References

1. Large synoptic survey. http://www.lsst.org/
2. Hey, T., Tansley, S., Tolle, K. (eds.): The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research (2009)
3. Bell, G., Gray, J., Szalay, A.S.: Petascale computational systems. IEEE Comput. **39**(1), 110–112 (2006)
4. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H.: Big data: the next frontier for innovation, competition, and productivity. Technical report 1999–66 (2011)
5. Dean, J., Ghemawa, S.: MapReduce: simplified data processing on large clusters. In: OSDI (2004)
6. Yang, H., Dasdan, A., Hsiao, R., Parker, D.: Map-reduce-merge: simplified relational data processing on large clusters. In *SIGMOD* (2007)
7. Stonebraker, M.: The case for shared nothing. IEEE Database Eng. Bull. **9**(1), 4–9 (1986)
8. White, T.: Hadoop: The Definitive Guide. O'Reilly Media (2012)
9. Sakr, S., Liu, A., Fayoumi, A.G.: The family of mapreduce and large-scale data processing systems. ACM Comput. Surv. **46**(1) (2013)
10. Sakr, S., Pardede, E. (eds.): Graph Data Management: Techniques and Applications. IGI Global (2011)
11. Khan, A., Elnikety, S.: Systems for big-graphs. PVLDB **7**(13), 1709–1710 (2014)
12. O'reilly, T.: What is web 2.0: design patterns and business models for the next generation of software. Commun. Strateg. (1):17 (2007)
13. Scott, J.: Social Network Analysis. Sage (2012)
14. Linden, G., Smith, B., York, J.: Amazon. com recommendations: item-to-item collaborative filtering. IEEE Internet Comput. **7**(1), 76–80 (2003)
15. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Wiener, J.: Graph structure in the web. Comput. Netw. **33**(1), 309–320 (2000)
16. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: bringing order to the web (1999)
17. Trochim, W.M.K., Donnelly, J.P.: Research methods knowledge base (2001)
18. Bizer, C., Heath, T., Berners-Lee, T.: Linked data-the story so far. In: Semantic Services, Interoperability and Web Applications: Emerging Concepts, pp. 205–227 (2009)

S. Sakr et al., *Large-Scale Graph Processing Using Apache Giraph*,
DOI 10.1007/978-3-319-47431-1

19. Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.: Dbpedia: A Nucleus for a Web of Open Data. Springer, Berlin (2007)
20. Urbani, J., Kotoulas, S., Oren, E., Van Harmelen, F.: Scalable Distributed Reasoning Using Mapreduce. Springer (2009)
21. Steenstra, J., Gantman, A., Taylor, K., Chen, L.: Location based service (IBS) system and method for targeted advertising. US Patent App. 10/931, 309 (2004)
22. Caragliu, A., Del Bo, C., Nijkamp, P.: Smart cities in Europe. J. Urban Technol. **18**(2), 65–82 (2011)
23. Zheng, Y., Chen, Y., Xie, X., Ma, W.-Y.: Geolife2.0: a location-based social networking service. In: Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09, pp. 357–358. IEEE (2009)
24. Trinajstic, N., et al.: Chemical Graph Theory. CRC Press (1992)
25. Jeong, H., Mason, S.P., Barabási, A.-L., Oltvai, Z.N.: Lethality and centrality in protein networks. Nature **411**(6833), 41–42 (2001)
26. Robinson, I., Webber, J., Eifrem, E.: Graph Databases: New Opportunities for Connected Data. O'Reilly Media (2015)
27. Partner, J., Vukotic, A., Watt, N., Abedrabbo, T., Fox, D.: Neo4j in Action. Manning Publications Company (2014)
28. Martinez-Bazan, N., Gomez-Villamor, S., Escale-Claveras, F.: Dex: A high-performance graph database management system. In: 2011 IEEE 27th International Conference on Data Engineering Workshops (ICDEW), pp. 124–127. IEEE (2011)
29. Iordanov, B.: Hypergraphdb: a generalized graph database. In: Web-Age Information Management, pp. 25–36. Springer (2010)
30. Bu, Y., Howe, B., Balazinska, M., Ernst, M.D.: The HaLoop approach to large-scale iterative data analysis. VLDB J. **21**(2) (2012)
31. Ekanayake, J., Li, H., Zhang, B., Gunarathne, T., Bae, S.-H., Qiu, J., Fox, G.: Twister: a runtime for iterative MapReduce. In: HPDC (2010)
32. Zhang, Y., Gao, Q., Gao, L., Wang, C.: iMapReduce: a distributed computing framework for iterative computation. J. Grid Comput. **10**(1) (2012)
33. Chen, R., Weng, X., He, B., Yang, M.: Large graph processing in the cloud. In: SIGMOD (2010)
34. Kang, U., Tsourakakis, C.E., Faloutsos, C.: PEGASUS: a peta-scale graph mining system. In: ICDM (2009)
35. Kang, U., Tong, H., Sun, J., Lin, C.-Y., Faloutsos, C.: GBASE: a scalable and general graph management system. In: KDD (2011)
36. Kang, U., Tsourakakis, C.E., Faloutsos, C.: PEGASUS: mining peta-scale graphs. Knowl. Inf. Syst. **27**(2) (2011)
37. Kang, U., Meeder, B., Faloutsos, C.: Spectral analysis for billion-scale graphs discoveries and Implementation. In: PAKDD (2011)
38. Valiant, L.G.: A bridging model for parallel computation. CACM **33**(8) (1990)
39. Hewitt, C., Bishop, P., Steiger, R.: A universal modular ACTOR formalism for artificial intelligence. In: IJCAI, pp. 235–245 (1973)
40. Malewicz, G., Austern, M.H., Bik, A.J.C., Dehnert, J.C., Horn, I., Leiser, N., Czajkowski, G.: Pregel: a system for large-scale graph processing. In: SIGMOD (2010)
41. Salihoglu, S., Widom, J.: GPS: a graph processing system. In: SSDBM (2013)
42. Khayyat, Z., Awara, K., Jamjoom, H., Kalnis, P.: Mizan: Optimizing Graph Mining in Large Parallel Systems
43. Khayyat, Z., et al.: Mizan: a system for dynamic load balancing in large-scale graph processing. In: EuroSys (2013)
44. Bu, Y., Borkar, V.R., Jia, J., Carey, M.J., Condie, T.: Pregelix: Big(ger) graph analytics on a dataflow engine. PVLDB **8**(2) (2014)

45. Borkar, V.R., Carey, M.J., Grover, R., Onose, N., Vernica, R.: Hyracks: A flexible and extensible foundation for data-intensive computing. In: ICDE (2011)
46. Tian, Y., Balmin, A., Corsten, S.A., Tatikonda, S., McPherson, J.: From "Think Like a Vertex" to "Think Like a Graph". PVLDB **7**(3) (2013)
47. Brin, S., Page, L.: Reprint of: the anatomy of a large-scale hypertextual web search engine. Comput. Netw. **56**(18), 3825–3833 (2012)
48. Salihoglu, S., Widom, J.: Optimizing graph algorithms on pregel-like systems. Proc. VLDB Endow. **7**(7), 577–588 (2014)
49. Tasci, S., Demirbas, M.: Giraphx: parallel yet serializable large-scale graph processing. In: Proceedings of the 19th International Conference on Parallel Processing. Euro-Par'13, pp. 458–469. Springer, Berlin (2013)
50. Wikipedia. Exponential backoff—wikipedia, the free encyclopedia (2015). Accessed 13 Nov 2015
51. Wikipedia. Mean time between failures—wikipedia, the free encyclopedia (2015). Accessed 13 Nov 2015
52. Zaharia, M., Chowdhury, M., Franklin, M.J., Shenker, S., Stoica, I.: Spark: cluster computing with working sets. In: HotCloud (2010)
53. Shvachko, K., Kuang, H., Radia, S., Chansler, R.: The Hadoop distributed file system. In: MSST (2010)
54. Hindman, B., Konwinski, A., Zaharia, M., Ghodsi, A., Joseph, A.D., Katz, R.H., Shenker, S., Stoica, I.: Mesos: a platform for fine-grained resource sharing in the data center. In: NSDI (2011)
55. Vavilapalli, V.K., et al.: Apache hadoop YARN: yet another resource negotiator. In: SOCC (2013)
56. Armbrust, M., Xin, R.S., Lian, C., Huai, Y., Liu, D., Bradley, J.K., Meng, X., Kaftan, T., Franklin, M.J., Ghodsi, A., Zaharia, M.: Spark SQL: relational data processing in spark. In: SIGMOD (2015)
57. Sparks, E.R., Talwalkar, A., Smith, V., Kottalam, J., Pan, X., Gonzalez, J.E., Franklin, M.J., Jordan, M.I., Kraska, T.: MLI: an API for distributed machine learning. In: ICDM (2013)
58. Gonzalez, J.E., Xin, R.S., Dave, A., Crankshaw, D., Franklin, M.J., Stoica, I.: GraphX: graph processing in a distributed dataflow framework. In: OSDI (2014)
59. Low, Y., et al.: Distributed GraphLab: a Framework for machine learning in the cloud. PVLDB **5**(8) (2012)
60. Gonzalez, J.E., Low, Y., Gu, H., Bickson, D., Guestrin, C.: PowerGraph: distributed graph-parallel computation on natural graphs. In: OSDI (2012)
61. Kyrola, A., Blelloch, G.E., Guestrin, C.: GraphChi: large-scale graph computation on just a PC. In: OSDI (2012)