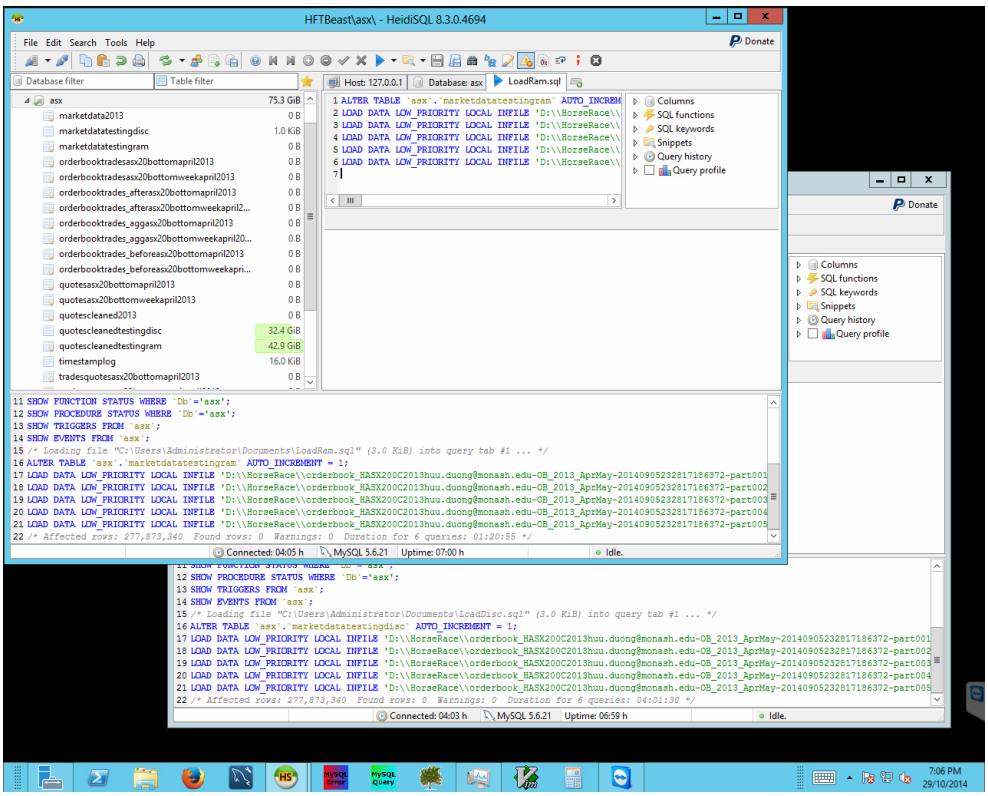
Stock Price Impact Models Study on R@CMon Phase 2

1 December 2014
Categories: MeRC
Tags: All Stories, Market Impact, Remote Desktop, Research Stories, Trading Costs, Using the Cloud, Windows

Paul Lajbcygier, Associate Professor from the Faculty of Business and Economics, Monash University is studying one of the important changes that affects the cost of trading in financial markets. This change relates to the effects of trading to prices, known as *"price impact"*, which is brought by wide propagation of algorithmic and high frequency trading and augmented by technological and computational advances. Professor Lajbcygier's group has recently published new results supported by R@CMon infrastructure and application migration activities, providing new insights into the trading behaviour of so-called "Flash Boys".

This study uses datasets licensed from Sirca and represents stocks in the S&P/ASX 200 index from year range 2000 to 2014. These datasets are pre-processed using Pentaho and later ingested into relational databases for detailed analysis using advanced queries. Two NeCTAR instances on R@CMon have been used initially in the early stages of the study. One of the instances is used as the *processing* engine where Pentaho and Microsoft Visual Studio 2012 are installed for pre-processing and post-processing tasks. The second instance is configured as the *database* server where the extraction queries are executed. Persistent volume storage is used to store reference datasets, pre-processed input files and extracted results. A VicNode merit application for research data storage allocation has been submitted to support the computational access to the preprocessed data supporting the analysis workflow running on the NeCTAR Research Cloud.



Ingestion of pre-processed data into the database running on the high-memory instance, for analysis.

Initially econometric analyses were done on just the lowest two groups of stocks in the S&P/ASX 200 index. Some performance hiccups were encountered when processing higher frequency groups in the index – some of the extraction queries, which require a significant amount of memory, would not complete when run on the exponentially higher stock groups. The release of R@CMon Phase 2 provided the analysis workflow the capability to attack the higher stock groups using a high-memory instance, instantiated on the new *"specialist"* kit. Parallel extraction queries are now running on this instance (close to 100% utilisation) to traverse the remaining stock groups from year range 2000 to 2014.

A recent paper by Manh Pham, Huu Nhan Duong and Paul Lajbcygier, entitled, *"A Comparison of the Forecasting Ability of Immediate Price Impact Models"* has been accepted for the *"1st Conference on Recent Developments in Financial Econometrics and Applications"*. This paper highlights the results of the examination of the lowest two groups of the S&P/ASX 200 index, i.e., just the initial results. Future research and publications include examination of the upper group of the index based on the latest reference data as they come available and analysis of other price impact models.

This is an excellent example of novel research empowered by specialist infrastructure, and a clear win for a build-it-yourself cloud (you can't get a 920GB instance from AWS). The researchers are able to use existing and well-understood computational methods, i.e., relational databases, but at much greater capacity than normally available. This has the effect of speeding up initial exploratory work and discovery. Future work may investigate the use of contemporary data-intensive frameworks such as Hadoop + Hive for even larger analyses.

This article can also be found, published created commons here [1].

1. Revote, Jerico; Bethwaite, Blair; Quenette, Steve; Aung, Swe Win (2014): Stock Price Impact Models Study on R@CMon. Monash University. Online resource. https://doi.org/10.26180/5f716a4f008fc