

Competitive Benchmarking at Netezza - an Introduction



Session Number **IDW-3073A**

Allan Edwards - IBM

IBM Software

Information On Demand **2011**

Introduction

- Overview of Objectives
 - To Raise Awareness & Inform
 - Clarify Role, How We Operate
 - Foster Communication – Solicit Needs, Opportunities, etc
- Define Context within IBM Ecosystem and Landscape
- Define Context outside of IBM
- Promote Transparency
- Create Good Will



Definition – “Benchmarking”

- Documented Performance Measurement for Comparison
 - Typically a pre-defined, scripted set of test packages
 - Data Model
 - Data
 - Workload
 - Workload Submission/Execution Mechanism (ie: user simulation)

- Desirable Characteristics:
 - Scripted
 - Transparent
 - Reproducible
 - Relevant
 - Realistic



Definition – “Benchmarking”

- Inter-Vendor Performance Measurement
 - “us vs them”
 - This is where the “Competitive” element comes in
- Intra-Vendor Performance Measurement
 - Various configurations, releases/versions, etc
 - Longitudinal performance & patterns
- Needs to be a Fair Comparison
 - Workload, Data & Execution Mechanism must be identical
 - System Under Test – typically DBMS – can vary





Speaker Background

- Career DBMS Engineer, 1984 to
 - Database Internals
 - Development
 - Testing
 - Deployment & application
- Performance Specialization
 - TPC Benchmarks
 - Customer Benchmarks
 - Customer Performance Optimization, Management
 - Theoretical Limitation & Capacity Testing
- Performance Lab Management & Experience
 - 3 Hardware Vendors
 - Oracle Corp
 - Integration & Implementation Consulting
- Personal Passion: Instrumentation, Quantification, Optimization



Defining “Our Space”

- Netezza is Not a “General Purpose” Platform

- Benchmarking is Focused on Core Value Prop’s of Netezza
 - Analytics
 - Huge Data Volumes (typically many TB and up)
 - Hunting for patterns across/within vast subsets
 - Complex computation & analysis
 - Multi-pass, multi-touch investigation
 - Unpredictable access paths, analysis step
 - “guided discovery”, frequently ad-hoc
 - Typically expressed in pure SQL
 - Advanced Analytics
 - Greater sophistication in analysis, computation
 - Expressed in SQL & other analytic languages (R, INZA, etc)



Related Reading – “Analytic Workload” Discussion



<http://www.netezza.com/analyticworkloads/>

Motivations – Why We Exist

- Performance is Central to our Value Proposition
 - But not the only element ...

- Performance is Difficult to Predict
 - Must be measured, verified & monitored

- We are the “Test Track”
 - Another place to test limits

- Quantitative Data for Selling & Positioning

- Feedback to Engineering, Marketing, Management, etc



Other Performance Work at NZ

- Performance Architecture Group
 - Forward-looking performance research, often pre-release
 - How should/do proposed designs perform?
 - How do latest developments perform? As expected?
- System Quality Assurance
 - Performance testing in the release verification process
 - Are there any performance regressions, changes?
- Customer PoC's (**P**roof **o**f **C**oncept)
 - Many occurring concurrently at any moment
 - Some of the most important & greatest coverage
 - Lifelike by definition; real usage in real settings
- Partners
 - Similar to much of the above, in parallel



Resources of Benchmark Group

- Dedicated Computing Lab for Performance Measurement
 - Servers
 - Storage
 - Peripheral systems
 - Execution & Replay
 - Monitoring, performance data collection
 - Performance data analysis & investigation
 - Historical data preservation
- Personnel
 - Engineers dedicated to testing
 - Preparation & Execution
 - Results Analysis & Documentation
 - Tuning & Optimization



IBM Counterparts - CPO

- IBM has a Well-Established Competitive Group
 - Competitive Project Office
 - Deep resources in many technologies
 - Versatile, capable, seasoned performance engineers
 - Vast array of hardware resources, technologies
- We Work Together Cooperatively for Common Good
- IBM “Network” Affords Access to Vast Resources & Opportunities
 - Customers, workloads, infrastructure, assets, etc



Vendor Benchmark Landscape

- All of our Competitors have Benchmarking Teams
 - Frequently the “best-of-the-best” for that vendor

- Serving Multiple Needs
 - Internal product performance measurement
 - External, customer-facing performance
 - External, Industry-facing (standardized industry benchmarks)

- Important Differences
 - NZ Comp Bench team rarely involved in pre-sales benchmarks
 - Performed by pre-sales engineers & customer resources
 - Typically on-site in customer’s actual infrastructure
 - Appliance nature of our solution
 - Fewer layers, smaller spectrum of expertise



Reinforcing Some Important Unity

- We are a Critical Intersection Point of Multiple Groups

- Marketing

- Limited first-hand experience with products

- Engineering

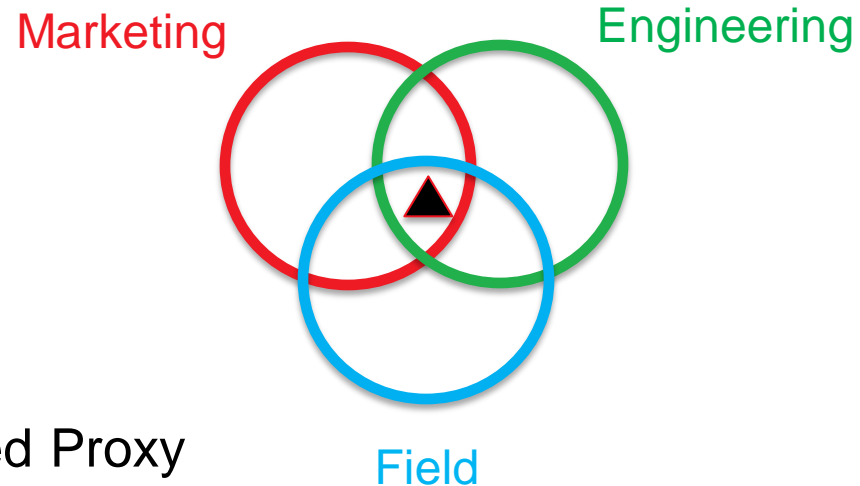
- Limited exposure/awareness of how customers use products

- Field

- Customers
- Sales Engineers
- Support

- We are Often a Factory-Based Proxy

- What customers experience



Requirements for Success & Benefit

- Sense of How Customers Really Use Our Products
 - Testing needs to be relevant, representative
 - How many care how fast car goes in reverse?
 - Some, but not most

- Empirical Methods
 - Rigorous Scientific Methods – Test & Control
 - Isolation from external factors, variables
 - Integrity and purity

- Fact-based Objectivity, Rigorous Documentation

- Knowledge of How Things Work Internally



Requirements for Valid Comparison

- Identical Data
 - Data Model
 - Raw Data
 - Platform-specific load mechanism & scripting is expected
- Identical Workload
 - SQL
 - Bind/run-time values
- Replay Fidelity & Consistency
 - User volumes
 - Execution sequence
 - Think times, human factors, etc
- Performance & Execution Measurement
 - Common, equivalent performance metrics
 - End-user performance experience at the very least



Challenges

- No Single Test is Universal, Indicative
 - Multiple benchmarks provide better coverage
- The Application Details DO Matter
- The Overall Ecosystem Matters, too
 - Behavior in isolation differs from live infrastructure
 - We are attempting to study a system, in a network of systems
- Nearly All Environments are Materially Different
 - Usage patterns
 - User loads
 - Workload types
- Real, Production Usage is Very Hard to Simulate
 - A synthetic test can easily mislead



Widely-Adopted Analytics Benchmarks : None

- Relevant, Mainstream Benchmarks for Analytics are Lacking
 - Unlike the transactional space, ex: TPC-C
 - TPC-D, TPC-H were attempted, then deprecated
 - Various shortcomings, issues, concerns
 - TPC-DS proposed, but not ratified nor completed
 - Still evolving, since 2007
- Lack of Consensus & Official Results = No “Off-the-Shelf” Results
 - You must test a competitors platform yourself
 - Better yet, customer must test
- Standardized, Audited, Public Results Do Not Exist



Importance of Transportability

- An Ideal Workload is Entirely Transportable
 - SQL is great example of a transportable, standard language*
 - Yes, there are minor challenges

- Advanced Analytics are Often NOT Entirely Transportable
 - Sometimes expressed through SQL extensions
 - Sometimes through package-specific proprietary calls
 - Ex: SAS vs SPSS vs R vs ... -> all different

- As a Result, Advance Analytics are Harder to Compare Fairly
 - But still worth doing ...



Recent Activity

- Group & Role Officially Launched in 2010
- Soon Thereafter, IBM Acquisition Announcement
- Several Initiatives Currently in Motion:
 - Development of a TPC-DS test suite for internal use
 - Several threads of IBM integration & performance testing
 - A Cognos-Based performance benchmark for Internal Use by IBM
 - Dialogue with customers to procure data & workloads



Status & Goals for Coming Year

- Continue Development & Testing of Existing Programs
 - Customer Workloads
 - TPC-DS implementation
 - Other active workstreams
- Identification and Pursuit of Initial Advance Analytics Tests
- Performance Characterization of Competitor Systems
- Creation of Prototype Bundle(s) for Field-Based Performance Tests



Parting Thoughts

- Trust But Verify
- Beware of Gamesmanship & Self-Promotion
- Don't Take Anyone's Word – Find Out for Yourself
 - Your Workload
 - Your Data
 - Your Infrastructure
 - Your Staff
- Beware of Perfection
 - All systems are complex & imperfect



Don't Bet Your Business on Someone Else's Word ...



megawoosh

Search

Browse

Movies

MEGAWOOSH - Bruno Kammerl jumps

MrKammerl

3 videos

Subscribe



Like Add to Share

5,794,717

Uploaded by MrKammerl on Aug 3, 2009

8,696 likes, 609 dislikes



Don't Bet Your Business on Someone Else's Word ...



HOME NEWS THE PROJECT SOFTSLIDE CONTACT



What is MEGAWOOSH???

MEGAWOOSH is my idea for the most unusual water slide in the world. Thought up and designed by me. It is my dream to make it a reality on a really high mountain.

Investors sought!

Whoever knows investors or has any interest whatsoever in courageously making the world a better place: Apply here!
info@megawoosh.com



softslide test!



05/08 Softslide also passes second series of tests successfully

All doubts were unfounded. The 35.2 meter distance jump was the first practical Softslide test with real humans (me) under real conditions. I think words would fail even the last remaining skeptics at this point.



Install Microsoft Silverlight to view this video in full screen mode.
(4 MB Download, 20 Second Installation)



Thank You!

Your Feedback is Important to Us

- Access your personal session survey list and complete via SmartSite
 - Your smart phone or web browser at: iodsmartsite.com
 - Any SmartSite kiosk onsite
 - Each completed session survey increases your chance to win an Apple iPod Touch with daily drawing sponsored by Alliance Tech

allan.edwards@us.ibm.com