ФГБОУ ВО

«Сибирский Государственный Университет Телекоммуникаций и Информатики»

Кафедра вычислительных систем

Расчетно-графическое задание

по дисциплине

«Архитектура вычислительных систем»

Вариант 2

Выполнил:

студент группы ИП-814

Краснов И.В.

Проверил:

доцент кафедры ВС

Ефимов А.В.

**Оглавление**

[Постановка задачи 3](#_Toc59879819)

[Общие сведения 4](#_Toc59879820)

[Основные цели создания 5](#_Toc59879821)

[Технические характеристики 6](#_Toc59879822)

[Производительность 7](#_Toc59879823)

[Вычислительные узлы 8](#_Toc59879824)

[Система охлаждения 9](#_Toc59879825)

[Список литературы 10](#_Toc59879826)

# **Постановка задачи**

Выполнить анализ архитектуры супер ВС из списка Top 500. В соответствии с моделью коллектива вычислителей выделить и описать уровни мультиархитектуры супер ВС. В том числе для каждого уровня показать функциональную структуру, сущность вычислителя, топологию сети связей, доступные технологии программирования и область эффективного применения, а также структурные характеристики.

Расчет структурных характеристик (диаметр, средний диаметр, бисекционная пропускная способность) выполнить для одного из уровней мультиархитектуры.

# **Общие сведения**

**Summit** (OLCF-4) — [суперкомпьютер](https://ru.wikipedia.org/wiki/%D0%A1%D1%83%D0%BF%D0%B5%D1%80%D0%BA%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80), разработанный компанией IBM для [Ок-Риджской национальной лаборатории](https://ru.wikipedia.org/wiki/%D0%9D%D0%B0%D1%86%D0%B8%D0%BE%D0%BD%D0%B0%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F_%D0%BB%D0%B0%D0%B1%D0%BE%D1%80%D0%B0%D1%82%D0%BE%D1%80%D0%B8%D1%8F_%D0%9E%D0%BA-%D0%A0%D0%B8%D0%B4%D0%B6). Обладает [вычислительной мощностью](https://ru.wikipedia.org/wiki/%D0%92%D1%8B%D1%87%D0%B8%D1%81%D0%BB%D0%B8%D1%82%D0%B5%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F_%D0%BC%D0%BE%D1%89%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D0%BA%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80%D0%B0) 122,3 [ПФлопс](https://ru.wikipedia.org/wiki/%D0%A4%D0%BB%D0%BE%D0%BF%D1%81), продемонстрированной на тесте [HPL](https://ru.wikipedia.org/wiki/High_Performance_Linpack). Суперкомпьютер Summit являлся самым высокопроизводительным компьютером по состоянию на 2019 год в открытом рейтинге суперкомпьютеров. По состоянию на ноябрь 2020 года занимает 2 место в рейтинге суперкомпьютеров Top 500.

Контракт на $325 млн Министерства Энергетики США на построение суперкомпьютера в 2014 году получили IBM (серверные узлы), Mellanox (межсоединения) и Nvidia (графические ускорители).



(рис. 1) Суперкомпьютер Summit

# **Основные цели создания**

Поскольку некоторые научно-технические задачи, в том числе точное моделирование физических процессов, требуют огромных вычислительных мощностей, их невозможно проводить на одном или даже сотнях обычных компьютеров. Для этого используются суперкомпьютеры, состоящие из тысяч мощных компьютеров с десятками вычислительных ядер в каждом. Они могут за приемлемое время проводить ресурсоемкие вычисления, на которые у обычных компьютеров уходит на несколько порядков большее время.

Суперкомпьютер Summit был создан для Ок-Риджской научной лаборатории. Руководитель IBM Джинни Рометти говорит, что с помощью Summit ученые могут за день обсчитать задачи, которые отняли бы у них «от 27 до 13 000 лет». Огромная вычислительная мощь будет использована для создания новых материалов и лекарств. В ближайшее время компьютер планируют использовать для анализа взрыва сверхновых звезд, маркеров болезни Альцгеймера и исследования рака.

По словам Рометти, дело даже не только в мощности, но и в интеллекте. Ученые разработали Summit, с тем чтобы он стал первым суперкомпьютером, способным решать сложные задачи в области искусственного интеллекта. Его конфигурация сочетает в себе огромную вычислительную мощность, графические процессоры, а также инновации в области памяти и транспортировки данных. Директор Ок-Риджской национальной лаборатории Томас Закария отметил, что Summit может автоматизировать важнейшие этапы процесса научных открытий.

# **Технические характеристики**

Процессоры: 9216 22-ядерных IBM POWER9 22C (2 на узел)

Архитектура процессора: IBM POWER

Графические ускорители: 27,648 NVIDIA Volta V100s (6 на узел)

Количество узлов: 4608

Память узла: 512 ГБ DDR4 + 96 ГБ HBM2

Энергозависимая память узла: 1600 ГБ

Общая память системы: >10 ПБ DDR4 + HBM + энергозависимая память

Межсетевые соединения: Mellanox EDR 100G InfiniBand

Энергопотребление: 15 МВт

Требуемый объём воды для системы охлаждения: 15 150 литров

Занимаемая площадь: 520 м2

# **Производительность**

Summit построила команда в составе представителей Ок-Риджской национальной лаборатории Министерства энергетики США и американских компаний IBM и Nvidia Corporation, и он обладает удивительными характеристиками. Summit развивает вычислительную производительность до 200 петафлопс. Это означает, что компьютер может осуществлять 200 тысяч триллионов операций в секунду, если использовать стандартную меру измерения мощности суперкомпьютеров. В ходе майских [испытаний](https://www.olcf.ornl.gov/2018/06/08/genomics-code-exceeds-exaops-on-summit-supercomputer/) 2018 года Summit заставили анализировать миллионы геномов, и суперкопьютер продемонстрировал быстродействие в 1,88 exaops — миллиард миллиардов операций в секунду.

Основные характеристики производительности:

* Процессор: IBM POWER9 22C
* Достигнутая вычислительная производительность для задач с плавающей точкой: 122.3 ПФлопс
* Теоретический пик производительности для задач с плавающей точкой: ≈ 200 ПФлопс
* HPL\*: 148 600 ТФлопс
* Теоретический пик производительности HPL: 200 795 ТФлопс
* Размер матрицы NMax\*\*: 16 473 600
* HPCG\*\*\*: 2 925,75 ТФлопс

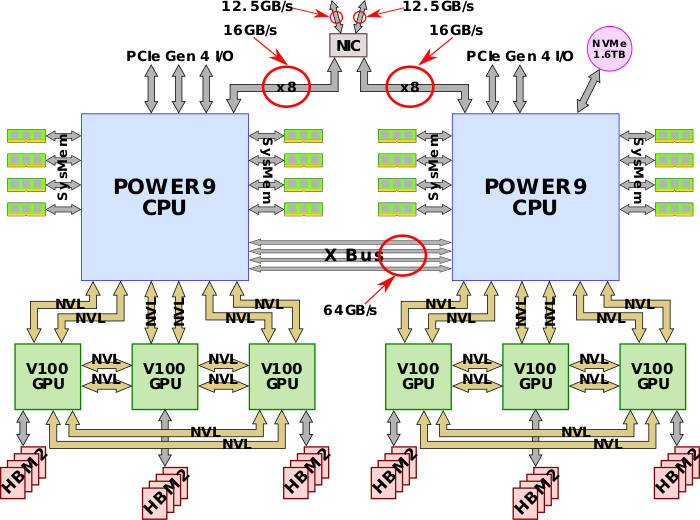
\*High Performance Linpack (HPL) – тест с использованием бенчмарка Linpack, который в процессе работы осуществляет перемножение матриц.

\*\*NMax – максимальный размер матрицы, при котором была достигнута максимальная производительность теста HPL.

\*\*\* High Performance Gradients (HPCG) – тест с использованием бенчмарка HPCG, который в процессе работы осуществляет решение задач класса Big Data и моделирование физических процессов (например, моделирование распределения давлений в потоке жидкости).

# **Уровень вычислительного узла**

Узел состоит из двух процессорных сокетов IBM Power9 и 6 видеокарт NVIDIA Tesla V100 (Volta). К каждому процессору подключается 256 Гб оперативной памяти DDR4 и 3 графических ускорителя Tesla V100. Графические ускорители имеют в своём распоряжении 96 Гб памяти HBM2 на весь узел (16 Гб на ускоритель). Графические ускорители взаимодействуют друг с другом и с процессорами с помощью мостов NVIDIA NVLink (NVL). Процессоры осуществляют взаимодействие через шину X-Bus. У каждого сокета имеется по 4 слота PCIe Gen 4, состоящие из двух x16 (с поддержкой CAPI), одного x8 (также с поддержкой CAPI) и одного слота x4.



(рис. 2) Узел суперкомпьютера IBM Summit

# **Уровень процессора**

Вначале 2018 года IBM представила горизонтально масштабируемое семейство серверов IBM POWER9. Название является аббревиатурой от оптимизации производительности с помощью улучшенного компьютера с сокращенным набором команд (RISC).

Характерные особенности RISC-процессоров:

* Элемент маркированного списка.
* Фиксированная длина машинных инструкций, простой формат команды.
* Специализированные команды для операций с памятью — чтения или записи. Операции вида Read-Modify-Write («прочитать-изменить-записать») отсутствуют. Любые операции «изменить» выполняются только над содержимым регистров (т.н. архитектура load-and-store).
* Большое количество регистров общего назначения (32 и более).
* Отсутствие поддержки операций вида «изменить» над укороченными типами данных — байт, 16-битное слово. Так, например, система команд DEC Alpha содержала только операции над 64-битными словами, и требовала разработки и последующего вызова процедур для выполнения операций над байтами, 16- и 32-битными словами.
* Отсутствие микропрограмм внутри самого процессора.

Преимущества RISC:

* Повышение производительности
* Новый микропроцессор может быть разработан и испытан более быстро.
* Операционная система и прикладные программисты, которые используют инструкции микропроцессора могут легче разрабатывать код с меньшим набором команд.
* Простота RISC обеспечивает большую свободу выбора, как использовать пространство на микропроцессоре.

Основные характеристики процессора IBM POWER9 22C:

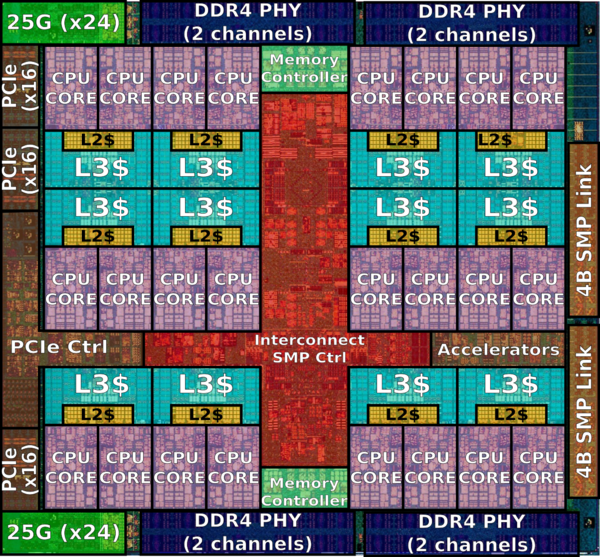
* Количество ядер процессора: 22
* Техпроцесс: 14 нм
* Количество транзисторов: 8 000 000 000
* Частота процессора: 3,1 ГГц
* Кэш L1: 64 КБ на ядро
* Кэш L2: 512 КБ на ядро
* Кэш L3: 120 МБ на чип

Процессор IBM POWER9 выпускается в 4 вариантах:

* Технология Scale-Out. Версия процессора SO позиционируется как решение для двухпроцессорных конфигураций с встроенным контроллером памяти на 8 каналов небуферизованной DDR4.
* Технология Scale-Up. Версия процессора SU (Scale-up) предложены как решение для конфигураций с числом процессоров 4 и более и доступом к буферизованной DDR4 памяти через внечиповый контроллер памяти Centaur.

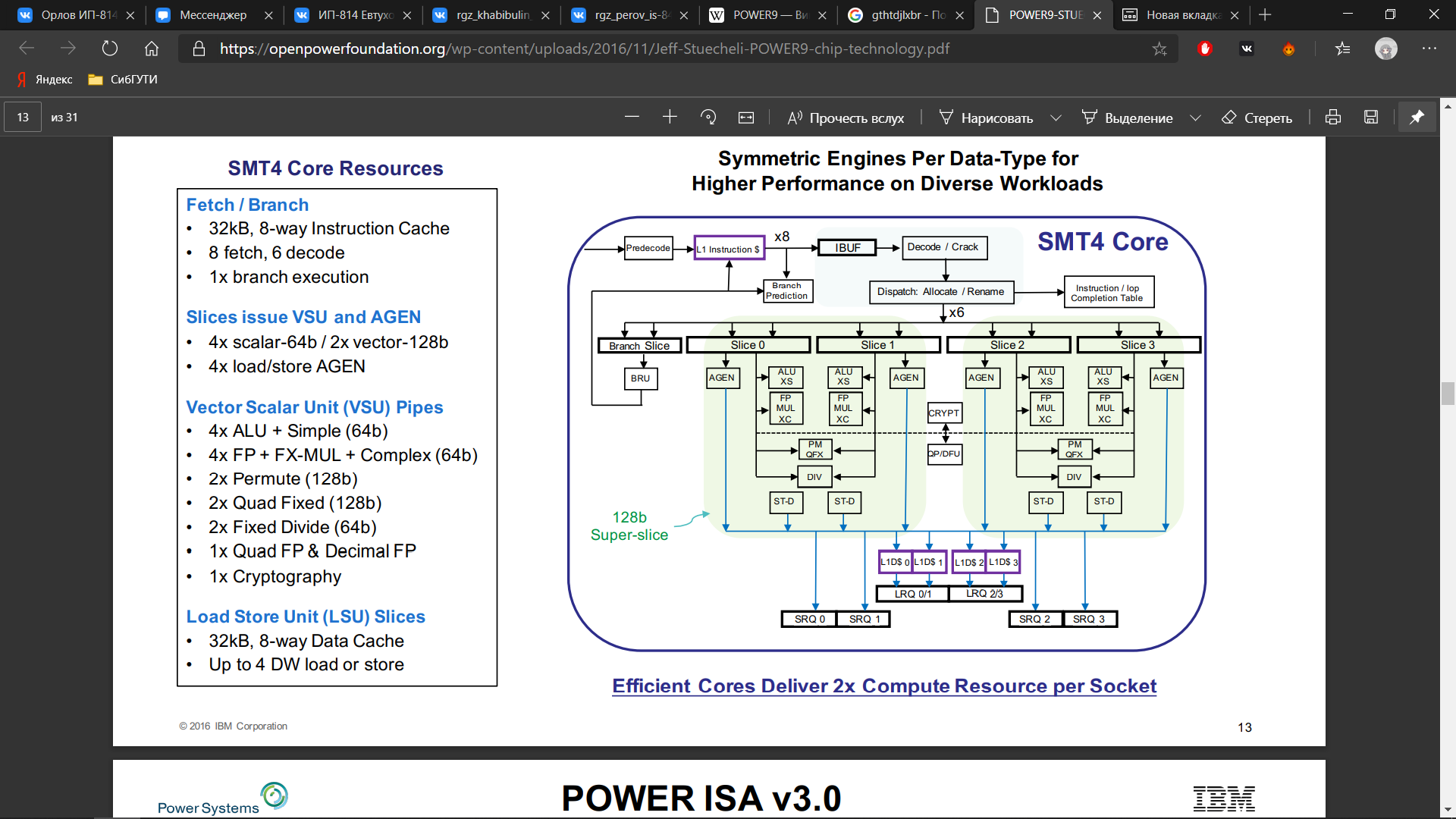
Для каждой технологии доступны два варианта архитектуры:

* До 12 ядер с 8 потоками на ядро(SMT8) – предназначен для платформ с расширенной виртуализацией.
* До 24 ядра с поддержкой 4 потоков на ядро (SMT4) – предназначен для решения обычных задач в среде Linux.



(рис. 3) Кристалл процессора IBM POWER9 (24 ядра)

# **Уровень ядра процессора**



(рис. 4) Архитектура ядра процессора POWER9 с технологией SMT4

Ядро процессора POWER9 - это 64-битная реализация архитектуры набора команд IBM Power (ISA) версии 3.0.

Базовым строительным элементом ядер SMT4 и SMT8 является срез. Срез - это элементарный 64-битный однопоточный обрабатывающий элемент с блоком сохранения загрузки (LSU), целочисленным блоком (ALU) и векторным скалярным блоком (VSU, выполняющим операции SIMD и операции с плавающей точкой). Два фрагмента объединяются в 128-битный суперсрез. Ядра SMT4 и SMT8 содержат одинаковое количество потоков = 96. POWER9 имеет меньшую длину конвейера выборки по сравнению с POWER8, продолжительность вычислений уменьшена на 5 циклов

Инструкций за цикл для SMT4: 64.

# **Файловая система**

Summit использует файловую систему IBM Spectrum Scale. Это не обычная система. Она специально создана для использования графических процессоров чтобы преуспеть в развитии искусственного интеллекта (ИИ). ИИ отличается от традиционных высокопроизводительных вычислительных (HPC) рабочих нагрузок. Большая часть HPC посвящена использованию компьютеров для имитации реального мира с помощью математических моделей. Есть много импортных приложений, от разработки лекарств до прогнозов погоды.

В отличие от этого, программы ИИ обучаются на исторических данных и предсказывают новые результаты по мере передачи данных. В большинстве случаев данные неструктурированные, плотные и разнообразные. Они могут состоять из изображений, файлов, документов и наборов данных. IBM Spectrum Scale была разработана для доставки файлов для одного из самых быстрых суперкомпьютеров в мире, который был специально построен для рабочих нагрузок ИИ с более чем 27 500 графическими процессорами NVIDIA.

Совокупные цифры ошеломляют. 250-петабайтная система хранения данных Summit поставляется кластером из 77 систем хранения данных IMB ESS, которые доставляет 2,5 TB данных. Саммит имеет емкость 30 миллиардов файлов и 30 миллиардов каталогов и сможет создавать файлы со скоростью более 2,6 миллиона I / O файловых операций в секунду. То есть открытие всех книг Библиотеке Конгресса США займёт 10 секунд.

# **Программное обеспечение**

Операционная система:

* Red Hat Enterprise Linux (RHEL) 7.4

MPI\*:

* IBM Spectrum MPI

Компиляторы:

* + - IBM XLC
    - NVCC
    - GCC

Подключаемые математические библиотеки:

* IBM ESSL
* CUBLAS 9.2

Языки программирования:

* C
* C++
* CUDA

\*Message Passing Interface — программный интерфейс для передачи информации, который позволяет обмениваться сообщениями между процессами, выполняющими одну задачу

# **Расчёт структурных характеристик**

Расчет структурных характеристик включает в себя:

1. Расчет диаметра структуры

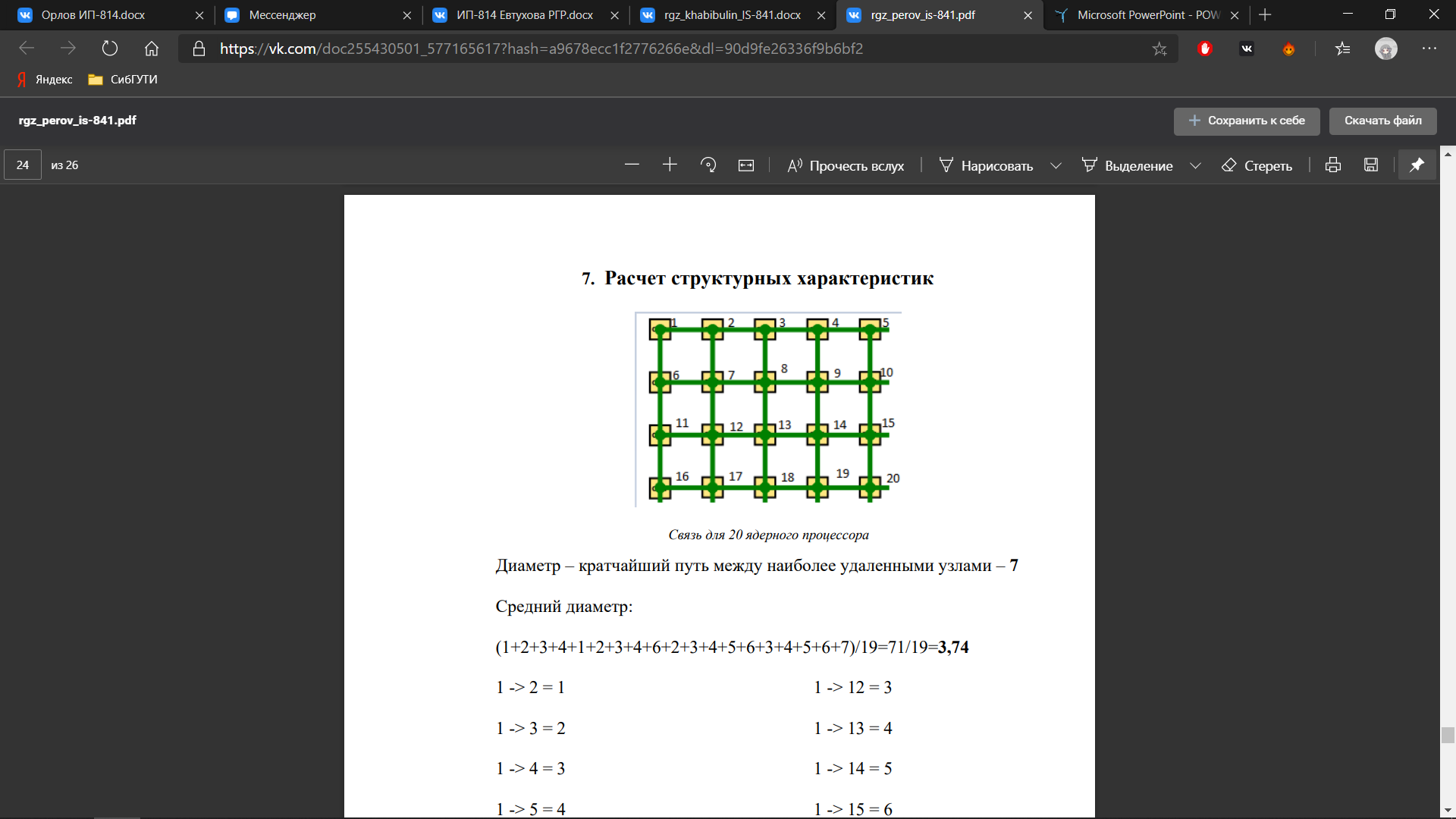
2. Расчет среднего диаметра

3. Расчет бисекционной пропускной способности

Расчет структурных характеристик производится для уровня вычислительного узла.

## **Расчёт диаметра структуры**

В системе установлен 22-ядерный процессор. « ядра зарезервированы под операционную систему, потому для вычисления используется только по 20 ядер каждого процессора.



(рис. 5) Связь для 20-ядерного процессора

𝑑=𝑚𝑎𝑥𝑖𝑗{𝑑𝑖𝑗}

𝑑𝑖𝑗 - расстояние, т.е. минимальное число рёбер, образующих путь из вершины i в вершину j; i, j € {0, 1, ..., N −1};

Ответ: d = 7

## **Расчёт среднего диаметра**

𝑑 = (𝑁−1)−1  nl

nl – число вершин, находящихся на расстоянии l от любой выделенной вершины (однородного) графа.

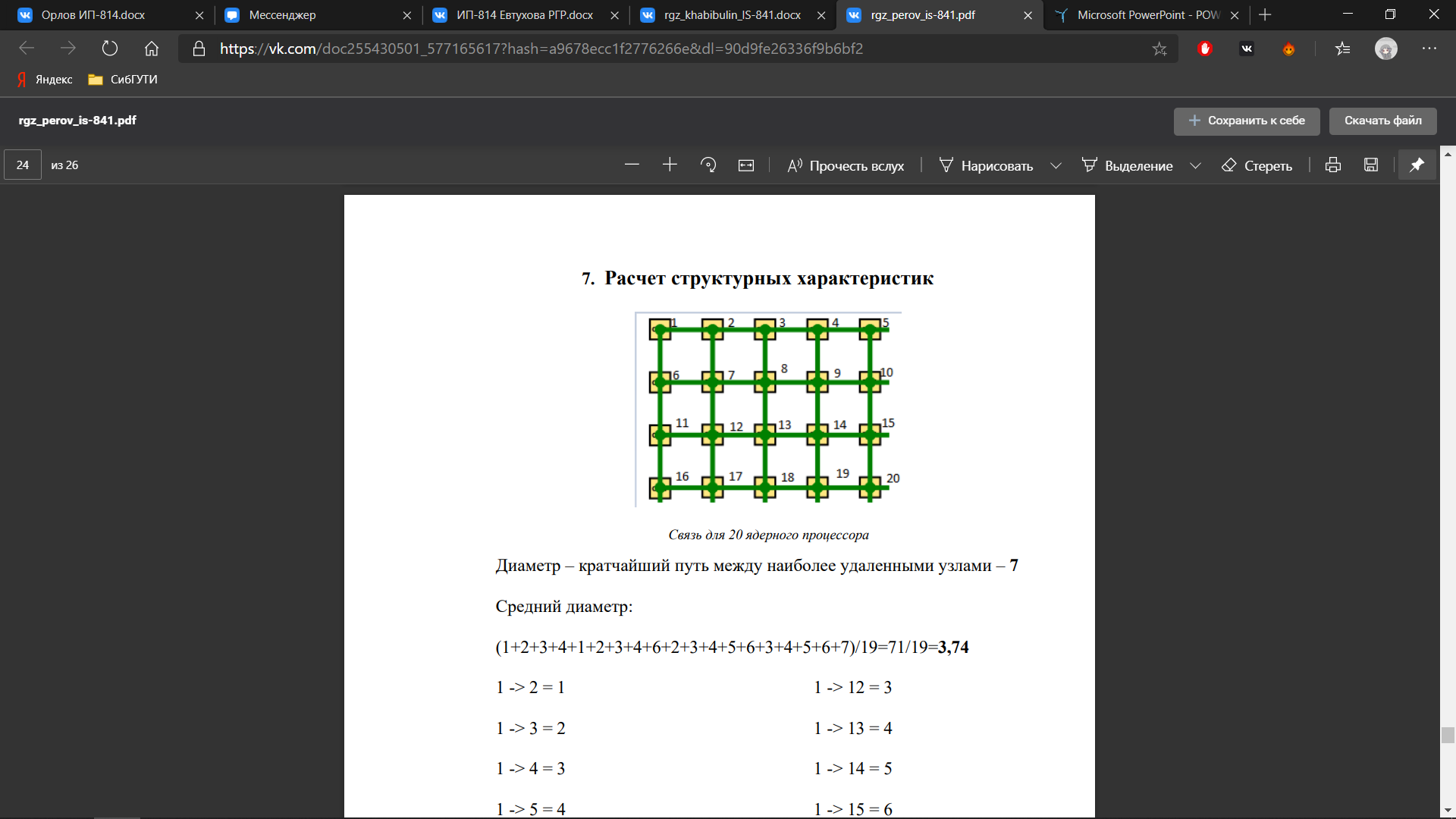
В графе, изображенном на рисунке выделим вершину 0 и рассчитаем на основе нее средний диаметр, тогда

𝑑 = 1/19 \* (1\*2 + 2\*3 + 3\*4 + 4\*4 + 5\*3 + 6\*2 + 7) = 71 / 19 = 3,6842

Ответ: 3,6842

## **Расчёт бисекционной пропускной способности**

Бисекционная пропускная способность – суммарная пропускная способность каналов связи между двумя непересекающимися подмножествами машин системы (для худшего разбиения, минимальное значение).



(рис. 6) Схема 20-ядерного процессора для расчёта бисекционной пропускной способности

Бисекционная пропускная способность равна 4, это минимальное возможное значение для данной схемы

Ответ: 4

# **Список литературы**

1. <https://nplus1.ru/news/2018/06/09/summit>
2. <https://hightech.plus/2018/06/09/superkompyuter-iz-ssha-vdvoe-obognal-kitaiskii>
3. <https://www.olcf.ornl.gov/olcf-resources/compute-systems/summit/>
4. <https://ru.wikipedia.org/wiki/Summit_(%D1%81%D1%83%D0%BF%D0%B5%D1%80%D0%BA%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80)>
5. <https://www.top500.org/system/179397/>
6. <https://servernews.ru/967110>
7. <https://parallel.ru/computers/computers.html>
8. <https://plc.ucoz.net/index/shina_bus_x/0-7>
9. <https://www.nvidia.com/ru-ru/design-visualization/nvlink-bridges/>
10. <https://fuse.wikichip.org/news/1351/ornls-200-petaflops-summit-supercomputer-has-arrived-to-become-worlds-fastest/>
11. <https://mipt.ru/upload/pr/Summit.pdf>
12. <https://en.wikichip.org/wiki/ibm/microarchitectures/power9>
13. <https://openpowerfoundation.org/wp-content/uploads/2016/11/Jeff-Stuecheli-POWER9-chip-technology.pdf>
14. [https://hpc.llnl.gov/training/tutorials/using-lcs-sierra-system#POWER9](https://hpc.llnl.gov/training/tutorials/using-lcs-sierra-system%23POWER9)