

Б А К А Л А В Р И А Т

*А.А. Григорьев, Е.А. Исаев, П.А. Тарасов*

**ПЕРЕДАЧА, ХРАНЕНИЕ  
И ОБРАБОТКА  
БОЛЬШИХ ОБЪЕМОВ  
НАУЧНЫХ ДАННЫХ**

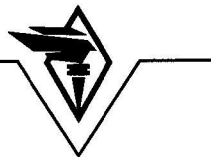
У Ч Е Б Н О Е П О С О Б И Е



Электронно-  
Библиотечная  
Система  
**znanium.com**



Данная книга доступна в цветном  
исполнении в электронно-библиотечной  
системе Znanium.com



**А.А. ГРИГОРЬЕВ**

**Е.А. ИСАЕВ**

**П.А. ТАРАСОВ**

# **ПЕРЕДАЧА, ХРАНЕНИЕ И ОБРАБОТКА БОЛЬШИХ ОБЪЕМОВ НАУЧНЫХ ДАННЫХ**

**УЧЕБНОЕ ПОСОБИЕ**

*Рекомендовано Межрегиональным учебно-методическим советом профессионального образования в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлениям подготовки 01.03.02 «Прикладная математика и информатика», 09.03.01 «Информатика и вычислительная техника», 38.03.05 Бизнес-информатика» (квалификация (степень) «бакалавр») (протокол № 10 от 12.10.2020)*

**Электронно-  
Библиотечная  
Система**  
**znanium.com**

Москва  
ИНФРА-М  
2022

**УДК 004.6(075.8)**  
**ББК 32.973.202я73**  
**Г83**

Рецензенты:

*Сиденко А.В.*, доктор экономических наук, профессор, президент, директор Института статистики Российской академии диалектико-системных исследований и разработок;

*Тихвинский В.О.*, доктор экономических наук, кандидат технических наук, старший научный сотрудник, профессор кафедры «Цифровая экономика, управление и бизнес-технологии» факультета «Цифровая экономика и массовые коммуникации» Московского технического университета связи и информатики

**Григорьев А.А.**

**Г83** Передача, хранение и обработка больших объемов научных данных : учебное пособие / А.А. Григорьев, Е.А. Исаев, П.А. Тарасов. — Москва : ИНФРА-М, 2022. — 207 с. — (Высшее образование: Бакалавриат). — DOI 10.12737/1073525.

ISBN 978-5-16-015985-0 (print)

ISBN 978-5-16-108415-1 (online)

В учебном пособии рассматриваются крупные научные проекты и объемы генерируемых ими данных, дается обзор научных компьютерных сетей, позволяющих производить высокоскоростную передачу больших объемов данных для этих проектов. Рассматриваются вычислительные системы, предлагаемые ведущими производителями компьютерной техники для обработки больших объемов данных и предоставляющие как возможности хранения больших объемов данных, в том числе распределенных, так и средства аналитики и параллельной обработки данных в реальном масштабе времени. Особое внимание уделено безопасности передаваемой научной информации.

Соответствует требованиям федеральных государственных образовательных стандартов высшего образования последнего поколения.

Для студентов технических специальностей бакалавриата, магистратуры, специалитета, обучающихся по направлениям подготовки «Прикладная математика и информатика», «Бизнес-информатика» и «Информатика и вычислительная техника».

**УДК 004.6(075.8)**  
**ББК 32.973.202я73**



Данная книга доступна в цветном исполнении в электронно-библиотечной системе Znanium.com

ISBN 978-5-16-015985-0 (print)  
ISBN 978-5-16-108415-1 (online)

© Григорьев А.А., Исаев Е.А.,  
Тарасов П.А., 2021

# Оглавление

<b>Список принятых сокращений.....</b>	<b>3</b>
<b>Введение .....</b>	<b>7</b>
<b>Глава 1. Передача больших объемов научных данных.....</b>	<b>12</b>
1.1. Научные проекты, генерирующие большие объемы данных.....	12
1.2. Высокоскоростные научные и коммерческие компьютерные оптоволоконные линии связи .....	15
1.2.1. Общеввропейская сеть науки и образования GEANT .....	15
1.2.2. Проект European Grid Infrastructure.....	20
1.2.3. Национальная научная сеть США Energy Sciences Network .....	22
1.2.4. Корпоративная сеть Российской академии наук RASNet.....	24
1.2.5. Федеральная университетская компьютерная сеть России RUNNet.....	24
1.3. Перспективные технологии и разработки по передаче данных.....	27
1.4. Передача информации при помощи «закрученного» света.....	29
1.5. Оптические солитоны.....	31
1.6. Проблемы космической дальней радиосвязи.....	33
<i>Контрольные задачи .....</i>	<i>35</i>
<i>Контрольные вопросы и задания .....</i>	<i>40</i>
<b>Глава 2. Обработка больших объемов научных данных.....</b>	<b>41</b>
2.1. Большие данные и проблемы их обработки.....	41
2.2. Центры обработки данных.....	58
2.3. Модульные центры обработки данных .....	65
2.4. Компьютерные кластеры .....	67
2.5. Современная биоинформатика и большие данные .....	68
2.6. Программное обеспечение для работы с большими данными.....	70
2.7. Форматы данных.....	73
2.8. Хранение и обмен данными.....	74
<i>Контрольные задачи .....</i>	<i>75</i>
<i>Контрольные вопросы и задания .....</i>	<i>78</i>
<b>Глава 3. Хранение больших объемов научных данных. Пример CERN .....</b>	<b>79</b>
<i>Контрольные вопросы и задания .....</i>	<i>86</i>
<b>Глава 4. Безопасная передача научной информации по линиям связи .....</b>	<b>87</b>
4.1. Современные методы обеспечения информационной безопасности волоконно- оптических линий связи .....	87
4.2. Способы защиты информации, передаваемой с помощью ВОЛС от НСД.....	89
4.2.1. Обнаружение и локализация факта НСД.....	89
4.2.2. Методы защиты передаваемой информации.....	91
4.3. Перспективные разработки в области квантовой криптографии .....	95
4.3.1. Проблема создания высокоскоростных линий квантовой связи.....	96

4.3.2. Проблема увеличения расстояния передачи по квантовым линиям связи.....	97
<i>Контрольные вопросы и задания</i> .....	103
<b>Приложение 1.</b> Обработка больших экспериментальных данных на примере Пушинского научного центра (ПНЦ) РАН .....	105
<b>Приложение 2.</b> Организация технологии обработки данных и пример применения трехзвенной технологии в реальной информационной системе.....	121
<b>Приложение 3.</b> Основные угрозы для надежности и безопасности дата-центров и методы борьбы с ними.....	127
<b>Приложение 4.</b> Основные методы обеспечения безопасности беспроводных сенсорных сетей.....	141
<b>Приложение 5.</b> Обеспечение информационной безопасности облачных вычислений .....	150
<b>Заключение</b> .....	164
<b>Темы рефератов</b> .....	165
<b>Тесты для аттестации знаний по изучаемому курсу</b> .....	167
<b>Ответы на тесты</b> .....	176
<b>Библиографический список</b> .....	177
<b>Глоссарий</b> .....	203