

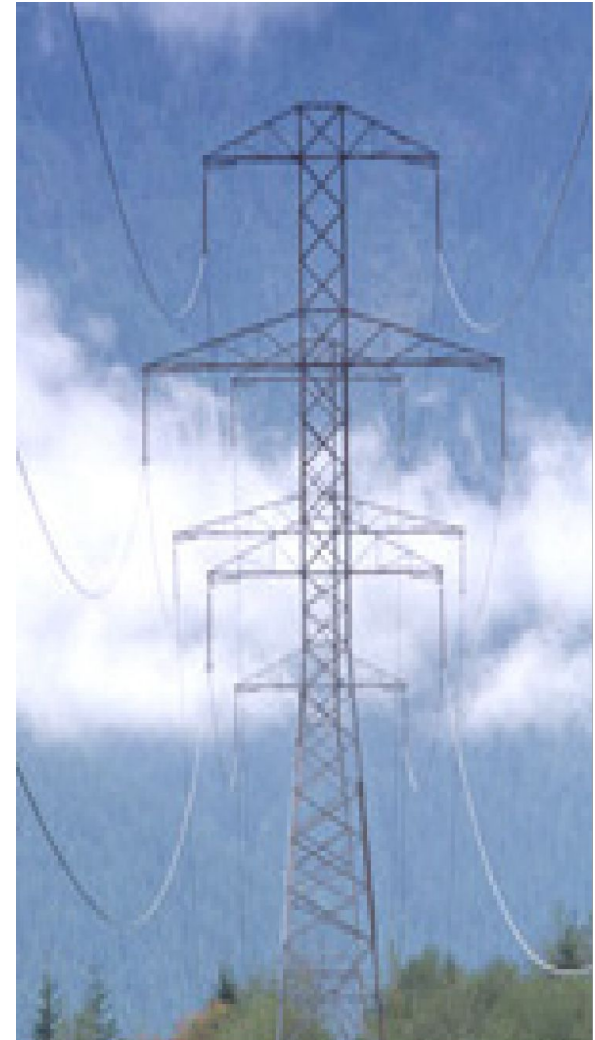
# Введение в ГРИД

*Новодворский Е.Г., Олешко С.Б.  
Петербургский институт ядерной физики  
г.Гатчина*

Термин **ГРИД** был выбран по аналогии с энергетическими сетями (**electric power grid**)

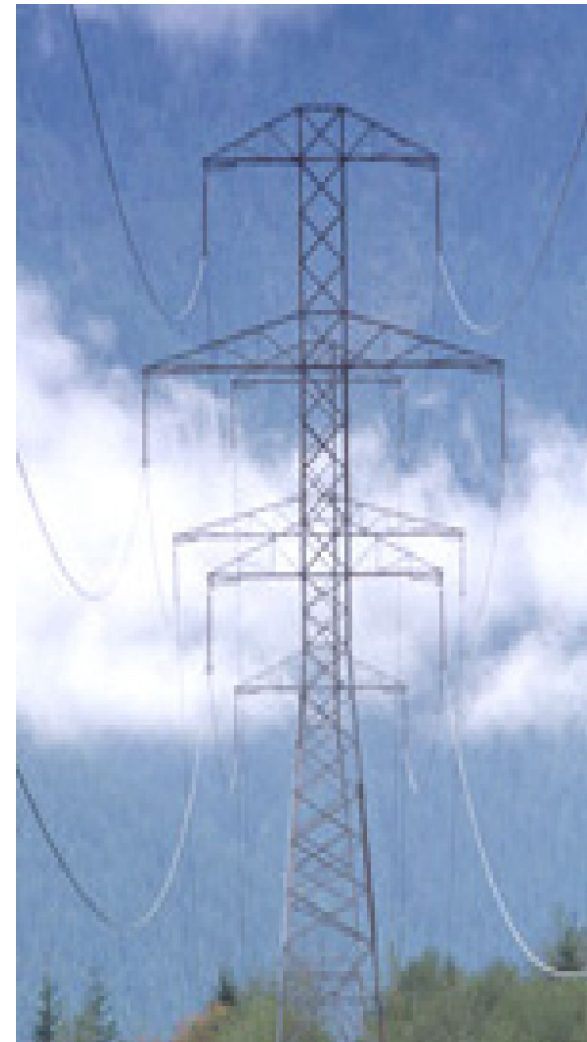
подобно тому, как при использовании энергосистем мы не интересуемся – какой конкретный электрогенератор выработал ток, который мы потребляем,

при использовании компьютерного грида мы можем не заботиться о том - какой конкретно компьютер (или устройство хранения/передачи данных) в грид-системе выполнил нашу задачу



## Короче

- доступ “прозрачен” для конечного пользователя: достаточно подключиться и получить необходимые ресурсы, не заботясь о том, откуда они берутся
- доступность “отовсюду”

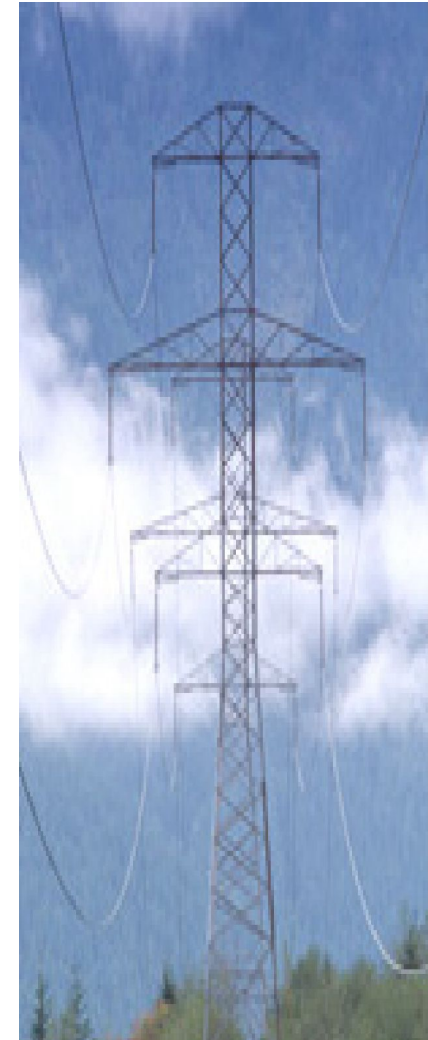


## Продолжим аналогию

Современный статус компьютерных инфраструктур можно сравнить с состоянием электрических систем в 1910 году. Тогда необходимость для каждого пользователя иметь свой генератор тормозила развитие отрасли. Революционным шагом было возникновение электросетей, создание технологий передачи и распределения электроэнергии, создание **стандартизированной службы** универсального и гарантированного доступа к электроэнергии. В результате не только резко повысилась эффективность использования ресурсов, но и стали возможны принципиально новые направления развития.

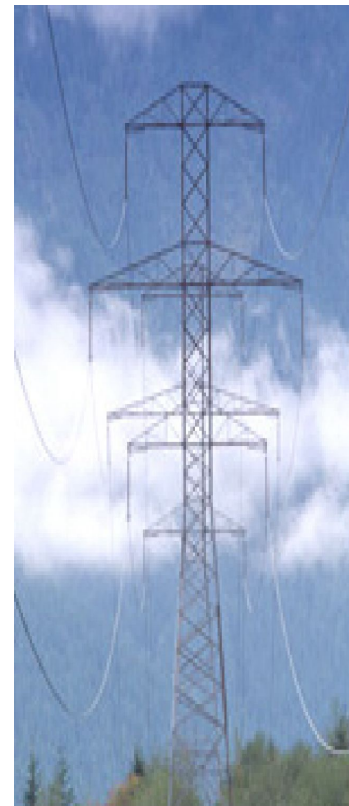


**Грид призван обеспечить  
возможность делать  
компьютерные вычисления  
«по требованию»  
просто подключившись  
к «решетке»  
вычислительных ресурсов.**



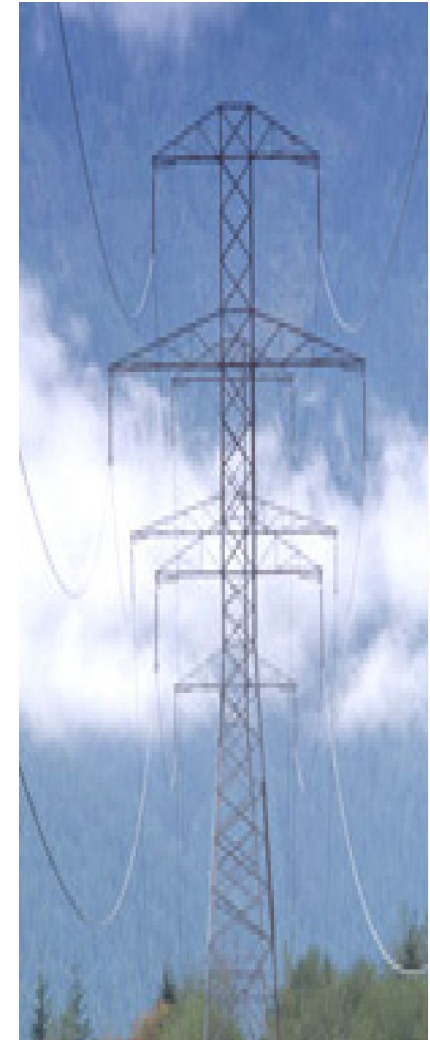
Также как и электрические сети,  
грид это соединение

- *технологии,*
- *инфраструктуры*
- *стандартов.*



## Технология

это специальное программное обеспечение, которое позволяет организациям или частным лицам **предоставлять ресурсы** (компьютеры, хранилища данных, сети и другие) **в общее пользование**, а потребителям – использовать их, когда необходимо.



## Инфраструктура

СОСТОИТ ИЗ

- **аппаратных средств и**
- **служб**

(на основе людских и программных ресурсов), которые должны быть организованы, и постоянно поддерживаться для того, чтобы ресурсы могли совместно использоваться.



## Стандарты

должны определять

- **формат и**
- **протоколы обмена сообщениями,**

как между службами, так и между службами и пользователями, а также правила работы грида.

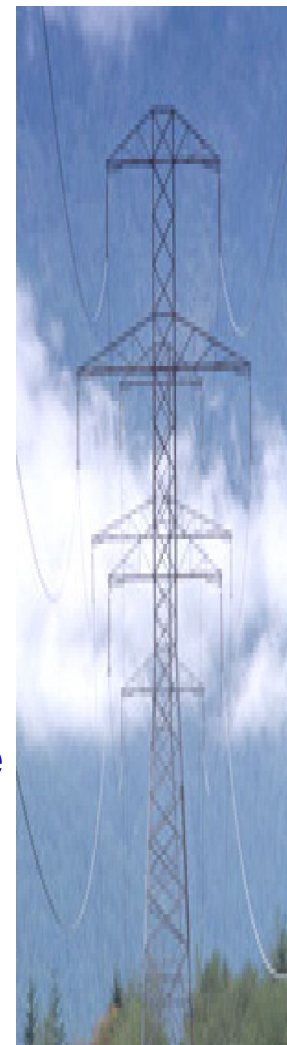
## Опять аналогия

В энергосетях определяющую роль играют:

**технологии** передачи и (пере)распределения электроэнергии на большие расстояния;

**инфраструктура** - аппаратные средства (повышающие/понижающие подстанции, линии электропередач, другое электрооборудование, позволяющее электростанциям предоставлять энергию потребителям) и соответствующие службы (ремонта, контроля и т.д.);

**стандарты** – параметры электрического тока (напряжение в сети, частота), типы и размеры вилок/розеток, позволяющие без проблем подключать любой электроприбор к единой сети.



**1985–1995** Программа Национального Научного Фонда США «National Science Foundation (NSF) Supercomputer Centers».

Результат – создание развитой коммуникационной инфраструктуры и нескольких суперкомпьютерных центров для поддержки академических работ и исследований.

**1985–1995** Программа Национального Научного Фонда США «National Science Foundation (NSF) Supercomputer Centers».

**Октябрь 1997** NSF инициировал новую программу развития информационных технологий Partnerships for Advanced Computational Infrastructure (PACI).

Возникла концепция Грид-инфраструктуры, как некая аналогия электроэнергетической инфраструктуры. Предполагалось, что Grid будет использоваться в основном для того, чтобы задействовать **простаивающие** вычислительные компьютерные ресурсы.

**1985–1995** Программа Национального Научного Фонда США «National Science Foundation (NSF) Supercomputer Centers».

**Октябрь 1997** NSF инициировал новую программу развития информационных технологий – Partnerships for Advanced Computational Infrastructure (PACI).

**1998 Создан** (и успешно развивается)  
инструментальный пакет **Globus Toolkit**.

Является технологической базой создания Грид-инфраструктуры (*стандарт de facto*)

- 1985–1995** Программа Национального Научного Фонда США «National Science Foundation (NSF) Supercomputer Centers».
- Октябрь 1997** NSF инициировал новую программу развития информационных технологий – Partnerships for Advanced Computational Infrastructure (PACI).
- 1998** Создан (и успешно развивается) инструментальный пакет Globus Toolkit.
- 1999** Сформировалось (и активно действует) международное научное Грид-сообщество Global Grid Forum (GGF)

**1985–1995** Программа Национального Научного Фонда США «National Science Foundation (NSF) Supercomputer Centers».

**Октябрь 1997** NSF инициировал новую программу развития информационных технологий – Partnerships for Advanced Computational Infrastructure (PACI).

**1998** Создан (и успешно развивается) инструментальный пакет Globus Toolkit.

**1999** Сформировалось (и активно действует) международное научное Грид-сообщество – Global Grid Forum (GGF)

**2002** GGF и IBM была представлена новая системная разработка – Open Grid Service Architecture (OGSA)

Введены в грид понятия и стандарты веб-сервисов

1985–1995 Программа Национального Научного Фонда США «National Science Foundation (NSF) Supercomputer Centers».

Октябрь 1997 NSF инициировал новую программу развития информационных технологий – Partnerships for Advanced Computational Infrastructure (PACI).

1998 Создан (и успешно развивается) инструментальный пакет Globus Toolkit.

1999 Сформировалось (и активно действует) международное научное Грид-сообщество – Global Grid Forum (GGF)

2002 GGF и IBM была представлена новая системная разработка – Open Grid Service Architecture (OGSA).

## 2003 Создано объединение Enterprise Grid Alliance (EGA)

адаптация академических разработок по Грид-технологиям к потребностям производителей (Fujitsu, Siemens, HP, Intel, NEC, Oracle, Sun, EMC)



- 1985–1995** Программа Национального Научного Фонда США «National Science Foundation (NSF) Supercomputer Centers».
- Октябрь 1997** NSF инициировал новую программу развития информационных технологий – Partnerships for Advanced Computational Infrastructure (PACI).
- 1998** Создан (и успешно развивается) инструментальный пакет Globus Toolkit.
- 1999** Сформировалось (и активно действует) международное научное Грид-сообщество – Global Grid Forum (GGF)
- 2002** GGF и IBM была представлена новая системная разработка – Open Grid Service Architecture (OGSA).
- 2003** Создано объединение Enterprise Grid Alliance (EGA)
- 2006** GGF и EGA объявили о слиянии и образовании Open Grid Forum (OGF)



## Отдельный компьютер (PC)

Прикладное ПО

Операционная система PC

Ресурсы PC

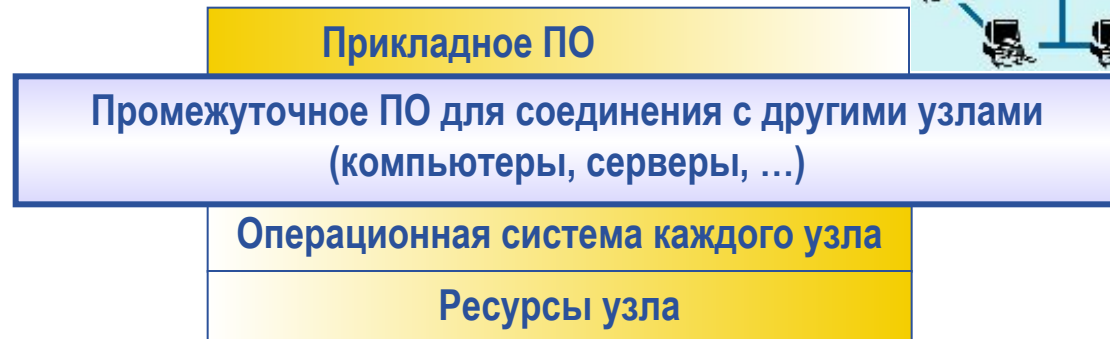
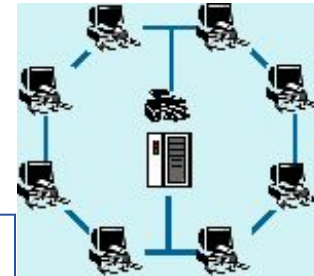
Диски, процессор, память ...

## Отдельный компьютер

- Прикладное ПО
- Операционная система
- Ресурсы PC



## Локальная сеть



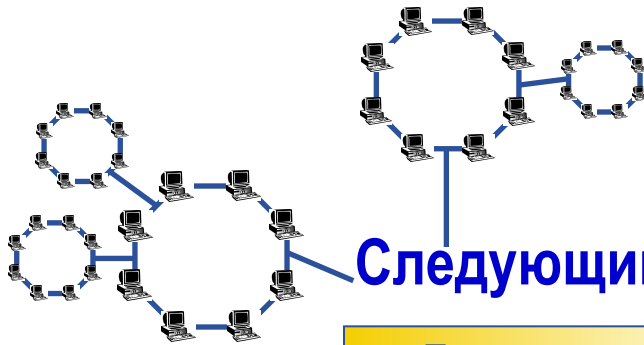
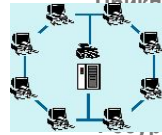
## Отдельный компьютер

Прикладное ПО  
Операционная система  
Ресурсы PC

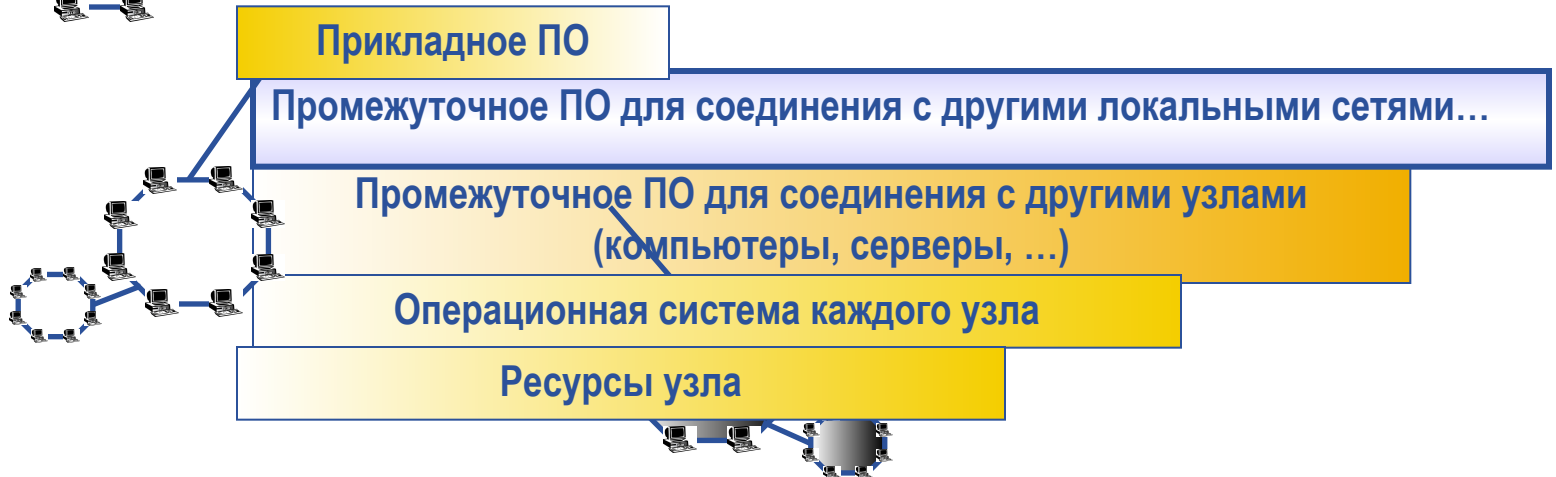


## Локальная сеть

Прикладное ПО  
Промежуточное ПО для соединения с другими ресурсами  
Операционная система каждого узла  
Ресурсы узла



## Следующий шаг - Грид

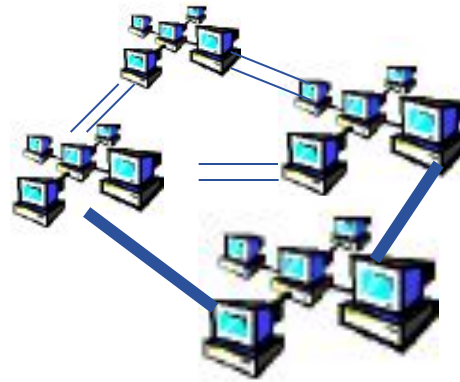


## Гетерогенные ресурсы

Собственность  
организаций и  
индивидуумов

Различные  
требования  
безопасности

Потенциальная  
недоступность  
ресурсов



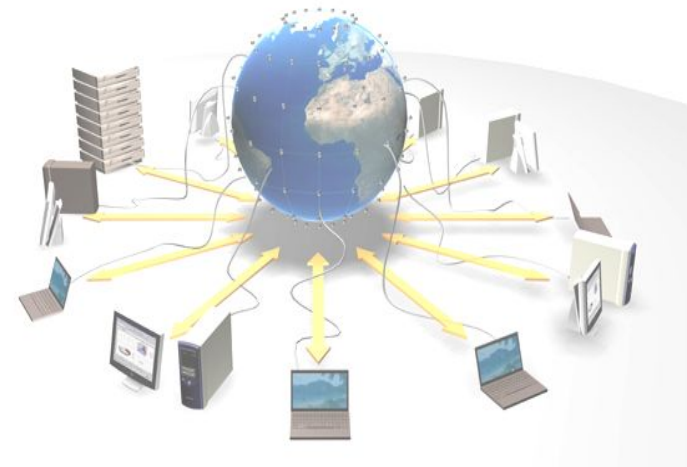
Соединение гетерогенных  
многоуровневых сетей

Различные политики  
управления  
ресурсами

Географическая  
распределенность

**World Wide Web** обеспечивает доступ к информации, которая находится в миллионах различных серверов, географически распределённых по всему миру

В отличие от WWW: **ГРИД** – это новая вычислительная инфраструктура, которая обеспечивает бесперебойный доступ к вычислительным мощностям и ресурсам хранения данных, распределённых по всему миру

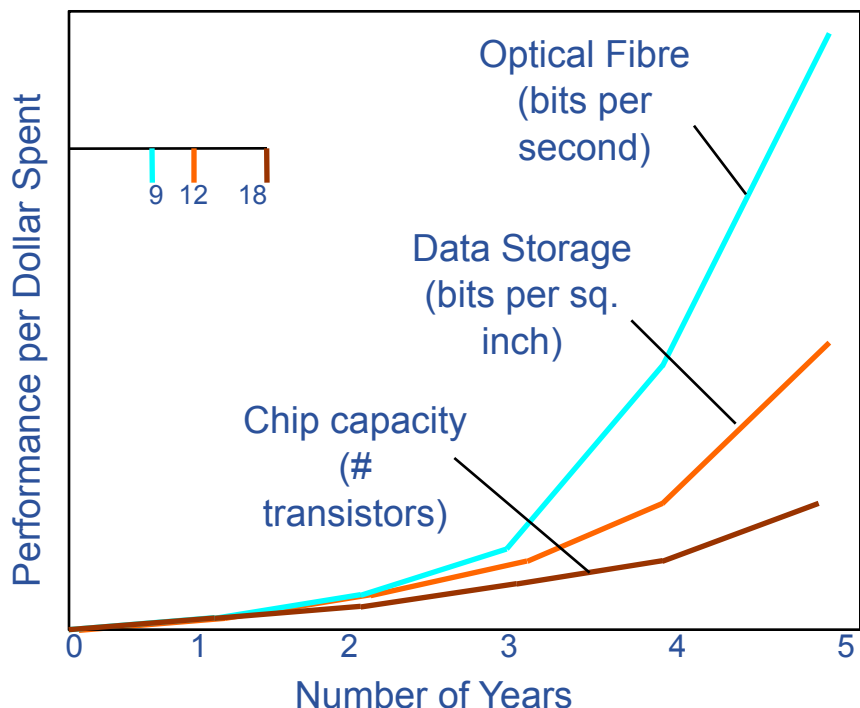


- Стремительное развитие сетевой транспортной среды и технологий высокоскоростной передачи данных.

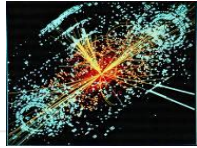
- скорости сетей удваиваются каждые 9 месяцев

- 1986 – 2000 : x 340 000

- 2001 – 2010 : x 4000



- Наличие во многих организациях высокопроизводительных вычислительных ресурсов (часто кластеров)
  - вычислительные скорости удваиваются каждые 18 месяцев
    - 1986 – 2000 : x 500
    - 2001 – 2010 : x 60
- Развитие традиционных Интернет-технологий
- Необходимость решения сложных научных, инженерных и бизнес-задач



Необходимость решения сложных научных, производственных, инженерных и бизнес-задач

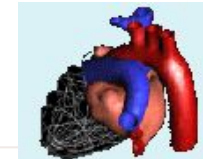
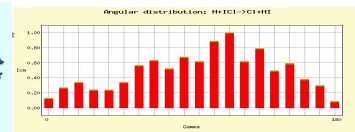
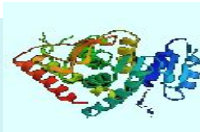
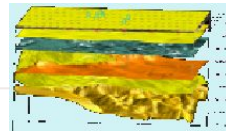
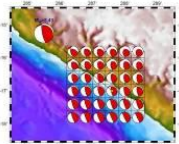


Наука все больше делается глобальные *коллаборации*



через распределенные использование *internet*

Желание расширить границы научного открытия вычислительным анализом и моделированием (**e-Science**)



ГРИД предоставляет *доступ* к:

- Очень большим наборам данных
- Огромным (тера...) вычислительным ресурсам
- Высокопроизводительной визуализации
- Высокоскоростным сетям



## Распределённые вычислительные инфраструктуры существовали и до Грид, но ....

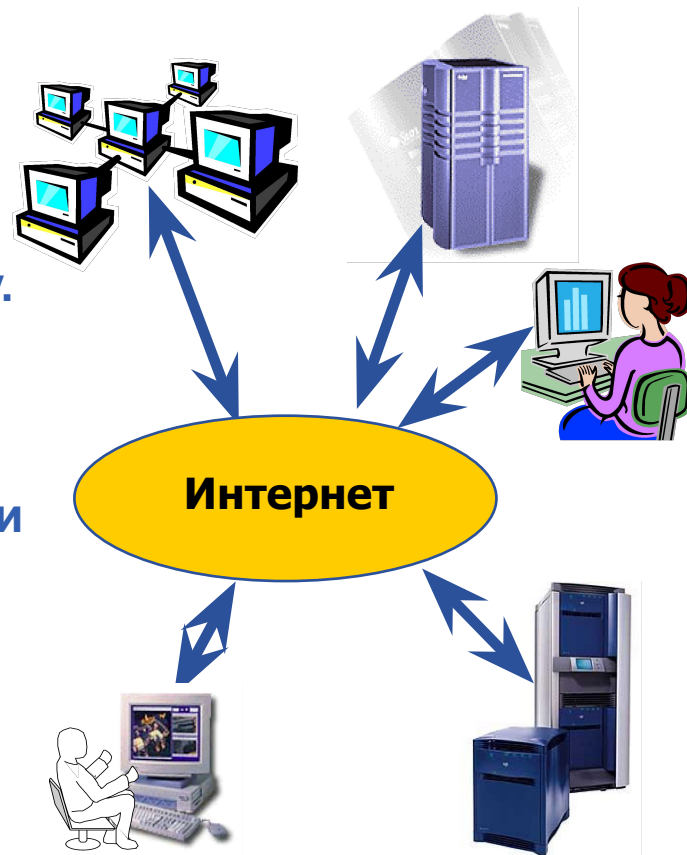
- обычно они представляют собой **локальные или специализированные системы**
  - созданные для одной определённой цели или для отдельной группы пользователей
  - обычно количество пользователей ограничено
  - не допускается совместное использование ресурсов других организаций

## Грид идёт дальше и предполагает:

- **Различные типы ресурсов**
  - не всегда одинаковые аппаратная часть, данные, приложения и политика администрирования
- **Различные способы взаимодействия**
  - различные группы приложений хотят взаимодействовать с Грид различными способами
- **Доступ к ресурсам через границы административных доменов для неограниченного количества не локальных пользователей**
- **Динамическая природа ресурсов**
  - Ресурсы часто добавляются/удаляются/изменяются
- **Масштабируемость до всемирных размеров**

Грид – это множество компьютеров

- (суперкомпьютеры, кластеры, персональные компьютеры, ...), средств хранения данных, специальных устройств, служб, которые могут динамически входить и покидать Грид-систему.
- Они гетерогенны во всех аспектах
- Они могут быть географически распределены и соединены скоростными сетями передачи данных
- Они могут быть доступны по запросу для некоторого множества пользователей



(Ian Forster: “What is Grid”, 2002 г.)

**«Грид – это система, которая:**

**координирует** использование ресурсов при отсутствии централизованного управления этими ресурсами

**использует** стандартные, открытые, универсальные протоколы и интерфейсы

**обеспечивает высококачественное обслуживание»**

- “С помощью Web мы можем взять из Интернета информацию и документы, а с помощью Grid мы можем взять оттуда и другие ресурсы: вычислительная мощность, хранилища данных, программное обеспечение и т.д. ...”.
- Грид-компьютинг – это скоординированное разделение ресурсов и решение задач в динамически меняющихся виртуальных организациях со многими участниками
- Виртуальная организация – объединение (коллаборация) специалистов из некоторой прикладной области, которые объединяются для достижения общей цели

**располагает** определенным количеством ресурсов, которые предоставлены зарегистрированными в ВО владельцами

**знает** кто что делает, и кому не будет позволено делать это снова...

**контролирует** доступ к описанным ресурсам

**устанавливает** самостоятельно правила работы для своих участников

**может образовываться** динамически и иметь ограниченное время существования

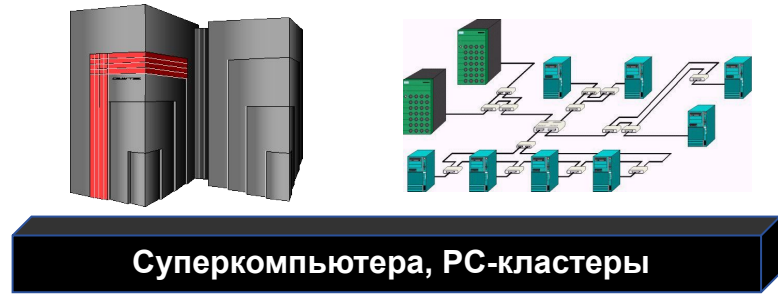
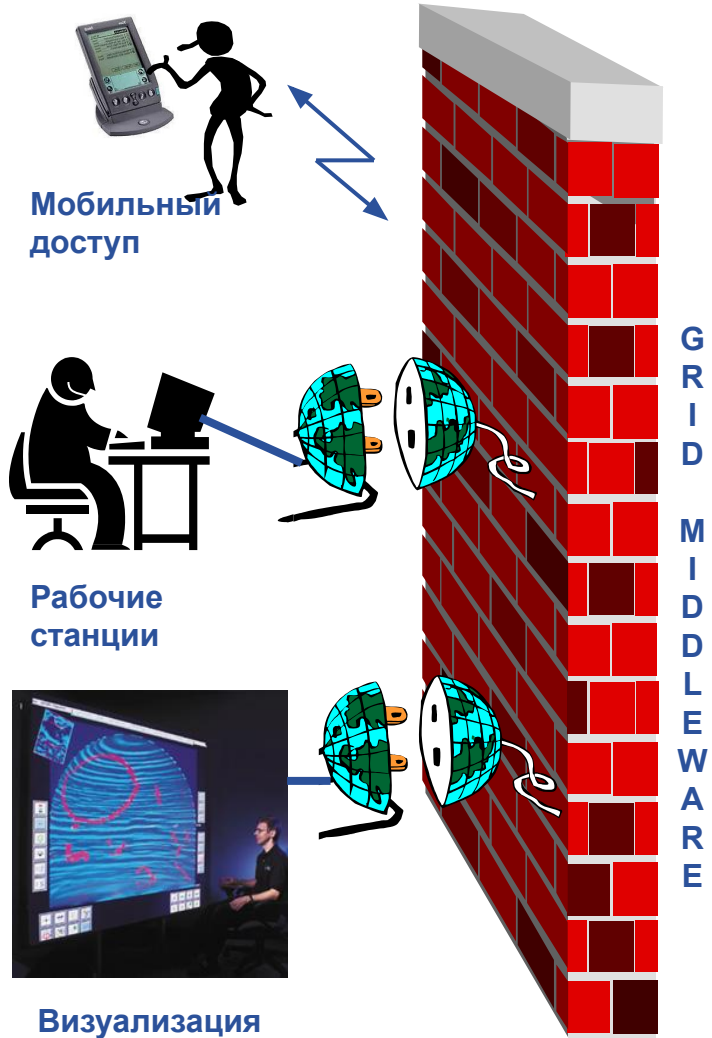
обычно **включает** географически распределенных пользователей

”Прозрачное” взаимодействие между гетерогенными ресурсами (принадлежащих различным, географически распределённым организациям), приложениями и пользователями возможно только при помощи...

- использования специализированного программного обеспечения, называемого **middleware**

Middleware “**скрывает**” технические детали Грид-инфраструктуры и обеспечивает безопасную интеграцию/распределение ресурсов

*интернет-протоколы не обеспечивают безопасный механизм доступа к разделяемым ресурсам*





- **Используя ПК или workstation**

- Для входа – имя (account) и пароль (“Аутентификация”)
- Использование прав, разрешенных вам (“Авторизация”)
- Запуск заданий
- Управление файлами: создание, чтение/запись, список директорий

- **Компоненты системы связываются по шине**

- **Операционная система**

- **Один административный домен**

- **Используя Грид**

- Для входа – единый цифровой сертификат (“Аутентификация”)
- Использование прав, делегированных вам (“Авторизация”)
- Запуск заданий
- Управление файлами: создание, чтение/запись, список директорий

- **Сервисы связываются через Интернет**

- **ППО (middleware)**

- **Множество административных доменов**



**User Interface (UI)**: Для пользователя - точка входа в Грид



**Resource Broker (RB) (Workload Management System (WMS))**:  
Определяет соответствие между требованиями задания и доступными Грид-ресурсами



**Information System**: Содержит характеристики и статус ресурсов



**File and replica catalog**: местоположение файлов в Грид и их реплик



**Logging and Bookkeeping (LB)**: Статус заданий, выполняющихся в Грид



**Computing Element (CE)**: Очередь заданий на вычислительном узле, где может выполняться задание пользователя



**Storage Element (SE)**: Сервер хранения данных, где сохраняются Грид файлы (чтение/запись/копирование) или их реплики





Необходимость решения  
сложных задач

e-Science

Наличие вычислительных  
ресурсов

Увеличение  
производительности  
микропроцессоров

доступность сетевой  
транспортной среды

Технологии Интернет

**Физика высоких энергий**  
*симуляция, реконструкция, анализ, ...*

**Медицина / Здоровоохранение**  
*отображение, диагностика и лечение*

**Биоинформатика** *геном человека, ...*

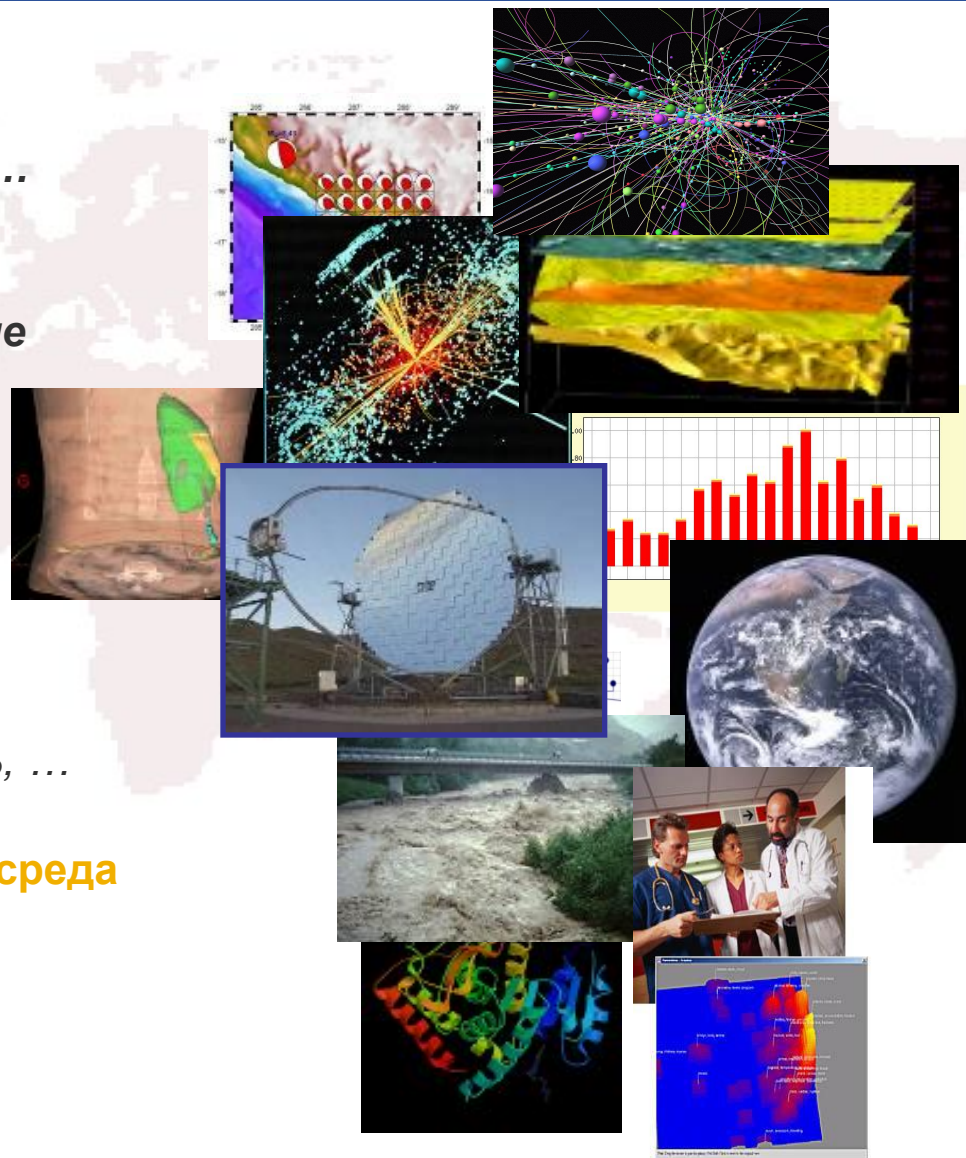
**Нанотехнологии**  
*разработка новых материалов на молекулярном уровне, ...*

**Инженерия** *авиационная безопасность, ...*

**Природные ресурсы и Окружающая среда**  
*прогноз погоды, ...*

**Искусство**

**Материаловедение**

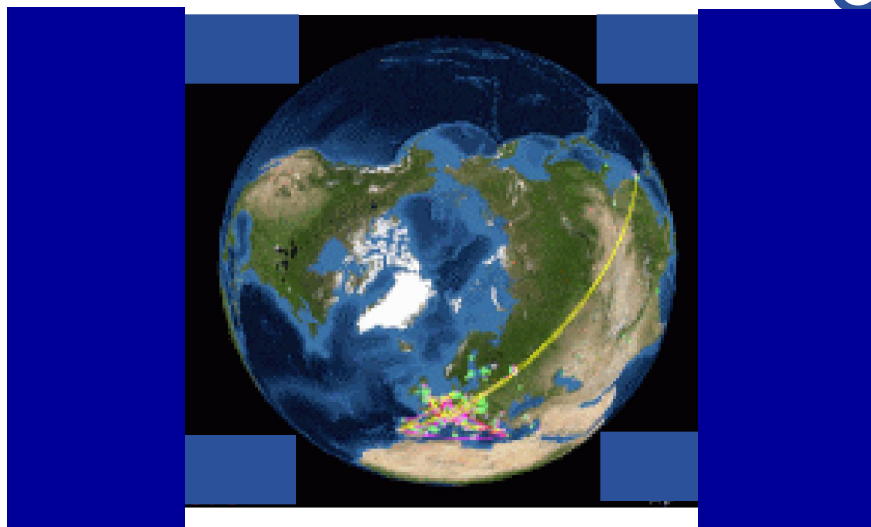


- Англия UK National Grid Service <http://www.ngs.ac.uk>
- Австрия Austrian Grid Initiative <http://www.austriangrid.at>
- Бельгия BEgrid <http://www.begrid.be/>
- Болгария Bulgarian Grid Consortium, BGGC, <http://www.grid.bas.bg/consortium.htm>
- Хорватия CRO-GRID <http://www.srce.hr/crogrid/>
- Чехия METACentrum <http://meta.cesnet.cz>
- Эстония Estonian Grid, working under Estonian Educational and Research Network (EENet) <http://grid.eenet.ee>
- Финляндия CSC – Scientific Computing Ltd., <http://www.csc.fi>
- Франция EGEE-FranceGermanyD-Grid, <https://www.d-grid.de/index.php?id=1&L=1>
- Греция HellasGrid, HG, <http://www.hellasgrid.gr>
- Ирландия Grid-Ireland <http://www.grid.ie/>
- Израиль Israel Academic Grid (IAG) <http://iag.iucc.ac.il>
- Италия Italian Grid Infrastructure, IGI, <http://www.italiangrid.org>
- Литва Lithuanian Grid LitGrid <http://www.litgrid.lt>
- Голландия Netherlands National Science Grid, internally known as BIGGRID project
- Норвегия Norwegian Grid (NorGrid), <http://www.norgrid.no>
- Польша PL-Grid, <http://plgrid.cyfronet.pl> (public part in preparation)
- Португалия Iniciativa Nacional Grid – INGRID <http://www.gridcomputing.pt/>
- Румыния RoGrid URL: <http://www.rogrid.ro> (old version)
- **Россия Russian Data Intensive Grid (RDIG): <http://www.egee-rdig.ru>**
- Сербия Serbia Academic and Educational Grid Initiative of Serbia AEGIS <http://aegis.phy.bg.ac.yu>
- Украина Ukrainian Grid, UGrid <http://grid.ntu-kpi.kiev.ua/>
- Швеция Swedish Grid, SweGrid, <http://www.swegrid.se/> and <http://www.snrc.vr.se/>
- Турция TR-Grid National Grid Initiative, TR-Grid NGI <http://www.grid.org.tr>
- Германия(D-initiative)

- <http://www.eu-egee.org> - проект EGEE-II
- <http://egee.lib.ed.ac.uk/> - репозиторий материалов (EGEE Digital Library)
- <http://www.egee-rdig.ru/> - RDIG консорциум
- <http://roston.jinr.ru:8080/accounting/> - мониторинг RDIG
- <http://www.gridclub.ru/> - GRIDCLUB.RU :: Интернет-портал по грид-технологиям
- <http://egee.sinp.msu.ru> - НИИЯФ МГУ
- <http://egee.pnpi.nw.ru/> - ПИЯФ РАН

*Добро пожаловать*

*работатъ в ГРИД!  
Спасибо за внимание*



*Спасибо за внимание*