Lessons from Adopting Cloud-like Architectures in Real-life Financial Applications

Bernard Lee, PhD, CFA
Founder and Chief Executive, HedgeSPA Pte. Ltd.
(Hedge Funds and Sophisticated Products Advisors)

at the 5th International Open Cirrus Summit

Hosted by the Russian Academy of Sciences

June 01 – 03, 2011, Moscow, Russia

Outline

- Introduction
- Platform Design
- Lessons learned
- Future challenges
- Conclusion

Introduction

Who We Are?

- Funded by the Singapore Government to strengthen asset management infrastructure in Singapore.
- Front-to-back platform, doing the plumbing to cater to any "2 traders with a Bloomberg"
- Basic value proposition: 'Democratize' the access to high end financial analytics platform

Why Using Clouds?

- High cost of sophisticated financial analytics platform
  - → Previously, only major institutional investors can afford these platforms
  - → Primarily used to create artificial barrier to market entry
- Advanced calculations adopted by our platform
  - → Scalability is key to potential success in business model
  - → To on-line users, computational time of each iteration needs to be cut from hours to minutes and even seconds via the use of massively parallel computing

Platform Design

Key Functionalities:

- Portfolio Construction
- What-if Analysis
- Portfolio Fine-tuning
- Risk and Performance Reporting

Unique Features:

- Customizable analytics for specific market analyses and trading strategies
- Optimize analytical algorithms using MPI deployed on a cloud computing infrastrcuture
- Computing resource demand can be intensive but only on a need basis, which fits the "metering" model of cloud computing

3-Layer Architecture

- Layered Service-oriented Architecture

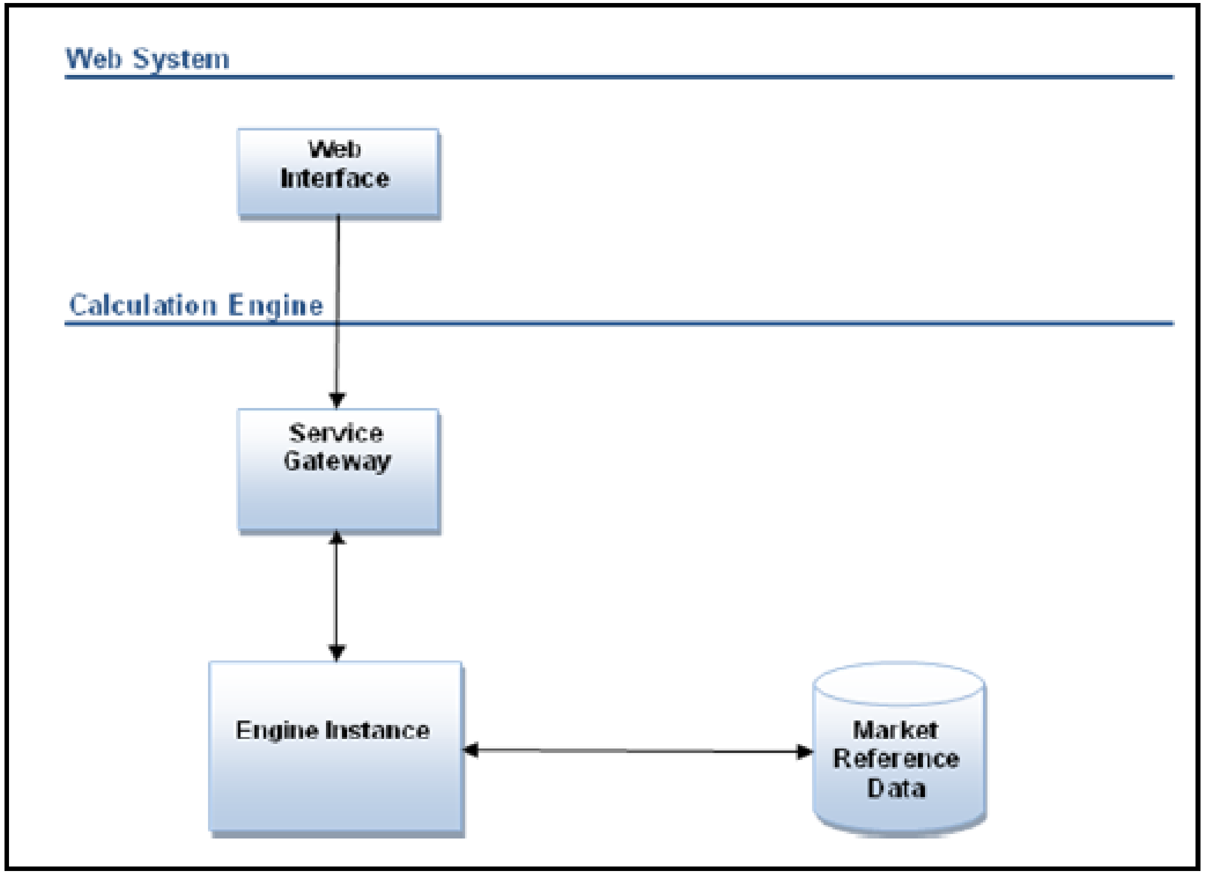


Figure 1. Top-level System Architecture

Lessons Learned

- **Regulatory Issues**
  - Cloud computing is not an explicitly approved application paradigm for finance-related implementations
  - Banking secrecy and ata confidentiality
  - Market regulators in many such jurisdictions do not respond to generic questions on cloud computing
- **Business Issues**
  - not particular keen on the public cloud architecture, because of the potential regulatory complications
  - More suitable for retail side
- **Technology Issues**

Challenges - Technology Issues

- Large Matrix Computations
- Optimization Algorithm
- Data-Centric Algorithm

Basic Idea

- Cornish-Fisher Expansion

$$z_{cf} = z_C + \frac{1}{6}\left(z_C^2 - 1\right)S + \frac{1}{24}\left(z_C^3 - 3z_C\right)K - \frac{1}{36}\left(2z_C^3 - 5z_C\right)S^2$$

- S – Skewness, K – Kurtosis

$$\frac{\partial \mathrm{SR}_{cf}}{\partial \pi_1} = \frac{e_1 z_{cf}\sigma - \sum_i \pi_i e_i \frac{\partial z_{cf}\sigma}{\partial \pi_1}}{\left(z_{cf}\sigma\right)^2} = 0 \quad \text{where } \frac{\partial z_{cf}\sigma}{\partial \pi_1} = \sigma \frac{\partial z_{cf}}{\partial \pi_1} + z_{cf}\frac{\partial \sigma}{\partial \pi_1}$$

$$\Leftrightarrow \quad e_1 z_{cf}\sigma = \sum_i \pi_i e_i \frac{\partial z_{cf}\sigma}{\partial \pi_1}$$

$$\Leftrightarrow \quad \frac{e_1}{\frac{\partial z_{cf}\sigma}{\partial \pi_1}} = \frac{\sum_i \pi_i e_i}{z_{cf}\sigma}$$

Fourth-Order Objective Function

- "Obvious" definition of the stochastic-term and tail-risk adjusted Sharpe Ratio:

$$SR_{cf}^{*} = \frac{\sum_i e_i \pi_i}{z_{cf}\sigma_\pi} + \frac{1}{2}\frac{\sum_i \pi_i \sigma_i^2}{z_{cf}\sigma_\pi} - \frac{\sigma_\pi}{2z_{cf}}$$

- B. Lee and Y. Lee, "Alternative Sharpe Ratio," in B. Schachter (ed), *Intelligent Hedge Fund Investing*, Risk Books, 2004

$$ASR \equiv \frac{\sum_i e_i \pi_i}{z_\pi^- \sigma_\pi} + \frac{1}{2}\frac{\sum_i \pi_i \left(z_i^+ \sigma_i\right)^2}{z_\pi^- \sigma_\pi} - \frac{1}{2} z_\pi^- \sigma_\pi$$

$$z^+ = \frac{\max\left(z_{cf}\left(z_C^+\right), 0\right)}{z_C^+}$$ where z_C^+ is critical value for probability α and

$$z^- = \frac{\max\left(z_{cf}\left(z_C^-\right), 0\right)}{z_C^-}$$ where z_C^- is critical value for probability $1-\alpha$

(e.g. z_C^+ = 2.33 at 1%, z_C^- = -2.33 at 99%)

Challenges 1: Large Matrix Computation

- Tail-Risk Contribution

$$\frac{\partial z_{cf}\sigma}{\partial \pi_1} = \sigma \frac{\partial z_{cf}}{\partial \pi_1} + z_{cf}\frac{\partial \sigma}{\partial \pi_1}$$

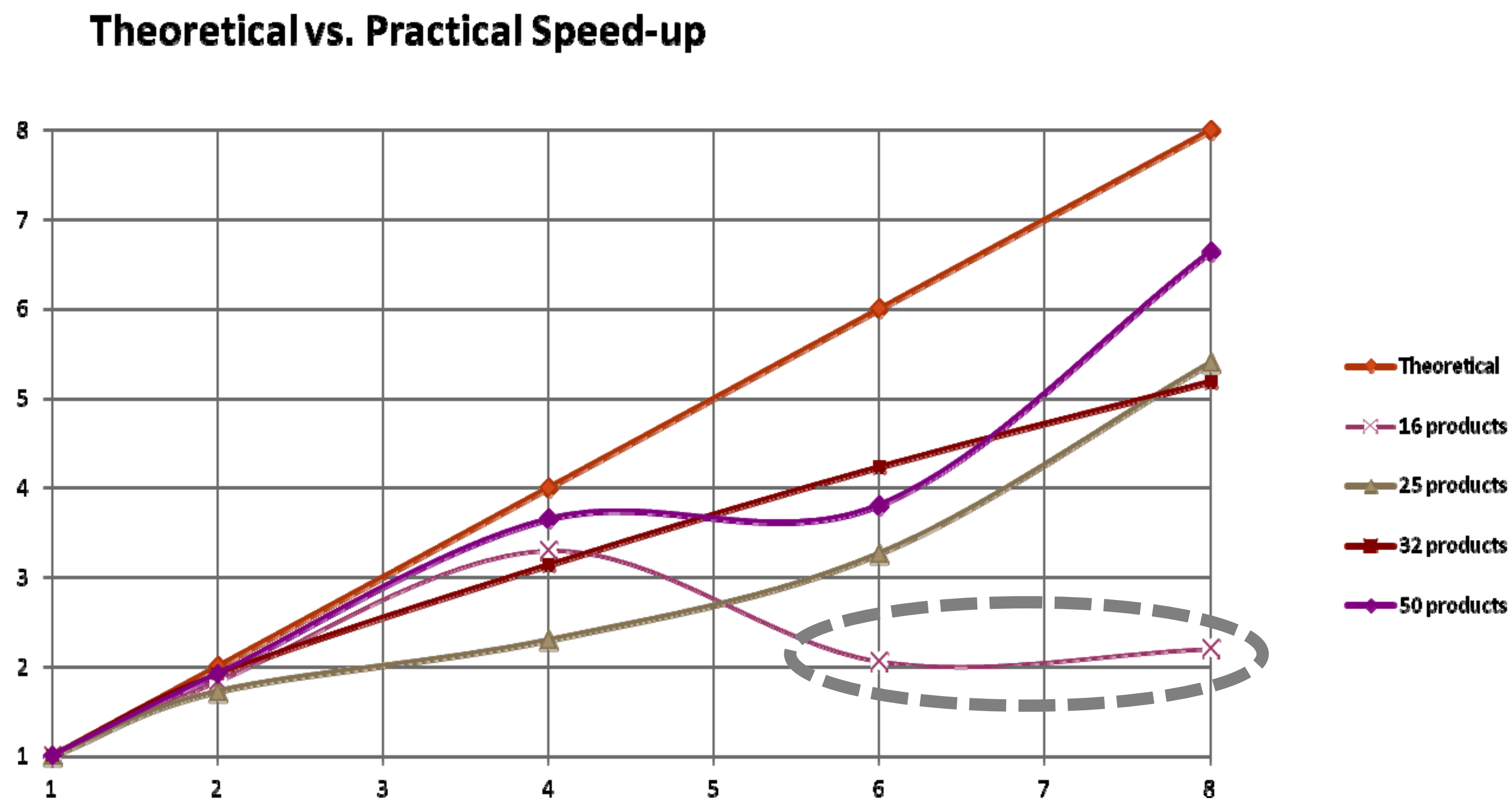
where

$$\frac{\partial z_{cf}}{\partial \pi_1} = \frac{1}{6}\left(z_C^2 - 1\right)\frac{\partial S}{\partial \pi_1} + \frac{1}{24}\left(z_C^3 - 3z_C\right)\frac{\partial K}{\partial \pi_1} - \frac{2}{36}\left(2z_C^3 - 5z_C\right)S\frac{\partial S}{\partial \pi_1} + \cdots$$

$$\frac{\partial S}{\partial \pi_1} = 3\sum_i \sum_j \pi_i \pi_j E\left\{\left[\frac{R_{i,t} - E\left(R_{i,t}\right)}{\sigma}\right]\left[\frac{R_{j,t} - E\left(R_{j,t}\right)}{\sigma}\right]\left[\frac{R_{1,t} - E\left(R_{1,t}\right)}{\sigma}\right]\right\} - 3\frac{S}{\sigma}\frac{\partial \sigma}{\partial \pi_1}$$

$$\frac{\partial K}{\partial \pi_1} = 4\sum_i \sum_j \sum_k \pi_i \pi_j \pi_k E\left\{\left[\frac{R_{i,t} - E\left(R_{i,t}\right)}{\sigma}\right]\left[\frac{R_{j,t} - E\left(R_{j,t}\right)}{\sigma}\right]\left[\frac{R_{k,t} - E\left(R_{k,t}\right)}{\sigma}\right]\left[\frac{R_{1,t} - E\left(R_{1,t}\right)}{\sigma}\right]\right\} - 4\frac{K+3}{\sigma}\frac{\partial \sigma}{\partial \pi_1}$$

Speed-up Graph

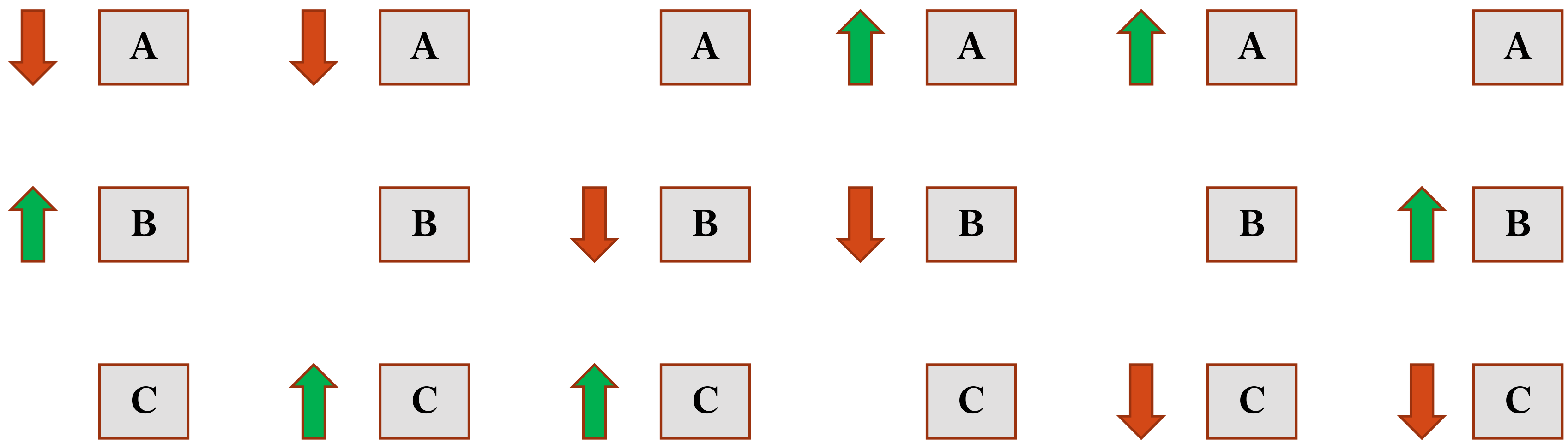


Share Memory Challenges

- Order(N^4)
- Carving up the work – may not worth doing this for "small" problems
- Minimize recomputation of "symmetrical" values
- Return matrix is (relatively) small, but aggregation problematic if written into shared memory
- Load balancing (frequent access to problem "pool" vs. everyone waiting for single threat to complete)

Challenge 2:
Optimization Algorithm - Combinatoric Search

- Fourth-Order Objective Function – Local vs. Global Optimum
- Use Interior Point Method to get to close enough neighborhood
- Must check different combinations of "perturbing the solution" to ensure that the solution is not a local optimum
- Total Portfolio = 100%; Not negative weights; Min Step size
- Conceptually,

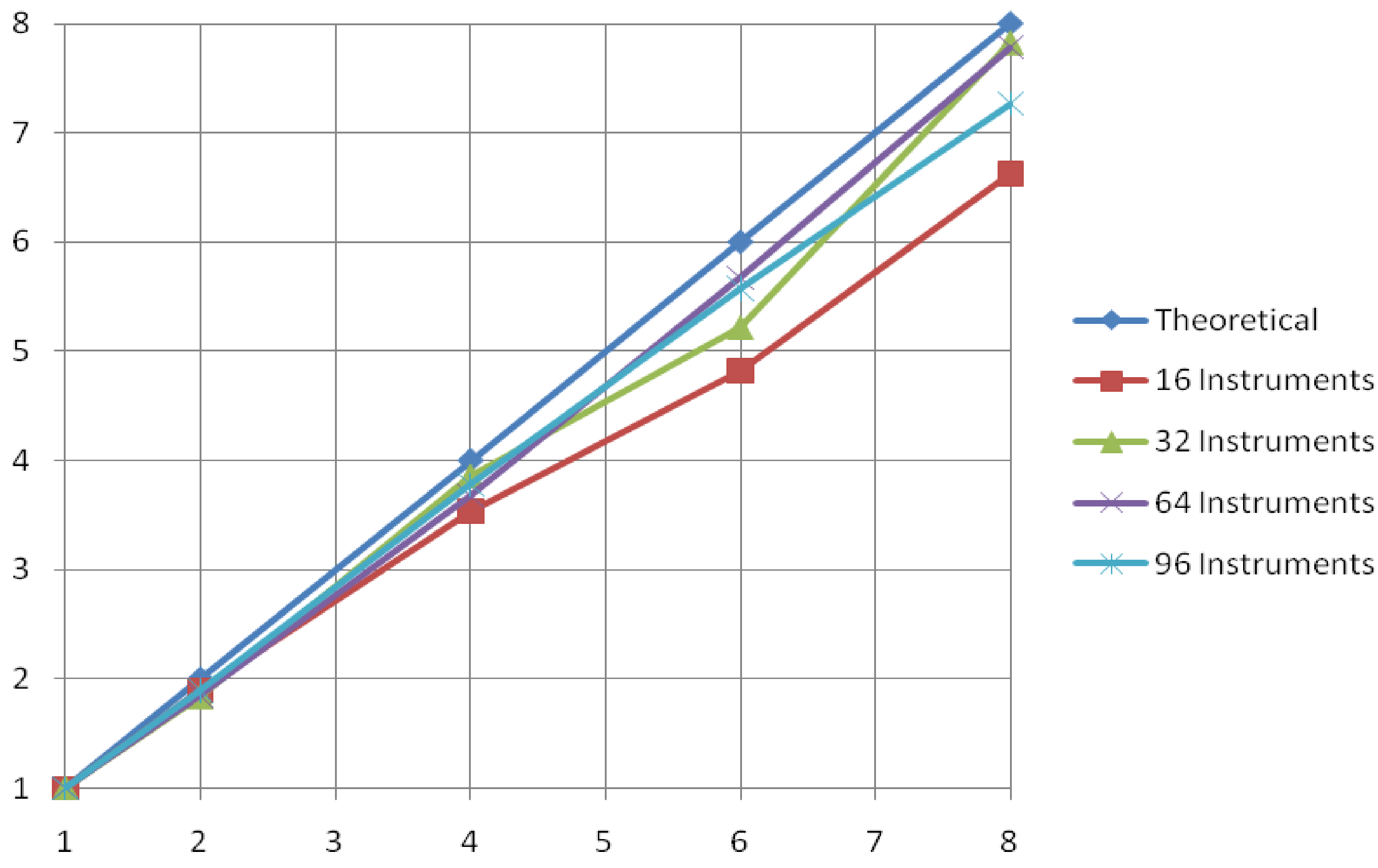


Computational Complexity

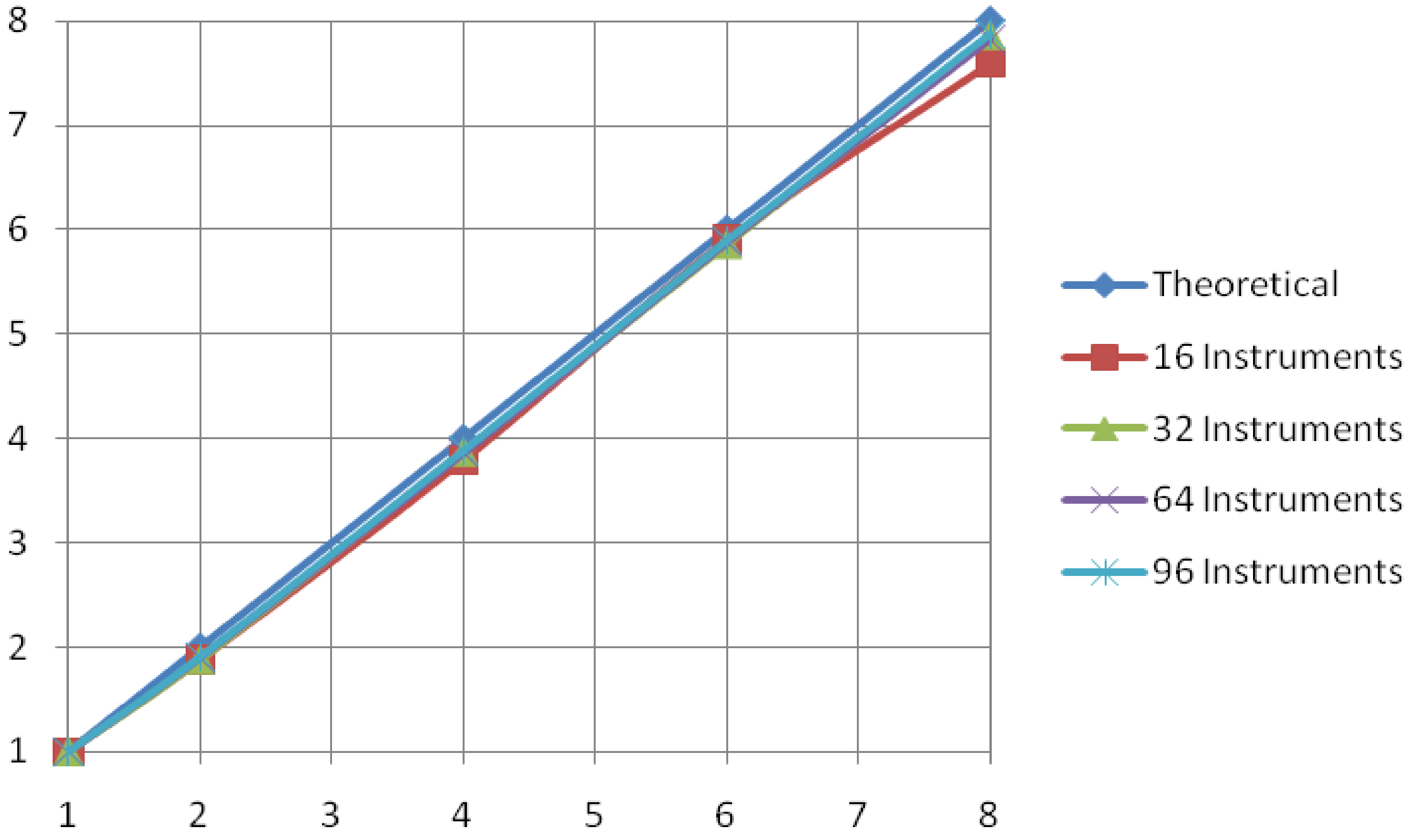
- Time for 16 instruments on single processor = **132 Seconds**
- Approx Time for 32 instruments on single processor = 132 * 65537 = 8650884 seconds = **100 Days**
- Approx Time for 100 instruments on single processor = 132 * 1.93E+25 = 2.55E+27 seconds = **8.09E+19 Years**

Number of Assets	Number of Combinations
100	1,267,650,600,228,230,000,000,000,000,000
50	1,125,899,906,842,620
32	4,294,967,295
25	33,554,431
16	65,535
15	32,767
14	16,383
13	8,191
12	4,095
11	2,047
10	1,023

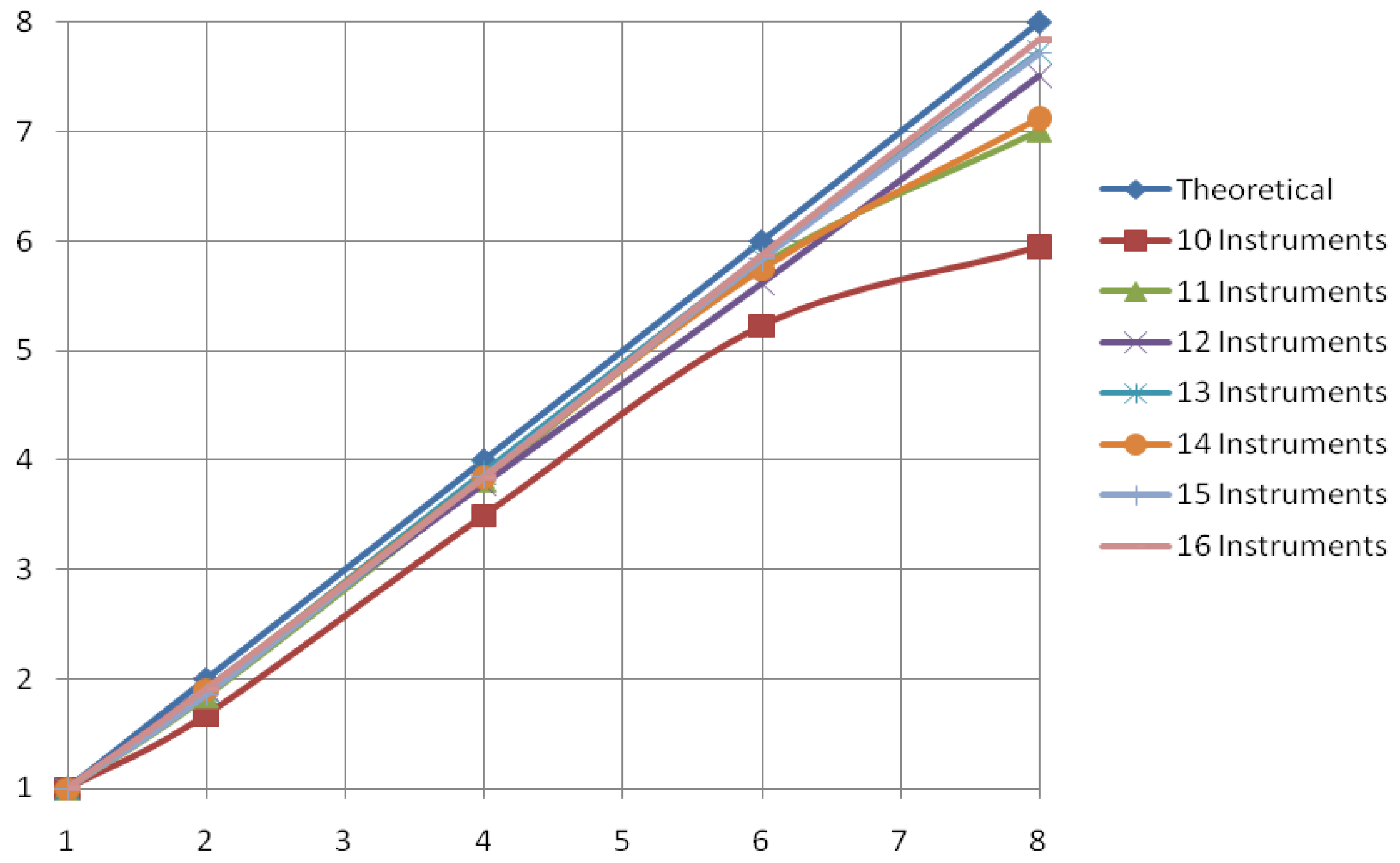
Speed-up Graph – Level 2



Speed-up Graph – Level 3



Speed-up Graph – All Levels



Message-Passing Challenges

- Factorial Time
- Processors no share memory needs
- Carving up the work – "Vertical" vs. "Horizontal" Partitions
- How often the slaves should check back for the problem queue to ensure that a new optimum has not been found, or is it more effective to send messages to the slaves?
- Pt-to-Pt Messages vs. Interrupts
- Possible to eliminate master (since there is no aggregation)
- Stopping rule vs. acceptable accuracy

Challenge 3:
Data-Centric Algorithm

- Data-centric batch processing algorithm that can run overnight
- Algorithms are not that "heavy-duty"
- While this may be an interesting for Hadoop, benefits limited
- Processors tend to be idle during significant periods of time overnight

Future Challenges

- Regulatory Approval
- Industry Acceptance
- Technology Advances
  1. Running Hadoop or a Hadoop-like protocol
  2. Incorporating Data-Driven Features

Conclusions

- Larry Ellison: “The interesting thing about cloud computing is that we've redefined cloud computing to include everything that we already do. I can't think of anything that isn't cloud computing with all of these announcements. The computer industry is the only industry that is more fashion-driven than women's fashion”
- Credible application → Regulatory Approval → Business Acceptance → Industry-wide Adoption of Cloud Computing in Finance

Contact Information:

- **Website:** www.hedgespa.com
- **Address:** SMU Business Innovation Generator
  60 Stamford Road #B1-40, Singapore 178900
- **Email:** bernardl@alumni.princeton.edu
  bernard.lee@hedgespa.com
  info@hedgespa.com

Thank you !