# Composite Metrics for System Throughput in HPC

John D. McCalpin, Ph.D.
IBM Corporation
Austin, TX

SuperComputing 2003
Phoenix, AZ
November 20, 2003

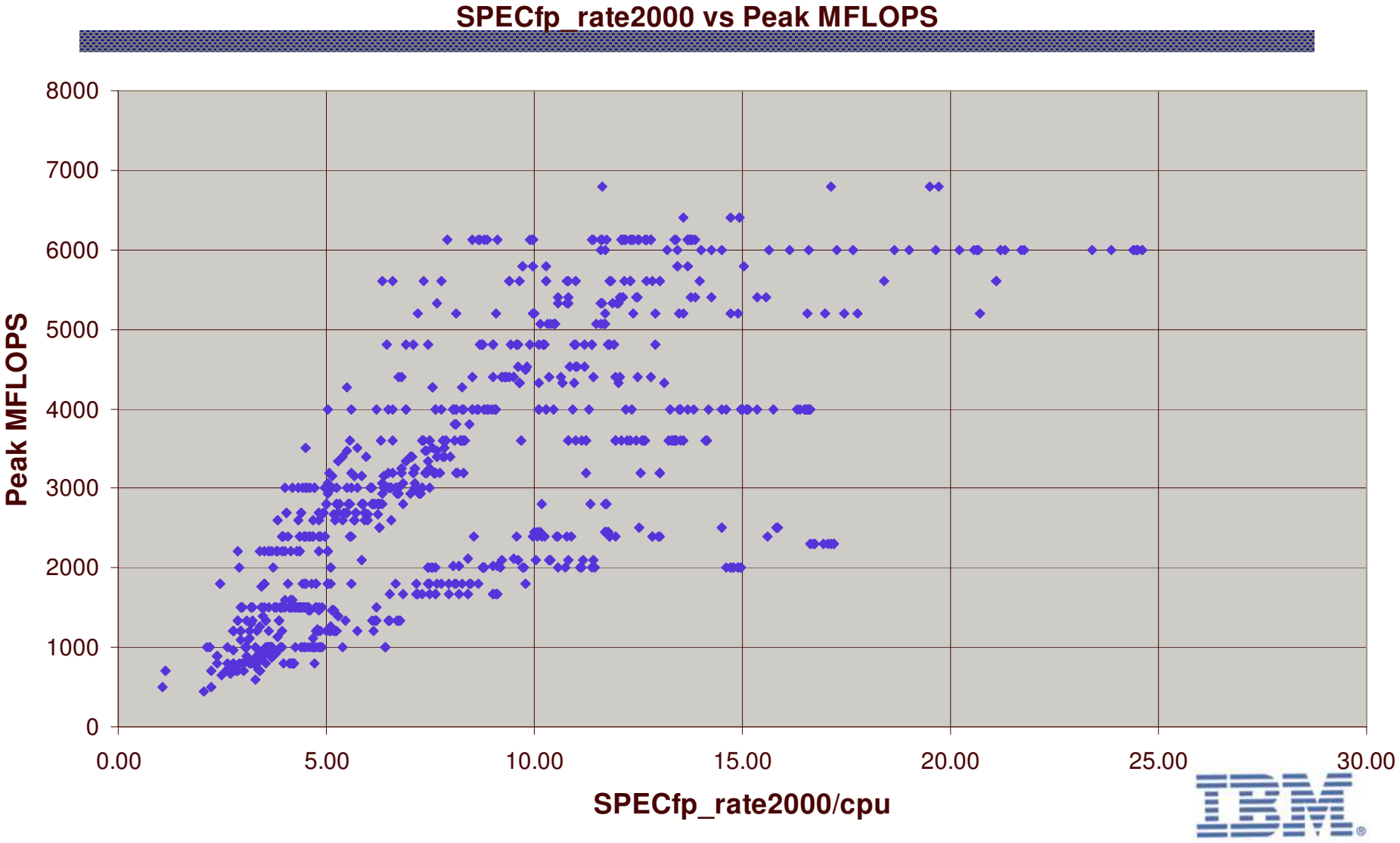
# Overview

- The HPC Challenge Benchmark was announced last night at the TOP500 BOF
- The HPC Challenge Benchmark consists of
  - **LINPACK** (HPL)
  - **STREAM**
  - **PTRANS** (transposing the array used by HPL)
  - **GUPS**
  - and some low-level MPI latency & BW measurements
- No single figure of merit is defined

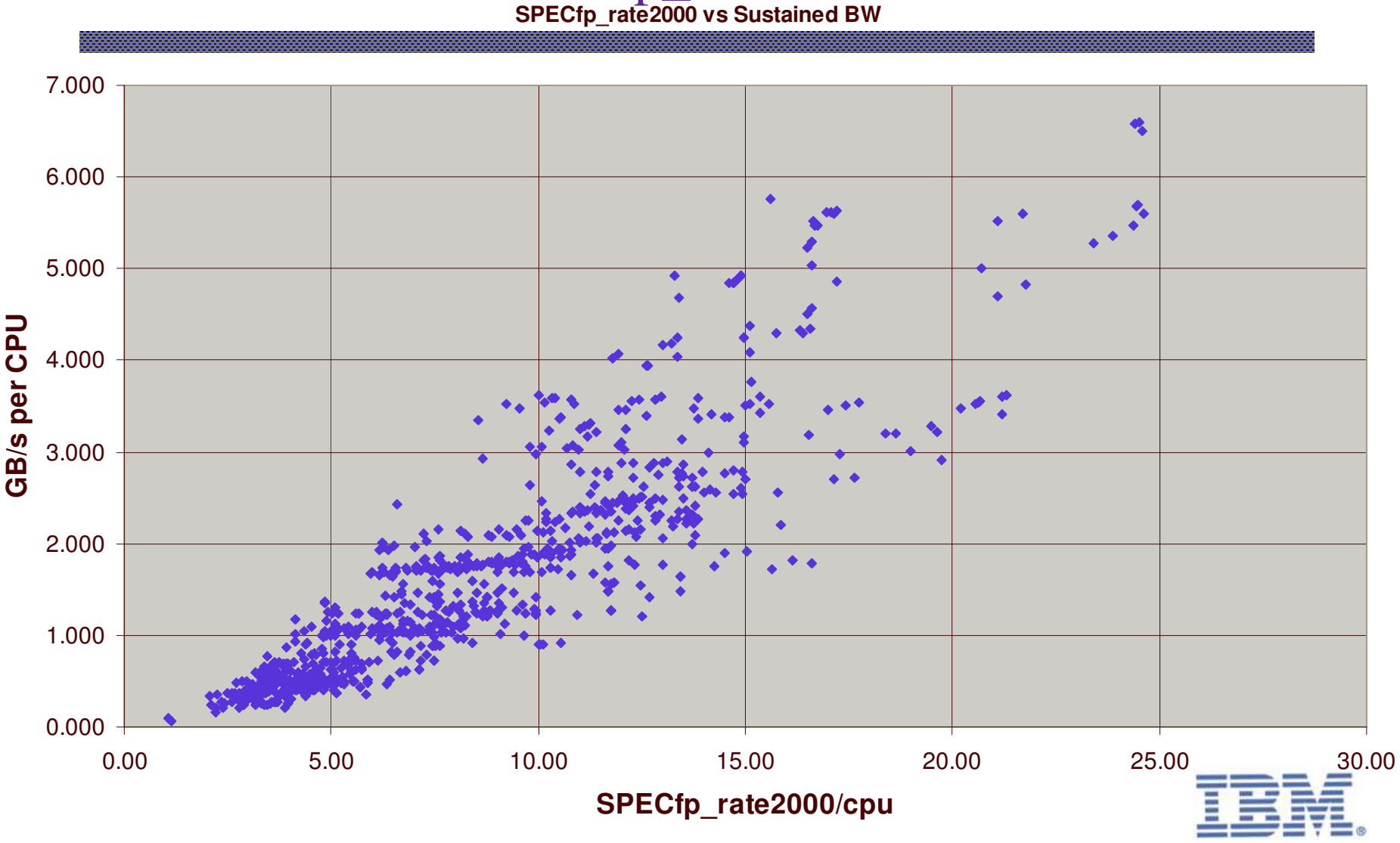
# The Big Question

- Q: How should one think about composite figures of merit based on such a collection of low-level measurements?
- A: Composite Figures of Merit must be based on "time" rather than "rate"
  - i.e., weighted harmonic means of rates
- Why?
  - Combining "rates" in any other way fails to have a "Law of Diminishing Returns"

# Does Peak GFLOPS predict SPECfp_rate2000?



# Does Sustained Memory Bandwidth predict SPECfp_rate2000?



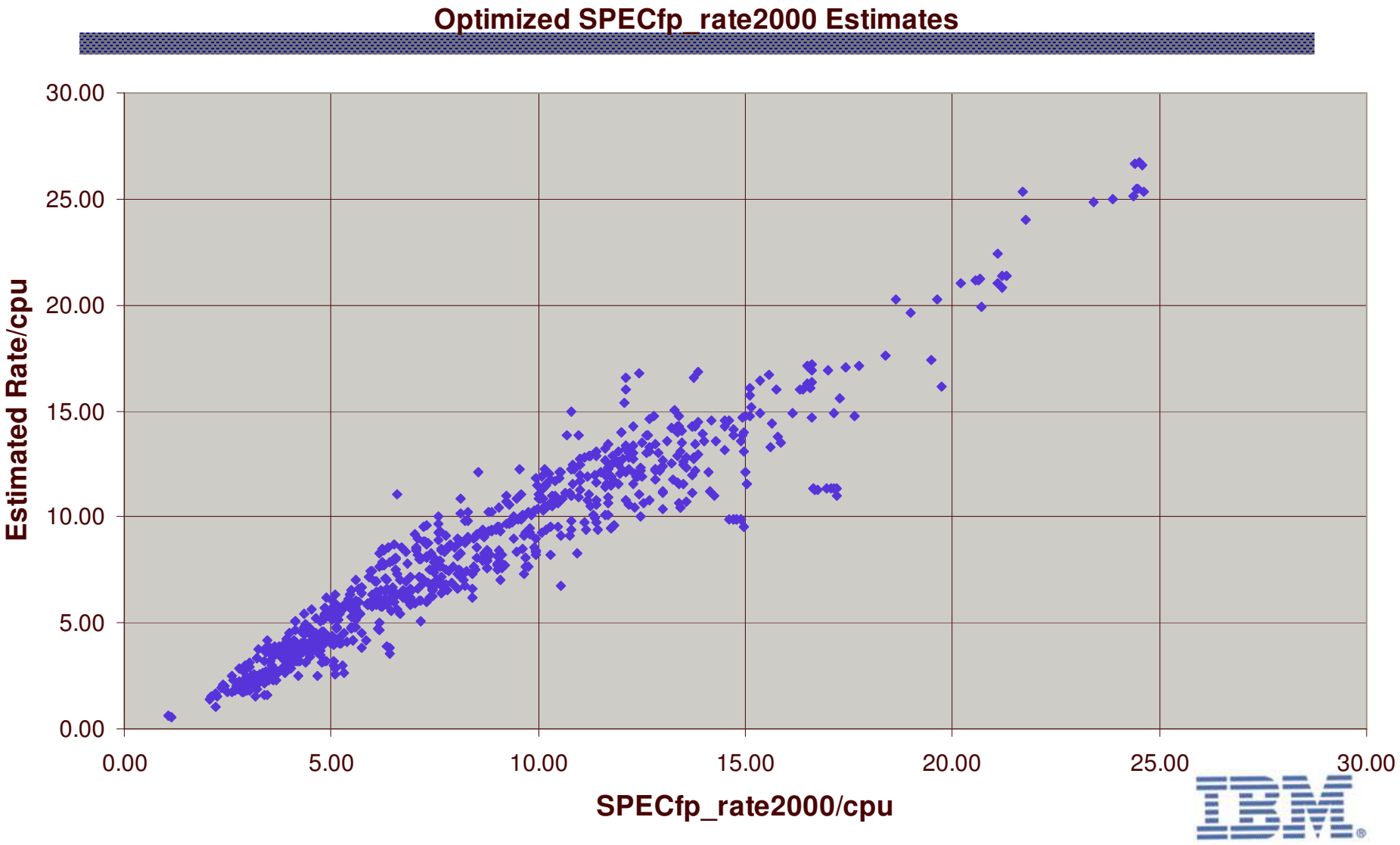
# A Simple Composite Model

- Assume the time to solution is composed of a compute time proportional to peak GFLOPS plus a memory transfer time proportional to sustained memory bandwidth
- Assume “x Bytes/FLOP” to get:

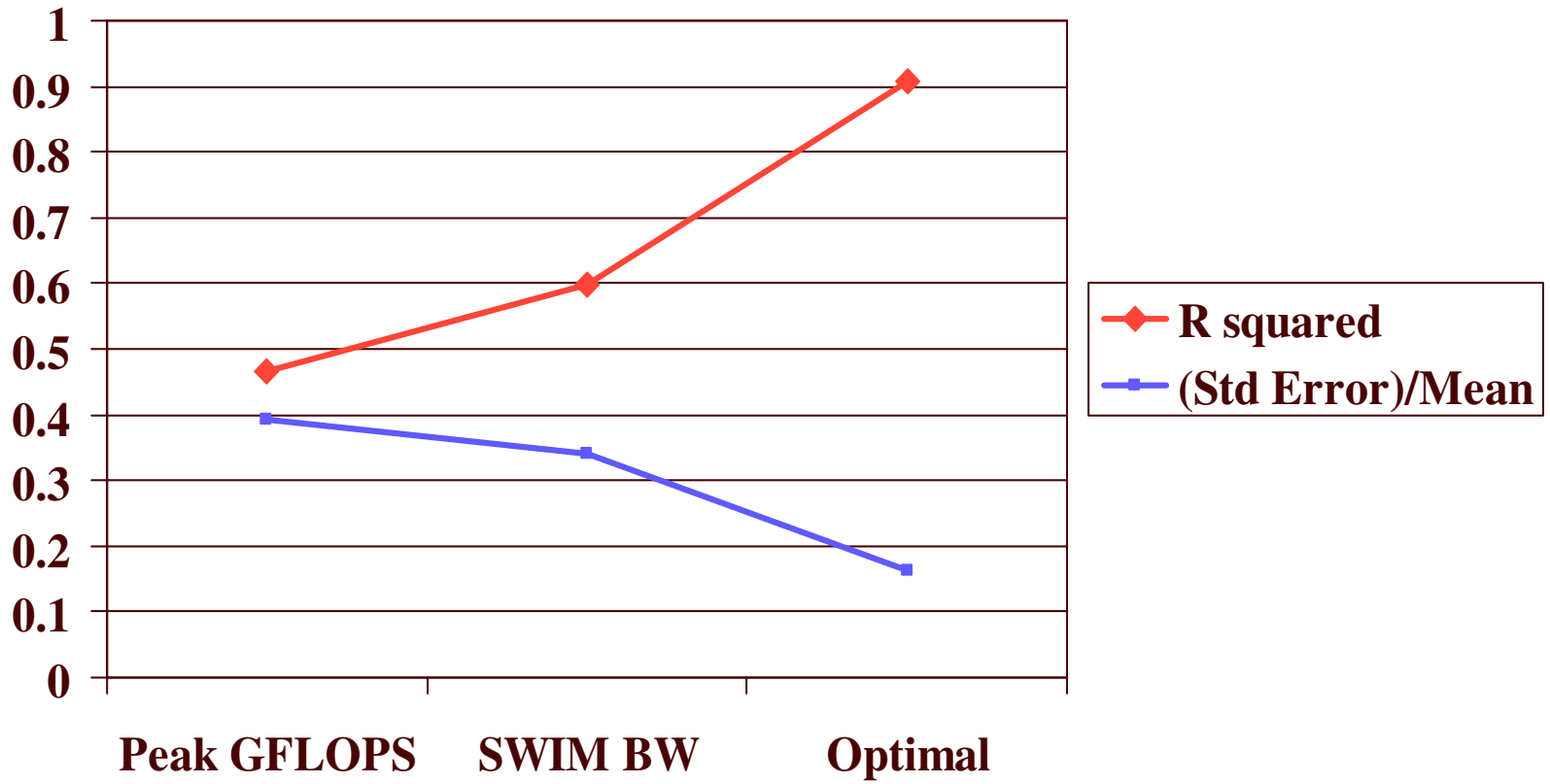
$$\text{"Balanced GFLOPS"} \equiv \frac{1\ \text{"Effective FP op"}}{\left(\frac{1\ \text{FP op}}{\text{Peak GFLOPS}}\right) + \left(\frac{x\ \text{Bytes}}{\text{Sustained GB/s}}\right)}$$

IBM

# Does this Revised Metric predict SPECfp_rate2000?

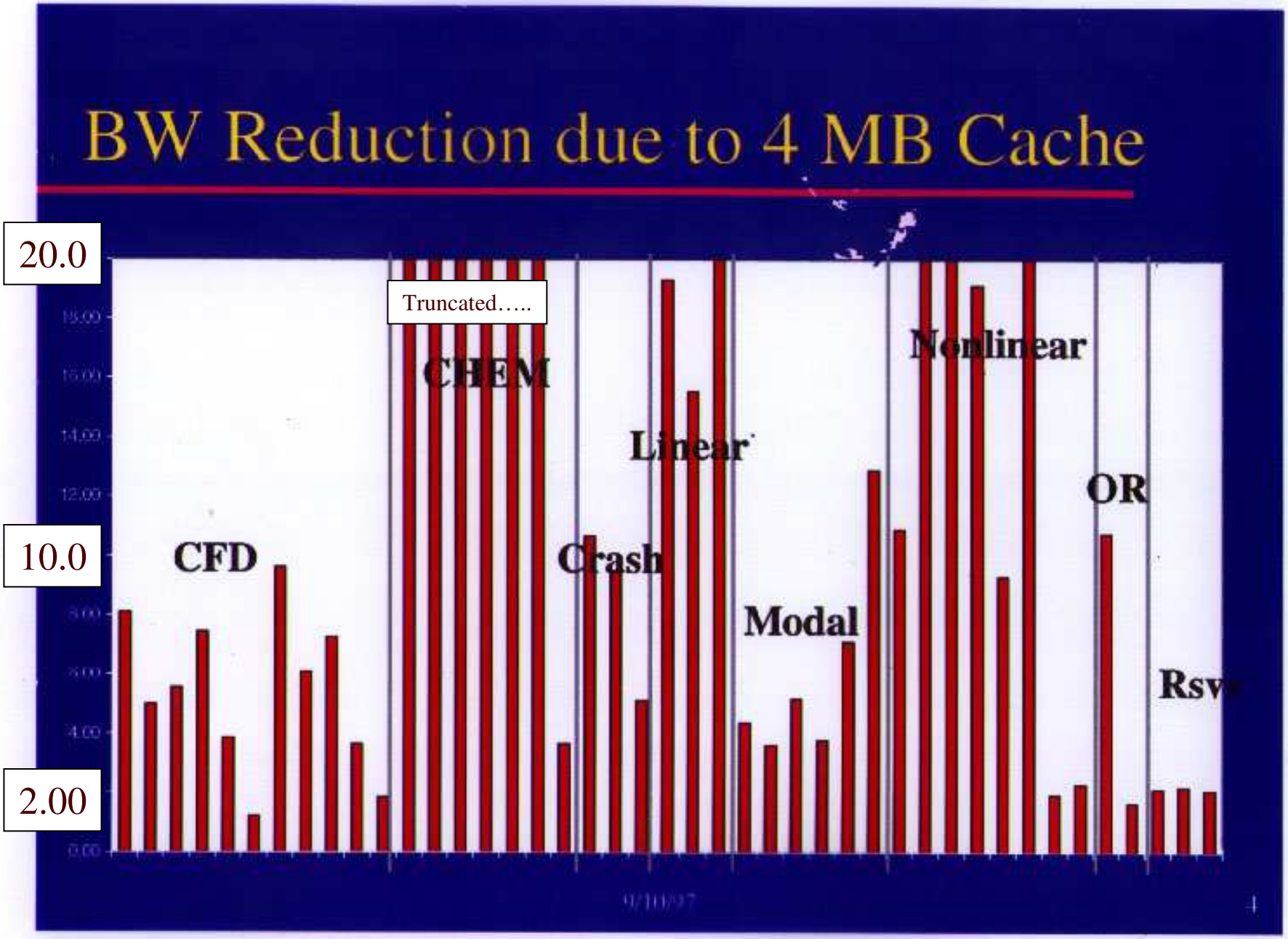


# Statistical Metrics



# What about other applications?

- Effectiveness of caches varies by application area
- Requirements for interconnect performance vary by application area
  - Some apps are short-message dominated
  - Some apps are long-message dominated
- Composite models can be tuned to specific application areas – if app properties known


BW Reduction due to 4 MB Cache
20.0
10.0
2.00
Truncated…..
CHEM
Linear
Nonlinear
OR
CFD
Crash
Modal
Rsv
9/10/97
4

# An Example Model tuned for CFD

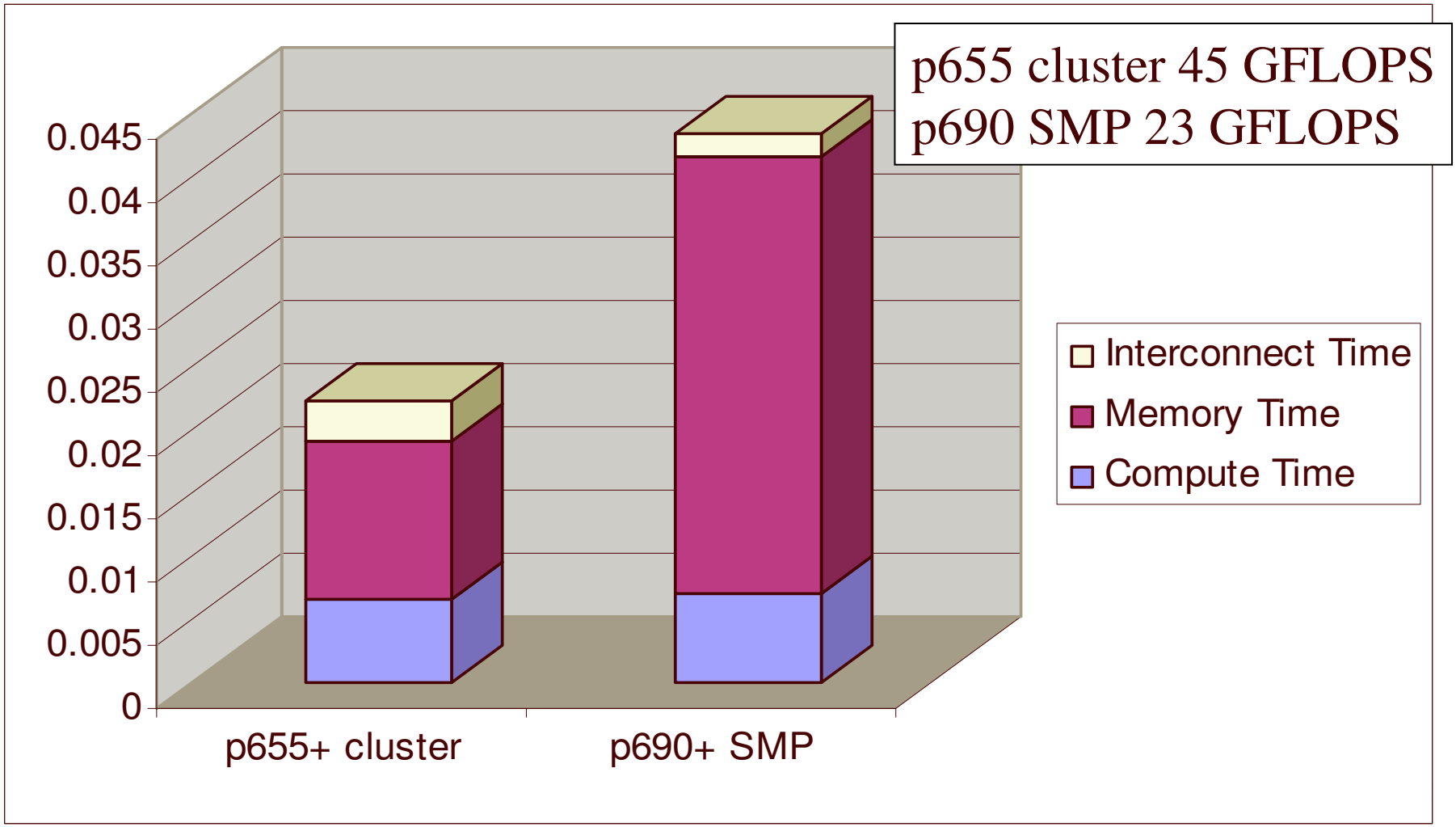
- Analyze applications and pick reasonable values:

$$\text{"Balanced GFLOPS"} \equiv \frac{1\ \text{"Effective FP op"}}{\left(\frac{1\,\text{FP op}}{\text{LINPACK GFLOPS}}\right)+\left(\frac{2\,\text{Bytes}}{\text{STREAM GB/s}}\right)+\left(\frac{0.1\,\text{Bytes}}{\text{Network GB/s}}\right)}$$

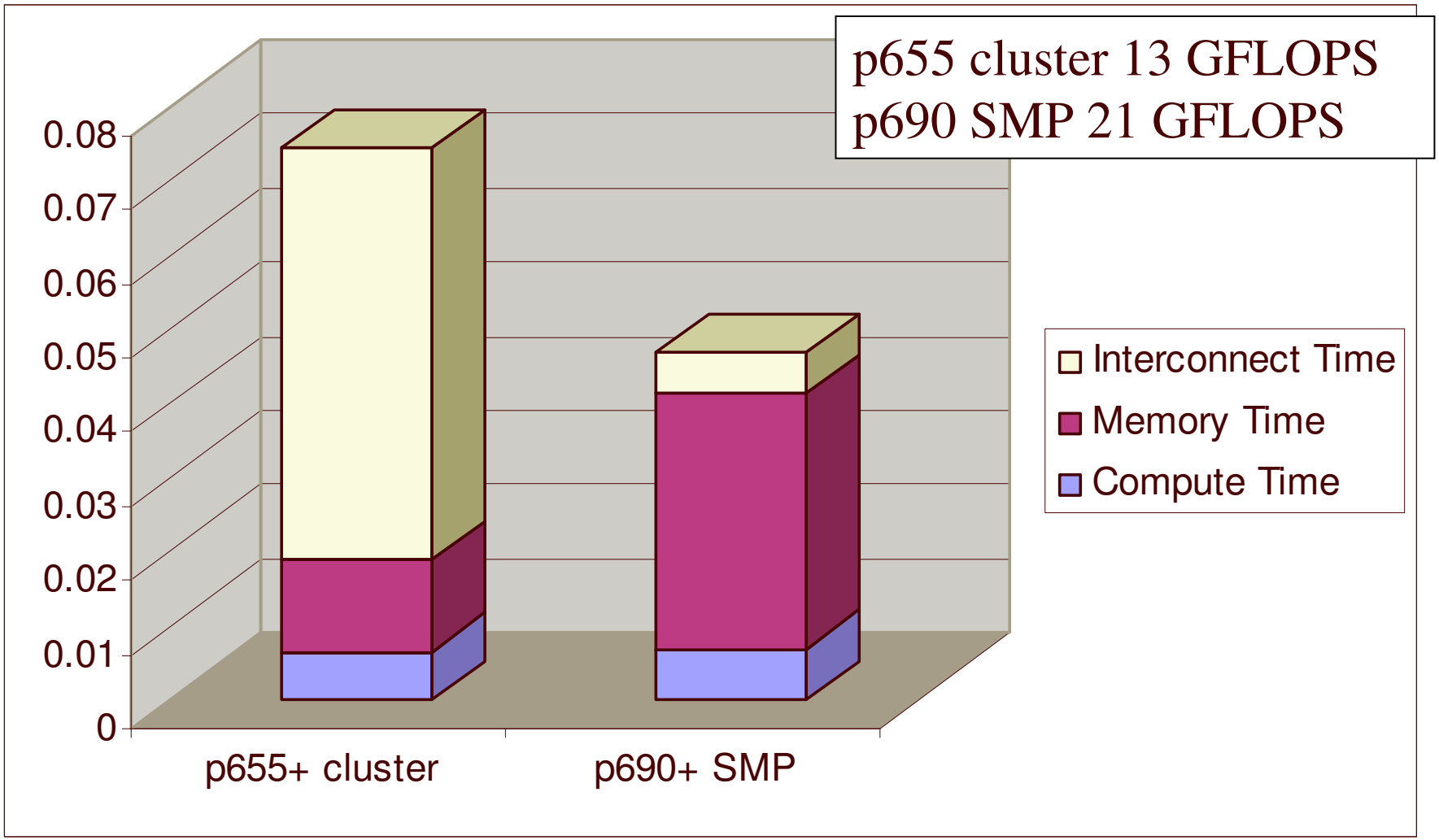
- Two cases tested:
  - Assume long messages (network BW tracks PTRANS)
  - Assume short messages (network BW tracks GUPS)

- The relative time contributions will quickly identify applications that are poorly balanced for the target workload

# Comparing p655 cluster vs p690 SMP Assumes long messages



# Comparing p655 cluster vs p690 SMP
# Assumes short messages



# Summary

- The composite methodology is
  - Simple to understand
  - Based on the components of the HPC Challenge Benchmark
  - Based on a mathematically correct model of performance

- Much work remains on documenting the work requirements of various application areas in relation to the component microbenchmarks