Characterizing Policies \textit{with} Optimal Response Time Tails \textit{under Heavy-Tailed Job Sizes}

Ziv Scully \hspace{1cm} CMU
Lucas van Kreveld \hspace{1cm} UvA
Onno Boxma \hspace{1cm} TU/e
Jan-Pieter Dorsman \hspace{1cm} UvA
Adam Wierman \hspace{1cm} Caltech
M/G/1 Queue
M/G/1 Queue
M/G/1 Queue
M/G/1 Queue

queue

server

job

size

(unknown)
M/G/1 Queue

queue

server

job

size (unknown)
M/G/1 Queue
M/G/1 Queue

queue

server

job

size

(unknown)
M/G/1 Queue

random arrivals

queue

server

job

remaining size

size (unknown)

age
M/G/1 Queue

random arrivals

queue

server

job

size

remaining size

age

(unknown)
M/G/1 Queue

random arrivals

\[ X = \text{job size r.v.} \]
\[ \lambda = \text{arrival rate} \]
\[ \rho = \lambda E[X] < 1 \]

size (unknown)

remaining size

age

server

job

queue
M/G/1 Queue

$X = \text{job size r.v.}$
$\lambda = \text{arrival rate}$
$\rho = \lambda E[X] < 1$

Scheduling policy:
- picks which job to serve

random arrivals

queue

server

job

size (unknown)

remaining size

age

X = \text{job size r.v.} \\
\lambda = \text{arrival rate} \\
\rho = \lambda E[X] < 1
M/G/1 Queue

- $X = \text{job size r.v.}$
- $\lambda = \text{arrival rate}$
- $\rho = \lambda E[X] < 1$

**Random arrivals**

**Queue**

- **Size** (unknown)
- **Age**
- **Remaining size**

**Server**

**Scheduling policy:** picks which job to serve
M/G/1 Queue

\[ X = \text{job size r.v.} \]
\[ \lambda = \text{arrival rate} \]
\[ \rho = \frac{\lambda \mathbb{E}[X]}{1} < 1 \]

random arrivals

queue

server

Scheduling policy:
- picks which job to serve

remaining size

size (unknown)

age

job
M/G/1 Queue

\[ X = \text{job size r.v.} \]
\[ \lambda = \text{arrival rate} \]
\[ \rho = \lambda \mathbb{E}[X] < 1 \]

Scheduling policy: picks which job to serve

random arrivals

\begin{align*}
\text{size} & \quad \{ \text{remaining size} \} \\
\text{(unknown)} & \quad \{ \text{age} \}
\end{align*}
$X = \text{job size r.v.}$
$\lambda = \text{arrival rate}$
$\rho = \lambda E[X] < 1$

**M/G/1 Queue**

- **Random arrivals**
- **Queue**
- **Server**
- **Job**
- **Remaining size**
- **Size** (unknown)
- **Age**

**Scheduling policy:**
- Picks which job to serve
Response Time
Response Time
Response Time

= T = response time
Response Time

= $T = response\ time$
Response Time

Goal: schedule to minimize two metrics
Response Time

Goal: schedule to minimize two metrics
- mean response time $\mathbb{E}[T]$
Response Time

\[ T = \text{response time} \]

**Goal:** schedule to minimize two metrics

- *mean* response time \( E[T] \)
- *tail* of response time \( P[T > t] \)
Response Time

\[ \begin{align*}
T &= T = \text{response time} \\
\text{Goal: schedule to minimize two metrics} \\
\text{• mean response time } \mathbb{E}[T] \\
\text{• tail of response time } \mathbb{P}[T > t] & \xrightarrow{t \to \infty} \text{ limit}
\end{align*} \]
Response Time

\[ T = \text{response time} \]

\[ \text{Goal: schedule to minimize two metrics} \]

- mean response time \( \mathbb{E}[T] \)
- tail of response time \( \mathbb{P}[T > t] \) \( t \to \infty \) limit

\[ \text{Setting: heavy-tailed job size } X \]
Response Time

Goal: schedule to minimize two metrics

• mean response time $E[T]$
• tail of response time $P[T > t]$ $t \rightarrow \infty$ limit

Setting: heavy-tailed job size $X$
Scheduling with Heavy Tails
Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $\mathbb{E}[T]$</th>
<th>Tail $\mathbb{P}[T &gt; t]$</th>
</tr>
</thead>
</table>

$t \rightarrow \infty$ limit
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Come, First Served (FCFS)</td>
<td>$t \rightarrow \infty$ limit</td>
<td></td>
</tr>
</tbody>
</table>
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>$t \rightarrow \infty$ limit</td>
</tr>
</tbody>
</table>

First Come, First Served
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
</tbody>
</table>

$t \rightarrow \infty$ limit

First Come, First Served
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
</tbody>
</table>

For $t \to \infty$ limit:

$$P[T > t] = \Theta(t) \cdot P[X > t]$$
<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
</tbody>
</table>

Processor Sharing

$t \to \infty$ limit
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td></td>
</tr>
</tbody>
</table>

$t \to \infty$ limit

Processor Sharing

Policy

Mean $E[T]$

Tail $P[T > t]$
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
</tbody>
</table>

$t \rightarrow \infty$ limit

**Processor Sharing**
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
</tbody>
</table>

$t \to \infty$ limit

$P[T > t] = \Theta(1) \cdot P[X > t]$
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
</tbody>
</table>

$t \rightarrow \infty$ limit

**Foreground-Background** FB
Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
</tbody>
</table>

$\lim_{t \to \infty}$

Foreground-
Background

serves job
of least age
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td></td>
</tr>
</tbody>
</table>

**Foreground-Background**

- Serves job of least *age*

$t \rightarrow \infty$ limit
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
</tbody>
</table>

- Foreground-Background
  - serves job of least age

$t \rightarrow \infty$ limit
Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
</tbody>
</table>

$P[T > t] = \Theta(1) \cdot P[X > t]$

$t \to \infty$ limit

Foreground-Background

serves job of least age
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
</tbody>
</table>

Prioritize by Gittins rank

$t \to \infty$ limit
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>$t \to \infty$ limit</td>
</tr>
</tbody>
</table>

Prioritize by Gittins rank
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
</tbody>
</table>

$t \to \infty$ limit

Prioritize by Gittins rank
Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$t \to \infty$ limit

Monotonic Shortest Expected Remaining Processing Time
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td></td>
</tr>
</tbody>
</table>

$t \to \infty$ limit

---

*Monotonic Shortest Expected Remaining Processing Time*
Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>???</td>
</tr>
</tbody>
</table>

$t \rightarrow \infty$ limit

Monotonic Shortest Expected Remaining Processing Time
## Scheduling with Heavy Tails

$t \rightarrow \infty$ limit

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>???</td>
</tr>
<tr>
<td>RMLF</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Randomized Multi-Level Feedback
### Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>???</td>
</tr>
<tr>
<td>RMLF</td>
<td>best (X unknown)</td>
<td>???</td>
</tr>
</tbody>
</table>

$t \to \infty$ limit
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $\mathbb{E}[T]$</th>
<th>Tail $\mathbb{P}[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>???</td>
</tr>
<tr>
<td>RMLF</td>
<td>best $(X$ unknown)</td>
<td>???</td>
</tr>
</tbody>
</table>

$t \to \infty$ limit

Randomized Multi-Level Feedback
# Scheduling with Heavy Tails

## Table of Performance

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>???</td>
</tr>
<tr>
<td>RMLF</td>
<td>best (X unknown)</td>
<td>???</td>
</tr>
</tbody>
</table>
Question:
can we optimize both mean and tail of response time?
## Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>???</td>
</tr>
<tr>
<td>RMLF</td>
<td>best (X unknown)</td>
<td>???</td>
</tr>
</tbody>
</table>
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>???</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>???</td>
</tr>
<tr>
<td>RMLF</td>
<td>best (X unknown)</td>
<td>???</td>
</tr>
</tbody>
</table>
# Scheduling with Heavy Tails

<table>
<thead>
<tr>
<th>Policy</th>
<th>Mean $E[T]$</th>
<th>Tail $P[T &gt; t]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCFS</td>
<td>bad</td>
<td>worst</td>
</tr>
<tr>
<td>PS</td>
<td>okay</td>
<td>best</td>
</tr>
<tr>
<td>FB</td>
<td>good</td>
<td>best</td>
</tr>
<tr>
<td>Gittins</td>
<td>best</td>
<td>best*</td>
</tr>
<tr>
<td>M-SERPT</td>
<td>5-approx.</td>
<td>best</td>
</tr>
<tr>
<td>RMLF</td>
<td>best (X unknown)</td>
<td>best</td>
</tr>
</tbody>
</table>

$t \to \infty \text{ limit}$

$new!$
Our contribution: a sufficient condition for \textit{optimal} response time tail.
Our contribution: a sufficient condition for optimal response time tail

Gittins, M-SERPT, RMLF, and more…
Our contribution: a sufficient condition for optimal response time tail

Gittins, M-SERPT, RMLF, and more…

… all asymptotically optimize $\mathbb{P}[T > t]$
Our contribution: a sufficient condition for optimal response time tail

Gittins, M-SERPT, RMLF, and more…

... all asymptotically optimize $P[T > t]$
Part 1: formally state results

Part 2: sketch proof techniques
Part 1: formally state results
Part 1: formally state results

easy version of
Our contribution:

a sufficient condition for **optimal** response time tail
Our contribution:
a sufficient condition for \textbf{optimal} response time tail

Question: What does a sufficient condition look like?
Our contribution:
a sufficient condition for **optimal** response time tail

**Question**: What does a sufficient condition look like?

- “Don’t let small jobs get stuck behind large jobs”
Our contribution: a sufficient condition for **optimal** response time tail

**Question**: What does a sufficient condition look like?

- “Don’t let small jobs get stuck behind large jobs”
- How to formalize?
Scheduling policy:
picks which job to serve
Describing Policies with SOAP

Scheduling policy: picks which job to serve

SOAP scheduling policy: picks which job to serve using a *rank* function
Describing Policies with SOAP

Scheduling policy:
picks which job to serve

SOAP scheduling policy:
picks which job to serve using a rank function

\[ r : \text{age} \rightarrow \text{rank} \]
Describing Policies with **SOAP**

Scheduling policy:
- picks which job to serve

**SOAP** scheduling policy:
- picks which job to serve using a *rank* function

\[
\begin{align*}
    r : \text{age} &\rightarrow \text{rank} \\
    \{ \text{age} \} &\rightarrow 
\end{align*}
\]
Describing Policies with **SOAP**

**Scheduling policy:**
- picks which job to serve

**SOAP scheduling policy:**
- picks which job to serve using a *rank* function

\[
 r : \text{age} \rightarrow \text{rank}
\]

*a job’s priority* (lower is better)
Describing Policies with SOAP

Scheduling policy:
picks which job to serve

**SOAP scheduling policy:**
picks which job to serve using a \textit{rank} function

\[ r : \text{age} \rightarrow \text{rank} \]

\text{a job’s priority} (lower is better)
Describing Policies with **SOAP**

**Scheduling policy:**
picks which job to serve

**SOAP scheduling policy:**
picks which job to serve using a *rank function*

\[ r : \text{age} \rightarrow \text{rank} \]

*a job’s priority (lower is better)*

serves job of least age
Describing Policies with **SOAP**

**Scheduling policy:**
picks which job to serve

**SOAP scheduling policy:**
picks which job to serve using a *rank function*

\[ r : \text{age} \rightarrow \text{rank} \]

*a job’s priority (lower is better)*

\[ r(a) = a \]

serves job of least age

lower is better

age

rank
Wide Range of SOAP Policies

One rule of SOAP:
always serve job of minimum rank
(break ties FCFS)
Wide Range of **SOAP** Policies

One rule of **SOAP**: always serve job of *minimum rank* (break ties FCFS)

$$r(a) = a$$

![Diagram showing rank vs. age with a line representing the relationship $$r(a) = a$$]
Wide Range of SOAP Policies

One rule of SOAP:
always serve job of minimum rank
(break ties FCFS)

FCFS

FB
\[ r(a) = a \]
Wide Range of **SOAP** Policies

One rule of **SOAP**: always serve job of *minimum rank* (break ties FCFS)

**FCFS**

\[ r(a) = 1 \]

**FB**

\[ r(a) = a \]
Wide Range of **SOAP** Policies

One rule of **SOAP**: always serve job of *minimum rank* (break ties FCFS)

### FCFS

- **r(a) = 1**
- *worst tail*

### FB

- **r(a) = a**
- *best tail*
Wide Range of **SOAP** Policies

One rule of **SOAP**:
always serve job of *minimum rank*
(break ties FCFS)

- **FCFS**
  \[ r(a) = 1 \]
  ![Graph showing FCFS with worst tail](image)

- **RMLF**
  ![Graph showing RMLF](image)

- **FB**
  \[ r(a) = a \]
  ![Graph showing FB with best tail](image)
Wide Range of **SOAP** Policies

One rule of **SOAP**: always serve job of *minimum rank* (break ties FCFS)

- **FCFS**
  \[ r(a) = 1 \]
  \[ \text{worst tail} \]

- **RMLF**
  \[ r(a) = 2^{|\log_2 a|} \]

- **FB**
  \[ r(a) = a \]
  \[ \text{best tail} \]
Wide Range of **SOAP** Policies

One rule of **SOAP**: always serve job of *minimum rank* (break ties FCFS)

- **FCFS**
  
  \[ r(a) = 1 \]

- **RMLF**
  
  \[ r(a) = 2^{\lfloor \log_2[a] \rfloor} \]

- **FB**
  
  \[ r(a) = a \]

... with some randomization

**worst** tail

**best** tail
Wide Range of SOAP Policies

One rule of SOAP:
always serve job of minimum rank
(break ties FCFS)

\[ r(a) = 1 \]  
**FCFS**

\[ r(a) = 2^{\left\lfloor \log_2 a \right\rfloor} \]  
**RMLF**

\[ r(a) = a \]  
**FB**

**worst** tail

???

**best** tail
Our contribution:
a sufficient condition for *optimal* response time tail
Our contribution: a sufficient condition for **optimal** response time tail
Our contribution:
a sufficient condition for **optimal** response time tail

For **SOAP** policies: want a condition on the *rank* function
Sufficient Condition

rank

age
Sufficient Condition
Sufficient Condition

my worst ever rank

my size

rank

age
Sufficient Condition

my worst ever rank

my size

big jobs

age
Sufficient Condition

my worst ever rank

Big jobs get in my way!
Sufficient Condition

my size
big jobs
my worst ever rank

Big jobs get in my way!
Sufficient Condition

Big jobs get in my way!
Sufficient Condition

Big jobs get in my way!
Sufficient Condition

Big jobs get in my way!

No big jobs bothering me
Sufficient Condition

No big jobs bothering me
Suppose for some $\delta \geq \gamma > 0$:

$$\Omega(a^\gamma) \leq r(a) \leq O(a^\delta)$$
Sufficient Condition

Suppose for some $\delta \geq \gamma > 0$:

$$\Omega(a^\gamma) \leq r(a) \leq O(a^{\delta})$$
Sufficient Condition

Suppose for some $\delta \geq \gamma > 0$:

$$\Omega(a^\gamma) \leq r(a) \leq O(a^\delta)$$

Want $\gamma$ and $\delta$ to be close

No big jobs bothering me.
Main theorem: Consider an M/G/1 queue whose job size distribution $X$ is *intermediate regularly varying* and satisfies

$$\Omega(k^{-\beta}) \leq \frac{P[X > kx]}{P[X > x]} \leq O(k^{-\alpha}), \quad (\beta \geq \alpha > 1, \alpha \to \infty)$$
**Main theorem:** Consider an M/G/1 queue whose job size distribution $X$ is *intermediate regularly varying* and satisfies

$$
\Omega(k^{-\beta}) \leq \frac{P[X > kx]}{P[X > x]} \leq O(k^{-\alpha}), \quad (\beta \geq \alpha > 1, a \to \infty)
$$

and suppose a **SOAP** policy with **rank** function $r$ satisfies

$$
\Omega(a^{\gamma}) \leq r(a) \leq O(a^{\delta}), \quad (\delta \geq \gamma > 0, a \to \infty)
$$
Main theorem: Consider an M/G/1 queue whose job size distribution $X$ is *intermediate regularly varying* and satisfies

$$
\Omega(k^{-\beta}) \leq \frac{\mathbb{P}[X > kx]}{\mathbb{P}[X > x]} \leq O(k^{-\alpha}), \quad (\beta \geq \alpha > 1, a \to \infty)
$$

and suppose a SOAP policy with rank function $r$ satisfies

$$
\Omega(a^{\gamma}) \leq r(a) \leq O(a^{\delta}), \quad (\delta \geq \gamma > 0, a \to \infty)
$$
Main theorem: Consider an M/G/1 queue whose job size distribution \( X \) is \textit{intermediate regularly varying} and satisfies

\[
\Omega(k^{-\beta}) \leq \frac{\mathbb{P}[X > kx]}{\mathbb{P}[X > x]} \leq O(k^{-\alpha}), \quad (\beta \geq \alpha > 1, a \to \infty)
\]

and suppose a \textbf{SOAP} policy with \textbf{rank} function \( r \) satisfies

\[
\Omega(a^{\gamma}) \leq r(a) \leq O(a^{\delta}), \quad (\delta \geq \gamma > 0, a \to \infty)
\]

Then if

\[
\frac{\delta}{\gamma} < \frac{\alpha - 1}{2\beta} + \sqrt{1 + \left( \frac{\alpha - 1}{2\beta} \right)^2},
\]
Main theorem: Consider an M/G/1 queue whose job size distribution $X$ is *intermediate regularly varying* and satisfies

$$\Omega(k^{-\beta}) \leq \frac{\mathbb{P}[X > kx]}{\mathbb{P}[X > x]} \leq O(k^{-\alpha}), \quad (\beta \geq \alpha > 1, \alpha \to \infty)$$

and suppose a SOAP policy with rank function $r$ satisfies

$$\Omega(a^\gamma) \leq r(a) \leq O(a^\delta), \quad (\delta \geq \gamma > 0, \delta \to \infty)$$

Then if

$$\frac{\delta}{\gamma} < \frac{\alpha - 1}{2\beta} + \sqrt{1 + \left(\frac{\alpha - 1}{2\beta}\right)^2}$$

the SOAP policy is *tail-optimal* for $X$, meaning

$$\mathbb{P}\left[T > \frac{x}{1-\rho}\right] \sim \mathbb{P}[X > x]. \quad (x \to \infty)$$
Main theorem: Consider an $M/G/1$ queue whose job size distribution $X$ is intermediate regularly varying and satisfies

$$\Omega(k^{-\beta}) \leq \frac{P[X > kx]}{P[X > x]} \leq O(k^{-\alpha}), \quad (\beta \geq \alpha > 1, a \to \infty)$$

and suppose a SOAP policy with rank function $r$ satisfies

$$\Omega(a^{\gamma}) \leq r(a) \leq O(a^{\delta}), \quad (\delta \geq \gamma > 0, a \to \infty)$$

Then if

$$\gamma = \delta \text{ suffices}$$

$$\frac{\delta}{\gamma} < \frac{\alpha - 1}{2\beta} + \sqrt{1 + \left(\frac{\alpha - 1}{2\beta}\right)^2},$$

the SOAP policy is tail-optimal for $X$, meaning

$$P\left[T > \frac{x}{1-\rho}\right] \sim P[X > x]. \quad (x \to \infty)$$
Applying the Condition

**FCFS**
\[ r(a) = 1 \]

**RMLF**
\[ r(a) = 2^{|\log_2 a|} \]

**FB**
\[ r(a) = a \]

-worst tail

-??

-best tail
Applying the Condition

**FCFS**

\[ r(a) = 1 \]

**RMLF**

\[ r(a) = 2^{\lfloor \log_2 a \rfloor} \]

**FB**

\[ r(a) = a \]

worst tail

???

best tail
Applying the Condition

**FCFS**
\[ r(a) = 1 \]

**RMLF**
\[ r(a) = 2^{[\log_2(a)]} \]

**FB**
\[ r(a) = a \]

worst tail

best tail
Applying the Condition

FCFS
\[ r(a) = 1 \]

RMLF
\[ r(a) = 2^{\lfloor \log_2[a] \rfloor} \]

FB
\[ r(a) = a \]

worst tail

best tail
Applying the Condition

**FCFS**
\[ r(a) = 1 \]

**RMLF**
\[ r(a) = 2^{[\log_2 a]} \]

**FB**
\[ r(a) = a \]

**worst** tail

**best** tail
Applying the Condition

**FCFS**
\[ r(a) = 1 \]
worst tail

**RMLF**
\[ r(a) = 2^{\lfloor \log_2 a \rfloor} \]
\[ \gamma = \delta = 1 \]

**FB**
\[ r(a) = a \]
\[ \gamma = \delta = 1 \]
best tail
Applying the Condition

**FCFS**

\[ r(a) = 1 \]

- **worst tail**

**RMLF**

\[ r(a) = 2^{\lfloor \log_2 a \rfloor} \]

- **best tail**

**FB**

\[ r(a) = a \]

- **best tail**

\[ \gamma = \delta = 1 \]
M-SERPT

\[ r(a) = \max_{0 \leq b \leq a} \mathbb{E}[X - b \mid X > b] \]
M-SERPT

\[ r(a) = \max_{0 \leq b \leq a} E[X - b \mid X > b] \]
M-SERPT

\[ r(a) = \max_{0 \leq b \leq a} \mathbb{E}[X - b \mid X > b] \]
M-SERPT

\[ r(a) = \max_{0 \leq b \leq a} \mathbb{E}[X - b \mid X > b] \]
M-SERPT

\[ r(a) = \max_{0 \leq b \leq a} \mathbb{E}[X - b \mid X > b] \]

\( r(a) = \Theta(a) \)

\[ \gamma = \delta = 1 \Rightarrow \text{M-SERPT is tail-optimal} \]
\[ \inf_{b \geq a} \frac{1}{h_X(b)} \leq r_{\text{Gittins}}(a) \leq r_{\text{M-SERPT}}(a) \]
Gittins

\[ \inf_{b \geq a} h_X(b) \leq r_{\text{Gittins}}(a) \leq r_{\text{M-SERPT}}(a) \]
Gittins

\[
\inf_{b \geq a} h_X(b) \leq r_{\text{Gittins}}(a) \leq r_{\text{M-SERPT}}(a)
\]

hazard rate of \( X \)
Gittins

**Theorem:** Gittins is *tail-optimal* if $X$’s hazard rate obeys

$$h_X(a) = O(a^{-\gamma})$$

for some

$$\gamma > \sqrt{1 + \left(\frac{\alpha - 1}{2\beta}\right)^2} - \frac{\alpha - 1}{2\beta}.$$
Theorem: Gittins is tail-optimal if $X$’s hazard rate obeys

$$h_X(a) = O(a^{-\gamma})$$

for some $\gamma = 1$ suffices

$$\gamma > \sqrt{1 + \left(\frac{\alpha - 1}{2\beta}\right)^2} - \frac{\alpha - 1}{2\beta}.$$
Part 2:
sketch proof techniques

$E[B^p]$
Proof Outline
Proof Outline

tail-optimal
Proof Outline

\[ \mathbb{E}[T(x)^p] \text{ small} \]

Núñez-Queija’s method (2002)

\text{tail-optimal}
Proof Outline

Núñez-Queija’s method (2002)

\[ \mathbb{E}[T(x)^p] \text{ small} \]

response time of job of size \( x \)

tail-optimal
Proof Outline

$E[B^p]$ small

SOAP (2018)

$E[T(x)^p]$ small

Núñez-Queija’s method (2002)

tail-optimal
Proof Outline

\[ \mathbb{E}[B^p] \text{ small} \]

SOAP (2018)

\[ \mathbb{E}[T(x)^p] \text{ small} \]

Núñez-Queija's method (2002)

tail-optimal

M/G/1 busy period

\[ B \]
Proof Outline

\[ \mathbb{E}[B^p] \text{ small} \]

\[ \mathbb{E}[T(x)^p] \text{ small} \]

Núñez-Queija’s method (2002)

\[ \text{tail-optimal} \]

uses **rank** function

SOAP (2018)

M/G/1 busy period

\[ B \]
Proof Outline

- **E**[^B^p^] small
- **E**[^T(x)^p^] small
- **M/G/1** busy period

- SOAP (2018)
- Núñez-Queija’s method (2002)
- Uses rank function

New bound on fractional moments
Proof Outline

- E\[B^p\] small
- E\[(T(x))^p\] small
- Núñez-Queija’s method (2002)
- tail-optimal
- SOAP (2018)
- new version
- M/G/1 busy period
- new bound on fractional moments

uses rank function
Proof Outline

\[ \mathbb{E}[B^p] \text{ small} \]

\[ \mathbb{E}[T(x)^p] \text{ small} \]

Núñez-Queija’s method (2002) uses rank function

\[ \mathbb{E}[T(x)^p] \text{ small} \]

\[ M/G/1 \text{ busy period} \]

new version

new bound on fractional moments

tail-optimal

uses rank function

SOAP (2018)
Summary

Result: sufficient condition for **tail-optimality**
Summary

**Result:** sufficient condition for **tail-optimality**

**Key idea #1:** condition stated using *rank* function of **SOAP** policy
Summary

**Result:** sufficient condition for **tail-optimality**

**Key idea #1:** condition stated using *rank* function of SOAP policy

**Key idea #2:** new bound on *fractional moments* of M/G/1 *busy periods*
Summary

Result: sufficient condition for **tail-optimality**

Key idea #1: condition stated using *rank* function of SOAP policy

Key idea #2: new bound on *fractional moments* of M/G/1 busy periods

Get in touch: zscully@cs.cmu.edu
Bonus Slides
Prior Sufficient Conditions
Prior Sufficient Conditions

Núñez-Queija (2002): policy is tail-optimal if moments of $T(x)$ are small.
Prior Sufficient Conditions

Núñez-Queija (2002): policy is *tail-optimal* if moments of $T(x)$ are small
Prior Sufficient Conditions

Núñez-Queija (2002): policy is tail-optimal if moments of $T(x)$ are small

- Hard to verify!
Prior Sufficient Conditions

Núñez-Queija (2002): policy is tail-optimal if moments of $T(x)$ are small
  • Hard to verify!

NWZ (2008): SMART policies are tail-optimal
Prior Sufficient Conditions

Núñez-Queija (2002): policy is \textit{tail-optimal} if moments of $T(x)$ are small
• Hard to verify!

NWZ (2008): \textit{SMART} policies are \textit{tail-optimal}
Prior Sufficient Conditions

Núñez-Queija (2002): policy is tail-optimal if moments of $T(x)$ are small

- Hard to verify!

NWZ (2008): SMART policies are tail-optimal

- Easy to verify...
Prior Sufficient Conditions

Núñez-Queija (2002): policy is tail-optimal if moments of $T(x)$ are small
• Hard to verify!

NWZ (2008): SMART policies are tail-optimal
• Easy to verify…
• … but only applies with known job sizes
Prior Sufficient Conditions

Núñez-Queija (2002): policy is \textit{tail-optimal} if moments of $T(x)$ are small

\begin{itemize}
  \item Hard to verify!
\end{itemize}

NWZ (2008): \textit{SMART} policies are \textit{tail-optimal}

\begin{itemize}
  \item Easy to verify…
  \item … but only applies with known job sizes
\end{itemize}

Wanted:
\textit{easy-to-verify} condition for systems with \textit{unknown} job sizes