1. Response Time

Definition: Time taken from start to finish of a specific task.
Also called execution time.
Includes all delays, such as:

- CPU processing time

- Memory access time

- Disk I/O

- Operating system overhead

Useful for evaluating how fast a system completes individual tasks.

2. Throughput

Definition: The number of tasks completed in a given time.
Think of it as work done per second, like:

- "Jobs per hour"

- "Requests per second"

Useful for evaluating overall system productivity, especially in servers or multitasking systems.

Relation Between Response Time and Throughput

Generally:

Decreasing response time (faster individual tasks) $\Rightarrow$ Increases throughput (more tasks completed in

But not always strictly linear—depends on the workload and resource management.

3. Performance and Execution Time

Formula:
$$\text{Performance}_x = \frac{1}{\text{Execution Time}}$$
Meaning: Shorter execution time = better performance.

Inverse relationship: Faster computers have lower execution time, so higher performance.

4. Comparing Two Computers (X and Y)

If:
$$\text{Performance}_x > \text{Performance}_y$$

Then:

$$\frac{1}{\text{Execution Time}_x} > \frac{1}{\text{Execution Time}_y} \quad \Rightarrow \quad \text{Execution Time}_y > \text{Execution Time}_x$$

So, X completes tasks faster than Y.

5. Speedup (n times faster)

If computer X is n times faster than computer Y:

$$\frac{\text{Performance}_x}{\text{Performance}_y} = n \quad \text{OR} \quad \frac{\text{Execution Time}_y}{\text{Execution Time}_x} = n$$

Example: If:

- Y takes 10 seconds
- X is 2 times faster

Then:
$$X \text{ takes only 5 seconds.}$$

Summary

Term	Definition	Goal
Response Time	Time to finish one task	Minimize
Throughput	Number of tasks completed per unit time	Maximize
Performance	Inversely related to execution time	Higher is better
Speedup (n×)	One system is n times faster than another	Indicates performance gap