Scheduling: Performance and Asymptotics - part I

Bert Zwart
CWI, VU University, Eurandom, Georgia Tech

November 19, 2009

YEQT-III

Scheduling

- Area at the core of OR and CS
- Many applications (not only in computer systems)
- Playground for theoreticians, leading to different paradigms
  - Worst-case analysis (many books, for e.g. Pinedo (2008)), optimization/complexity/approximation algorithms, ...
  - Average case analysis (closer to queueing/performance analysis)
  - "something in between"

Focus: "something in between"

- Main scheduling paradigms: average case analysis and worst case analysis
- Many impressive results are available
- Disjoint communities
- Focus in this tutorial: probabilistic (not worst-case) but focusing on bad events (i.e. not average case)
- Main performance measure: response time (a.k.a. processing time) of a job

Typical questions

- How likely is it that a long sojourn time will occur?
- If a long sojourn time occurs, how does it occur? How can we avoid this?
- Can we avoid long-sojourn times effectively if we don't know much about the distribution of the job sizes?
- How do we deal with multiple job classes?
- Central question: how do we choose a scheduling policy in order to effectively mitigate the negative impact of rare events?
- Use tools from large deviations theory

Some common scheduling disciplines

Key notions:

- Work-conserving
- Preemptive vs. non-preemptive
- Blind vs. size based
- Adaptive vs. non-adaptive

Key examples:

- FCFS/FIFO
- ROS (Random order of Service)
- LCFS/LIFO (Last come first served)
- PS (Processor Sharing)
- SRPT (Shortest Remaining Processing Time)
- FB (Foreground background PS)

Overview

- Tutotial on intersection of:
  - Scheduling
  - Queueing/Performance analysis
  - Large deviations
- Today:
  - Introduction
  - Basics on large deviations (light and heavy tails)
  - Rare events in FIFO queues
- Tomorrow:
  - LIFO, PS, SRPT, ...
  - Robustness and optimality issues
  - Multi-class and multi-node systems

Let $X_i, i \geq 1$ be an i.i.d. sequence with $\mu = E[X_1]$.

Let $S_0 = 0$ and for $n \geq 1$, let $S_n = X_1 + \ldots + X_n$.

We first recall some more basic limit theorems:

Laws of Large Numbers:
The strong law of large numbers (SLLN) states that

$$P(\lim_{n\to\infty} S_n/n = \mu) = 1,$$

we often say that $S_n/n \to \mu$ almost surely (a.s.).

The weak law of large numbers (WLLN) states that

$$\lim_{n\to\infty} P(|S_n/n - \mu| > \epsilon) = 0$$

for every $\epsilon > 0$.

The LLN says that $S_n \approx \mu n$ for n large.

Two basic questions:

1. Can we refine $S_n - \mu n$?
2. How fast is the convergence to 0 in the WLLN?

Answer to question 1: Central limit theorem. Suppose $E[X_1^2] < \infty$.

$$\frac{S_n - \mu n}{\sigma\sqrt{n}} \xrightarrow{d} U \stackrel{d}{=} N(0,1),$$

as $n \to \infty$.

Naive answer to question 2:

$$P(S_n - \mu n > \epsilon n) \approx P(\sigma U > \epsilon\sqrt{n})$$

Not accurate if X_1 is non-normal (no uniform convergence)!

Rare events: simple estimates

We are interested in the event $\{S_n > an\}$ with $a > \mu$. Assume $P(X_1 > a) > 0$. Observe:

$$P(S_n > an) \geq P(X_i > a, i = 1, ..., n) = P(X_1 > a)^n.$$

Thus, rate of convergence in WLLN is at most exponentially fast.

To get a simple upper bound, observe

$$P(S_n > an) \leq E[e^{sX_1}]^n e^{-san}.$$

Optimizing over s leads to the Chernoff bound

$$P(S_n > an) \leq e^{-n \sup_{s \geq 0}[as - \log E[\exp\{sX_1\}]]},$$

Note: two exponential bounds do not agree with one another.

Note: bound only non-trivial if $a > \mu$ and $E[e^{\epsilon X_1}] < \infty$ for some $\epsilon > 0$.

Some notation:

- Cumulant generating function of X_1: $\Lambda(s) = \log E[e^{sX_1}]$
- Convex conjugate of Λ: $\Lambda^*(a) = \sup_{s \geq 0}[as - \Lambda(s)]$.
- Λ^* is convex, and strictly convex if X_1 is non-deterministic
- Let s^* be such that $\Lambda^*(a) = as^* - \Lambda(s^*)$.

Cramérs theorem states that the Chernoff bound is in some sense sharp:

$$\lim_{n \to \infty} \frac{1}{n} \log P(S_n > an) = -\Lambda^*(a).$$

Proof: need sharp lower bound.

We will continue to assume $P(X > a) > 0$.

Lower bound

Define the subset A_n of $\mathbb{R}^n$ by

$$A_n = \{(x_1, \ldots, x_n) : x_1 + \ldots + x_n > an\}.$$

Observe that

$$P(S_n \geq an) = \int_{A_n} dF(x_1) \ldots dF(x_n),$$

with $F(x) = P(X \leq x)$. Given s^*, we define the so-called tilted distribution $\tilde{F}$ of F as follows:

$$d\tilde{F}(x) = \frac{e^{s^* x}}{E[e^{s^* X}]} dF(x).$$

Let $\tilde{X}_i, i \geq 1$ be an i.i.d. sequence with distribution function $\tilde{F}$ and define $\tilde{S}_n = \tilde{X}_1 + \ldots \tilde{X}_n$.

Lower bound (2)

We see that

$$
\begin{aligned}
& P(S_n > an) \\
= & \int_{A_n} dF(x_1) \dots dF(x_n) \\
= & \int_{A_n} e^{n\Lambda(s^*)} e^{-s^*(x_1+\ldots+x_n)} d\tilde{F}(x_1) \dots d\tilde{F}(x_n) \\
= & \int_{\mathbb{R}^n} I(x_1 + \ldots + x_n > an) e^{n\Lambda(s^*)} e^{-s^*(x_1+\ldots+x_n)} d\tilde{F}(x_1) \dots d\tilde{F}(x_n) \\
= & E[e^{-s^*\tilde{S}_n + n\Lambda(s^*)} I(\tilde{S}_n > an)].
\end{aligned}
$$

A second assumption we make is that $E[e^{(s^*+\epsilon)X}] < \infty$ for some $\epsilon > 0$. Since $E[e^{s\tilde{X}}] = E[e^{(s^*+s)X}]/E[e^{s^*X}]$. We see that $E[e^{\epsilon\tilde{X}}] < \infty$, so that all moments of $\tilde{X}$ are finite.

Verify that $E[\tilde{X}] = a$! Thus, $\tilde{F}$ is designed in such a way that the mean has increased from μ to a, which makes the event of interest more likely.

Lower bound (3)

We see that

$$P(an < \tilde{S}_n < an + \sqrt{n}) \to P(0 < U < 1) > 0, \tag{0.1}$$

with U a normal random variable with zero mean and the same variance as $\tilde{X}$.

We are now ready to derive a sharp enough lower bound.

$$\begin{aligned} P(S_n > an) &= E[e^{-s^*\tilde{S}_n + n\Lambda(s^*)} I(\tilde{S}_n > an)] \\ &\geq E[e^{-s^*\tilde{S}_n + n\Lambda(s^*)} I(an < \tilde{S}_n < an + \sqrt{n})] \\ &\geq e^{-s^*(an+\sqrt{n}) + n\Lambda(s^*)} P(an < \tilde{S}_n < an + \sqrt{n}). \end{aligned}$$

Taking logarithms, dividing by n and letting $n \to \infty$ completes the proof.

Comments

- Large deviations and queueing: Schwarz & Weiss (1995), Ganesh, O'Connell, Wischik (2004).
- General text: Dembo & Zeitouni (1998).
- Extensions in DZ can be found on
  - Cramér for more general sets: $P(S_n/n \in A)$ [generally one needs to tackle topological issues, which I circumvented]
  - No regularity condition on distribution of X_1
  - More general spaces
- "logarithmic asymptotics": $\log P(S_n > an) \sim -n\Lambda^*(a)$:.
- "Exact asymptotics":

$$P(S_n > an) \sim \frac{C_1}{\sqrt{n}} e^{-n\Lambda^*(a)}.$$

(Bahadur-Rao (1976). See also DZ or Asmussen (2003))

The GI/GI/1 queue

Consider a GI/GI/1 FIFO queue with i.i.d. inter-arrival times (A_i), i.i.d. service times (B_i), working at speed 1. $\rho = E[A]/E[B] < 1$.

Let W be the steady-state waiting time, and V the steady-state sojourn time, $V = W + B$.

Well-known result:

$$W \stackrel{d}{=} \sup_{n \geq 0} S_n,$$

with $S_n = \sum_{i=1}^{n} X_i$ and $X_i = B_i - A_i$.

Main question: what is the behavior of

$$P(W > x) = P(\sup_{n \geq 0} S_n > x)$$

as $x \to \infty$?

Assume $E[e^{\epsilon X}] < \infty$ for some $\epsilon > 0$ and $E[X] < 0$.

The following crude bounds turn out to be sharp enough!

$$\sup_n P(S_n > x) \leq P(\sup_n S_n > x) \leq \sum_{n=0}^{\infty} P(S_n > x).$$

Upper bound: Let $u > 0$ be such that $E[e^{uX}] < 1$, and observe that

$$\begin{aligned} \sum_{n=0}^{\infty} P(S_n > x) &\leq \sum_{n=0}^{\infty} E[e^{uS_n}]e^{-ux} \\ &= \frac{1}{1 - E[e^{uX}]} e^{-ux}. \end{aligned}$$

Define $\gamma_F = \sup\{u : E[e^{uX}] \leq 1\}$.
Since the above bound is valid for all $u < \gamma_F$, we see that

$$\limsup_{x\to\infty} \frac{1}{x} \log P(W > x) \leq -\gamma_F.$$

Lower bound

Use $P(S_n > x)$ with n chosen cleverly: $n = \lceil bx \rceil$ with $b = 1/\Lambda'(\gamma_F)$.

Intuitively, this makes sense, since under the exponential tilting with this particular γ_F, we have $E[\tilde{X}_1] = \Lambda'(\gamma_F)$.
Under this new probability distribution, the random walk S_n reaches level x at time $x/\Lambda'(\gamma_F)$.

We see that

$$\liminf_{x\to\infty} \frac{1}{x} \log P(W > x) \geq \liminf_{x\to\infty} \frac{1}{x} \log P(S_{\lceil bx \rceil} > x).$$

The latter liminf can be analyzed by transforming it in a problem of the type we have seen before.

Assume $E[e^{\gamma_F X_1}] = e^{\Lambda(\gamma_F)} = 1$.

Define $n = \lceil bx \rceil$ and observe that

$$P(S_{\lceil bx \rceil} > x) \geq P(S_n > n/b).$$

From Cramér's theorem, we conclude that

$$\begin{aligned}
\liminf_{x\to\infty} \frac{1}{x} \log P(S_{\lceil bx \rceil} > x) &\geq b \liminf_{n\to\infty} \frac{1}{n} \log P(S_n > n/b) \\
&= -b \sup_{s\geq 0} [s/b - \Lambda(s)] \\
&= -\sup_{s\geq 0} [s - \Lambda(s)/\Lambda'(\gamma_F)].
\end{aligned}$$

It can be shown that the optimal value of this optimization problem is γ_F. Since $\Lambda(\gamma_F) = 0$, the corresponding value is γ_F.

We conclude

$$\lim_{x\to\infty} \frac{-\log P(W > x)}{x} = \gamma_F.$$

We call γ_F the decay rate of W.

Example

Consider the $M/M/1$ queue with arrival rate λ and service rate μ. In this case

$$E[e^{sX}] = E[e^{sB}]E[e^{-sA}] = \frac{\mu}{\mu - s}\frac{\lambda}{\lambda + s}$$

It can be shown that $\gamma_F = \mu - \lambda$. The tilted distribution in the lower bound corresponds to switching λ and μ.

Consistent with

$$P(W > x) = \rho e^{-(\mu-\lambda)x}.$$

Comments

- The limit

$$\lim_{x\to\infty} \frac{-\log P(W > x)}{x} = \gamma_F = \sup\{s : E[e^{sX}] \leq 1\}$$

holds also if $E[e^{\gamma_F X_1}] < 1$.

- Important interpretation: rare events under light tails typically occur by a temporary change of the underlying distribution, from F to some exponentially tilted $\tilde{F}$.
- In a queueing context, this causes the drift to change from negative to positive.
- Choosing $\tilde{F}$ typically relates to a minimization problem. In GI/GI/1: trade off between the slope of the new drift, and the duration of the change.
- bx can be interpreted as the time to overflow.
- $P(W > x) \leq P(V > x) \leq E[e^{\gamma_F B}]P(W > x)$, so V and W have the same decay rate.

Precise asymptotics

Assume that $E[Xe^{\gamma_F X}] < \infty$ and set $\tau(x) = \inf\{n : S_n > x\}$, and that X is non-lattice.
As in the lower bound leading to Cramérs theorem, we can show that

$$P(W > x) = e^{-\gamma_F x} E[e^{-\gamma_F (\tilde{S}_{\tau(x)} - x)}],$$

with $\tilde{F}(dx) = e^{\gamma_F x} F(dx)$.

It can be shown, using the key renewal theorem for random walks (McDonald & Ney (1978)), that the expected value converges to a limit C_F.

Thus, we have exact asymptotics

$$P(W > x) \sim C_F e^{-\gamma_F x}.$$

Heavy tails

The results obtained so far are not very meaningful if

$$E[e^{\epsilon X}] = \infty$$

for all $\epsilon > 0$.

In this case, we say that X has a heavy (right) tail.

Examples of heavy tails:

- Pareto: $P(X > x) \sim x^{-\alpha}$
- Lognormal: $P(X > x) \sim e^{-(\log x)^2}$
- Weibull: $P(X > x) \sim e^{-x^{\alpha}}$, $\alpha \in (0, 1)$.
- Any df with a hazard rate decreasing to 0.

Classes of HT distributions

X is called long-tailed (and write $X \in \mathcal{L}$) if

$$\frac{P(X > x + y)}{P(X > x)} \to 1$$

as $x \to \infty$ for every $y > 0$.

An equivalent way to write this is

$$P(X > x + y \mid X > x) \to 1.$$

A non-negative random variable X is subexponential if for two independent copies X_1, X_2 of X,

$$\frac{P(X_1 + X_2 > x)}{P(X_1 > x)} \to 2,$$

as $x \to \infty$. We write $X \in \mathcal{S}$. Note that

$$P(\max\{X_1, X_2\} > x) \sim 2P(X_1 > x),$$

and that $X_1 + X_2 > \max\{X_1, X_2\}$. Thus, X is subexponential, if the inequality in

$$P(X_1 + X_2 > x) \geq P(\max\{X_1, X_2\} > x)$$

can be replace by "$\sim$", and

$$P(X_1 + X_2 > x; \max\{X_1, X_2\} < x) = o(P(X_1 > x)).$$

Thus, a large sum is most likely due to a large maximum.

Properties

- $X \in \mathcal{L}$ and $Y \geq 0$ then $P(X > x + Y) \sim P(X > x)$.
- If $X \in \mathcal{S}$, then $\bar{F}^{n*}(x)/\bar{F}(x) \to n$.
- If $X \in \mathcal{S}$ and $P(Y > x) = o(P(X > x))$ and Y is independent of X, then $P(X + Y > x) \sim P(X > x)$.
- If $X \in \mathcal{S}$ then $X \in \mathcal{L}$. This may not hold if X can be negative!
- Kesten: If X is subexponential, then its distribution function F satisfies the following: for every $\epsilon > 0$ there exists $K < \infty$ such that for every $n \geq 2$ and $x \geq 0$:

$$\frac{\bar{F}^{n*}(x)}{\bar{F}(x)} \leq K(1+\epsilon)^n$$

Random sums

Let N be a random variable independent of the $i.i.d.$ sequence $X_1, X_2, \ldots$. Define the random sum $Z = X_1 + \ldots + X_N$. What can we say about the tail behavior of Z if $X_1 \in \mathcal{S}$?

Under the above conditions, if N is also such that $E[(1+\epsilon)^N] < \infty$ for some $\epsilon > 0$, then

$$\lim_{x\to\infty} \frac{P(Z > x)}{P(X_1 > x)} = E[N].$$

Proof: Write $p_n = P(N = n)$ and observe that

$$\frac{P(Z > x)}{P(X_1 > x)} = \sum_{n=0}^{\infty} p_n \frac{\bar{F}^{n*}(x)}{\bar{F}(x)}.$$

Interchange of limit and sum is allowed: combine "Kesten" and dominated convergence.
Result is not true if N has heavier tail than X (tomorrow)!

Application to FIFO waiting time

Let B^* be a r.v. with density $P(B > x)/E[B]$.

Theorem (Korshunov (1997)). The following are equivalent:

1. $W \in \mathcal{S}$,
2. $B^* \in \mathcal{S}$,
3. $P(W > x) \sim \frac{\rho}{1-\rho} P(B^* > x)$.

Proof: Using Wiener-Hopf factorization for W. For $M/G/1$:

$$W \stackrel{d}{=} \sum_{i=1}^{N} B_i^*,$$

with $P(N = n) = (1 - \rho)\rho^n$ so $E[N] = \frac{\rho}{1-\rho}$.

It can be shown that $P(B > x) = o(P(B^* > x))$, so

$$P(V > x) = P(W + B > x) \sim P(W > x).$$

Example: Regular variation

Part of the previous theorem have been shown by Borovkov (1971), Cohen (1973), Pakes (1975), Veraverbeke (1977). Extensions to non-iid input processes exist as well.

An important special case is

$$P(B > x) = L(x)x^{-\alpha}, \alpha > 1,$$

with L slowly varying, i.e.

$$L(ax)/L(x) \to 1$$

for any $a > 0$. In this case

$$P(B^* > x) = \frac{1}{E[B]} \int_x^\infty P(B > u)du \sim \frac{1}{E[B](\alpha - 1)} L(x)x^{1-\alpha}.$$

$$P(W > x) \sim \frac{\rho}{1-\rho} \frac{1}{E[B](\alpha - 1)} L(x)x^{1-\alpha}.$$

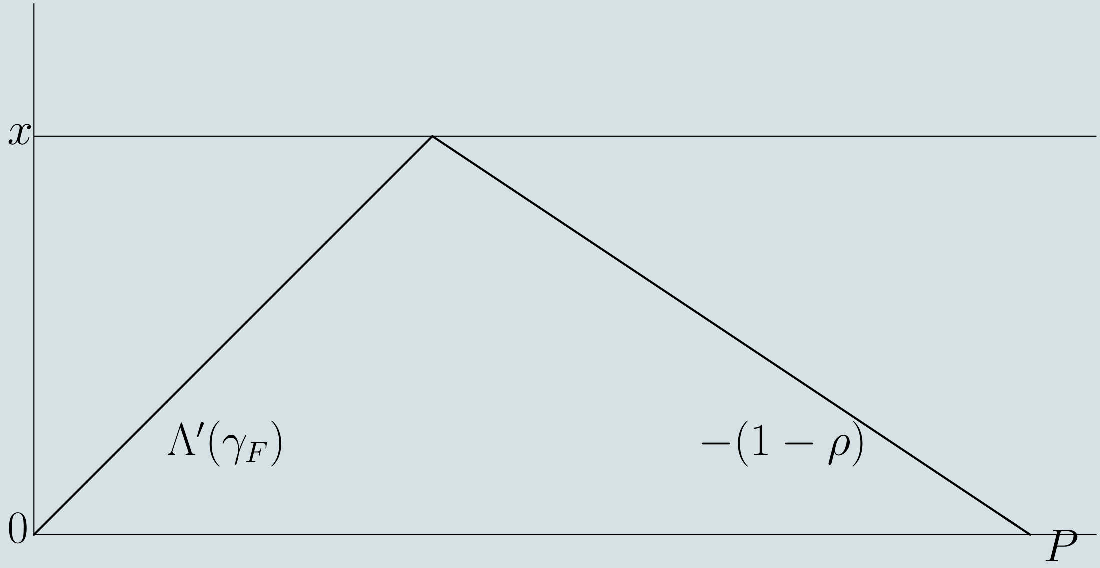
The principle of a single big jump

Most likely way to obtain the event $W > x$ (Zwart01, Baccelli+Foss02, Zachary04):

- At some time n, the random walk S_n has the typical value $-an$, $a = -E[X]$.
- $X_{n+1} = B_{n+1} - A_{n+1}$ is so large that $S_{n+1} > x$. For this to happen, we need $X_n > an + x$.
- This can happen at any time n.

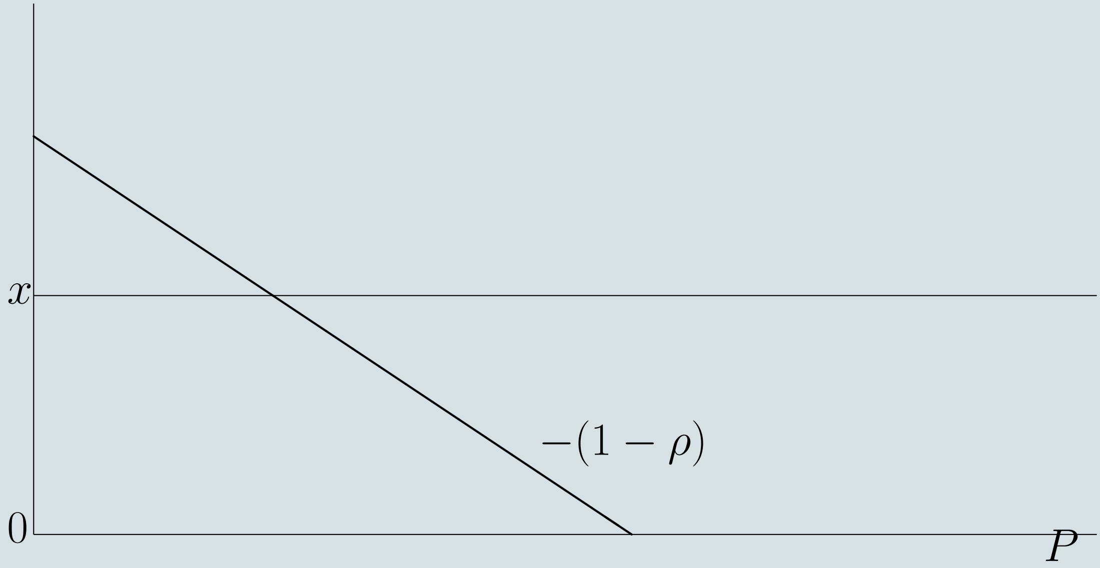
$$
\begin{aligned}
P(W > x) &\approx P(\cup_{n=1}^{\infty}\{S_n \approx -an; X_{n+1} > an + x\}) \\
&\approx \sum_{n=0}^{\infty} P(X_{n+1} > an + x) = \sum_{n=0}^{\infty} \bar{F}(an + x) \\
&\sim \frac{1}{a}\int_x^{\infty} \bar{F}(u)du \\
&\sim \frac{1}{a}\int_x^{\infty} \bar{P}(B > u)du = \frac{\rho}{1-\rho} P(B^* > x).
\end{aligned}
$$

Summary: The light-tailed case



- In beginning of busy period: Sample from exponentially(γ_F) tilted distribution until level x is crossed.
- Maximum in busy cycle: $x + O(1)$

Summary: The heavy-tailed case



- In beginning of busy period (after $O(1)$ time): Huge job arrives
- Maximum in busy cycle: $x + O(x)$ (in case of regular variation) $x + O(a(x))$ (in general, $a(x)$ is "auxiliary function" from EVT)

Optimality (1)

For light tails, the main factor $e^{-\gamma_F x}$ cannot be improved!

For any scheduling discipline π:

$$\begin{aligned} P(V_\pi > x) &= \frac{1}{E[N]} E\left[\sum_{i=1}^{N} I(V_{\pi,i} > x)\right] \\ &\geq \frac{1}{E[N]} E\left[\sum_{i=1}^{N} I(V_{\pi,i} > x) I(C_{max} > x)\right] \\ &\geq \frac{1}{E[N]} P(C_{max} > x). \end{aligned}$$

Iglehart (1972): $P(C_{max} > x) \sim K e^{-\gamma_F x}$.

More generally: $-\log P(C_{max} > x) \sim \gamma_F x$.

Optimality of FIFO w.r.t. decay rates is shown by Ramanan & Stolyar (2001) for single nodes and by Stolyar (2003) for networks.

Optimality (2)

If service time tails are regularly varying, we see

$$P(V > x) = O(xP(B > x)).$$

Intuition: $O(x)$ jobs in a busy period get stuck behind single large job.

Consistent with well known result: $E[V] < \infty \Leftrightarrow E[B^2] < \infty$.

Can we do better?

$$E[V_{PS}] = \frac{E[B]}{1-\rho}, \qquad E[V_{SRPT}] \leq E[V_{PS}].$$

Anantharam (1999): Any nonpreemptive scheduling discipline is not optimal in case of pareto job sizes.

To be continued...

Tomorrow 13.30