SPECmarks Considered Harmful

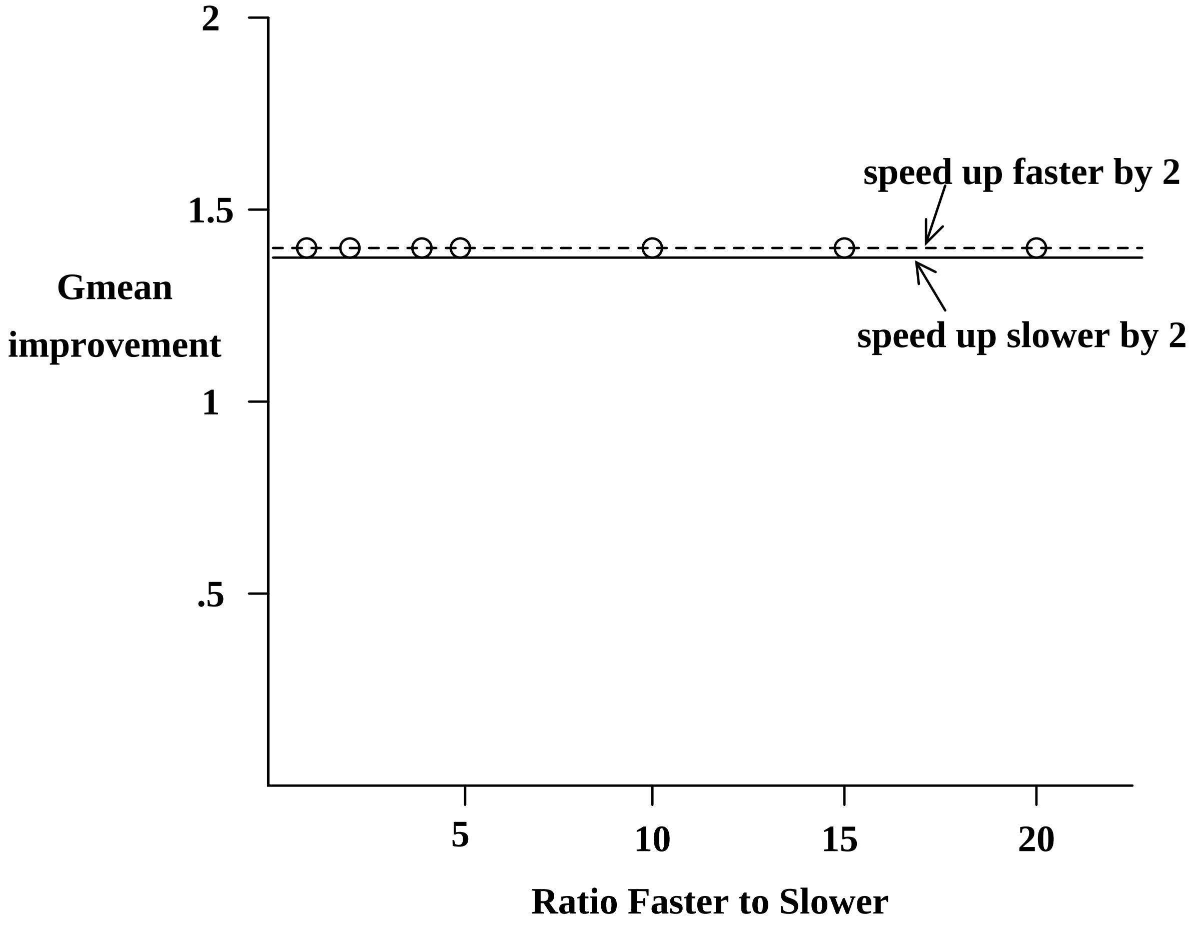
for all the usual reasons

plus

Geometric mean

Consider: Two benchmarks, one faster than the other

- If performance of the slower is doubled,
  what happens to the Gmean?
- If the faster is doubled, what happens?



=> doubling performance of faster gives
as much "reward" as doubling the slower.

Performance Limitations:
Latency and Bandwidth

Latency: limits highly serial, inherently scalar code

Bandwidth: limits highly parallel code

To improve *latency:*
faster technology and cleverness

To improve *bandwidth:*
faster technology and throw transistors
at the problem (i.e. less cleverness)

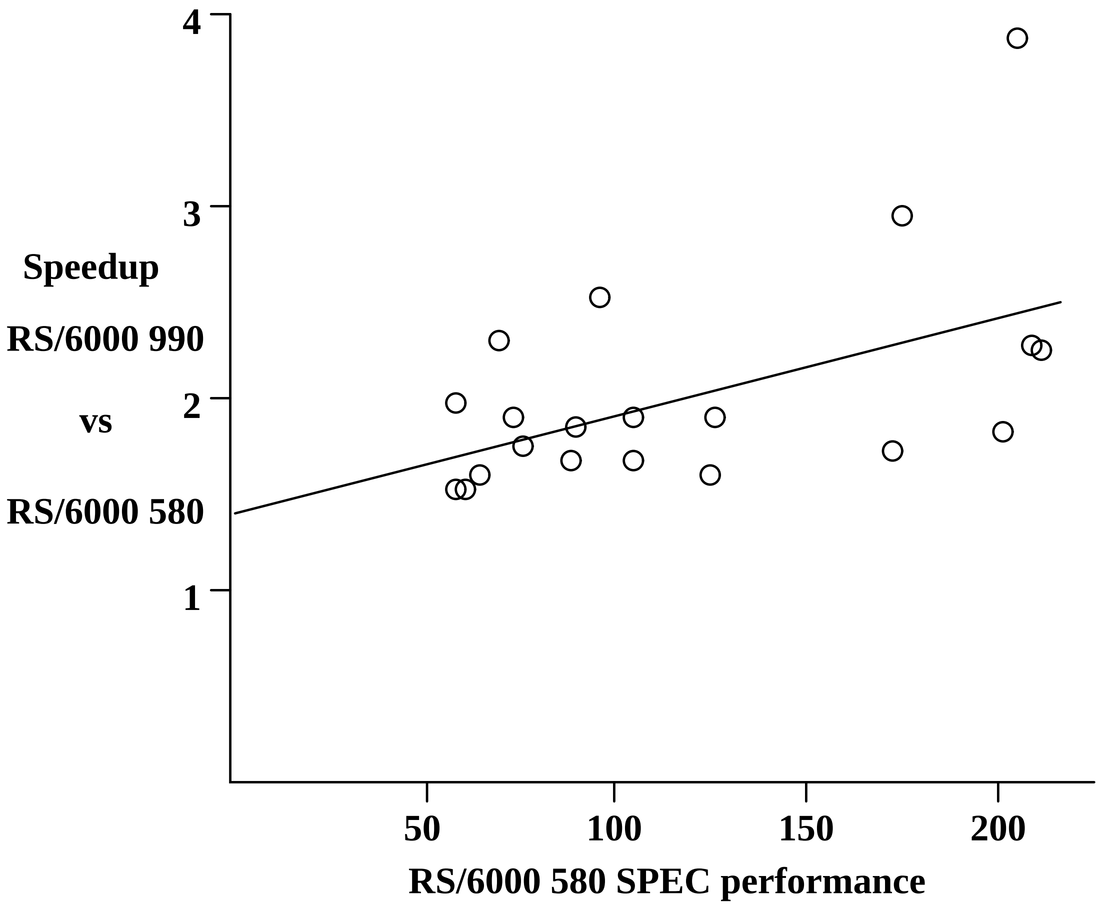
Now,
- with SPECmarks being given the level
  of importance they are,
- with technology being as it is,
- Gmean being as it is,
- and human nature being as it is...

After a certain level of maturity is reached,

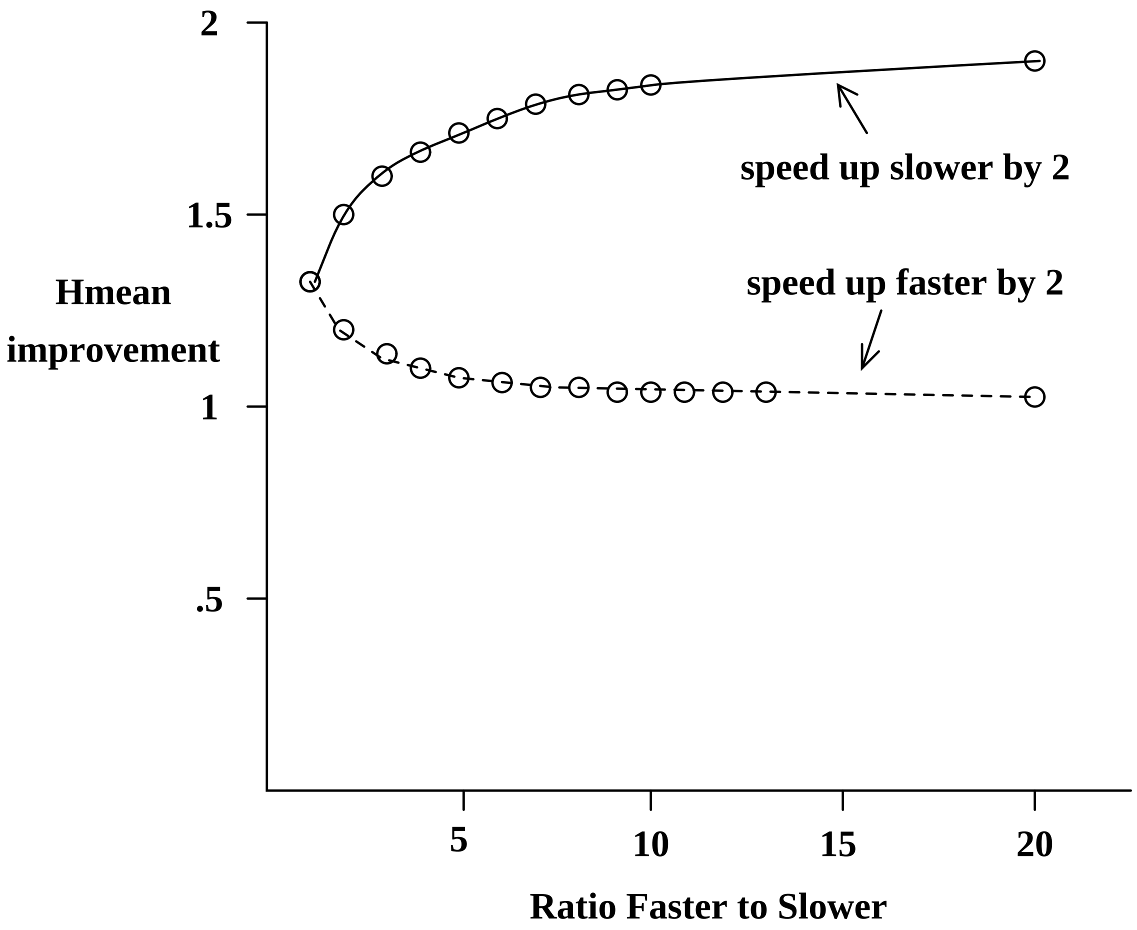
the fast codes will be sped up more than slow codes.

Example

IBM RS/6000 580 (POWER1) and RS/6000 990 (POWER2)



Suggestion: use Harmonic Mean



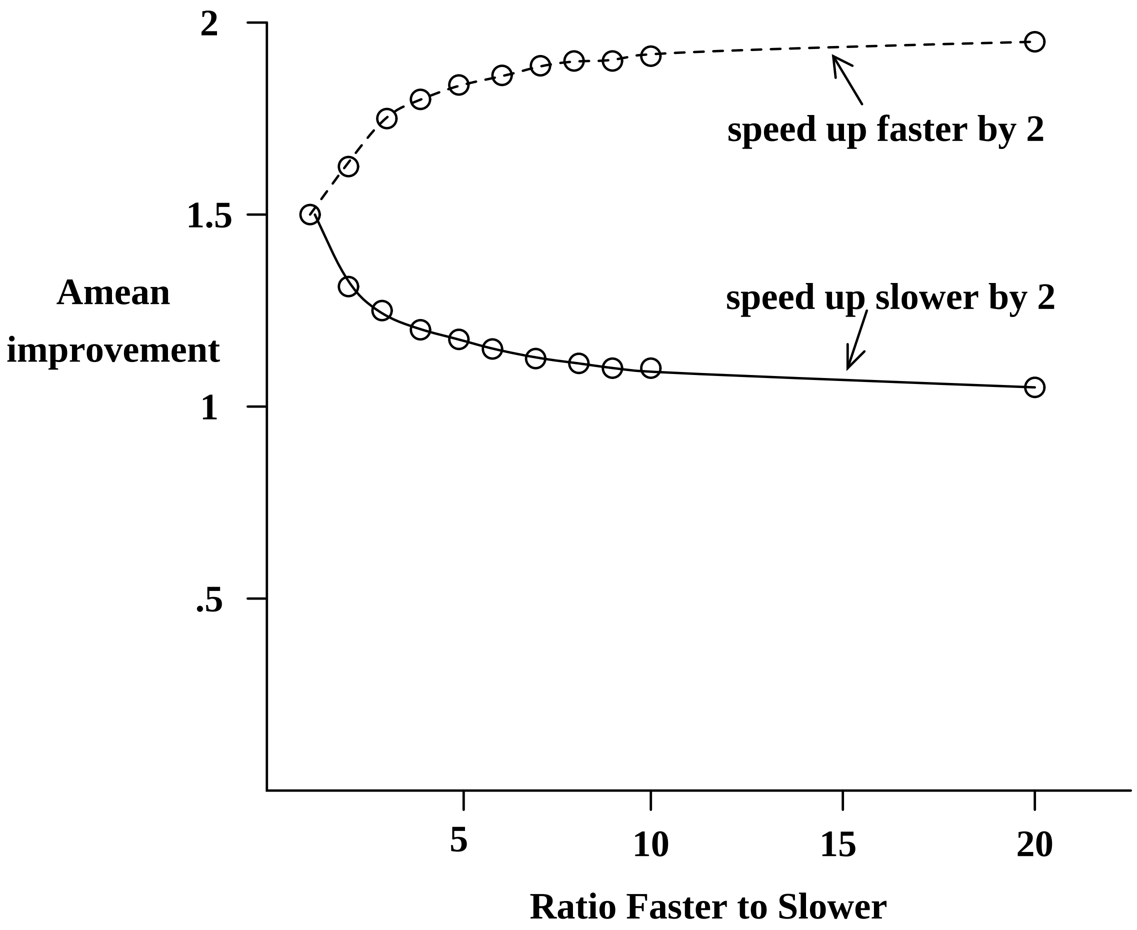
- speeding up slower benchmarks gives greater reward
- this will tend to drive speedups on all benchmarks, not just the easy (fast) ones
- observation: Matrix 300 "problem" was a problem with gmean, *not* the benchmark.

CRAY RESEARCH, INC.

Suggestion: use Harmonic Mean

- Speeding up slower benchmarks gives greater reward
- this will tend to drive speedups on all benchmarks, not just the easy (fast) ones
- harmonic mean reflects total execution time when benchmarks all do an equivalent amount of work.
- observation: Matrix 300 "problem" was a problem with gmean, *not* the benchmark.

FYI: Arithmetic Mean



CRAY RESEARCH, INC.

Geometric Mean

- Speeding up any benchmark by **S** has same effect on Gmean

- *Latency:* limits highly serial, inherently scalar code

- *Bandwidth:* limits highly parallel code

- To improve *latency:*
  faster technology and cleverness

- To improve *bandwidth:*
  faster technology and throw transistors
  at the problem (i.e. less cleverness)

Now,
- with SPECmarks being given the level
  of importance they are,
- with technology being as it is,
- Gmean being as it is,
- and human nature being as it is...

After a certain level of maturity is reached,

the fast codes will be sped up more than slow codes.

CRAY RESEARCH, INC.