

## СЕТЬ НАУКИ И ОБРАЗОВАНИЯ КАК ИНФРАСТРУКТУРА ДЛЯ GRID - ПРИЛОЖЕНИЙ

Секриеру Г.В., Богатенков П.П., Ильюха Н.П.  
Институт Математики и Информатики АНМ, Кишинев, Молдова  
[secrieru@renam.md](mailto:secrieru@renam.md), [bogatencov@renam.md](mailto:bogatencov@renam.md), [nick@renam.md](mailto:nick@renam.md)

**Abstract.** *Modern GRID-system have been widely developed and the creation of applications in GRID - infrastructure is a critical factor for determining progress in many areas, which require the use of intensive data processing and storing large amounts of data. In the proposed paper examines the main aspects of the GRID-Infrastructure development in Moldova that is focusing on running various complex applications in different research areas.*

**Ключевые слова:** *Grid-технологии, Grid-систем, распределенная обработка информации, высокопроизводительные вычисления, MD-Grid NGI.*

### 1. Введение

Характерной особенностью последних лет является развитие технологии организации распределенной обработки информации и высокопроизводительных вычислений. Широкое развитие получили Grid-технологии, представляющие возможность создания согласованной, стандартизированной и открытой технологической среды, обеспечивающей гибкое и скоординированное совместное использование вычислительных и информационных ресурсов для решения сложных и ресурсоемких задач в важнейших областях современной науки и техники. В e-science, наряду с вычислительными гридами, требуются информационные Grid-структуры, обеспечивающие доступ к неоднородным, распределенным хранилищам данных большого объема.

Grid-инфраструктуры служат для обеспечения доступа к экспериментальным данным, вычислительным мощностям, и широкому набору других ресурсов, которые доступны при помощи глобальной сети Интернет. Если типы Grid распределить на классы, то можно различать следующие Grid - классы:

- Computational Grid – grid, ориентированный на распределенные вычисления, включая параллельные вычисления;
- Data grid – grid, ориентированный на обработку больших потоков данных;
- Informational Grid – grid, ориентированный на интеграцию распределенных хранилищ данных и сервисов;
- Hibrid grid – grid, сочетающий в себе архитектуры всех предыдущих типов.

Развитие современных Grid-систем обусловлено широкими перспективами их применения в науке, образовании и других областях человеческой деятельности. Владение навыками работы и создания приложений в Grid-инфраструктуре является важнейшим фактором, определяющим прогресс во многих областях, в которых возникают потребности в высокопроизводительных вычислениях и обработке больших объемов данных.

Главным образом Grid-технологии предназначены для решения сложных научных, производственных и инженерных задач, которые невозможно решить в разумные сроки на отдельных вычислительных установках. Grid-инфраструктура претендует на роль универсальной вычислительной инфраструктуры для обработки данных, в которой функционирует множество служб (Grid Services), которые позволяют решать не только конкретные приклад-

ные задачи, но и предлагают различные сервисы: поиск необходимых ресурсов, сбор информации о состоянии ресурсов, хранение и доставку данных.

В докладе представлены некоторые результаты реализации концепции построения Grid-инфраструктуры в Молдове. Описана иерархическая структурно-организационная модель инфраструктуры и приведены некоторые направления внедрения Grid-технологий.

## II. Основные направления создания Grid-систем

По своей топологии Grid-система представляется как географически распределенная инфраструктура, объединяющая множество ресурсов разных типов (процессоры, долговременная и оперативная память, базы данных и т.п.), доступ к которым пользователь может получить из разных точек, независимо от места их расположения. Концепция Grid предполагает коллективный разделяемый режим доступа к ресурсам и к связанным с ними услугам в рамках глобально распределенных виртуальных организаций, состоящих из предприятий, групп пользователей и отдельных специалистов, совместно использующих общие ресурсы.

В каждой виртуальной организации имеется своя собственная политика поведения ее участников, которые должны соблюдать установленные правила. Виртуальная организация может образовываться динамически и иметь ограниченное время существования.

Потенциал использования Grid-технологии уже сейчас оценивается высоко: он имеет стратегический характер, и по своему функциональному назначению в перспективе должен стать вычислительным инструментарием для развития высоких технологий в науке, образовании и других сферах человеческой деятельности [1-3]. Такие оценки можно объяснить способностью Grid на основе безопасного и надежного удаленного доступа к ресурсам глобально распределенной инфраструктуры решить следующие принципиальные проблемы:

- создание Grid-инфраструктуры высокой пропускной способности из серийно выпускаемого оборудования при одновременном повышении эффективности (до 100%) имеющегося парка вычислительной техники путем предоставления в Grid временно не используемых ресурсов;
- создание распределенных вычислительных систем, способных обеспечивать решение сложных научных, инженерных и производственных задач;
- создание широкомасштабных систем мониторинга, управления, комплексного анализа и обслуживания с глобально распределенными источниками данных, повышающих жизнедеятельность научных учреждений, организаций и других структур общества.

Национальная Grid инициатива в Молдове (MD-Grid NGI) была сформирована в 2007-2010 годах в рамках международных проектов [2,3]. В качестве организационной структуры была использована Ассоциация RENAM (Research and Educational Networking Association of Moldova), которая построила и администрирует общую сетевую инфраструктуру с доступом в Интернет для научно-образовательного сообщества Молдовы с одноименным названием - RENAM. Эта сеть объединяет оптоволоконными каналами институты Академии наук Молдовы (АНМ), ведущие университеты и другие научно – образовательные учреждения. Министерство Информационного Развития Молдовы и руководство АНМ, учреждения сферы науки, образования и медицины поддержали инициативу создания Grid-инфраструктуры и Национальной Grid инициативы в Молдове.

Участие Молдовы в ряде международных проектов способствовало разработке принципов функционирования MD-Grid NGI и ее базовых элементов [4-8].

Одним из важнейших условий построения Grid-инфраструктуры является наличие скоростного и надежного канала доступа в Internet. Для этой цели используется сетевая инфраструктура RENAM, которая объединяет ресурсы АНМ, ведущих университетах и других учреждений сферы науки и образования в единую сеть с доступом в Internet. По своей тополо-

гии сеть RENAM представляет собой трех уровневую архитектуру. Первый уровень это локальные сети кампусов, организаций и учреждений. Второй уровень это сетевые узлы с оптоволоконными каналами (пропускная способность 1Gbps) для подключения локальных сетей исследовательских институтов АНМ и университетов к сети RENAM. Третий уровень это центральный узел сети и внешний магистральный оптоволоконный канал Кишинев - Яссы (Румыния) (пропускная способность оптического оборудования 10 Gbps) с последующим выходом на Транс - Европейскую академическую сеть GEANT. Создание и ввод в эксплуатацию в 2010 году оптоволоконного канала с пропускной способностью 10 Gbps для доступа к сети GEANT было осуществлено в рамках двух международных проектов:

- RI211338 «SEE-GRID eInfrastructure for regional eScience (SEE-GRID-SCI)» (финансированный Европейской Комиссией);
- NIG982702 «New RENAM-RoEduNet gateway based on CWDM technologies implementation» (финансированный НАТО).

### III. Актуальность применения Grid-технологий

Главной задачей внедрения и развития распределенных вычислений в Grid-средах является повышение эффективности фундаментальных и прикладных исследований проводимых научно-исследовательскими институтами и университетами и требующих значительных вычислительных ресурсов. Необходимыми условиями для реализации этой задачи являются:

- надежное функционирование и развитие скоростной, защищенной информационно-вычислительной сетевой инфраструктуры (в нашем случае – RENAM);
- развитие и надежное функционирование распределенной высокопроизводительной вычислительной инфраструктуры;
- информационная, алгоритмическая и программная поддержка научно-технических учреждений, создающих прикладные системы, работающие в распределенной вычислительной инфраструктуре;
- надежное функционирование и развитие Grid-сегмента Молдовы как элемента глобальной Grid-инфраструктуры;
- внедрение технологий распределенной обработки информации и доступа к распределенной информации;
- разработка и адаптация существующих методов "гридификации" прикладного программного обеспечения и обеспечение взаимодействия различных Grid-систем.

Развитие региональной научно-исследовательской вычислительной инфраструктуры координируется такой общеевропейской инициативой, как European Grid Initiative (EGI). EGI является координатором проекта EGI-InSPIRE (European Grid Initiative: Integrated Sustainable Pan-European Infrastructure for Researchers in Europe).

Основная цель EGI-InSPIRE — скоординировать действия участников проекта на последовательном процессе перехода от инфраструктур, поддерживаемых и устойчивых в рамках проектов к выходящей за рамки проектов устойчивой общеевропейской инфраструктуре.

Проект охватывает такие Распределенные Вычислительные Инфраструктуры (Distributed Computing Infrastructures (DCIs), как облачные структуры (clouds), сети суперкомпьютеров (supercomputing networks) и Гриды, объединяющие ресурсы настольных компьютеров (desktop grids). Объединенная инфраструктура поддерживает различные типы высокопроизводительных вычислений:

- требующих больших ресурсов для быстроты вычислений (high-performance computing — HPC);
- требующих больших ресурсов для долговременных вычислений (high-throughput

computing — HTC);

- требующих больших ресурсов для многоэтапных вычислений (использующих первые два типа ресурсов и файловые структуры для обмена управляющей информацией - many-task computing MTC).

В рамках проекта EGI-InSPIRE проводится разнообразная работа и большое внимание уделено работе с пользователями:

- собираются и анализируются требования и предложения пользователей;
- поддерживаются существующие виртуальные сообщества и оказывается помощь в организации новых виртуальных сообществ;
- особое внимание уделяется так называемым «тяжелым» пользователям — пользователям, чьи приложения и задачи требуют вычислительных ресурсов выше среднего (физика высоких энергий, вычислительная химия и науки о жизни). Этим пользователям помогают оптимально адаптировать их задачи в той инфраструктурной модели ресурсов, которую поддерживают их виртуальные сообщества.

Создание национального сегмента Grid-инфраструктуры в Молдове преследует цель расширения внедрения информационных технологий для науки, образования, медицины и других сфер. Приоритетными направлениями для этого сегмента являются:

- Интеграция необходимых элементов единого национального Grid-сегмента: коммуникационных, компьютерных и программных ресурсов;
- Внедрение современных технологий в научных исследованиях и образовании, интегрирование научных учреждений в европейское и мировое научное пространство и привлечение молдавских ученых к участию в международных проектах и виртуальных научных сообществах;
- Обеспечение компьютерной обработки больших объемов результатов метеорологических, геофизических, экологических и других измерений;
- Создание условий для внедрения новых современных методов медицинского обслуживания с использованием распределенных баз диагностических данных и распределенной обработки медицинских данных;
- Разработка и реализация системы подготовки и повышения квалификации специалистов для работы в Grid-системах и разработчиков Grid-приложений.

В настоящее время Grid сегмент Молдовы объединяет вычислительные ресурсы различных учреждений. В Таблице 1 представлены основные параметры Grid-сайтов MD-GRID NGI.

Таблица 1

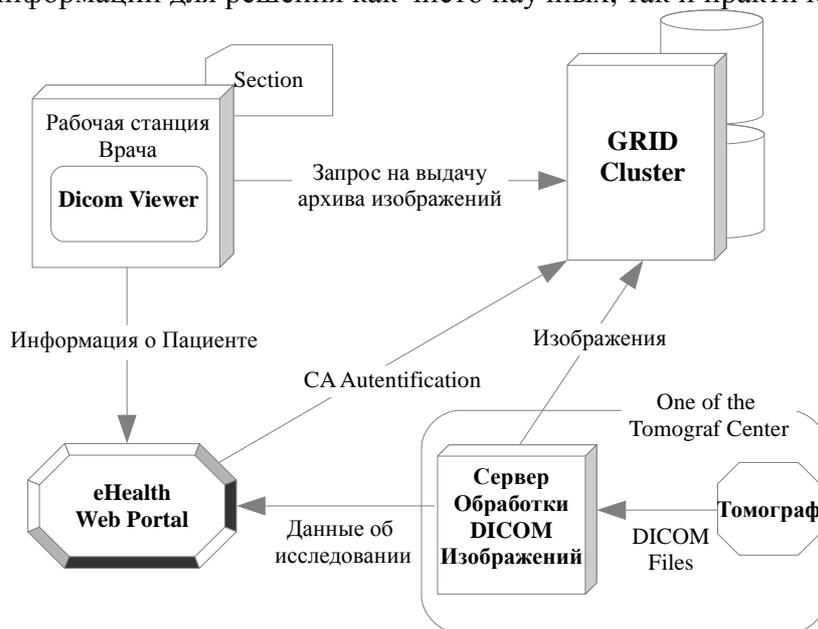
MD-GRID NGI site	Available CPUs	Available storage	Network (External)
<b>Certified sites</b>			
<b>MD-03-SUMP</b>	5 x CPU AMD Athlon 64 X2 6000+ (3.0GHz)	650 GB on Storage Element	100 Mbps Ethernet
<b>MD-04-RENAM</b>	6 Quad Core Xeon 5130 CPUs	2 TB on Storage Element	100 Mbit Ethernet
<b>MD-02-IMI</b>	12 Quad Core Xeon 5130 CPUs	3,5 TB on Storage Element	100 Mbit Ethernet
<b>Planned to be integrated into MD-GRID NGI</b>			
<b>MD-05-SUM</b>	4x2xAMD 275 Dual-Core 2.2GHz and 3x2xAMD 280 Dual-Core 2.4GHz CPUs	2x500GB 7.2k SATA and 4x80 GB 7.2k SATA	100 Mbit Ethernet

Ассоциация RENAM координирует функционирование и развитие MD-GRID NGI, активными участниками которой в настоящее время являются:

- ИГС АНМ — Институт Геологии и Сейсмологии Академии Наук Молдовы;
- ИМИ АНМ — Институт Математики и Информатики Академии Наук Молдовы;
- ИПФ АНМ — Институт Прикладной Физики Академии Наук Молдовы;
- ФРТ ТУМ — Факультет Радиоэлектроники и Телекоммуникаций Технического Университета Молдовы;
- ГГСМ — Государственная Гидрометеорологическая Служба Молдовы;
- НЦМСП – Национальный научно-практический центр медицины скорой помощи.

Приложения, активно использующие Grid-технологии сейчас развиваются по трем направлениям - сейсмология, экология-метеорология и медицина [7]. Последнее направление связано с созданием, используя технологии Grid, системы для обмена и хранения информации в формате DICOM (стандарт изображений, получаемых с медицинского оборудования). Создание такой системы призвано обеспечить автоматизацию хранения, обработки и доступа к результатам медицинских исследований, накапливаемым в формате DICOM [5]. Общая структура такой системы показана на *Рис.1*. При решении задач информационного характера, когда потребитель должен получать интересующую его информацию в момент порождения или когда в ней возникает потребность, Grid-технологии применяются для интеграции всех накапливаемых данных.

Основной чертой развития информационных технологий в Молдове является стремление эффективно использовать распределенные разнородные вычислительные ресурсы и системы хранения информации для решения как чисто научных, так и практических задач.



**Рис. 1**

Исходя из этого, интерес представляет поддержка и расширение научно-исследовательской и прикладной деятельности в сфере Grid-технологий и High-Performance Computing в следующих основных направлениях:

1. Развитие тестовой Grid-структуры:
  - обмен опытом по установке и эксплуатации систем виртуализации для развертывания Grid-сайтов;
  - обмен опытом по настройке и эксплуатации middleware, систем мониторинга и учета

ресурсов;

2. Апробация и адаптация приложений в Grid (гридификация приложений):

- обмен опытом по созданию приложений в областях распределенных и параллельных вычислений, систем визуализации и распределенных баз данных;
- разработка и создание приложений в области параллельных компьютерных технологий и высокопроизводительных вычислений;
- консультации и помощь пользователям в создании Grid-ориентированных приложений;
- Разработка, адаптация и внедрение удобного для пользователей сервиса доступа к ресурсам Grid.

3. Привлечение новых пользователей ресурсов Grid-инфраструктуры:

- подготовка и издание учебно-методических материалов для популяризации Grid и HPC технологий в научно-образовательных и других структурах общества;
- обучение пользователей работе в Grid-инфраструктурах;
- включение пользователей в существующие и создаваемые Виртуальные Организации учебно-тестовой и реальной Grid-инфраструктур;

4. Осуществление совместных международных научно-технических программ и проектов в сфере Grid-технологий и High-Performance Computing.

В процессе реализации таких инфраструктур основное внимание сосредоточено на организации среды вычислений высокой производительности или на создание среды для обеспечения доступа к большому числу распределенных информационных ресурсов и их интероперабельности при решении конкретных задач. При этом, важное значение приобретают вопросы семантики предметных областей задач и ее связи с семантикой информационных ресурсов грида. Это связано с тем, что в последнее время наблюдается существенный рост объема получаемых экспериментальных данных в различных областях науки. Возникает определенный разрыв между исследователями и источниками данных и сервисов, который приводит к необходимости создания новых информационных систем.

Поэтому в условиях бурного развития информационных и коммуникационных технологий владение навыками работы и умение создания приложений в Grid-средах является важной и актуальной задачей. Заключенный в начале 2010 года договор о сотрудничестве между Лабораторией Информационных Технологий Объединенного Института Ядерных Исследований (ЛИТ ОИЯИ) в Дубне и MD-GRID NGI призван вносить существенный вклад в решение этой задачи. Успешная реализация процесса обучения пользователей обусловлена наличием в ЛИТ ОИЯИ современного учебного полигона, который доступен и для пользователей из Молдовы, и который позволяет продемонстрировать работу в различных Grid-средах.

#### IV. Заключение

В настоящей работе рассмотрены современные Grid-инфраструктуры и их разнообразие по своему функциональному назначению. Различаются Computational Grid, ориентированный на распределенные вычисления, Data grid, ориентированный на обработку больших потоков данных, а также другие классы Grid. Показана актуальность, основные компоненты Grid-систем и важность применения Grid-технологий при решении сложных научно-технических и производственных задач.

Поэтому следует отметить, что владение навыками работы и создания приложений в Grid-инфраструктурах является важнейшим фактором, определяющим прогресс во многих областях, в которых возникает необходимость применения высокопроизводительных вычислений и массовая обработка значительных объемов информации.

## V. Литература

1. Добрецов В.Ю., Ильин В.А., Кореньков В.В., Крюков А.П., Рябов Ю.Ф. Проектирование и разработка Грид -инфраструктуры для национальной нанотехнологической сетию. Proceedings of the 4th International Conference, (Dubna, June 28 – July44, 2010).-Dubna: JINR, 2010/- 452p. (pp. 357-363).
2. South-East European GRid eInfrastructure Development. <http://www.see-grid.eu/>
3. SEE-GRID eInfrastructure for regional eScience. <http://www.see-grid-sci.eu/>
4. Altuhov A., Bogatencov P., Secrieru G., Sidorencu V., Pocotilenco V. Participation of scientific - educational community of Moldova in the European project of Grid development in the Easernt Europe countries. Proceedings of All-Russia scientific conference "Scientific service in Internet: Technologies of parallel programming". Novorossisk, September 18-23, 2006. Moscow University, M. 2006, pp. 127-128. (in Russian).
5. Golubev, A.; Bogatencov, P.; Secrieru, G.; Iliuha, N. Dicom Network – Solution for Medical Imagistic Investigations Exchange. În: Proceedings IIS "International Workshop on Intelligent Information Systems", September 13-14, 2011, Chişinău, 179-182. ISBN 978-9975-4237-0-0.
6. Sidorencu V., Bogatencov P., Altuhov A. MD-GRID JRU Consortium and its Role in SEE-GRID-SCI Project. Proceedings of the Third International Conference, Dubna, June 30 – July 4, 2008.
7. Алтухов А.А., Богатенков П.П., Васюкова Е.В., Секриеру Г.В. MD-GRID NGA: Современное состояние и перспективы развития GRID-технологий в Молдове. Proceedings of the 4th International Conference, (Dubna, June 28 – July44, 2010).-Dubna: JINR, 2010/- 452p. (pp. 279-283).
8. Астафьев А. С., Астафьев А. П., Лазарев И.В., Сухорослов О.В., Татасов А.С. Научная сервис -ориентированная среда на основе технологий Web и распределенных вычислений. // Научный сервис в сети Интернет: масштабируемость, параллельность, эффективность: Труды Всероссийской суперкомпьютерной конференции (21-26 сентября 2009 г., Новороссийск). М.: Изд-во МГУ, 2009, 524 с. (с. 463-467).