исследований их функциональных возможностей необходимо комплексно оценивать принятые технические решения: тип архитектуры суперкомпьютерного ресурса (например, кластерная), тип вычислительных узлов суперкомпьютера, типы используемых процессоров (x86, GPU, APU и др.), сетевые решения и т.п.

В настоящей статье приведены результаты практического использования АПК-ГРИД для решения различных ресурсоемких прикладных задач. В качестве инструментальной среды для проведения исследований использовался экспериментальный образец ЦОД в составе национального ГРИД-сегмента на базе кластера «СКИФ-GPU» [1]. В кластере «СКИФ-GPU» 34 вычислительных и 2 управляющих узла. Каждый ВУ кластера содержит два 4-х ядерных процессора Nehalem Intel Xeon X5570 2,93 GHz с архитектурой x86-64 и графический процессор NVIDIA GeForce GTX 295. Пиковая производительность кластера «СКИФ-GPU» для графических процессоров при выполнении вычислений с плавающей запятой одинарной точности – не менее 60,0 Тфлопс, двойной точности – не менее 5,0 Тфлопс. Реальная производительность опытного образца суперкомпьютера «СКИФ-GPU» на тесте Linpack на 36 узлах без графических процессоров - 3025 Гфлопс при показателе эффективности кластера $U_{eff} = 89,62\%$ от пиковой 3375.36 Гфлопс. По результатам тестирования кластер «СКИФ-GPU» был включен в 13 редакцию списка Top50 от 21.09.2010 г. под номером 45.

2. Суперкомпьютерное моделирование конструктивно-технологических решений при проектировании перспективных конкурентоспособных шин отечественной разработки

Экспериментальный образец специализированного ЦОД в составе национального ГРИД-сегмента первоначально был создан для выполнения работ в рамках хоздоговора между ОИПИ НАН Беларуси и ОАО «Белшина». При создании ЦОД был разработан комплекс технологических решений для обеспечения функционирования суперкомпьютерных ресурсов и грид-сервисов (корневых сервисов) национальной грид-сети [2], задействованных в численном решении задач договора. Для выполнения работ был создан UNICORE-сайт SKIF_GPU, использующий в качестве вычислительных ресурсов 30 8-ядерных узлов кластера «СКИФ-GPU», который входит в состав национальной грид-сети. Для обеспечения возможности проведения су-

перкомпьютерного моделирования для обработки решений по созданию шинных изделий были выполнены установка и настройка пакета конечно-элементного анализа LS-DYNA на грид-сайте SKIF_GPU, создан репозиторий данных задач компьютерного моделирования, разработанная технология запуска задач компьютерного моделирования на грид-сайте SKIF_GPU.

В результате выполнения работ были созданы технологические решения, обеспечивающие предоставление пользователям ОИПИ НАН Беларуси и ОАО «Белшина» надёжного и безопасного инструментария для проведения инженерных расчётов в национальной грид-сети. Была проведена проверка функционирования следующих компонентов ЦОД:

- 1) Клиент – Пользовательский интерфейс (User Interface);
- 2) Шлюз (gateway);
- 3) Сайт REGISTRY (Shared Registry Site);
- 4) Сервер UNICORE/X;
- 5) База данных пользователей UNICORE XUADB;
- 6) Интерфейс исполнительной системы TSI (Target System Interface);
- 7) Система мониторинга UNICORE-сайта.

Проверка функционирования сервисов UNICORE-сайта SKIF_GPU выполнялась на управляющей машине (УМ) кластера «СКИФ-GPU».

На UNICORE-сайте SKIF_GPU запущен интерфейс исполнительной системы (TSI) `tsi_linux_torque`, позволяющий выполнять задания через систему управления заданиями `torque` на 30 8-ядерных ВУ кластера «СКИФ-GPU» следующей конфигурации: ОС Linux Fedora 12 (x86_64); два 4-ядерных процессора Intel Xeon X5570 2.93GHz; 24 Гб ОЗУ; InfiniBand QDR. На пользовательских компьютерах было установлено клиентское ПО UNICORE, включающее графический клиент UNICORE RichClient и файловый менеджер UNICORE (UFM).

В результате проведенных работ удалось решить связанную задачу образования температурных полей и определения напряженно-деформированных состояний крупногабаритных шин, выпускаемых ОАО «Белшина». Также на суперкомпьютерах национальной грид-сети было успешно проведено моделирование процесса стендовых испытаний шин.

3. АПК на основе библиотеки параллельной обработки изображений RIPL

Библиотека RIPL (Parallel Image Processing Library), предназначенная для использования в высокопроизводительных многопроцессорных системах с параллельной архитектурой, разработана в ОИПИ НАН Беларуси. Основное назначение – обработка изображений, полученных в системах дистанционного зондирования Земли [3].

Возможности программного комплекса на основе библиотеки параллельной обработки изображений RIPL тестировались в рамках программы СГ «Космос-НТ» на базе кластерной конфигурации POLYGON_GPU для отработки технологии параллельной обработки больших объемов информации данных ДЗЗ, представляющих собой высококачественные мультиспектральные изображения. В число задач цифровой обработки изображений входят: восстановление зашумленных изображений; реконструкция объектов; повышение качества и сглаживание; обнаружение яркостных перепадов и линий; текстурный анализ; сегментация области и распознавание образов; преобразование полутоновых изображений; картографическая генерализация; геометрические трансформации и другие.

Кластерная система «POLYGON_GPU» была создана на основе 4 вычислительных узлов (ВУ) суперкомпьютера «СКИФ-GPU». В графическом режиме управление работой программного комплекса осуществляется с помощью графического редактора Gimp.

Библиотека параллельных функции содержит функции, позволяющие улучшить качество изображений, т.е. изменить яркость, контрастность, убрать зашумление изображений, усилить одну и подавить другую информацию, подчеркнуть края, а также осуществить поиск особенных зон на изображениях.

Программно-аппаратная реализация на базе RIPL имеет архитектуру клиент-сервер, где серверная компонента реализует запрашиваемые операции параллельной обработки исходного изображения, а клиентская компонента предоставляет оператору пользовательский интерфейс.

В комплексе реализованы следующие основные операции над растровыми снимками: усредняющий фильтр (Average), медианный фильтр (Median), гауссово размытие (Gaussian blur), обострение/подчеркивание краев (Unsharpen mask), гауссово искажение (Gaussian Noise), выравнивание гистограммы (Equilize Histogram), растягивание диапазона (Stretch Range), эрозия (Erode), дилатация (Dilatation), фильтр на основе преобразования Фурье (Fourier), корреляционные методы поиска изображений по шаблону. Возможно добавление любого пользовательского фильтра. Для решения задачи поиска эталонного фрагмента на изображении была использована открытая библиотека OpenCV.

В таблицах 1 и 2 приведены результаты исследования масштабируемости задачи поиска нужных участков на изображениях при удаленном запуске задачи на полигоне.

В эксперименте использованы два полутоновых файла (файл 1, файл 2, градации серого) формата tif. В таблице 1 приведены результаты (время поиска эталонных зон в секундах) при последовательном увеличении количества параллельных процессов на управляющей машине полигона. Как видно из таблицы с увеличением задействованных ресурсов время поиска существенно уменьшается и решение задачи масштабируется.

Таблица 1. Время выполнения при увеличении числа процессов

Кол-во процессов	1	2	3	4	5	6	7
Файл 1 13152*1478	23,972	15,049	12,401	9,867	9,375	9,032	7,185
Файл 2 16384*1509	32,525	21,061	16,598	15,020	13,607	12,267	12,007

В таблице 2 приведены результаты эксперимента (время поиска эталонных зон в секундах) с последовательным добавлением вычислительных узлов полигона. Из таблицы видно, что 16 параллельных процессов на двух узлах достаточно для оптимального решения данной задачи. Дальнейшее увеличение количества процессов не дает уменьшения времени выполнения задачи.

Таблица 2. Время выполнения при увеличении числа узлов

Количество процессов	8	16	24	32	40
Файл 1 13152*1478	7,352	4,422	4,322	4,65	4,52
Файл 2 16384*1509	12,010	8,093	8,227	8,675	8,288

Результаты тестирования программного комплекса на основе библиотеки PIPL, предназначенного для параллельной реализации типичных операций, осуществляемых над изображениями ДЗЗ или над произвольными большими массивами данных, позволяют сделать следующие выводы:

- комплекс содержит полностью распараллеленный программный код с хорошими показателями масштабируемости;
- комплекс построен по схеме клиент-сервер с возможностью удаленной работы пользователей;
- комплекс имеет возможность интеграции функций комплекса в пользовательский интерфейс программ обработки изображений (GIMP, различные ГИС и т. п.);
- функционирование комплекса может быть осуществлено как на отдельном многоядерном компьютере, кластерном суперкомпьютере, так и в локальной сети организации;
- комплекс предоставляет возможность разработки эффективных пользовательских функций обработки данных без знания технологий параллельного программирования.

Наиболее эффективным является скриптовый режим работы комплекса, позволяющий обработать по заранее заданному набору фильтров большое количество снимков (микрокадров)

для улучшения их качества. Также применение RPL оправдано при поиске объектов по шаблонам в больших базах данных.

4. АПК обработки и предоставления данных ДЗЗ на базе суперкомпьютерных и грид-технологий (АПК Грид-ДЗЗ)

Основная цель работы на первом этапе – проведение предварительных исследований и анализа принципов выполнения прикладных ресурсоемких задач обработки данных ДЗЗ на основе реальных цифровых фотографий поверхности Земли.

В качестве программного обеспечения промежуточного уровня (ПО ПУ) в АПК Грид-ДЗЗ используется ПО ПУ UNICORE [4]. Для данного ПО ПУ в ОИПИ НАН Беларуси в рамках программы СГ «СКИФ-ГРИД» был разработан комплект дистрибутива для установки на вычислительные ресурсы национальной грид-сети, работающие под операционными системами Linux и Windows.

Минимальный набор сервисов грид-инфраструктуры должен включать следующие службы: вычислительные сервисы; информационные сервисы; сервисы мониторинга; сервисы учета ресурсов; ресурс-брокеры; элементы хранения данных; сервисы управления виртуальными организациями. Должны также быть предусмотрены корневые грид-сервисы, которые включают в себя следующие основные службы: единый реестр грид-ресурсов (реестр грид-сервисов, реестр пользователей, сервисов и ресурсных центров); служба Shared Registry; служба тестирования; служба мониторинга; служба учета ресурсов.

Главным потребителем ресурсов создаваемой в рамках настоящего проекта грид-инфраструктуры является белорусское Научно-инженерное республиканское унитарное предприятие «Геоинформационные системы» (НИРУП ГИС). На этом предприятии создан Национальный центр приема и первичной обработки космической информации, что позволяет, по сути, отрабатывать функциональные возможности проектируемой грид-инфраструктуры на реальных массивах данных ДЗЗ. В процессе выполнения предварительных исследований определено, что наиболее ресурсоемкой прикладной задачей из предварительного состава, сформированного на основании данных НИРУП ГИС, является задача восстановления 3D изображений на основе 2D цифровых фотографий поверхности Земли. В связи с этим анализ технологических операций обработки данных ДЗЗ проводился на примере технологии построения 3D моделей по изображению.

4.1. Удостоверяющий центр грид-сети ДЗЗ

Удостоверяющий центр формирует на основе инфраструктуры открытых ключей цифровые сертификаты сервисов и пользователей грид-сети ДЗЗ. Удостоверяющий центр грид-сети ДЗЗ входит в состав Операционного центра национальной грид-сети.

UNICORE-сайт Monitoring_RS использует в качестве вычислительных ресурсов 4 8-ядерных узла следующей конфигурации: ОС Linux Fedora 15 (x86_64); два 4-ядерных процессора Intel Xeon X5570 2.93GHz; однопроцессорный видеоадаптер NVIDIA GeForce GTX 480; 24 ГБ ОЗУ; InfiniBand QDR.

На UNICORE-сайте Monitoring_RS запущены следующие сервисы: шлюз (gateway); сервер UNICORE/X; база данных пользователей XUADB; интерфейс целевой системы TSI; система мониторинга (рисунок 1).

Интерфейс исполнительной системы (TSI) tsi_linux_torque позволяет выполнять задания на целевой системе через систему управления заданиями torque на 4-х 8-ядерных вычислительных узлах кластера АПК ГРИД ДЗЗ.



Рис. 1. Система мониторинга грид-сайта Monitoring_RS

4.2. Технологии визуализации данных

Неотъемлемой и важнейшей частью современного анализа различных процессов в любой отрасли естественных наук является визуализация данных. Существует значительное число средств визуализации научных данных, обладающих широкими функциональными возможностями. К сожалению, объём визуализируемых с их помощью данных естественным образом ограничен ресурсами персонального компьютера. В проекте АПК ГРИД-ДЗЗ система визуализации, реализованная на базе технологии клиент-сервер, включает: суперкомпьютерные вычислительные ресурсы в рамках распределенной грид-инфраструктуры; удаленное рабочее место (клиент системы); локальную (или глобальную) сеть. В качестве технологических решений, обеспечивающих работу системы визуализации, был выбран графический протокол VNC и программный пакет x11vnc [5].

Отличительными особенностями данного решения являются:

- 1) возможность запуска серверной компоненты визуализации на любой аппаратной платформе, поддерживающей работу сервера XWindows (NVIDIA, ATI, Intel и т.п.);
- 2) аппаратная поддержка работы приложений, реализованных с использованием технологий OpenGL, CUDA и OPENCL;
- 3) авторизация пользователей с использованием x509 сертификатов, выданных удостоверяющим центром национальной грид-сети;
- 4) возможность работы клиентской компоненты в качестве отдельного приложения ssnvc или через web-интерфейс;
- 5) возможность полноценной работы в высоком разрешении со сложными 3D моделями при использовании обычных 3G каналов мобильной связи;

б) наличие полных исходных кодов продукта с возможностью их модификации.

По сути, используемые средства визуализации позволяют создавать для пользователя полноценный удаленный графический стол, работающий на суперкомпьютерных ресурсах (сервере визуализации) распределенной вычислительной среды. При этом используются единые механизмы информационной безопасности (пользовательские и серверные цифровые сертификаты) уже реализованные в национальной грид-сети. Использование данной системы визуализации позволяет целиком или частично избежать копирования громоздких результатов суперкомпьютерного моделирования на рабочее место пользователя.

5. Экспериментальная технология восстановления 3D-моделей местности

На данном этапе технологические схемы создания 3D-моделей местности реализованы на основе использования программного комплекса – Agisoft PhotoScan Professional 1.0.4, представляющего собой специализированное программное обеспечение для фотограмметрии и работы с аэрофотоснимками.

Программа Agisoft PhotoScan позволяет по набору снимков объекта восстановить его текстурированную 3D модель. Эта цель достигается посредством выполнения следующих этапов обработки:

- 1) Загрузка фотографий, определение положений и параметров внешнего и внутреннего ориентирования камер.
- 2) Выравнивание фотографий. На этой стадии программа определяет положение и ориентацию камеры для каждого кадра и строит разреженное облако общих точек в 3D пространстве модели и данные о положении и ориентации камер.
- 3) Построение плотного облака точек на основании рассчитанных положений камер и используемых фотографий.
- 4) Построение трехмерной полигональной модели объекта, которая описывает форму объекта на основании плотного облака точек.
- 5) Текстурирование объекта и/или построение ортофотоплана (рисунок 2).

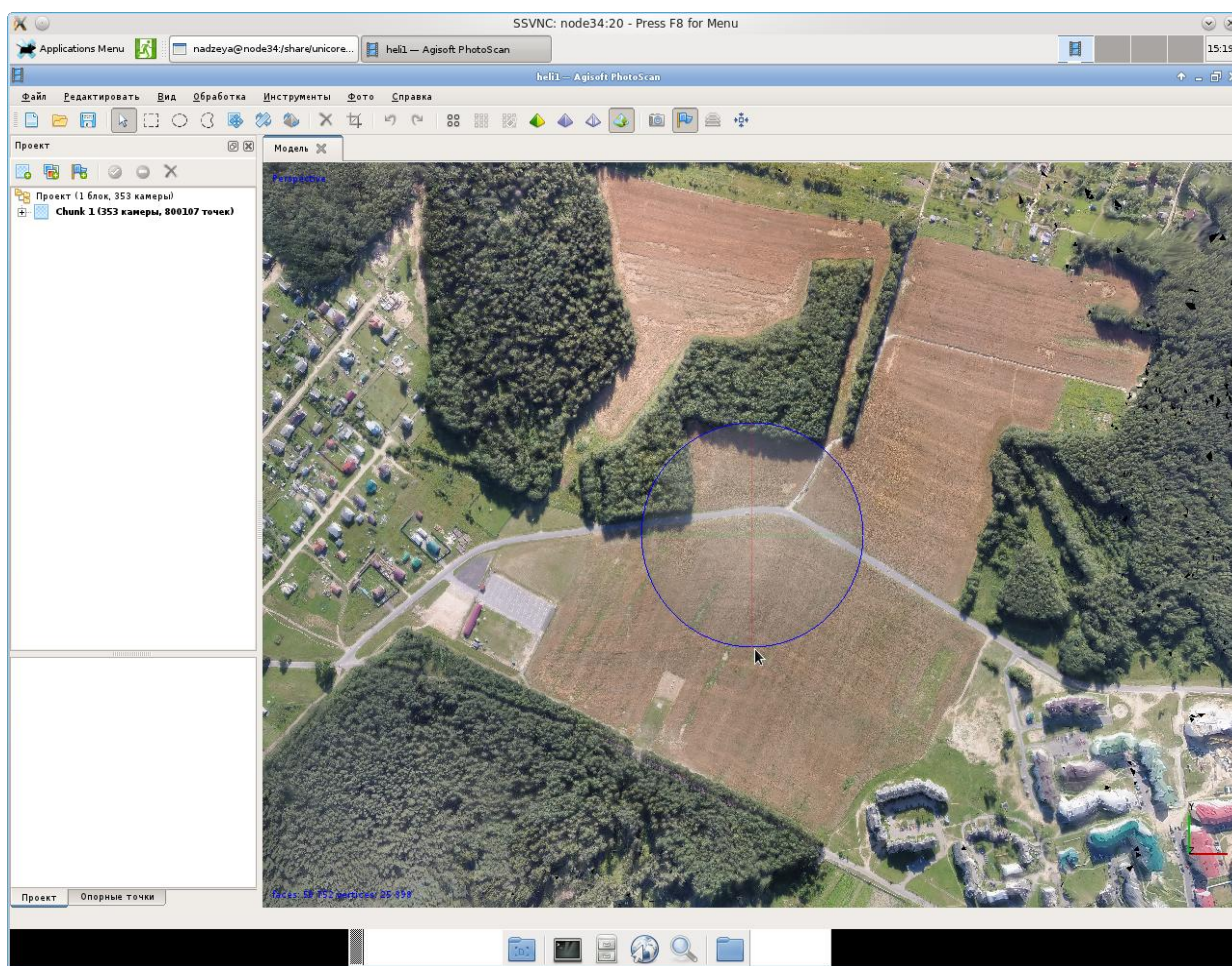


Рис. 2. Текстурированная модель объекта (вид из системы визуализации)

Применяемый программный продукт имеет все необходимые функциональные возможности для создания 3D-моделей местности по аэрокосмическим снимкам. Создаваемая с его помощью 3D – модель имеет приемлемое для многих потребителей качество. Процесс построения трехмерной модели на экспериментальном стенде при работе на одном узле с графическим процессором выполнялся от 20 до 100 мин в зависимости от количества обрабатываемых снимков и заданного качества модели (150-300 снимков). При этом использование на узле графических ускорителей с технологией OPENCL позволяет значительно (в 8-10 раз) сократить время выполнения отдельных этапов построения 3D-моделей (наиболее эффективно при построении плотного облака точек).

Применение грид-технологий для организации доступа к суперкомпьютерным ресурсам и наличие у конечного пользователя НИРУП ГИС собственного беспилотного аппарата позволяют реализовать полную цепочку построения 3D-моделей: от выезда на местность для получения снимков до редактирования на месте построенной модели средствами 3D-редакторов или других ГИС.

6. Заключение

В представленной работе приведены практические результаты применения аппаратно-программных комплексов на базе распределенной грид-инфраструктуры суперкомпьютерных ресурсов для решения ресурсоемких прикладных задач. В качестве таких прикладных задач выделены следующие: моделирование напряженно-деформированных состояний крупногабаритных шин, обработка больших массивов растровых данных и построение трехмерных моделей местности по снимкам ДЗЗ. Приведено описание технологических решений для удаленной визуализации данных суперкомпьютерного моделирования.

Литература

1. Анищенко В.В., Криштофик А.М., Парамонов Н.Н., Чиж О.П. Белорусские кластеры семейства «СКИФ»: состояние и перспективы развития // Новороссийск: Труды Всероссийской суперкомпьютерной конференции Научный сервис в сети Интернет: экзафлопсное будущее. 19-24 сентября 2011. С. 23-27.
2. Операционный центр национальной грид-сети Республики Беларусь. URL: <http://noc.grid.basnet.by/> (дата обращения: 25.05.2015).
3. Парамонова Н.И., Чиж О.П. Программный комплекс для параллельной обработки данных ДЗЗ // Минск: Доклады четвертой Международной научной конференции «Суперкомпьютерные системы и их применение» SSA'2012. 23-25 октября 2012 г. Мн: ОИПИ НАН Беларуси, С. 95-102.
4. UNICORE. Distributed computing and data resources. URL: <http://www.unicore.eu/> (дата обращения: 25.05.2015).
5. x11vnc. VNC server for real X displays. URL: <http://www.karlrunge.com/x11vnc/> (дата обращения: 25.05.2015).