

Решение задач фотореалистичной компьютерной графики на базе защищённой инфраструктуры суперкомпьютерного центра коллективного пользования*

Б.М. Шабанов¹, А.П. Овсянников¹, А.В. Баранов¹, Е.А. Киселёв¹,
С.А. Лещев¹, Б.В. Долгов¹, Д.Г. Гуменный², Д.Л. Шуров²

¹Межведомственный суперкомпьютерный центр РАН – филиал ФГУ ФНЦ НИИСИ
РАН, ²Студия визуальных эффектов CGF

В настоящей работе рассматриваются вопросы организации защищённой инфраструктуры суперкомпьютерного центра. Авторами предложен способ защиты конфиденциальных данных пользователей путем организации выделенного сегмента суперкомпьютера на основе виртуальных локальных сетей Ethernet и разделов сети Infiniband. В работе предлагается технология построения масштабируемой гетерогенной вычислительной среды на ресурсах суперкомпьютерного центра в контексте решения задач по созданию фотореалистичной компьютерной графики. Приводятся результаты экспериментальных расчетов.

Ключевые слова: фотореалистичная компьютерная графика, суперкомпьютеры, вычислительные среды, защищенная инфраструктура, параллельная обработка

1. Введение

В кинематографе для обеспечения качественной и фотореалистичной компьютерной графики требуются значительные вычислительные ресурсы. Как правило, вычислительного потенциала студий, занимающихся визуальными эффектами, не всегда достаточно для выполнения крупномасштабных, с точки зрения объёма выполняемой работы, проектов. В таких случаях часть вычислительной нагрузки, связанной с обработкой графики, перераспределяется на ресурсы суперкомпьютерных центров коллективного пользования (ЦКП). Как известно, создаваемые кинематографические материалы носят конфиденциальный характер. Потеря или кража таких материалов может нанести значительный материальный и репутационный ущерб как правообладателю, так и разработчику. Поэтому при решении вышеуказанных задач на суперкомпьютерных ресурсах ЦКП важно обеспечивать защиту обрабатываемой информации.

Межведомственный суперкомпьютерный РАН (МСЦ РАН) является одним из крупнейших центров коллективного пользования (ЦКП), обеспечивающим исследователей современными высокопроизводительными вычислительными ресурсами. Ресурсы МСЦ РАН используются для решения задач сотрудников научных учреждений, подведомственных ФАНО России, участников научных программ Министерства образования и науки Российской Федерации, грантодержателей РФФИ и федеральных научных центров.

При предоставлении вычислительных ресурсов пользователям важно учитывать специфику решаемых ими задач. Если для решения задачи достаточно стандартного набора программного обеспечения и конфигурации вычислительного и коммуникационного оборудования, пользователи получают доступ к вычислительным ресурсам МСЦ РАН на

* Работа выполнена за счёт бюджетных средств в рамках государственного задания.

правах общей очереди. Очередность выполнения заданий пользователей организуется системами пакетной обработки, например, СУППЗ или SLURM [1].

Если к составу и структуре программно-аппаратного обеспечения предъявляются особые требования, возникает необходимость выделения изолированной инфраструктуры для ограниченного круга пользователей.

К числу задач, требующих создания изолированной инфраструктуры, относятся задачи создания фотореалистичной компьютерной графики для кинематографа: шейдинга (создание объектов графики), текстурирования (реалистичная передача объектов и их освещённости в зависимости от положения источников света, расположения теней и фактуры поверхности), расчёт отображения и задач поведения различных физических явлений и объектов (огонь, дым, лава, вода и т.д.).

Необходимость создания изолированной инфраструктуры для таких задач обусловлена в первую очередь требованиями конфиденциальности обрабатываемой на вычислительных ресурсах информации, а также необходимостью использования специализированного программного обеспечения. При этом с целью предотвращения возможных утечек информации, изоляция выделяемых для решения задач ресурсов должна осуществляться на уровне коммуникационной среды, на уровне системы хранения данных и на уровне вычислительных ресурсов.

При создании изолированной инфраструктуры важно обеспечить возможность её масштабирования по запросу студии. Указанное требование обусловлено недетерминированным технологическим процессом создания фотореалистичной компьютерной графики для кино.

В настоящей работе рассмотрена технология построения масштабируемой изолированной среды для решения задач создания фотореалистичной компьютерной графики. Представленная в работе технология не требует перекоммутации сетевого оборудования и внесения изменений в структуру вычислительных ресурсов ЦКП, позволяет динамично изменять объём используемых вычислительных ресурсов по требованию, а также применима для решения других задач конфиденциального характера.

2. Ресурсы МСЦ РАН для высокопроизводительных вычислений

В настоящее время в МСЦ РАН представлены вычислительные ресурсы, отличающиеся по производительности, по способу обработки данных и архитектуре. Традиционно все пользователи ЦКП получают доступ к двум суперкомпьютерам: МВС-100К и МВС-10П.

Суперкомпьютер представляет собой совокупность вычислительных узлов, объединённых рядом сетей: обмена информацией во время счёта, доступа к системе хранения, управления заданиями, управления вычислительными узлами (техническими средствами) и их мониторинга. Помимо вычислительных узлов, в состав суперкомпьютера входят узлы управления, узлы доступа пользователей, узлы доступа системы хранения данных, также подключённые к вышеперечисленным сетям.

Процесс обработки информации пользователем на ресурсах суперкомпьютера заключается в подготовке задания (программ и данных), загрузки его в систему хранения, запуска задания на выполнение, в процессе которого информация считывается и записывается на систему хранения, выгрузки информации с системы хранения.

МВС-100К состоит из 1275 вычислительных узлов, каждый из которых оснащён двумя четырёхядерными/шестиядерными процессорами Intel Xeon, 8/16 ГБ оперативной памяти. Для объединения узлов кластера в единое решающее поле используется технология Infiniband QDR.

Суперкомпьютер МВС-10П имеет несколько сегментов, сегмент на узлах RSC Tornado и сегмент на узлах RSC Petastream. В состав сегмента МВС-10П Tornado входит 207

вычислительных узлов. Каждый вычислительный узел имеет в своём составе 2 процессора Xeon E5-2690, 64 ГБ оперативной памяти и два сопроцессора Intel Xeon Phi 7110X.

В состав сегмента МВС-10П Petastream входят 8 вычислительных узлов под управлением Intel Xeon E5-2667 с 48 ГБ оперативной памяти и сопроцессором Intel Xeon Phi 7120D с 16 Гбайт оперативной памяти.

Вычислительные ресурсы МВС-10П имеют значительно большую производительность в сравнении с ресурсами МВС-100К, но ввиду их небольшого количества они сильно загружены заданиями пользователей. Поэтому разработка и отладка параллельных алгоритмов, как правило, производится на ресурсах МВС-100К, а трудоёмкие вычисления – на ресурсах МВС-10П.

Домашние каталоги пользователей и проектные каталоги размещаются на СХД NetAPP FAS и доступны с любого суперкомпьютера ЦКП.

3. Защита данных

Выбор методов защиты данных пользователей при их обработке на суперкомпьютерной системе общего пользования прежде всего связан с тем, как сам пользователь определяет угрозы безопасности своей информации, какова возможность нарушения конфиденциальности, целостности или доступности обрабатываемой информации, и приведёт ли нарушение хотя бы одного из указанных свойств безопасности информации к неприемлемому ущербу обладателю информации.

Для выбора необходимых мер защиты информации была построена модель угроз безопасности информации, которая содержит систематизированный перечень угроз безопасности данных, обусловленных преднамеренными или непреднамеренными действиями физических лиц, создающими условия для нарушения безопасности данных, приводящему к ущербу.

При рассмотрении модели безопасности были выделены следующие возможные нарушители безопасности информации (люди, группы людей, имеющие целью нанесение ущерба владельцам информации — пользователям вычислительной системы, владельцам вычислительной системы, а также программы (процессы) ими запущенные):

внутренние — сотрудники вычислительного центра, обслуживающие вычислительный кластер в соответствии с их правами и полномочиями (системные и сетевые администраторы, операторы, инженеры и др.);

внешние — сотрудники вычислительного центра, которым не предоставлены права по доступу к вычислительному кластеру; пользователи вычислительного кластера из других организаций, не имеющие прав доступа к защищаемой информации; все остальные нарушители, не имеющие доступа к вычислительному кластеру.

Для защиты от внутреннего нарушителя прежде всего используются регламентирующие действия, такие как установление ответственности за разглашение конфиденциальной информации для сотрудников суперкомпьютерного центра и заключение соглашения о нераспространении конфиденциальной информации между суперкомпьютерным центром и организацией (пользователем), установление ограничений полномочий при действиях с обслуживаемой системой, и др. Необходимо также строгий учёт, в том числе техническими средствами, действий системных администраторов, операторов и других уполномоченных лиц по изменению параметров (настроек) вычислительной системы.

Для защиты от внешнего нарушителя, перехватывающего трафик по линии связи пользователя (киностудии) с суперкомпьютерным центром, следует использовать криптографические средства защиты, например, средства построения виртуальных частных сетей.

Для защиты от внешних нарушителей – пользователей суперкомпьютера из других организаций и сотрудников центра, не относящихся к категории внутренних нарушителей, –

необходимо максимально закрыть доступ к информационным потокам и обрабатывающим их системам.

Для всех остальных пользователей ЦКП должен быть закрыт доступ к хранящимся в домашних и проектных каталогах данным, должна быть исключена сама возможность их доступа на узлы, предназначенные для проекта или перехват передаваемых данных, в случае если злоумышленник, притворяющийся легальным пользователем, взломает узел доступа пользователей или узел, на котором выполняются пользовательские задания.

Проведённый авторами анализ актуальных публикаций [5-8] показал, что в современных ЦКП применяются два способа изоляции вычислительных ресурсов: на основе физической или логической сегментации сети.

Наиболее эффективным методом защиты является физическая изоляция части вычислительной системы, предназначенной для конфиденциальных расчётов. Для этого требуется разделить все сети на физически независимые сегменты, что довольно затратно: для отделения сегментов сети Ethernet могут потребоваться дополнительные коммутаторы, а для сетей Infiniband и OmniPath, топология которых представляет собой «толстое дерево», - переключения очень большого числа линий связи. При физической сегментации сети невозможно обеспечить масштабируемость вычислительных ресурсов в режиме реального времени.

В нашем случае был выбран метод логической сегментации сетей. Сети, построенные по технологии Ethernet, были сегментированы на основе механизма виртуальных локальных сетей (VLAN). Коммуникационная сеть Infiniband была сегментирована на основе разделов (partition). Такой способ сегментации сети позволяет обеспечить требуемый уровень изоляции ресурсов пользователей и позволяет масштабировать вычислительные ресурсы без необходимости перекоммутации оборудования.

4. Структура вычислительной среды

С учётом выбранной модели угроз безопасности информации была разработана структура изолированной вычислительной среды (Рис.1), включающая в себя:

2 вычислительных узла (node1 и node2) MBC-100K под управлением 2-х микропроцессоров Intel Xeon E5450 (8/16 Гб оперативной памяти);

2 вычислительных узла (hps05 и hps06) MBC-10П Petastream под управлением Intel Xeon E5-2667 (48 Гб оперативной памяти) и 8-ю Intel Xeon Phi 7120D (16 Гб оперативной памяти на 1 плате), подсистемой хранения данных NetAPP FAS;

1 сервер под управлением 4-х микропроцессоров AMD Opteron 6176 (256 Гб оперативной памяти);

2 сервера удаленных рабочих столов Citrix XenServer [2] под управлением Intel Xeon X5670 (48 Гб оперативной памяти).

Все компоненты вычислительной среды, включая каждый сопроцессор Intel Xeon Phi 7120D, объединены двумя виртуальными локальными сетями (VLAN): сетью управления и сетью передачи данных.

Сеть управления используется для организации процесса вычисления и доступа к вычислительным ресурсам. Сеть передачи данных используется для циркуляции и хранения обрабатываемой информации между вычислительными узлами и СХД NetApp.

При создании выделенного сегмента суперкомпьютера требуется выделение части СХД для работы в качестве изолированного хранилища. Изоляция должна производиться как на уровне логических объектов хранения (томов или LUN-ов), так и на сетевом (как правило на 2-м уровне модели OSI, с помощью технологии VLAN).

При построении вычислительной среды изоляция части СХД была реализована с использованием технологии виртуализации систем хранения данных Storage Virtual Machine (SVM) для СХД NetApp под управлением ОС Clustered Data ONTAP версии 8.1 и выше.

На СХД суперкомпьютера была создана SVM nfs_cgf с одним vlan-интерфейсом в виртуальной локальной сети передачи данных. В изолированном пространстве SVM созданы 2 тома - для домашних каталогов пользователей и для хранения данных разрабатываемых проектов. Доступ к сетевым каталогам осуществляется по протоколу NFSv3 с ограничением доступа только для диапазона IP-адресов VLAN передачи данных.

Для доступа к СХД с вычислительных узлов суперкомпьютера MBC-10П используется решение на основе Virtual Protocol Interface (VPI) [3]. На интерфейсах Infiniband вычислительных узлов поднят протокол IPoIB, и некоторые коммутаторы Mellanox sx6036 используются как мосты (VPI-Gateway) между IPoIB и сетью Ethernet. Программное обеспечение MLX-OS позволяет организовать мост между разделом (partition) IPoIB коммуникационной сети Infiniband и виртуальной локальной сетью (VLAN).

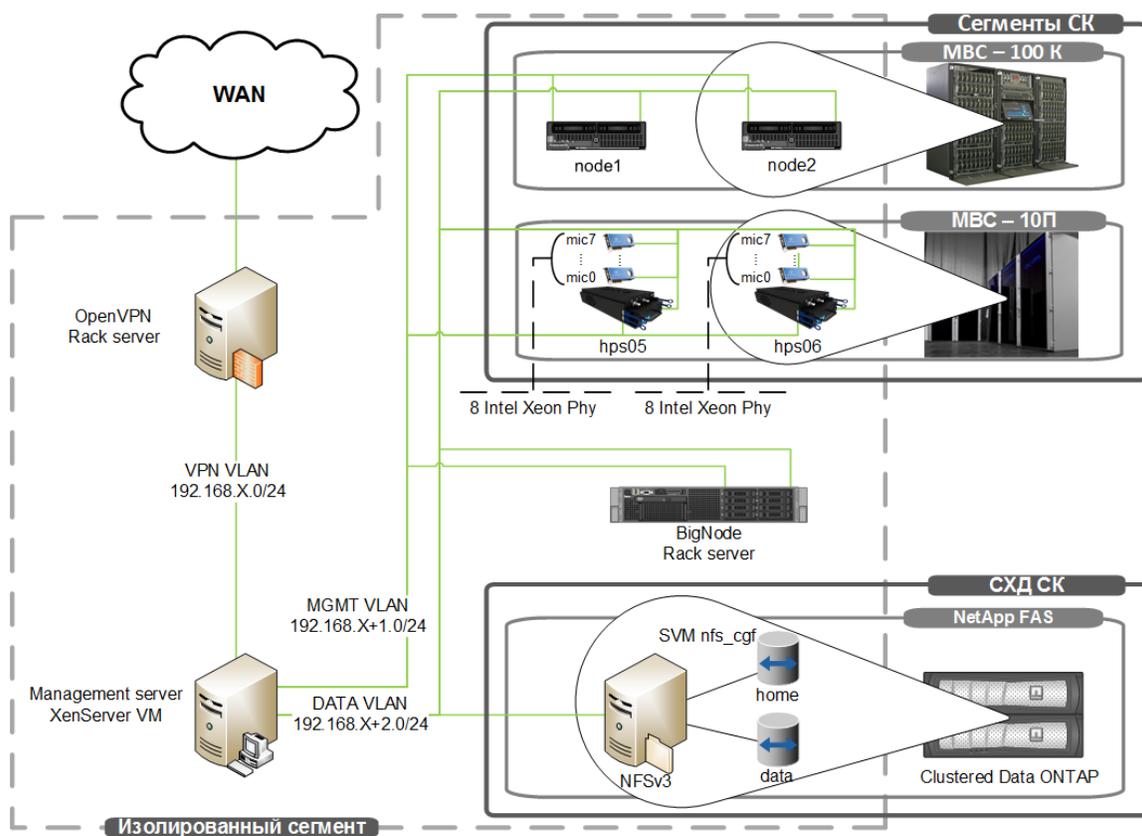


Рис.1. Структура изолированной вычислительной среды для задач фотореалистичной компьютерной графики

Виртуальная локальная сеть (VLAN) передачи данных проекта заведена на интерфейс Ethernet VPI-Gateway. Со стороны коммуникационной сети настройками задан дополнительный раздел (partition), в который вошли выделенные для проекта узлы, эти же узлы были исключены из раздела по умолчанию. Этим в решающем поле вычислительного кластера были созданы два раздела IPoIB с разными IP-сетями. На VPI-Gateway настроены проху-интерфейсы между разделами коммуникационной сети и соответствующими виртуальными сетями. В результате получился изолированный сегмент, включающий выделенные узлы MBC-10П и MBC-100К и виртуальный сервер хранения данных. Созданный сегмент подключен к серверу виртуальной частной сети, на котором для

сотрудников студии визуальных эффектов организованы учётные записи для доступа к вычислительной среде.

Таким образом, созданием выделенного на уровне виртуальных сетей сегмента, с изоляцией внутреннего и шифрованием внешнего трафика удалось решить задачу защиты данных при их обработке на вычислительном кластере ЦКП МСЦ РАН.

5. Оценка времени решения задач фотореалистичной компьютерной графики на ресурсах вычислительной среды

Для проверки работоспособности созданной среды и оценки времени решения задач фотореалистичной компьютерной графики с использованием различных вычислительных модулей, были решены две задачи:

Задача №1 – получение 3D-изображения объекта со сложной геометрией, состоящего из 2025768 полигонов. При решении данной задачи изображение целиком размещается в оперативной памяти вычислительного узла. На время построения изображения влияют только характеристики микропроцессора и объём оперативной памяти на вычислительном узле.

Задача №2 – получение изображения лица крупным планом из 29,343 полигонов и 94,813 кривых на волосах, просчет подповерхностного рассеивания (subsurface scattering). При решении данной задачи изображение строится по частям. Каждая часть изображения подгружается с сетевого файлового хранилища и размещается в оперативной памяти вычислительного узла для дальнейшей обработки. На время построения изображения значительное влияние оказывают характеристики коммуникационной среды и микропроцессора, а также объём оперативной памяти на вычислительном узле.

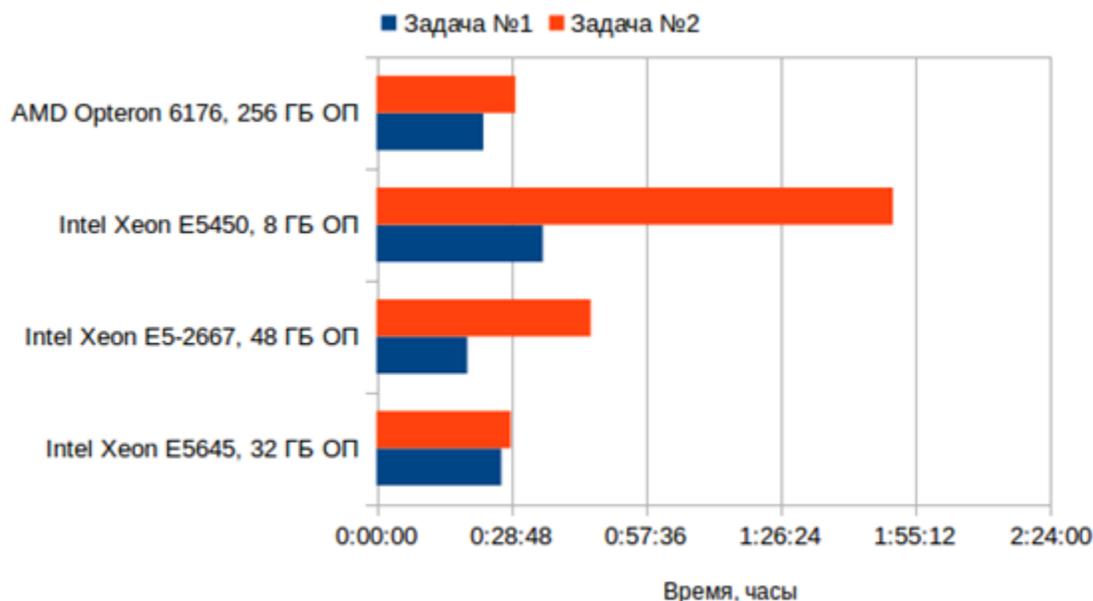


Рис. 2. Сравнение времени решения задач на вычислительных ресурсах киностудии и созданной вычислительной среды

Запуск задач осуществлялся из студии CGF через защищённое VPN-соединение с помощью планировщика рендер-задач с открытым кодом «Afanasy», который был дополнен программными модулями собственной разработки. Работа с промежуточными результатами

и оценка временных характеристик проводилась с использованием удалённых рабочих столов Citrix XenServer.

Полученные результаты (рис. 2) сравнивались со временем решения указанных задач на вычислительных узлах студии CGF (Intel Xeon E5645, 32 ГБ).

Анализ результатов тестирования показал, что наименьшее время решения задачи №1 было получено на вычислительных модулях Intel Xeon E5-2667 и AMD Opteron 6176 (МСЦ РАН). Полученные результаты объясняются тем, что вычислительные модули МСЦ РАН с микропроцессорами Intel Xeon E5-2667 и AMD Opteron 6176 обладают большей производительностью в сравнении с вычислительными узлами студии CGF с микропроцессорами Intel Xeon E5645.

Наименьшее время решения задачи №2 было получено на узлах Intel Xeon E5645 (студия CGF) и AMD Opteron 6176 (МСЦ РАН). Полученные результаты объясняются тем, что в выделенном сегменте сети МСЦ РАН для доступа к сетевому хранилищу используется сеть с пропускной способностью 1 Гбит/сек, в то время как в сети студии CGF для доступа к сетевому хранилищу используется сеть с пропускной способностью 10 Гбит/сек. При решении задачи 2 объём оперативной памяти вычислительного узла с AMD Opteron 6176 позволил сразу загрузить все данные, необходимые для построения изображения. Благодаря этому построение изображения выполнялось практически без обращения к сетевому хранилищу.

Оценка перспектив использования более производительных вычислительных модулей МВС-10П и сопроцессоров Intel Xeon Phi 7120D для решения указанных задач является предметом дальнейшего исследования.

Выводы

В ходе выполнения настоящей работы была разработана и реализована на практике технология создания защищённой вычислительной среды для решения задач фотореалистичной компьютерной графики. Предложенная технология позволяет обеспечить изоляцию данных на уровне сети, системы хранения данных и вычислительных ресурсов.

Благодаря наличию механизма удалённых рабочих столов, сегмент вычислительной среды может применяться и для решения других задач имитационного моделирования (газодинамики, гидродинамики, инженерных расчетов и др.).

Предложенная технология создания изолированной вычислительной среды позволяет создавать любое число сегментов и включать в их состав одновременно вычислительные ресурсы как суперкомпьютера МВС-100К, так и МВС-10П.

Результаты проведенных экспериментов свидетельствуют о необходимости увеличения пропускной способности сети хранения данных для выделенного изолированного сегмента. В то же время показана возможность перераспределения задач студии на ресурсы МСЦ РАН без существенных потерь во времени обработки, а для некоторых задач – с сокращением времени выполнения.

Литература

1. Баранов А.В., Ляховец Д.С. Сравнение качества планирования заданий в системах пакетной обработки SLURM и СУППЗ. Научный сервис в сети Интернет: все грани параллелизма: Труды Международной суперкомпьютерной конференции (23-28 сентября 2013 г., г. Новороссийск). М.: Изд-во МГУ, 2013. С. 410-414. URL: <http://agora.guru.ru/abrau2013/pdf/410.pdf> (дата обращения: 12.04.2017).

2. Citrix XenDesktop. Технические характеристики. URL: http://www.ocs.ru/OCS/media/Products/Citrix/DataSheet_XenDesktop_rus.pdf (дата обращения: 12.04.2017).
3. Mellanox Virtual Protocol Interconnect Creates the Ideal Gateway Between InfiniBand and Ethernet. URL: http://www.mellanox.com/related-docs/case_studies/CS_VPI_GW.pdf (дата обращения: 12.04.2017).
4. CGRU Afanasy. Open source render farm manager. URL: <http://cgru.info/afanasy/afanasy> (дата обращения: 12.04.2017).
5. Valentin Del Piccolo, Ahmed Amamou, William Dauchy, and Kamel Haddadou. Multi-tenant isolation in a trill based multicampus network. In Cloud Networking (CloudNet), 2015 IEEE 4th International Conference on, pages 5157, Oct 2015.
6. Wei Jing, Nan Guan, and Wang Yi. Performance isolation for real-time systems with xen hypervisor on multi-cores. In Embedded and Real-Time Computing Systems and Applications (RTCSA), 2014 IEEE 20th International Conference on, pages 17, Aug 2014.
7. Albert Greenberg, James R. Hamilton, Navendu Jain, Srikanth Kandula, Changhoon Kim, Parantap Lahiri, David A. Maltz, Parveen Patel, and Sudipta Sengupta. VI2: A scalable and exible data center network. In Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication, SIGCOMM '09, pages 5162, New York, NY, USA, 2009. ACM. <http://doi.acm.org/10.1145/1592568.1592576>.
8. Fang Hao, T. V. Lakshman, Sarit Mukherjee, and Haoyu Song. Secure cloud computing with a virtualized network infrastructure. In Proceedings of the 2Nd USENIX Conference on Hot Topics in Cloud Computing, HotCloud'10, pages 1616, Berkeley, CA, USA, 2010. USENIX Association. <http://dl.acm.org/citation.cfm?id=186310>.

Solving the problems of photorealistic computer graphics using the secure infrastructure of the supercomputer center

B.M. Shabanov¹, A.P. Ovsyannikov¹, A.V. Baranov¹, E.A. Kiselev¹, S.A. Leshev¹, B.V. Dolgov¹, D.G. Gumenniy², D.L. Shurov²

¹Joint Supercomputer Center of the Russian Academy of Sciences - Branch of Federal State Institution «Scientific Research Institute for System Analysis of the Russian Academy of Sciences», Moscow, Russia ²CGF - Post Production and Visual Effects Studio, Moscow, Russia

In the article, we consider the organization of the protected infrastructure of the supercomputer center. The authors propose a method for protecting confidential user data by organizing a dedicated segment of a supercomputer based on virtual Ethernet networks and InfiniBand network partitions. The paper proposes a technology for building a scalable heterogeneous computing environment on the resources of a supercomputer center in the context of solving problems of creating photorealistic computer graphics. The results of experimental calculations are presents.

Keywords: Photorealistic computer graphics, supercomputers, computing environments, secure infrastructure, parallel processing

References

1. Baranov A.V., Lyakhovets D.S. Sravnenie kachestva planirovaniya zadaniy v sistemah paketnoj obrabotki SLURM i SUPPZ [Comparison of the quality of job scheduling in workload management systems SLURM and SUPPZ]. Nauchnyj servis v seti Internet: vse grani parallelizma: Trudy mezhdunarodnoy superkomp'yuternoy konferentsii (Novorossiysk, 23-28 sentyabrya 2013 g.) [Scientific Services & Internet: All Facets of Parallelism: Proceedings of the International Supercomputing Conference (Novorossiysk, Russia, September, 23–28, 2013)]. Moscow, Publishing of Lomonosov Moscow State University, 2013. P. 410–414. URL: <http://agora.guru.ru/abrau2013/pdf/410.pdf> (accessed: 12.04.2017).
2. Citrix XenDesktop. DataSheet. URL: <http://www.scc.com/wp-content/uploads/2015/11/Xendesktop-Data-Sheet.pdf> (accessed: 12.04.2017).
3. Mellanox Virtual Protocol Interconnect Creates the Ideal Gateway Between InfiniBand and Ethernet. URL: http://www.mellanox.com/related-docs/case_studies/CS_VPI_GW.pdf (accessed: 12.04.2017).
4. CGRU Afanasy. Open source render farm manager. URL: <http://cgru.info/afanasy/afanasy> (accessed: 12.04.2017).
5. Valentin Del Piccolo, Ahmed Amamou, William Dauchy, and Kamel Haddadou. Multi-tenant isolation in a trill based multicampus network. In Cloud Networking (CloudNet), 2015 IEEE 4th International Conference on, pages 5157, Oct 2015.
6. Wei Jing, Nan Guan, and Wang Yi. Performance isolation for real-time systems with xen hypervisor on multi-cores. In Embedded and Real-Time Computing Systems and Applications (RTCSA), 2014 IEEE 20th International Conference on, pages 17, Aug 2014.
7. Albert Greenberg, James R. Hamilton, Navendu Jain, Srikanth Kandula, Changhoon Kim, Parantap Lahiri, David A. Maltz, Parveen Patel, and Sudipta Sengupta. VI2: A scalable and exible data center network. In Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication, SIGCOMM '09, pages 5162, New York, NY, USA, 2009. ACM. <http://doi.acm.org/10.1145/1592568.1592576>.
8. Fang Hao, T. V. Lakshman, Sarit Mukherjee, and Haoyu Song. Secure cloud computing with a virtualized network infrastructure. In Proceedings of the 2Nd USENIX Conference on Hot Topics in Cloud Computing, HotCloud'10, pages 1616, Berkeley, CA, USA, 2010. USENIX Association. <http://dl.acm.org/citation.cfm?id=186310>.