Operational Laws Review

Little's Law for Open Systems:

- $\mathbf{E}[N] =$ ______________________
- $\mathbf{E}[N_Q] =$ ______________________
- $\mathbf{E}[N_{red}] =$ ______________________

Little's Law for Closed Systems:

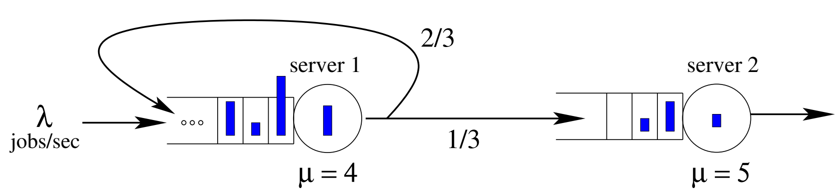
- $N =$ ______________________
- Fix $\mathbf{E}[Z]$. When $\mathbf{E}[R]$ goes down, what happens to X?

Utilization Law (applies to single device i):

- $\rho_i =$ ______________________

Forced Flow Law:

- X = system throughput
- X_i = throughput at device i
- V_i = number of visits to device i per job
- Forced Flow Law: ______________________



Example

Suppose we have an interactive system with the following characteristics:

- 25 user terminals ($N = 25$)
- 18 seconds average think time ($\mathbf{E}[Z] = 18$)
- 20 visits to a specific disk per interaction on average ($\mathbf{E}[V_{disk}] = 20$)
- 30% utilization of that disk ($\rho_{disk} = .3$)
- .025 sec average service time per visit to that disk ($\mathbf{E}[S_{disk}] = .025$)

That is all the information we have.

Question: What is the mean response time, $\mathbf{E}[R]$?

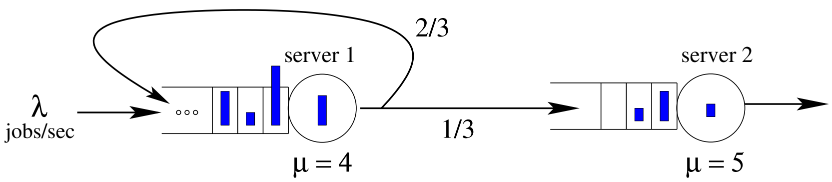
Device Demands

Defn: D_i = total demand on device i for all visits of a single arriving job.

$$D_i = \underline{\hspace{4cm}}$$

$$\mathbf{E}[D_i] = \underline{\hspace{4cm}}$$

EXAMPLE:



Ques: What is $\mathbf{E}[D_1]$?

Bottleneck Law

Ques: How can we express ρ_i in terms of $\mathbf{E}[D_i]$?

PROOF OF BOTTLENECK LAW:

Modification Analysis

REST OF TODAY and FRIDAY:

- Closed Systems Only – most effective there

- We will see how to estimate X and $\mathbf{E}[R]$ as a function of N
  – "asymptotic bounds"

- Modification analysis, a.k.a. "what-if analysis"
  – work of Systems Consultant

Asymptotic bounds theorem

Theorem: Given closed interactive system with N users and m devices. Let

$$D = \sum_{i=1}^{m} \mathbf{E}[D_i] \qquad\qquad D_{max} = \max_i \{\mathbf{E}[D_i]\}$$

Then

$$X \;\leq\; \min\left(\frac{N}{D + \mathbf{E}[Z]}\;,\;\frac{1}{D_{max}}\right)$$

$$\mathbf{E}[R] \;\geq\; \max\left(D\;,\; N D_{max} - \mathbf{E}[Z]\right)$$

where the first term in each is an asymptote for small N, and the second term in each is an asymptote for large N.

PROOF FOR LARGE N ASYMPTOTE:

Asymptotic bounds theorem

Theorem: Given closed interactive system with N users and m devices. Let

$$D = \sum_{i=1}^{m} \mathbf{E}[D_i] \qquad\qquad D_{max} = \max_i \{\mathbf{E}[D_i]\}$$

Then

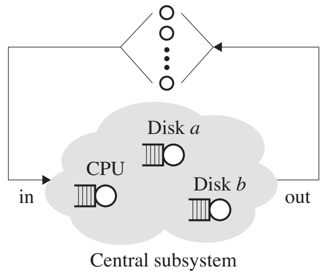
$$X \;\leq\; \min\left(\frac{N}{D + \mathbf{E}[Z]}\;,\;\frac{1}{D_{max}}\right)$$

$$\mathbf{E}[R] \;\geq\; \max\left(D\;,\;ND_{max} - \mathbf{E}[Z]\right)$$

where the first term in each is an asymptote for small N, and the second term in each is an asymptote for large N.

PROOF FOR SMALL N ASYMPTOTE:

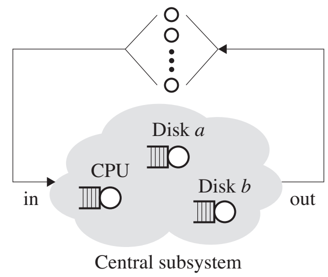
Bounds Example



$$\mathbf{E}[Z] = 18s \qquad \mathbf{E}[D_{cpu}] = 5s \qquad \mathbf{E}[D_A] = 4s \qquad \mathbf{E}[D_B] = 3s \ .$$

Determine X and $\mathbf{E}[R]$ as a function of N:

Bounds Example, cont.



$\mathbf{E}[Z] = 18s \qquad \mathbf{E}[D_{cpu}] = 5s \qquad \mathbf{E}[D_A] = 4s \qquad \mathbf{E}[D_B] = 3s$.

Ques: Where is the "knee", N^*, of these curves?

Ques: What does N^* represent?

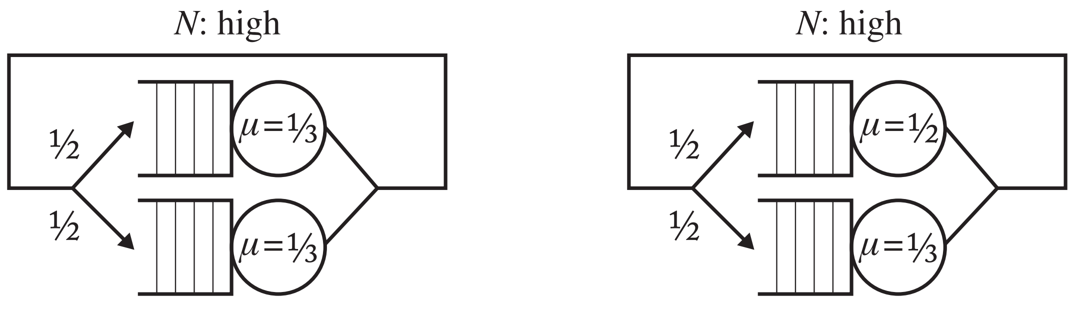
Ques: If $N > N^*$, what do we need to do to increase X?

Ques: What if we instead decrease $D_{\text{next-to-max}}$?

Ques: How does this change when $\mathbf{E}[Z] = 0$ (batch system)?

Easy Modification Analysis Example

Coming up on your next homework ...



(a) Original **(b) "Improved"**

Question: Does the improvement help increase X? Does it reduce $\mathbf{E}[T]$?

Commercial Break: Announcements

1. Mor's office hours today! 5:30 p.m. - 7 p.m. in GHC 7207.
2. Keerthana's office hours tomorrow! 5:30 p.m. - 7 p.m. in GHC 7004.
3. Special class Friday! Zhouzi will cover new 2024 paper on Modification Analysis used to improve the performance of a caching system implementation.

Harder Modification Analysis Example

The following measurements were obtained for an interactive system over some time:[1]

- $B_{\text{cpu}} = 400$ seconds
- $B_{\text{slowdisk}} = 100$ seconds
- $B_{\text{fastdisk}} = 600$ seconds
- $C = C_{\text{cpu}} = 200$ jobs
- $C_{\text{slowdisk}} = 2,000$ jobs
- $C_{\text{fastdisk}} = 20,000$ jobs
- $\mathbf{E}[Z] = 15$ seconds
- $N = 20$ users

Your job is to examine four possible improvements (modifications):

1. **Faster CPU:** Replace the CPU with one that's twice as fast.
2. **Balancing slow and fast disks:** Shift some files from the fast disk to the slow disk, balancing their demand.
3. **Second fast disk:** Buy a second fast disk to handle half the load of the busier existing fast disk.
4. **Balancing among three disks plus faster CPU:** Make all three improvements together: Buy a second fast disk, balance the load across all three disks, and also replace the CPU with a faster one.

[1] Just as most fairy tales start with "once upon a time," most performance analysis problems begin with "the following measurements were obtained."

Start by filling in these quantities:

- $D_{\text{cpu}} =$

- $D_{\text{slowdisk}} =$

- $D_{\text{fastdisk}} =$

- $\mathbf{E}\left[V_{\text{cpu}}\right] =$

- $\mathbf{E}\left[V_{\text{slowdisk}}\right] =$

- $\mathbf{E}\left[V_{\text{fastdisk}}\right] =$

- $\mathbf{E}\left[S_{\text{cpu}}\right] =$

- $\mathbf{E}\left[S_{\text{slowdisk}}\right] =$

- $\mathbf{E}\left[S_{\text{fastdisk}}\right] =$

- $N^* =$

Now evaluate each of the 4 potential modifications:

Why we like modification analysis

1. It's easy – "back of envelope"

2. Distribution independent

3. Bounds only, but sufficient to make decisions

Ques: Why doesn't asymptotic analysis make sense for *open* systems?