IEEE BigData 2014 Tutorial on Big Data Benchmarking

Dr. Tilmann Rabl
Middleware Systems Research Group, University of Toronto
tilmann.rabl@utoronto.ca

Dr. Chaitan Baru
San Diego Supercomputer Center, University of California San Diego (currently at NSF)
baru@sdsc.edu

Outline

Tutorial overview (5 mts)
Introduction to Big Data benchmarking issues (15 mts)
Different levels of benchmarking (10 mts)
Survey of some Big Data Benchmarking initiatives (15 mts)

BREAK (5 mts)

Discussion of *BigBench* (30 mts)
Discussion of the *Deep Analytics Pipeline* (10 mts)
Next Steps, Future Directions (10 mts)

About Us

Dr. Tilmann Rabl

- PostDoc at MSRG, UofT; CEO at bankmark
- Developer of Parallel Data Generation Framework (PDGF)
- Member of Steering Committee, WBDB, BDBC; Chair of SPEC RG Big Data Working Group; TPC professional affiliate

Dr. Chaitan Baru

- Associate Director, Data Initiatives, San Diego Supercomputer Center, UC San Diego
- Previously worked on DB2 Parallel Edition at IBM (18 years ago!)
  - At that time, helped with TPC-D spec; helped deliver industry's first audited TPC-D result
- Member of WBDB Steering Committee; Co-Chair of SPEC Big Data RG
- Now Senior Advisor for Data Science, NSF, Arlington VA.

Resources

Specifying Big Data Benchmarks, edited by T. Rabl, M. Poess, C. Baru, H.-A. Jacobsen, Lecture Notes in Computer Science, Springer Verlag, LNCS 8163, 2014.

Advancing Big Data Benchmarks, edited by T. Rabl, R. Nambiar, M. Poess, M. Bhandarkar, H.-A. Jacobsen, C. Baru, Lecture Notes in Computer Science, Springer Verlag, LNCS 8585, 2014.

Workshops on Big Data Benchmarking (WBDB), see http://clds.sdsc.edu/bdbc/workshops.

SPEC Research Group on Big Data Benchmarking, see http://research.spec.org/en/working-groups/big-data-working-group.html

TPCx-HS Benchmark for Hadoop Systems, http://www.tpc.org/tpcx-hs/default.asp

BigBench Benchmark for Big Data Analytics, https://github.com/intel-hadoop/Big-Bench

Overview…1

Introduction to Big Data benchmarking issues (15 mts)

- Motivation
  - Lack of standards; vendor frustration; opportunity to define the set of big data application “classes”, or range of scenarios
- Which Big Data?
  - The V’s; warehouse vs. pipelines of processing; query processing vs. analytics
- Introduction to benchmarking issues
  - How does industry standard benchmarking work?
  - TPC vs. SPEC model
  - The need for audited results
- Summary of the Workshops on Big Data Benchmarking (WBDB)
  - Who attends; summary of ideas discussed

Overview…2

Benchmarking at different levels (10 mts)

- Micro-benchmarking, e.g. IO-level
- Functional benchmarks, e.g. Terasort, Graphs
  - Overview of TPCx-HS. What does the TPC process bring?
  - Graph 500: characteristics; results
- Application-level benchmarking, e.g. TPC-C, TPC-H, TPC-DS
  - History / success of TPC benchmarks
  - Description of how TPC benchmarks are constructed; data generation; ACID rules; auditing; power runs; throughput runs; metrics

Overview…3

Survey of some Big Data benchmarking efforts (15 mts)

- E.g. HiBench, YCSB, …

Break (5 mts)

Discussion of BigBench (30 mts)

- Extending the TPC-DS schema and queries
- Data Generation
- HIVE implementation
- Preliminary results

Overview…4

Discussion of the *Deep Analytics Pipeline* (10 mts)

Next Steps, Future Directions (10 mts)

- Platforms for Benchmarking
- SPEC Research Group for BigData
- Creating the *BigData Top100 List*

Tutorial Overview

Tutorial overview

Introduction to Big Data benchmarking issues

Different levels of benchmarking

Survey of some Big Data Benchmarking initiatives

BREAK

Discussion of *BigBench*

Discussion of the *Deep Analytics Pipeline*

Next Steps, Future Directions

Big Data Benchmarking Issues

Motivation

- Lack of standards; vendor frustration; opportunity to define the set of big data application "classes", or range of scenarios

Which Big Data?

- The V's; warehouse vs pipelines of processing; query processing vs analytics

Different approaches to benchmarking

- How does industry standard benchmarking work?
- TPC vs SPEC model
- The need for audited results

Summary of the Workshops on Big Data Benchmarking (WBDB)

- Who attends; summary of ideas discussed

Which Big Data ?

Benchmarks for big data could be defined by the "Vs"

Volume

- Can the benchmark test scalability of the system to very large volumes of data?

Velocity

- Can the benchmark test ability of the system to deal with high velocity of incoming data?

Variety

- Can the benchmark include operations on heterogeneous data, e.g. unstructured, semi-structured, structured?

And "Variety" also refers to different *data genres*

- Can the benchmark incorporate operations on graphs, streams, sequences, images, text, …

Approaches to Big Data Benchmarking: Data Science Workloads

Big data enables Data Science

Data science workloads incorporate not just queries, but also analytics and data mining

Data Science workloads are characterized as consisting of:

- **O**btain, **S**crub, **E**xplore, **M**odel, **I**nterpret set of steps
- Implies workloads that are *pipelines* of processing

Refs: [1] Hilary Mason and Chris Wiggins, *A Taxonomy of Data Science*, Sept 25th, 2010, dataists.com, [2] *Data Science Workloads for Big Data Benchmarking*, Milind Bhandarkar http://clds.sdsc.edu/sites/clds.sdsc.edu/files/wbdb2012/presentations/WBDB2012Presentation19Bhandarkar.pdf

In the beginning there was Sorting…

Early popularity of Terasort

- Sortbenchmark.org
- GraySort, MinuteSort, TeraSort, CloudSort, …

Pluses:

- Simple benchmark—easy to understand and easy to run
- Therefore, developed a “brand”
- Scalable model
- Good for “shaking out” large hardware configurations

TeraSort

Minuses:

- Not standardized
- “Flat” data distribution (no skews)
- Not application-level

Require more than just sorting for a Big Data benchmark

See presentation by Owen O’Malley, Hortwonworks at 1st WBDB, 2012
http://clds.sdsc.edu/sites/clds.sdsc.edu/files/wbdb2012/presentations/WBDB2012Presentation04OMalley.pdf

TPC Benchmarks

TPC: Transaction Processing Performance Council

Existing Benchmark Suite

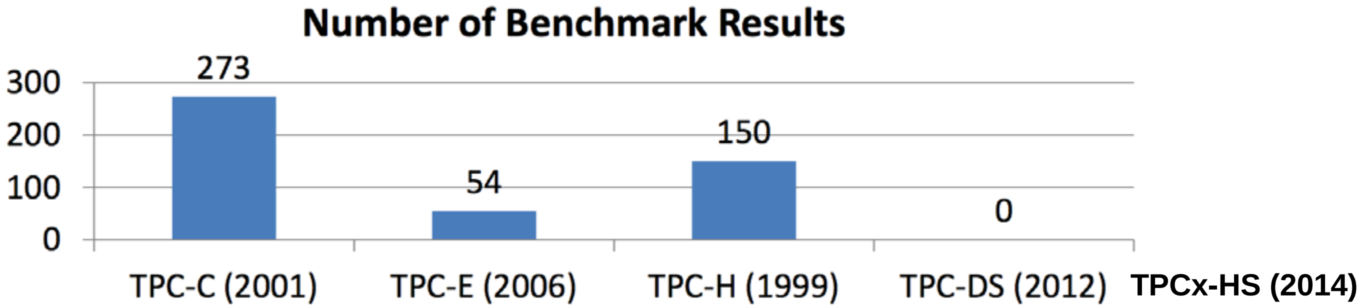
- OLTP: TPC-C, TPC-E
- BI: TPC-H, TPC-DS
- Applicable to all above
  - Energy, Pricing

Benchmarks under Development

- Virtualization
- Data Integration (ETL)

New benchmark

- TPCx-HS (for Hadoop systems)



TPC Benchmarks

Benchmarks are:

- Free for download
- Utilize standardized metrics
  - Price performance
  - Energy
- Test entire system performance, transactions or queries per unit of time
- Are software and hardware independent
- Have long shelf life

Benchmark publications are:

- Subject to a fee
- Require full disclosure
- Are independently audited

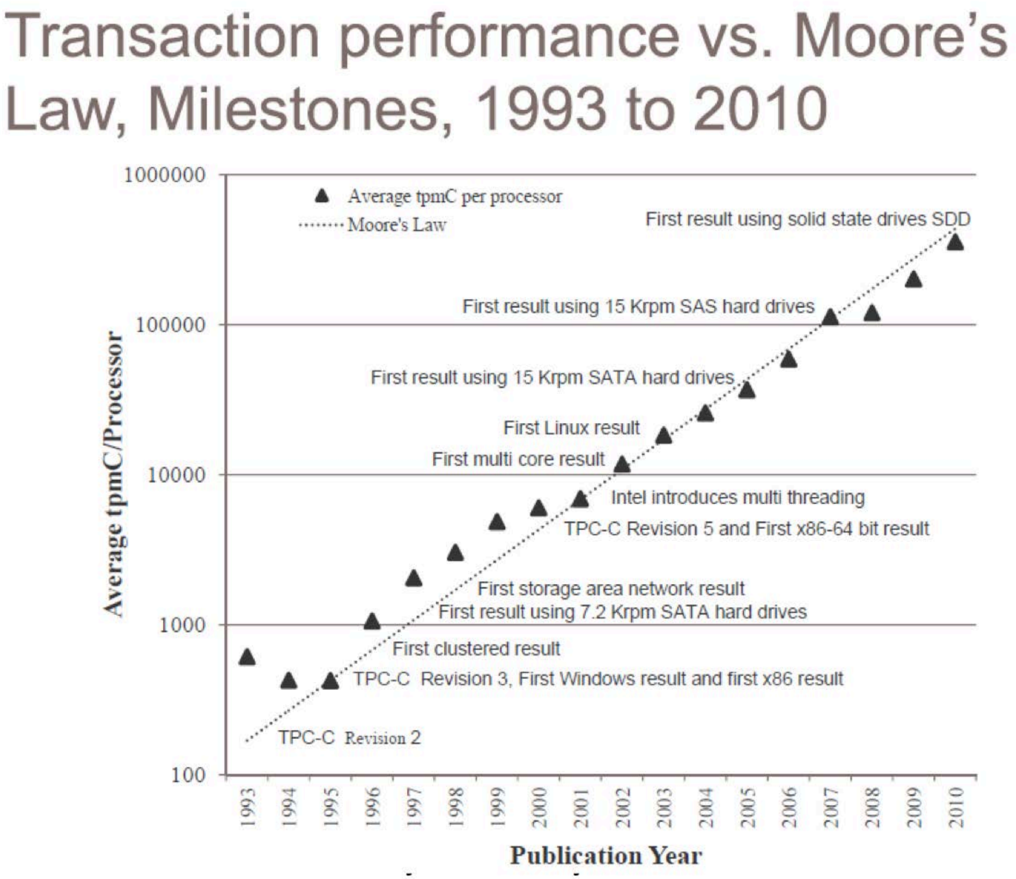
TPC-C: Transaction processing benchmark

Longevity and robustness of the TPC-C benchmark

- Benchmark measures transactions per minute for a scenario based on Order-Entry systems
- The transactions include entering and delivering orders, recording payments, checking the status of orders, and monitoring the level of stock at the warehouses

Data Scaling

- “Continuous scaling” model
  - The number of warehouses in the database need to scale up with the number of transactions



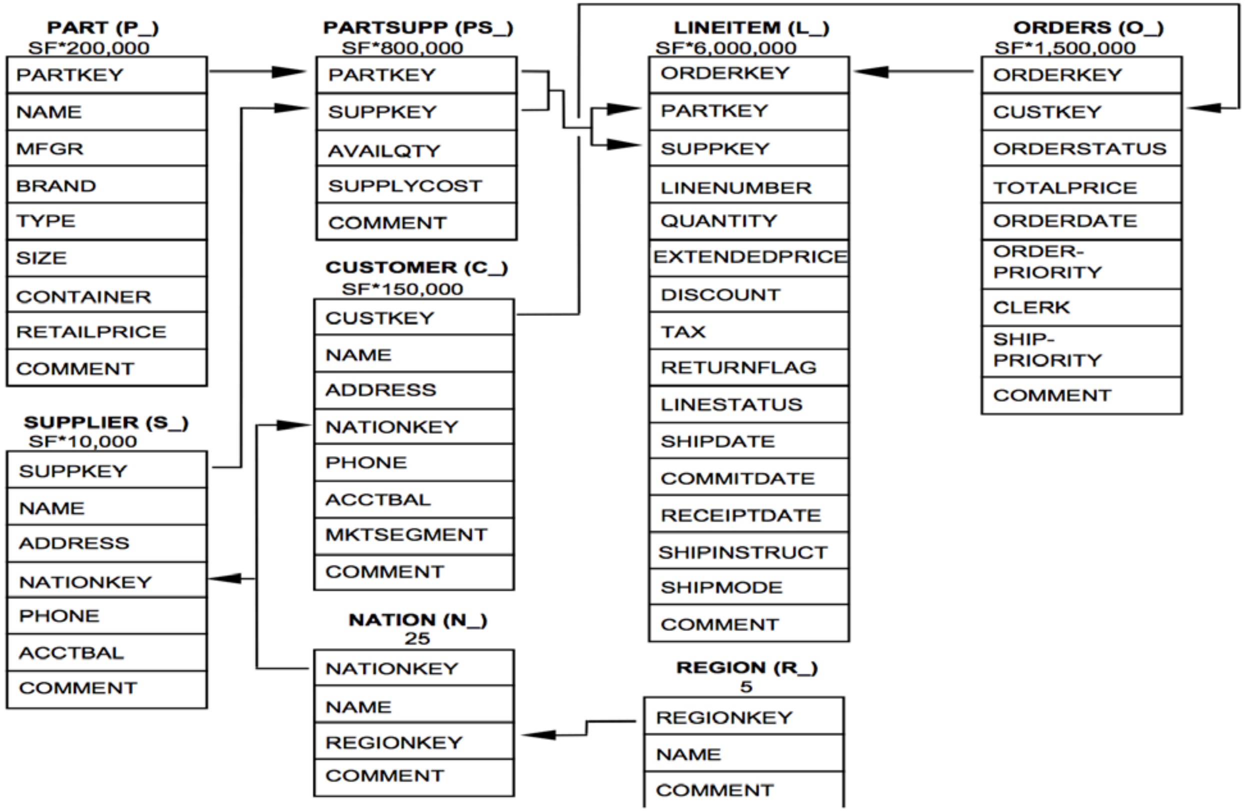
From presentation by Meikel Poess, 1st WBDB, May 2012

TPC-H: Data Warehousing Benchmark

Parts, Suppliers, Customers, Orders, Lineitems

Scaling TPC-H

- Scale Factors: From 1GB data size upwards
- Size = Table cardinality × SF, except Nation and Region (code tables)



TPC Scale Factors

Discrete scale factors, with corresponding DB sizes

ScaleFactor	1	10	30	100	300	1000	3000	10000	30000	100000
DB size in GB's	1	10	30	100	300	1,000	3,000	10,000	30,000	100,000

Most popular range

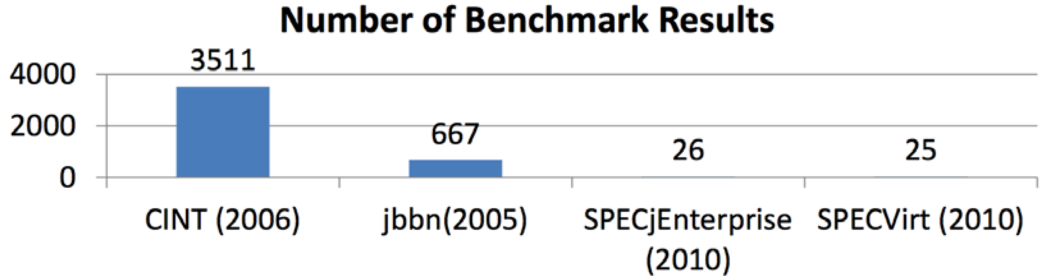
Recent result:
Dell @ 100TB, 9/23/2014
QphH: 11,612,395; Price/QphH= 0.37c

SPEC Benchmarks

SPEC: Standard Performance Evaluation Corporation

Existing Benchmark Suite

- CPU: new version every 3-5 years
- High Performance Computing: new version every 3-6 years
- Enterprise benchmark
- Power
- Virtualization
- Web



Aspects of SPEC Benchmarks

Benchmarks

- Can be downloaded for a fee
- Each benchmark defines its own metric
- Benchmarks test performance of small systems or components of systems
- Server-centric
- Have a short shelf life

Benchmark publications

- Are free to members and subject to a fee for non-members
- Are peer reviewed
- Require disclosure summary

TPC vs. SPEC

TPC Model	SPEC Model
Specification based	Kit based
Performance, price, energy in one benchmark	Performance and energy in separate benchmarks
End-to-end	Server-centric
Multiple tests (ACID, load, etc.)	Single test
Independent review	Peer review
Full disclosure	Summary disclosure
TPC Technology Conference	SPEC Research Group, ICPE (International Conference on Performance Engineering)

From presentation by Meikel Poess, 1st WBDB, May 2012

Dealing with elasticity and failures

TPC: ACID test are performed “out of band”

- Official TPC benchmarking requires performance of ACID test (to test availability of features that support Atomicity, Consistency, Isolation, and Durability)

Big Data platforms are expected to be “elastic”

- Can absorb, utilize new resources added to the system during run time
- Can deal with hardware failures during run time, e.g. via replication

Workshops on Big Data Benchmarking

Initiated as a industry-academia forum for developing big data benchmarking standards

First workshop held in May 2012, San Jose, CA

About 60 attendees from 45 different organizations:

Actian
AMD
BMMsoft
Brocade
CA Labs
Cisco
Cloudera
Convey Computer
CWI/Monet
Dell
EPFL
Facebook
Google

Greenplum
Hewlett-Packard
Hortonworks
Indiana Univ / Hathitrust Research Foundation
InfoSizing
Intel
LinkedIn
MapR/Mahout
Mellanox
Microsoft

NSF
NetApp
NetApp/OpenSFS
Oracle
Red Hat
SAS
Scripps Research Institute
Seagate
Shell
SNIA
Teradata Corporation

Twitter
UC Irvine
UC San Diego
Univ. of Minnesota
Univ. of Toronto
Univ. of Washington
VMware
WhamCloud
Yahoo!

Pune, India

2nd WBDB: http://clds.sdsc.edu/wbdb2012.in

Xi'an, China

3rd WBDB: http://clds.sdsc.edu/wbdb2013.cn

San Jose, USA

4th WBDB: http://clds.sdsc.edu/wbdb2013.us

Potsdam, Germany

6th WBDB…Time/Place TBD

SPEC RG on Big Data Benchmarking will meet at ICPE, Austin, TX, Jan 30-Feb 4, 2015

WBDB Outcomes

Big Data Benchmarking Community (BDBC) mailing list (~200 members from ~80 organizations)

- Organized webinars every other Thursday
- http://clds.sdsc.edu/bdbc/community

Paper from First WBDB

- *Setting the Direction for Big Data Benchmark Standards* C. Baru, M. Bhandarkar, R. Nambiar, M. Poess, and T. Rabl, published in *Selected Topics in Performance Evaluation and Benchmarking*, Springer-Verlag

WBDB Outcomes…

Selected papers in Springer Verlag, *Lecture Notes in Computer Science,* Springer Verlag

- Papers from 1st and 2nd WBDB published in Specifying Big Data Benchmarks, ISBN 978-3-642-53973-2, Editors: Rabl, Poess, Baru, Jacobsen
- Papers from 3rd and 4th WBDB published in Advancing Big Data Benchmarks, ISBN 978-3-319-10596-3, Editors: Rabl, Nambiar, Poess, Bhandarkar, Jacobsen, Baru
- Papers from 5th WBDB will be in Vol III

Formation of TPC Subcommittee on Big Data Benchmarking

- Working on TPCx-HS: TPC Express benchmark for Hadoop Systems, based on Terasort
- http://www.tpc.org/tpcbd/

Formation of a SPEC Research Group on Big Data Benchmarking

- http://research.spec.org/working-groups/big-data-working-group.html

Which Big Data? – Abstracting the Big Data World

1. Enterprise Data Warehouse + other tructured data
   - Extend data warehouse to incorpor semi-structured data from weblogs, customer reviews, etc.
   - Mixture of analytic queries, rep machine learning, and MR style processing

BigBench

2. Collection of heterogeneous data + pipel ng
   - Enterprise data processing as a pipeline fr transformation, extraction, subsetting, machine learning, pred
   - Data from multiple struc tured sources
   - "Runtime" sche application-driven schemas

Deep Analytics Pipeline (DAP)

Tutorial Overview

Tutorial overview

Introduction to Big Data benchmarking issues

Different levels of benchmarking

Survey of some Big Data Benchmarking initiatives

BREAK

Discussion of *BigBench*

Discussion of the *Deep Analytics Pipeline*

Next Steps, Future Directions

Benchmark Design Issues (from WBDB)

Audience: Who is the audience for the benchmark?

- Marketing (Customers / End users)
- Internal Use (Engineering)
- Academic Use (Research and Development)

Is the benchmark for innovation or competition?

- If a competitive benchmark is successful, it will be used for innovation

Application: What type of application should be modeled?

- TPC: schema + transaction/query workload
- BigData: Abstractions of a data processing pipeline, e.g. Internet-scale businesses

Benchmark Design Issues - 2

Component vs. end-to-end benchmark. Is it possible to factor out a set of benchmark "components", which can be isolated and plugged into an end-to-end benchmark?

- The benchmark should consist of individual components that ultimately make up an end-to-end benchmark

Single benchmark specification: Is it possible to specify a single benchmark that captures characteristics of multiple applications?

- Maybe: Create a single, multi-step benchmark, with plausible end-to-end scenario

Benchmark Design Issues - 3

Paper & Pencil vs Implementation-based. Should the implementation be specification-driven or implementation-driven?

- Start with an implementation and develop specification at the same time

Reuse. Can we reuse existing benchmarks?

- Leverage existing work and built-up knowledgebase

Benchmark Data. Where do we get the data from?

- Synthetic data generation: structured, non-structured data

Verifiability. Should there be a process for verification of results?

- YES!

Types of Benchmarks

Micro-benchmarks. To evaluate specific lower-level, system operations

- E.g., A Micro-benchmark Suite for Evaluating HDFS Operations on Modern Clusters, Panda et al, OSU

Functional \ component benchmarks. Specific high-level function.

- E.g. Sorting: Terasort
- E.g. Basic SQL: Individual SQL operations, e.g. Select, Project, Join, Order-By, …

Genre-specific benchmarks. Benchmarks related to type of data

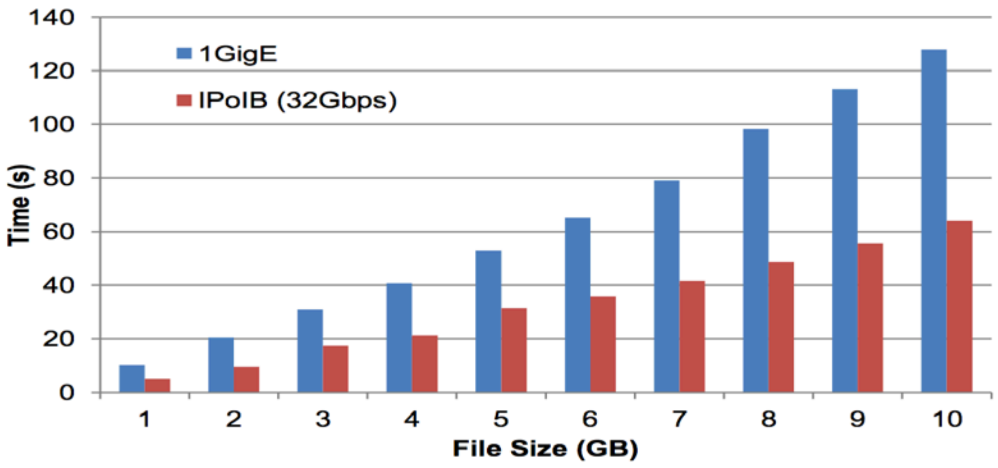
- E.g. Graph500. Breadth-first graph traversals

Application-level benchmarks.

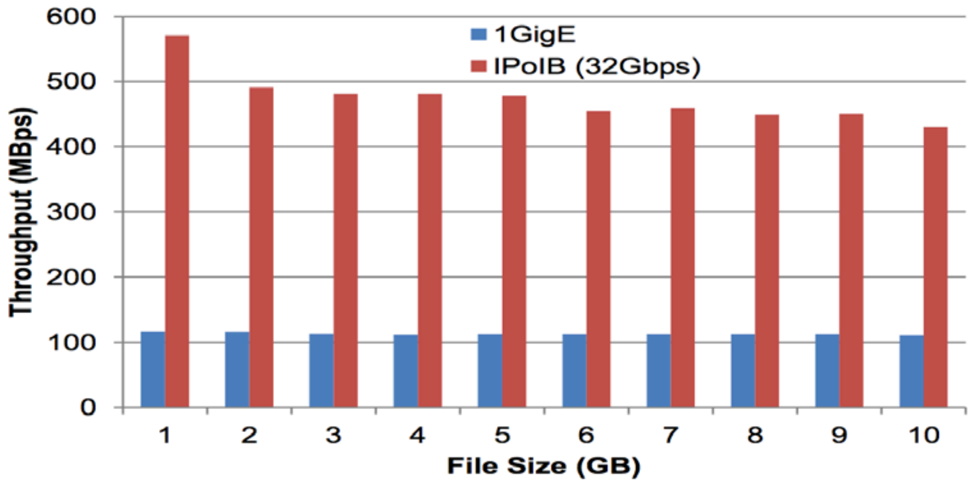
- Measure system performance (hardware and software) for a given application scenario—with given data and workload

Micro-benchmark: HDFS I/O operations

Islam, Lu, Rahman, Jose, Wang, Panda, *A Micro-benchmark suite for evaluating HDFS operations on modern cluster*, in Specifying Big Data Benchmark*, LNCS 8163, 2014*



File Write Latency with 32 DataNodes

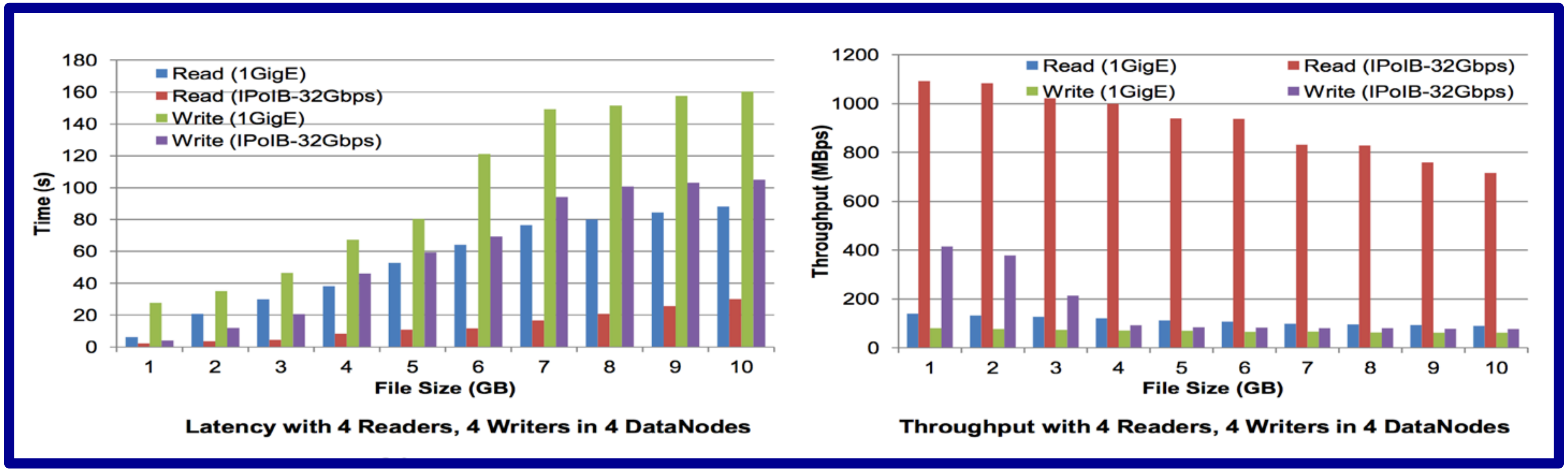


Write Throughput with 4 Writers in 32 DataNodes

Sequential writes

Micro-benchmark: HDFS I/O operations

Islam, Lu, Rahman, Jose, Wang, Panda, *A Micro-benchmark suite for evaluating HDFS operations on modern cluster*, in Specifying Big Data Benchmark*, LNCS 8163, 2014*



Sequential read/write throughput

Function-based: Sort

Sort Benchmark Home Page

http://sortbenchmark.org/

News: The 2014 sort benchmark contest is now closed. The official results for this year have not been announced. They will be announced shortly.

We have released the specification for a new sort benchmark, CloudSort, that measures the total cost of ownership for external sorts run in a public cloud.

Sort 100-byte records, first 10 bytes are the key

Benchmark variations:

- Minute Sort– # of records sorted in 1 minute
- Gray Sort – time taken to sort 100TB dataset
- CloudSort, PennySort, JouleSort

IBM InfoSphere BigInsights and Terasort

http://www-03.ibm.com/systems/platformcomputing/products/symphony/highperfhadoop.html
August 2012

Terasort Benchmark

- **Hardware**
- 200 IBM dx360M3 computers in iDataPlex racks
- 2 IBM dx360M3 computers in iDataPlex racks as master hosts
- 120 GB memory per host, 12 x 3 TB spindles per host
- 2,400 cores

- **Software**
- 1000 Virtual machines
- RHEL 6.2 with KVM
- IBM InfoSphere BigInsights 1.3.0.1
- IBM Platform Symphony Advanced Edition 5.2
- IBM Platform Symphony BigInsights Integration Path for 1.3.0.1

Fact Sheet

HP Unleashes the Power of Hadoop

Industry's first enterprise-ready Hadoop solution

HP achieves No. 1 performance benchmark for Hadoop

Based on results of the industry-standard Apache Hadoop Terasort benchmark, which is designed to demonstrate real world big data workloads, the HP Apache Hadoop solution is the first to deliver industry-leading performance for a 10-terabyte (TB) dataset processed in 5,128 seconds (approximately 1.5 hours). Built on HP Converged Infrastructure consisting of an 18-node HP ProLiant Generation 8 (Gen8) DL380 cluster and HP Networking, HP solutions proved to be 3.8 times and 2.6 times faster than Oracle and SGI Hadoop offerings, respectively.(1)

(1) As the first vendor to submit performance results for the 10TB Terasort benchmark, an 18-node cluster of HP ProLiant Gen8 DL380 servers sorted the 10TB data set in 5128 seconds, a rate of 1.99 gigabytes per second; it sorted the 100 gigabyte data set in 55 seconds at a rate of 1.82 gigabytes (GB) per second. On a per node basis, the HP ProLiant Gen8 DL380 was 3.8 times faster than Oracle's 2010 100GB result and 2.6 times faster than SGI's 100GB 2011 result. Hardware Configuration: 18 HP ProLiant DL380 Gen8 servers; Dual 6 core Intel® E5-2667 2.9GHz processors; 64 GB memory; 16 x 1 TB SAS 7.2K disks per node; 4 x 1GB Ethernet. Software Configuration: Red Hat Enterprise Linux 6.2; Java Platform, Standard Edition, JDK 6 Update 29-b11.

http://www.hp.com/hpinfo/newsroom/press_kits/2012/HPDiscover2012/Hadoop_Appliance_Fact_Sheet.pdf 2012

Cisco UCS Demonstrates Leading TeraSort Benchmark Performance

Performance Brief
August 2013

http://unleashingit.com/docs/B13/Cisco%20UCS/le_tera.pdf, August 2013

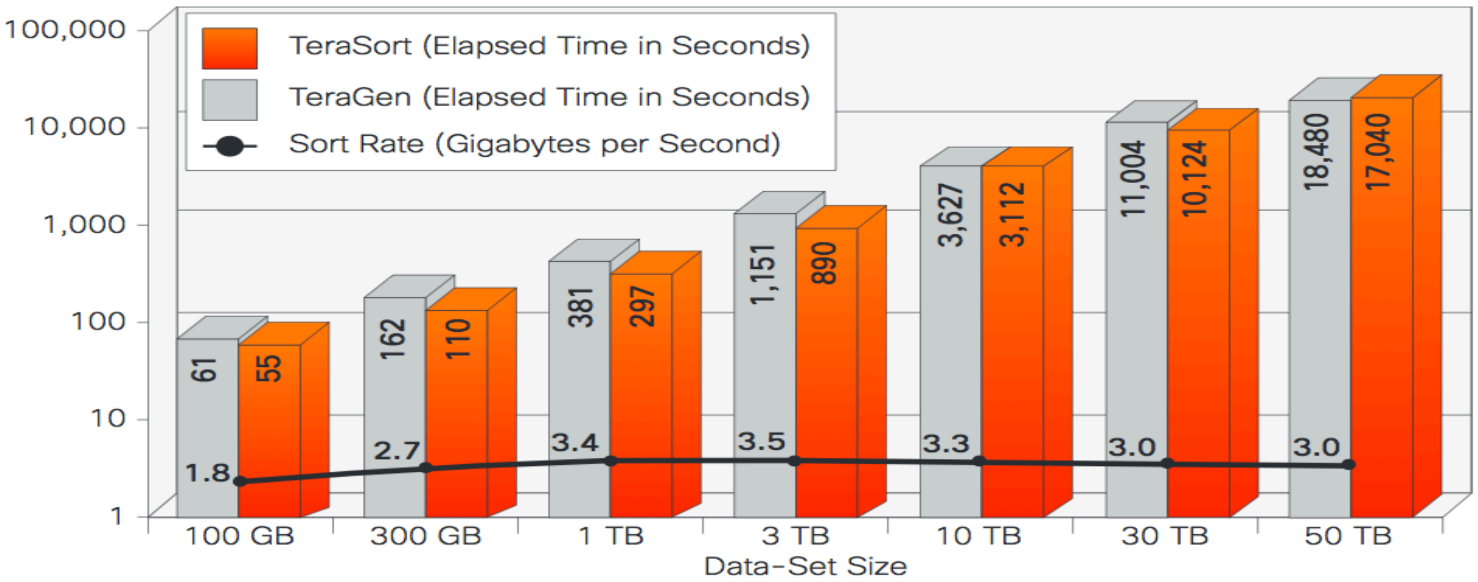


Figure 1. TeraGen and TeraSort on Cisco UCS CPA for Big Data.

TPCx-HS: Terasort-based TPC Benchmark

TPCx-HS: TPC Express for Hadoop Systems
- Based on kit; independent or peer review

Based on Terasort: Teragen, Terasort, Teravalidate

Database size / Scale Factors

ScaleFactor (in TBs)	1	3	10	30	100	300	1,000	3,000	10,000
# of 100-byte records (B)	10	30	100	300	1,000	3,000	10,000	30,000	100,000

Performance Metric
- HSph@SF = SF/T (total elapsed time in hours)

Price/Performance
- $/HSph, $ is 3-year total cost of ownership

Home | Complete Results | Benchmarks | Green Graph 500 | Log In

The Graph 500 List

June 2014

No.	Rank	Machine	Installation Site	Number of nodes	Number of cores	Problem scale	GTEPS
1	1	K computer (Fujitsu - Custom supercomputer)	RIKEN Advanced Institute for Computational Science (AICS)	65536	524288	40	17977.1
2	2	DOE/NNSA/LLNL Sequoia (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Lawrence Livermore National Laboratory	65536	1048576	40	16599
3	3	DOE/SC/Argonne National Laboratory Mira (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Argonne National Laboratory	49152	786432	40	14328
4	4	JUQUEEN (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Forschungszentrum Juelich (FZJ)	16384	262144	38	5848
5	5	Fermi (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	CINECA	8192	131072	37	2567

Graph benchmarking…

"… a memory efficient implementation for the NVM-based Hybrid BFS algorithm … demonstrate extremely fast BFS execution for large-scale unstructured graphs whose size exceed the capacity of DRAM on the machine.

Experimental results of Kronecker graphs compliant to the Graph500 benchmark on a 2-way INTEL Xeon E5-2690 machine with 256 GB of DRAM

Our proposed implementation can achieve 4.14 GTEPS for a SCALE31 graph problem with 231 vertices and 235 edges, whose size is 4 times larger than the size of graphs that the machine can accommodate only using DRAM with only 14.99 % performance degradation.

We also show that the power efficiency of our proposed implementation achieves 11.8 MTEPS/W.

Based on the implementation, we have achieved the 3rd and 4th position of the Green Graph500 list (2014 June) in the Big Data category.

--from *NVM-based Hybrid BFS with Memory Efficient Data Structure,* Keita Iwabuchi, Hitoshi Sato, Yuichiro Yasui, Fujisawa, and Matsuoka, IEEE BigData 2014

Tutorial Overview

Tutorial overview

Introduction to Big Data benchmarking issues

Different levels of benchmarking

Survey of some Big Data Benchmarking initiatives

BREAK

Discussion of *BigBench*

Discussion of the *Deep Analytics Pipeline*

Next Steps, Future Directions

Other Big Data Benchmark Initiatives

HiBench, Yan Li, Intel

Yahoo Cloud Serving Benchmark, Brian Cooper, Yahoo!

Berkeley Big Data Benchmark, Pavlo et al., AMPLab

BigDataBench, Jianfeng Zhan, Chinese Academy of Sciences

HiBench, Lan Yi, Intel, 4th WBDB

1 *Micro Benchmarks*

- Sor
- Wo
- Ter

2 *Web Search*

- xing

3 *Ma*

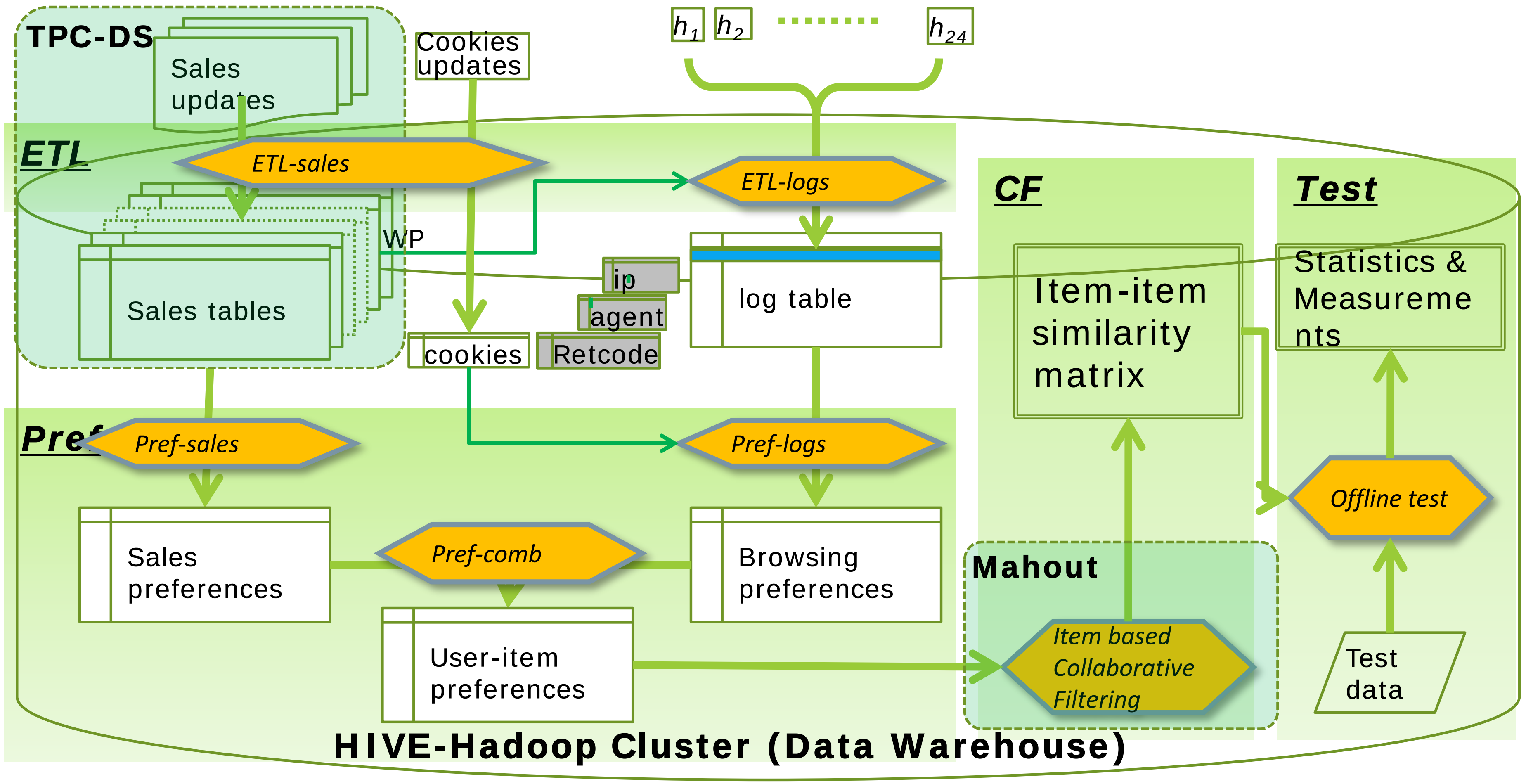
- Bayesian Classification
- K-Means Clustering

- Enhanced DFSIO

1. Different from GridMix, SWIM?
2. Micro Benchmark?
3. Isolated components?
4. End-2-end Benchmark?
5. We need ETL-Recommendation Pipeline

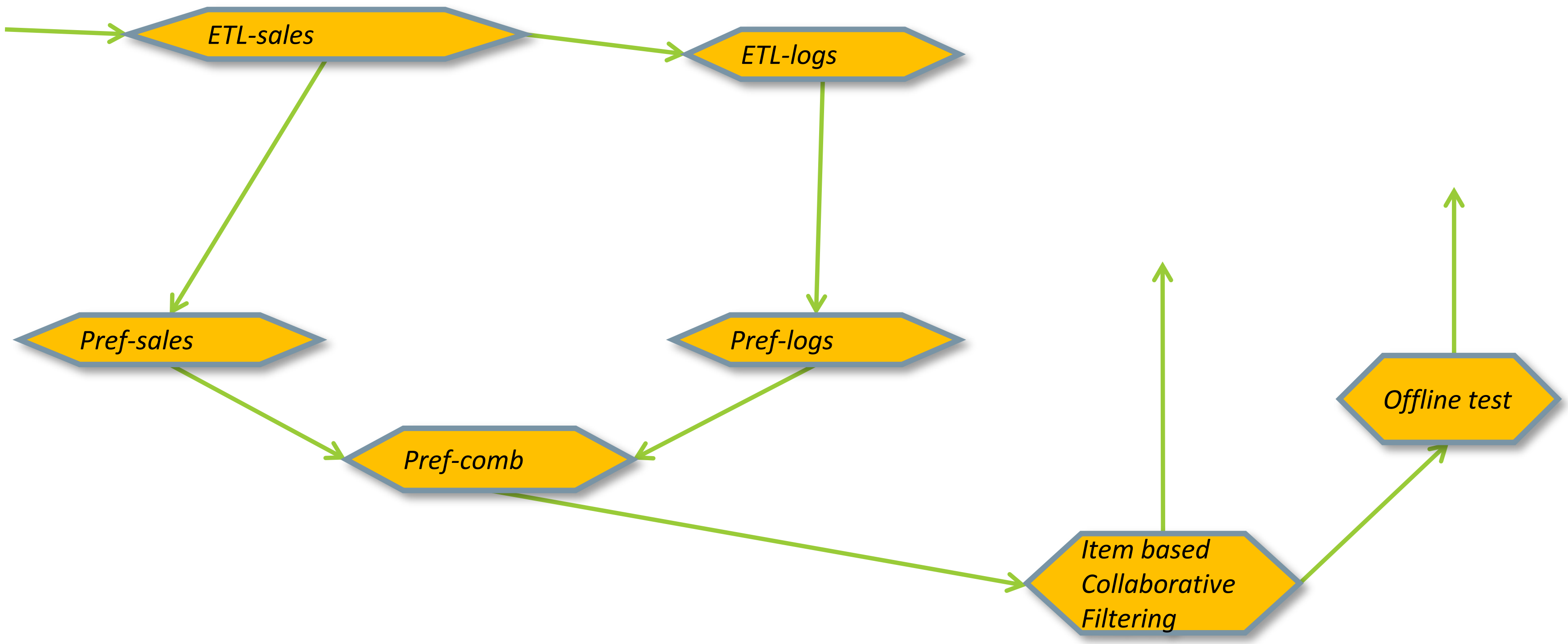
See paper "The HiBench Suite: Characterization of the MapReduce-Based Data Analysis" in ICDE'10 workshops (WISS'10)

ETL-Recommendation (hammer)



ETL-Recommendation (hammer)

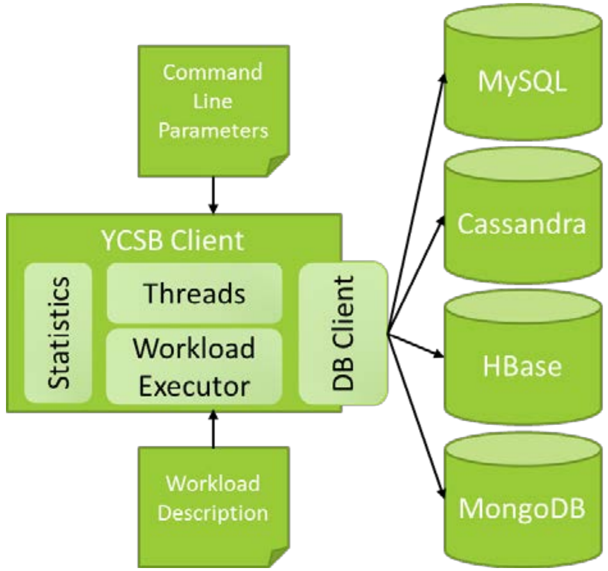
Task Dependences

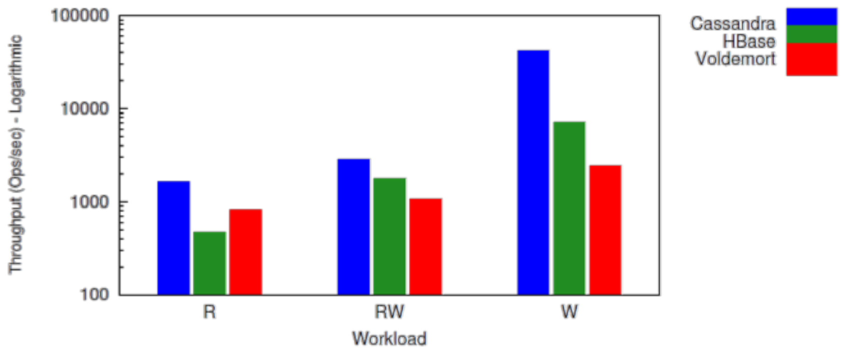


Yahoo! Cloud Serving Benchmark

Key-value store benchmark

- CRUD operations (insert, read, update, delete, scan)
- Single table
  - Usertable
  - user[Number], random string values
- Different access distributions
  - Uniform, Zipfian, latest, hot set
- Many database connectors
  - Accumulo, Cassandra, HBase, HyperTable, JDBC, Redis, …
- Many extensions
  - YCSB++, YCSB+T, various forks





https://github.com/brianfrankcooper/YCSB/

Berkeley Big Data Benchmark

"A comparison of approaches to large-scale data analysis"

A.k.a. CALDA

- 2 simple queries with varying result set size (BI-like, intermediate, ETL-like)
  - `SELECT pageURL, pageRank FROM rankings WHERE pageRank > X`
  - `SELECT SUBSTR(sourceIP, 1, X), SUM(adRevenue) FROM uservisits GROUP BY SUBSTR(sourceIP, 1, X)`
- Join query
  - Join Rankings and UserVisits
- UDF query
  - URL count on Documents

https://amplab.cs.berkeley.edu/benchmark/

SF	#Rankings	Rankings B	#UserVisits	UserVisits B	DocumentsB
small	1200	77.6KB	10000	1.7MB	6.8MB
1	18 Million	1.28GB	155 Million	25.4GB	29.0GB
5	90 Million	6.38GB	775 Million	126.8GB	136.9GB

Documents	Rankings	UserVisits
Unstructured HTML documents	Lists websites and their page rank	Stores server logs for each web page
	pageURL VARCHAR(300) pageRank INT avgDuration INT	sourceIP VARCHAR(116) destURL VARCHAR(100) visitDate DATE adRevenue FLOAT userAgent VARCHAR(256) countryCode CHAR(3) languageCode CHAR(6) searchWord VARCHAR(32) duration INT

BigDataBench

Mashup of many benchmarks

- Collection of popular data sets and workloads

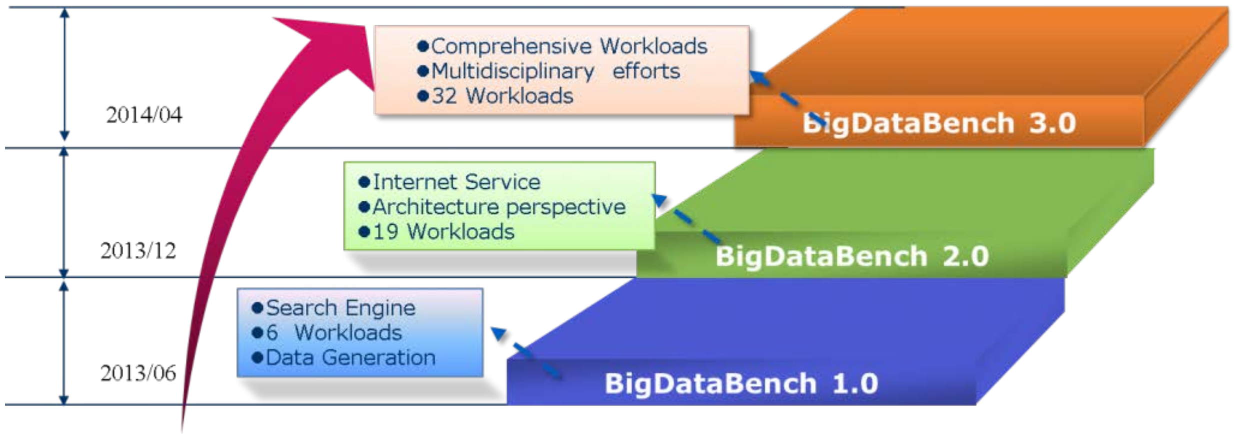
Synthetic and real data sets

- 6 real world
- 2 synthetic

32 workloads

Under active development and expansion

data sets	data size	
1	Wikipedia Entries	4,300,000 English articles
2	Amazon Movie Reviews	7,911,684 reviews
3	Google Web Graph	875713 nodes, 5105039 edges
4	Facebook Social Network	4039 nodes, 88234 edges
5	E-commerce Transaction Data	Table1: 4 columns, 38658 rows. Table2: 6 columns, 242735 rows
6	ProfSearch Person Resumes	278956 resumes
7	CALDA Data (synthetic data)	Table1: 3 columns. Table2: 9 columns.
8	TPC-DS Web Data (synthetic data)	26 tables



BigDataBench Workloads

Cloud OLTP
- YCSB like

Offline analytics
- HiBench
- MR batch jobs

OLAP and interactive analytics
- CALDA
- TPC-DS excerpt

Mix and match for your use-case

http://prof.ict.ac.cn/BigDataBench/

Application Types	Benchmark Types	Workloads	Data Sets	Software Stacks
Cloud OLTP	Micro Benchmarks	Read	*ProfSearch Person Resumes: Semi-structured Table*	HBase, Mysql
		Write		
		Scan		
	Application Benchmarks	Search Server	*Wikipedia Entries: Semi-structured Text*	HBase, Nutch
Offline Analytics	Micro Benchmarks	Sort	*Wikipedia Entries*	MPI, Spark, Hadoop
		Grep		
		WordCount		
		BFS	*Graph500 data set: Unstructured Graph*	MPI
	Application Benchmarks	Index	*Wikipedia Entries*	MPI, Spark, Hadoop
		PageRank	Unstructured Graph(*Google WebGraph)*	
		Kmeans	*Google Web Graph and Facebook Social Network: Unstructured Graph*	
		Connected Components		
		Collaborative Filtering	*Amazon Movie Reviews: Semi-structured Text*	
		Naive Bayes		
OLAP and Interactive Analytics	Micro Benchmarks	Project	*E-commerce Transaction data, CALDA data andTPC-DS Web data: Structured Table*	Mysql, Hive, Shark, Impala
		Filter		
		OrderBy		
		Cross Product		
		Union		
		Difference		
		Aggregation		
	Application Benchmarks	Join Query		
		Select Query		
		Aggregation Query		
		Eight TPC-DS Web Queries		

Tutorial Overview

Tutorial overview
Introduction to Big Data benchmarking issues
Different levels of benchmarking
Survey of some Big Data Benchmarking initiatives

BREAK

Discussion of *BigBench*
Discussion of the *Deep Analytics Pipeline*
Next Steps, Future Directions

Tutorial Overview

Tutorial overview

Introduction to Big Data benchmarking issues

Different levels of benchmarking

Survey of some Big Data Benchmarking initiatives

BREAK

Discussion of *BigBench*

Discussion of the *Deep Analytics Pipeline*

Next Steps, Future Directions

The BigBench Proposal

End to end benchmark
- Application level

Based on a product retailer (TPC-DS)

Focused on Parallel DBMS and MR engines

History
- Launched at 1st WBDB, San Jose
- Published at SIGMOD 2013
- Spec at WBDB proceedings 2012 (queries & data set)
- Full kit at WBDB 2014

Collaboration with Industry & Academia
- First: Teradata, University of Toronto, Oracle, InfoSizing
- Now: bankmark, CLDS, Cisco, Cloudera, Hortonworks, Infosizing, Intel, Microsoft, MSRG, Oracle, Pivotal, SAP, IBM

Derived from TPC-DS

Multiple snowflake schemas with shared dimensions

24 tables with an average of 18 columns

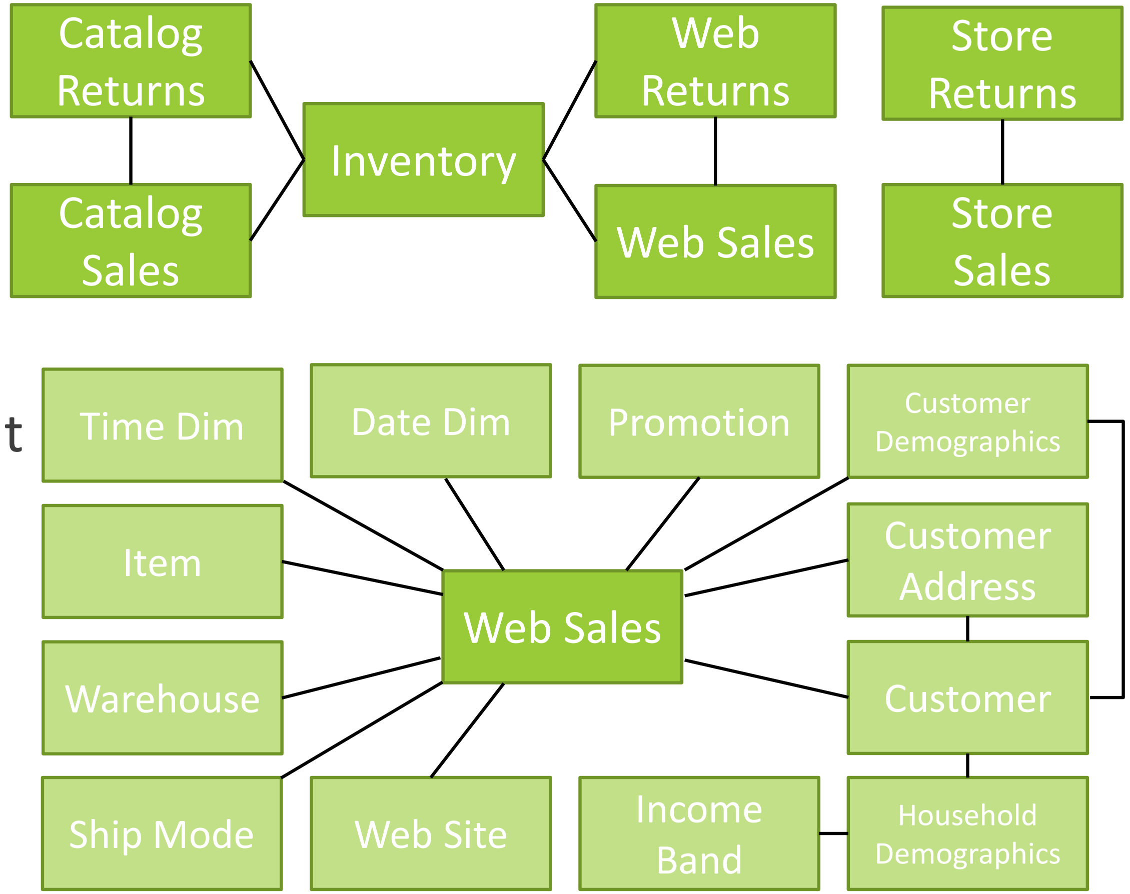
99 distinct SQL 99 queries with random substitutions

More representative skewed database content

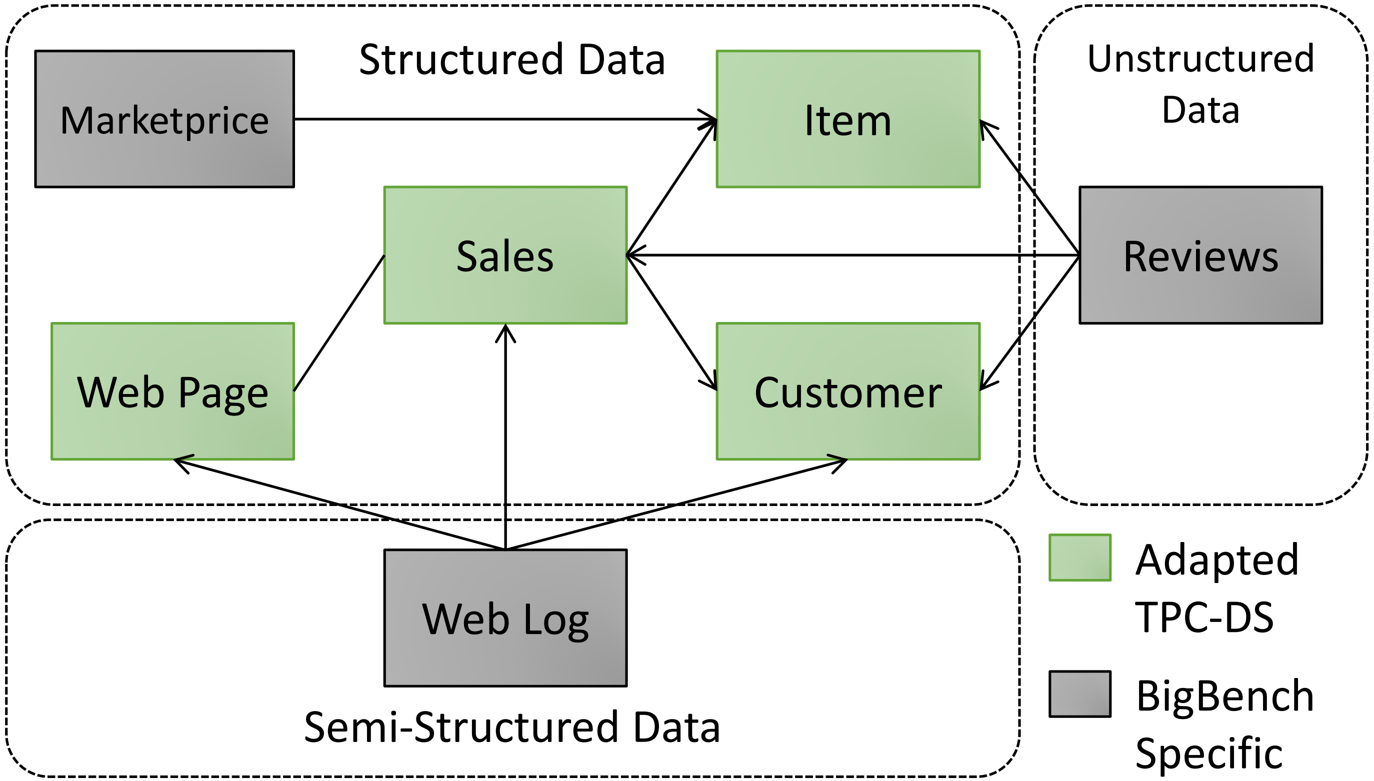
Sub-linear scaling of non-fact tables

Ad-hoc, reporting, iterative and extraction queries

ETL-like data maintenance



BigBench Data Model



- Structured: TPC-DS + market prices
- Semi-structured: website click-stream
- Unstructured: customers' reviews

Data Model – 3 Vs

Variety

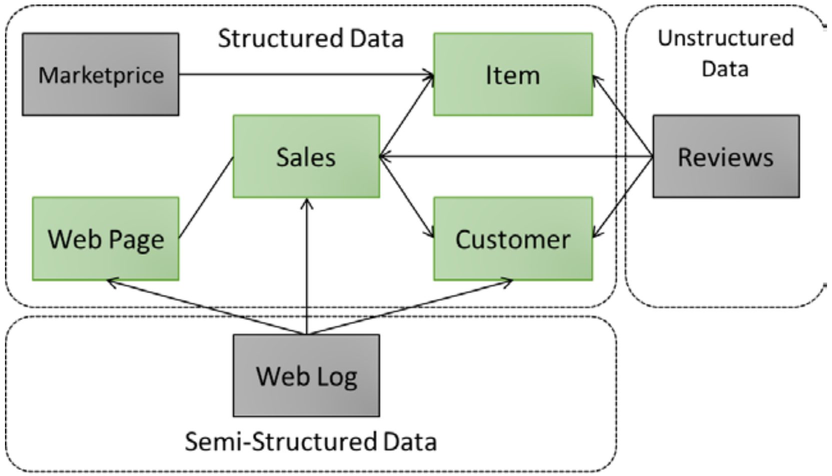
- Different schema parts

Volume

- Based on scale factor
- Similar to TPC-DS scaling, but continuous
- Weblogs & product reviews also scaled

Velocity

- Refresh for all data with different velocities



Scaling

Continuous scaling model

- Realistic

SF 1 ~ 1 GB

Different scaling speeds

- Adapted from TPC-DS
  - Static
  - Square root
  - Logarithmic
  - Linear (LF)

$$LF = SF + (SF - (\log_5(SF) * \sqrt{SF})) = 2SF - \log_5(SF) * \sqrt{SF}$$

Table Name	# Rows SF 1	Bytes/Row	Scaling
date	109573	141	static
time	86400	75	static
ship_mode	20	60	static
household_demographics	7200	22	static
customer_demographics	1920800	40	static
customer	100000	138	square root
customer_address	50000	107	square root
store	12	261	square root
warehouse	5	107	logarithmic
promotion	300	132	logarithmic
web_page	60	134	logarithmic
item	18000	308	square root
item_marketprice	90000	43	square root
inventory	23490000	19	square root * logarithmic
store_sales	810000	143	linear
store_returns	40500	125	linear
web_sales	810000	207	linear
web_returns	40500	154	linear
web_clickstreams	6930000	27	linear
product_reviews	98100	670	linear

Generating Big Data

Repeatable computation

- Based on XORSHIFT random number generators

Hierarchical seeding strategy

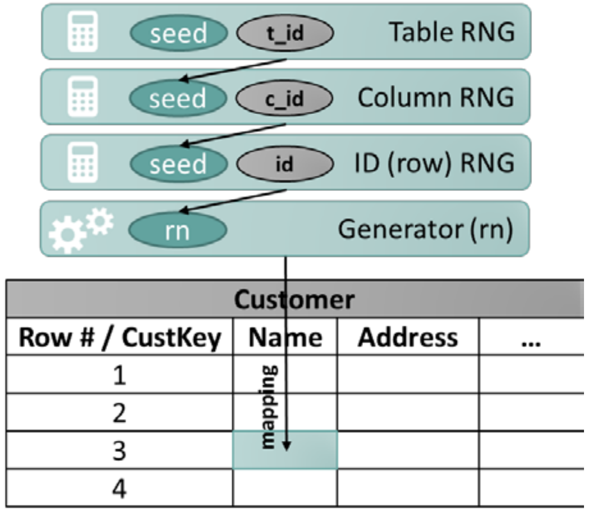
- Enables independent generation of *every* value in the data set
- Enables independent *re-generation* of every value for references

User specifies

- Schema – data model
- Format – CSV, SQL statements, …
- Distribution – multi-core, multi-node, partially

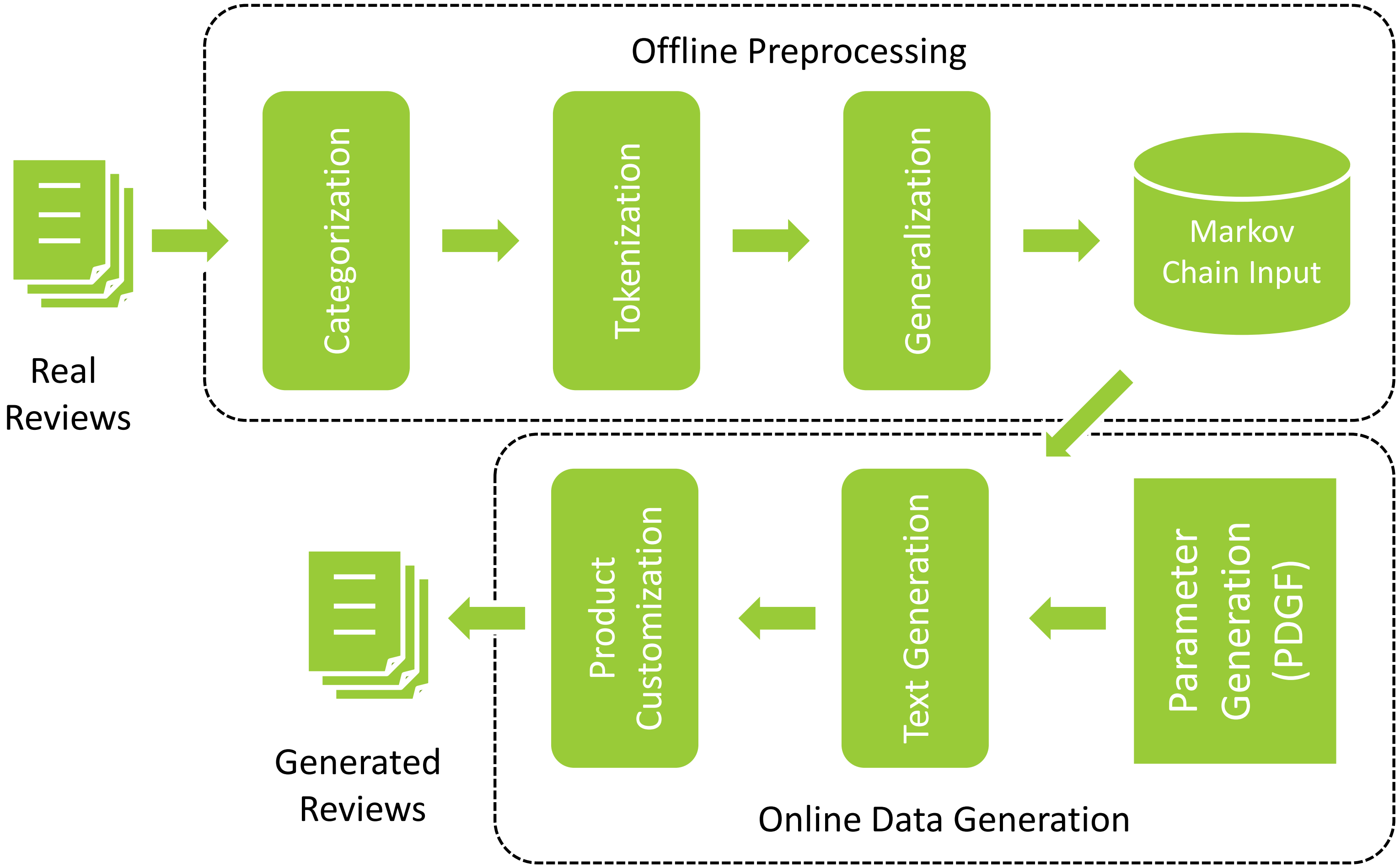
PDGF generates

- High quality data – distributed, in parallel, in the correct format
- Large data – terabytes, petabytes




```xml
<property name="METERS_DAY" type="double">1000000</property>
<table name="smartgrid">
  <size>${METERS_DAY}</size>
  <field name="id" size="" type="NUMERIC"
         primary="true" unique="true">
    <gen_IdGenerator/>
  </field>
  <field name="territory" type="NUMERIC">
    <gen_LongGenerator>
      <min>100000000</min>
      <max>999999999</max>
    </gen_LongGenerator>
  </field>
  <field name="regionnumber" type="NUMERIC">
    <gen_SequentialGenerator concatenate="true">
      <gen_OtherFieldValueGenerator>
        <reference field="territoryhierarchy" />
      </gen_OtherFieldValueGenerator>
      <gen_LongGenerator>
        <min>1000000000000.0</min>
        <max>9999999999999.0</max>
      </gen_LongGenerator>
    </gen_SequentialGenerator>
  </field>
[..]
```

Review Generation



Workload

Workload Queries
- 30 “queries”
- Specified in English (sort of)
- No required syntax (first implementation in Aster SQL MR)
- Kit implemented in Hive, HadoopMR, Mahout, OpenNLP

Business functions (Adapted from McKinsey)
- **Marketing**
  - Cross-selling, Customer micro-segmentation, Sentiment analysis, Enhancing multichannel consumer experiences
- **Merchandising**
  - Assortment optimization, Pricing optimization
- **Operations**
  - Performance transparency, Product return analysis
- **Supply chain**
  - Inventory management
- **Reporting (customers and products)**

Workload - Technical Aspects

GENERIC CHARACTERISTICS

Data Sources	#Queries	Percentage
Structured	18	60%
Semi-structured	7	23%
Un-structured	5	17%

HIVE IMPLEMENTATION CHARACTERISTICS

Query Types	#Queries	Percentage
Pure HiveQL	14	46%
Mahout	5	17%
OpenNLP	5	17%
Custom MR	6	20%

SQL-MR Query 1

```
SELECT category_cd1 AS category1_cd,
       category_cd2 AS category2_cd , COUNT (*) AS cnt
FROM basket_generator (
      ON
              ( SELECT i.i_category_id AS category_cd ,
                       s.ws_bill_customer_sk AS customer_id
              FROM web_sales s INNER JOIN item i
              ON s.ws_item_sk = i.item_sk )
       PARTITION BY customer_id
       BASKET_ITEM (‘category_cd')
       ITEM_SET_MAX (500)
)
GROUP BY 1,2
ORDER BY 1, 3, 2;
```

HiveQL Query 1

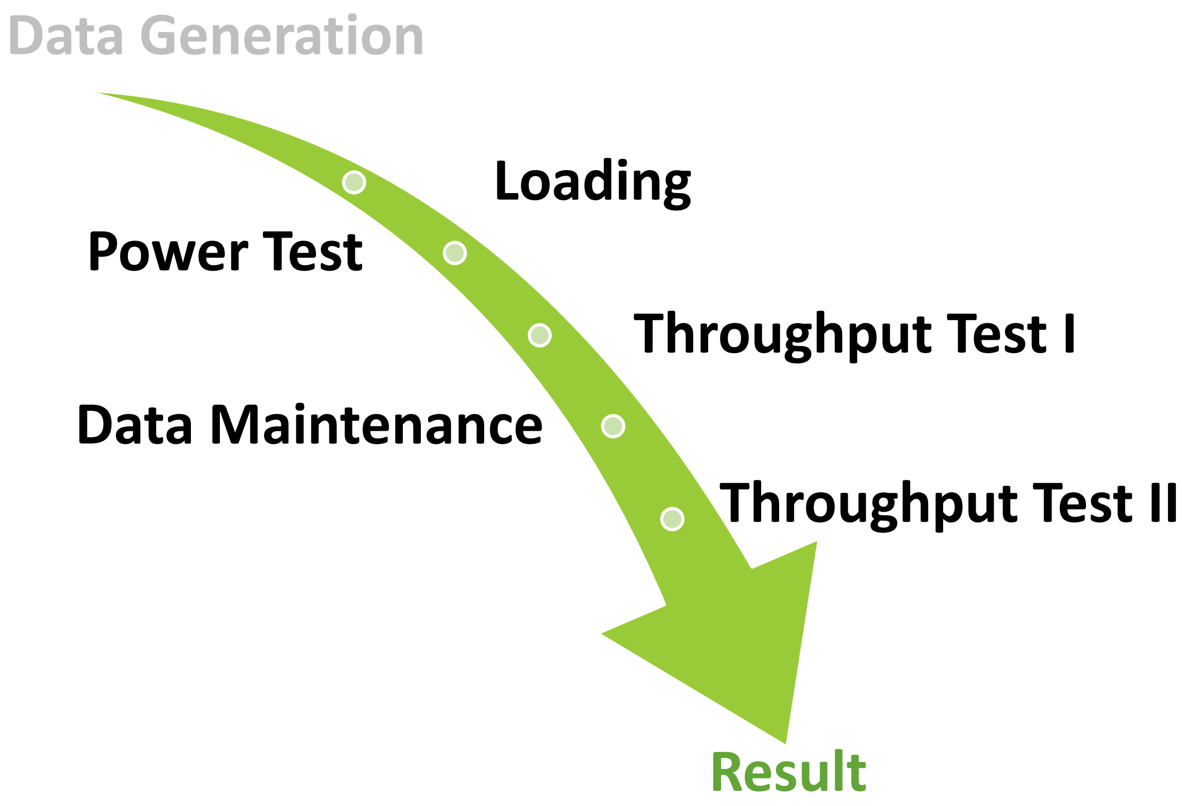
```
SELECT     pid1, pid2, COUNT (*) AS cnt
FROM (
           FROM (
                      FROM (
                                 SELECT s.ss_ticket_number AS oid , s.ss_item_sk AS pid
                                 FROM store_sales s
                                 INNER JOIN item i ON s.ss_item_sk = i.i_item_sk
                                 WHERE i.i_category_id in (1 ,2 ,3) and s.ss_store_sk in (10 , 20, 33, 40, 50)
                      ) q01_temp_join
                      MAP q01_temp_join.oid, q01_temp_join.pid
                      USING 'cat'
                      AS oid, pid
                      CLUSTER BY oid
           ) q01_map_output
           REDUCE q01_map_output.oid, q01_map_output.pid
           USING 'java -cp bigbenchqueriesmr.jar:hive-contrib.jar de.bankmark.bigbench.queries.q01.Red'
           AS (pid1 BIGINT, pid2 BIGINT)
) q01_temp_basket
GROUP BY pid1, pid2
HAVING COUNT (pid1) > 49
ORDER BY pid1, cnt, pid2;
```

Benchmark Process

Adapted to batch systems
- No trickle update

Measured processes
- Loading
- Power Test (single user run)
- Throughput Test I (multi user run)
- Data Maintenance
- Throughput Test II (multi user run)

Result
- Additive Metric



Metric

Throughput metric

- BigBench queries per hour

Number of queries run

- 30*(2*S+1)

Measured times

T_L: Execution time of the loading process;
T_P: Execution time of the power test;
T_{TT_1}: Execution time of the first throughput test;
T_{DM}: Execution time of the data maintenance task.
T_{TT_2}: Execution time of the second throughput test;

Metric

$$BBQpH = \frac{30 * 3 * 3600}{T_L + T_P + \frac{T_{TT1}}{S} + T_{DM} + \frac{T_{TT2}}{S}}$$

$$BBQpH = \frac{30 * 3 * S * 3600}{S * T_L + S * T_P + T_{TT1} + S * T_{DM} + T_{TT2}}$$

BigBench Experiments

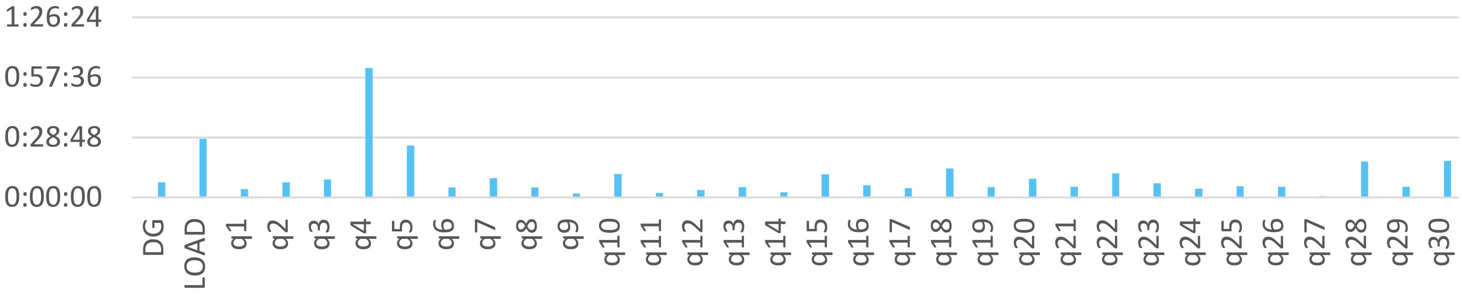
Tests on

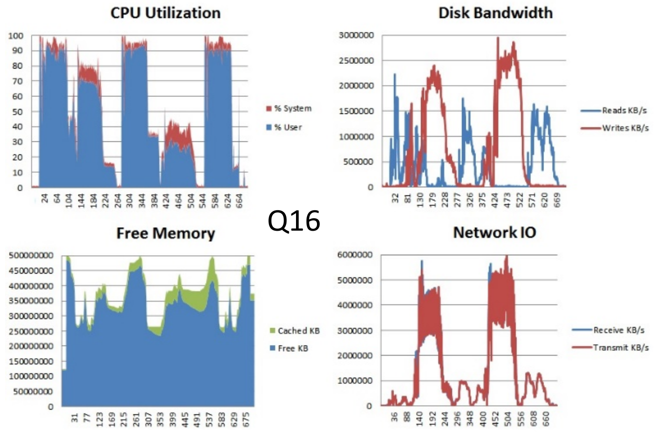
- Cloudera CDH 5.0, Pivotal GPHD-3.0.1.0, IBM InfoSphere BigInsights

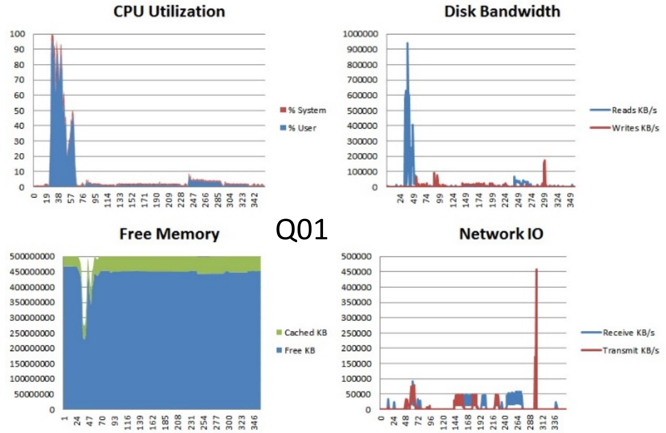
In progress: Spark, Impala, Stinger, …

3 Clusters (+)

- 1 node: 2x Xeon E5-2450 0 @ 2.10GHz, 64GB RAM, 2 x 2TB HDD
- 6 nodes: 2 x Xeon E5-2680 v2 @2.80GHz, 128GB RAM, 12 x 2TB HDD
- 546 nodes: 2 x Xeon X5670 @2.93GHz, 48GB RAM, 12 x 2TB HDD







Towards an Industry standard

Collaboration with SPEC and TPC

Aiming for fast development cycles

Enterprise vs express benchmark (TPC)

Enterprise	Express
Specification	Kit
Specific implementation	Kit evaluation
Best optimization	System tuning (not kit)
Complete audit	Self audit / peer review
Price requirement	No pricing
Full ACID testing	ACID self-assessment (no durability)
Large variety of configuration	Focused on key components
Substantial implementation cost	Limited cost, fast implementation

Specification sections (**open**)

- Preamble
  - High level overview
- Database design
  - Overview of the schema and data
- Workload scaling
  - How to scale the data and workload
- Metric and **execution rules**
  - Reported metrics and rules on how to run the benchmark
- **Pricing**
  - **Reported price information**
- **Full disclosure report**
  - **Wording and format of the benchmark result report**
- **Audit requirements**
  - **Minimum audit requirements for an official result, self auditing scripts and tools**

Use BigBench

Try BigBench

- https://github.com/intel-hadoop/Big-Bench
- Only Hive implementation for now
- More to come soon

Bring some time

- Full BigBench run on Hive takes 2 days+
- Will verify if your cluster is setup correctly

Future extensions (among others)

- Graph queries
- More procedural workload

Tutorial Overview

Tutorial overview

Introduction to Big Data benchmarking issues

Different levels of benchmarking

Survey of some Big Data Benchmarking initiatives

BREAK

Discussion of *BigBench*

Discussion of the *Deep Analytics Pipeline*

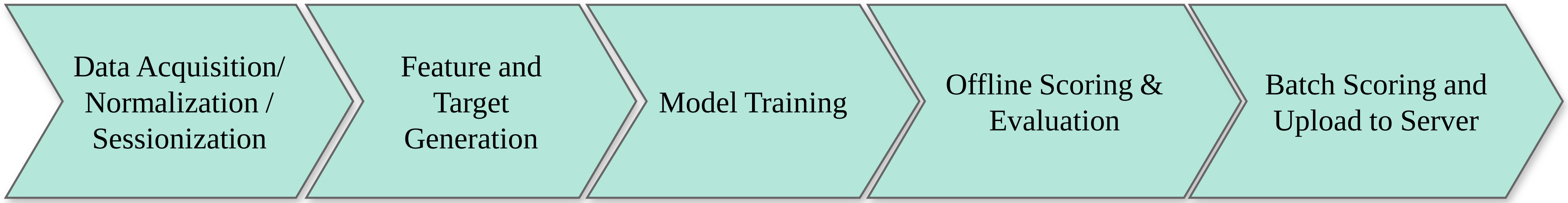
Next Steps, Future Directions

“Deep Analytics Pipelines”

“User Modeling” pipelines

Generic use case: Determine user interests or user categories by mining user activities

- Large dimensionality of possible user activities
- Typical user represents a sparse activity vector
- Event attributes change over time



Courtesy: Milind Bhandarkar, Pivotal

Analytics Pipelines: Example Application Domains

Retail
- Events: clicks on purchases, ad clicks, FB likes, …
- Goal: Personalized product recommendations

Datacenters
- Events: log messages, traffic, communications events, …
- Goal: Predict imminent failures

Healthcare
- Events: Doctor visits, medical history, medicine refills, …
- Goal: Prevent hospital readmissions

Telecom
- Events: Calls made, duration, calls dropped, location, social graph, …
- Goal: Reduce customer churn

Web Ads
- Events: Clicks on content, likes, reposts, search queries, comments, …
- Goal: Increase engagement, increase clicks on revenue-generation content

Courtesy: Milind Bhandarkar, Pivotal

User/event modeling pipelines

Acquisition and normalization of data
- Collate, consolidate data

Join targets and features
- Construct targets; filter out user activity without targets; join feature vector with targets

Model Training
- Multi-model: regressions, Naïve Bayes, decision trees, Support Vector Machines, …

Offline scoring
- Score features, evaluate metrics

Batch scoring
- Apply models to all user activity; upload scores

Courtesy: Milind Bhandarkar, Pivotal

Need fact (user event) data with appropriate distributions

Browsing page views
- E.g., UserID, geo-location, pageID, <list of pages>

Search queries
- E.g., userID, TS, geo-location, queryID, queryString, <ranked list of pageID’s>

Ads – search and display ads associated with queries and page views
- E.g., userID, TS, geo-location, pageID, <list of ad_ID>

Clicks – user clicks, with delays
- E.g., userID, TS, <ad_ID>

Social Graph data
- E.g. twitter user/follower

Courtesy: Vijay Narayanan, MSR

Application Classes & Scaling Factors

Widely varying number of events per entity

Multiple classes of applications, based on size, e.g.:

Size	#Unique Users	#Unique Pages	#Unique Queries	#Unique SearchAds	#Unique DisplayAds	Social Graph Edges	Avg # Events/user in time window
Tiny	100K	100	1K	3K	300	300K	10
Small	1M	1K	10K	10K	1K	5M	30
Medium	10M	5K	500K	500K	5K	100M	100
Large	100M	20K	2M	2.5M	10K	1B	1000
Huge	1B	50K	10M	10M	50K	100B	1000

Example Benchmark Metrics

Publish results for every stage in the pipeline

- Data pipelines for different application domains may be constructed by mix and match of various pipeline stages

Quality of result for a fixed budget, within a maximum time

Optimize the model quality within a fixed time

Optimize model quality for a fixed budget

Tutorial Overview

Tutorial overview

Introduction to Big Data benchmarking issues

Different levels of benchmarking

Survey of some Big Data Benchmarking initiatives

BREAK

Discussion of *BigBench*

Discussion of the *Deep Analytics Pipeline*

Next Steps, Future Directions

Big Data Needs Big Systems

Too much big data research is done on tiny data sets / systems / installations

This is not necessary!

Large Scale Hadoop Testbed – Greenplum Analytics Workbench

Test on the EMC/Pivotal 1000 node cluster
- Intel
  - Contributed 2,000 6-core CPUs.
- Mellanox
  - Contributed >1,000 network cards and 72 switches.
- Micron
  - Contributed 6,000 8GB DRAM modules.
- Seagate
  - Contributed 12,000 2TB Drives

Hardware
- Physical Hosts - More than 1,000 nodes
- Processors - Over 24,000 CPU's
- RAM – Over 48TB for memory
- Disk capacity – More than 24PB of raw storage

Year round applications: www.analyticsworkbench.com

Source: www.analyticsworkbench.com

HPI Future SOC Lab

Top notch hardware for free at theHasso Plattner Institute

- 1000 Core Cluster (25 nodes) with a total of 25 TB RAM and 75 TB SSD.
- Multi-Core servers with up to 2 TB RAM and 64 Cores.
- In-Memory Computing SAP HANA.
- Hewlett-Packard Converged Cloud.
- Coprocessors NVIDIA Tesla K20X and Intel Xeon Phi.

Apply twice per year for 6 months of access: www.hpi.de/future-soc-lab

HPI Hasso Plattner Institut

IT Systems Engineering | Universität Potsdam

Future SOC Lab

NSFCloud Platforms

NSF-funded Cloud Platforms

- *August 20, 2014*: The National Science Foundation (NSF) today announced two $10 million projects to create cloud computing testbeds--to be called "Chameleon" and "CloudLab"--that will enable the academic research community to develop and experiment with novel cloud architectures and pursue new, architecturally-enabled applications of cloud computing. http://nsf.gov/news/news_summ.jsp?cntn_id=132377

Community workshop, December 11-12, 2014

NSFCloud Workshop on Experimental Support for Cloud Computing

About | Participation | Organizers | Venue | Agenda

About

Workshop: NSFCloud Workshop on Experimental Support for Cloud Computing

Where: Waterview Conference Center in Arlington, VA

When: December 11-12, 2014

Formation of SPEC Research Big Data Working Group

Mission Statement

The mission of the Big Data (BD) working group is to facilitate research and to engage industry leaders for defining and developing performance methodologies of big data applications. The term ‘‘big data’’ has become a major force of innovation across enterprises of all sizes. New platforms, claiming to be the “big data” platform with increasingly more features for managing big datasets, are being announced almost on a weekly basis. Yet, there is currently a lack of what constitutes a big data system and any means of comparability among such systems.

Initial Committee Structure

- Tilmann Rabl (Chair)
- Chaitan Baru (Vice Chair)
- Meikel Poess (Secretary)

Replaces less formal BDBC group

SPEC RG Big Data Agenda

Topics of Interest include, but are not limited to, the following

- Big data vs. very large databases
- Measurements aspects and goals
- Run rules and metrics
- Tools and kits
- Collaboration and synergies in benchmarking efforts
- Filling gaps in current big data benchmarking

Sharing information on big data benchmarking related topics

- Weekly calls, internal and public calls alternating
- Wednesday, 10 am pacific time

SPEC RG Big Data Membership

BDBC continues to be free

- Public calls / presentations / collaborations / workshops / mailing list

SPEC RG Big Data is part of the SPEC Research Group

- Internal calls, voting, organization, mailing list require membership
- USD $200 p.a. (unless your company is already SPEC member)

Mailing lists

- bdbc@sdsc.edu -> “old” BDBC mailing list
  - We will continue to distribute open information there, ask Chaitan to be part
- bigdata@spec.org -> new SPEC internal mailing list
  - Requires SPEC membership, has to be activated by SPEC representative

http://research.spec.org/working-groups/big-data-working-group.html

BigData Top100 List

Modeled after Top500 and Graph500 in HPC community

Proposal presented at Strata Conference, February 2013

Based on application-level benchmarking

Article in inaugural issue of the Big Data Journal

- *Big Data Benchmarking and the Big Data Top100 List* by Baru, Bhandarkar, Nambiar, Poess, Rabl, Big Data Journal, Vol.1, No.1, 60-64, Anne Liebert Publications.

In progress

Thank You !

Slides will be available online: www.msrg.org

Contact:

Tilmann Rabl, tilmann.rabl@utoronto.ca

Chaitan Baru, baru@sdsc.edu