

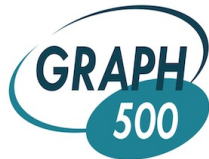
Graph 500

Renzo Gonzalo Gómez Diaz

Instituto de Matemática e Estatística
Universidade de São Paulo

17 de junho, 2015

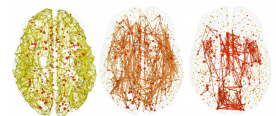
- Ranking de supercomputadoras (similar ao Top500).
- Focado em aplicações “*data-intensive*”.
- Desde 2010, aparece duas vezes por ano (Junho e Novembro).



- Iniciativa proposta por Richard Murphy e outros pesquisadores (Sandia Lab., E.E.U.U.).
- Primeira lista publicada em Novembro de 2010, na ACM/IEEE Supercomputing Conference.
- Comitê de direção formado por especialistas de diversas instituições (Sandia Lab., Google, NVIDIA, Georgia Tech, etc.)

Surgimento de problemas reais de natureza combinatória,

- Análise de dados em redes sociais (grafo praticamente ilimitado).
- Redes simbólicas (o cérebro humano contém 25 bill. de neurônios, cada um com 7K conexões)



Esses problemas exibem as seguintes características,

- Usa operações com inteiros.

Esses problemas exibem as seguintes características,

- Usa operações com inteiros.
- Baixa localidade espacial e temporal.

Esses problemas exibem as seguintes características,

- Usa operações com inteiros.
- Baixa localidade espacial e temporal.
- Grande processamento de dados.

Esses problemas exibem as seguintes características,

- Usa operações com inteiros.
- Baixa localidade espacial e temporal.
- Grande processamento de dados.
- Acesso a memória não segue um padrão estruturado.

Atualmente, o benchmark está composto por um teste de busca em largura concorrente em grafos (“Search”).

Alguns Parâmetros

- Escala: $\log_2(N)$, onde $N =$ núm. nós.
- Factor de arestas: Núm. arestas / Num. nós.
- TEPS: Núm. de arestas / tempo BFS.

Dependendo do tamanho do grafo classificamos a instância do benchmark como

Classe	Escala	Factor de arestas	Tamanho (Aprox. TB)
Brinquedo	26	16	0.0172
Mini	29	16	0.1374
Pequeno	32	16	1.0995
Médio	36	16	17.5922
Grande	39	16	140.7375
Enorme	42	16	1125.8999

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	98304	1572864	41	23751
2	2	K computer (Fujitsu - Custom supercomputer)	RIKEN Advanced Institute for Computational Science (AICS)	82944	663552	40	19585.2
3	3	DOE/SC/Argonne National Laboratory Mira (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Argonne National Laboratory	49152	786432	40	14982
4	4	JUQUEEN (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Forschungszentrum Juelich (FZJ)	16384	262144	38	5848
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
7	7	Turing (IBM - BlueGene/Q, Power BQC 16C 1.60GHz)	CNRS/IDRIS-GENCI	4096	65536	36	1427

Comparação com Top500

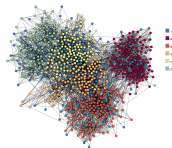
Graph500	Top500	Máquina
1	3	Sequoia (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)
2	4	K computer (Fujitsu - Custom supercomputer)
3	5	Mira (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)
4	8	JUQUEEN (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)
5	23	Fermi (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)
6	1	Tianhe-2 (MilkyWay-2) (National University of Defense Technology - MPP)
7	42	Turing (IBM - BlueGene/Q, Power BQC 16C 1.60GHz)

No futuro pretende-se adicionar mais problemas ao benchmark.

- **Otimização:** Caminhos mínimos.



- **Orientado a arestas:** Conjuntos independentes, Clustering.





Graph 500 site.

<http://www.graph500.org/>.

Accessed: 16-06-2015.



R.C. Murphy.

On the effects of memory latency and bandwidth on supercomputer application performance.

2013 IEEE International Symposium on Workload Characterization (IISWC), 0:35–43, 2007.



R.C. Murphy and P.M. Kogge.

On the memory access patterns of supercomputer applications: Benchmark selection and its implications.

Computers, IEEE Transactions on, 56(7):937–945, July 2007.



R.C. Murphy, K.B. Wheeler, W. Barret, B, and J.A. Ang.

Introducing the graph 500.

Cray User's Group (CUG), May 2010.

Obrigado!!
Gracias!!

