Metrics & Message
Lessons from the TPC

• Marketing wants to control the message
  • Bragging rights for one simple metric (print media centric)
  • Rules about its fair use
  • Complex environment with vendor supplied code
  • Expensive and time consuming for all parties

• Users (I think) actually want information to help make a real decision
  • Implies a far richer and more complex analysis covering many aspects of a run
  • To date has often been for simpler hw benchmarks (cpu, disk, net) with few if any fair use rules, and a fixed execution environment
  • Sustained by user demand

• So which type of benchmark do we want for big data?
Metrics & Message
Is a user centric, analysis centric benchmark feasible?

• If the tool suite served on the benchmark site becomes the de facto std for analysis of public big data results, marketing will come on board.
  • Give up control but attract eyeballs
  • EULAs make vendor support mandatory

• A data set of all timing data and setup (e.g. TPC ES) across an archive of runs should provide for ample data visualization opportunities (think Rosling’s Joy of Stats).
  • Major chicken and egg problem

• Producing such a data set implies a high degree of automation and thus cost savings for sponsors.
Metrics & Message
Configuration as a metric

• One item impacting cost for both sponsors and users is configuration complexity
• Ideally this would also be a metric
  • Whether defaulting everything out of the box Vs. making changes to cpu bios, storage bios, OS, FS, & dbms would be very useful information
• Having such as a comparable data item might help push the industry toward more automation.
  • Alas this seems extremely difficult
  • At least if the entire test is supplied and allows no vendor specific code this might cut down on the clutter in disclosure documents.