



Sandia
National
Laboratories

Metrics and Benchmarks for Quantum Computers

Travis Humble
Catherine McGeoch
Kevin Young



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

What are metrics and benchmarks?

- **Benchmark**

- (n) something that serves as a standard by which others may be measured or judged
- (n) a point of reference from which measurements may be made
- (n) a standardized problem or test that serves as a basis for evaluation or comparison (as of computer system performance)
- (v) *business* - to study (something, such as a competitor's product or business practices) in order to improve the performance of one's own company

- **Metrics**

- (n) *business* - a set of figures or statistics that measure results
- (n) synonym for benchmark

"benchmark." *Merriam-Webster.com*. 2019. <https://www.merriam-webster.com> (15 Oct 2019).
"metric." *Oxford-English Dictionary*. 2019. <https://www.lexico.com> (15 Oct 2019).

What are metrics and benchmarks?

- **Benchmark**

- (n) a standardized problem or test that serves as a basis for evaluation or comparison (as of computer system performance)

- **Metrics**

- (n) a set of figures or statistics that measure results

"benchmark." *Merriam-Webster.com*. 2019. <https://www.merriam-webster.com> (15 Oct 2019).
"metric." *Merriam-Webster.com*. 2019. <https://www.merriam-webster.com> (15 Oct 2019).

What are metrics and benchmarks?

- **Benchmark**

- (n) a standardized problem or test that serves as a basis for evaluation or comparison (as of computer system performance)
- *(n) ***a standardized test suite and method for postprocessing measurement results.***
 - For gate model quantum computers, the test suite is a family of (possibly random) quantum circuits.

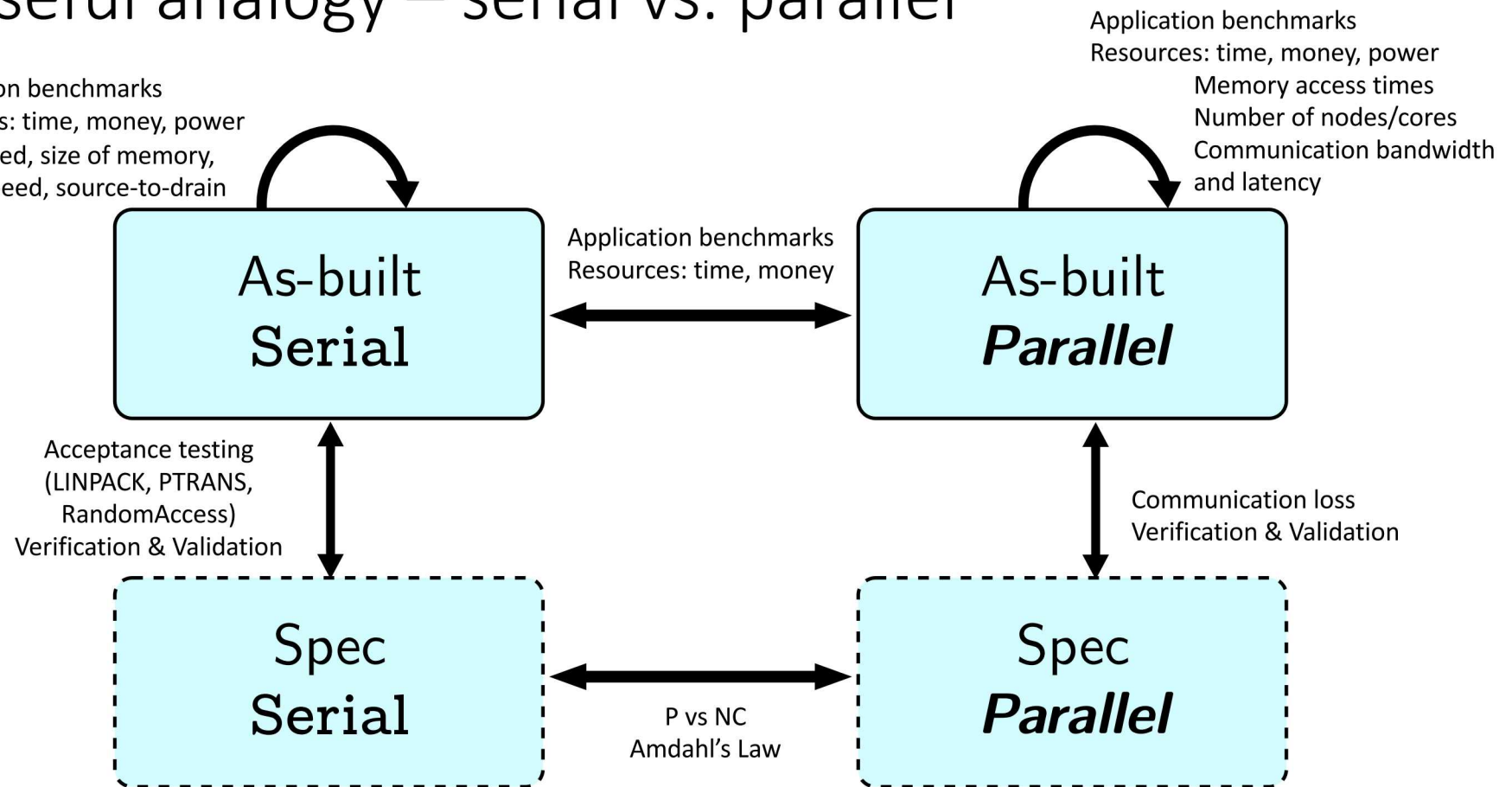
- **Metrics**

- (n) a set of figures or statistics that measure results
- *(n) **measurable quantities that may have meaning independent of a specific benchmark**
 - Eg., gate error rates, measurement fidelities, leakage current

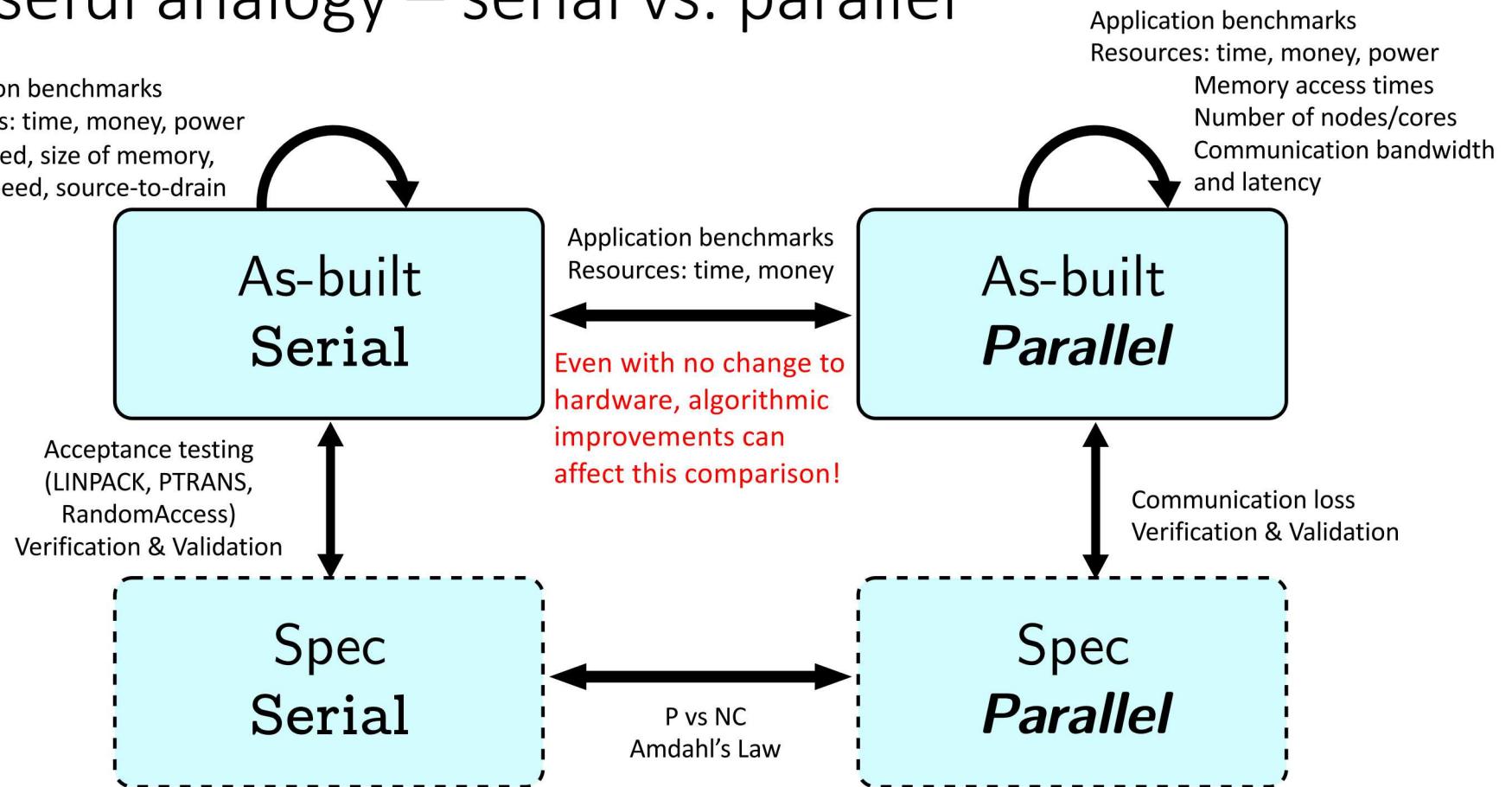
“benchmark.” *Merriam-Webster.com*. 2019. <https://www.merriam-webster.com> (15 Oct 2019).

“metric.” *Merriam-Webster.com*. 2019. <https://www.merriam-webster.com> (15 Oct 2019).

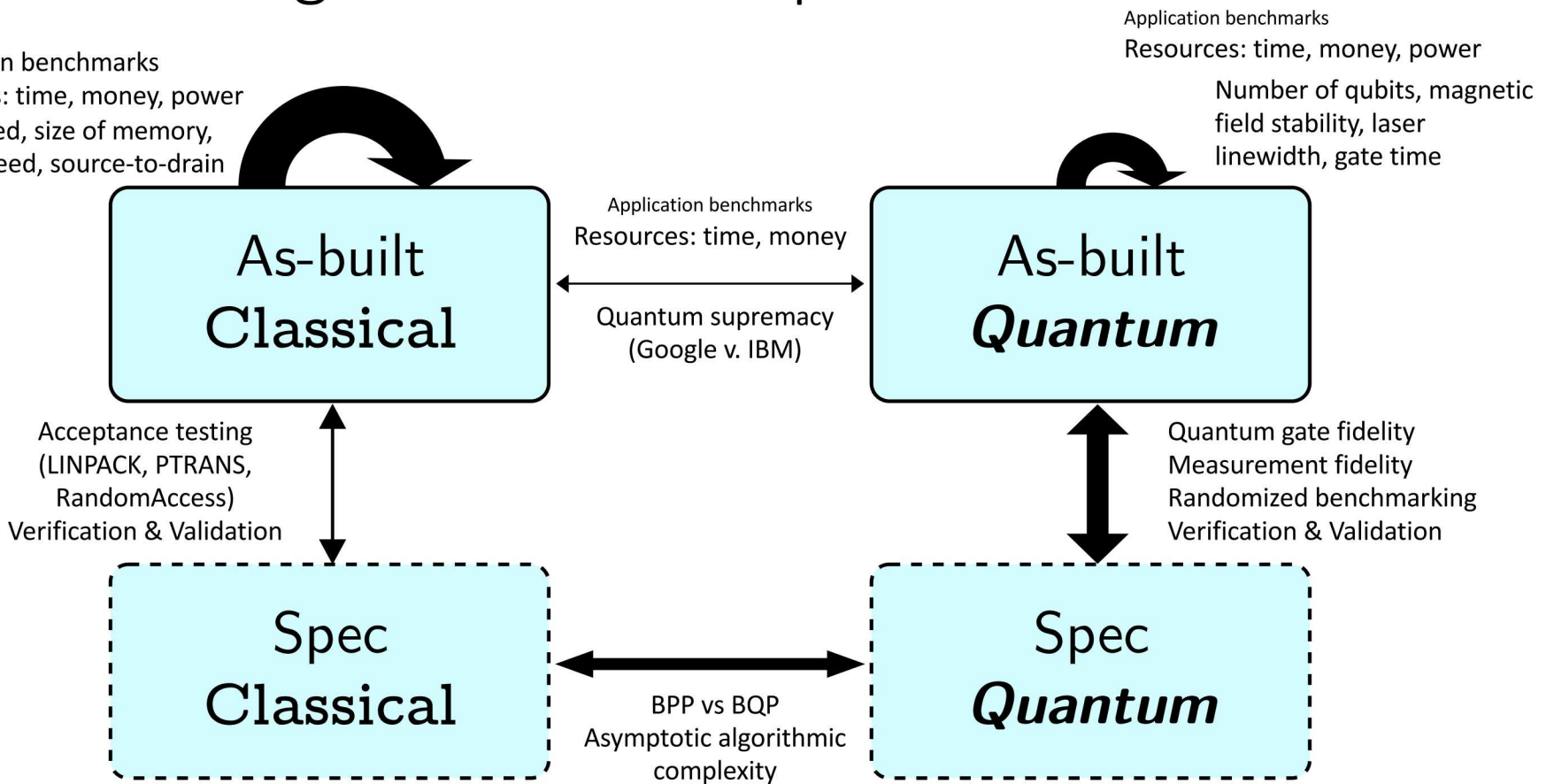
A useful analogy – serial vs. parallel



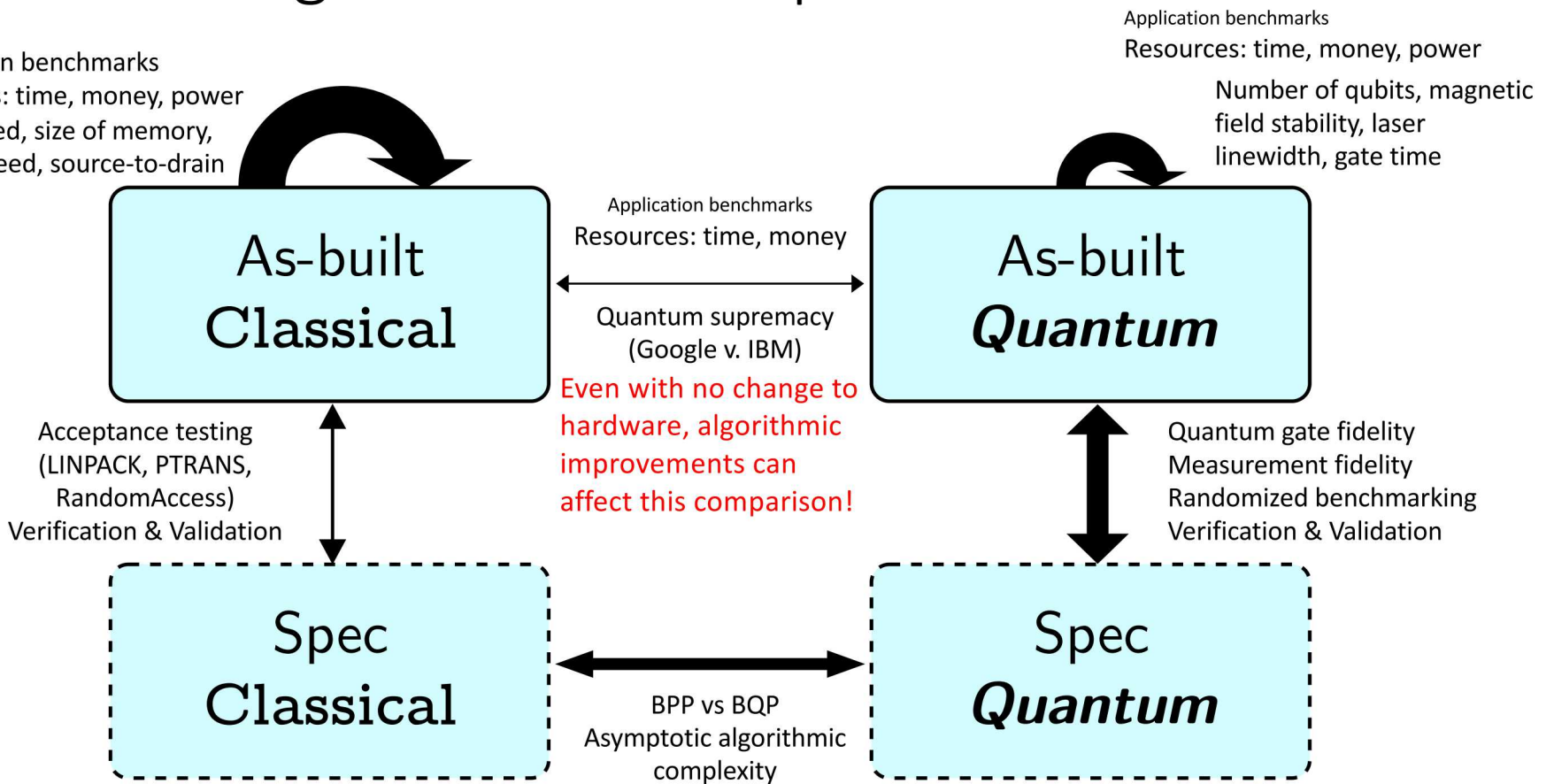
A useful analogy – serial vs. parallel



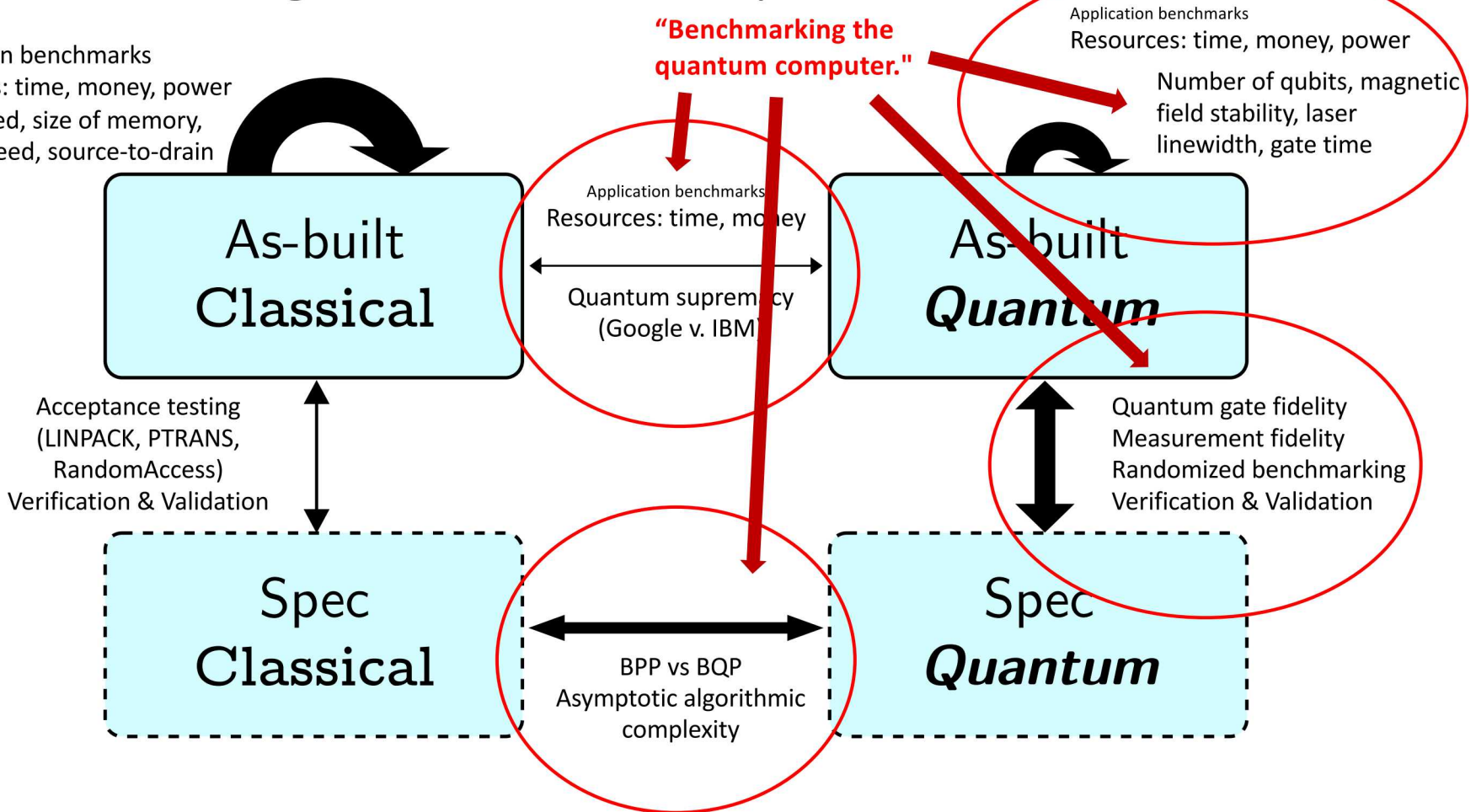
Benchmarking – classical vs. quantum



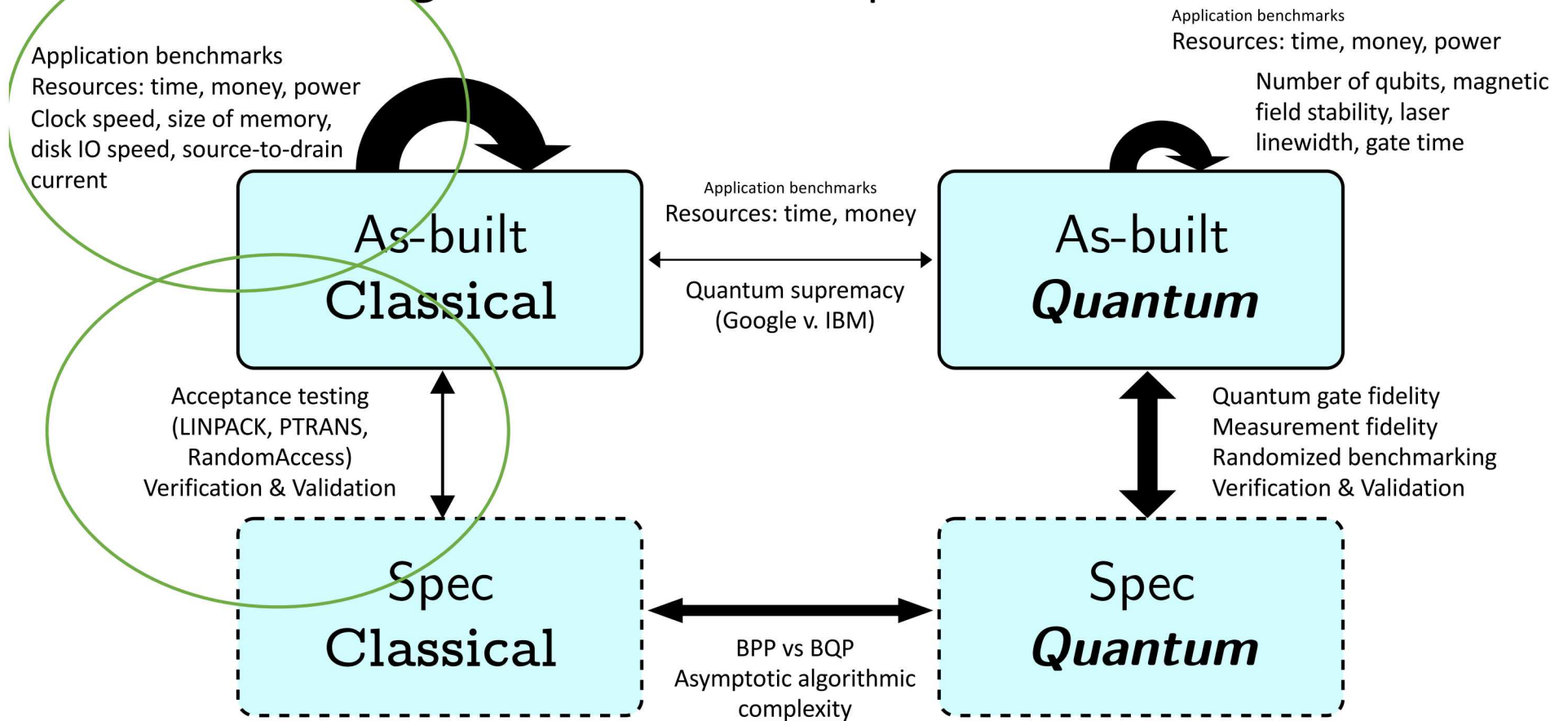
Benchmarking – classical vs. quantum



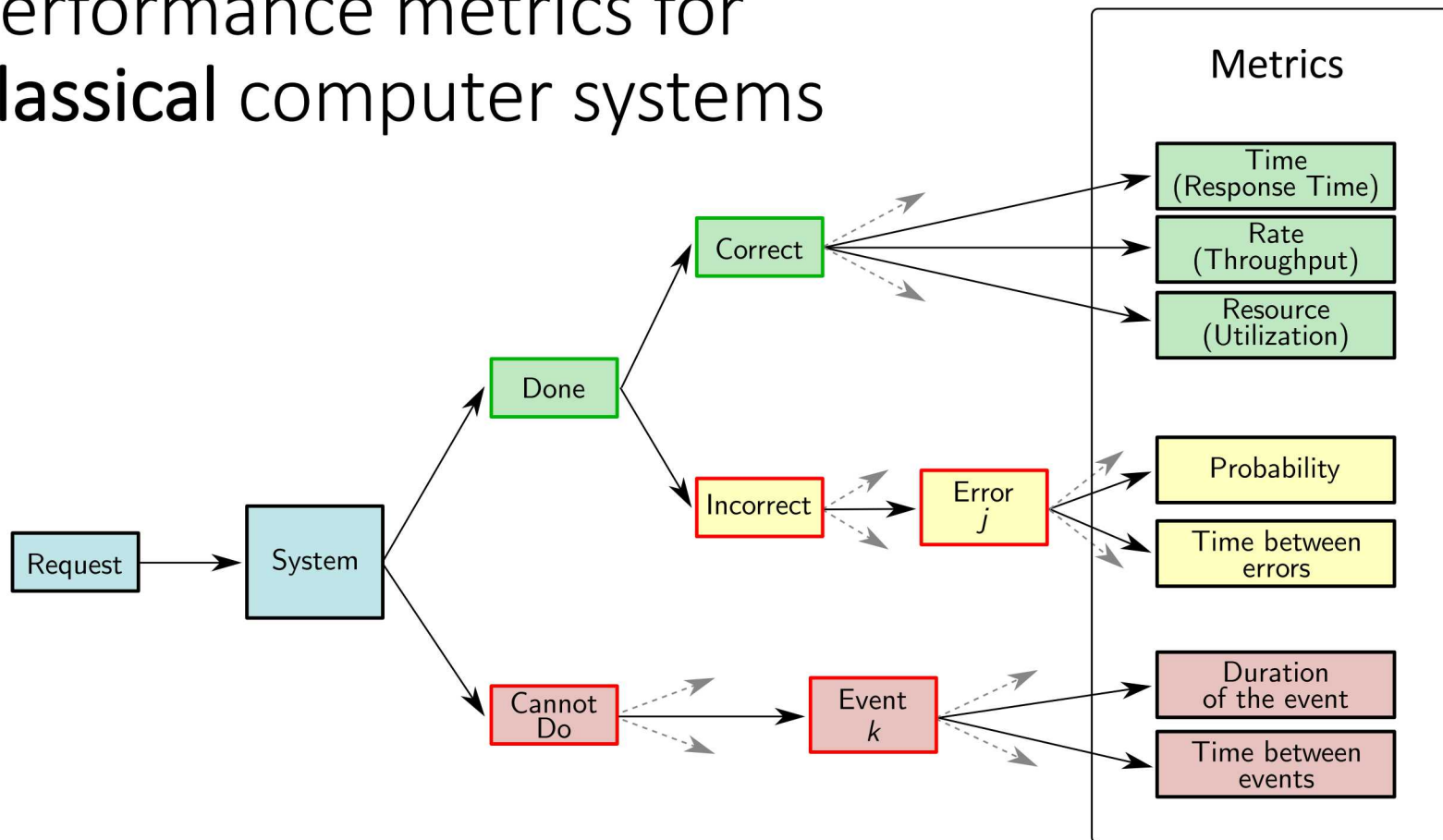
Benchmarking – classical vs. quantum



Benchmarking – classical vs. quantum

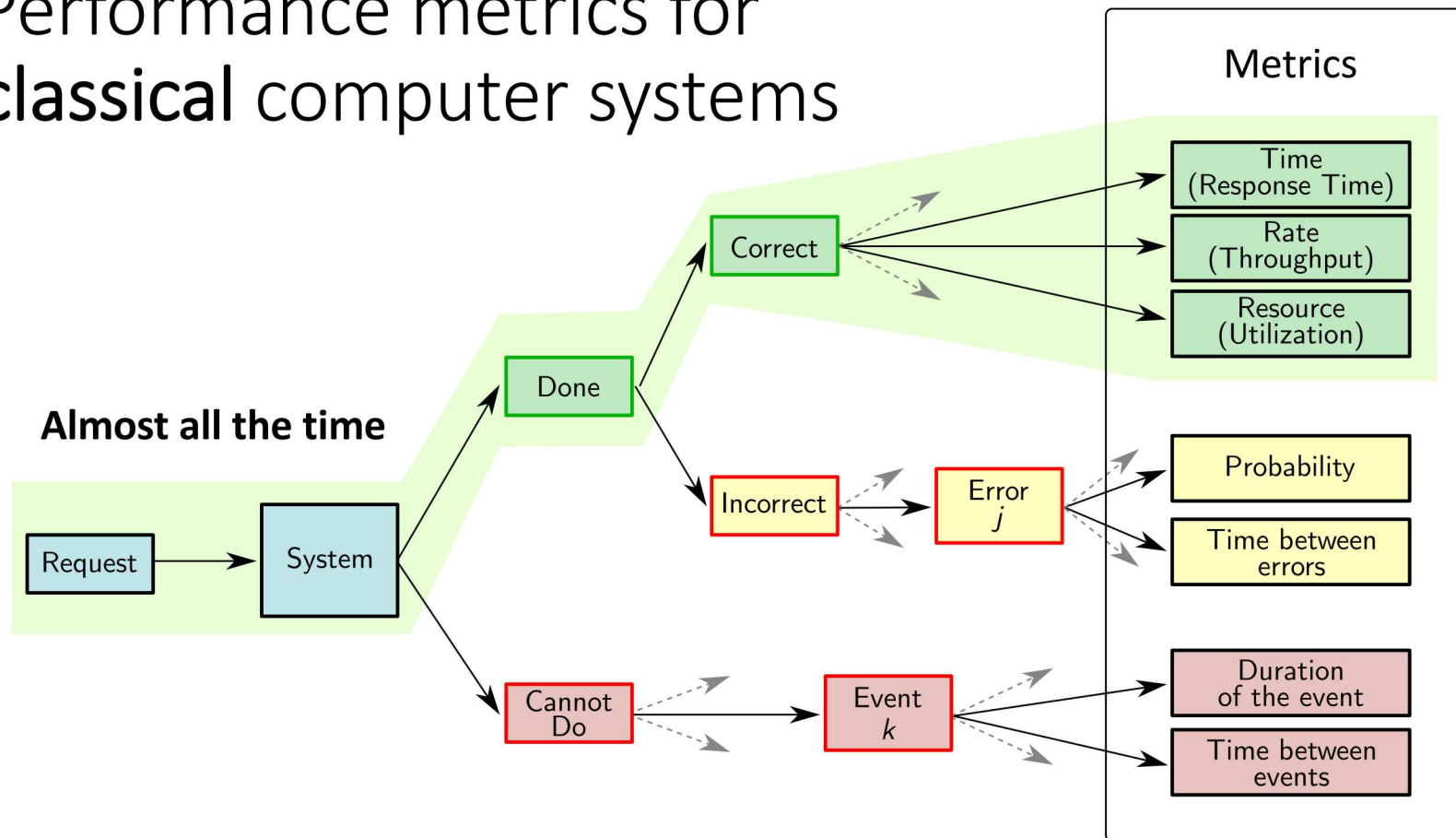


Performance metrics for classical computer systems



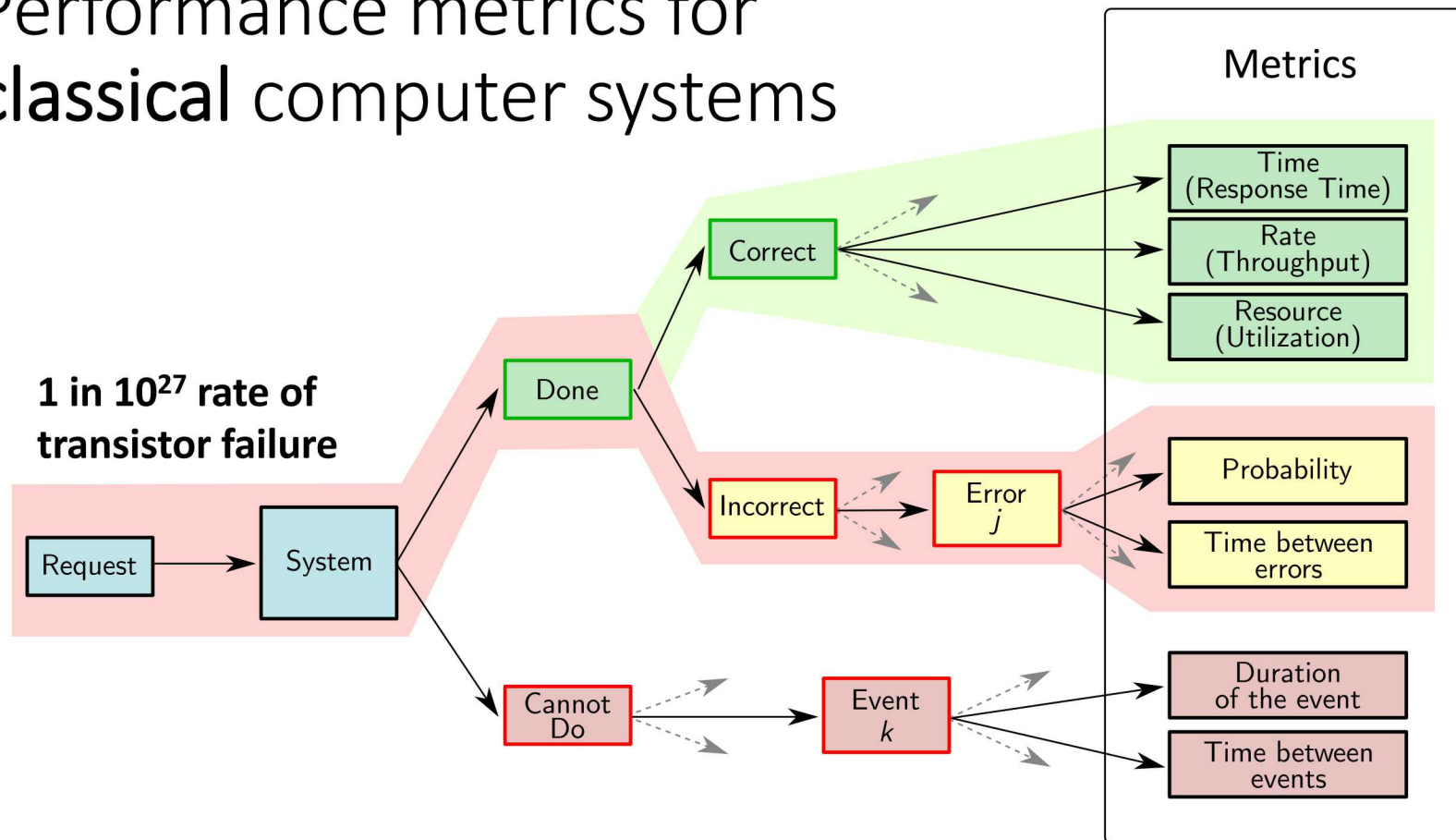
Adapted from R. Jain, *The Art of Computer Benchmarks*, Wiley, (1991)

Performance metrics for classical computer systems



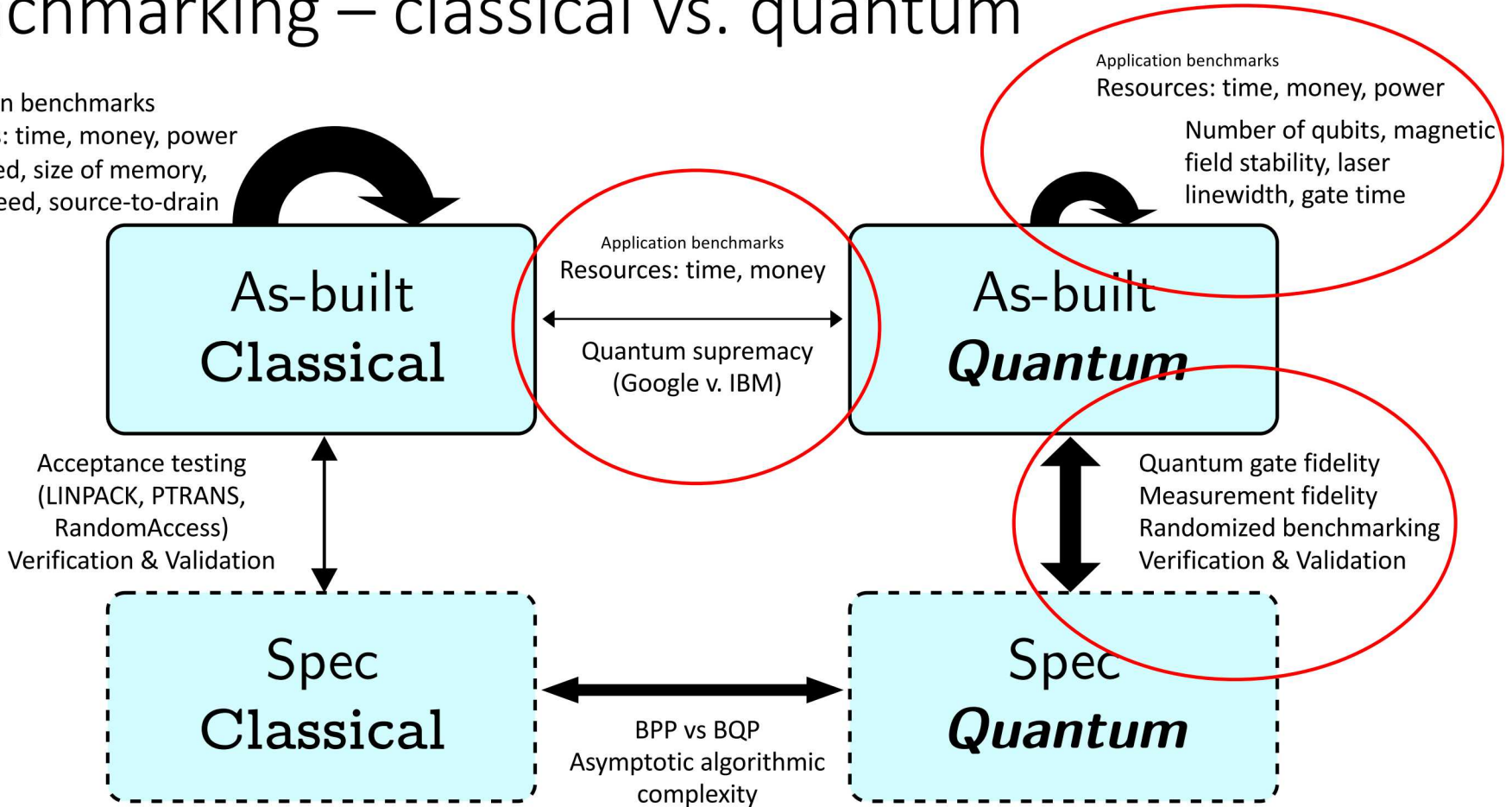
Adapted from R. Jain, *The Art of Computer Benchmarks*, Wiley, (1991)

Performance metrics for classical computer systems

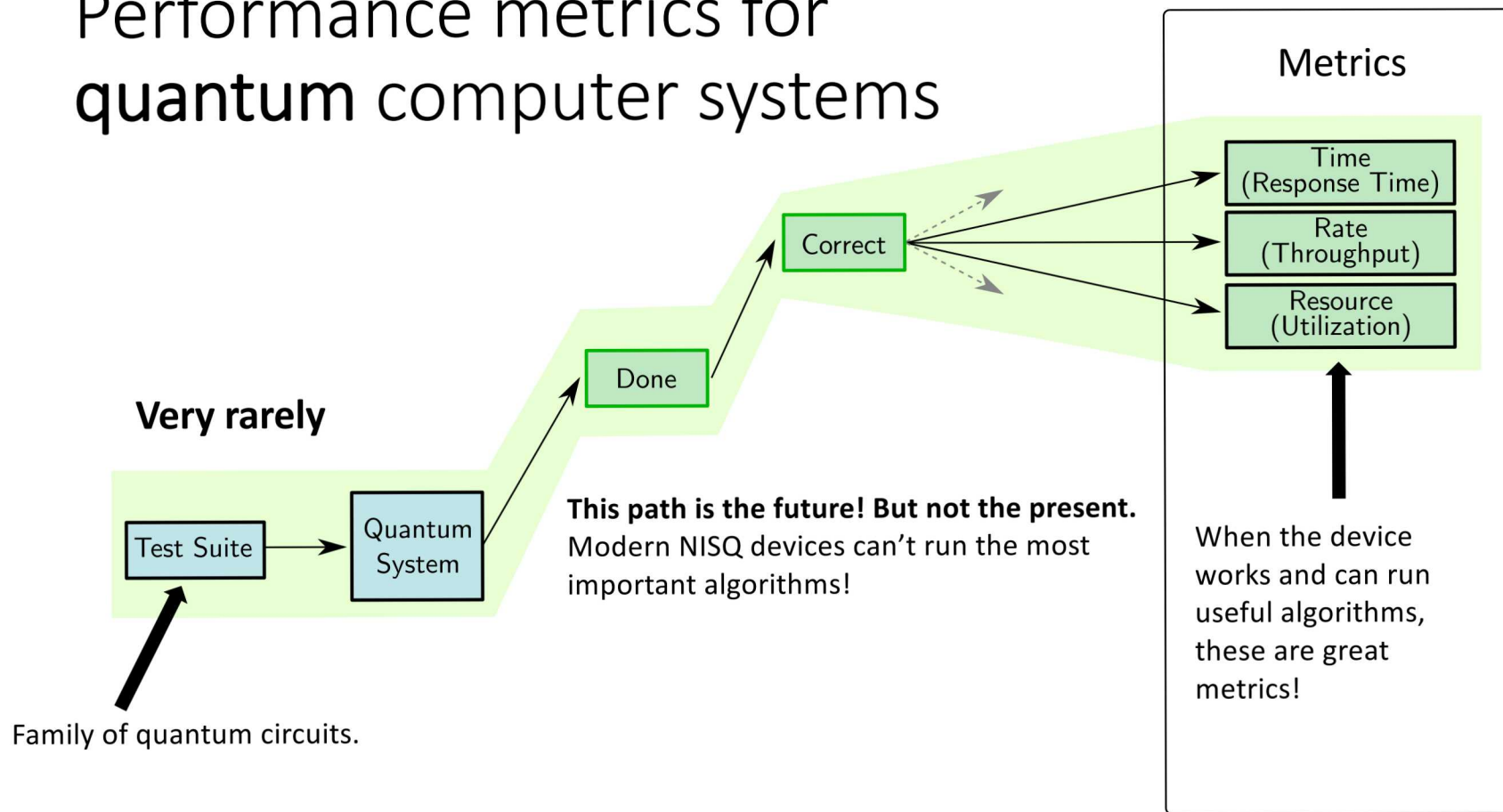


Adapted from R. Jain, *The Art of Computer Benchmarks*, Wiley, (1991)

Benchmarking – classical vs. quantum



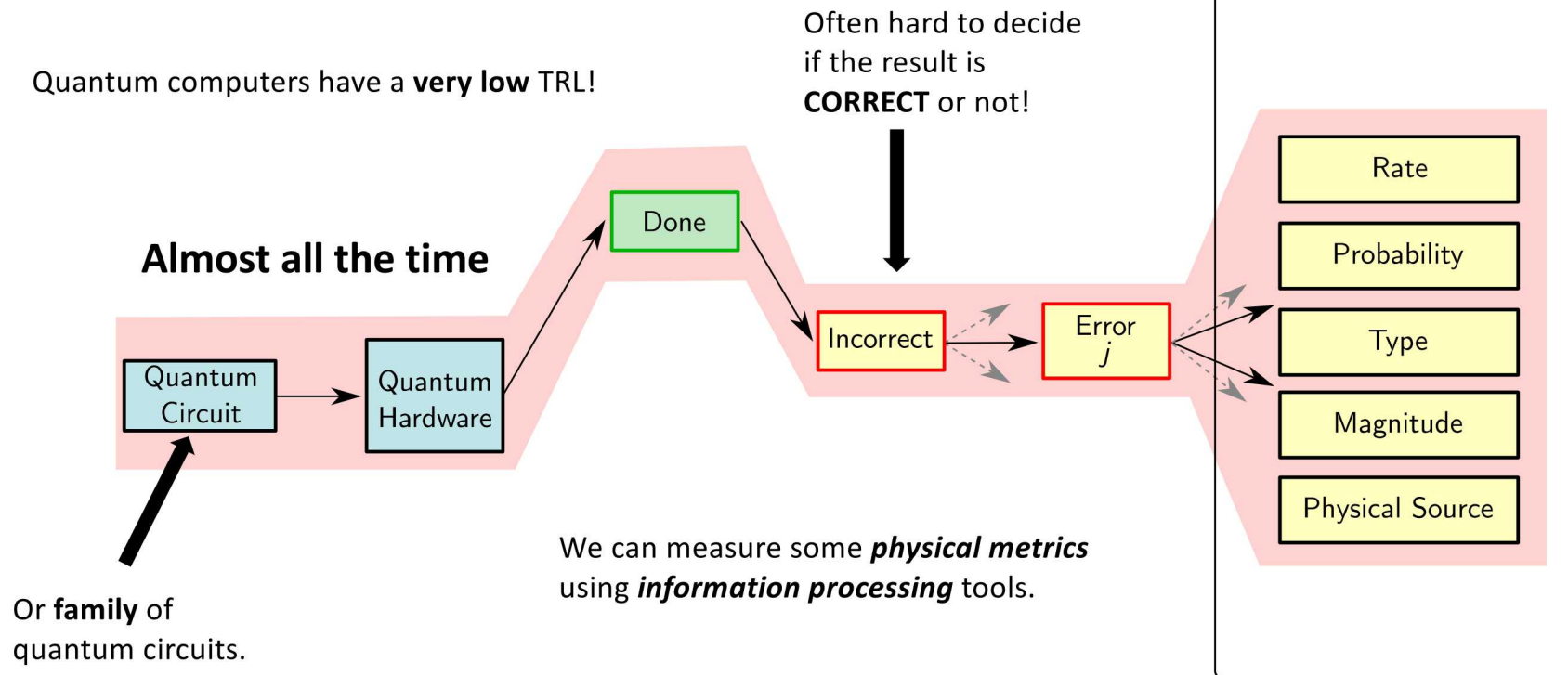
Performance metrics for quantum computer systems



Adapted from R. Jain, *The Art of Computer Benchmarks*, Wiley, (1991)

Performance metrics for quantum computer systems

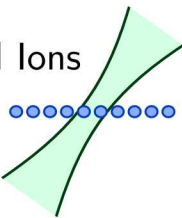
Quantum computers have a **very low** TRL!



Adapted from R. Jain, *The Art of Computer Benchmarks*, Wiley, (1991)

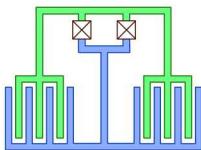
Low-level benchmarking ions vs. superconductors

Trapped Ions



Trap frequency stability
Scattering rate Heating rate
Laser linewidth Magnetic field stability
Laser amplitude stability Gate time
T1 Time Laser amplitude stability

Superconductors



Gate time
Crosstalk T1 time / stability
Josephson Energy
Leakage/seepage Rates
Inhomogeneity

Shared Metrics

Average gate infidelity/diamond distance
State infidelity/trace distance
Measurement fidelity
Randomized benchmarking score
Average cross entropy/TVD
Quantum Volume
Volumetric benchmarks
Time to solution / solution quality

Takeaways

1. “Benchmarking a quantum computer” can refer to several distinct, but related tasks:
 1. Quantum (as-built) vs. Quantum (as-built)
 2. Quantum (as-built) vs. Classical (as-built)
 3. Quantum (spec) vs. Classical (as-built)
 4. Quantum (spec) vs. Classical (spec)
2. The TRL of (gate-model) quantum computers is **TOO LOW** for them to be compared directly to classical computers.
3. Most quantum benchmarking effort goes towards **characterizing errors** in quantum primitives.