Weekly Report – 01

Xiufeng Liu

University of Waterloo, CA
xiufeng.liu@uwaterloo.ca

1 This Week

2 Next Week

3 In-database Analytics

We do the benchmark study for the birt model [1], which uses three linear regression lines to fit the relation between electricity consumption and the weather temperature. Using the three-line model, we can de-aggregate the energy consumption of a customer into five usage categories, including the base load, activity load, lowest external temperature at which a/c is used, the heating season gradient, and cooling season gradient.

The benchmark is conducted on TOSHIBA laptop with a Intel quad CPU, 8GB RAM, and installed with Ubuntu 12.04.

Table 1 shows the benchmark results using the in-database analytics technologies, KDB+ and Madlib+PostgreSQL. The program for birt model is written by `Q` programming language and `plpgsql` which are naively supported by KDB+ and PostgreSQL respectively. We scale up the size of household data from 5,000 to 25,000, and measure the time. The time only consists of the quantile (10% and 90%), doing linear regression, and find the optimal connecting points of the three lines, while the time of loading data into the databases is not included.

Table 1. The running time of birt model, seconds

No. of households	5,000	10,000	15,000	20,000	25,000
Madlib+PostgreSQL	1,043	1,481	1,798	2,615	3,494
KDB+	12	46	109	172	330

References

1. Birt, Benjamin J., et al. "Disaggregating categories of electrical energy end-use from whole-house hourly data." Energy and Buildings 50 (2012): 93-102.
2. Ten Million Meters Scalable to One Hundred Million Meters for Five Billion Daily Meter Readings. Sept. 2011.
3. J. Yang, Y. Zhai, D. Xu, et al. "SMO Algorithm applied in time series model building and forecast". In *Proc. of ICMLC*, 2007:2395-2400, 2012.
4. P. G. Brown. "Overview of SciDB: Large Scale Array Storage, Processing and Analysis". In Proc. of SIGMOD, pp. 963–968, 2010.
5. SciDB User's Guide. `http://scidb.org/HTMLmanual/13.3/scidb_ug/index.html`
6. SciDB `http://www.scidb.org/` as of 2013-12-07.
7. SciDBR. `https://github.com/Paradigm4/SciDBR` as of 2013-12-07.
8. KDB+. `www.kx.com` as of 2013-12-07.
9. KDB+ Customers. `http://kx.com/end-user-customers.php`