

---

# Performance Metrics for HPC

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



# How Close is the Holy Grail?

---

- It all depends on what you want this Holy Grail to do for you

# How Close is the Holy Grail?

---

- It all depends on what you want this Holy Grail to do for you
- Some environments have performance projection goals that are easy to obtain
  - Sites that want to run hardware that currently exists
  - Sites with simple workloads
  - Sites with workloads that are computationally dense and embarrassingly parallel



# How Close is the Holy Grail?

---

- It all depends on what you want this Holy Grail to do for you
- Some environments have performance projection goals that are easy to obtain
- **Some environments have performance projection goals that are hard to obtain, but feasible**
  - Sites that want to run hardware that is an evolutionary modification of existing hardware
  - Sites with more complex workloads that remain relatively stationary



# How Close is the Holy Grail?

---

- It all depends on what you want this Holy Grail to do for you
- Some environments have performance projection goals that are easy to obtain
- Some environments have performance projection goals that are hard to obtain, but remain tractable
- **Some environments have performance projection goals that remain intractable**
  - Sites wanting hardware that is not yet designed
  - Sites with unpredictable or rapidly varying workloads



# How about “good enough”?

---

- If you are willing to put in the effort to understand a lot about the performance characteristics of your applications

# How about “good enough”?

---

- If you are willing to put in the effort to understand a lot about the performance characteristics of your applications, **and**
- If you are willing to live with a fairly large degree of “fuzz” in the results (maybe ?20%)

# How about “good enough”?

---

- If you are willing to put in the effort to understand a lot about the performance characteristics of your applications, **and**
- If you are willing to live with a fairly large degree of “fuzz” in the results (maybe ?20%), **and**
- If you are willing to live to  $O(1)$  errors on occasion



## How about “good enough”?

---

- If you are willing to put in the effort to understand a lot about the performance characteristics of your applications, **and**
- If you are willing to live with a fairly large degree of “fuzz” in the results (maybe ?20%), **and**
- If you are willing to live to  $O(1)$  errors on occasion, **then**
- **Simple composite metrics will probably work OK**



# What needs to be done?

---

- Single node “balance” and its influence on performance is known fuzzily
- Interconnect “balance” and its influence on performance is known poorly, except for particular applications
- Predictions are hard

# What needs to be done?

---

- Single node “balance” and its influence on performance is known fuzzily
- Interconnect “balance” and its influence on performance is known poorly, except for particular applications
- Predictions are hard – especially when they are about the future

# What needs to be done?

---

- Future problems
  - Unexpected bottlenecks will show up at extreme scale
  - Unexpected bottlenecks will show up for future processors, future memory subsystems, future programming languages
- So we need to stay flexible
  - Computing and memory accesses are likely to be important for the future
  - The more fuzz you can tolerate, the less detail you need
  - Simple metrics might be more politically useful than complex ones

