



ЦИФРОВАЯ ТРАНСФОРМАЦИЯ ОСНОВНЫЕ ПОНЯТИЯ И ТЕРМИНОЛОГИЯ

УДК 004:001.4(082)

ББК 32.81я43

Ц75

Редакционная коллегия:

А. В. Тузиков (председатель), С. В. Кругликов, В. А. Лапицкий,
Р. Б. Григянец, Г. Н. Наumenко

Рецензенты:

академик НАН Беларуси, доктор технических наук, профессор С. В. Абламейко,
член-корреспондент НАН Беларуси, доктор физико-математических наук,
профессор М. Я. Ковалев

Цифровая трансформация. Основные понятия и терминология : сб.
Ц75 статей / редкол.: А. В. Тузиков (пред.) [и др.] ; Нац. акад. наук Беларуси,
Объед. ин-т проблем информатики. – Минск : Беларуская навука, 2020. –
267 с. : ил.

ISBN 978-985-08-2588-9.

В сборнике статей рассматриваются основные понятия и определения информатизации и цифровой трансформации. Пять тематических блоков охватывают различные аспекты знаний в области информационно-коммуникационных технологий, вычислительной и робототехники, грид-систем и суперкомпьютеров, облачных технологий и квантовых вычислений, информатики, разработки программного обеспечения, баз данных и искусственных нейронных сетей, цифрового контента, информационной безопасности, а также построения информационного общества.

Издание предназначено для научных работников, учащихся, студентов, магистрантов, аспирантов и других специалистов.

УДК 004:001.4(082)

ББК 32.81я43

ISBN 978-985-08-2588-9

© Оформление. РУП «Издательский дом
«Беларуская навука», 2020

Технической основой средств вычислительной техники в настоящее время являются микроэлектронные изделия – цифровые сверхбольшие интегральные схемы и системы на кристалле, осуществляющие хранение, передачу и переработку цифровой двоичной (0,1) информации, которая в закодированном виде представляет все ее многообразие – текстовую и графическую информацию, звуковую и видеоинформацию и т. д. Успехи микроэлектроники позволили создать мощные персональные компьютеры и серверы, а также мобильные устройства – смартфоны, которые не только используются как телефоны, но и выполняют большое число дополнительных функций, в том числе имеют доступ в Интернет. Создание быстрой и надежной сети Интернет, совершенствование программного обеспечения привели к развитию облачных технологий, благодаря которым пользователи получают удаленный доступ к различным вычислительным ресурсам и большим объемам данных. Однако размерности задач и сложность требуемых вычислений быстро возрастают, так как практике приходится решать все более сложные задачи моделирования физических и технологических процессов, автоматизации проектирования, прогноза погоды и т. д. Эффективным практическим подходом к решению данных проблем является создание суперкомпьютерных центров и грид-систем с использованием параллельных вычислений. Тем не менее имеются задачи, например, в области криптографии, для которых существующих вычислительных средств совершенно недостаточно, поэтому для решения такого класса задач начинают использоваться квантовые вычисления и квантовые компьютеры. Развитие высокоразвитых технологий взаимосвязано. Совершенствование программно реализованных методов искусственного интеллекта, успехи микроэлектроники, механики (машиностроения) привели к созданию автоматических производственных линий и промышленных роботов в тяжелой и автомобильной промышленности, приборостроении и других отраслях. В настоящее время активно разрабатываются самые разные автономно функционирующие роботы, например, пожарные, космические, сервисные (бытовые), в том числе человекоподобные, которые заменяют обслуживающий персонал в гостиницах, банках и т. д. Общеизвестно, что информационные вычислительные технологии, искусственный интеллект и роботы, наряду с другими перспективными технологиями, составят основу четвертой промышленной революции.

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

П. Н. Бибило, Н. Н. Парамонов

Под вычислительной техникой понимается совокупность технических и программных средств, методов и средств (алгоритмов, технологий), используемых для автоматизации процессов вычислений, хранения и обработки информации.

Основным содержанием понятия «*вычислительная техника*» являются компьютеры. Слово «*компьютер*» – производное от английских слов to compute, computer, которые переводятся как «вычислять», «вычислитель». Первоначально в английском языке это слово означало человека, производящего арифметические вычисления с привлечением или без привлечения механических устройств. Впервые трактовка слова *компьютер* появилась в 1897 г. в Оксфордском английском словаре. Его составители тогда понимали компьютер как механическое вычислительное устройство.

Человечество с глубокой древности пыталось облегчить себе жизнь, изобретая разнообразные приспособления и устройства для вычислений. Самые первые идеи создания программируемой механической вычислительной машины выдвинуты более 160 лет назад. Они принадлежали англичанину Чарльзу Бэббиджу, чьи идеи намного опередили свое время. «Элементарной базой» машины Бэббиджа служили «цифровые» колеса с различным числом зубьев. Созданная в 1834 г. Ч. Бэббиджем аналитическая машина стала первым прообразом современных компьютеров.

Свое название компьютеры получили по основной функции – проведению вычислений. Однако в настоящее время полагают, что основные функции компьютеров – обработка и управление информацией. Компьютеры настолько широко вошли в нашу жизнь, что мы уже и не представляем, как люди обходились без них раньше. На современных компьютерах пишут и рисуют, создают мультфильмы, рекламные ролики и специальные эффекты для фильмов. Особые программы помогают учиться, проектировать космические ракеты и города, предсказывать погоду. Компьютеры позволяют создавать и записывать музыку, производить сложные научные расчеты и эксперименты, восстанавливать облик вымерших животных, прогнозировать будущее. Самые интересные игры сегодня – это опять-таки компьютерные игры.

Наибольшее распространение среди компьютеров получили так называемые электронные вычислительные машины (ЭВМ). Теоретическим фунда-

ментом современных ЭВМ является булева алгебра, основы которой в 1854 г. разработал английский математик и философ Джордж Буль. Булева алгебра устанавливает правила выполнения логических операций сложения (ИЛИ), умножения (И) и отрицания (НЕ), которые технически реализуются электронными элементами. Математик Джон фон Нейман заложил основные принципы построения современных ЭВМ, состоящих из большого числа элементарных логических схем, действующих по двоичному принципу, сформулировав в 1946 г. основные принципы реализации архитектуры электронных вычислительных машин:

1. Принцип программного управления. Программа состоит из набора команд, которые выполняются процессором друг за другом в определенной последовательности.

2. Принцип однородности памяти. Как программы (команды), так и данные хранятся в одной и той же памяти (и кодируются в одной и той же системе счисления, чаще всего двоичной). Над командами можно выполнять такие же действия, как и над данными.

3. Принцип адресуемости памяти. Структурно основная память состоит из пронумерованных ячеек; процессору в произвольный момент времени доступна любая ячейка.

Компьютеры, построенные на этих принципах, относят к типу фоннеймановских. Очевидно, что наряду с аппаратными средствами важнейшим компонентом, необходимым для реализации архитектурных принципов ЭВМ, является программное обеспечение. В разговорной речи, говоря о программном обеспечении, часто используют слово «софт» (от англ. software). Программное обеспечение включает компьютерные программы и данные, предназначенные для решения определенного круга задач и хранящиеся на машинных носителях.

В вычислительной технике принято выделять этапы развития электронных вычислительных машин. ЭВМ относят к тому или иному поколению в зависимости от используемых при создании компьютеров технологий.

Самые ранние компьютеры были полностью механическими системами, в 1930-е гг. появились электромеханические компоненты (реле), а в 1940-х гг. были созданы первые полностью электронные компьютеры, имевшие в своей основе вакуумные электронные лампы. В 1950–60-е гг. на смену лампам пришли транзисторы, а в конце 1960-х – начале 1970-х гг. – полупроводниковые интегральные схемы (кремниевые чипы).

ЭВМ на электронных лампах принято относить к первому поколению компьютеров. Первая ЭВМ – ENIAC – была создана в США в 1948 г. и выполняла ограниченное число задач. Вместо реле в ENIAC использовались 18 тыс. электронных ламп, а для ввода-вывода информации – перфокарты. Первые компьютеры данного поколения в СССР были созданы немного позднее. К ним относятся такие ЭВМ, как БЭСМ-1, Урал-1 и др. Они содержали несколько десятков тысяч ламп и проделывали 5–10 тыс. операций в секунду. В Беларуси

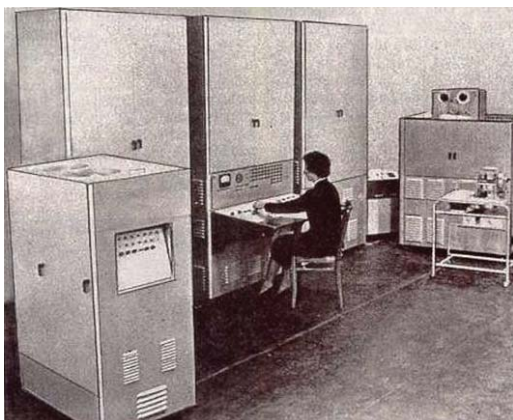


Рис. 1. ЭВМ «Минск-1»

была создана серийная ЭВМ первого поколения «Минск-1» производительностью 3000 операций в секунду (рис. 1).

Элементной базой ЭВМ второго поколения являются полупроводниковые приборы – транзисторы (рис. 2, б). Первые компьютеры на основе транзисторов появились в конце 1950-х гг., а к середине 1960-х гг. были созданы более компактные внешние устройства, что позволило фирме Digital Equipment выпустить в 1965 г. первый мини-компьютер PDP-8 размером с холодильник. Применение транзисторов в качестве основного элемента в ЭВМ привело к уменьшению размеров компьютеров в сотни раз и к повышению их надежности. Одновременно с процессом замены электронных ламп транзисторами совершенствовались методы хранения информации. Увеличился объем памяти, а магнитную ленту начали использовать как для ввода, так и для вывода информации. С середины 1960-х гг. получило распространение хранение информации на дисках.

Совершенствование архитектуры компьютеров позволило достичь быстродействия в миллион операций в секунду. В то время СССР шел в ногу со временем и выпускал ЭВМ мирового уровня. Одна из мощнейших машин этого класса – БЭСМ-6 (рис. 3) создана в СССР в 1967 г. Она совершала около 1 млн операций в секунду, содержала 60 тыс. биполярных транзисторов и 200 тыс. диодов.

Совершенствование архитектуры компьютеров позволило достичь быстродействия в миллион операций в секунду. В то время СССР шел в ногу со временем и выпускал ЭВМ мирового уровня. Одна из мощнейших машин этого класса – БЭСМ-6 (рис. 3) создана в СССР в 1967 г. Она совершала около 1 млн операций в секунду, содержала 60 тыс. биполярных транзисторов и 200 тыс. диодов.

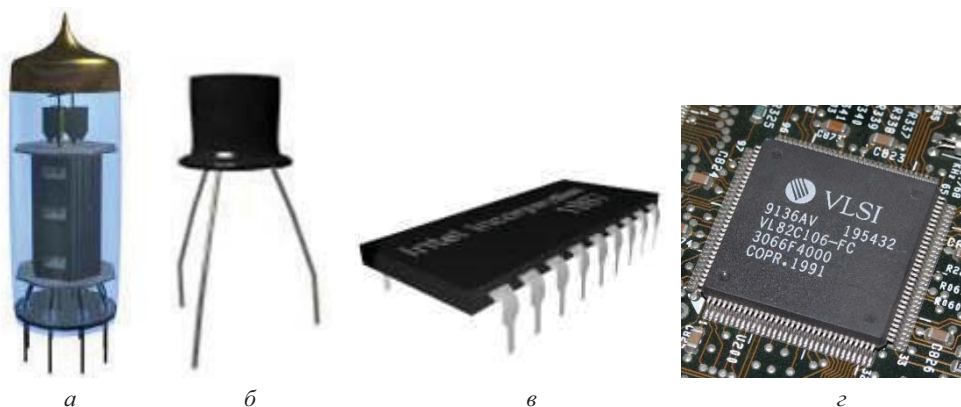


Рис. 2. Развитие элементной базы ЭВМ: а – электронная лампа; б – транзистор; в – интегральная схема малой степени интеграции; г – сверхбольшая интегральная схема



Рис. 3. ЭВМ БЭСМ-6



Рис. 4. ЭВМ ЕС-1060

Появление интегральных схем (ИС) ознаменовало собой новый этап в развитии вычислительной техники – рождение ЭВМ третьего поколения. Появление ИС означало подлинную революцию в вычислительной технике. Размеры и потребляемая мощность электронных схем сильно уменьшились, а их надежность значительно возросла. Выпускаемые в начале 1960-х гг. ИС содержали несколько десятков транзисторов, их называли интегральными схемами малой степени интеграции (рис. 2, в). Первая в СССР серийная ЭВМ третьего поколения ЕС-1020 была создана в Минске в 1971 г. Одной из самых мощных ЭВМ третьего поколения в СССР была ЭВМ ЕС-1060 (рис. 4). Она обладала быстродействием до 3 млн операций в секунду, и содержала более 10 млн транзисторов.

Бурное развитие полупроводниковой технологии позволило в начале 1970-х гг. перейти к созданию больших и сверхбольших интегральных схем (БИС и СБИС) (рис. 2, г), на основе которых были созданы ЭВМ четвертого поколения, выполняющие сотни миллионов операций в секунду. История четырех поколений электронных вычислительных машин является, по сути, историей развития больших универсальных ЭВМ, предназначенных для выполнения интенсивных вычислительных работ и организации централизованных хранилищ данных большой емкости.

Затем массовое распространение получили персональные ЭВМ, или персональные компьютеры (ПЭВМ, или ПК) (рис. 5). В чем причина феноменального успеха ПК? В 1981 г. появился ПК IBM PC (Personal Computer). Через несколько лет он стал стандартом персональных компьютеров. В ПК был реализован принцип, аналогичный принципам детского конструктора. Схемы, управляющие всеми устройствами (монитором, дисками, принтером, модемом и др.), реализованы на отдельных платах, которые вставляются в стандартные разъемы системной платы – слоты. Весь компьютер питается от единого блока питания. Этот принцип, названный принципом открытой архитектуры, наряду с другими достоинствами обеспечил потрясающий успех персональным компьютерам. Многие компании начали выпускать ПК, что

способствовало наступлению эры персональных компьютеров и компьютерной революции.

Развитие микроэлектроники и микропроцессорных технологий привело к созданию многоядерных универсальных процессоров и спецпроцессоров, число транзисторов в современных кристаллах достигает десятков миллиардов – это так называемые системы-на-кристалле. Сегодня лидерами в производстве процессоров являются две крупнейшие компании: Intel и AMD. Функциональные возможности процессоров стремительно развивались, увеличивалось количество ядер, появились гибридные процессоры на базе чипов, в которых интегрированы как ядра традиционных процессоров, так и ядра графических процессоров GPU (Graphic Processing Unit).

Появились малогабаритные накопители на жестких магнитных дисках (HDD) и твердотельные накопители (SSD), а также модули оперативной памяти большой емкости и мощные видеокарты с графическими спецпроцессорами. Все эти технические успехи дали новый толчок к развитию суперкомпьютерных технологий, базирующихся на принципе параллельных вычислений. Компьютерный мир изменился. Полувековая эпоха последовательной организации вычислений начала сменяться эрой параллелизма параллельных вычислительных технологий и параллельных вычислительных систем. Появление кластерных систем и переход всех ведущих производителей вычислительной техники на многоядерные процессоры сделали компьютерный мир параллельным. Большинство суперкомпьютеров имеет кластерную архитектуру. Кластер – это совокупность вычислительных узлов, объединенных между собой высокоскоростными связями для выполнения соответствующих задач под управлением операционной системы. Системное ПО кластерных суперкомпьютерных конфигураций обычно включает следующие основные программные компоненты: ОС (как правило, типа Linux); программные средства



Рис. 5. Персональный компьютер



Рис. 6. Суперкомпьютер «СКИФ К-500»

поддержки системной сети типа InfiniBand; средства разработки параллельных программ; системы пакетной обработки заданий. Тенденции развития информационных технологий, наличие и возрастающая доступность суперкомпьютерных ресурсов диктуют необходимость широкого внедрения принципов параллельных вычислений для решения сложнейших и наукоемких задач. Быстро развиваются облачные вычисления на основе обеспечения удобного доступа по сети Интернет к общему фонду конфигурируемых вычислительных ресурсов.

Созданные в Республике Беларусь кластерные суперкомпьютеры «СКИФ К-500» (рис. 6) и «СКИФ К-1000» в 2003–2006 гг. входили в рейтинг Топ-500 самых быстродействующих вычислительных конфигураций в мире. Поражают темпы роста производительности суперкомпьютерных систем. Первый Топ-500 самых мощных суперкомпьютеров в мире в июне 1993 г. возглавлял американский компьютер CM-5 с пиковой производительностью 131 гигафлопс (1024 процессоров). В публикации списка Топ-500 за ноябрь 2018 г. верхнюю строчку занимает американский суперкомпьютер Summit с пиковой производительностью 200 петафлопс (количество процессорных ядер – 2 397 824). Компьютерная революция продолжается.

СУПЕРКОМПЬЮТЕРНЫЕ И ГРИД-СИСТЕМЫ

Н. Н. Парамонов, О. П. Чиж

Суперкомпьютерные, или HPC (High Performance Computing) технологии являются, по сути, новым технологическим инструментом для постановки и решения принципиально и качественно новых научно-технических приложений в различных отраслях экономики. Использование суперкомпьютерных технологий обеспечивает получение новых знаний об объектах, сложных системах и явлениях как результат проведения серий вычислительных экспериментов. Суперкомпьютерные технологии позволяют использовать новые граничные условия и методы дискретизации многомерных объектов в динамически изменяемых условиях поведения, что практически невозможно при использовании обычных персональных компьютеров и рабочих станций.

В настоящее время очевидно, что суперкомпьютерные ресурсы относятся к факторам, определяющим стратегический потенциал страны оборонного, научно-технического и экономического значения. Темпы развития информационных технологий, наличие и возрастающая доступность суперкомпьютерных

ресурсов диктуют необходимость широкого внедрения принципов параллельных вычислений для решения сложнейших и наукоемких задач в ключевых экономически значимых секторах экономики.

Тенденции развития суперкомпьютерных технологий объективно отображаются в списках пятисот наиболее мощных суперкомпьютеров мира Топ-500, которые публикуются два раза в год начиная с 1993 г. В мае 2004 г. Межведомственный суперкомпьютерный центр РАН и Научно-исследовательский центр МГУ им. М. В. Ломоносова начали совместный проект по формированию списка 50 наиболее мощных компьютеров СНГ (Топ-50). В список включаются 50 вычислительных систем, установленных на территории СНГ и показавших к моменту выхода списка наибольшую производительность на тесте Linpack. Список обновляется два раза в год.

Формальное определение понятия «суперкомпьютер» было сформулировано на VII Национальном суперкомпьютерном форуме в Переславле-Залесском (28 ноября 2018 г.) член-корреспондентом РАН, директором ИПС РАН С. М. Абрамовым. Он заявил, что говорить про цифровую экономику нельзя, пока у страны нет развитой сети суперкомпьютеров. Только они могут быть фундаментом для цифровой экономики, цифрового проектирования и цифрового производства. С. М. Абрамов дал следующее определение понятия «суперкомпьютер»: «Главное свойство суперкомпьютера – огромная скорость расчетов, высокая производительность, причем не просто высокая, а самая высокая в мире. Дважды в год появляется рейтинг самых производительных машин в мире (Топ-500), и теперь суперкомпьютерами называют те машины, которые вошли в этот список. Однако здесь никогда не будет военных разработок, поэтому определение приходится расширить. Суперкомпьютер – это тот компьютер, который попал в рейтинг Топ-500 или мог бы попасть по своим характеристикам, но не подавался по разным причинам».

Производительность суперкомпьютерных систем непрерывно возрастает. На момент публикации первого списка Топ-500 лучших суперкомпьютеров в июне 1993 г. самым мощным компьютером в мире был CM-5/1024, расположенный в Лос-Аламосской национальной лаборатории в Калифорнийском университете. Он находился в ведении Министерства энергетики США и был изготовлен компанией Thinking Machine. CM-5/1024 состоял из 1024 процессоров SuperSPARC, работающих на тактовой частоте 32 МГц. Теоретическая вычислительная мощность системы составляла 131 гигафлопс, но на тесте Linpack, который использовался для получения результатов для рейтинга Топ-500, фактическая производительность (59,7 Гфлопс) оказалась менее половины заявленной.

В публикации списка Топ-500 за ноябрь 2018 г. верхнюю строчку занимает суперкомпьютер Summit из американской Окриджской национальной лаборатории. Согласно спецификациям он обеспечивает пиковую производительность Rpeak в 200 петафлопс двойной точности (FP64). В Топ-500 этот суперкомпьютер лидирует с показателем Rmax в 143,5 Пфлопс. Количество ядер – 2 397 824. Потребляемая мощность установки – 9783 кВт.

Вычислительные конфигурации, являющиеся основой корпоративных (отраслевых) центров обработки данных (ЦОД) и входящие в списки 50 наиболее мощных компьютеров СНГ (Топ-50), будем называть корпоративными (отраслевыми) суперкомпьютерами. При дальнейшем изложении будем использовать термин «суперкомпьютер», имея ввиду вхождение такой вычислительной конфигурации в Топ-500 и (или) Топ-50.

Специфика архитектурных и технических решений, реализованных в процессе развития НРС-технологий, отражена в данной статье на практических примерах создания белорусских суперкомпьютерных конфигураций семейства «СКИФ».

Суперкомпьютерное направление «СКИФ» отметило в 2018 г. 20-летний юбилей. Аббревиатура «СКИФ» (СуперКомпьютерная Инициатива Феникс) отражает основную цель совместных программ этого направления – внедрение суперкомпьютерных технологий в экономику Беларуси и России для решения ресурсоемких задач путем создания ряда программно совместимых суперкомпьютерных конфигураций с широким спектром производительности и прикладных комплексов на их основе.

Суперкомпьютерное направление «СКИФ» развивается в рамках программ Союзного государства (рис. 1). Государственный заказчик этих программ в Республике Беларусь – Национальная академия наук (НАН) Беларуси, головной исполнитель в Республике Беларусь – Объединенный институт проблем информатики НАН Беларуси. За это время созданы пять поколений кластеров семейства «СКИФ»: конфигурации Ряда 1, Ряда 2, Ряда 3, Ряда 4 и конфигурации «СКИФ-ГЕО», а также персональные кластеры семейства «СКИФ-Триада».

Первая программа направления «СКИФ» «Разработка и освоение в серийном производстве семейства моделей высокопроизводительных вычислительных систем с параллельной архитектурой (суперкомпьютеров) и создание прикладных систем на их основе» была утверждена постановлением Исполнительного Комитета Союза Беларуси и России № 43 от 22 ноября 1999 г. С учетом темпов роста мировой суперкомпьютерной отрасли и реальных бюджетов проектов белорусско-российских программ направления «СКИФ» главной задачей становится реализация в этих проектах технических решений, соответствующих современному мировому уровню суперкомпьютерных технологий.

Опытные образцы белорусских суперкомпьютерных конфигураций «СКИФ» реализованы на базе перспективных для своего времени технических решений. Например, в 2003 г. в рамках программы Союзного государства «СКИФ» был создан кластер «СКИФ К-500» с пиковой производительностью 716,6 Гфлопс. Создание этого кластера явилось качественно новым результатом, позволившим вплотную приблизиться к терафлопсному диапазону производительности.

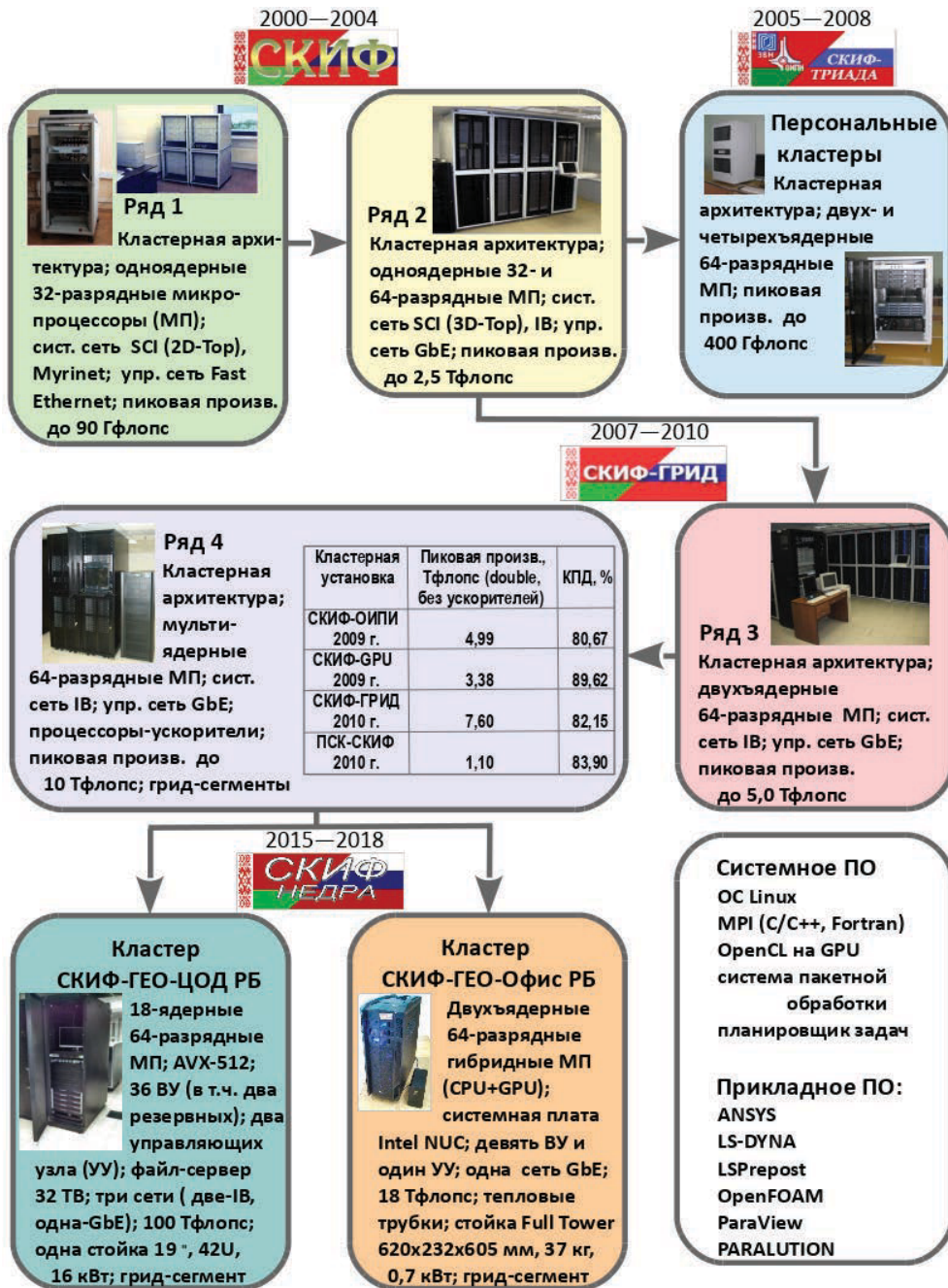


Рис. 1. Основные этапы создания суперкомпьютерного семейства «SKIIF» в Республике Беларусь

Реальная производительность кластера в 423,6 млрд операций в секунду, достигнутая на тесте Linpack, стала основанием для его включения под номером 407 в 22-й выпуск списка Топ-500. Включение кластера «СКИФ К-500» в Топ-500 наиболее мощных вычислительных установок в мире означало достижение уже в 2003 г. важного прямого политического эффекта: Республика Беларусь и Россия наравне с США, Японией и еще 11 странами стали создателями и обладателями престижной критической суперкомпьютерной технологии.

Старшая модель семейства «СКИФ» Ряда 2 – кластер «СКИФ К-1000» с пиковой производительностью 2,5 трлн операций в секунду – 9 ноября 2004 г. была включена в очередной 24-й выпуск списка Топ-500 под номером 98. В первую сотню рейтинга вошли суперкомпьютерные установки 16 стран, из них установки собственных разработчиков представили только пять стран: США, Япония, Китай, Беларусь и Россия. Суперкомпьютер «СКИФ К-1000» на 1 января 2005 г. являлся самым мощным компьютером на территориях СНГ и Восточной Европы и входил в четыре подряд редакции списка Топ-500.

Белорусские кластеры семейства «СКИФ» Ряда 4 до сих пор соответствуют современным базовым мировым тенденциям развития суперкомпьютерных технологий: кластерная архитектура, ОС Linux, сетевые интерфейсы – InfiniBand (IB) и Gigabit Ethernet (GbE), вычислительные узлы на основе многоядерных процессоров с архитектурой x86 и спецпроцессоров.

Эти кластеры неоднократно входили в списки Топ-50 самых мощных компьютеров СНГ. Так кластер «СКИФ ОИПИ» был включен в 11-ю редакцию списка Топ-50 от 22 сентября 2009 г. под номером 21, кластер «СКИФ-GPU» вошел в 13-ю редакцию списка Топ-50 от 21 сентября 2010 г. под номером 45, а кластер «СКИФ-ГРИД» включен в 14-ю редакцию списка Топ-50 от 29 марта 2011 г. под номером 40. Вошел в 29-ю редакцию от 24 сентября 2018 г. списка Топ-50 под номером 37 и кластер «СКИФ-ГЕО-ЦОД РБ» с пиковой производительностью 100 Тфлопс.

При оценке диапазона производительности перспективных белорусских суперкомпьютеров учитывались реальные объемы финансирования создания белорусских суперкомпьютерных конфигураций семейства «СКИФ», развитие грид-технологий в Республике Беларусь и реальная востребованность корпоративных (отраслевых) суперкомпьютерных ресурсов в нашей стране.

Бурное развитие суперкомпьютерных технологий приводит к существенному увеличению вычислительной плотности и снижению стоимости вычислений. Например, стоимость создания в 2004 г. кластера «СКИФ К-1000» (пиковая производительность 2,5 Тфлопс) составила 2 млн долларов. Стоимость 1 Гфлопс в этом случае составила 800 долларов. Оборудование кластера «СКИФ К-1000» располагалось в восьми 19” стойках, бесперебойное электро-

питание кластера обеспечивалось при помощи восьми стоек ИБП (источник бесперебойного питания).

Стоимость создания в 2018 г. кластера «СКИФ-ГЕО-ЦОД РБ» (пиковая производительность 100,0 Тфлопс) составила 0,55 млн долларов. Стоимость 1 Гфлопса в этом случае составила примерно 5,5 доллара. Оборудование кластера «СКИФ-ГЕО-ЦОД РБ» располагается в одной 19” стойке, бесперебойное электропитание кластера обеспечивается при помощи одной стойки ИБП. Потребляемая мощность – в пределах 18 кВт. Отметим, что с учетом экспертных оценок создание самых мощных суперкомпьютеров в мире (Топ-5) стоит 150–200 млн долларов. При этом стоимость 1 Гфлопса составляет 1,0–2,0 доллара. Потребляемая мощность – 10 и более мВт.

Базовые технические решения создания суперкомпьютеров семейства «СКИФ» непрерывно совершенствовались. Концепция создания моделей суперкомпьютеров семейства «СКИФ» базируется на масштабируемой кластерной архитектуре, реализуемой на классических кластерах из вычислительных узлов на основе компонент широкого применения (стандартных микропроцессоров, модулей памяти, жестких дисков и материнских плат, в том числе с поддержкой SMP). *Кластерный архитектурный уровень* – это тесно связанная сеть (кластер) вычислительных узлов (ВУ), работающих под управлением операционной системы (ОС). В семействе суперкомпьютеров «СКИФ» в качестве базовой классической системы поддержки параллельных вычислений выбран MPI (Message Passing Interface), что не исключает использования других средств. Базовая кластерная архитектура с использованием двух сетей (системной и командной) была основой для первых четырех поколений белорусских кластеров семейства «СКИФ». В кластере пятого поколения «СКИФ-ГЕО-ЦОД РБ» кластерный архитектурный уровень организован на базе трех сетей: системной сети для MPI-обменов, сети для файловых операций ввода/вывода и командной (вспомогательной) сети для оптимизации параллельного выполнения программ. Трехсетевая архитектура кластера «СКИФ-ГЕО-ЦОД РБ» представлена на рис. 2.

Первоначально белорусские кластерные конфигурации создавались на основе базовой кластерной архитектуры без использования спецпроцессоров. В процессе развития суперкомпьютерных технологий на рынке появились эффективные спецпроцессоры, которые использовались при создании белорусских суперкомпьютеров Ряда 4 (кластеры «СКИФ-ОИПИ», «СКИФ-GPU», «СКИФ-ГРИД»). В кластере «СКИФ-ОИПИ» использованы спецпроцессоры Cell, в кластерах «СКИФ-GPU» и «СКИФ-ГРИД» – графические процессоры GPU (Graphic Processing Unit). В этих кластерах используются гибридные ВУ на базе процессоров архитектуры x86-64 и процессоров-ускорителей.

Бурное развитие технологий НРС привело к естественной экспансии параллельной архитектуры (в основном кластерной) во все направления компьютерной отрасли: суперкомпьютеры, серверы, рабочие станции. Переход на

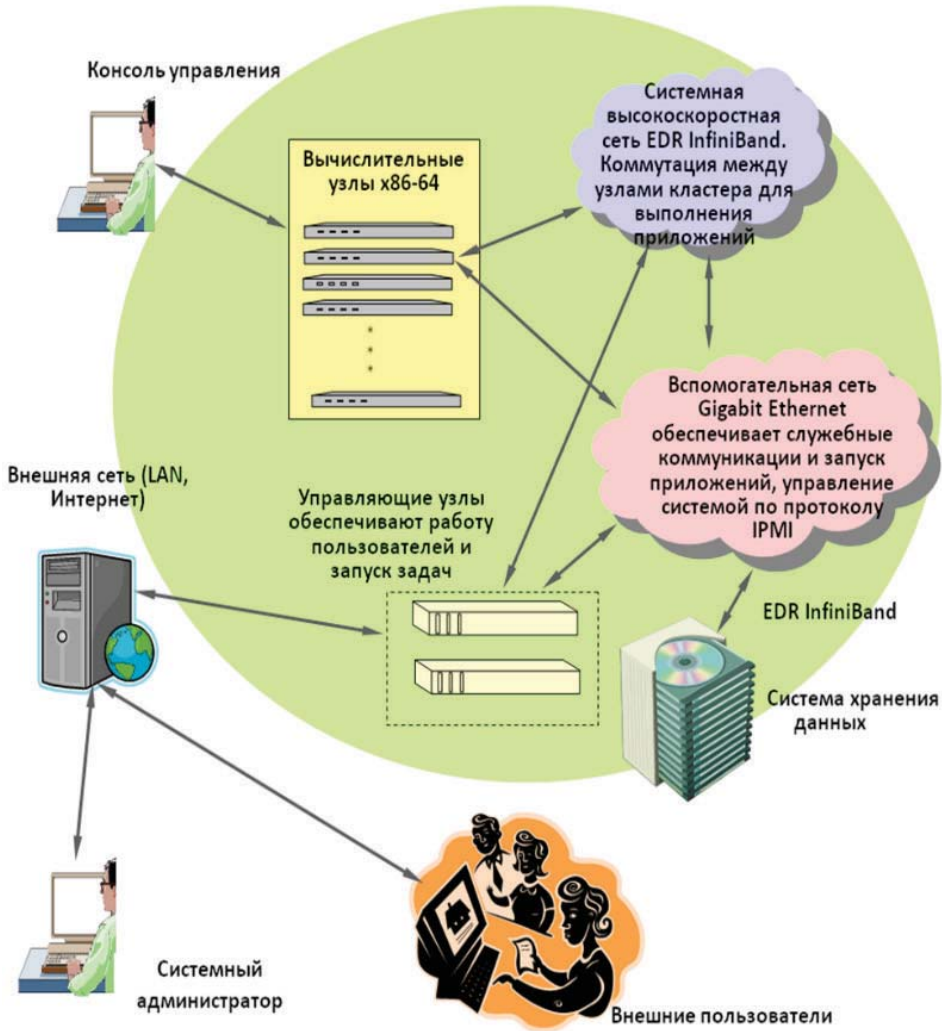


Рис. 2. Трехсетевая архитектура кластера «СКИФ-ГЕО-ЦОД РБ»

вычислительную технику с параллельной архитектурой вынуждает пользователей изменять весь привычный стиль взаимодействия с компьютерами. И это обстоятельство может в значительной степени снизить эффективность использования новой и к тому же дорогой суперкомпьютерной техники с параллельной архитектурой. Изложенную проблему может решить появление сравнительно недорогих малогабаритных вычислительных комплексов с кластерной архитектурой, которые можно устанавливать в обычных помещениях, лабораториях или офисах на рабочих местах пользователей. Персональные кластеры (персональные суперкомпьютеры) позволяют разрабатывать

методику применения параллельных вычислений для тех или иных прикладных задач, которые можно легко переносить на более высокопроизводительные кластеры. Адресной аудиторией персональных кластеров являются академические и научные общества, большие корпорации, учебные заведения и предприятия.

Примером реализации этого направления является малогабаритный офисный кластер «СКИФ-ГЕО-Офис РБ», созданный в 2018 г. в рамках программы Союзного государства «СКИФ-НЕДРА». Уровень акустических шумов, создаваемых офисным кластером, не должен превышать норм, установленных для офисных устройств.

Офисный кластер «СКИФ-ГЕО-Офис РБ» имеет встроенное воздушное охлаждение на базе отечественных тепловых трубок. В нем используются гибридные ВУ на базе процессоров APU (Accelerated Processing Unit), в которых интегрированы как ядра традиционных (классических) процессоров архитектуры x86-64, так и ядра GPU. Общий вид опытного образца кластера «СКИФ-ГЕО-Офис РБ» представлен на рис. 3, 4. В кластере «СКИФ-ГЕО-Офис РБ» предусмотрены 9 ВУ и 1 управляющий узел. Интегральная пиковая производительность кластера «СКИФ-ГЕО-Офис РБ» для 10 ВУ при выполнении операций с плавающей точкой с двойной точностью $P_{int} = 17,846$ Тфлопс.

Программное обеспечение (ПО) белорусских суперкомпьютеров семейства «СКИФ» включает следующие основные программные компоненты: ОС, программные средства поддержки системной сети InfiniBand (для кластеров ЦОД), средства разработки параллельных программ, системы пакетной обработки заданий.



Рис. 3. Общий вид опытного образца кластера «СКИФ-ГЕО-Офис РБ»



Рис. 4. Общий вид опытного образца кластера «СКИФ-ГЕО-Офис РБ» при снятых боковых панелях

В качестве базовой ОС используется ОС Linux – одна из самых надежных, эффективных и перспективных ОС, которую сегодня многие коммерческие и государственные организации выбирают в качестве базовой для приложений и перспективных разработок в области построения распределенных вычислительных систем и параллельных вычислений. ОС Linux свободно распространяется с исходными текстами. С учетом практического опыта было принято решение использовать в белорусских суперкомпьютерах семейства «СКИФ» последние версии дистрибутива Fedora Linux с обновлениями для повышения производительности кластеров.

В качестве средств разработки параллельных программ реализована связка из стандартных компиляторов дистрибутива Linux (C/C++, Fortran) и библиотеки MPI (Message Passing Interface).

В качестве программных средств пакетной обработки заданий используется свободно распространяемая версия системы пакетной обработки PBS (Portable Batch System) – Torque (Terascale Open-source Resource and QUEUE Manager), а в качестве планировщика задач используется программный пакет Maui.

На белорусских суперкомпьютерах семейства «СКИФ», как правило, устанавливаются следующие прикладные предметно ориентированные программные пакеты, обеспечивающие адекватность поведения моделируемых процессов и распараллеливание расчетов при разработке приложений в кластерной среде: ANSYS, LS-DYNA, LSPrepost, OpenFOAM, ParaView, PARALUTION.

Важнейшим практическим результатом реализации программы «СКИФ» является создание в ОИПИ НАН Беларуси Республиканского суперкомпьютерного центра коллективного пользования (РСКЦ КП) с возможностью удаленного доступа к его вычислительным ресурсам. Создание РСКЦ КП в ОИПИ НАН Беларуси обеспечило возможность предоставления услуг для решения сложнейших прикладных задач в различных отраслях экономики страны, требующих компьютерных и информационных ресурсов, владение которыми недоступно или экономически нецелесообразно для отдельных организаций.

Непрерывный рост производительности суперкомпьютерных систем дал толчок развитию распределенных вычислительных систем. Одним из перспективных направлений создания таких систем является грид-компьютинг. Грид – это современная технология создания эффективных распределенных гетерогенных сетей, обеспечивающих интеграцию информационных и вычислительных ресурсов. Основная задача грид – реализация гибкого, защищенного, скоординированного вычислительного пространства для совместного использования ресурсов динамически меняющимися сообществами пользователей. В распределенных грид-инфраструктурах реализуются механизмы обеспечения безопасности, в частности, аутентификация и авторизация пользователей, обмен сертификатами, обеспечение конфиденциальности и це-

лостности данных, а также аудит и мониторинг ресурсов и пользователей. Благодаря наборам стандартизированных служб (сервисов) обеспечивается надежный совместный доступ к территориально распределенным информационным и вычислительным ресурсам и данным (отдельным компьютерам, кластерам, хранилищам информации и сетям) независимо от места расположения потребителей. Практической реализацией идеи грид-компьютинга является создание аппаратно-программных комплексов на базе грид-инфраструктур с суперкомпьютерными ресурсами (АПК-ГРИД).

АПК-ГРИД фактически представляет собой систему обработки, хранения и предоставления данных на базе распределенной грид-инфраструктуры с суперкомпьютерными ресурсами и соответствующим ПО, предназначенную для интеграции распределенных информационных и вычислительных ресурсов с целью эффективного решения различных ресурсоемких задач. Распределенную инфраструктуру АПК-ГРИД, по сути, можно представить в виде «облака», включающего разнообразные ресурсы для обработки, хранения и передачи данных и предоставляющего соответствующие услуги (сервисы) по запросам абонентов. Абонентами АПК-ГРИД могут быть как сотрудники предприятия, в котором установлен этот комплекс, так и иные пользователи, имеющие соответствующие сертификаты (права) доступа к ресурсам комплекса. Абоненты АПК-ГРИД, оснащенные соответствующими средствами обработки и приема-передачи данных, взаимодействуют с АПК-ГРИД через соответствующие каналы связи локальной или глобальной сетей.

В общем случае АПК-ГРИД должен создаваться на базе соответствующей грид-сети или входить в состав некоторой грид-сети того или иного масштаба. Можно выделить три основные категории грид-сетей: относительно простая вычислительная грид-сеть, предоставляющая ресурсы пользователям одной рабочей группы, одного департамента, одного проекта отраслевого предприятия, – групповая грид-сеть (Cluster Grid); грид-сеть, охватывающая несколько групп, работающих над различными проектами в рамках одного отраслевого предприятия, – грид-сеть предприятия (Enterprise Grid); сеть, в которой участвуют несколько независимых предприятий отрасли, предоставляющих друг другу свои ресурсы в соответствии с установленными между ними определенными правилами обмена ресурсами и протоколами взаимодействия, – отраслевая грид-сеть (Branch Grid, отраслевой ГРИД). В принципе возможна организация грид-сети республиканского (национального) масштаба. Грид-сети высокого уровня (национальные, отраслевые, а также грид-сети предприятий, имеющих собственные ЦОД с суперкомпьютерными ресурсами), как правило, имеют операционный и удостоверяющий центры (ОЦ и УЦ). Цели и задачи работы ОЦ: разработка структуры управления грид-сетью и разработка процедур управления грид-сетью; развертывание серверов и служб, обеспечивающих работу сервисов ОЦ и корневых сервисов грид-се-

ти; ведение и учет пользователей и ресурсов грид-сети; осуществление управления грид-сетью.

УЦ грид-сети формирует на основе инфраструктуры открытых ключей цифровые сертификаты сервисов и пользователей грид-сети. УЦ грид-сети может входить в состав ОЦ грид-сети. В основе инфраструктуры открытых ключей (PKI – Public Key Infrastructure) лежит использование криптографической системы с открытым ключом. Закрытый ключ известен только его владельцу, УЦ создает сертификат открытого ключа, удостоверяя этот ключ. Грид-сети нижнего уровня (групповая грид-сеть, грид-сеть предприятия) могут иметь только УЦ для обеспечения безопасного доступа. В принципе возможно делегирование функций УЦ сети более высокого уровня с ОЦ и УЦ.

В рамках программ Союзного государства проводились практические работы по созданию конкретных специализированных АПК-ГРИД. С учетом этого опыта на рис. 5 представлена обобщенная структура АПК-ГРИД.



Рис. 5. Обобщенная структура АПК-ГРИД

Основой структуры АПК-ГРИД являются вычислительные ресурсы с соответствующим системным и прикладным ПО, программное обеспечение промежуточного уровня (ПО ПУ) и грид-сервисы. В случае использования кластерных конфигураций ПО ПУ устанавливается на управляющих узлах

этих кластеров. ПО ПУ обеспечивает, по сути, функционирование всей распределенной сетевой инфраструктуры, включая прозрачную работу приложений в неоднородной сетевой среде. Задача ПО ПУ – создать единую виртуальную среду для выполнения распределенного приложения, не зависящую от сетевых служб, аппаратных платформ, операционных сред и географической удаленности.

Для нормального функционирования распределенной системы ПО ПУ должно как минимум обеспечить идентификацию выполняемой программы; авторизацию пользователя; описание ресурсов; выполнение распределенных алгоритмов; доступ к удаленным данным; распределение ресурсов; определение работоспособности ресурсов и др. Реализация этих функций обеспечивается соответствующими грид-службами (грид-сервисами). Ключевую роль играют корневые грид-сервисы – набор сервисов верхнего уровня, обеспечивающий работу как грид-сети в целом, так и отдельных грид-сервисов. Для предоставления пользователям надежного, безопасного и унифицированного доступа к ресурсам кластерных конфигураций, входящих в состав АПК-ГРИД, в составе грид-сети создаются грид-сайты. Использование системы визуализации позволяет целиком или частично избежать копирования громоздких результатов суперкомпьютерного моделирования на рабочее место пользователя.

Примером практической реализации АПК-ГРИД является аппаратно-программный комплекс распределенной грид-инфраструктуры обработки и предоставления данных дистанционного зондирования Земли (АПК-ГРИД-ДЗЗ). Работы по созданию АПК-ГРИД-ДЗЗ завершены в 2017 г. в рамках программы Союзного государства «Мониторинг-СГ». Целью работы было создание АПК-ГРИД-ДЗЗ и проведение исследований и анализа принципов выполнения прикладных ресурсоемких задач обработки данных ДЗЗ на основе реальных цифровых фотографий поверхности Земли.

Другой пример – создание испытательного полигона (стенда) и совместного российско-белорусского участка ГРИД-сегмента для проведения испытаний программно-аппаратных комплексов, проведения исследований и анализа прикладного ПО для решения геолого-геофизических задач, созданного в 2018 г. в рамках программы Союзного государства «СКИФ-НЕДРА».

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ	4
ВВЕДЕНИЕ	6
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА. РОБОТОТЕХНИКА	10
Вычислительная техника (<i>Бибило П. Н., Парамонов Н. Н.</i>)	12
Суперкомпьютерные и грид-системы (<i>Парамонов Н. Н., Чиж О. П.</i>)	17
Облачные технологии. Облачные вычисления (<i>Венгеров В. Н.</i>)	29
Квантовые вычисления (<i>Могилевцев Д. С.</i>)	34
Робототехника (<i>Прокопович Г. А.</i>)	36
ИНФОРМАТИКА	48
Алгоритмы (<i>Мельников О. И., Шафранский Я. М.</i>)	51
Технология разработки программного обеспечения. Производство программного продукта и услуг (<i>Романов В. И.</i>)	57
Базы данных и системы управления базами данных (<i>Григянец Р. Б.</i>)	64
Искусственные нейронные сети и машинное обучение (<i>Павленко Д. А.</i>)	73
ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ	82
Информационные технологии (<i>Черемисинова Л. Д.</i>)	85
Информационные системы и ресурсы (<i>Григянец Р. Б., Венгеров В. Н.</i>)	93
Большие данные (<i>Венгеров В. Н.</i>)	97
SALS-ERP-технологии и Индустрия 4.0 (<i>Муха Н. П.</i>)	102
Правовое регулирование построения информационного общества (<i>Абламейко М. С.</i>) ..	109
ИНФОРМАТИЗАЦИЯ. ЦИФРОВАЯ ТРАНСФОРМАЦИЯ	114
Информационно-коммуникационная инфраструктура. Компьютерные сети. Интернет (<i>Нозик В. М.</i>)	118
Индустрия программного обеспечения (<i>Черемисинов Д. И.</i>)	128
Цифровой контент (<i>Липницкий С. Ф.</i>)	134
Электронные библиотеки (<i>Григянец Р. Б.</i>)	140
Библиотека в цифровую эпоху (<i>Дрозд О. М., Ахремчик М. П.</i>)	144
Электронное правительство (<i>Науменко Г. Н.</i>)	148
Электронное здравоохранение (<i>Липницкий В. А.</i> , Том II. Э.)	151
Электронное обучение (<i>Воротницкий Ю. И.</i>)	156
Электронная торговля и логистика (<i>Решетняк А. В.</i>)	159
Цифровой банкинг (<i>Ковалев М. М.</i>)	164
Цифровая трансформация в промышленности (<i>Губич Л. В., Муха Н. П.</i>)	171
Интернет вещей. Промышленный Интернет (<i>Решетняк А. В.</i>)	179
Цифровая трансформация ЖКХ. Интеллектуальное здание (<i>Ивашко В. М., Черновец В. Н., Потетенко С. В.</i>)	184

Геоинформатика. Геоинформационные системы (<i>Крючков А. Н.</i>)	191
Информационное общество (<i>Науменко Г. Н.</i>)	203
Цифровая экономика (<i>Енин С. В.</i>)	208
Цифровой регион (<i>Кругликов С. В., Матюшенко Г. П.</i>)	213
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ	224
Современное состояние и актуальные проблемы информационной безопасности (<i>Дмитриев В. А., Степанян А. Б., Максимович Е. П.</i>)	227
Методы и средства защиты информации (<i>Дмитриев В. А., Степанян А. Б., Макси- мович Е. П.</i>)	234
ПРИЛОЖЕНИЕ	240