МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ

ЗАТВЕРДЖЕНО
на засіданні кафедри
інформаційних систем
Протокол № 1 від 22.08.2023 р.

ПОГОДЖЕНО
Проректор з навчально-методичної роботи

_________________ Каріна НЕМАШКАЛО

РОЗПОДІЛЕНІ ТА ПАРАЛЕЛЬНІ ОБЧИСЛЕННЯ

робоча програма навчальної дисципліни (РПНД)

Галузь знань	**12 "Інформаційні технології"**
Спеціальність	**121 "Інженерія програмного забезпечення"**
Освітній рівень	**перший (бакалаврський)**
Освітня програма	**"Інженерія програмного забезпечення "**

Статус дисципліни	**обов’язкова**
Мова викладання, навчання та оцінювання	**англійська**

Розробник:
д.т.н., професор підписано КЕП Сергій МІНУХІН

Завідувач кафедри
інформаційних систем _________________________ Дмитро БОНДАРЕНКО

Гарант програми _________________________ Олег ФРОЛОВ

Харків
2024

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
SIMON KUZNETS KHARKIV NATIONAL UNIVERSITY OF ECONOMICS

APPROVED
at the meeting of the department
information systems
Protocol № 1 of 22.08.2023

AGREED
Vice-rector for educational and methodical work

_________________ Karina NEMASHKALO

DISTRIBUTED AND PARALLEL COMPUTING

Program of the course

Field of knowledge	**12 "Information Technology"**
Specialty	**121 "Software engineering"**
Study cycle	**first (bachelor)**
Study programme	**"Software engineering"**

Course status	**Mandatory**
Language	**English**

Developers:
Doctor (Technical sciences), Professor — digital signature — Serhii MINUKHIN

Head of Information systems department:
Ph.D. (Technical sciences), associate professor _________________ Dmytro BONDARENKO

Head of Study Programme:
Ph.D. (Technical sciences), associate professor _________________ Oleg FROLOV

Kharkiv
2024

INTRODUCTION

The conditions for the growth of data volumes and the increase in the dependence of business processes of enterprises on data flows determine the need for the creation of distributed information systems (DIS) of various levels, which must ensure a sufficient level of data processing efficiency in the conditions of system scalability and increased data processing intensity. Such tasks are solved on the basis of the development of DIS using appropriate technologies and software in accordance with the standards and architectures of distributed computing systems (DCS) using parallel programming.

The technologies of distributed systems and parallel computing are the basis of the construction of distributed IS - from the level of the computing cluster to the level of domains of grid systems and services of cloud computing platforms. The principles of construction, methods and technologies of creation, deployment of DCS and the use of parallel computing are the basis for solving computationally complex and time-consuming tasks in various subject areas.

The course. "Distributed and parallel computing" is studied by students of specialty 121 "Software engineering" of all forms of education in the 4th year during the seventh semester. Studying the course "Distributed and parallel computing" involves acquiring theoretical knowledge and mastering practical skills related to the use of technologies for performing labor-intensive computing processes. The course is aimed at the formation of students of the general foundations for understanding the essence of parallel and distributed computing, mastering the basic concepts of programming within the paradigms of parallel and distributed computing.

The purpose of teaching the course "Distributed and parallel computing" is to provide higher education students with a system of special knowledge and the acquisition of practical skills and abilities in the use of DCS technologies, installation and configuration of appropriate software for launching and performing tasks on a computing cluster and use of parallel programming technologies and tools according to existing standards.

The tasks of the course. are:

- familiarization with the basic principles of construction and information processing technologies in DCS;
- familiarization with the standards of distributed computing systems and software for configuring resource and task management systems in DCS;
- installation and configuration of computer cluster software, obtaining practical work skills for creating and running programs (applications) on the cluster;
- familiarization with the main paradigms of parallel programming in high-performance computing systems - DCS (Grid) and computing cluster levels;
- study of parallel programming technologies - OpenMP and MPI and their implementations when performing practical tasks;
- mastering the practical skills of using parallel programming to solve scientific and engineering tasks.

The subject of the course. is standards, technologies and methods of development and implementation of distributed and parallel computing technologies for solving scientific and engineering tasks.

The object of the course. is computational processes in distributed systems using parallel programming to improve the quality of management of enterprises and institutions based on modern information technologies for the performance of labor-intensive tasks.

The learning outcomes and competence formed by the course, are defined in table. 1.

Table 1

Learning outcomes and competencies formed by the course

Learning outcomes	Competencies
LO 06	SK 3, SK 11
LO 07	SK 11, SK 13, SK 14
LO 13	SK 3, SK 14
LO 18	SK 3
LO 25	GK 1, GK 2, SK 13, SK 15

where, LO 06. The ability to choose and use a software development methodology appropriate to the task.

LO 07. To know and apply in practice the fundamental concepts, paradigms and basic principles of the functioning of language, instrumental and computing tools of software engineering.

LO 13. To know and apply methods of developing algorithms, designing software and data and knowledge structures.

LO 18. To know and be able to apply information technologies for processing, storing and transmitting data.

LO 25. Have knowledge and skills in software development using technologies of distributed data processing and standards of parallel computing on cluster computing systems.

GK 1. The ability to think abstractly, analyze and synthesize.

GK 2. The ability to apply knowledge in practical situations.

SK 3. Ability to develop architectures, modules and components of software systems.

SK 11. The ability to implement phases and iterations of the life cycle of software systems and information technologies based on appropriate software development models and approaches.

SK 13. The ability to reasonably choose and master software development and maintenance tools.

SK 14. The ability to think algorithmically and logically.

SK 15. The ability to use technologies and tools of distributed data processing and parallel computing in software development.

COURSE CONTENT

Content module 1. Technologies of distributed systems.

Topic 1. The concept and classification of distributed computing systems (DCS). Classification of Grid Systems. Composition and purpose of levels of Open Grid Architecture.

1.1. Basic terms and definitions. Classification of distributed computing systems (DSC). Concepts and types of Grid Systems. Multicluster systems and computing clusters. Basic terms and definitions. Areas of application of the latest information technologies in systems of collective use of information systems (IS) resources.

1.2. Composition and assignment of levels of open architecture of distributed computing systems (DCS). The composition and functions of the 5-level DCS architecture: application level, collective level, connectivity level, access level, hardware level. Trends in the development of modern architectures of distributed computing systems.

Topic 2. Principles of organization of data processing in Distributed Systems. The concept and composition of DCS middleware (on the example of a Grid System).

2.1. Organization of data processing in DCS. Types of grid resources: computing resources, data storage, information resources (directories), network resources.

2.2. The concept and purpose of DCS middleware levels.

Topic 3. OGSA architecture for DCS. Concept and classification of resource management systems. Task schedulers. Local resource management systems.

3.1. OGSA architecture. OGSA is a standard for supporting Open Grid Services Infrastructure (OGSI) agreements. Types of services. Basic standard interfaces of Grid Services.

3.2. Concept and classification of types and levels of task planners in DCS.

3.3. Concept and classification of types and levels of resource management systems in DCS.

Topic 4. Information services and systems of DCS. Designation and construction of Information Systems In DCS. Organization of information systems based on R-GMA and MDS technologies.

4.1. Concept of DCS information service. Models of data and access to DCS information services. Tasks of the information service and information systems of DCS. Information service as a basis for implementation of the "supplier-user" paradigm.

4.2. Organization of the data model based on the R-GMA relational model. Peculiarities of implementing the MDS data model in the Globus Toolkit system.

4.3. Types of information services. Semantic information service.

Topic 5. Principles of work and organization of work of monitoring systems in DCS.

5.1. Peculiarities of solving the tasks of monitoring the status of nodes, tasks and services in ROS.

5.2. Tasks of the monitoring system. ROS monitoring based on Nagios, Icinga, Ganglia.

Content module 2. Parallel Computing Technologies.

Topic 6. The concept and classification of parallel computing systems (PCS). Multiprocessor and multicomputer systems. Flynn's classification. Principles of the organization of work of PCS.

6.1. PCS classifications.

6.2. Paradigms of programming in PCS. Task parallelism and data parallelism.

6.3. Principles of construction and organization of PCS with shared and distributed memory.

6.4. Types and main characteristics of topologies of data transmission environments in PCS.

Topic 7. Models of parallel programming.

7.1. Graph models of parallel programming. The concept and representation of a graph for the implementation of a parallel program. Informational and algorithmic dependencies.

7.2. Stages of developing a parallel program (algorithm). Multi-threaded applications.

7.3. A paradigm of message passing between program fragments and results. Multithreading as a means of implementing a parallel program on multiprocessor and multicore architectures.

Topic 8. Execution of multi-threaded programs in the OS.

8.1. Concept of process and flow. Management of threads and processes in the OS.

8.2. Means of data exchange in POS: message transfer in systems with distributed memory. Shared variables in shared memory systems.

Topic 9. Parallel Programming based on OpenMP.

9.1. Functional debugging of the OpenMP program.

9.2. Basic concepts: directives, functions and clauses. The concept of a structural block. The parallel directive. Environment variables that control the execution of OpenMP programs. Classes of variables.

9.3. Directives for defining a parallel domain. The for directive, sections directive, single directive, workshare directive.

9.4. Directives for the distribution of calculations within a parallel area: directives for, sections, single. OpenMP function library. Compilation of the OpenMP program.

Topic 10. Parallel Programming based on MPI.

10.1. The essence of message passing: processes interact by sending and receiving messages.

10.2. Standards MPI 1.1, MPI 2.0, MPI 2.1, MPI 2.2, MPI 3.0.

10.3. Collective communication functions: synchronization (barrier). Collective interactions. Types of communicators. Block parallel algorithms, implementation features using MPI functions. Functions supporting distributed operations: performing global operations with returns.

10.4. Application of OpenMPI for solving scientific and engineering problems.

The list of laboratory studies in the course is given in table 2.

Table 2

The list of laboratory studies

Name of the topic and/or task	Content
Topic 1-3, 6. Laboratory work No. 1	Creation and configuration of computing cluster nodes.
Topic 1-3, 5. Laboratory work No. 2	Configuring the SLURM system on a computing cluster
Topic 1-4, 6. Laboratory work No. 3	Study of the construction of a computing cluster and principles of working with it.
Topic 7-9. Laboratory work No. 4	Parallel programming in the OpenMP open standard and its application in solving a system of linear algebraic equations (SLAE).
Topic 7, 8, 10. Laboratory work No. 5	Parallel programming in the MPI standard.

The list of self-studies in the course is given in table 3.

Table 3

List of self-studies

Name of the topic and/or task	Content
Topic 1 – 10	Studying lecture material
Topic 1 – 10	Preparation for laboratory classes
Topic 1 – 10	Preparation for the exam

The number of hours of lectures, laboratory classes and hours of self-study are given in the technological card for the course.

TEACHING METHODS

In the process of teaching the course, in order to acquire certain learning outcomes, to activate the educational process, it is envisaged to use such teaching methods as:

Verbal (lecture-discussion (Topic 1-10), problematic lecture (Topic 7, 8), provocative lecture (Topic 3).

Visual (demonstration (Topic 1 – 10)).

Laboratory work (Topic 1–10).

FORMS AND METHODS OF ASSESSMENT

The University uses a 100-point cumulative system for assessing the learning outcomes of students.

Current control is carried out during lectures, laboratory classes is aimed at checking the level of readiness of the student to perform a specific job and is evaluated by the amount of points scored:

– for courses with a form of semester control as an exam: maximum amount is 60 points; minimum amount required is 35 points.

The final control includes current control and an exam.

Semester control is carried out in the form of a semester exam.

The final grade in the course *is determined:*

– for the course's with a form of exam, the final grade is the amount of all points received during the current control and the exam grade.

During the teaching of the course, the following control measures are used:

Current control:

defense of **laboratory work** (60 points);

Semester control: Grading including **Exam** (40 points)

An example of an exam card and assessment criteria.

An example of an examination paper

Simon Kuznets Kharkiv National University of Economics
First (bachelor) level of higher education
Specialty "Software Engineering"
Educational and professional program "Software engineering"
Semester 7 (6)
Educational the course. "DISTRIBUTED AND PARALLEL COMPUTING"

EXAM CARD

Task 1 (diagnostic, 10 points).
List and characterize the composition of Open Grid Service Architecture services. State its differences in comparison with the 5-level architecture of distributed computing systems.
Task 2 (heuristic, 15 points).
To implement the solution of the problem based on the OpenMP open standard: create two matrices of dimensions 60 x 60 each. Fill them with random numbers. Output the values of the obtained matrices. Compile the obtained matrices with the help of 3 sections. Output the number of threads and the final matrix.
Task 3 (heuristic, 15 points).
Implement solutions based on OpenMPI technology:
create two one-dimensional arrays of dimension 30 in processes 0 and 3, respectively. Fill them with random numbers. Output process numbers and received arrays. Transfer the values of processes 0 and 3 to process 1. In process 1, calculate their scalar product and return the result of process 0. Output the process number and the result.

Protocol No. __1__ dated "_23__"__08___2023_ was approved at the meeting of the Department of Information Systems

Examiner Professor Minukhin S.

Chief of Department PhD, Associate Professor Bondarenko D.

Assessment criteria

The final marks for the exam consist of the sum of the marks for the completion of all tasks, rounded to a whole number according to the rules of mathematics.

The algorithm for solving each task includes separate stages that differ in complexity, time-consumingness, and importance for solving the task. Therefore, individual tasks and stages of their solution are evaluated separately from each other as follows.

Task 1.

This task is evaluated on a 10-point scale.

A score of 10 points is given if the acquirer has given the composition of services of the Open Grid Service Architecture in full and compared it with the 5-level architecture in accordance with the specified features.

A score of 9 points is given if the acquirer provides the full range of services of the Open Grid Service Architecture. However, the answer has certain inaccuracies in defining the differences compared to the 5-level architecture of ROS.

An assessment of 8 points is given if the acquirer does not provide the full range of services of the Open Grid Service Architecture. A comparison was made with the features of the 5-level architecture of ROS, but there are inaccuracies regarding their justification.

A score of 7 points is given if the acquirer does not provide the full range of services of the Open Grid Service Architecture. The composition of the levels of the 5-level architecture is not fully characterized.

A score of 6 points is given if the acquirer does not provide the complete list of services of the Open Grid Service Architecture in full and with errors. The composition of the levels of the 5-level architecture of ROS has not been fully characterized.

A score of 5 points is given if the acquirer does not provide the composition of the services of the Open Grid Service Architecture in full and with errors. The levels of the 5-level architecture of ROS have been characterized incompletely and with inaccuracies.

A score of 4 points is given if the acquirer has given the composition of Open Grid Service Architecture services with errors. The levels of the 5-level architecture of ROS are given incompletely and with errors.

A score of 3 points is given if the acquirer provides the composition of Open Grid Service Architecture services with errors. There is a significant number of errors and inaccuracies in the description of the 5-level architecture of ROS.

An assessment of 2 points is given if the acquirer has incorrectly specified the composition of services of the Open Grid Service Architecture. There is a significant number of errors and inaccuracies in the description of the 5-level architecture of ROS.

A score of 1 point is given if the acquirer has incorrectly specified the composition of services of the Open Grid Service Architecture. There is an incorrect description of the levels of the 5-level architecture of ROS.

A score of 0 points is given for failure to complete the task in general.

Task 2.

This task is evaluated on a 15-point scale.

For a completely correctly completed task 1, the applicant receives an assessment of 15 points, which consists of the following:

if there is a program code for solving the problem without parallelization in OpenMP, the winner receives 3 (three) points;

if there are parallelization directives and functions of the OpenMP open standard in the program, according to the tasks specified, the winner receives 3 (three) points;

if the program compilation results are available on the cluster, the applicant receives 3 (three) points;

if there is a script for running the task in batch mode on the cluster according to the conditions of the task, the winner receives 1 (one) point;

if there are correct results of the task performance based on the use of the OpenMP open standard in the form of a report from the resource manager, the applicant receives 5 (five) points.

In the absence of fulfillment of certain points, the assessment for this task is reduced by the specified number of points for the corresponding point.

Task 3.

This task is evaluated on a 15-point scale.

For a completely correctly completed task 3, the applicant receives an assessment of 15 points, which consists of the following:

if there is a program code for solving the problem without parallelization in MPI, the winner receives 3 (three) points;

in the presence of MPI functions for parallelization and data transfer according to the conditions of the task, the winner receives 3 (three) points;

if the program compilation results are available on the cluster, the applicant receives 3 (three) points;

if there is a script for running the task in batch mode on the cluster according to the conditions of the task, the winner receives 1 (one) point;

if there are correct results of the task performance based on the use of MPI technology in the form of a report from the resource manager, the applicant receives 5 (five) points.

In the absence of fulfillment of certain points, the assessment for this task is reduced by the specified number of points for the corresponding point.

RECOMMENDED LITERATURE

Main

1. Малашонок Г. І., Сідько А. А. Паралельні обчислення на розподіленій пам’яті: OpenMPI, Java, Math Partner : підручник. / Г. І. Малашонок, А. А. Сідько. – Київ : НаУКМА, 2020. – 266 с.

2. Минайленко Р. М. Паралельні та розподілені обчислення: навч. посіб. – Кропивницький : Видавець Лисенко В. Ф., 2021. – 153 с.

3. Коцовський В. М. Теорія паралельних обчислень: навчальний посібник. / В. М. Коцовський. – Ужгород: ПП «АУТДОР-Шарк», 2021. – 188 с. http://195.230.140.114/jspui/bitstream/123456789/10630/1/Par_rozp_obch_2021.pdf.

4. Комп’ютерні мережі. Лабораторний практикум для студентів спеціальності 122 "Комп’ютерні науки" освітньої програми "Комп’ютерні науки" першого (бакалаврського) рівня [Електронний ресурс] / уклад. С. В. Мінухін, В. П. Коцюба, Ю. В. Савін; Харківський національний економічний університет ім. С. Кузнеця. – Електрон. текстові дан. (9,03 МБ). — Харків : ХНЕУ ім. С. Кузнеця, 2023. – 121 с. – Загол. з титул. екрану. http://www.repository.hneu.edu.ua/handle/123456789/30794.

5. Наконечна О. А., Ярмоленко Т. А., Алексеєнко В. В., Якимчук Б. М. Інструктивно-методичні рекомендації з дисципліни «Технології розподілених систем та паралельних обчислень» / уклад.: Оксана Наконечна, Тетяна Ярмоленко, Вікторія Алексеєнко, Богданна Якимчук. Житомир: Вид-во ЖДУ ім. Івана Франка, 2023. – 74 с.

6. Інформатика в сфері комунікацій [Електронний ресурс] : навч.-практ. посіб. : у 3-х ч. Ч. 3 : Використання web-технологій у сфері комунікацій / С. Г. Удовенко, В. А. Затхей, О. В. Гороховатський [та ін.] ; за заг. ред. С. Г. Удовенка; Харківський національний економічний університет ім. С. Кузнеця. - Електрон. текстові дан. (10.5 МБ). - Харків : ХНЕУ ім. С. Кузнеця, 2020. - 154 с. : іл. - Загол. з титул. екрану. - Бібліогр.: с. 153. http://www.repository.hneu.edu.ua/handle/123456789/24506.

Additional

7. Сучасні інформаційні технології та системи [Електронний ресурс] : монографія / Н. Г. Аксак, Л. Е. Гризун, С. В. Мінухін [та ін.] ; за заг. ред. Пономаренка В. С. – Харків : ХНЕУ ім. С. Кузнеця, 2022. – 270 с. http://www.repository.hneu.edu.ua/handle/123456789/29233.

8. Голубничий, Д. Ю. Операційні системи [Електронний ресурс] : навч. посіб. / Д. Ю. Голубничий, А. В. Холодкова ; Харківський національний економічний університет ім. С. Кузнеця. - Електрон. текстові дан. (12,5 МБ). - Харків : ХНЕУ ім. С. Кузнеця, 2018. - 316 с. : іл. - Загол. з титул. екрану. - Бібліогр.: с. 313-315. http://www.repository.hneu.edu.ua/handle/123456789/23844.

9. Klemm M., Cownie J. High Performance Parallel Runtimes: Design and Implementation. – Walter de Gruyter GmbH & Co KG, 2021.

Information resources

10. The site of the personal educational system for the educational the course. "Distributed and parallel computing" - The access mode : https://pns.hneu.edu.ua/course/view.php?id=3833.

11. The OpenMP API specification for parallel programming. – The access mode : http://openmp.org/wp/ openmp-specifications/.

12. Parallel Programming in MPI and OpenMP. – The access mode : https://web.corral.tacc.utexas.edu/CompEdu/pdf/pcse/EijkhoutParallelProgramming.pdf.

13. Eijkhout Victor. Parallel Programming in MPI and OpenMP. 2nd edition 2022. – The access mode : https://web.corral.tacc.utexas.edu/CompEdu/pdf/pcse/EijkhoutParallelProgramming.pdf.