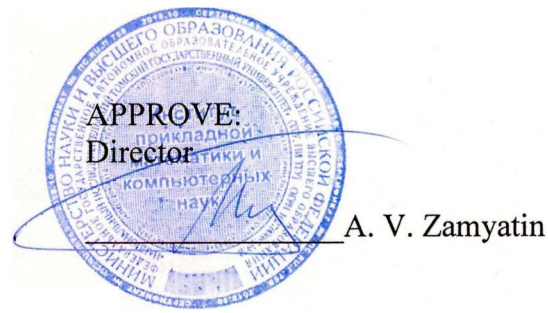
Appendix 1

Ministry of Science and Higher Education of the Russian Federation
NATIONAL RESEARCH
TOMSK STATE UNIVERSITY (NR TSU)

Institute of Applied Mathematics and Computer Science



Evaluation materials of the current control and intermediate certification in the discipline

(Evaluation tools by discipline)

High Performance Computing

in the major of training

01.04.02 Applied mathematics and informatics

Orientation (profile) of training:

Big Data and Data Science

Tomsk – 2025

ET was implemented:
Ph.D. tech. Sciences,
Associate Professor of the Department
of Theoretical Foundations of Computer Science D.V. Druzhinin

Reviewer:
Dr. Tech. sciences, professor,
Professor of the Department of Theoretical Foundations of Informatics Yu.L. Kostyuk

Evaluation tools were approved at a meeting of the educational and methodological commission of the Institute of Applied Mathematics and Computer Science (EMC IAMCS).

Protocol dated 22.05.2025 № 1

Chairman of the EMC IAMCS,
Dr. tech. Sciences, Professor S.P. Sushchenko

Evaluation tools (ET) are an element of the system for assessing the formation of competencies among students in general or at a certain stage of its formation.

The ET is developed in accordance with the work program (WP) of the discipline.

1. Competencies and training outcomes, obtained upon the discipline mastery

Competencies	**Competence indicator**	**Code and name of planned training outcomes that characterize the stages of competency formation**	**Criteria for evaluating training outcomes**			
			Excellent	Good	Satisfactory	Unsatisfactory

UK-1. Able to critically analyze problem situations based on a systematic approach and develop an action strategy	IAA-1.1. Identifies a problematic situation and, based on a systematic approach, carries out its multifactor analysis and diagnosis. IUC-1.2. Searches, selects and systematizes information to determine alternative strategic solutions in a problem situation. IUC-1.3. Proposes and justifies an action strategy taking into account limitations, risks and possible consequences.	OP-1.1.1: Know the relevance, basic terminology, architecture of high-performance systems. OP-1.2.1: Know methods and algorithms for high-performance data processing. OP-1.3.1: Possess skills in using in-depth theoretical and practical knowledge in the use of high-performance data processing technologies to solve data mining and bioinformatics problems. OP-1.3.2: Have the skills to develop and apply mathematical methods of system and application software to solve problems in scientific, design and technological activities	Demonstration of a high level of knowledge; ability to independently analyze and implement acquired knowledge. Knowledge of terminology, architecture of high-performance systems; methods and algorithms for high-performance data processing. Possession of skills in using in-depth theoretical and practical knowledge in the field of high-performance data processing technologies to solve problems of data mining and bioinformatics and the development and application of mathematical methods of system and application software to solve problems of scientific and design-	Generally successful, but containing some gaps in knowledge of terminology, architecture of high-performance systems; methods and algorithms for high-performance data processing. Generally successful, but containing some gaps, the skills to use practical knowledge in the field of high-performance data processing technologies to solve problems of data mining and the development and application of mathematical methods of system and application software to solve problems of scientific and design-technological	Fragmentary, incomplete knowledge of terminology and architecture of high-performance systems without gross errors; methods and algorithms for high-performance data processing. Fragmented, incomplete possession without gross errors of skills in using practical knowledge in the field of high-performance data processing technologies to solve data mining problems.	Does not have a clear understanding of the material being studied and makes serious mistakes when using practical knowledge in the field of high-performance data processing technologies to solve data mining problems.

PC-6. Able to manage the receipt, storage, transmission, processing of large	IPK-6.1. Monitors and evaluates the performance of big data processing. IPK-6.2. Uses methods and tools for receiving, storing, transmitting, processing big data IPK-6.3. Develops proposals to improve the performance of big data processing	OP-6.1.1: Be able to use system and application software to solve scientific and practical problems of technological activities in the field of data mining and bioinformatics. OP-6.2.1: Be able to select methods and means of high-performance data processing to solve research and applied problems. OP-6.3.1: Be able to develop and apply mathematical methods to solve scientific research problems.	Demonstration of a high level of knowledge; ability to use system and application software to solve scientific and practical problems of technological activities in the field of data mining and bioinformatics. Ability to select methods, tools and mathematical methods for solving high-performance data processing problems for solving scientific research and applied problems.	Generally successful, but containing some gaps in the ability to use system and application software to solve scientific and practical problems of technological activities in the field of data mining and bioinformatics. Generally successful, but containing some gaps in the ability to select methods, tools and mathematical methods for solving high-performance data processing problems for solving scientific research and applied problems.	Fragmentary, incomplete ability without gross errors to use system and application software to solve scientific and practical problems of technological activities in the field of data mining and bioinformatics.	Does not have a clear idea of the material being studied and makes gross mistakes when choosing methods and means of high-performance data processing to solve research and applied problems.

2. Stages of competency formation and types of evaluation tools

№	Stages of competency formation (discipline sections)	Code and name of training outcomes	Type of evaluation tool (tests, assignments, cases, questions, etc.)
1	Relevance, basic terminology and development trends. Parallelism of computer calculations	OP-1.1.1, OP-1.2.1	Current control issues
2	High Performance Computing Systems Architecture	OP-6.2.1, OP-6.1.1, OP-6.3.1	Current control issues
3	Cloud technologies, their properties and types	OP-1.3.1, OP-1.3.2	Current control issues
4	MapReduce	OP-6.2.1; OP-6.1.1; OP-6.3.1; OP-1.3.1	Current control issues
5	Distributed file systems	OP-6.2.1; OP6.1.1; OP-6.3.1; OP-1.3.1; OP-1.3.2	Current control issues
6	Programming for high performance computing. Methodology for designing parallel algorithms	OP-6.2.1; OP-61.1; OP-6.3.1; OP-1.3.1; OP-1.3.2; OP-6.1.1; OP-6.2.1;OP-6.3.1	Laboratory work, current control issues

3. Typical control tasks or other materials necessary for the assessment of educational training outcomes

3.1. Typical tasks for conducting ongoing monitoring of progress in the discipline:

Current control in the discipline is carried out in the form of questions on lecture material, laboratory work and is recorded in the form of a control point at least once a semester.

3.2. Typical tasks for conducting intermediate certification in the discipline.

Questions for intermediate certification:

Topic 1. Relevance, basic terminology and development trends. Parallelism of computer calculations

1. Define the term “high performance computing.”

2.Name the main ways to increase the performance of a computing system.

3.What are the difficulties associated with further increasing the processor clock frequency?

4. Define the term “multiprocessing”.

5.Tell about vector data processing.

6.Is vector data processing used in modern personal computers? Name the relevant processor technologies/extensions.

7.Tell about multifunctional data processing.

8.Tell about the command pipeline.

9.What is the difference between peak and real performance of a supercomputer?

10.Which supercomputer currently has the highest peak performance?

Topic 2. Architecture of high-performance computing systems.

Classification of computing systems.

1.By what criterion are computing systems classified in the Flynn classification?

2.How do they differ from other SISD systems?

3.How do they differ from other SIMD systems?

4.How are MISD systems different from other systems?

5.How are MIMD systems different from other systems?

6.What is the difference between multiprocessors and multicomputers?

7. How do UMA systems differ from NUMA systems?

8.How are CC-NUMA systems different from other systems?

9.How are NCC-NUMA systems different from other systems?

10.How do they differ from other SMP systems?

11. How do clusters differ from other multiprocessor computing systems?

Topic 3. Cloud technologies, their properties and types.

1.What is "cloud technology"?

2.Name the main properties of cloud technologies.

3.What is the reason for the widespread adoption of cloud technologies?

4.Name the main types of cloud services when classifying according to the type of resource provided.

5.Give examples of cloud services.

6.Why is HaaS (Hardware as a Service) not a cloud technology?

7.What is the difference between DaaS and SaaS?

8.Which types of cloud services are intended primarily for end users, and which for software developers?

9.What models of deployment of cloud systems do you know?

10.What is the purpose of hybrid clouds?

Topic 4. MapReduce technology.

1. Define "higher order function" in programming.

2.What are the higher order functions used in MapReduce cloud computing technology?

3.Name the steps of the general MapReduce algorithm.

4.What types of nodes are created when using MapReduce technology and what is their purpose?

5.What characteristics must a computational problem have in order for it to be effectively solved using MapReduce technology?

6.Give an example of a problem that can be effectively solved using MapReduce technology?

7.Write down the signature of the map() function in a pseudo-programming language for the example given.

8.Write down the signature of the reduce() function in a pseudo-programming language for the example given.

9.Name the most famous implementations of MapReduce.

10.For what types of computing systems in Flynn's classification does it make sense to use MapReduce technology?

Topic 5. Distributed file systems.

1. Define "distributed file system".

2. Define "data replication."

3.What is the purpose of data replication?

4.What are the areas of application of distributed file systems?

5.What is the difference between a distributed file system and a distributed data storage?

6.What distributed file systems do you know?

7.Describe how Google File System works.

8.What happens in the Google File System if the data node holding the master copy of a file fragment fails?

9.Describe the operating principle of Hadoop distributed file system.

10.What happens in Hadoop distributed file system if the data node holding the master copy of a file fragment fails?

11.Name the similarities and differences between Google File System and Hadoop distributed file system.

Topic 6. Programming for high-performance computing. Methodology for designing parallel algorithms.

1.Name the main problems of parallel programming.

2.Tell us about the methodology for organizing parallel computing for SIMD architecture.

3.Tell us about the methodology for organizing parallel computing for MIMD architecture.

4.Name the main quality indicators of parallel methods.

5. How are indicators such as acceleration and efficiency related to each other?

6.Can the speedup value exceed the number of processors used to run a parallel program?

7.What is the difference between strong and weak scalability?

8.What is the MPI library used for, what main capabilities does it provide?

9.What is OpenMP, what capabilities does it provide?

10.Does it make sense to use MPI on shared memory computers?

4. Methodological materials that determine the procedures for evaluating training outcomes

4.1. Methodological materials for assessing the current control of progress in the discipline.

Rating system for assessing current student performance

Table - Scoring for control elements

Elements of learning activity	Maximum score since the beginning of the semester	Assessed competence
Poll on topics in class	20	UK-1, PC-6
Laboratory works	50	UK-1, PC-6

4.2. Methodological materials for conducting intermediate certification in the discipline.

Methodological materials include: the procedure for forming an assessment when using a point-rating system; criteria for assessing results for intermediate certification, taking into account assessments for competencies.

Rating system for assessing students' intermediate performance

Table - Scoring for control elements

Elements of learning activity	Maximum score since the beginning of the semester	Assessed competence
Poll on topics in class	20	UK-1, PC-6
Laboratory works	50	UK-1, PC-6
Exam	30	UK-1, PC-6

Recalculation of scores into intermediate performance assessments

Points on the checkpoint date	Mark
≥ 90% from the maximum points	5 (passed)
from 70% till 89% rom the maximum points	4 (passed)
Ofrom 60% till 69% rom the maximum points	3 (passed)
< 60% rom the maximum points	2 (failed)