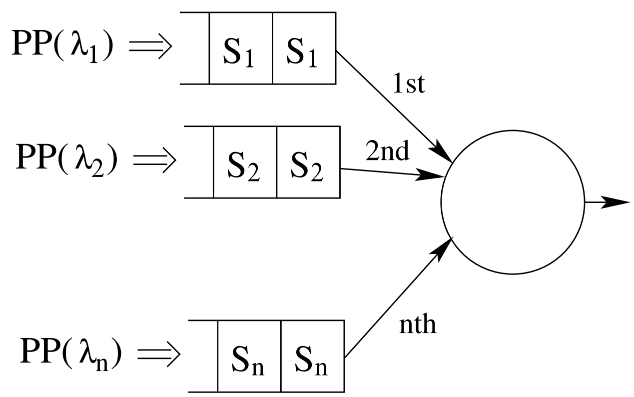
As usual: M/G/1, where job size S has finite mean and variance. Assume $\rho < 1$, and some work-conserving scheduling policy.

	Non–Preemptive	Preemptive
Don't Use Size		
Use Size/ Class		

TODAY: Preemptive policies which use size/class (Chpt 32).

1 Preemptive Priority Queue

Let $\lambda_k = \lambda \cdot p_k$



Question: What is S?

Question: What is ρ_k?

Question: What is $\sum_{k=1}^{n} \rho_k$?

GOAL: $\mathbf{E}[T(k)]$

Question: What are the 3 things that have to happen before type k job can leave?

(1):

(2):

(3):

Derivation of (2):

$$
\begin{aligned}
(\mathbf{2}) &= \mathbf{E}\left[\text{Remaining work in P-Prio system of class 1 through } k\right] \\[2em]
&= \mathbf{E}\left[\text{Total rem. work in P-Prio system assuming } \underline{\hspace{6cm}}\right] \\[2em]
&= \mathbf{E}\left[\text{Total work in } \underline{\hspace{2.5cm}} \text{ system assuming } \underline{\hspace{6cm}}\right] \\[2em]
&= \mathbf{E}\left[\underline{\hspace{1.2cm}} \text{ in } \underline{\hspace{2.5cm}} \text{ system assuming } \underline{\hspace{7cm}}\right] \\[2em]
&=
\end{aligned}
$$

where we use the notation: $F_k = \sum_{i=1}^{k} p_i$.

which can be simplified to ...

$$(\mathbf{2}) \;=\; \frac{\sum_{i=1}^{k} \rho_i}{1 - \sum_{i=1}^{k} \rho_i} \quad \cdot \quad \frac{\sum_{i=1}^{k} p_i \mathbf{E}\left[S_i^2\right]}{2 \sum_{i=1}^{k} p_i \mathbf{E}\left[S_i\right]}$$

Now use $\rho_i = \lambda \cdot p_i \cdot \mathbf{E}\left[S_i\right]$

$$= \frac{\lambda \sum_{i=1}^{k} p_i \mathbf{E}\left[S_i\right]}{1 - \sum_{i=1}^{k} \rho_i} \quad \cdot \quad \frac{\sum_{i=1}^{k} p_i \mathbf{E}\left[S_i^2\right]}{2 \sum_{i=1}^{k} p_i \mathbf{E}\left[S_i\right]}$$

$$= \frac{\lambda \sum_{i=1}^{k} p_i \mathbf{E}\left[S_i^2\right]}{2\left(1 - \sum_{i=1}^{k} \rho_i\right)}$$

$$= \frac{\sum_{i=1}^{k} \rho_i \frac{\mathbf{E}\left[S_i^2\right]}{2\mathbf{E}\left[S_i\right]}}{1 - \sum_{i=1}^{k} \rho_i}$$

Derivation of (3):

$(\mathbf{3}) =$

Combining: $\mathbf{E}\left[T(k)\right]^{P-Prio} = (1) + (2) + (3)$

Question: How does $\mathbf{E}\left[T(k)\right]^{P-Prio}$ compare with $\mathbf{E}\left[T(k)\right]^{NP-Prio}$?

2 Another Interpretation of P-Prio system

$$\mathbf{E}\left[T(k)\right]^{P-Prio} = \frac{\sum_{i=1}^{k} \rho_i \frac{\mathbf{E}\left[S_i^2\right]}{2\mathbf{E}\left[S_i\right]}}{\left(1-\sum_{i=1}^{k} \rho_i\right)\left(1-\sum_{i=1}^{k-1} \rho_i\right)} + \frac{\mathbf{E}\left[S_k\right]}{1-\sum_{i=1}^{k-1} \rho_i}$$

$$\mathbf{E}\left[T(k)\right]^{P-Prio} = \mathbf{E}\left[Wait(k)\right]^{P-Prio} + \mathbf{E}\left[Res(k)\right]^{P-Prio}$$

3 Commercial Break: Announcements

1. I got **bigger index cards**: 5×7. Feel free to grab these in class or outside my office. These are for the Dec 4th exam, which will be held in our classroom starting at 5 p.m. The exam covers Chpts 14-33, but only the material covered in class or homework.

2. HW 12 is due **Tuesday 5 p.m.** outside Mor's office: GHC 7207.

3. Go to **Zhouzi's office hours today**, right after class, to get help on HW 12 – GHC 6003.

4. Mor will have **extra office hours** on **Tuesday 12 p.m. - 1 p.m.** in GHC 7207 to help you all with unfinished HW 12 questions. You can also ask to copy any of my lecture notes.

4 Preemptive-Shortest-Job-First (PSJF)

PSJF is the preemptive version of SJF. Priority is based on *original size*.

$$\mathbf{E}\left[T(x)\right]^{PSJF} = \mathbf{E}\left[Wait(x)\right]^{PSJF} + \mathbf{E}\left[Res(x)\right]^{PSJF}$$

- $Wait(x)$ represents the time until job x receives its first bit of service.
- $Res(x)$ is the time from when job x receives first receives service until it completes.

5 Prep Work

Let S denote the size of a job. Density $f_S(t)$, Cumulative Distrib. $F_S(t)$.
Let $S_{\leq x}$ denote the size of a job of size $\leq x$. That is, $S_{\leq x} = [S \mid S \leq x]$.
Let $\lambda_{\leq x} = \lambda \cdot F_S(x)$ denote the arrival rate of jobs of size $\leq x$.
Let $\rho_{\leq x}$ denote the load made up of jobs of size $\leq x$.

Question: What is $\mathbf{E}\left[S_{\leq x}\right]$?

Question: What is $\rho_{\leq x}$?

Let $B_{\leq x}$ denote a busy period made up of only jobs from $S_{\leq x}$.

Question: What is $\mathbf{E}\left[B_{\leq x}\right]$?

Prep Work, Continued

Let $B_{\leq x}(y)$ denote a busy period started by a job of size y, made up of only jobs from $S_{\leq x}$.

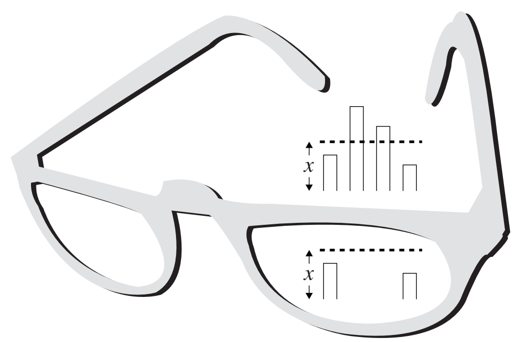
Question: What is $\mathbf{E}\left[B_{\leq x}(y)\right]$?

6 Computing mean residence: $\mathbf{E}\left[Res(x)\right]^{PSJF}$

7 Computing mean waiting time: $\mathbf{E}[Wait(x)]^{PSJF}$

Let $W_{\leq x}$ = work in PSJF system made up of only jobs of size $\leq x$.

Question: How can we express $\mathbf{E}[Wait(x)]^{PSJF}$ in terms of $W_{\leq x}$?



Question: What is $\mathbf{E}[W_{\leq x}]^{PSJF}$?

8 Derive $\mathbf{E}[T(x)]^{PSJF}$

Question: What is $\mathbf{E}[Wait(x)]^{PSJF}$?

Question: What is the final expression for $\mathbf{E}[T(x)]^{PSJF}$?

Notice that a job of size x is affected by only the variability of S due to jobs of size $\leq x$. This is the benefit of PSJF over non-preemptive policies. This is the same benefit we saw for P-PRIO over NP-PRIO.

Question: Why does PSJF win over PS?