Predictability in Physical Design

- Overestimates = User frustration
- Underestimates = Missing opportunities

Methodology

- Compare different physical database designers in terms of their predictability.

Predictability = difference between estimated and actual improvement

Algorithm:

1. Run
2. Tune
3. Apply
4. Re-run
5. Calculate...

Questions:

a) Does space budget matter?
b) What about workload size?
c) Introducing updates?
d) What about database size?
e) Do statistics matter?

Impact of Space Budget

- Improvement typically higher than estimated
- Applied designs degrade performance

Impact of Workload Size

- Improvement lower than estimated
- Wrong cardinality estimates hurt performance

What About Updates?

- Cannot balance improvement and maintenance

Summary

Proposed designs can be unpredictable

- System A: $R_{EE}$ below 46%, performance hurt for 18% only with updates
- System B: $R_{EE}$ up to 92%, performance hurt up to 776%
- System C: $R_{EE}$ up to 87%, performance hurt up to 219%