

67-я студенческая научная конференция ЮУрГУ

Реализация теста Graph500 для параллельной СУБД PargreSQL

Сафонов А. Ю. группа ВМИ-456

Научный руководитель:
кандидат физ.-мат. наук Цымблер М. Л.


[Home](#)
[Complete Results](#)
[Benchmarks](#)
[Green Graph 500](#)
[Log In](#)

The Graph 500 List

Top 10 (November 2013)

Rank	Machine
1	DOE/NNSA/LLNL Sequoia - Lawrence Livermore National Laboratory (65536 nodes, 1048576 cores)
2	DOE/SC/Argonne National Laboratory Mira - Argonne National Laboratory (49152 nodes, 786432 cores)
3	JUQUEEN - Forschungszentrum Juelich (FZJ) (16384 nodes, 262144 cores)
4	K computer - RIKEN Advanced Institute for Computational Science (AICS) (65536 nodes, 524288 cores)
5	Fermi - CINECA (8192 nodes, 131072 cores)
	Tianhe-2

November 2013

No.	Rank	Machine	Installation Site	Number of nodes	Number of cores	Problem scale	GTEPS ▼
1	1	DOE/NNSA/LLNL Sequoia (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Lawrence Livermore National Laboratory	65536	1048576	40	15363
2	2	DOE/SC/Argonne National Laboratory Mira (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Argonne National Laboratory	49152	786432	40	14328
3	3	JUQUEEN (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Forschungszentrum Juelich (FZJ)	16384	262144	38	5848
4	4	K computer (Fujitsu - Custom supercomputer)	RIKEN Advanced Institute for Computational Science (AICS)	65536	524288	40	5524.12
5	5	Fermi (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	CINECA	8192	131072	37	2567
6	6	Tianhe-2 (MilkyWay-2) (National University of Defense Technology - MPP)	Changsha, China	8192	196608	36	2061.48
		Turing (IBM -					

Цель и задачи исследования

- Цель

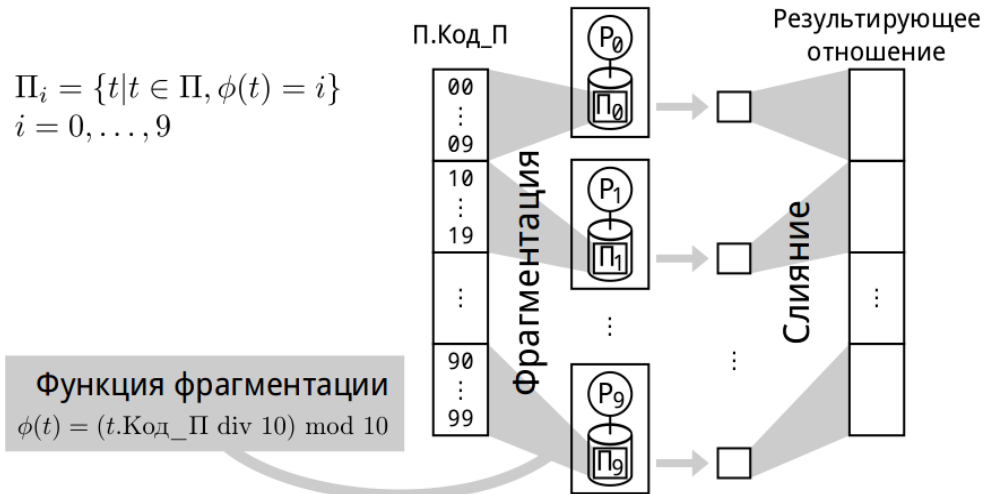
- Оценка эффективности параллельной СУБД PargreSQL на приложениях с интенсивной обработкой данных с помощью сравнительного теста Graph500

- Задачи

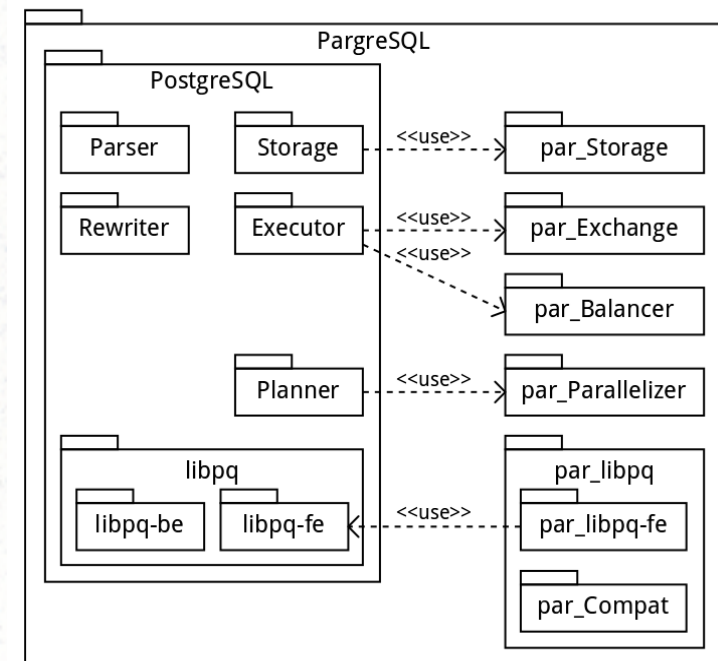
- Изучить теоретический базис параллельных систем баз данных, архитектуру параллельной СУБД PargreSQL и спецификацию теста Graph500
- Разработать схему базы данных для хранения графа и промежуточных данных в соответствии с спецификацией теста Graph500 а также реализовать способ фрагментации, обеспечивающий эффективную реализацию теста
- Выполнить проектирование и разработку алгоритмов на языке SQL, реализующих тест Graph500 для параллельной СУБД PargreSQL
- Провести вычислительные эксперименты на суперкомпьютере “Торнадо ЮурГУ”, исследующие эффективность параллельной СУБД PargreSQL на тесте Graph500

Параллельная СУБД PargreSQL

- Фрагментный параллелизм



- Архитектура PargreSQL

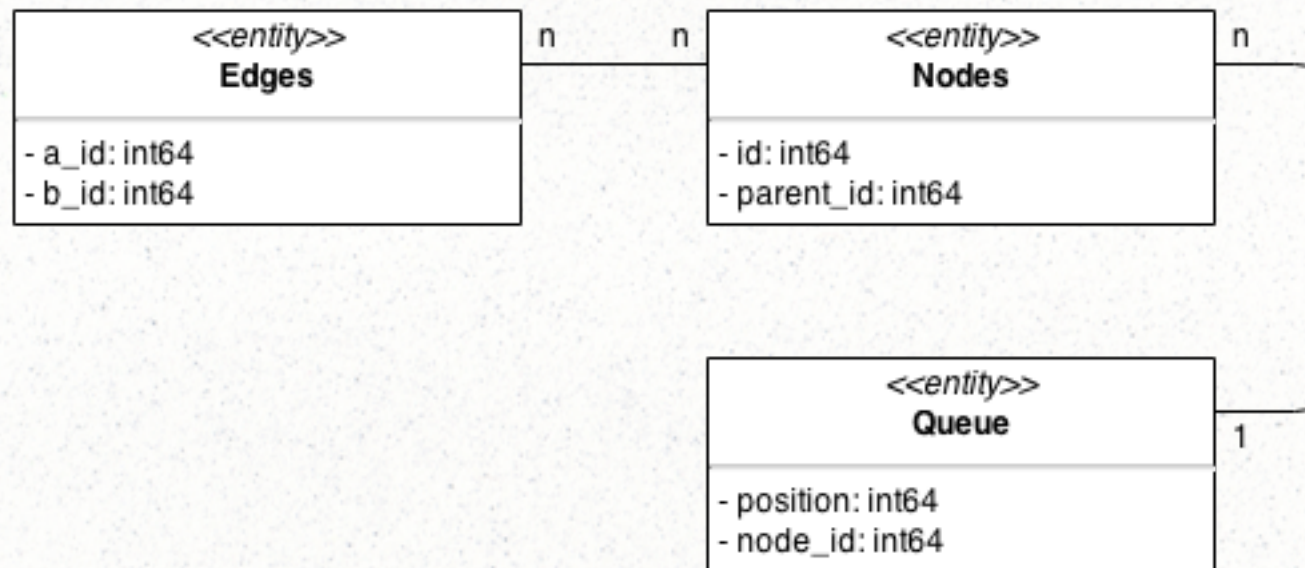


Алгоритм работы теста Graph500

Тест Graph500



Схема хранения данных



Алгоритм поиска в ширину

```
BFS(start_node, goal_node) {  
  for(all nodes i) visited[i] = false; // изначально список посещённых узлов пуст  
  queue.push(start_node);              // начиная с узла-источника  
  visited[start_node] = true;  
  while(! queue.empty() ) {            // пока очередь не пуста  
    node = queue.pop();                 // извлечь первый элемент в очереди  
    if(node == goal_node) {  
      return true;                     // проверить, не является ли текущий узел целевым  
    }  
    foreach(child in expand(node)) {    // все преемники текущего узла, ...  
      if(visited[child] == false) {    // ... которые ещё не были посещены ...  
        queue.push(child);             // ... добавить в конец очереди...  
        visited[child] = true;         // ... и пометить как посещённые  
      }  
    }  
  }  
  return false;                        // Целевой узел недостижим  
}
```


Реализация на SQL поиска в ширину в графе

Текущие результаты

- Изучен теоретический базис параллельных систем баз данных, архитектуру параллельной СУБД PargreSQL и спецификацию теста Graph500
- Разработана схема базы данных для хранения графа и промежуточных данных в соответствии с спецификацией теста Graph500, а также реализован эффективный способ фрагментации
- Выполнено проектирование и разработка алгоритмов на языке SQL, реализующих тест Graph500 для параллельной СУБД PargreSQL