Good morning, everyone. I’m appreciated to be here today. Now I’ll talk about our project: Financial Data Analysis and Monitoring Based on Spark and ELK.

First, look at this question: What differentiates today’s thriving companies? Data.

As a data management team of Prime Broker, sometimes we don’t know clearly for data the business really care. We can see there are some duplicated data to save day by day, which looks like unreasonable, but we are not sure it’s meaningful for business. If we have a tool to analyze data and its usage, we can better understand and optimize database. Because we have some senior developers who are familiar with business data flow and have database experience to analyze data. So we created this analysis tool together. This tool can provide prospective report, high-quality data analysis service for manager and data analyst, real-time data monitoring for support team.

OK. Let’s look at the architecture of the project. From this slide, we can see 2 lines from data producers to business.

First, take a look at the line above. We had created a simple and easy analysis tool, by combing the popular Elasticsearch, Logstash and Kibana.

Next, the line below can be divided into 4 layers: data collection, data transmission, data processing and data storage.

* Flume collects data in different scenarios.
* Kafka, as a cache, can receive data from flume.
* Spark, as a data processing engine, pulls data from Kafka cluster and process in real time. It also saves data processed to Mongo Database.
* Mongo Database stores data in real-time. All of data is also synchronized to the Elasticsearch for subsequent search analysis.

Now, you might have a question: why we have 2 lines? Why?

These 2 technical lines correspond to different application scenarios.

You know, ELK is easy to use. It can quickly analyze log data through configuration. QA and other non-developers can also skilled to use it.

On the other hand, the line below is only available to developers. But it has high-availability, high-security and real-time. Kafka can cache data without causing data loss. It can be applied to different scenarios, not only log analysis but also any other data. This is why we have 2 lines.

So this project will help you seize data dividend. You will be a data hero.

Finally, talk is too cheap, show the DEMO.

Thank you for your attention. I’d be glad to answer any question you might have.

Dear Lee:

I have pleasure in inviting you to attend our special conference to be held at the Clifton Hotel, London on Tuesday/Wednesday 8/9 October 2016.

This intensive, practical conference for professional secretaries aims to:

* Increase your managerial and office productivity
* Improve your communication skills
* Bring you up to data with the latest technology and techniques.
* Enable networking with secretaries

The seminar is power-packed with a distinguished panel of professional speakers who will give expert advise on many useful topics. A programme is enclosed giving full details of this seminar, which know you will not want to miss.

If you would like to join us, please complete the enclosed registration form and return it to me before 30 June with your fee of $50 per person.

I look forward to seeing you again at this exciting conference,

Yours sincerely