



7th Annual

2009 High Performance Linux

April 6, 2009, Monday

Roosevelt Hotel, New York, NY

on Wall Street

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

Database analysis has today become a core strategic competency for many enterprises.

Knowledge derived from analytics have become the key drivers of corporate strategy.

Evidence of the importance of analytics can be found in the ever-growing mountains of data being collected by enterprises across many markets. For example, banks, hedge funds, credit card and insurance companies, and others in the Financial Services industry depend heavily on analytics to handle processes such as risk profiling, portfolio pricing, marketing campaign management, and fraud detection. Data volumes today are easily in the terabyte-scale, even for medium-size businesses. Analysis of terabyte-scale databases requires planning and careful choice of computing/storage systems to achieve reasonable performance.

In recent years, the legacy "big Unix box" solutions are being challenged by two competing alternatives: fully-integrated "Appliances" and large Linux clusters. This panel will discuss the challenges, the emerging Appliance vendors, and the trade-offs versus commodity clusters.

Geno Valente

Vice President - Sales and Marketing
XtremeData, Inc.



XtremeData builds solutions for high performance ad hoc

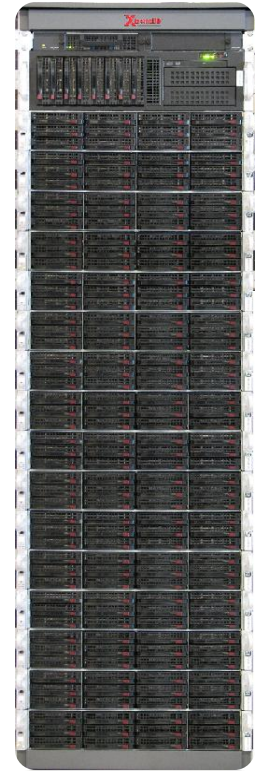
- Patented SQL on a chip!
- Software + processors + storage + interconnect
- Horizontally scalable from 4 to 1024 nodes
- SQL and other industry standard interfaces

The sustainable goals of dbX

- Best performance/\$
- Best performance/watt
- Best performance/square foot

Business benefits

- Deeper wider data exploration - >10x data sizes
- Cycle time compression - 1/10th of the time
- Increase analyst productivity



John Bair
Chief Technical Officer
LaunchPoint



John works as CTO LaunchPoint, helping information-centric startups to innovate and thrive. Before joining LaunchPoint, John worked as the lead technologist for HP's Information Management practice, responsible for BI and IM technology strategies, reference architectures and methodologies. John previously worked as Senior Principal and CTO West for Knightsbridge Solutions, which HP acquired in 2006.

John joined Knightsbridge from Amazon.com, where, as data warehouse architect, he created high performance scalable e-commerce business intelligence systems including Amazon's enterprise data warehouse and analytic services.

Gary Bhattacharjee

Business Intelligence Services

HP



Gary brings a wealth of experience in the field of information management, business intelligence and technology strategy. Over the past couple of decades he has provided thought leadership to support, enhance and improve business strategies primarily in the Financial Services and the Retail industry sector. Gary started his career at IBM Australia, where he was involved with building some core technologies in the area of data management. He played a significant role in Citigroup in the late nineties to lead the consolidation of information across the global corporate banking portfolio. As the leader of the Enterprise Architecture Group at A&P, a US Grocery Retailer, he drove an information oriented IT strategy to achieve substantive savings and increased productivity in operations. Currently he leads the business intelligence services portfolio for the Financial Services (Capital Markets) vertical at Hewlett-Packard.

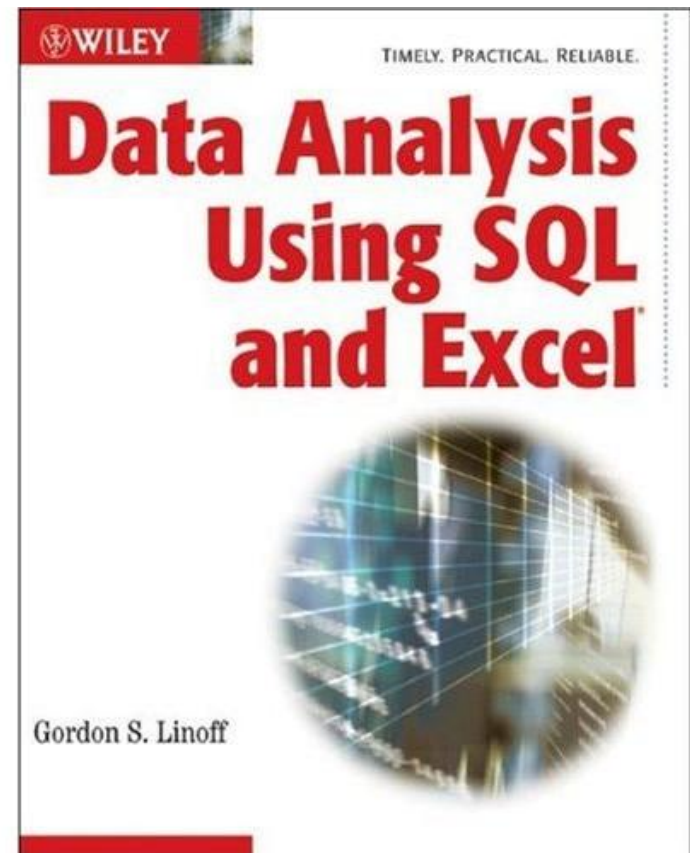
Gordon Linoff

Founder

Data Miners



Eleven years ago, Gordon started a consulting company with his business partner and co-author Michael Berry to focus on their passion: doing, teaching about, and writing about data mining and advanced analytics. Since then, Gordon has taught thousands of students the fundamentals of data mining and written four books. His most recent book, *Data Analysis Using SQL and Excel* provides practical examples of data analysis using readily available tools. His earlier books, written with Michael Berry, include *Data Mining Techniques for Marketing, Sales, and Customer Relationship Management*, and cover the fundamentals and algorithms of data mining in a business context.



Anil Sharma,
Executive Director, IDEAS Practice Areas
Morgan Stanley

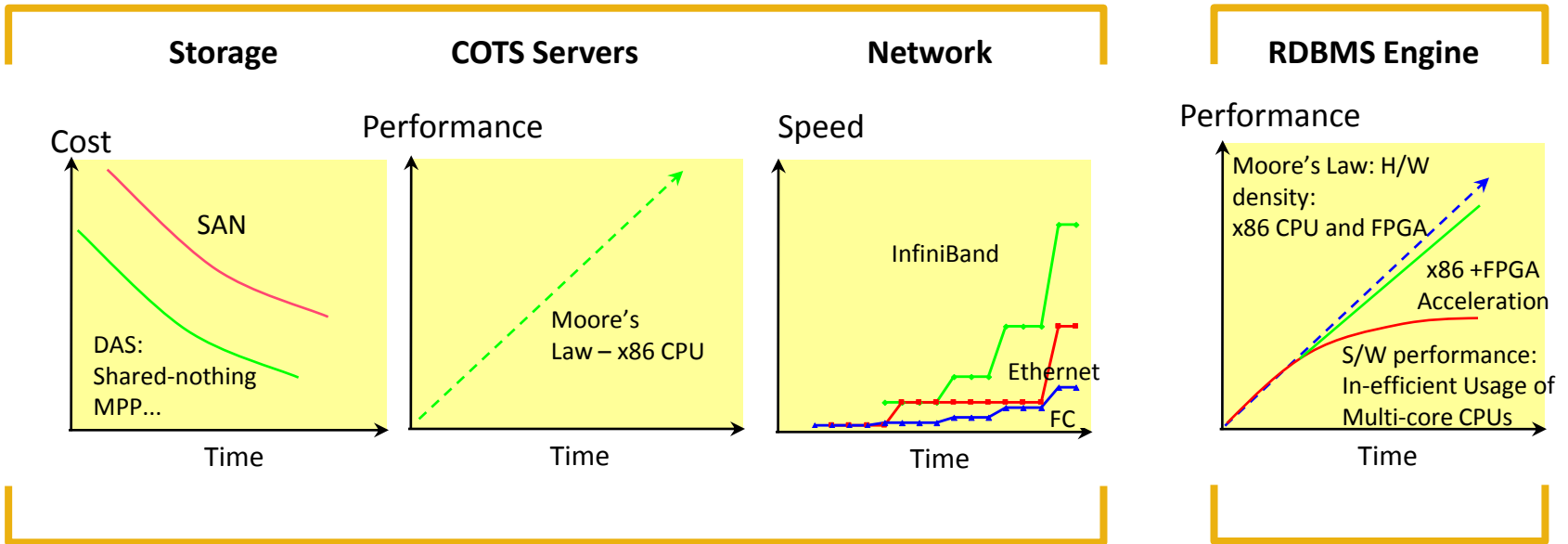


Anil is an Executive Director at Morgan Stanley and is currently leading the Information Management group focusing on Global Operations Risk and Control. He has over 15 years of experience in building enterprise wide Middleware and Database systems. He joined Morgan Stanley in 2000 and was responsible for designing multiple Rapid Application Development toolkits that were instrumental in numerous renovation efforts for many ageing Mainframe systems. For the past 5 years, his primary focus has been on development of large-scale active data warehouses on Linux clusters to provide cost effective, scalable solutions for meeting risk and regulatory requirements. As a key member of the database strategy group, he helped shape the current technology platform for database development at Morgan Stanley.

Views: Data Warehousing Industry

H/W

S/W



The industry challenges:
Incumbents are overly expensive; Appliances are typically proprietary/lock-in
Linux Clusters suffer from CPU slowdown and thus power issues at scale

The winning architecture will be the one that leverages the best of all industry trends...

Let's get started...

- 40 minutes of panel discussion
- 5 minutes of audience questions
- Available during break at ~3:20



Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

Where in FSI are we seeking the largest growth in data analytics and why?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

How could data analytics provide competitive advantage to its users?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

What are the important distinguishing characteristics between a database “appliance” and a customer-integrated solution comprised of a linux cluster and either open source or a commercial database?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

What benefits do customers expect from an appliance and, what liabilities do they perceive?

What benefits do customers expect from a home-grown linux cluster, and what liabilities do they perceive?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

Isn't scalability difficult to achieve with a home-grown solution?

Do customers have an expectation that their human resources be applied to solving business problems and not technology problems that could be solved by a vendor?

Home-grown solutions generally utilize SANs for their flexibility; do you perceive IO bandwidth to be an issue at the high end?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

To what extent is the analytics profile today pre-defined vs. ad-hoc, and does that need to change?

Is all the data preparation, data modeling, data placement, query tuning, etc. seen as a limitation imposed by the technologies?

Is ad-hoc analysis a more powerful vehicle for competitive advantage?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

Has technology been an impediment to
unconstrained ad-hoc data exploration?

How many users would typically access an
analytics platform?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

We constantly hear about data volumes increasing exponentially. Is scalability a real issue?

Do clients expect predictable price/performance over a large scale range?

Do clients expect linear or predictable performance over a large scale range?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

Is true scalability a technically difficult thing to achieve?

How complicated is a scale-up event today?

Do clients have to redistribute data on a larger storage footprint when scaling up?

Is there significant down-time during scale-up?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

What do customers expect to be their
“interface” to the analytics platform?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

What should new entrants in the mostly-read oriented database market do to be successful?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

What data analytics trends are FSI clients paying close attention to today?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

What sort of fault-tolerance is the customer expecting from an analytics platform?

Terabyte-scale Database Analytics: Integrated Appliances vs Linux Clusters

When evaluating data analytics platform price, what characteristics do the buyers focus on?

Performance

Footprint

Power consumption

Maintenance costs

Open Audience Questions



Thank you for Attending



7th Annual

2009 High Performance Linux

April 6, 2009, Monday

Roosevelt Hotel, New York, NY

on Wall Street