

Вычислительная инфраструктура исследовательских ядерно-физических лабораторий

(введение)

- **Современные запросы на вычисления**
 - наука/техника/образование
 - бизнес
 - государственные структуры
- **История вычислителей.**
- **Организация крупномасштабных вычислений**
 - оборудование, программы
 - Вычисления в сетях Grid и «облачных» сетях
 - Вычисления в крупных физических экспериментах
- **Особенности вычислений в ФВЭ**
 - Большие объёмы данных
 - Коллективная работа

Роль вычислений в обществе (не только в ядерной физике)

- Развитие человечества как природного феномена невозможен без экономического роста.
- Экономический рост в настоящее время и будущем может основываться лишь на использовании инноваций в массовом порядке.
- Любые инновации возможны лишь с усложняющейся переработкой больших объёмов информации (проще - вычислений).

Наука/техника/образование

- **Расчёты/моделирование в**
 - сейсмологии (землетрясения, проч.);
 - гидрологии (плотины, проч.);
 - органической и неорганической химии;
 - **ядерной физике (физика высоких энергий, реакторы);**
 - прогнозах (атмосфера, океаны, цунами, штормы, проч.);
 - биологии/медицине (томограммы);
 - поиске полезных ископаемых;
 - мировой экономике (по отраслям, глобально, по регионам, поведение биржи, кризисы);
 - космических исследованиях;
 - в коммуникациях (технических и человеческих);
 - инженерных расчётах (транспорт, двигатели, горючее, механизмы, сооружения, проч.).

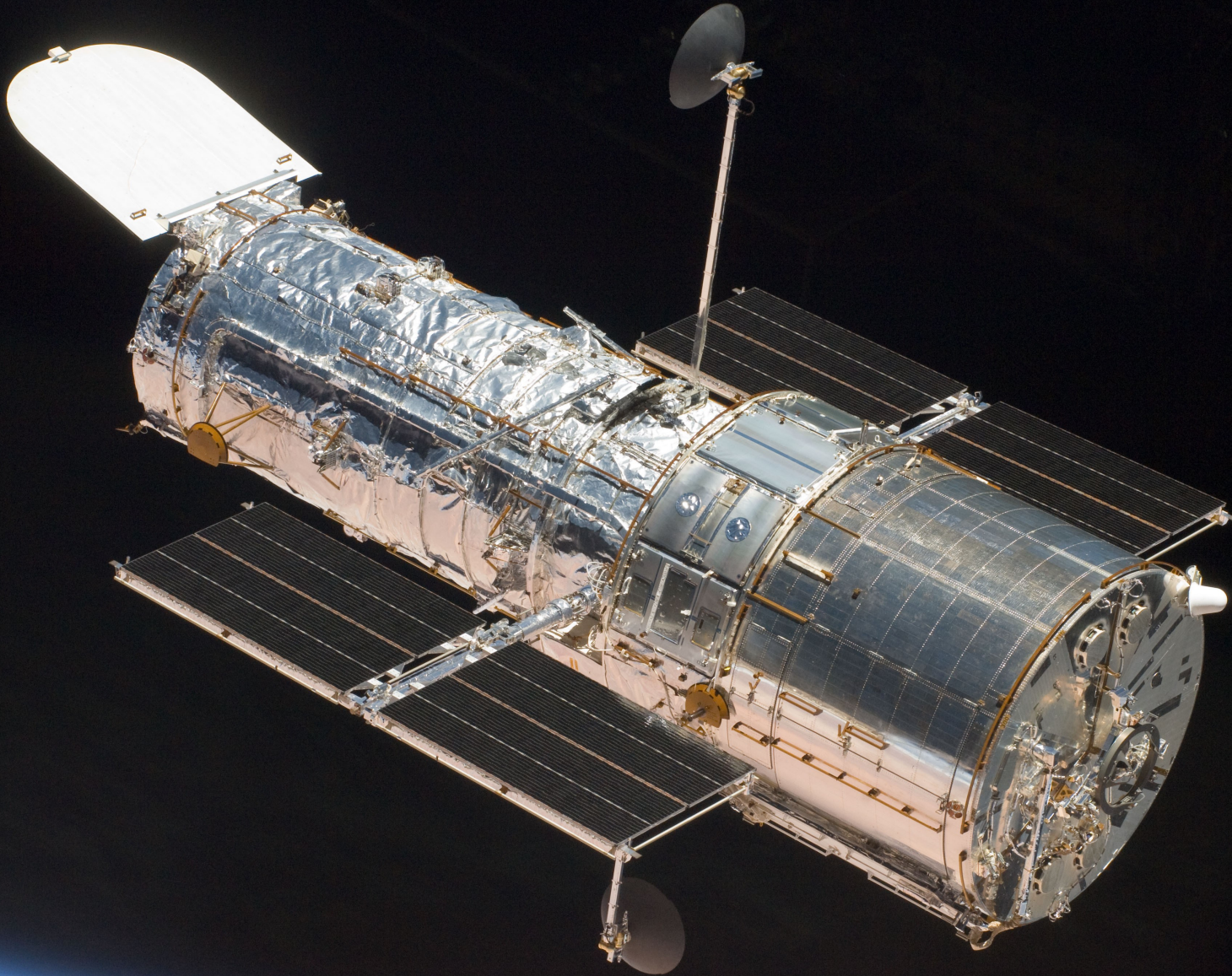
Государство/бизнес

- Делопроизводство.
- Экономические планы/отчёты/анализ по регионам, глобально, по отдельным отраслям/продуктам.
- Статистика.
- Комплексное управление бизнесом.
- Демография, социальные явления.
- Специальные разделы.

Финансовые и родственные приложения

- Выполнение традиционных банковских операций.
- «Электронные деньги»
- Кредитные карты, банкоматы, кассовые аппараты.
- Учёт движения товаров.
- Автоматизированные складские помещения.

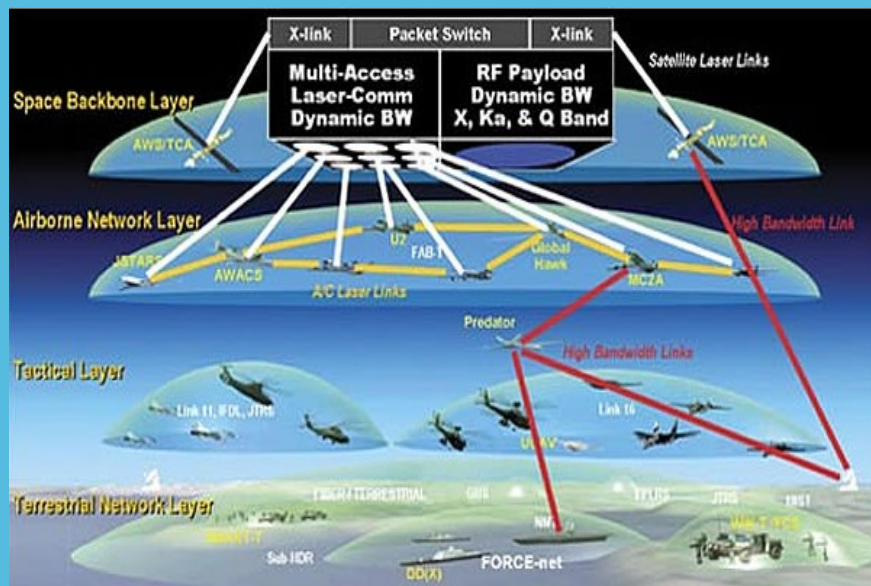
Исследования галактик: телескоп Хаббл (Hubble)



10 Dec 2011

Andrey.Shevel@pnpi.spb.ru

Военные приложения



10 Dec 2011

Andrey.Shevel@pnpi.spb.ru

Программы «электронное правительство»

- **Москва** - <http://www.mos.ru/>
- **С. Петербург** - <http://www.gov.spb.ru/>
- **Ленинградская область** - <http://www.lenobl.ru/>
- **Государственная власть России** - <http://www.gov.ru/>

Можно ли использовать один большой компьютер для любых (всех) отраслей?

- Различная терминология в разных отраслях
- Разный характер вычислений
- Различные потребности в оборудовании
- Различные условия использования (окружающая среда)
- Разные прикладные программы
- Вывод – каждая область имеет специфические потребности в вычислениях

Charles Babbage (1791 - 1871) –
Изобретатель первого
Программируемого вычислителя

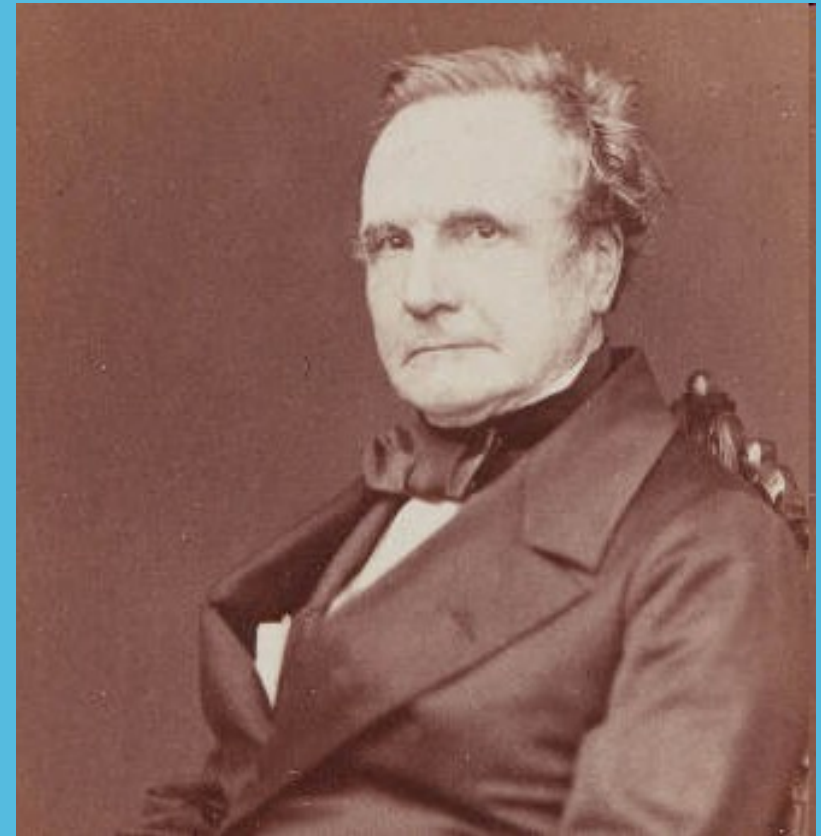


Ada Lovelace (Byron)

1815 - 1852

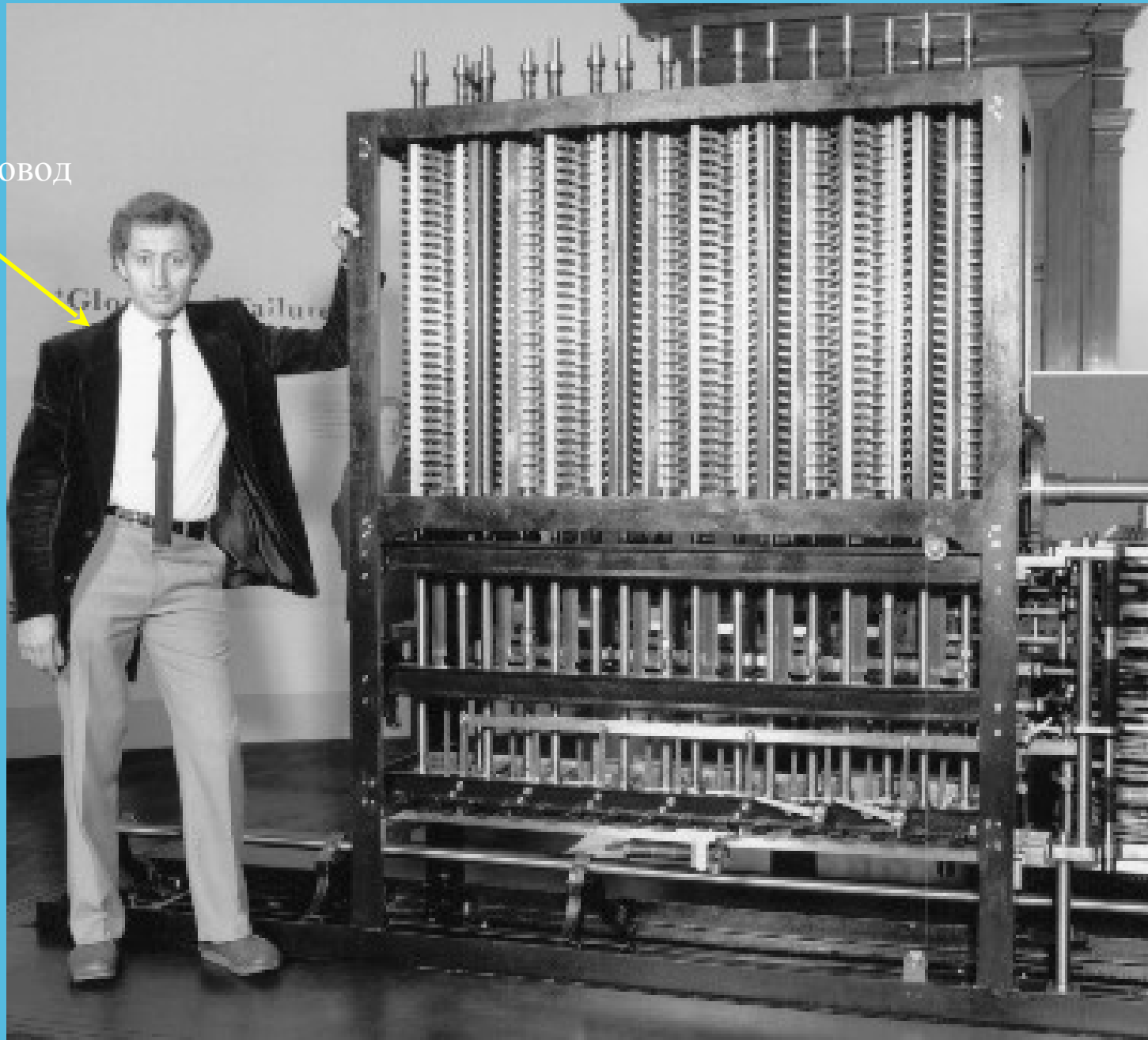
первый программист мира

10 Dec 2011



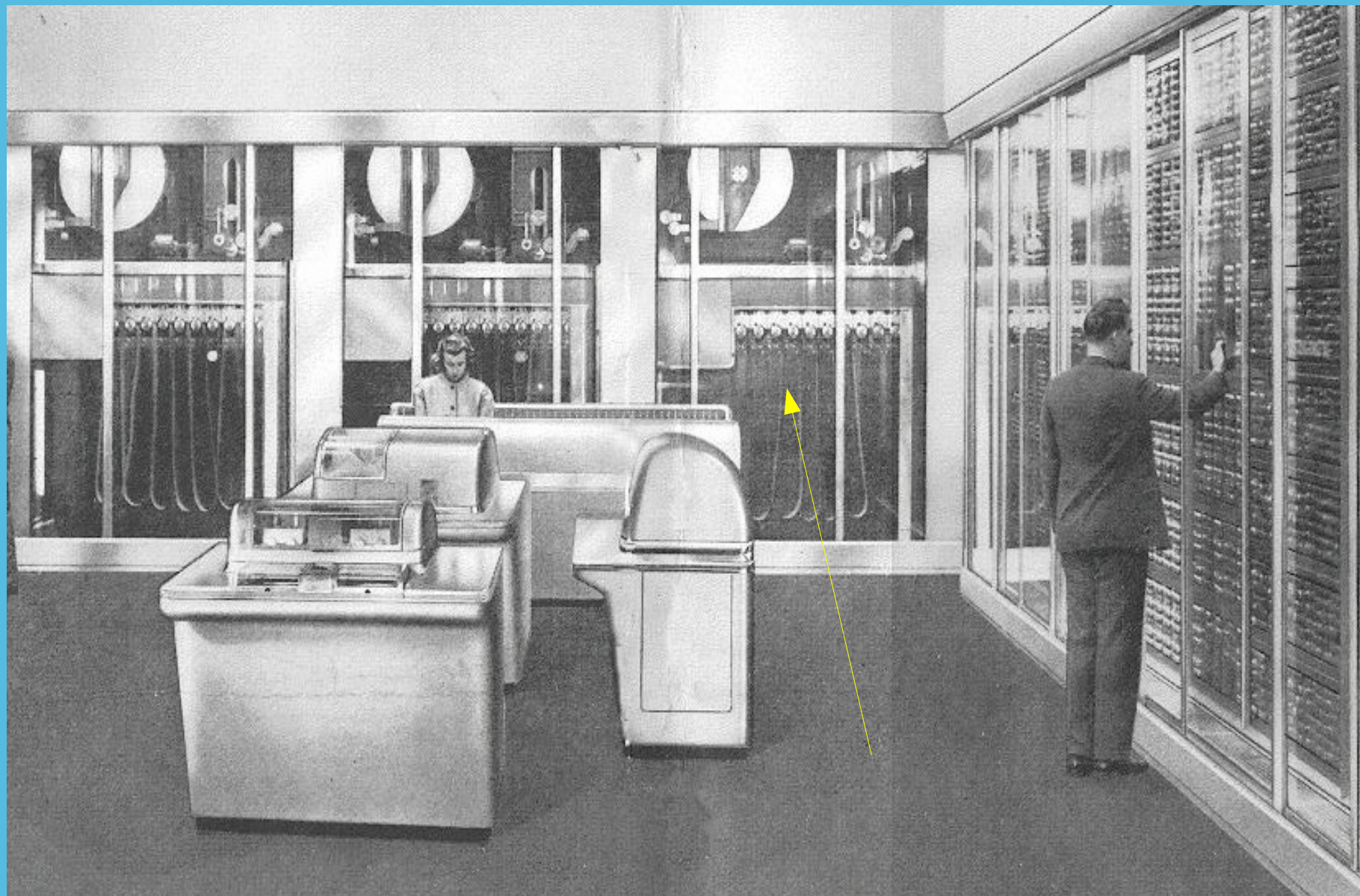
Andrey.Shevel@pnpi.spb.ru

Первый программируемый компьютер



экскурсовод

Первая ЭВМ - Electronic Numerical Integrator And Computer



10 Dec 2011

Andrey.Shevel@npi.spb.ru

Современные вычислительные системы

- Характерной особенностью современных коммунальных компьютерных архитектур является использование гетерогенных/гибридных компьютерных кластеров.
 - Вычислительные системы;
 - Пакетные системы выполнения заданий;
 - Системы памяти;
 - Серверы баз данных, крупные Вэб и мейл серверы. Серверы социальных сетей.

«Пейзаж» крупного кластера



10 Dec 2011

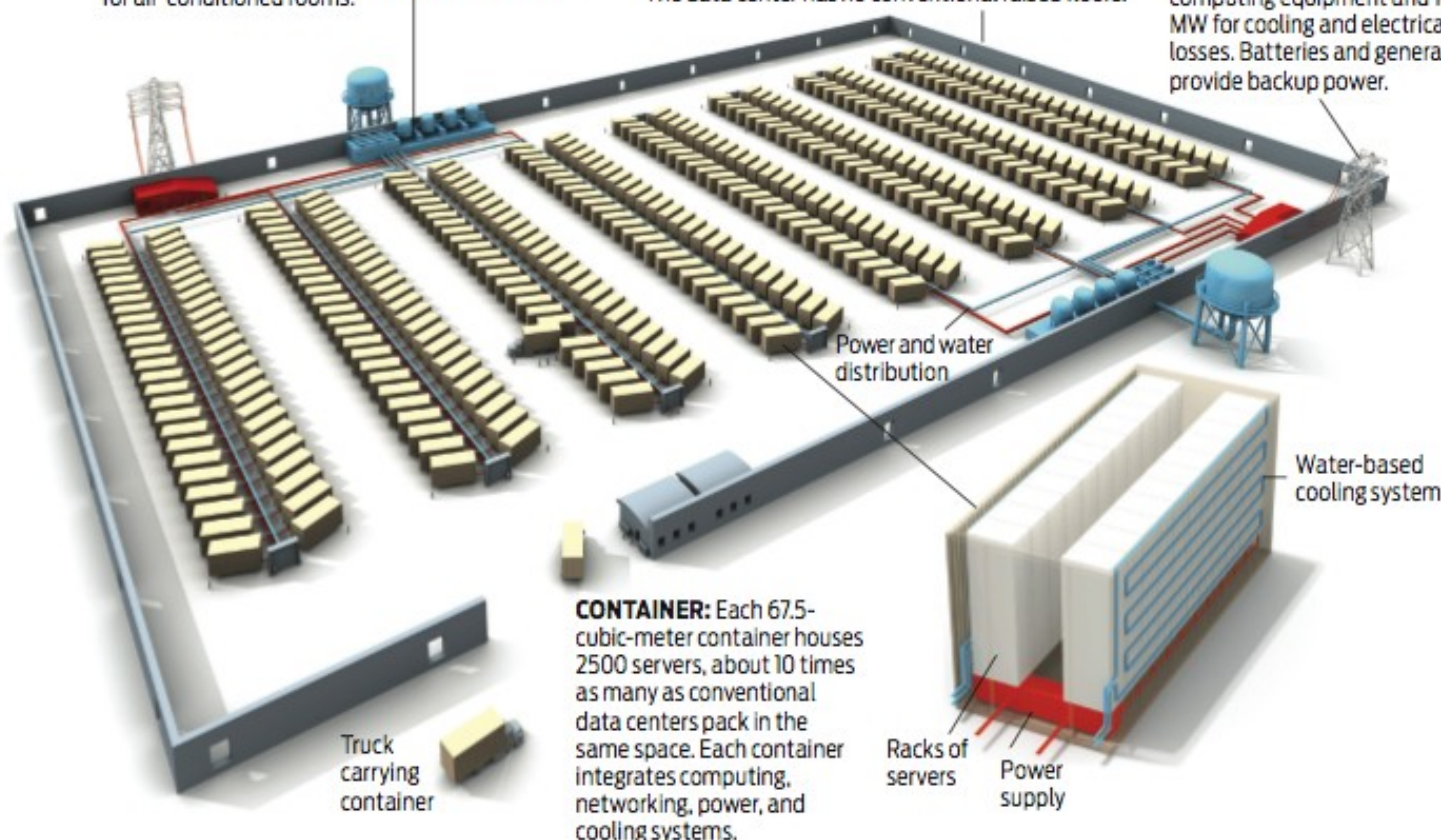
Andrey.Shevel@pnpi.spb.ru

План большого центра данных (кластера)

COOLING: High-efficiency water-based cooling systems—less energy-intensive than traditional chillers—circulate cold water through the containers to remove heat, eliminating the need for air-conditioned rooms.

STRUCTURE: A 24 000-square-meter facility houses 400 containers. Delivered by trucks, the containers attach to a spine infrastructure that feeds network connectivity, power, and water. The data center has no conventional raised floors.

POWER: Two power substations feed a total of 300 megawatts to the data center, with 200 MW used for computing equipment and 100 MW for cooling and electrical losses. Batteries and generators provide backup power.



CONTAINER: Each 67.5-cubic-meter container houses 2500 servers, about 10 times as many as conventional data centers pack in the same space. Each container integrates computing, networking, power, and cooling systems.

Большие центры обработки данных



WAREHOUSE-SIZE COMPUTERS: Google has built a sprawling data center on the banks of the Columbia River, in The Dalles, Ore. The site, with two server-packed buildings and space for a third, houses tens of thousands of computers—the exact number is a closely guarded secret. Microsoft, Yahoo, and Amazon are also building data centers in the region, enticed by its readily available fiber-optic connectivity and cheap electricity. *PHOTO: MELANIE CONNER*

Наиболее обсуждаемые компьютерные проблемы - 2011

- Использование облачных вычислений.
- Многоядерные процессоры.
- Увеличение объёма вычислительной мощности на ватт потребляемой мощности.
 - Важность подтверждается тем, что, например, Google имеет только в США более миллиона установленных серверов. Большой центр обработки данных может потреблять десятки MW и более.
- **Безопасность в сетях** (предотвращение недружественных вторжений, вирусы, спам) **НЕ ТОЛЬКО В кластерах.**
- Дополнительно см. конференцию sc11.supercomputing.org

Крупнейшие компьютерные кластеры мира

- **В мире**

- <http://www.top500.org> (список крупнейших кластеров мира – обновляется дважды в год)

- **В России**

- <http://www.supercomputers.ru/> (список крупнейших кластеров в России и СНГ)

Основные тенденции в программном обеспечении 2011

- **Использование Свободно распространяемого Программного Обеспечения (СПО) [некоммерческого]**
 - Операционные системы Линукс (различные дистрибутивы)
- **Виртуализация**
 - настольных компьютеров;
 - серверов;
 - компьютерной инфраструктуры.

Перспективные вычислительные структуры 2011

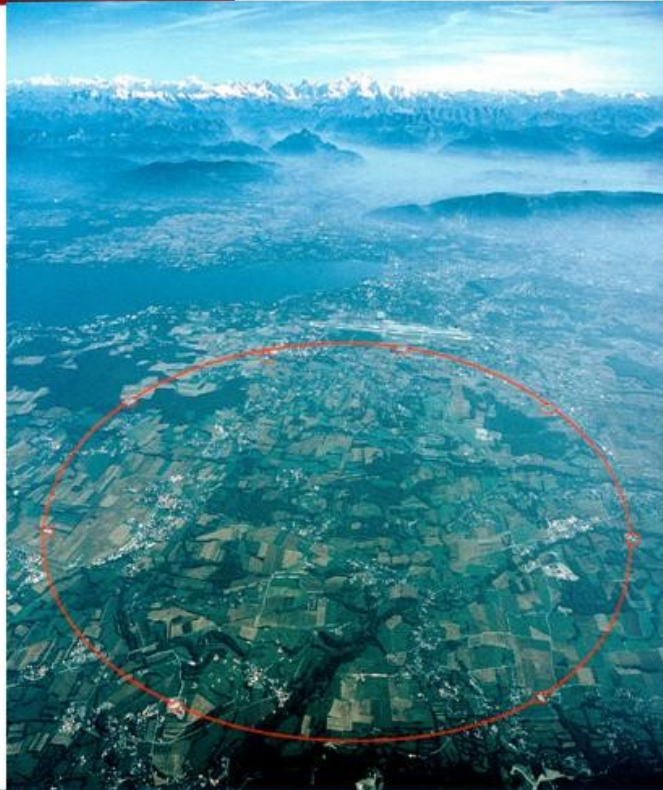
- **Грид (grid) системы (много проектов).** Полезная информации о таких системах на сайтах:
 - <https://sites.google.com/site/clustergateorg>
 - <http://www.gridclub.ru/>
 - <http://parallel.ru/>
- **Облачные вычисления (cloud computing), например, Elastic Cloud Computing** <http://www.amazon.com/ec2/>

Использование Grid для науки EGI

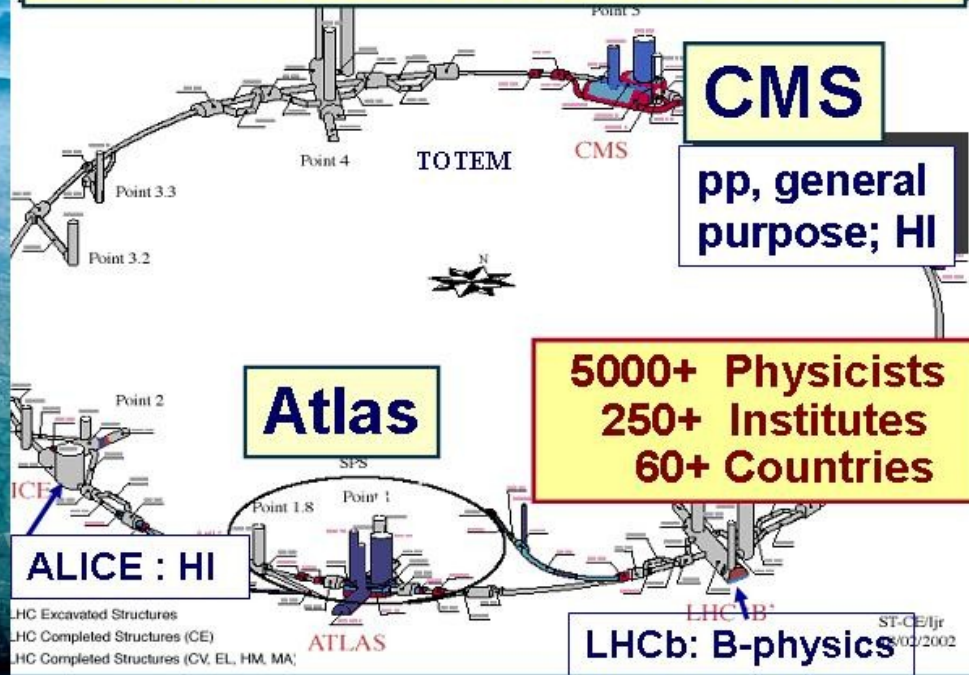
- **European Grid Initiative site –**
- *<http://www.egi.eu/>*
 - Основные области использования
 - Биология
 - Исследование земли и окружающего пространства
 - Физика Высоких Энергий
- **Российский сайт Grid - <http://www.egee-rdig.ru/>**



Large Hadron Collider CERN, Geneva: 2007 Start



* $pp \sqrt{s} = 14 \text{ TeV}$ $L = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
* 27 km Tunnel in Switzerland & France



**Challenges: Analyze petabytes of complex data cooperatively
Harness global computing, data & network resources**

Вид тоннеля ускорителя БАК (LHC)

600 Million Proton Collisions/second

SC Magnets cooled to -193.2°C (80 K),
 -268.7°C (4.5K), -271.3°C (1.9k)
Using 10'000 tons of liquid
nitrogen and 120 tonnes of liquid helium

26659m in Circumference (but varies with the moon!)

Вид внутри трубы



10 Dec 2011

Andrey.Shevel@pnpi.spb.ru



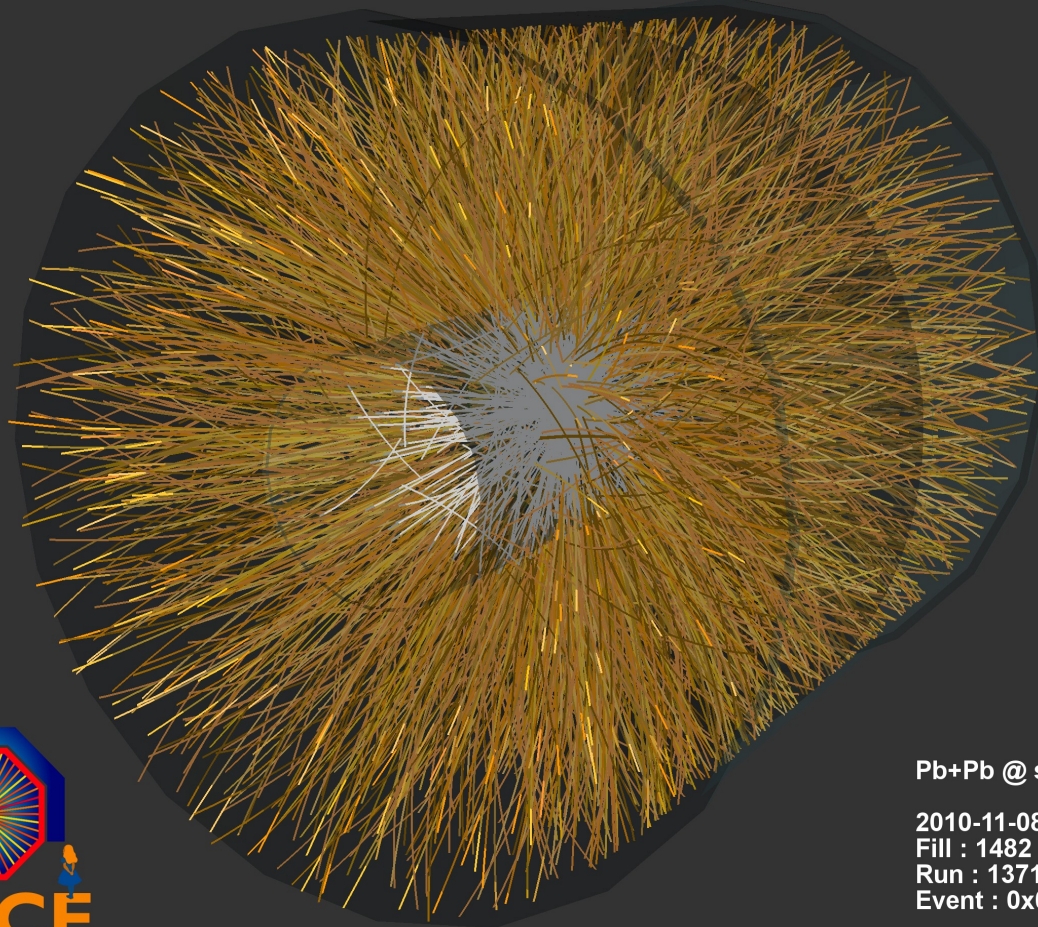
10 Dec 2011

Andrey.Shevel@pnpi.spb.ru

Пример размеров детектора на LHC



Столкновения частиц



Pb+Pb @ $\sqrt{s} = 2.76$ ATeV

2010-11-08 11:29:52

Fill : 1482

Run : 137124

Event : 0x0000000042B1B693

10 Dec 2011

Andrey.Shevel@pnpi.spb.ru

DISUN: CMS Global Data Grid

CMS Experiment



Online System

0.1 - 1.5 GB/s

CERN/Outside Resource Ratio ~1:4

Tier0/(Σ Tier1)/(Σ Tier2) ~1:2:2

Harvey B. Newman

Tier 1



Tier 2

Physics caches across Tier 2

2.5-10 Gb/s



Tier 3



Tier 4



CERN/Outside Ratio Smaller; Expanded Role of Tier1s & Tier2s: Greater Reliance on Networks

4 of 7 US CMS Tier2s Shown With ~8 MSi2k; 1.5 PB Disk by 2007 >100 Tier2s at LHC

Основные типы и вычислений в ФВЭ

- **Online** («в линию») вычисления (определяются условиями измерений в эксперименте)
 - должны быть выполнены максимально быстро (основной критерий - продуктивность для физики). Коллективная работа: программисты, электронщики, физики, студенты.
- **Offline** («не на линии») вычисления (определяются логикой обработки экспериментальных измерений) – это обработка большого объёма данных:
 - часто выполняется с использованием распределённых по миру компьютерных систем;
 - большой объём данных, состоящих из статистически независимых фрагментов данных (**событий**), описывающих отдельные физические события в измерительной установке.
 - подготовкой обработки и обсуждением результатов обработки данных занимается большое количество лиц (десятки или сотни).
 - Результаты компьютерной обработки данных эксперимента – это как правило коллективный труд, а не достижения отдельных персон.
- **Сопутствующие вычисления:** всё остальное, что выполняется на компьютерах в исследовательских учреждениях.

Вычислительный кластер в лаборатории ядерной химии

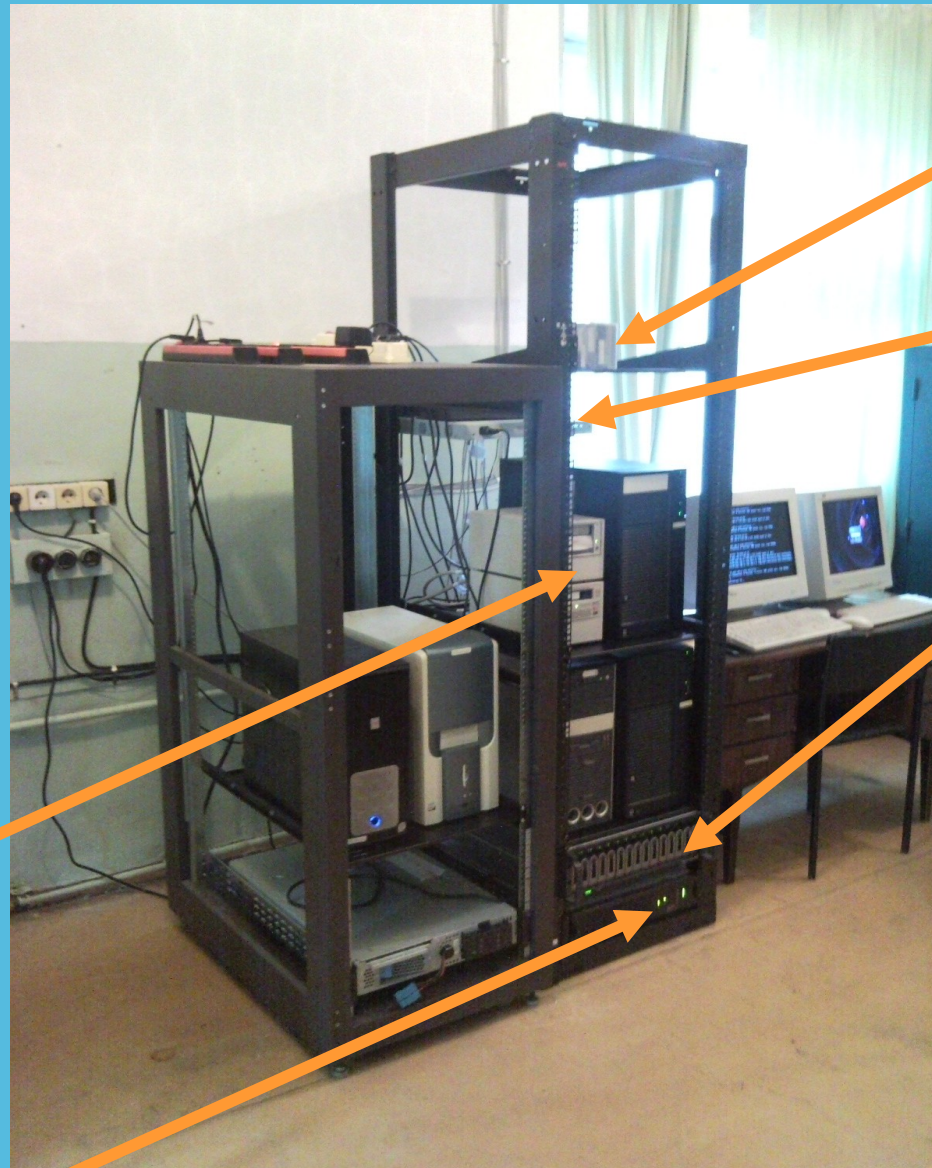
Stony Brook



10 Dec 2011

Andrey.Shevel@pnpi.spb.ru

Внешний вид кластера в Отделении ФВЭ ПИЯФ



KVM

Gbit switch

Disk array

SDTL

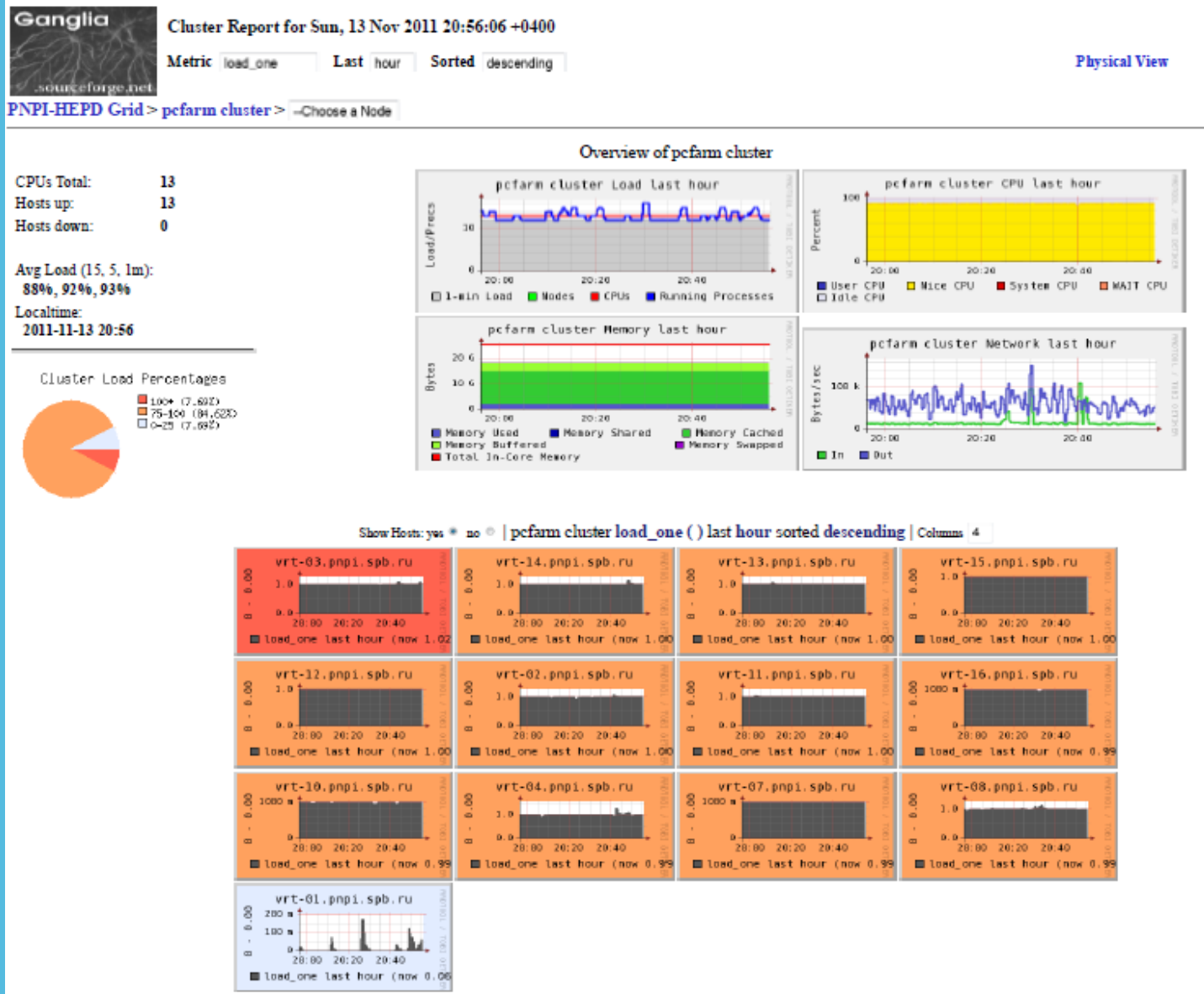
UPS

10 Dec 2011

Andrey.Shevel@pnpi.spb.ru

HEPD PNPI

Состояние кластера HEPD PNPI



10 Dec 2011

Andrey.Shevel@pnpi.spb.ru

Заключение - 1

- Компьютеры используются в любых отраслях человеческой деятельности, но в каждой имеются особенности использования.
- Современные особенности компьютерных архитектур (*всё используется в ядерной физике*)
 - гетерогенные компьютерные кластеры
 - свободное программное обеспечение
 - виртуализация
 - облачные и Грид архитектуры

Заключение - 2

- **Компьютерная инфраструктура экспериментальной ядерной физики/физики высоких энергий характеризуется:**
 - Большим объёмом данных (часто измеряется в ТВ - 10^{12} байтов и даже в PB – 10^{15} байтов)
 - Высокой скоростью поступления измерительных данных (много MB/sec)
 - Большими запросами на вычислительные мощности (много тысяч компьютеров)
- **Два вида компьютерной инфраструктуры**
 - On-line – вычисления во время проведения физических измерений, где вычисления являются неотъемлемой частью физических экспериментов
 - Off-line – вычисления связанные с обработкой измерительных данных и анализом результатов обработки. Такие вычисления могут выполняться независимо от проведения физических экспериментов.

Заключение - 3

- Данные состоят из отдельных фрагментов (могут иметь размер измеряемый в МВ) , которые содержат информацию о физических событиях. Эти фрагменты данных именуются “события”.
- В большинстве экспериментов события являются статистически независимыми, что даёт возможность обрабатывать каждое событие независимо от других. Например, параллельно, или в одно и то же время.

Заключение - 4

- **Обработка данных полученных в эксперименте может длиться продолжительное время (до нескольких месяцев или лет). Может производиться удалённо от места хранения данных.**
- **Большое распространение получает удалённое наблюдения за проведением эксперимента.**

Конец лекции

Можно ли принять участие в моделировании событий на LHC ?

- Ответ - “Да” для любого человека с компьютером, см.

<http://lhathome.web.cern.ch/LHCathome/>