HPTS 2009
Map/Reduce and Parallel DBs

HPTS 2009
Map/Reduce and Parallel DBs

What's the big deal?

Participants

- Eugene Shekita (IBM)
- Michael Olson (Cloudera)
- Mike Stonebraker (Company du jour)
- Daniel Abadi (Yale)

More information about your participants.

- Two panelists sent me (at least) three different versions of their slides.
- One panelist sent me at least two copies of slides.
- One panelist sent me slides on Monday hoping I hadn’t yet gotten on the plane to come to HPTS.

A Brief History

A Brief History

- 17 mail messages to decide on content of submission

A Brief History

- 17 mail messages to decide on content of submission
- “We’re all going to say the same thing.”

A Brief History

- 17 mail messages to decide on content of submission
- “We’re all going to say the same thing.”
- 12 mail messages to decide on format.

A Brief History

- 17 mail messages to decide on content of submission
- “We’re all going to say the same thing.”
- 12 mail messages to decide on format.
- I told them that they must have data to support claims.

A Brief History

- 17 mail messages to decide on content of submission
- “We’re all going to say the same thing.”
- 12 mail messages to decide on format.
- I told them that they must have data to support claims.
- 1 panelist has data; 1 panelist has an anecdote.

Rules of Engagement

- Each panelist gets 10 minutes.
- I own the questions for the next 10-20 minutes.
- Audience participation.

- NOTE: This was their idea.

Question #1

- Eugene -- so customers say they like Map/Reduce; how has that changed what you're doing in IBM?

Question #2

- Mike O: How much of the $15 Billion RDBMS market do you think MapReduce can take? Why?

Question 3:

- Mike S: You say customers don't ask for more scaling in the DB. Is it really because they don't need it or because we've trained them to design around it.

Question #4

- Daniel: You say that databases don’t scale, but WHY don’t they scale?

Question #5

- Let's assume that MapReduce is here to stay, two questions:
  1. Can they leverage the RDBMS research in useful ways?
- 2. Will they?

Question #6

- Let's say that MapReduce is here to stay, are the RDBMS guys going to address the restartability, reliability, issues?