

Вопросы к экзамену по курсу «Параллельная обработка больших графов», весенний семестр 2021/2022 уч. года

1. Алгоритмы обхода графов: поиск вширь, поиск вглубь. Сложность алгоритмов.
2. Алгоритмы поиска всех кратчайших путей от заданной вершины: Дейкстры, Беллмана-Форда. Сложность алгоритмов.
3. Алгоритм дельта-степпинг поиска всех кратчайших путей от заданной вершины. Сложность алгоритма.
4. Алгоритмы поиска минимального остовного дерева в графе: Прима, Крускала. Сложность алгоритмов.
5. Алгоритм Борувки поиска минимального остовного дерева в графе. Сложность алгоритма.
6. Организация и оптимизация структуры данных система непересекающихся множеств.
7. Алгоритм GHS построения минимального остовного дерева.
8. Алгоритмы поиска сообществ в графе: лувенский алгоритм, алгоритм label propagation. Сложность алгоритмов.
9. Алгоритмы расчета Betweenness Centrality: наивный, алгоритм Брандеса. Сложность алгоритмов.
10. Программные модели BSP и vertex-centric.
11. Программная модель и модель выполнения технологии параллельного программирования Charm++.
12. Программные модели MapReduce и Spark.
13. Математические основы GraphBLAS: данные, операции.
14. Виды синтетических графов: Random Uniform, small-world, RMAT, LFR, SSCA2.
15. Основные проблемы анализа больших графов.
16. Способы представления графов: списки смежности, матрица смежности. Форматы представления разреженных матриц: CRS, Coordinate list, JDS.
17. Форматы представления разреженных матриц ELLPACK, SELLPACK.
18. Особенности архитектуры современного вычислительного узла, влияющие на задач обработки больших графов: NUMA, многоядерность, длинные вектора, иерархия памяти.
19. Принципы организации динамической памяти.
20. Проблемы и подходы к решению задач обработки больших графов в рамках одного вычислительного узла.
21. Особенности архитектуры современной вычислительной системы с распределенной памятью: высокоскоростная коммуникационная сеть, ее топология, характеристики (задержка, пропускная способность, темп выдачи сообщений, бисекционная пропускная способность). Технология RDMA.
22. Топологии высокоскоростных коммуникационных сетей: жирное дерево, тор, Flattened Butterfly, Dragonfly.
23. Односторонняя и двусторонняя модели передачи сообщений.
24. Протоколы передачи сообщений Eager и Rendezvous при реализации библиотеки MPI.
25. Проблемы и подходы к решению задач обработки больших графов в рамках распределенной вычислительной системы: распределение данных, организация внутриузлового параллелизма, основной коммуникационный паттерн, сжатие сообщений, балансировка нагрузки.