Crafting Benchmarks for Big Data

Tilmann Rabl
Middleware Systems Research Group & bankmark UG
ISC’14, June 26, 2014

MIDDLEWARE SYSTEMS
RESEARCH GROUP
MSRG.ORG

bankmark

Outline

- **Big Data Benchmarking Community**
  - Our approach to building benchmarks
- **Big Data Benchmarks**
  - Characteristics
  - BigBench
  - Big Decisions
  - Hammer
  - DAP
- ***Slides borrowed from Chaitan Baru***

Big Data Benchmarking Community

- **Genesis of the Big Data Benchmarking effort**
  - Grant from NSF under the Cluster Exploratory (CluE) program (Chaitan Baru, SDSC)
  - Chaitan Baru (SDSC), Tilmann Rabl (University of Toronto), Milind Bhandarkar (Pivotal/Greenplum), Raghu Nambiar (Cisco), Meikel Poess (Oracle)
- **Launched Workshops on Big Data Benchmarking**
  - First WBDB: May 2012, San Jose. Hosted by Brocade
- **Objectives**
  - Lay the ground for development of industry standards for measuring the effectiveness of hardware and software technologies dealing with big data
  - Exploit synergies between benchmarking efforts
  - Offer a forum for presenting and debating platforms, workloads, data sets and metrics relevant to big bata
- **Big Data Benchmark Community (BDBC)**
  - Regular conference calls for talks and announcements
  - Open to anyone interested, free of charge
  - BDBC makes no claims to any developments or ideas
  - clds.ucsd.edu/bdbc/community

1st WBDB: Attendee 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
- San Diego Supercomputer Center
- SAS
- Scripps Research Institute
- Seagate
- Shell
- SNIA
- Teradata Corporation
- Twitter
- UC Irvine
- Univ. of Minnesota
- Univ. of Toronto
- Univ. of Washington
- VMware
- WhamCloud
- Yahoo!

Further Workshops

...ne, India

Xi'an, China

San Jose, CA, USA

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

Potsdam, Germany

5th WBDB: http://clds.sdsc.edu/wbdb2014.de

First 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

Further Outcomes

- **Selected papers in Springer Verlag, *Lecture Notes in Computer Science,* Springer Verlag**
  - LNCS 8163: Specifying Big Data Benchmarks (covering the first and second workshops)
  - LNCS 8585: Advancing Big Data Benchmarks (covering the third and fourth workshops, in print)
  - 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**
- **Proposal of BigData Top100 List**
- **Specification of BigBench**

TPC Big Data Subcommittee

- **TPCx-HS**
  - TPC Express for Hadoop Systems
- **Based on Terasort**
  - Teragen, Terasort, Teravalidate
- **Database size / Scale Factors**
  - SF: 1, 3, 10, 30, 100, 300, 1000, 3000, 10000 TB
  - Corresponds to: 10B, 30B, 100B, 300B, 1000B, 3000B, 10000B, 30000B, 100000B 100-byte records
- **Performance Metric**
  - HSph@SF = SF/T (total elapsed time in hours)
- **Price/Performance**
  - $/HSph, $ is 3-year total cost of ownership

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)
- **To replace less formal BDBC group**

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**

Big Data Benchmarks

Types of Big Data 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 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

Application-Level 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

App Level 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

App Level 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!

Abstracting the Big Data World

1. **Enterprise Data Warehouse + Other types of data**
   - Structured enterprise data …
   - Extend to incorpora… …ured data, e.g. from weblogs, machine logs, clickstream, custome… …ews, …
   - “Design time” schemas

BigBench

2. **Collection of heterogeneous data + Pipeli… …ssing**
   - Enterprise data processing as a pipelin… …gestion to transformation, extraction, sub… …e learning, predictive analytics
   - Data from multiple stru… …n-structured sources
   - “Runtime” sch… …ding, application-driven schemas

Deep Analytics Pipeline (DAP)

Other Benchmarks discussed at WBDB

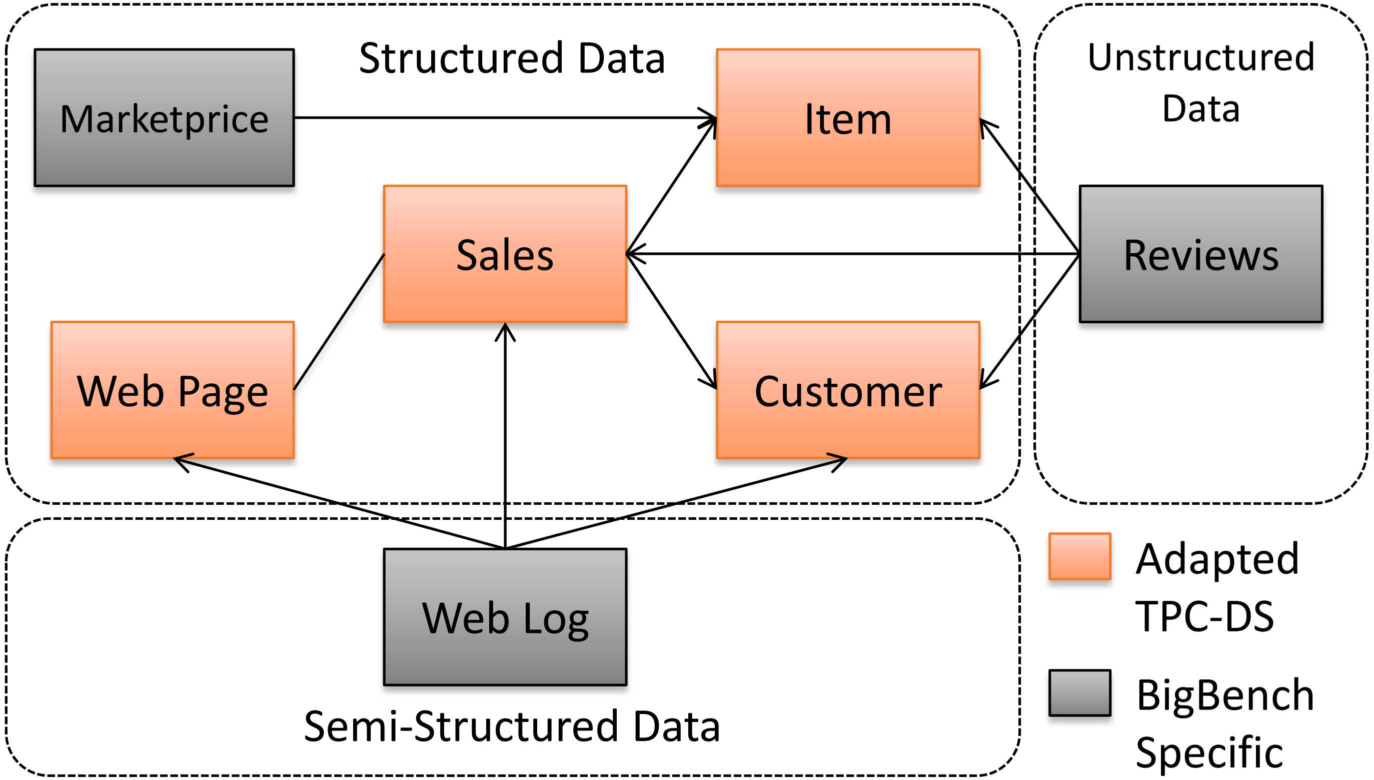
- Big Decision, Jimmy Zhao, HP
- HiBench/Hammer, Lan Yi, Intel
- BigDataBench, Jianfeng Zhan, Chinese Academy of Sciences
- CloudSuite, Onur Kocberber, EPFL

- Genre specific benchmarks
- Microbenchmarks

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
  - Full spec at WBDB proceedings 2012
  - Full kit at WBDB 2014
- **Collaboration with Industry & Academia**
  - First: Teradata, University of Toronto, Oracle, InfoSizing
  - Now: UofT, bankmark, Intel, Oracle, Microsoft, UCSD, Pivotal, Cloudera, InfoSizing, SAP, Hortonworks, Cisco, …

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**
  - Refreshes for all data
  - Different velocity for different areas
    - $V_{structured} < V_{unstructured} < V_{semistructured}$

Workload

- **Workload Queries**
  - 30 “queries”
  - Specified in English (sort of)
  - No required syntax
- **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)**

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;
```

BigBench Current Status

- All queries are available in Hive/Hadoop

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

- New data generator (continuous scaling, realistic data) available
- New metric available
- Complete driver available
- Refresh will be done soon
- Full kit at WBDB 2014
- https://github.com/intel-hadoop/Big-Bench

Big Decision, Jimmy Zhao, HP, 4th WBDB

- **Benchmark for A DSS/Data Mining solutions**
  - Everything running in the same system
  - Engine of Analytics
  - Reflecting the real business model
- **Huge data volume**
  - Data from Social
  - Data from Web log
  - Data from Comments
- **Broader Data support**
  - Semi-structured data
  - Un-structured data
- **Continuous Data Integration**
  - ETL just a normal job of the system
  - Data Integration whenever there's data
- **Big Data Analytics**

Big Decision – Big TPC-DS!

TPC-DS

- Mature and proved workload for BI
- Mix workloads
- Well defined scale factors

Semi + unstructured TPC-DS

- Additional data and dimension from new data
- Semi-structured and unstructured data
- TB to PB or even Zeta Byte support

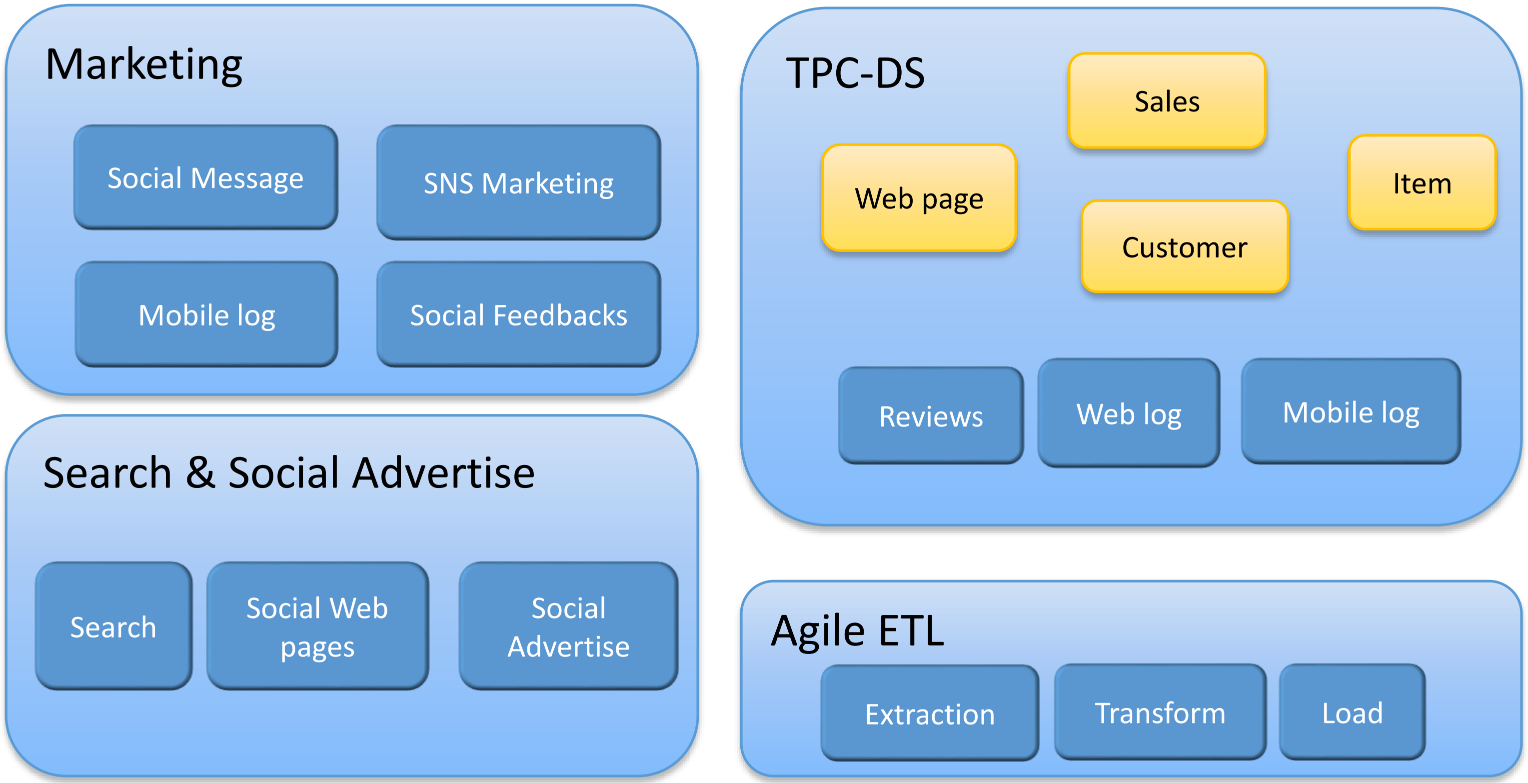
NEW TPC-DS generator – Agile ETL

- Continuously data generation and injection
- Consider as part of the workloads

New massive parallel processing technologies

- Convert queries to SQL liked queries
- Include interactive & regular Queries
- Include Machine Learning jobs

Big Decision Block Diagram



HiBench, Lan Yi, Intel, 4th WBDB

Micro Benchmarks

- Sor
- Wo
- Ter

Web Search

- xing

Ma

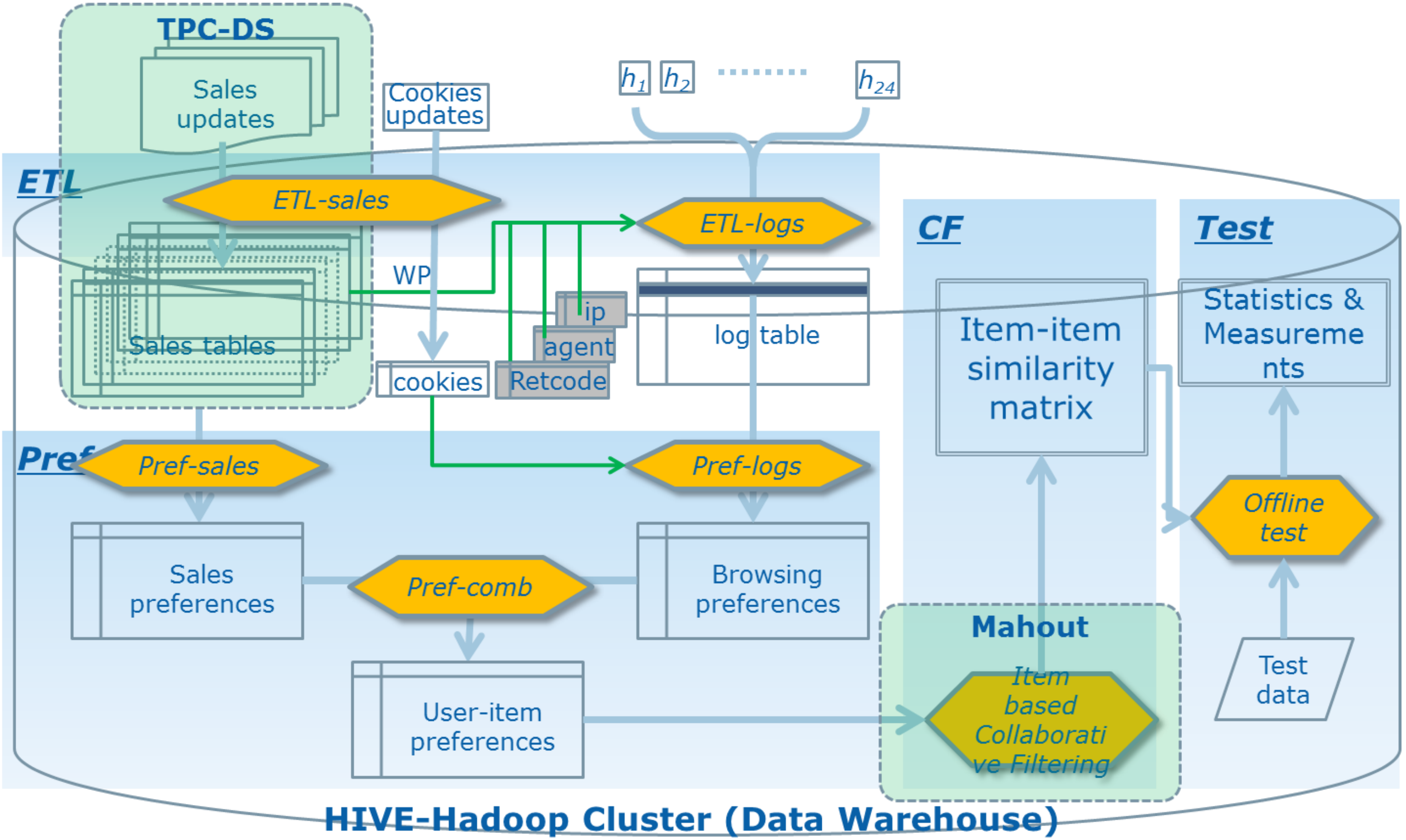
- Bayesian Classification
- K-Means Clustering

- Enhanced DFSIO

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

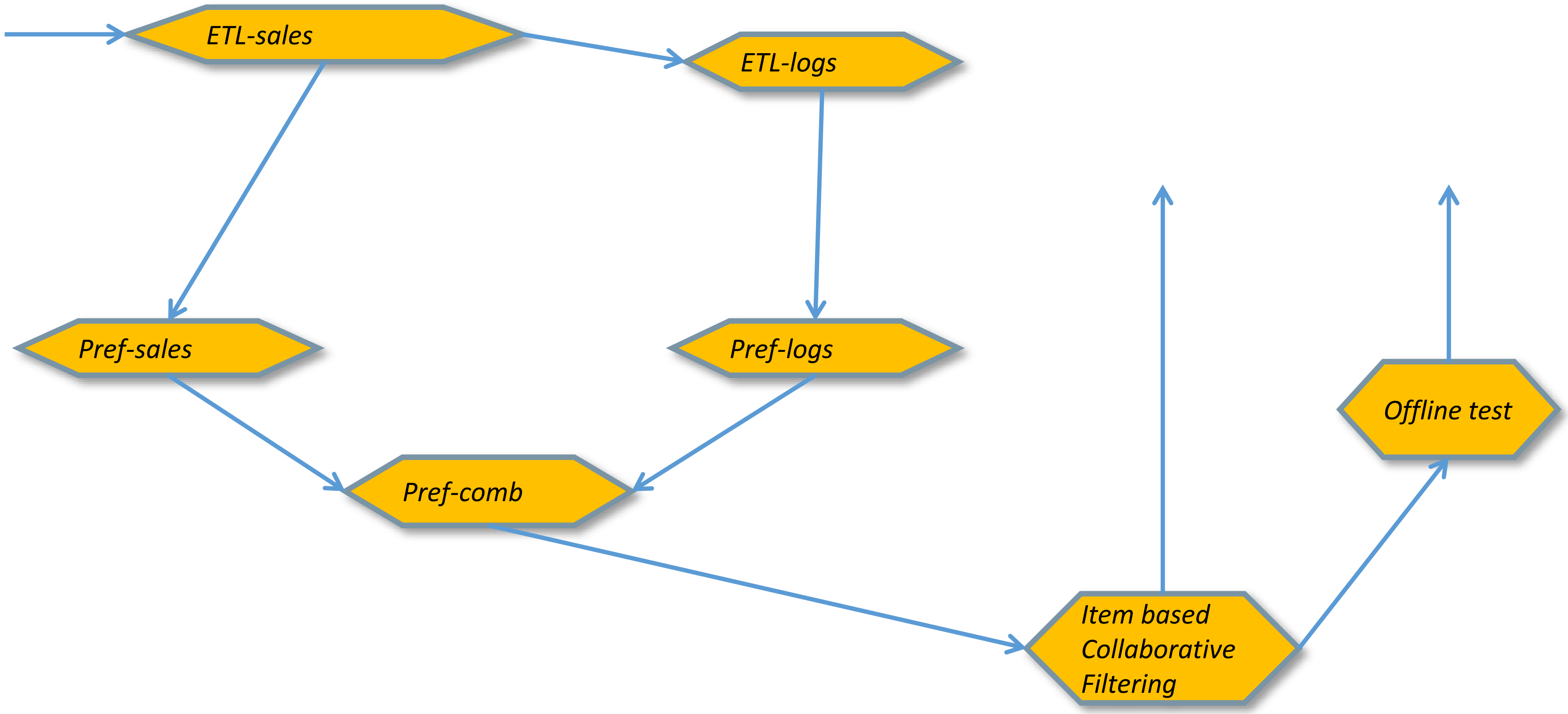
See our paper "The HiBench Suite: Characterization of the MapReduce-Based Data Analysis" in ICDE'10 workshops (WISS'10)

ETL-Recommendation (hammer)



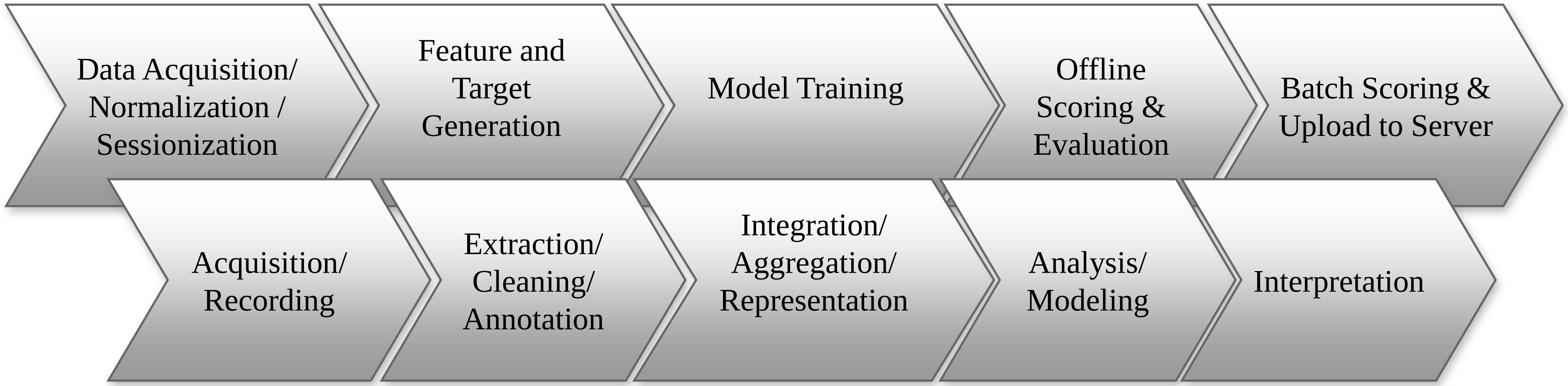
ETL-Recommendation (hammer)

- **Task Dependences**



The Deep Analytics Pipeline, Bhandarkar (1st WBDB)

- **“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



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

Steps in the Pipeline

- **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

Application Classes

- **Widely varying number of events per entity**
- **Multiple classes of applications, based on size, e.g.:**
  - Tiny (100K entities, 10 events per entity)
  - Small (1M entities, 10 events per entity)
  - Medium (10M entities, 100 events per entity)
  - Large (100M entities, 1000 events per entity)
  - Huge (1B entities, 1000 events per entity)

Proposal for Pipeline Benchmark Results

- 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
- Different modeling techniques per class
- So, need to publish performance numbers for every stage

Get involved

- **Workshop on Big Data Benchmarking (WBDB)**
  - Fifth workshop: August 6-7, Potsdam, Germany
  - clds.ucsd.edu/wbdb2014.de
  - Proceedings will be published in Springer LNCS
- **Big Data Benchmarking Community**
  - Biweekly conference calls (sort of)
  - Mailing list
  - clds.ucsd.edu/bdbc/community
- **Coming up next: BDBC@SPEC Research**
  - We will join forces with SPEC Research
- **Try BigBench:**
  - https://github.com/intel-hadoop/Big-Bench

Questions?

Thank You!

Contact:

Tilmann Rabl

tilmann.rabl@bankmark.de

tilmann.rabl@utoronto.ca

MIDDLEWARE SYSTEMS RESEARCH GROUP
MSRG.ORG

bankmark

VELUT ARBOR fiVO