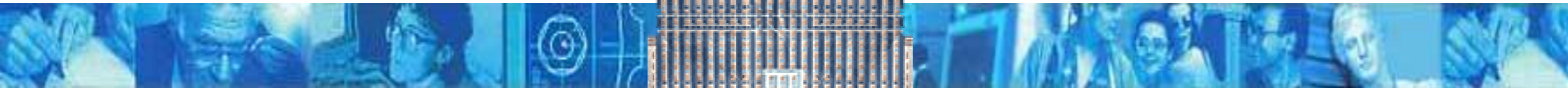


Распределенные объектные технологии

Лекция 1 Введение



Содержание

- Введение
- История возникновения вычислительных сетей
- История и виды распределенных вычислений
- Грид

Введение

Объем курса

- Лекционные занятия:
 - 9 (базовые лекции)
- Практические занятия:
 - 9 докладов
 - 9 пар на разработку распределенной вычислительной системы
- Экзамен:
 - Тест в системе «Компетентум» по лекциям и докладам

Темы основных лекций

- История и развитие распределенных вычислений.
- Грид: определение, применение, стандартизация.
- Слои архитектуры грид, основные сервисы грид.
- Основные технологии и стандарты распределенных вычислений и грид: WSRF, OGSA;
- Технологии разработки и использования ресурсов Грид (Система Globus)
- Основы облачных вычислений

Темы докладов

№	Число	Тема доклада
1	11.02.2012	Платформа UNICORE
2	11.02.2012	Платформа BOINC
3	25.02.2012	Windows Communication Foundation
4	25.02.2012	Платформа Windows Azure
5	10.03.2012	Платформа Google App Engine
6	10.03.2012	Платформа Amazon Web Services
7	24.03.2012	Фреймворк MapReduce
8	24.03.2012	Система HADOOP

Дополнительные темы:

- концепция REST

Еще?

История становления вычислительных сетей

Начало... ARPANET (1969)

- Предшественник Internet.
- Основывалась как «Сеть исследователей».
- Начинаясь со скорости 56 kbps.
- Основывалась на технологии TCP/IP.

Основные задачи

- **E-mail**
- Удаленный доступ к ресурсам
- Удаленное управление заданиями

Удаленный доступ к ресурсам

- Обеспечение распределенного доступа к ограниченным ресурсам mainframe-компьютеров для географически-распределенных пользователей.
- Это позволяло отказаться от дорогих переездов к компьютерным центрам для постановки задач и обеспечивало удаленное использование ресурсов.

Начало... Основа ARPANET - TCP/IP

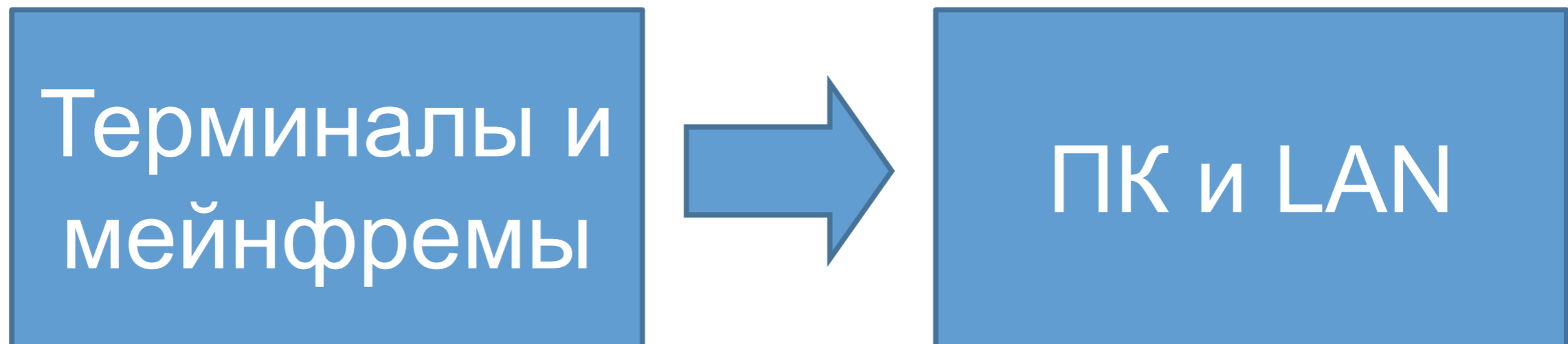
- TCP/IP развивался более активно, чем другие технологии связи (FidoNet, UUCP, OSI, X.25...)
- Причины определения TCP/IP как стандарта:
- Открытая документация протокола;
- Открытый исходный код;
- **Использование архитектура ARPANET/Internet и протокола TCP/IP для обучения => большой объем исследований в университетских лабораториях.**

Развитие и включение ПК в сеть (1)

- Изначально, LAN не был предусмотрен в структуре ARPANET. Доступ в сеть предоставлялся с терминала mainframe-компьютера либо с «миникомпьютера» DEC PDP-11.
- Но разработка прототипа Ethernet (3 mbps) и рабочей станции в Xerox Palo Alto Research Center (PARC) явилось первым шагом к формированию LAN.

Развитие и включение ПК в сеть (2)

- Рост мощности ПК привел к тому, что они смогли обеспечить доступ в сеть.
- Таким образом, модель сети сменилась:



- Мейнфреймы никуда не исчезли, но преобразовались в серверы.

История и виды распределенных вычислений

Развитие распределенных вычислений и грид

- Первый этап.
 - Середина 1990-х – становление и развитие систем распределенных вычислений.
- Второй этап.
 - 1998 г. – определение термина Grid (грид) в рамках книги «Грид. Новая инфраструктура вычислений». Развитие распределенных систем, ориентированных на массивные объемы передачи информации и вычислительные затраты.
- Третий этап.
 - 2001 г. – уклон в сторону “Виртуальных организаций”. Развитие сервисно-ориентированных подходов (SOA), автоматизация методов управления ресурсами.

Первый этап – становление распределенных вычислений

- Первые проекты по распределенным вычислениям (начало 1990-х) основывались на объединении вычислительных ресурсов суперкомпьютеров.
- Проект FAFNER: Factoring via Network-Enabled Recursion (Сетевое разложение на множители посредством рекурсии).
- Проект I-WAY: Information Wide Area Year (Год Информации Глобальных Сетей) – экспериментальная высокопроизводительная сеть, которая объединяла множество высокопроизводительных компьютеров и передовые средства визуализации. Прообраз Globus.

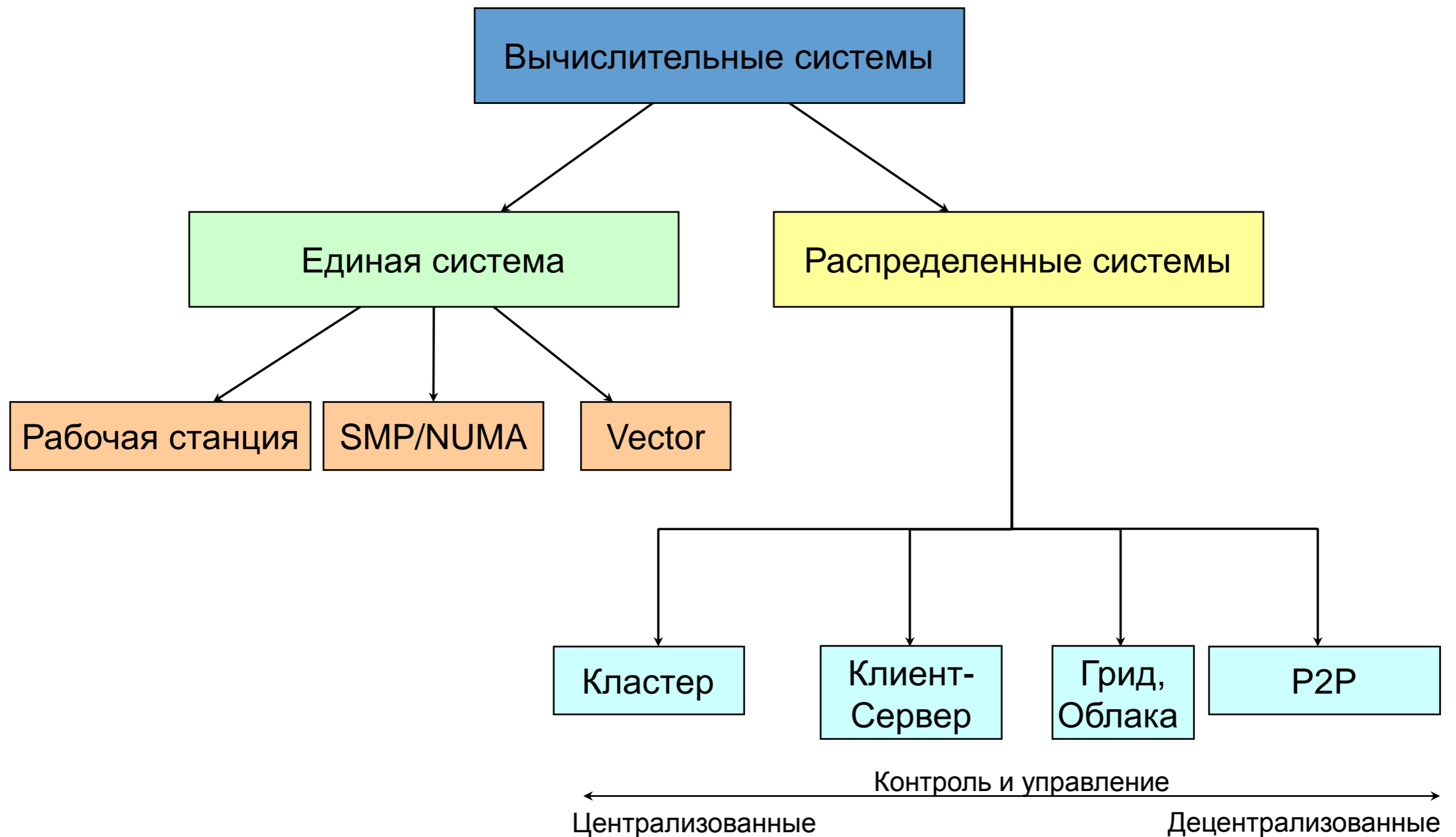
Второй этап – развитие средств создания грид

- Ближе к концу 1990-х годов наблюдается бурное развитие Grid-систем, начинаются разработки средств создания и управления Grid-сетями и вычислительными ресурсами.
- В ходе исследований систем распределенных вычислений, Ян Фостер вывел 3 основных требования, которым они должны удовлетворять.
 - Гетерогенность.
 - Масштабируемость.
 - Адаптируемость.
- Основные проекты: Globus, SETI@home (P2P)...

Третий этап – грид как виртуальная организация

- Развитие технологий в начале 21-го века приводит к переосмыслению понятия Grid. Теперь оно неразрывно связано с понятием “Виртуальной организации”.
- Развитие сервисно-ориентированных подходов позволяет гибко использовать одни и те же вычислительные ресурсы многими пользователями.
- Основные проекты и технологии: Globus, OGSA, WSRF...

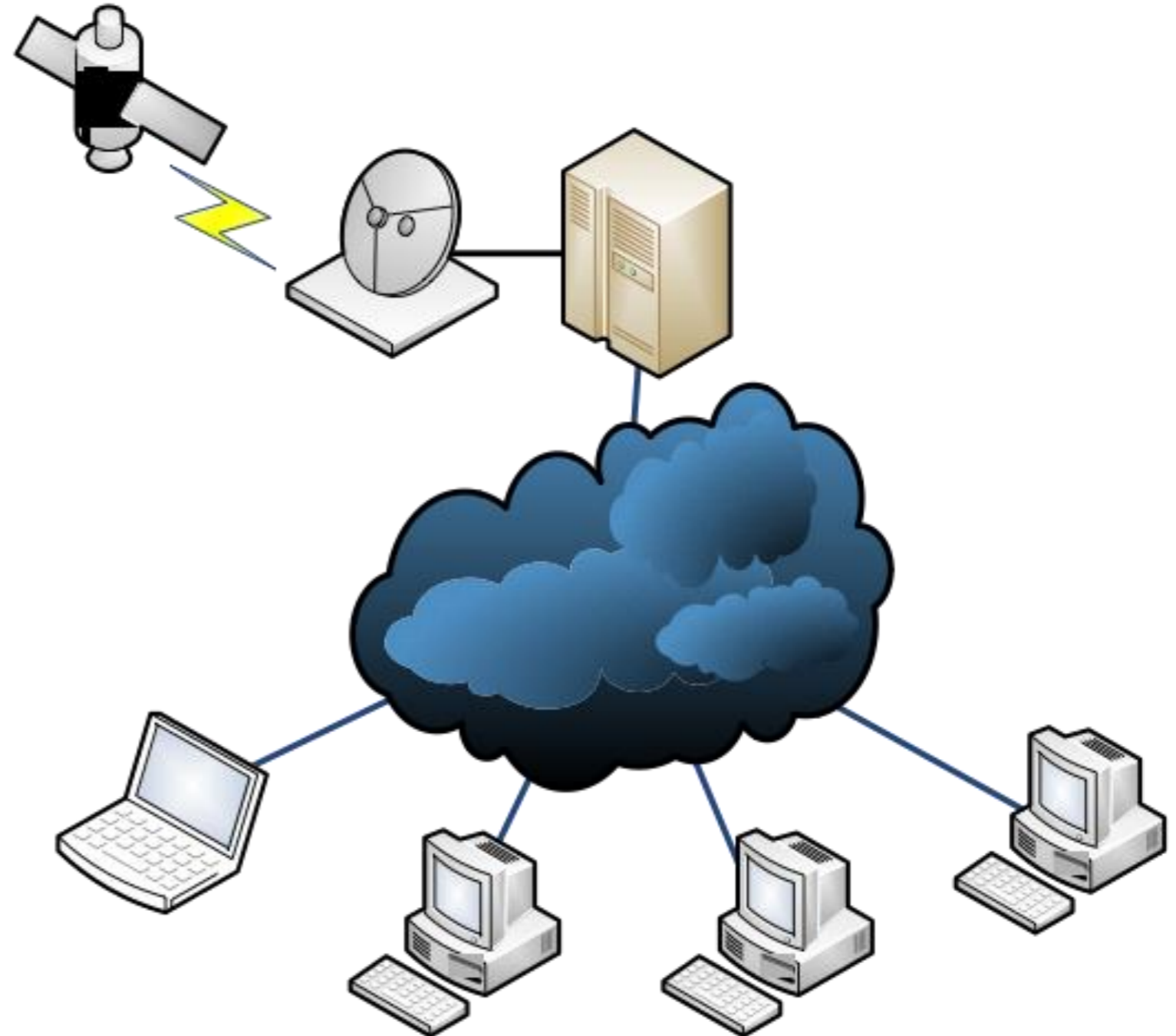
Виды вычислительных систем



Интернет вычисления

Примеры проектов:

- ⑩ Платформа BOINC (Berkeley Open Infrastructure for Network Computing):
 - ⑩ SETI@home
 - ⑩ Genome@home
 - ⑩ Folding@home

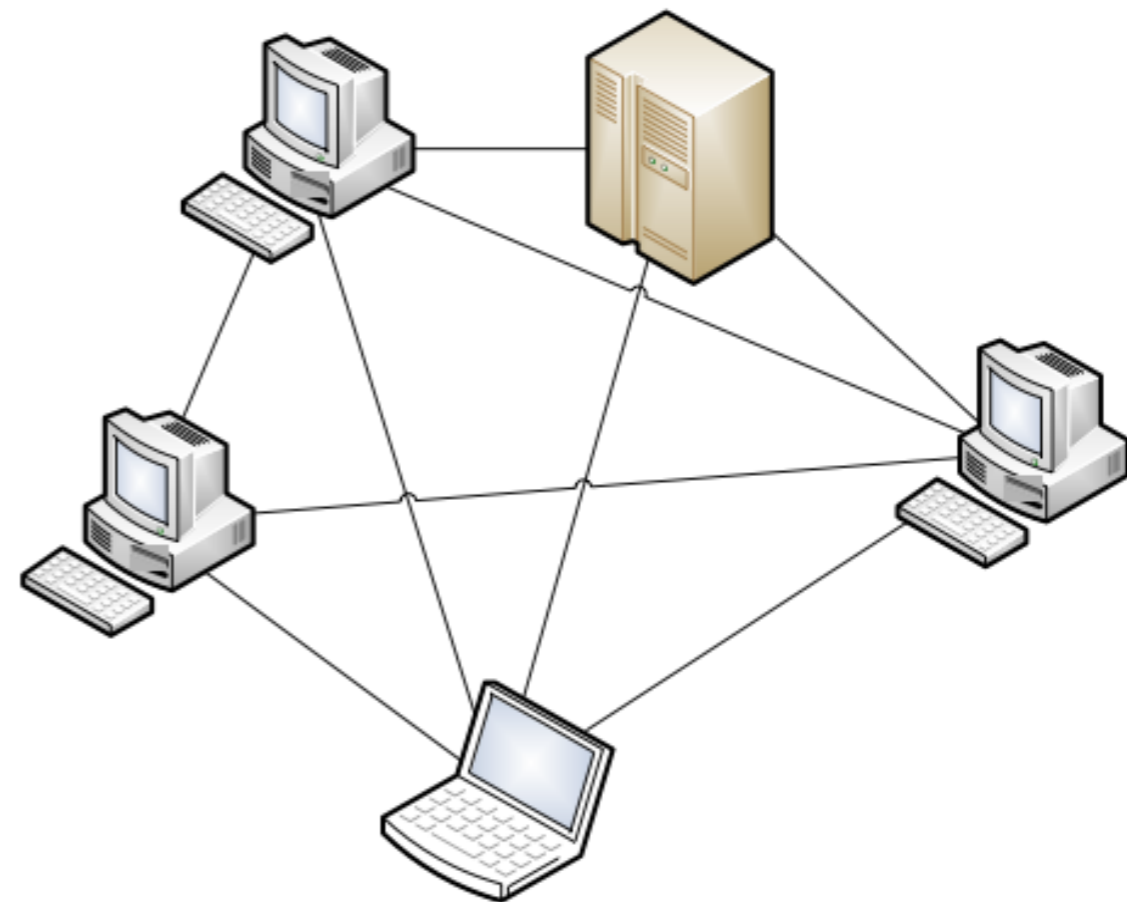
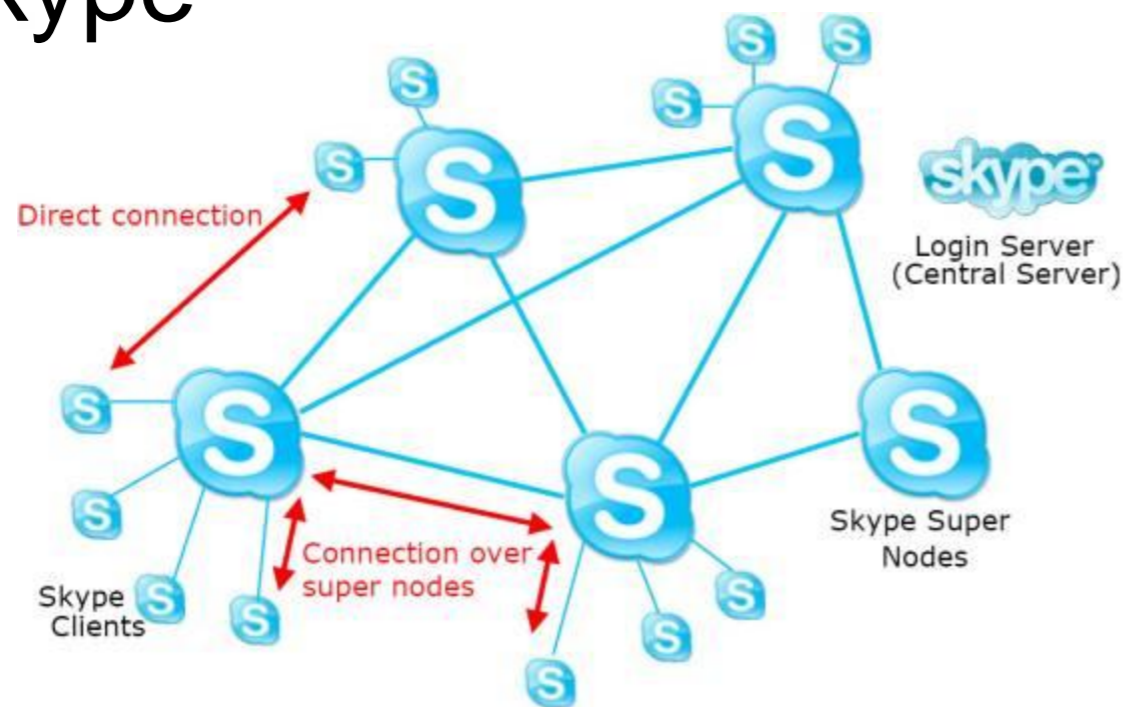


Клиент-серверная архитектура

P2P

Примеры проектов:

- eDonkey
- Kazaa, Napster (RIP)
- BitTorrent
- Jabber
- Skype



Грид

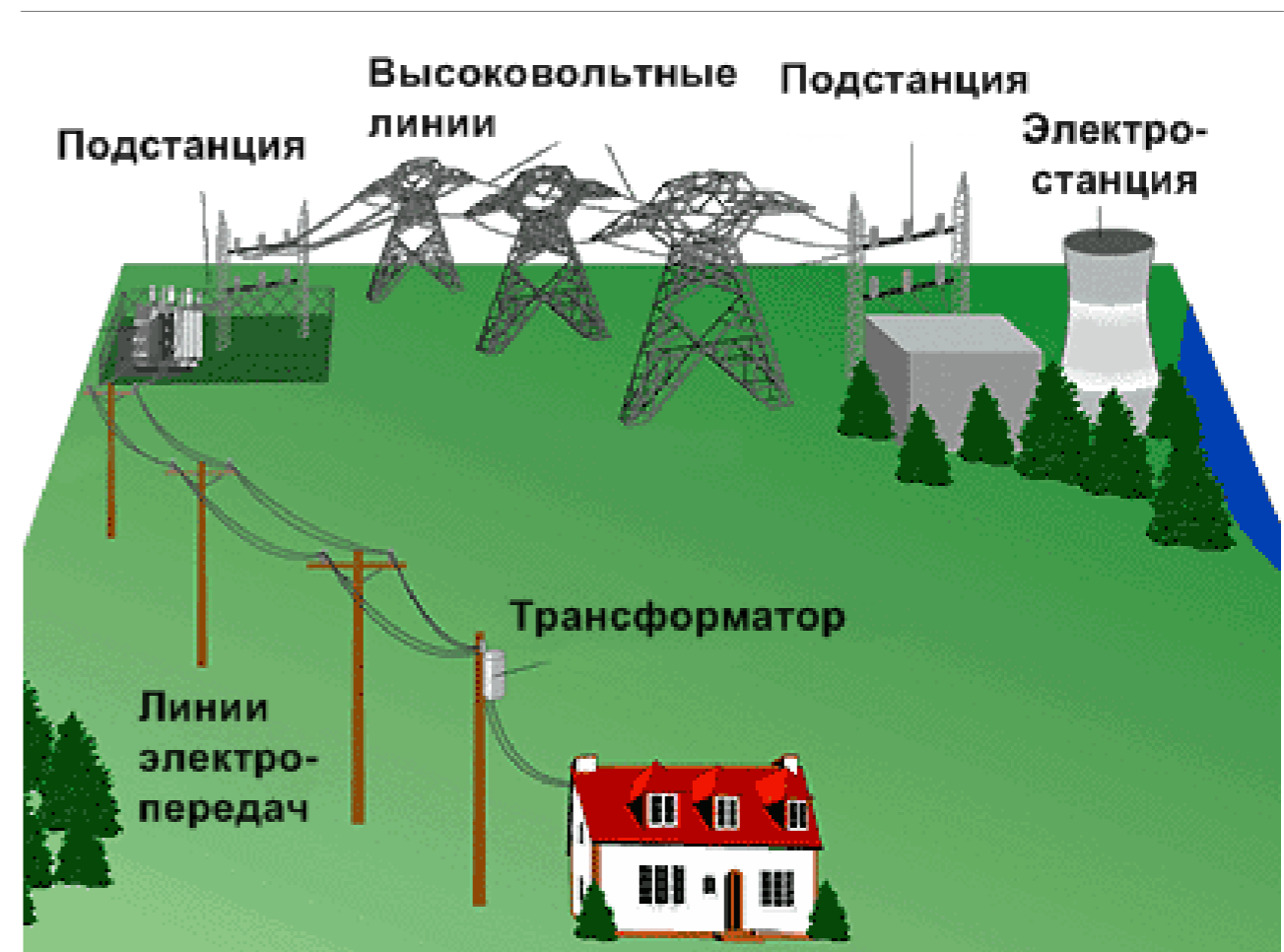
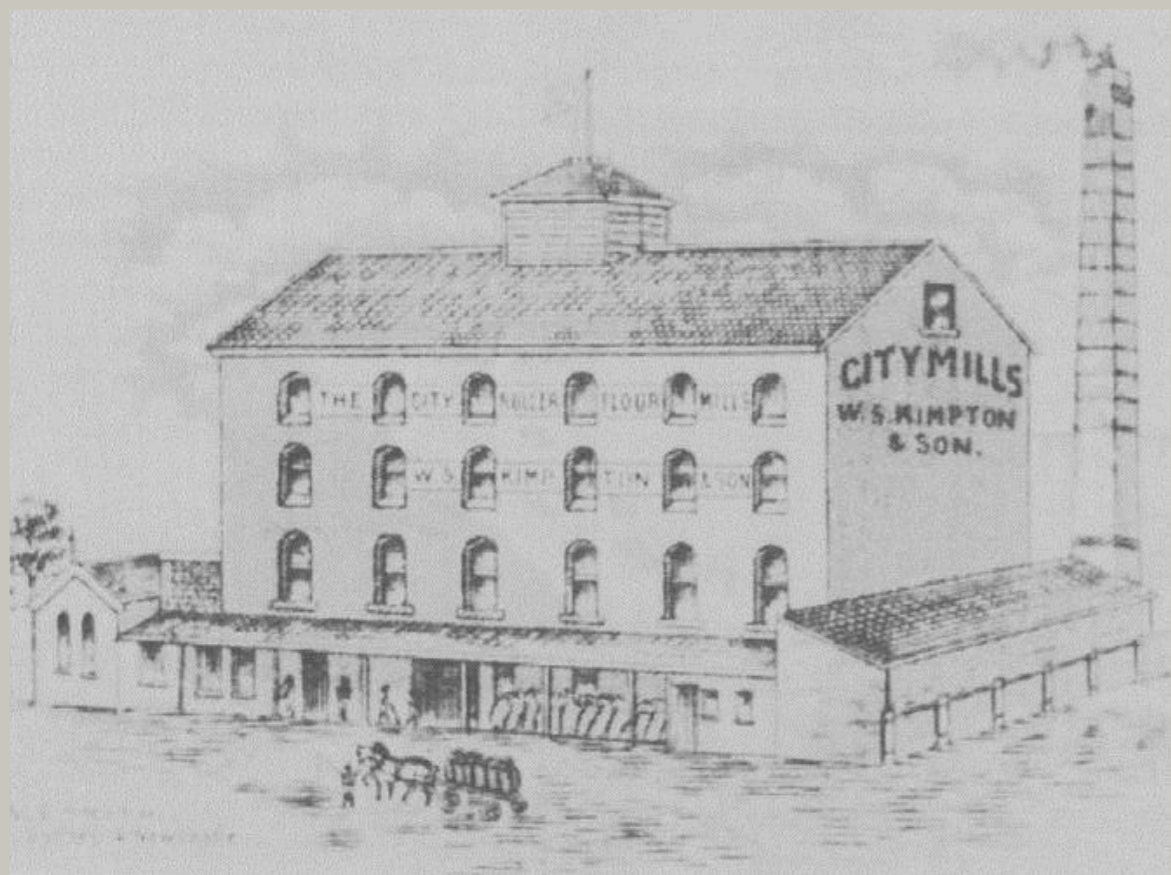
Что такое грид?

- Грид – как электросеть:
 - 1) координирует ресурсы, не подчиняющиеся централизованному контролю...
 - 2) ...используя стандартные, открытые протоколы и интерфейсы общего назначения ...
 - 3) ...чтобы предоставить возможность работы с нетривиальными свойствами служб.

Метафора грид

- Grid (англ.) – сеть, энергетическая система.

Либо покупать (или переезжать к) источнику энергии



Либо передавать энергию туда, где мы находимся

Обеспечение доступа по требованию и интеграция различных ресурсов и служб, независимо от местонахождения

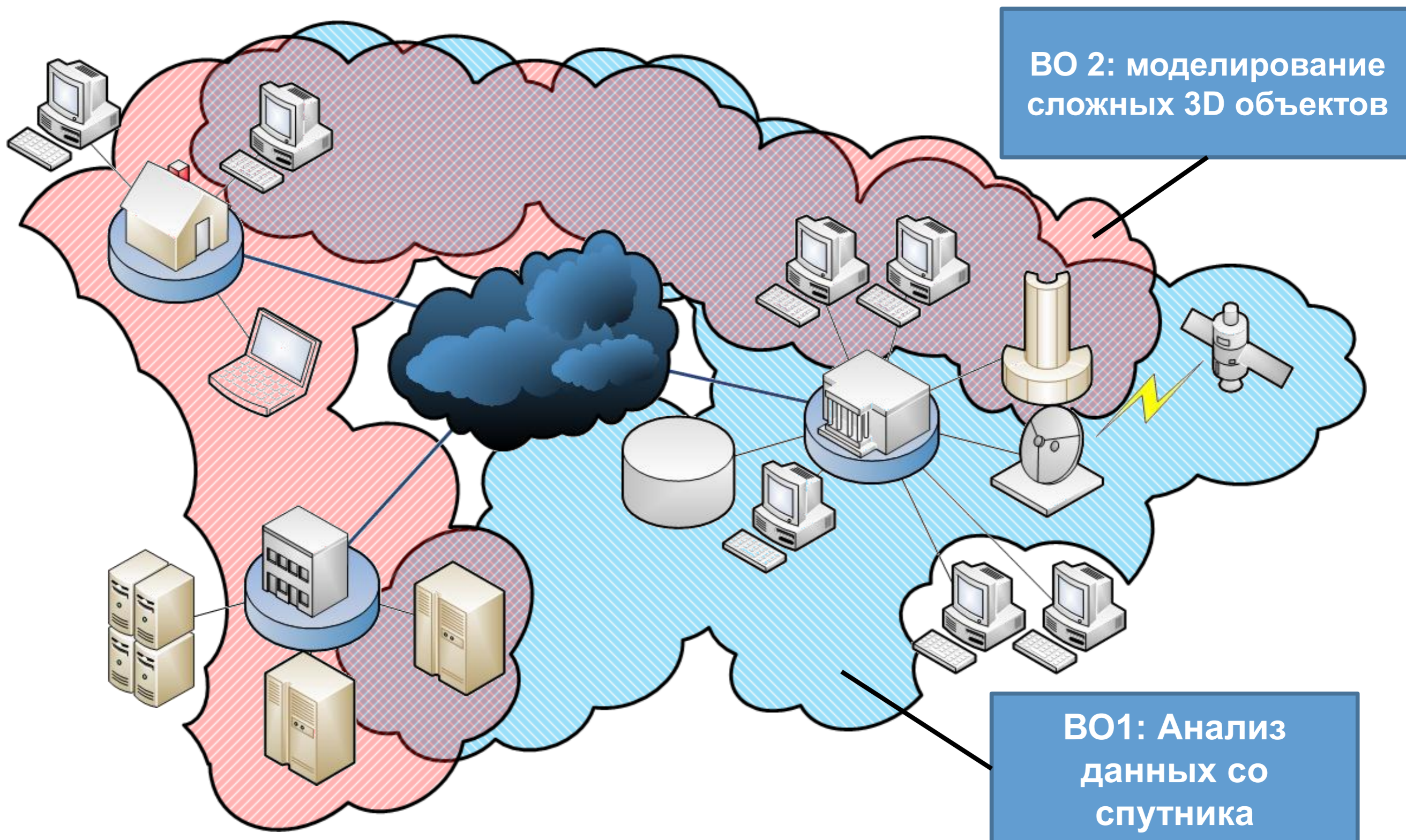
Концепция Грид

- Аналогия с энергетической системой (Power Grid)
 - Повсеместный, стандартный, надежный и дешевый доступ к ресурсам
 - Сложность инфраструктуры скрыта от пользователя (прозрачность)
 - Легкость интеграции новых ресурсов
 - Динамическое перераспределение нагрузки и отказоустойчивость
 - Управление сложным ансамблем ресурсов

Концепция Виртуальной Организации (ВО)

- ВО - это ряд людей и/или организаций, объединенных общими правилами коллективного доступа к определенным вычислительным ресурсам.
- Например:
 - провайдеры прикладных услуг, провайдеры услуг хранения;
 - участники промышленного консорциума, финансирующие создание нового самолёта;
 - участники многолетних, крупных международных объединений в области физики высоких энергий.

Пример ВО



Грид VS Все остальное

- Три критерия грид:
 - Осуществляется координация ресурсов, не управляемых централизованно
 - С использованием стандартных, открытых и универсальных протоколов и интерфейсов
 - Для предоставления нетривиального качества обслуживания
- Что не является Grid?
 - Кластер, сетевое хранилище данных, научный прибор, сеть, ...
 - Важные компоненты Grid, но сами им не являются

Сопоставление кластера, грид и P2P

Характеристика	Кластер	Грид	P2P
Вычислительные узлы	Обычные компьютеры	профессиональные компьютеры	Настольные PC
Собственники	один	многие	многие
Обнаружение	службы членства	централизованный индекс и децентрализованная информация	децентрализовано
Распределение/ планирование/ управление	централизованное	децентрализовано	децентрализовано
Интероперабельность	на базе VIA	более развитое (например, WSRF)	стандартов нет
Образ одной системы	да	нет	нет
Масштабируемость	100	1000	миллионы
Возможности	гарантированы	переменны, но ВЫСОКИ	переменны

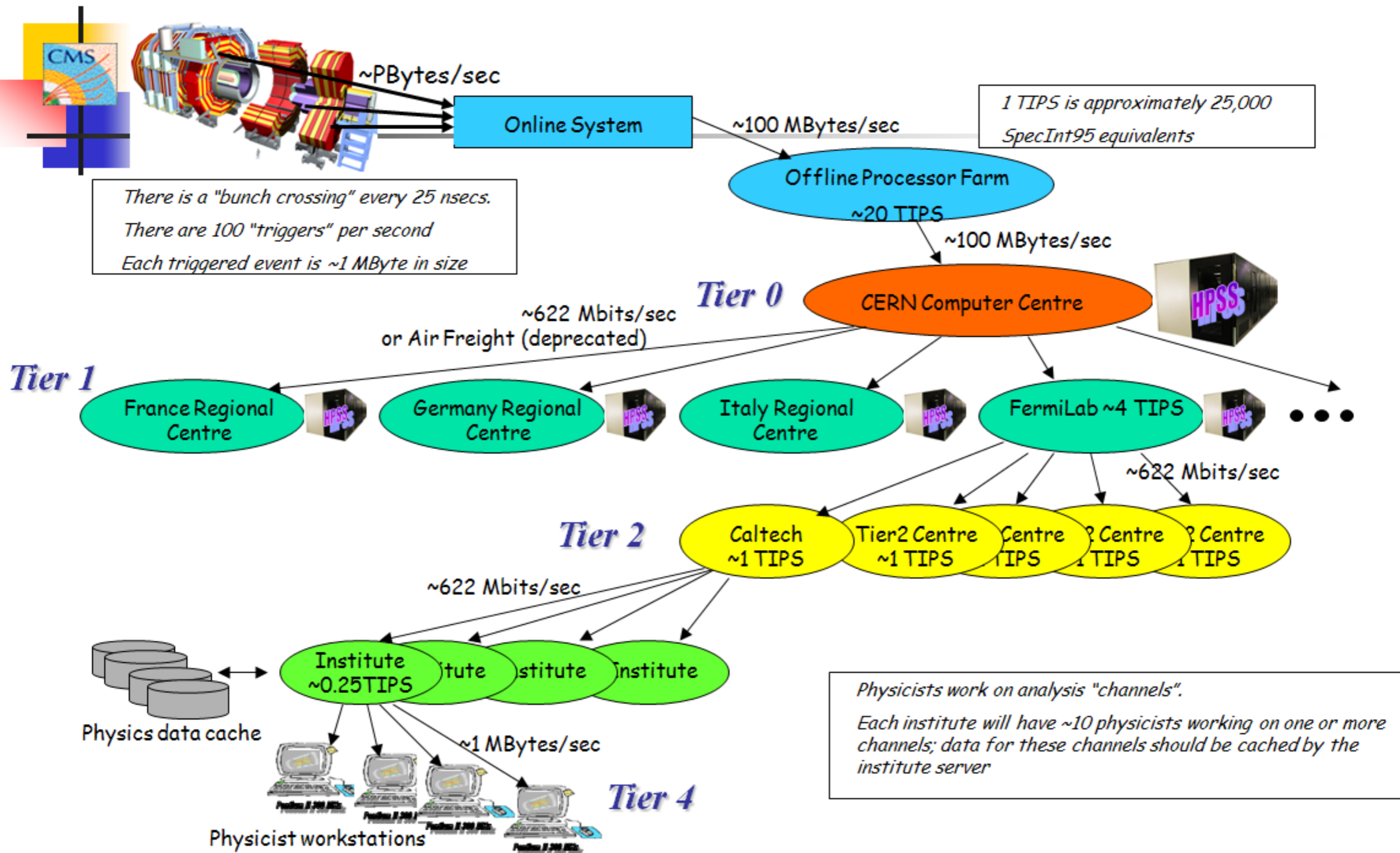
Популярные приложения Грид

- Высокопроизводительные вычисления
 - Интерактивное моделирование (климат)
 - Имитационные эксперименты с высоким разрешением / большим числом объектов (формирование галактик, гравитационные волны, имитация боевых действий)
 - Проектирование (прогон вариантов, объединение моделей компонентов)
- Работа с большими массивами данных
 - Анализ экспериментальных данных (физика высоких энергий)
 - Анализ изображений и данных зондирования (астрономия, изучение климата, экология)
- Совместная работа в рамках распределенных коллективов
 - Удаленное управление приборами (микроскопы, рентгенокопия , ...)
 - Удаленная визуализация (изучение климата, биология)
 - Engineering (large-scale structural testing, chemical engineering)
- Сложность задач требует совместной работы людей из различных организаций и разделяемого использования ресурсов, данных и приборов.

Пример применения грид : Большой Адронный Коллайдер (1)



Пример применения грид : Большой Адронный Коллайдер (2)



Облака

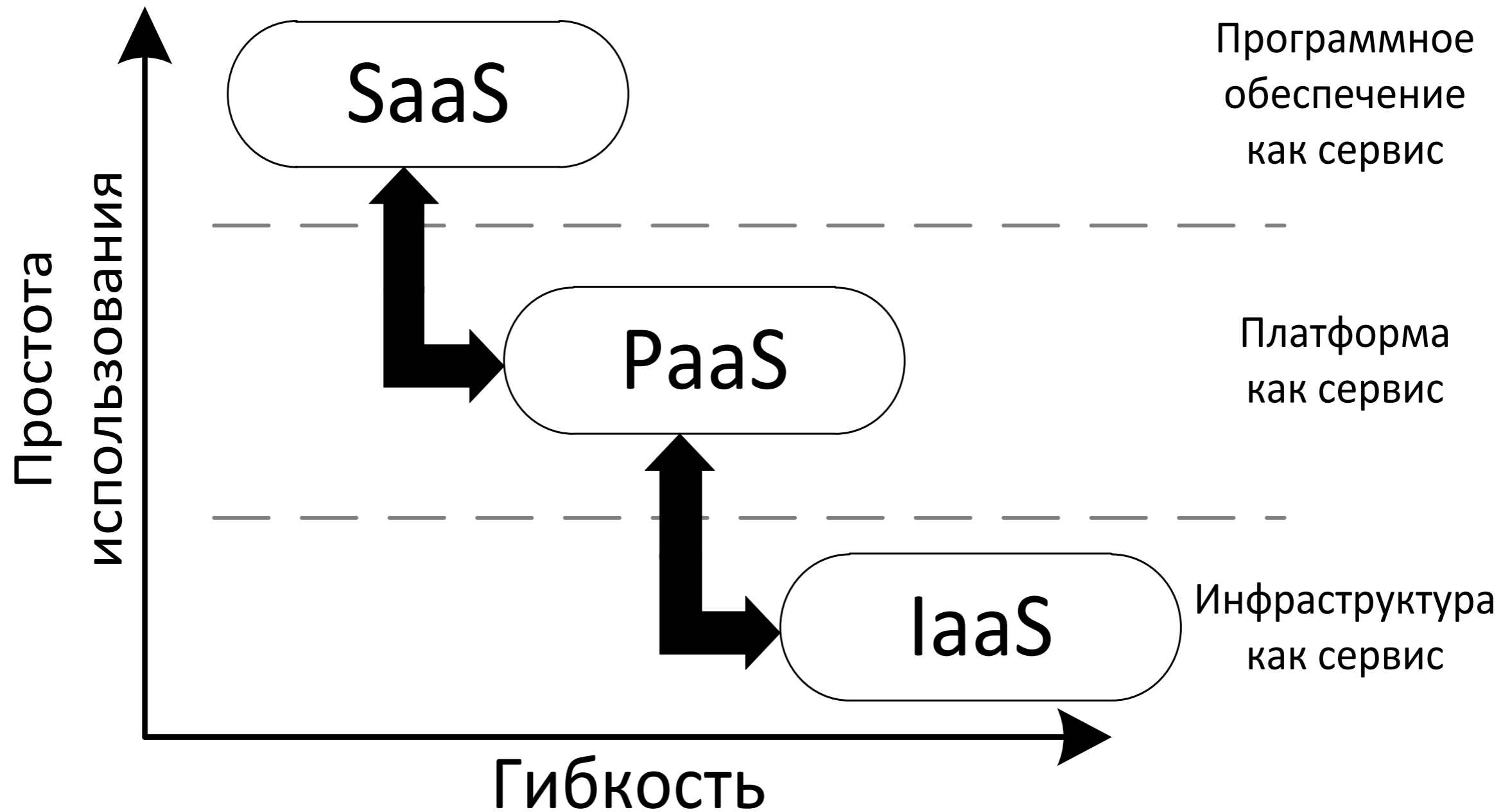
Коммунальные вычисления

- Предвестник облачных вычислений – «*utility computing*» (коммунальные вычисления) - предоставление данных и процессорных мощностей организовано по принципам коммунальных услуг, по принципу «оплата по мере использования».

Облачные вычисления

- *Облако* – это большой пул легко используемых и легкодоступных виртуализованных ресурсов.
 - Эти ресурсы динамически масштабируются для подстройки под текущую нагрузку, обеспечивая оптимальное использование и обычно предоставляется по принципу «оплата по мере использования».

Классификация облаков



Грид VS Облачные вычисления

- Термин «Облачные вычисления» пришел в мир информационных технологий в 2007 г., начиная с совместного проекта «Cloud Computing» компаний Google и IBM.
- Сейчас облачные вычисления – это множество платформ, предоставляемых компаниями Google, Microsoft, Amazon

Грид VS Облако

Характеристика	Грид	Облако
Вычислительные узлы	профессиональные компьютеры (кластеры)	Объединенные кластера и сетевые хранилища
Операционные системы	Любая стандартная ОС (доминирует Unix-подобные)	Гипервизор виртуальных машин со множеством ОС
Право собственности	Децентрализовано	Единый владелец
Интерконнект	Интернет со средней пропускной способностью	Централизованный, с малой латентностью и высокой пропускной способностью
Безопасность	На основе открытых/закрытых ключей	Каждому пользователю выделяется виртуальная машина
Стандарты	Стандарты OGF	Веб-сервисы
Стоимость	Обычно бесплатно в академических кругах	Оплата по факту использования ресурсов
Возможности разработки собственных решений	Ограничены в связи с ориентацией на научное использование	Высокие – ориентированы на предоставление сервисов с динамическим изменением нагрузки

Информация о Грид

- На Русском языке: портал <http://www.gridclub.ru/>
 - Статьи Яна Фостера
 - [«Что такое грид? Три критерия»](#)
 - [«Анатомия грид: создание масштабируемых виртуальных организаций»](#)
 - Статья «Определение понятия грид: фиксируем точку зрения экспертов»
 - Книги с Amazon.com