

Визуализация данных на параллельных вычислительных комплексах

Карпов А. Н.
GDT Software Group,
Тула, 300034, ул. Демонстрации, 27-79
Телефон: +7 0872 20 53 44 / FAX: +7 0872 20 68 77
URL: <http://www.cfd.ru/>
E-mail: karpov@cfdr.ru

Аннотация

Существует проблема визуализации большого объема данных в реальном времени. Источником данных может являться расчет задачи, выполняемый на вычислительном кластере. Проблема визуализации заключается в необходимости синхронного сбора и обработки данных, получаемых со всех узлов вычислительного кластера. Она была решена путем динамического прореживания данных на узлах вычислительного комплекса и последующей их синхронной передачи на удаленную машину для визуализации. Это дало возможность системе визуализации ScientificVR (SVR) в реальном времени отображать данные, поставляемые кластерной версией пакета газодинамического моделирования GasDynamicsTool (GDT).

Ключевые слова располагать сразу за аннотацией.

Ключевые слова: Визуализация, on-line визуализация, параллельный вычислительный комплекс.

1. ВВЕДЕНИЕ

При компьютерном моделировании важную роль играет задача интерпретации данных. Визуализация является одним из мощных средств, позволяющих эффективно решать эту задачу. Графические представления в форме цветовых распределений или функциональных зависимостей значительно упрощают восприятие блоков данных человеческим мозгом. Это особенно важно при обработке больших трехмерных массивов, с которыми, в основном, приходится иметь дело при численном моделировании.

Большой интерес визуализация представляет для пакетов прикладных программ численного моделирования. Одной из задач визуализации является отображение результатов непосредственно в процессе математического моделирования, возможность наблюдать развитие моделируемого физического процесса на протяжении всего расчета. Такой вид визуализации также называется on-line визуализация. Но задача визуализации в реальном времени и запись видеороликов, иллюстрирующих динамику моделируемых процессов, достаточно сложна в случае распределенной вычислительной сети.

В данной статье рассмотрены преимущества визуализации результатов расчета в реальном времени, а также пути построения системы визуализации в случае распределенного вычислительного комплекса. Примером успешной реализации подобной системы является пакет визуализации ScientificVR (SVR), а также кластерный вариант пакета

газодинамического моделирования GasDynamicsTool (GDT), являющийся источником данных для визуализации. [1]

2. ПРЕИМУЩЕСТВА ОТОБРАЖЕНИЯ ДАННЫХ В РЕАЛЬНОМ ВРЕМЕНИ

Визуализация в реальном времени позволяет оперативно оценивать корректность производимых вычислений и, в случае необходимости, прерывать их с целью доработки счетной модели. Это важное средство для параллельных пакетов численного моделирования. Такие пакеты, как правило, работают с большими задачами, обработка которых может занимать многие дни. Если не производить наблюдение на промежуточных этапах, то ошибку, допущенную в процессе постановки задачи, можно будет обнаружить только после завершения расчета. Что в свою очередь ведет к потере машинного времени, весьма дорогого в случае использования кластерных систем. Останов задачи, сохранение рассчитанных данных и их просмотр в эти промежутки времени также нерационален, так как требует слишком много ручной работы. При этом происходят дополнительные потери времени. По этому наиболее рациональным является отображение результатов расчета в режиме on-line.

Следует отметить, что отображение данных на каждом шаге расчета не имеет смысла. Вполне достаточны промежутки в несколько секунд или минут, в зависимости от счетной задачи. С одной стороны этого достаточно для наблюдения за динамикой развития процессов, а с другой - экономит машинное время. Изменяя частоту обновления изображения, можно в наиболее интересные моменты наблюдать за расчетом более детально. Именно такой подход осуществляется в системе визуализации ScientificVR (SVR) при работе с кластерной версией пакета газодинамического моделирования GasDynamicsTool (GDT).

Еще одним важным преимуществом визуализации в реальном времени является возможность записи видеороликов. В случае отсутствия механизма on-line визуализации, получение видеороликов технически сложно в своей реализации. Возникает необходимость периодически сохранять данные, с целью в дальнейшем получить из них соответствующие кадры видеоролика. Или писать отдельную подсистему сбора данных и ее записи. В случае наличия механизма on-line визуализации, запись видеоролика является простым его расширением.

Иногда on-line визуализация является удобным средством вспомогательных расчетов. Ряд задач не требует никаких выходных данных кроме изображения или видеоролика, отображающего процесс в динамике. В таких задачах

требуется качественно посмотреть на процессы, происходящие при заданных условиях. А уже затем, оценив качественную картину, создать более адекватный проект для расчета и снятия различных характеристик.

В качестве примера преимущества визуализации в реальном времени можно привести следующую задачу. Имеются данные эксперимента, но необходимо получить более детальную информацию по ряду параметров. Для начала следует проверить, насколько созданная модель согласуется с данными эксперимента. При такой постановке задачи визуализация в реальном времени - наиболее подходящее решение. Так как имеется возможность сразу наблюдать результаты расчета и оценивать их достоверность. А в случае явной ошибки - прекратить вычисления, не увеличивая затрат времени на продолжение расчета. На рисунке 1 приведено сравнение результатов эксперимента и расчетной картины смоделированного процесса в различные моменты времени. Очевидно, что результаты моделирования качественно совпадают с экспериментом, а, следовательно, в данном случае можно продолжить расчет с целью более подробного изучения протекающего процесса (как в количественном, так и в качественном аспекте).

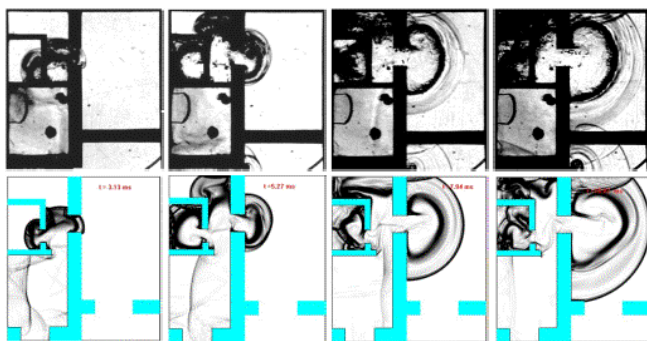


Рис. 1. Сравнение эксперимента и результатов расчета

3. ВЫБОР МЕХАНИЗМА ВИЗУАЛИЗАЦИИ В РЕАЛЬНОМ ВРЕМЕНИ НА ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ КОМПЛЕКСАХ

Вычислительные мощности современных компьютеров растут, их стоимость падает, и всё большее распространение получают высокопроизводительные многопроцессорные вычислительные системы (МВС). На таких системах возможно проведение расчётов, требующих десятков и даже сотен гигабайт оперативной памяти. Огромные объемы данных, генерируемые в ходе параллельного счета, безусловно требуют очень больших затрат вычислительных ресурсов при графической обработке. Эти затраты связаны с передачей информации на сервер визуализации, а также с обработкой "сырых" данных и непосредственным рендерингом сложных изображений. Особенно сложна визуализация в случае ее динамического отображения в режиме on-line.

Рассмотрим несколько подходов к осуществлению визуализации на параллельных вычислительных комплексах.

3.1 Визуализация на выделенном узле

При первом подходе визуализацию данных осуществляет выделенный узел. Данные при этом целиком передаются на визуализирующий узел (рисунок 2).

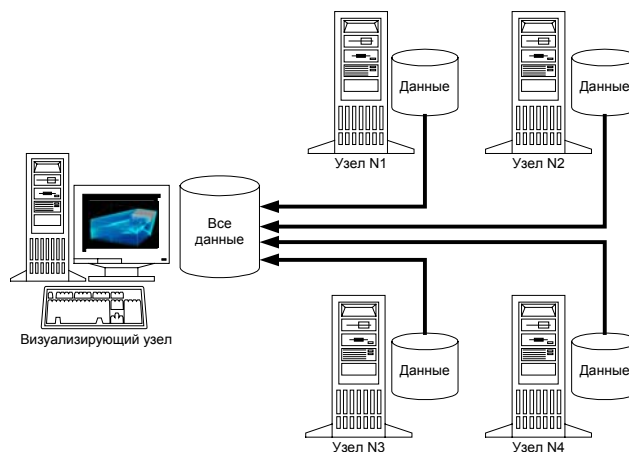


Рис. 2. Визуализация на выделенном узле

Этот способ можно применять, только если объём визуализируемых данных относительно невелик. Дело в том, что если объём данных превысит размер памяти визуализирующего узла (а на практике ещё раньше), визуализация станет невозможной или (при использовании виртуальной памяти) недопустимо медленной. Кроме того, сама обработка большого количества данных потребует значительного времени. И, наконец, для обеспечения быстрой передачи данных с вычислительных узлов на визуализирующий потребуется сетевой канал с высокой производительностью.

Для большей наглядности перечислим преимущества и недостатки данного метода по пунктам. К достоинствам такого подхода можно отнести:

- наибольшую простоту в реализации;
- возможность динамической визуализации.

Однако данный подход обладает наибольшим количеством недостатков:

- низкая производительность визуализации из-за больших объёмов данных;
- невозможность визуализации объёмов данных, превышающих размер памяти визуализирующего узла;
- низкая скорость передачи данных между МВС и визуализирующим узлом из-за большого объёма данных.

3.2 Последовательная визуализация на выделенном узле

Второй подход является разновидностью первого. В этом случае данные с узлов передаются на визуализирующий узел не все сразу, а добавляются в очередь. Визуализирующий узел осуществляет визуализацию по частям и далее формирует общее изображение (рисунок 3).

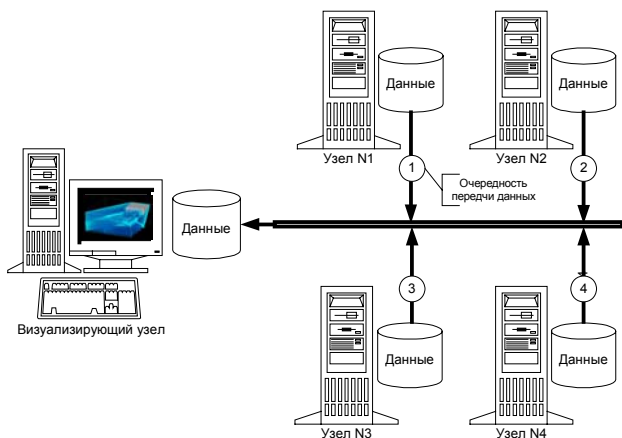


Рис. 3. Последовательная визуализация на выделенном узле

При этом снимается проблема с переполнением памяти на визуализирующем узле, однако остальные недостатки предыдущего подхода сохраняются.

Достоинства:

- возможность визуализации больших объёмов данных.

Недостатки:

- низкая производительность визуализации из-за больших объёмов данных;
- низкая скорость передачи данных между МВС и визуализирующим узлом из-за большого объёма данных;
- снижение производительности счётных узлов из-за ожиданий при постановке данных в очередь визуализации;
- сложность объединения изображений в единое целое.

3.3 Визуализация на счётных узлах

При третьем подходе каждый счётный узел осуществляет визуализацию своей собственной части данных (рисунок 4), а затем части изображения передаются на визуализирующий узел, где они объединяются в единое изображение.

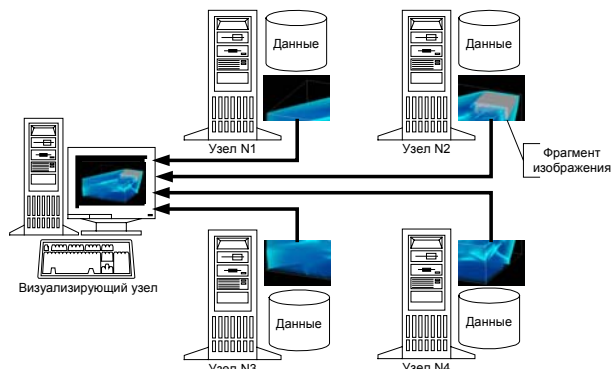


Рис. 4. Визуализация на счётных узлах

Таким образом, не тратится время на пересылку большого объёма данных на визуализирующий узел, а осуществляется только пересылка полученной части изображения. Проблема переполнения памяти на визуализирующем узле также разрешается. Кроме того, при подготовке изображения на

вычислительном узле можно принимать во внимание физические и иные особенности визуализируемых процессов.

Однако данному подходу присущи и недостатки. Во-первых, объединение изображения в единое целое представляет собой в общем случае весьма нетривиальную задачу [2]. Например, при визуализации изолиниями, одна и та же изолиния может проходить через разные части счётной области, а на изображении это должна быть одна единственная изолиния. Подобные проблемы возникают и при визуализации линий тока, изоповерхностей и т.п. Так на рисунке 5 приведена грубая иллюстрация проблемы склеивания изображений с изоповерхностями.

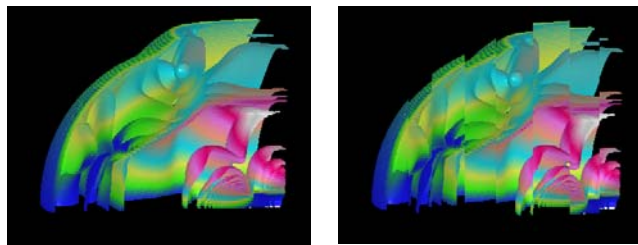


Рис. 5. Проблема объединения изображения изоповерхностей (слева - правильное изображение; справа - склеенное изображение)

При осуществлении визуализации на счётных узлах с одной стороны ускоряется процесс визуализации за счёт распараллеливания, а с другой - замедляется процесс расчёта.

При реализации визуализации на счётных узлах приходится также сталкиваться с трудностями технического характера. Как правило, счётные узлы не оборудованы программно-аппаратными средствами для визуализации, например, отсутствуют видеокарты и не установлены соответствующие библиотеки программ. Это особенно актуально при использовании на узлах специально облегчённых ядер операционных систем, обеспечивающих только самые необходимые средства. Все это ведёт к уменьшению производительности визуализации.

Достоинства:

- при визуализации можно учесть характерные особенности задачи;
- возможна визуализация больших объёмов данных, в том числе динамическая;
- высокая скорость визуализации.

Недостатки:

- приостановка расчёта для визуализации, снижение производительности вычислений;
- сложность объединения изображений в единое целое;
- технические трудности с осуществлением визуализации на счётных узлах.

3.4 Визуализация с предварительной обработкой данных на счётных узлах

Часть проблем, связанных с большими объемами данных при визуализации, можно решить за счет подхода, заключающегося в рассылке на вычислительные узлы, где размещены данные, некоторых общих правил их обработки,

например, правил выборки из распределенного массива. Тогда объем пересылаемой информации резко уменьшается. Схематически четвертый способ представлен на рисунке 6.

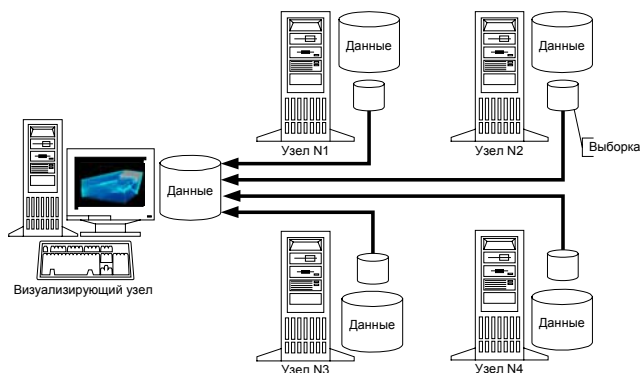


Рис. 6. Визуализация с предобработкой на счётных узлах

Поскольку при этом подходе сама визуализация осуществляется на отдельном узле, а на счётных узлах происходит только первичная обработка, то не возникает необходимости в специальных программно-аппаратных средствах на счётных узлах. Предварительная обработка осуществляется параллельно, что уменьшает время её выполнения, однако это снижает эффективность счёта.

Достоинства:

- высокая скорость визуализации;
- при визуализации можно учесть характерные особенности задачи;
- возможна визуализация больших объёмов данных, в том числе динамическая.

Недостатки:

- снижение производительности вычислений из-за затрат на предварительную обработку данных.

4. РЕАЛИЗАЦИЯ ВЫБРАННОЙ СХЕМЫ ВИЗУАЛИЗАЦИИ И РЕЗУЛЬТАТЫ

Схема визуализации с предобработкой на счётных узлах успешно реализована в версии пакета GasDynamicsTool, предназначенной для работы на многопроцессорных вычислительных комплексах, и позволяющей осуществлять динамическую визуализацию данных объёмом, значительно превышающим размер оперативной памяти визуализирующего узла.

Так с ее помощью был записан видеоролик выхода снаряда из ствола, рассчитанный на кластерной системе [3]. Объем данного проекта составил несколько сотен миллионов трехмерных ячеек (рисунок 7).

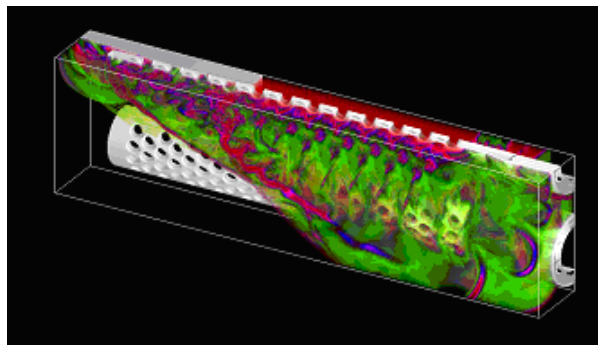


Рис.7. Распределение давления при выходе снаряда из системы со спаренными стволами

5. ЗАКЛЮЧЕНИЕ

Подводя итоги, можно сказать, что визуализация является одним из больших, важных и быстро развивающихся научных инструментов анализа объектов, как реального мира, так и математически моделируемых процессов. Достаточно интересной и сложной проблемой визуализации является ее применение в случае распределенной вычислительной сети. В исследовании этого направления сделано много шагов, но многие задачи как теоретического, так и инженерного плана ждут своего решения.

6. БИБЛИОГРАФИЯ

- [1] А.В. Зибаров, Д.Б. Бабаев, А.А. Миронов, А.В. Медведев, А.Н. Карпов, И.Ю. Комаров, П.В. Константинов. Применение пакета ScientificVR для визуализации результатов трехмерных газодинамических расчетов. // XII Международная конференция по вычислительной механике и современным прикладным программным системам (ВМСППС'2003) 30 июня - 5 июля 2003 г., Владимир, Россия
- [2] D.B. Babayev, A.A. Mironov, A.N. Karpov, I.Yu. Komarov, P.V. Konstantinov, A.V. Zibarov. Semitransparent voxel graphics realization in the ScientificVR visualisation package // 4 Pacific Symposium on Flow Visualization and Image Processing, 3-5 June 2003, Chamonix, France.
- [3] Zibarov A.V., Babayev D.B., Mironov A.A., Komarov I.J., Konstantinov P.V., Medvedev A.V. and Karpov A.N. Twin Barrel Artillery System Function // Journal of Visualization, Vol. 6, No. 2 (2003) 94

Об авторе

Карпов Андрей Николаевич – магистр техники и технологии по направлению «Системный анализ и управление». Является инженером-программистом в АНО «GDT Software Group». Адрес: г. Тула, Тула, 300034, ул. Демонстрации, 27-79. Телефон: +7 0872 20 53 44 / FAX: +7 0872 20 68 77 E-mail: karpov@cfд.ru

Data visualization on parallel computer systems

Abstract

There is a problem of large data array real time visualization. The source of the data can be the problem calculation produced on a cluster. The problem of visualization concerns the importance of synchronous accumulation and treatment of the data collected from all the cluster nodes. The problem was solved with the help of the dynamic data retrieval on the calculating complex nodes followed by their synchronous transference to the visualizing machine. This permits the visualization system ScientificVR (SVR) to reflect in the real time the data supplied by the cluster version of the gas-dynamic modeling package GasDynamicsTool (GDT).

Keywords: *Visualization, on-line visualization, computer cluster.*

About the author

Karpov Andrey has master's degree of technics and technologies in the field of system analysis and management. He is a software Designer at GDT Software Group.

Contact e-mail: karpov@cfed.ru.