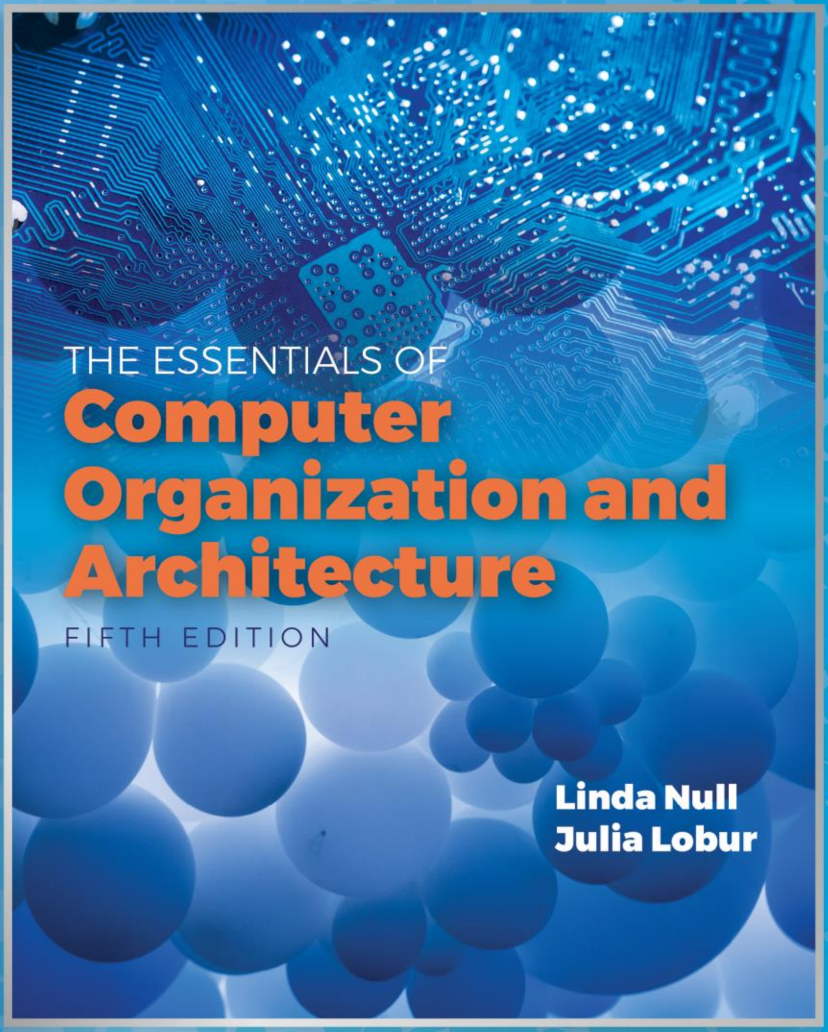
This is the second lecture of Chapter 11

Chapter 11

Performance Measurement and Analysis (B)



Quick review of last lecture

- Introduction
- Computer Performance Equations
  - CPU time
  - Amdahl's Law
- Quantitative Principles of Computer Design
  - Amdahl's Law and Parallel Speedup
  - The CPU Performance Equation
    - Average CPI, IC, Clock Cycle Time, Clock Rate

11.3 Mathematical Preliminaries (1 of 17)

- Measures of system performance depend upon one's point of view.
  - A computer user is most often concerned with response time: How long does it take the system to carry out a task?
  - System administrators are usually more concerned with throughput: How many concurrent tasks can the system handle before response time is adversely affected?
- These two ideas are related: If a system carries out a task in k seconds, then its throughput is $1/k$ of these tasks per second.

11.3 Mathematical Preliminaries (2 of 17)

- In comparing the performance of two systems, we measure the time that it takes for each system to do the same amount of work.
- Specifically, if System A and System B run the same program, System A is n times as fast as System B if:

$$\frac{\text{running time on system B}}{\text{running time on system A}} = n$$

- System A is x% faster than System B if:

$$\left[\frac{\text{running time on system B}}{\text{running time on system A}} - 1\right] \times 100\% = x\%$$

11.3 Mathematical Preliminaries (3 of 17)

- Suppose we have two racecars that have just completed a 10 mile race. Car A finished in 3 minutes, and Car B finished in 4 minutes. Using our formulas, Car A is 1.33 times as fast as Car B, and Car A is also 33% faster than Car B:

$$\frac{\text{time for Car B to travel 10 miles}}{\text{time for Car A to travel 10 miles}} = \frac{4}{3} = 1.33.$$

$$\left[\frac{\text{running time on system B}}{\text{running time on system A}} - 1\right] \times 100\% = \left(\frac{4}{3} - 1\right) \times 100 = 33\%$$

11.3 Mathematical Preliminaries (4 of 17)

- When we are evaluating system performance we are most interested in its expected performance under a given workload.
- We use statistical tools that are measures of central tendency.
- The one with which everyone is most familiar is the arithmetic mean (or average), given by:

$$\frac{\sum_{i=1}^{n} x_i}{n}$$

11.3 Mathematical Preliminaries (5 of 17)

- The arithmetic mean can be misleading if the data are skewed or scattered.
  - Consider the execution times given in the table below. The performance differences are hidden by the simple average.

Program	System A Execution Time	System B Execution Time	System C Execution Time
v	50	100	500
w	200	400	600
x	250	500	500
y	400	800	800
z	5,000	4,100	3,500
Average	1,180	1,180	1,180

11.3 Mathematical Preliminaries (6 of 17)

- If execution frequencies (expected workloads) are known, a weighted average can be revealing.
  - The weighted average for System A is:
    - $50 \times 0.5 + 200 \times 0.3 + 250 \times 0.1 + 400 \times 0.05 + 5000 \times 0.05 = 380$.

Program	Execution Frequency	System A Execution Time	System C Execution Time
v	50%	50	500
w	30%	200	600
x	10%	250	500
y	5%	400	800
z	5%	5,000	3,500
Weighted Average		380 seconds	695 seconds

11.3 Mathematical Preliminaries (7 of 17)

- However, workloads can change over time.
  - A system optimized for one workload may perform poorly when the workload changes, as illustrated below.

Program	Execution Time	Execution Frequency
v	50	25%
w	200	5%
x	250	10%
y	400	5%
z	5,000	55%
Weighted Average	2,817.5 seconds	

11.3 Mathematical Preliminaries (8 of 17)

- When comparing the relative performance of two or more systems, the geometric mean is the preferred measure of central tendency.
  - It is the *n*th root of the product of *n* measurements.

$$G = (x_1 \times x_2 \times x_3 \times \cdots \times x_n)^{\frac{1}{n}}$$

- Unlike the arithmetic means, the geometric mean does not give us a real expectation of system performance. It serves only as a tool for comparison.

Normalized Radio Using System B as a Reference

- The geometric mean of System A, when using System B as a reference is as below.

$$\left(\frac{50}{100}\times\frac{200}{400}\times\frac{250}{500}\times\frac{400}{800}\times\frac{5000}{4100}\right)^{\frac{1}{5}} = 0.59765$$

Program	System A Execution Time	System B Execution Time	System C Execution Time
v	50	100	500
w	200	400	600
x	250	500	500
y	400	800	800
z	5,000	4,100	3,500
Average	1,180	1,180	1,180

11.3 Mathematical Preliminaries (9 of 17)

- The geometric mean is often using normalized ratios between a system under test and a reference machine.
  - We have performed the calculation in the table below.

Program	System A Execution Time	Execution Time Normalized to B	System B Execution Time	Execution Time Normalized to B	System C Execution Time	Execution Time Normalized to B
v	50	0.5	100	1	500	5
w	200	0.5	400	1	600	1.5
x	250	0.5	500	1	500	1
y	400	0.5	800	1	800	1
z	5,000	1.22	4,100	1	3,500	0.85366
Geometric Mean		0.59765		1		1.44967

- $$\frac{Geometric\ Mean\ A}{Geometric\ Mean\ B} = \frac{0.59765}{1} = 0.59765$$

11.3 Mathematical Preliminaries (10 of 17)

- When another system is used for a reference machine, we get a different set of numbers.

Program	System A Execution Time	Execution Time Normalized to C	System B Execution Time	Execution Time Normalized to C	System C Execution Time	Execution Time Normalized to C
v	50	0.1	100	0.2	500	1
w	200	0.3333	400	0.66667	600	1
x	250	0.5	500	1	500	1
y	400	0.5	800	1	800	1
z	5,000	1.42857	4,100	1.17143	3,500	1
Geometric Mean		0.41223		0.68981		1

- $\frac{Geometric\ Mean\ A}{Geometric\ Mean\ B} = \frac{0.41223}{0.68981} = 0.597599$

11.3 Mathematical Preliminaries (11 of 17)

- The real usefulness of the normalized geometric mean is that no matter which system is used as a reference, the ratio of the geometric means is consistent.
- This is to say that the ratio of the geometric means for System A to System B, System B to System C, and System A to System C is the same no matter which machine is the reference machine.
- Using the results that we got when using System B and System C as reference machines we find that the ratio of the geometric means for System A to System B are 0.59765/1 = $0.41223/0.68981$

11.3 Mathematical Preliminaries (12 of 17)

- The inherent problem with using the geometric mean to demonstrate machine performance is that all execution times contribute equally to the result.
- So shortening the execution time of a small program by 10% has the same effect as shortening the execution time of a large program by 10%.
  - Shorter programs are generally easier to optimize, but in the real world, we want to shorten the execution time of longer programs.
- Also, if the geometric mean is not proportionate. A system giving a geometric mean 50% smaller than another is not necessarily twice as fast!

11.3 Mathematical Preliminaries (13 of 17)

- The harmonic mean provides us with a way to compare execution times that are expressed as a rate, such as operations per second.
- The harmonic mean allows us to form a mathematical expectation of throughput, and to compare the relative throughput of systems and system components.
- To find the harmonic mean, we add the reciprocals of the rates and divide them into the number of rates:

$$H = n \div (1/x_1 + 1/x_2 + 1/x_3 + \ldots + 1/x_n)$$

Harmonic Mean Example

- Driving 30 miles
  - First 10 miles, driving at 30 miles per hour
  - Second 10 miles, driving 40 miles per hour
  - Last 10 miles, droving 60 miles per hour
- What is average speed
  - 30 miles /(time_first10miles + time_second10miles + time_last_10miles)

    = 30 /(10/30 + 10/40 + 10/60)

    = 3 / (1/30 + 1/40 + 1/60) = 40 miles per hour
- Arithmetic average gives wrong result
  - (30+40+60)/3 = 43 miles per hour

11.3 Mathematical Preliminaries (14 of 17)

- The harmonic mean holds two advantages over the geometric mean.
- First, it is a suitable predictor of machine behavior.
  - So it is useful for more than simply comparing performance.
- Second, the slowest rates have the greatest influence on the result, so improving the slowest performance—usually what we want to do—results in better performance.
- The main disadvantage is that the harmonic mean is sensitive to the choice of a reference machine.

11.3 Mathematical Preliminaries (15 of 17)

- This chart summarizes when the use of each of the performance means is appropriate.

Mean	Uniformly Distributed Data	Skewed Data	Data Expressed as a Ratio	Indicator of System Performance Under a Known Workload	Data Expressed as a Rate
Arithmetic	X			X	
Weighted Arithmetic		X		X	
Geometric		X	X		
Harmonic				X	X