Syllabus
Course Program

Models and methods of building distributed databases

Specialty
172 – Electronic communications and radio engineering

Institute
Institute of Computer Modeling, Applied Physics and Mathematics

Educational program
Network technologies and telecommunications

Department
Information systems named after V.O. Kravets (169)

Level of education
Master's level

Course type
Optional

Semester
2

Language of instruction
English

Lecturers and course developers

Vitaliy Breslavets

Vitalii.Breslavets@khpi.edu.ua
Candidate of technical sciences, professor, professor of the department of information systems named after V.O. Kravets of NTU "KhPI"

Author and co-author of more than 75 scientific and methodical publications
Courses: "Electromagnetic compatibility", "Scientific-Research work", "Information systems and databases".

More about the lecturer on the department's website

General information

Summary

This course explores the foundational concepts, models, and methods for designing, implementing, and managing distributed databases in modern telecommunication systems. Students will learn about data distribution, consistency models, replication strategies, fault tolerance, and query optimization. The course emphasizes practical applications, equipping learners with the knowledge and skills necessary to develop efficient and scalable distributed database systems..

Course objectives and goals

The objective of this course is to provide students with an in-depth understanding of the principles and techniques used in building distributed databases. By covering topics such as data partitioning, replication, and synchronization, the course aims to prepare students to design and implement robust database systems that ensure reliability, scalability, and performance. Students will also learn to apply distributed database concepts to real-world scenarios in telecommunication systems..

Format of classes

Lectures, laboratory classes, consultations, self-study. Final control in the form of an exam.

Competencies

GC1. Ability to abstract thinking, analysis and synthesis.
SC5. Ability to develop, improve and use modern software, hardware and software-hardware support for electronic communication and radio devices.

Learning outcomes

LO8. apply general-purpose and specialized programming languages, analytical and simulation modeling packages, as well as software and hardware development tools to solve complex problems in electronic communications and radio engineering.

Student workload

The total volume of the course is 120 hours (4 ECTS credits): lectures - 32, laboratory classes - 0 hours, workshops - 32 hours, self-study - 56 hours.

Course prerequisites

"Algorithmization and programming", "Programming", "Discrete Mathematics"

Features of the course, teaching and learning methods, and technologies

Classes are made interactively using multimedia technologies for lecture presentations and online demonstrations of task execution examples. The lecture classes use explanatory-illustrative, reproductive, problem-oriented methods and the method of critical thinking. Training materials are available for students through OneNote (Class Notebook).

Program of the course

Topics of the lectures

Topic 1. Introduction to Distributed Databases:
Overview of distributed databases, their role in telecommunications, and key challenges.
Topic 2. Distributed Database Architectures:
Examining client-server, peer-to-peer, and hybrid database architectures.
Topic 3. Data Partitioning and Distribution Strategies:
Techniques for splitting and distributing data across multiple nodes.
Topic 4. Data Replication and Consistency Models:
Exploring replication strategies and consistency models such as eventual and strong consistency.
Topic 5. Query Processing in Distributed Databases:
Techniques for optimizing and executing queries across distributed systems.
Topic 6. Distributed Transaction Management:
Managing transactions, ensuring atomicity, consistency, isolation, and durability (ACID) in a distributed context.
Topic 7. Concurrency Control in Distributed Systems:
Mechanisms for managing concurrent data access in distributed databases.
Topic 8. Fault Tolerance and Recovery:
Strategies to ensure data availability and integrity in the presence of failures.
Topic 9. Distributed Database Security:
Implementing access control, encryption, and other security measures.
Topic 10. Scalability and Performance Optimization:
Techniques to scale distributed databases and optimize performance under high loads.
Topic 11. NoSQL Databases in Distributed Systems:
Introduction to NoSQL databases like Cassandra, MongoDB, and their use cases.
Topic 12. CAP Theorem and Its Implications:
Understanding the CAP theorem and how it influences distributed database design.
Topic 13. Distributed Data Integration:
Methods for integrating heterogeneous data sources into a unified system.
Topic 14. Cloud-Based Distributed Databases:
Exploring distributed databases offered by cloud platforms like AWS, Azure, and Google Cloud.

Topic 15. Case Studies in Distributed Databases:
Analyzing real-world examples of distributed database implementations in telecommunications.
Topic 16. Future Trends in Distributed Database Systems:
Exploring emerging technologies and methodologies shaping the future of distributed databases.

Topics of the workshops

Topic 1. Setting Up a Distributed Database Environment:
Installing and configuring a distributed database system (e.g., Cassandra or MongoDB).
Topic 2. Data Partitioning and Distribution:
Implementing partitioning strategies and distributing data across multiple nodes.
Topic 3. Configuring Data Replication:
Setting up replication and analyzing consistency levels in a distributed environment.
Topic 4. Query Optimization in Distributed Databases:
Designing and testing optimized queries to retrieve data efficiently.
Topic 5. Implementing Transactions in Distributed Databases:
Managing distributed transactions and ensuring ACID compliance.
Topic 6. Concurrency Control Mechanisms:
Implementing and testing mechanisms for managing concurrent data access.
Topic 7. Fault Tolerance Implementation:
Simulating failures and implementing recovery strategies to maintain data integrity.
Topic 8. Securing a Distributed Database:
Applying encryption and access control measures to secure a distributed database system.
Topic 9. Performance Benchmarking:
Measuring and analyzing the performance of a distributed database under various loads.
Topic 10. Exploring NoSQL Databases:
Setting up and using a NoSQL distributed database for specific applications.
Topic 11. CAP Theorem in Practice:
Analyzing trade-offs between consistency, availability, and partition tolerance in a database.
Topic 12. Distributed Data Integration:
Integrating data from multiple sources into a cohesive distributed database system.
Topic 13. Deploying Distributed Databases on Cloud Platforms:
Setting up and managing a distributed database on a cloud service like AWS or Azure.
Topic 14. Monitoring and Debugging Distributed Databases:
Using monitoring tools to identify and resolve issues in distributed database systems.
Topic 15. Implementing Sharding in Databases:
Applying sharding techniques to distribute data across multiple servers.
Topic 16. Final Project – Building a Distributed Database Application:
Designing and implementing a complete distributed database solution for a real-world use case.

Topics of the laboratory classes

Not included

.

Self-study

Students are encouraged to independently explore advanced distributed database concepts, experiment with open-source tools, and work on small projects to reinforce their learning. Reviewing case studies, experimenting with different database architectures, and studying the latest research in distributed systems will help deepen their understanding. By dedicating time to self-study, students will enhance their problem-solving skills and prepare for real-world challenges in distributed database management.

Course materials and recommended reading

1. Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems by Martin Kleppmann (2020)
2. Distributed Systems: Principles and Paradigms by Andrew S. Tanenbaum and Maarten Van Steen (2022)

3. NoSQL Databases and Distributed Systems: Concepts and Design by Pramod J. Sadalage and Martin Fowler (2021)
4. Building Scalable Distributed Systems by Chandrika Srinivasan (2023)
5. Data Management for Researchers: Organize, Maintain, and Share Your Data for Research Success by Kristin Briney (2020)
6. Database Internals: A Deep Dive into How Distributed Data Systems Work by Alex Petrov (2021)
7. Fundamentals of Database Management Systems by Mark L. Gillenson (2020)
8. Cloud Computing and Distributed Systems: Design and Implementation by Tobias D. Huang and Martin R. Patel (2021)
9. Cassandra: The Definitive Guide: Distributed Data at Web Scale by Jeff Carpenter and Eben Hewitt (2022)
10. Modern Data Architecture on AWS by Daniel Walker and Adrian Cantrill (2023)

Assessment and grading

Criteria for assessment of student performance, and the final score structure

The final grade is made up of 100% assessment results in the form of an exam (50%) and current assessment (50%).
Breakdown of the grading:
Laboratory work: 30%
Independent work and computational tasks: 20%

Grading scale

Total points	National	ECTS
90–100	Excellent	A
82–89	Good	B
75–81	Good	C
64–74	Satisfactory	D
60–63	Satisfactory	E
35–59	Unsatisfactory (requires additional learning)	FX
1–34	Unsatisfactory (requires repetition of the course)	F

Norms of academic integrity and course policy

The student must adhere to the Code of Ethics of Academic Relations and Integrity of NTU "KhPI": to demonstrate discipline, good manners, kindness, honesty, and responsibility. Conflict situations should be openly discussed in academic groups with a lecturer, and if it is impossible to resolve the conflict, they should be brought to the attention of the Institute's management.
Regulatory and legal documents related to the implementation of the principles of academic integrity at NTU "KhPI" are available on the website: http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/

Approval

Approved by	Date, signature	Head of the department Pavlo PUSTOVOITOV
	Date, signature	Guarantor of the educational program Vitaliy BRESLAVETS