

# **Високопродуктивні розподілені обчислювальні системи**

Основні положення побудови  
учбового курсу

# «Грід-технології для розподілених обчислень та обробки даних»

Згідно з вимогами освітньо-професійної програми студенти після засвоєння навчальної дисципліни мають продемонструвати такі результати навчання:

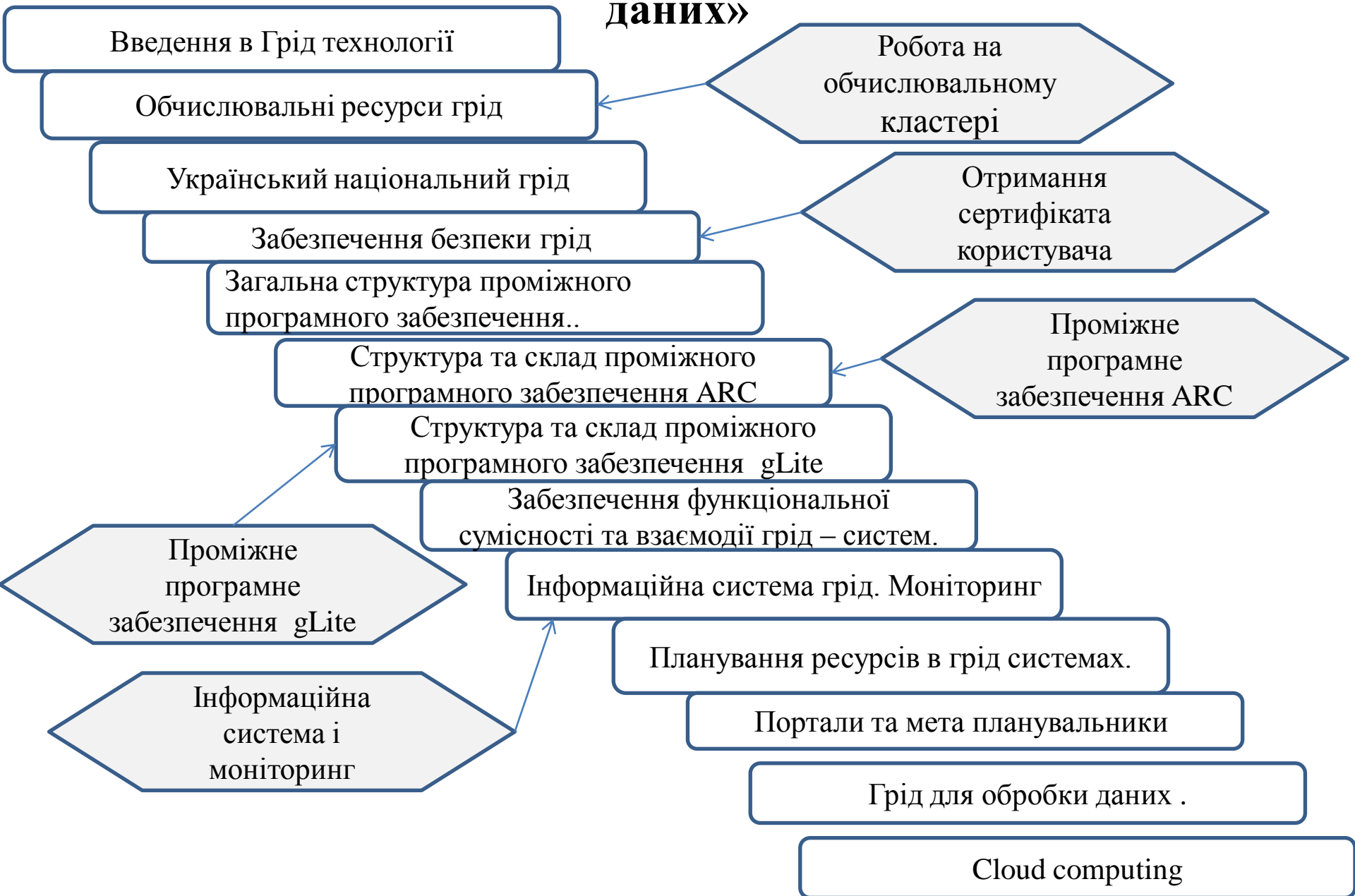
## *знання:*

- **основи Грід – технологій**, які дозволяють об'єднати обчислювальні ресурси та ресурси зберігання даних в єдину систему,
- **архітектуру Грід – систем**, які використовуються в Україні,
- **принципи функціонування основних складових частин Грід – системи**,
- **технологію підготовки завдань для використання Грід – середовища**,

## *вміння:*

- **Проведення керованих обчислень на кластері** та створювати опис завдання і даних для відправки завдання в грід-систему під керуванням проміжного програмного забезпечення ARC та gLite;
- **Відслідковувати та корегувати хід обчислювального процесу в грід-системах** під керуванням проміжного програмного забезпечення gLite, ARC;
- **Проводити інтелектуальну обробку великих масивів даних в грід-системах**;
  - Проводити хмарні обчислення з використанням архітектури Cloud computing.

# «Грід-технології для розподілених обчислень та обробки даних»

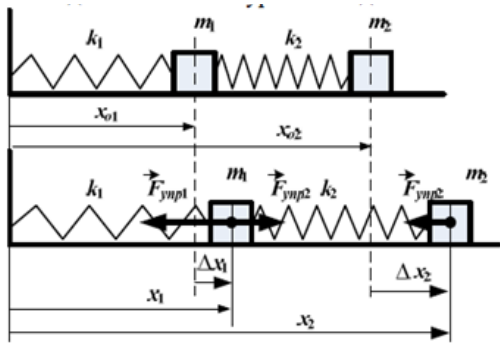


# «Грід-технології для розподілених обчислень та обробки даних»



Отмеченные лекции переделывались каждый семестр...

# «Грід-технології для розподілених обчислень та обробки даних»

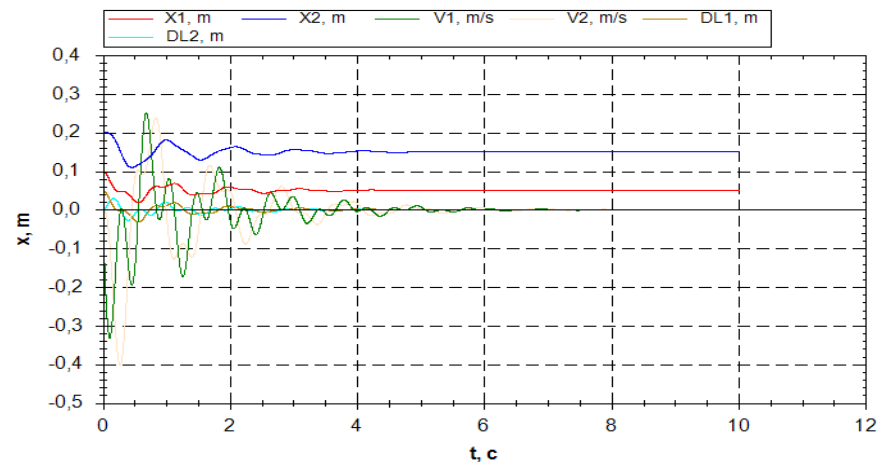


$$\begin{cases} m_1 \frac{d^2 x_1}{dt^2} = k_2((x_2 - x_1) - (x_{o2} - x_{o1})) - k_1(x_1 - x_{o1}), \\ m_2 \frac{d^2 x_2}{dt^2} = -k_2((x_2 - x_1) - (x_{o2} - x_{o1})). \end{cases} \quad (2.1)$$

Здесь:  $(x_2 - x_1) - (x_{o2} - x_{o1}) = \Delta x_2 - \Delta x_1 = \Delta \ell_2$  – удлинение (деформация) второй пружины в любой момент времени,  $(x_1 - x_{o1}) = \Delta x_1 = \Delta \ell_1$  – удлинение первой пружины.

1. Методом численного моделирования (выбирается в соответствии с вариантом) получить вынужденные колебания тела (уравнение 1.1), т.е. рассчитать скорость тела и его координату в соответствующие моменты времени в выбранном временном интервале.

2. Для исходных начальных данных и  $\omega v = 0$  (на тело действует постоянная сила величиной  $F = F_{\max}$ ). Изменяя коэффициент трения от нуля, подобрать такое его значение, начиная с которого колебания прекращаются. По результатам расчетов построить графики колебаний.



**Внедрены новые лабораторные работы, которые охватывали несколько курсов**

# «Грід-технології для розподілених обчислень та обробки даних»

## Выводы

- Даний курс являється спеціалізованим і розглядає тільки частину технології розподілених обчислень.
- За 5 років викладання лекційний матеріал суттєво змінився в відповідності з зміною технології.
- В лекційному курсі потрібно розглядати як загальні питання побудови систем розподілених обчислень, так і детально хмарні обчислення.
- В даний момент більше підходить назва «Хмарні та грід технології для розподілених обчислень та обробки даних»

## Distributed & cloud computing (*Louisiana Tech University*)

### *The Characterization of distributed computing and cloud computing.*

- System Models.
  - **Physical Model** (represents underlying hardware elements of a distributed system that abstracts away from specific details of the computer and networking technologies employed),
  - **Architectural model** defines the way in which the components of the system are placed and how they interact with one another and the way in which they are mapped onto the underlying network of computers,
  - **Fundamental models: Interaction model** (deals with communication details among the components and their timing and performance details), **Failure model** (gives specification of faults and defines reliable communication and correct processes), **Security model** (specifies possible threats and defines the concept of secure channels).
- Networking and inter-process communication.
- OS supports and Virtualization.
  - Clients servers architecture.
  - Peer-to-peer architecture.
  - Middleware and network operating systems
  - Different Virtualization Concepts (Xen, VMware, Qemu, KVM)
- RAS, Performance & Reliability Modeling. Security.

# Distributed & cloud computing (*Louisiana Tech University*)

## *Introduction to Cloud Computing*

- Various models and applications.
- Deployment models
- Service models (SaaS, PaaS, IaaS, XaaS)
- Public Cloud: Amazon AWS
- Private Cloud: OpenStack.
- Cost models ( between cloud vs host your own)

## Case studies

- Introduction to HPC
- Multicore & OpenMP
- Many core, GPGPU & CUDA
- Cloud-based Electrocardiography system (Medical Application)
- Distributed Object & Web Services
- Distributed File system
- Replication & Disaster Recovery



# DISTRIBUTED SYSTEMS (*Engineering at Illinois*)

- Introduction
  - *Introduction to Cloud Computing*
  - *MapReduce and Hadoop*
- Failure Detectors
- Time and Ordering
- Multicast Communications
- P2P Systems
- Leader Election (Replications)
- Networking and Routing
- RPCs and Marshalling
- Replication Control,
- Key-value Stores/NoSQL
- Stream Processing and Graph Processing in Clouds
- Measurements and Characteristics of Real Distributed Systems
- Sensor Networks, Security
- Distributed Shared Memory
- Self-Stabilization
- Distributed File Systems
- Datacenter Disasters - Case Studies

# **DISTRIBUTED AND HIGH-PERFORMANCE COMPUTING**

## **(Department of Computer Science University of Adelaide)**

- **Introduction to Distributed and High-Performance Computing:**
- **High-Performance Computing architectures** (Overview of the major classes of HPC architectures and their evolution; MIMD machines; cluster computing and Beowulf PC clusters)
- **Parallel programming models and performance analysis** (Parameterization, modelling, performance analysis, efficiency, and benchmarking of DHPC systems. )
- **Programming parallel computers** (Overview of parallel programming, parallel languages, parallelizing compilers, message passing and data parallel programming models)
- **Message passing programming and MPI** (programming using the Message Passing Interface )
- **Data parallel programming and HPF** (Data parallel programming paradigm; High Performance Fortran (HPF))
- **Shared memory programming, threads and OpenMP** (Use of shared memory machines; threads; mutual exclusion; locks, semaphores and monitors; parallel Java; programming using OpenMP).
- **High-performance distributed computing** (Distributed and concurrent computing on loosely coupled distributed systems; motivation and applications; transparency and metacomputing; networks and clusters of workstations; cluster management systems.)
- **Distributed computing middleware** (Middleware; RPC; DCE; CORBA; Java RMI.)
- **Grid computing** (The Grid; Grid computing (metacomputing over wide-area networks); grid computing environments)
- **Cloud computing** (architectures, Middleware, privat and public Cloud)
- **Causality and time in distributed systems**( The problem of time and ordering in distributed systems; vector time; fault tolerance: Byzantine Generals problem)

# DISTRIBUTED AND HIGH-PERFORMANCE COMPUTING

*(Department of Computer Science University of Adelaide)*

**Focus in particular on the following DHPC technologies:**

- **Data Parallelism** ( Uses; Current systems and work in progress (HPF, HPC, HPJava...)
- **Message Passing** ( Uses; Current systems and research. (PVM, MPI,...)
- **Shared Memory Systems** ( Architectures; Multi-threading; OpenMP).
- **Distributed Environments** ( Uses; Current systems and developments (DCE, CORBA, Java RMI...).
- **Cluster Computing** (Beowulf PC clusters, networks of workstations, network technologies, cluster management software)
- **Metacomputing or Grid Computing Environments** (Globus).

The course is assessed as follows:

- **Practical programming exercise** (choice of several) : 40%
- **Research a topic (choice of several) and write a report** : 20%

## **Некоторые выводы по анализу аналогичных лекционных курсов**

- Нужно определить основные акценты тематики курса (общее описание технологий, технологии построения, технологии программирования, пользовательский подход ).
- Нужно согласовать тематику курса с тематикой дисциплин, которые уже читаются.
- Нужно определить практическую часть курса (лабораторные работы, расчетно – графическая работа, самостоятельная работа)

## Високопродуктивні розподілені обчислювальні системи

По результатам анализа имеющейся литературы и по анализу аналогичных учебных курсов можно выделить три фундаментальных блока, которые составляют основу курса:

Сучасні архітектури  
високопродуктивних розподілених  
обчислювальних систем

```
graph TD; A[Сучасні архітектури високопродуктивних розподілених обчислювальних систем] --> B[Технології та програмне забезпечення]; B --> C[Використання та управління ресурсами у розподілених системах.]
```

Технології та програмне забезпечення

Використання та управління ресурсами  
у розподілених системах.

# Високопродуктивні розподілені обчислювальні системи

Сучасні архітектури високопродуктивних розподілених обчислювальних систем

Кластер

Организация вычислений в кластерной системе

Клиент - серверная

Организация распределенных вычислений с использованием серверов приложений

Организация распределенных вычислений а грид

Распределенная система

Организация распределенных вычислений с использованием облачной инфраструктуры

# Високопродуктивні розподілені обчислювальні системи

Технології та програмне забезпечення

Кластер

Клиент - сервер

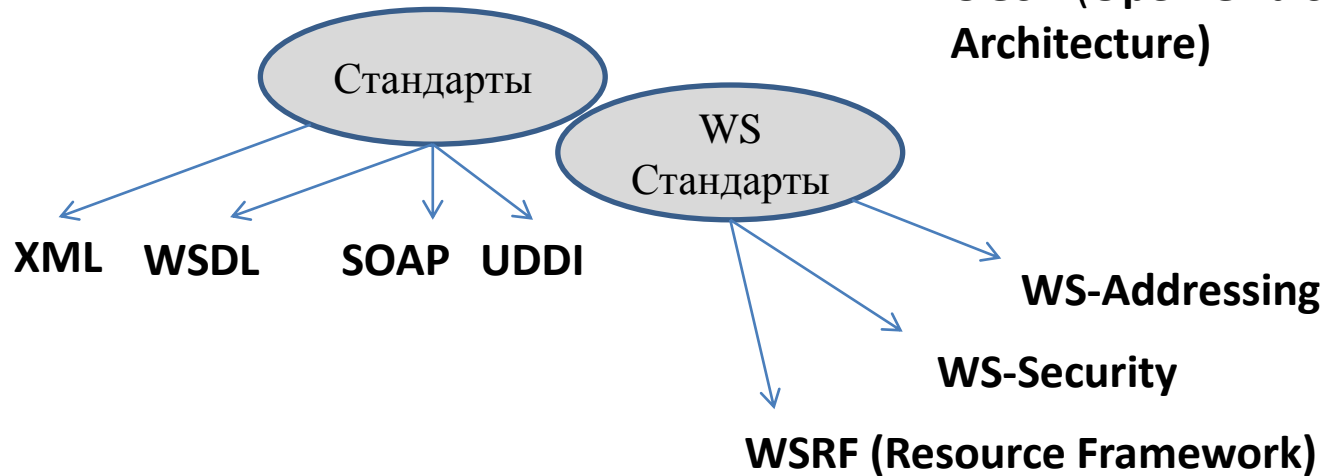
Распределенная система

Сервис-ориентированная архитектура

Программное обеспечение построения кластерной системы

*Веб-сервисы*

OGSA (Open Grid Services Architecture)



# Високопродуктивні розподілені обчислювальні системи

Технології та програмне забезпечення

Кластер

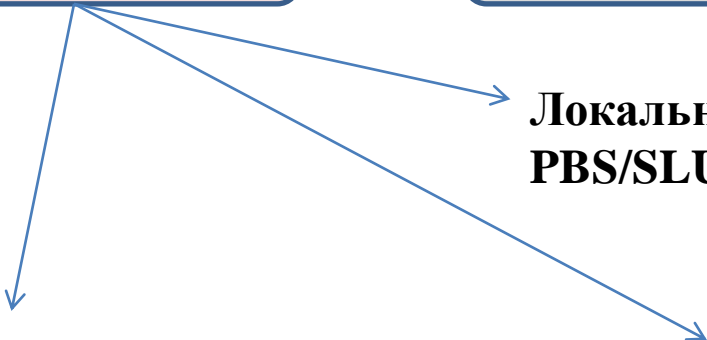
Клиент - сервер

Распределенная  
система

Локальные системы управления ресурсами  
**PBS/SLURM**

Библиотеки для организации  
параллельных вычислений на  
кластере : MPI, PVM, OpenMP

Методи та засоби сучасних  
бібліотек та мов програмування





# Високопродуктивні розподілені обчислювальні системи

Технології та програмне забезпечення

Кластер

Клиент - сервер

Распределенная  
система

Сервер приложений

IBM - WebSphere Application Server

Microsoft - .NET Framework

Oracle WebLogic Server (WLS)

Red Hat - JBoss Application Server

Java Platform, Enterprise Edition

Remote Procedure Call — RPC

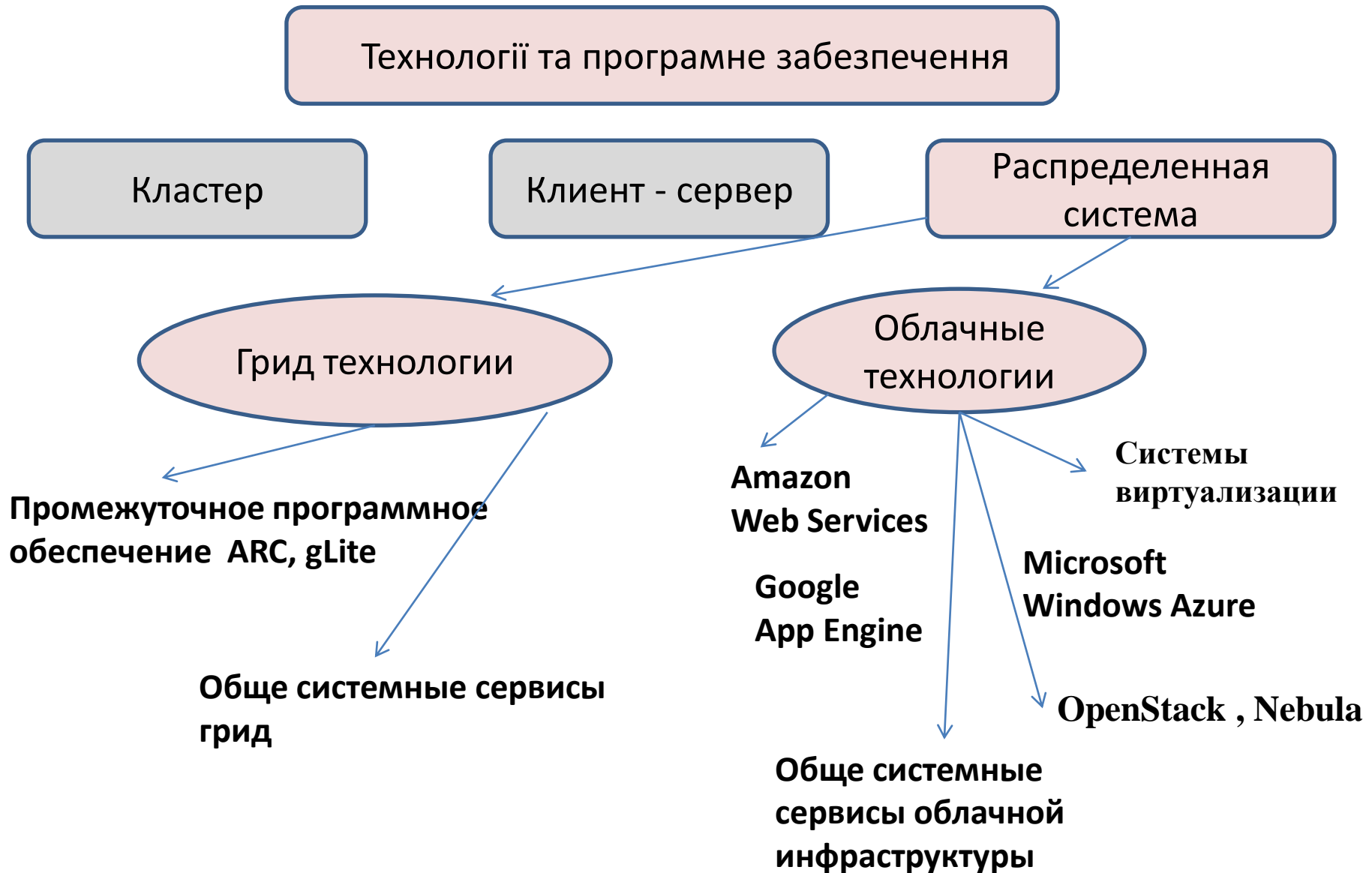
CORBA (Common Object Request Broker  
Architecture)

COM (Component Object Model)

DCOM (Distributed COM)

Enterprise JavaBean

# Високопродуктивні розподілені обчислювальні системи



# Високопродуктивні розподілені обчислювальні системи

**Використання та управління ресурсами у розподілених системах.**

```
graph TD; A[Використання та управління ресурсами у розподілених системах.] --> B[Обеспечение безопасности в распределенных системах. Сертификаты, виртуальные организации]; A --> C[Сервисы управления потоком задач в распределенной системе]; A --> D[Взаимодействие общесистемных сервисов грид и облачной структуры]; A --> E[Методы оцінювання станів процесів у просторово розподілених об'єктах]; A --> F[Сервисы мониторинга];
```

**Обеспечение безопасности в распределенных системах. Сертификаты, виртуальные организации**

**Сервисы управления потоком задач в распределенной системе**

**Взаимодействие общесистемных сервисов грид и облачной структуры**

**Методы оцінювання станів процесів у просторово розподілених об'єктах**

**Сервисы мониторинга**

# Високопродуктивні розподілені обчислювальні системи

**Сучасні архітектури високопродуктивних розподілених обчислювальних систем**

Комп'ютерне моделювання складних об'єктів

Архітектура обчислювальних систем

Параллельное программирование

**Технології та програмне забезпечення**

Базы данных

Веб-технології

Інтелектуальних аналіз даних

Клиент серверная архитектура

**Використання та управління ресурсами у розподілених системах.**

Грід технології для розподілених обчислень

Основи хмарних обчислень

Хмарні технології та сервіси

Отмечены смежные курсы, которые тесно связаны

# **Високопродуктивні розподілені обчислювальні системи**

## **Согласно расписанию учебной нагрузки**

- Курс *«Високопродуктивні розподілені обчислювальні системи»* читается на 5 курсе и состоит: 9 лекций (18 часов), 36 часов лабораторных работ, 1 расчетно-графическая работа, контрольные работы, экзамен.

**Каково содержание практической работы?**

# Високопродуктивні розподілені обчислювальні системи

## Выводы

- Курс в основном по своим разделам дублируется существующими курсами;
- По содержанию – в основном описательный, за 9 лекций изложить связанный материал не возможно;
- Практическая работа студентов требует подготовки вычислительной базы;
- Чисто описательный курс студенты, к сожалению, слушать не будут

# **Високопродуктивні розподілені обчислювальні системи**

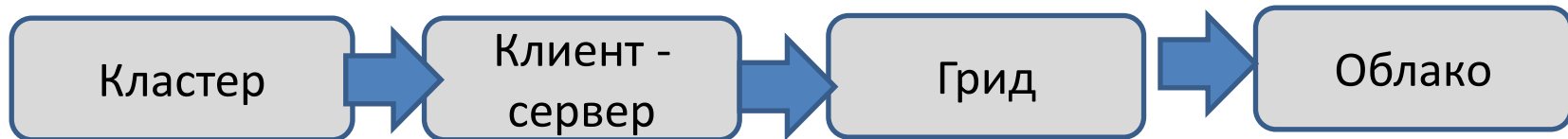
## **Основные требования к курсу:**

- Должен представлять собой целостный материал иллюстрирующий все технологии высокопродуктивных распределенных вычислений;**
- Должен содержать практическую работу студентов иллюстрирующую реальное применение изложенных знаний на практике;**
- Не должен повторять и дублировать уже имеющиеся учебные курсы;**
- Должен представлять современное состояние высокопродуктивных распределенных вычислений**

# Високопродуктивні розподілені обчислювальні системи

## Что предлагается

- Разбить весь курс на 4 модуля.



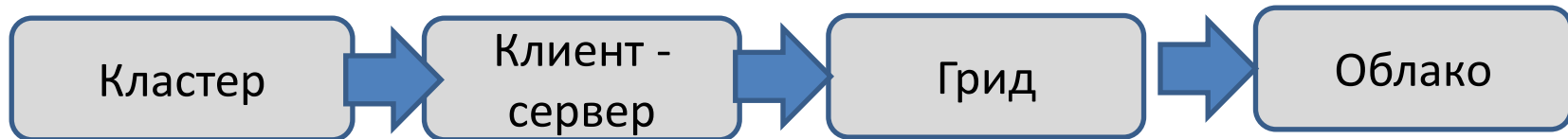
- Условно связывая с применяемыми технологиями
- Лекционный материал охватывает краткий обзор архитектуры и технологий по блоку.
- Для самостоятельной работы дается более детальное рассмотрение.



# Високопродуктивні розподілені обчислювальні системи

## Что предлагается

- Разбить весь курс на 4 модуля.

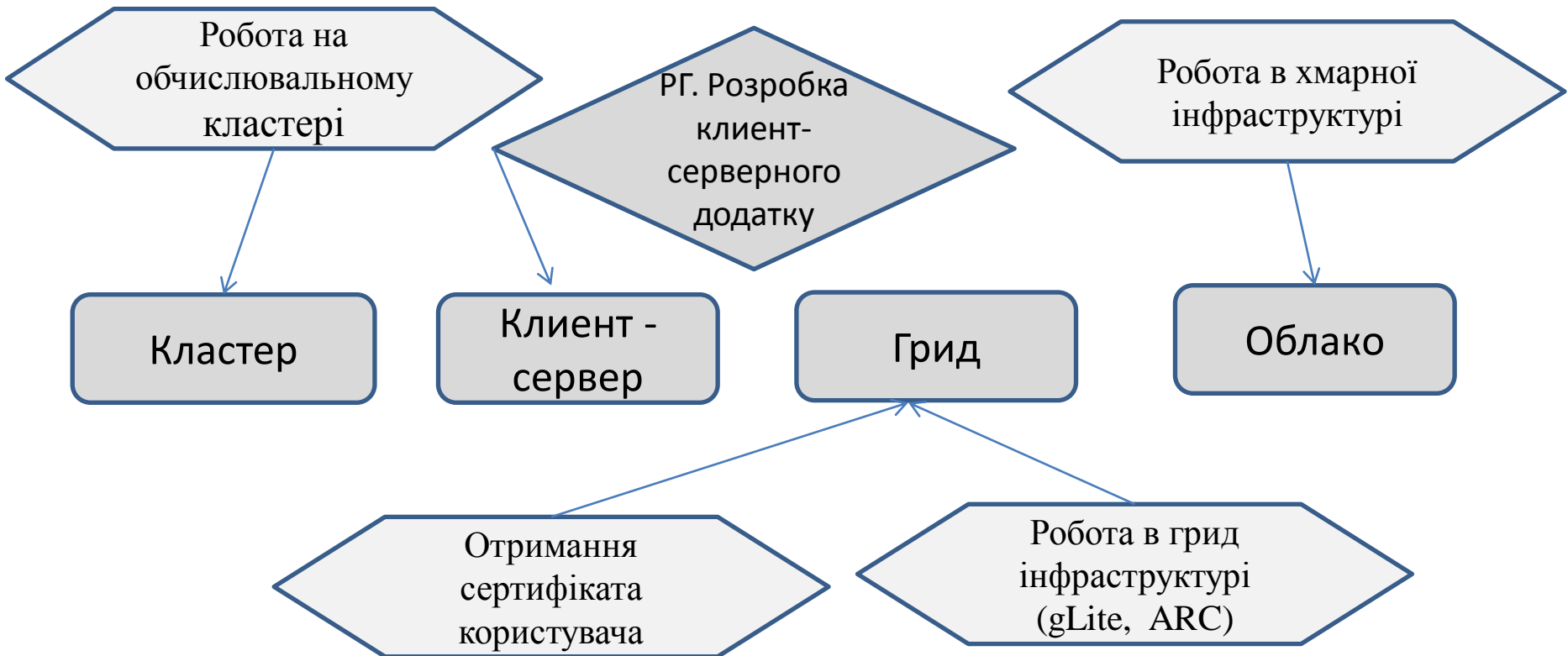


- Учитывая посещаемость лекций студентами старших курсов, предлагается «начитку» лекционного материала сделать блоком (4-6 часов) в начале месяца.
- После лекций студенту дается перечень вопросов, которые войдут в контрольную работу в конце месяца.
- Консультации в течении месяца – дистанционно.

# Високопродуктивні розподілені обчислювальні системи

## Что предлагается

- Практическая работа ориентирована на тематику блока



- Содержание практических работ нужно будет разрабатывать заново.

# Високопродуктивні розподілені обчислювальні системи

## Что нужно сделать

- Согласовать тематический план дисциплины со всеми учебными программами, которые уже читаются для исключения повторения материала.
- Разработать «рабочий план» дисциплины с перечнем лекций и кратким их содержанием.
- Разработать новые лабораторные работы для обеспечения практики.
- Подготовить вычислительную инфраструктуру для выполнения лабораторных работ и расчетно- графической работы.

# Високопродуктивні розподілені обчислювальні системи

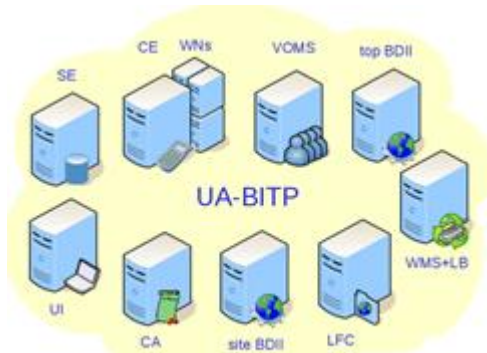
## Самое сложное из перечисленных работ

- Подготовить вычислительную инфраструктуру для выполнения лабораторных работ и расчетно-графической работы.

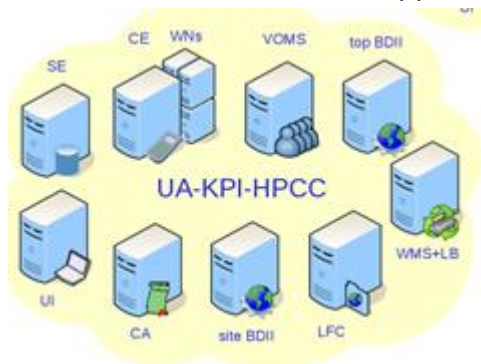
Инсталлирован учебный грид кластер ИТФ. Идет работа студентов ТЭФ

**Нет структуры для разработки клиент – серверных приложений**

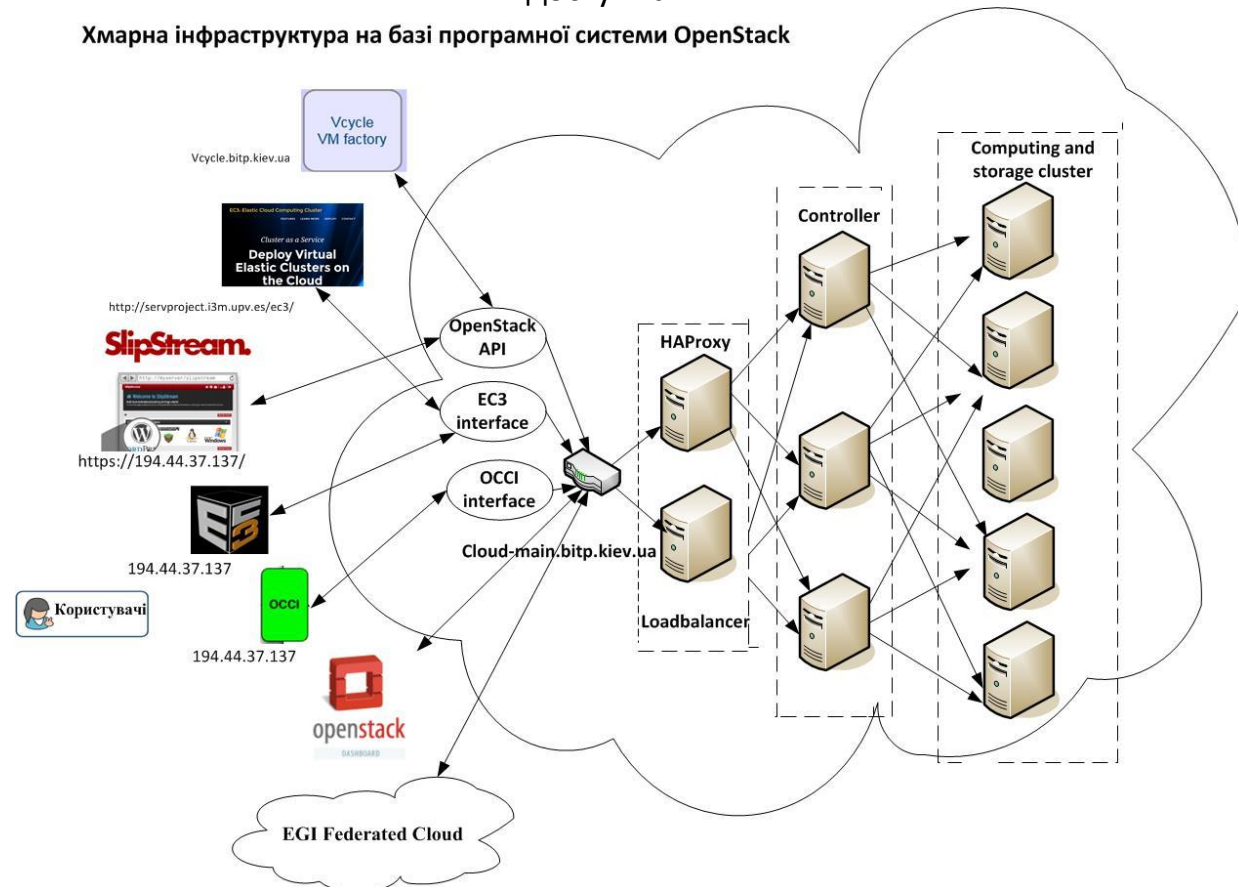
Инсталлирована промышленная облачная инфраструктура, но для студентов пока не доступна



Учебный грид кластер КПИ не работает с начала 2016 года



Хмарна інфраструктура на базі програмної системи OpenStack

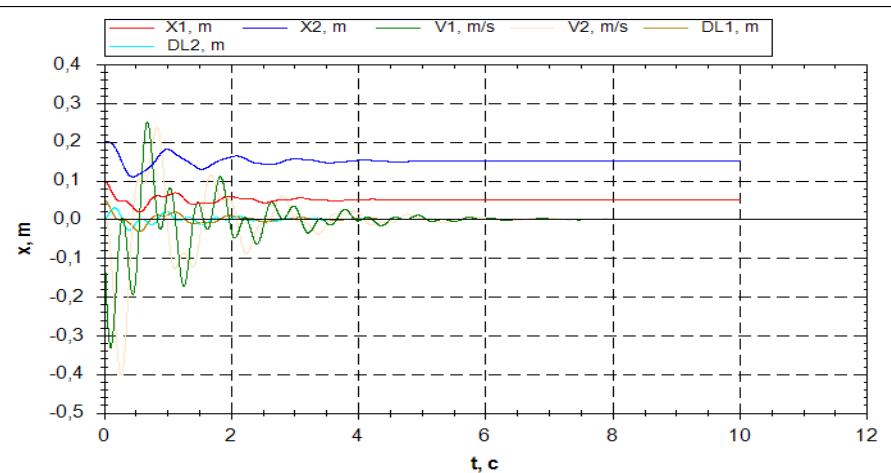
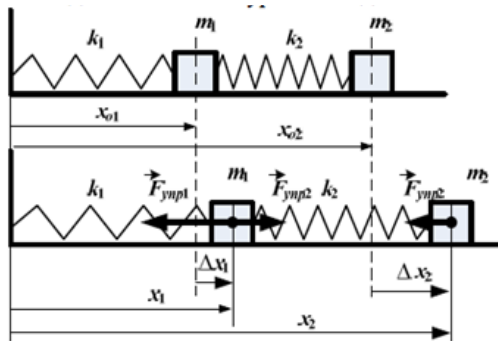


# Високопродуктивні розподілені обчислювальні системи

## Самое сложное из перечисленных работ

- Подготовить вычислительную инфраструктуру для выполнения лабораторных работ и расчетно-графической работы.

**Необходимо разработать новые задания для лабораторных работ, которые иллюстрируют применение технологий высокопроизводительных вычисление на практике**



**Задания должны показывать необходимость применения той или иной технологии для эффективного решения**

**Задания должны иллюстрировать преимущества и недостатки применяемой технологии**

**Спасибо за внимание!**

**Вопросы?**