ECE 454
Computer Systems Programming
What is Performance

Ding Yuan
ECE Dept., University of Toronto
http://www.eecg.toronto.edu/~yuan

Announcement

- HW schedule (tentative)

	Release	Due	Lab hour	TA
HW1	16-Sep	30-Sep	9/19 * 1 and 9/26 * 2	Michelle Wong
HW2	30-Sep	14-Oct	10/10 * 2	Michelle Wong
HW3	10-Oct	11-Nov	10/31 * 1 and 11/7 * 2	Xu Zhao
HW4	07-Nov	22-Nov	11/21 * 2	Yongle Zhang
HW5	20-Nov	06-Dec	12/5 * 2	Yongle Zhang

- Please log-in to Piazza
  - When you ask questions, mark it with appropriate label
    - e.g., lecture, final_exam, logistic, HW1, HW2..
    - Instructors are responsible for questions with different label

Before we go on…

- What do you exactly mean by "*performance*"?
  - Simple program: speed -- how fast your program runs
  - Unix "time" command

- Server program
  - Is "speed" the only important thing?
  - What is the "speed" for long running programs?
  - *Latency* vs. *throughput*

9/10/13   3   *Ding Yuan, ECE454*

Latency vs. throughput

- Latency
  - How fast the server respond my request?
    - Sometimes also called response time

- Throughput
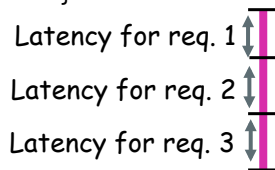 - Number of requests served/unit time

- Relationship?

9/10/13   4   *Ding Yuan, ECE454*

Positive correlation example

```
void dummy_server () {
  while (request = next_request ()) {
    respond (request);
  }
}
```

Latency for req. 1

Latency for req. 2

Latency for req. 3

$$\text{Throughput} = \frac{3}{L1 + L2 + L3}$$

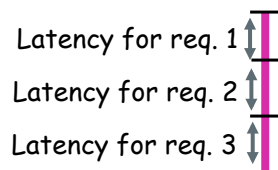*If we have a faster CPU, both latency and throughput will improve (smaller latency, higher throughput)!*
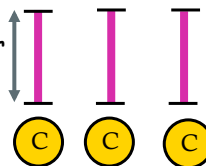
Negative correlation example

```
void dummy_server () {
  while (request = next_request ()) {
    thread_create(respond, request);
  }
}
```

Latency will be worse (lower), why?
Throughput will be better (higher), why?

Latency for req. 1

Latency for req. 2

Latency for req. 3

✂ Latency for each req

C C C

Before parallelization

After parallelization

Real life analogy

Plane	Toronto to Paris	Speed	Passenger	Throughput (pmph)
Boeing 747	8 hours	610 mph	470	286,700
Concorde	4 hours	1350 mph	132	178,200

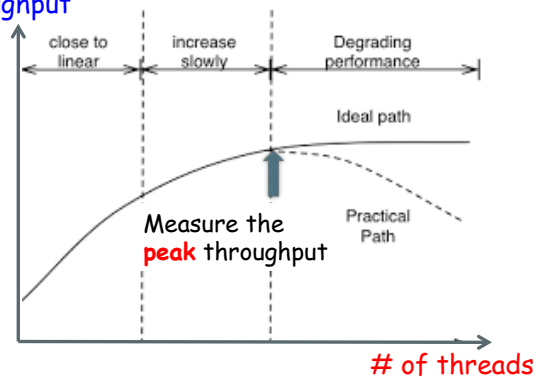Which plane has higher performance?

9/10/13   7   Ding Yuan. ECE454

Parallelism vs. Throughput

- Will more *parallelism* always improve *throughput*?

throughput

close to linear | increase slowly | Degrading performance

Ideal path

Measure the **peak** throughput

Practical Path

\# of threads

9/10/13   8   Ding Yuan. ECE454

Performance measurement is a very complicated problem

- Other metrics: bandwidth, jitter, etc.

- Extra considerations: best case? worst case? average?

- Different applications have different requirements
  - Netflix
  - Google/Facebook/Amazon
  - Online gaming
  - Flight control software on airplane

- ACM special interest group on performance evaluation (SIGMETRICS)

9/10/13   9   Ding Yuan, ECE454