Surprising results on task assignment for high-variability workloads

Mor Harchol-Balter, CMU, Comp. Sci.
Alan Scheller-Wolf, CMU, Tepper Business
Andrew Young, Morgan Stanley
Q: What is a good Assignment Policy?

(high-variability jobs)

Server farm model

Goal: Minimize mean response time: $E[T]$

Assignment Policy

$n$ servers

general i.i.d.
job sizes $\sim X$

$C^2 = \frac{\text{var}(X)}{E[X]^2}$

$\rho = \lambda E[X] \leq n$

POLICY MATTERS!
Good Answers

LWL (Least Work Left)
- High throughput
- Splits job to host with least remaining work.

M/G/2

SITA (Split Interval)
- Protects against high variability
- Isolation for smalls

+ Preemptive splitting on size
Prior Work on SITA

SITA in Practice
- **Supercomputing Centers**
  [Hotovy, Schneider, O'Donnell 96]
  [Schroeder, Harchol-Balter 00]
- **Manufacturing Centers**
  [Buzacott, Shanthikumar 93]
- **File Server Farms**
  [Cardellini, Colajanni, Yu 01]
- **Supermarkets**

SITA variants
- [Harchol-Balter 00]
- [Harchol-Balter 02]
- [Thomas 08]
- [Tari, Broberg, Zomaya, Baldoni 05]
- [Fu, Broberg, Tari 03]

Optimizing SITA cutoffs
- [Harchol-Balter, Crovella, Murta 98]
- [Bachmat, Sarfati 08]
- [Sarfati 08]
- [Harchol-Balter, Vesilo 08]

SITA vs. LWL
- [Broberg, Tari, Zeephongsek 09]
- [Harchol-Balter 99]
- [Ciardo, Risano 99]
- [Crovella, Murta 99]
- [Tari, Broberg, Zomaya, Baldoni 05]
- [Thomas 08]

All conclude SITA far superior for high variability
In search of a proof of SITA’s total dominance.

OK, so not optimal, but definite win for high variability.

Should at least beat all commonly used policies when variability is high enough.

Can’t prove anything because it’s not true!
The TRUTH about SITA, under very high job size variability

\[ C^2 = \frac{\text{var}(X)}{E[X]^2} \rightarrow \infty \quad \text{while} \quad E[X]: \text{fixed} \]
Q: In this talk we will show ... as $C^2 \to \infty$

a) SITA diverges & LWL diverges?
b) SITA converges & LWL diverges ?
c) SITA diverges & LWL converges?
d) SITA converges & LWL converges?

A: All of the above
Q: In this talk we will show ... as $C^2 \to \infty$

<table>
<thead>
<tr>
<th></th>
<th>Convergent LWL</th>
<th>Divergent LWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convergent SITA</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Divergent SITA</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Looking for simple job size distributions to illustrate each.
Results (2 server system)

\[ \text{depends on } p \cdot a \text{ & } (1-p) \cdot b \]

Conv. LWL

Diverg. LWL

Conv. SITA

Diverg. SITA

Bimodal

\[ \text{Exp}(\mu_a) \]

\[ \text{Exp}(\mu_b) \]
Results (2 server system)

Trimodal
\[ \rho < 1 \]

\[ c = b^m > 1 \]

or

\[ H_3 \]
\[ \rho < 1 \]

\[ \text{Exp}(\mu_a) \]
\[ \text{Exp}(\mu_b) \]
\[ \text{Exp}(\mu_c) \]

Conv. LWL

Diverg. LWL

depends on \( m \)
Results (2 server system)

SITA

Conv. LWL

Diverg. LWL

depends on $\alpha$

Bounded Pareto($\alpha$)

$1 < \alpha < 2$
**Bimodal Results**

\[ p_a = QE[X] \]

\[ p \rightarrow a \]

\[ 1-p \rightarrow b \]

\[ (1-p)b = (1-Q)E[X] \]

**Lemma:** As \( C^2 \rightarrow \infty \), but \( E[X] \), \( Q \): const, 
a's get little smaller \( \rightarrow QE[X] \)
b's get much bigger \( \rightarrow \infty \)
\( p \rightarrow 1 \)

**THM:** If \( \rho_a < 1 \) & \( \rho_b < 1 \) \( \Rightarrow \) Convergent SITA

**THM:** LWL always diverges.
Isn’t LWL always bad for high $C^2$?

It depends ...

But shorts stuck behind longs, so $E[T] \to \infty$

Need 2 longs for this to be a problem!

So we need: $\Pr\{2 \text{ longs}\} \times E[T \mid 2 \text{ longs}]$?

Suffices to just look at $E[X^{3/2}]$. 
Understanding LWL

**Thm:** [Scheller-Wolf, Sigman 97], [Scheller-Wolf, Vesilo 06] (2 SERVERS)

If $E[X^{3/2}] < \infty$ & $\rho < 1 \implies E[T]^{LWL} < \infty$

1 spare server

**Thm:**

$$E[ X^{3/2} ] : \text{bounded}$$ & $\rho < 1 \implies \text{LWL converges}$

I can make both happen!
Bimodal Results

Lemma: As $C^2 \to \infty$, but $E[X]$, $Q$: const, $a \to QE[X]$, $b \to \infty$, $p \to 1$

THM: LWL always diverges.

$$E[X^{\frac{3}{2}}] = pa^{\frac{3}{2}} + (1 - p)b^{\frac{3}{2}}$$
$$= QE[X]\sqrt{a} + (1 - Q)E[X]\sqrt{b}$$
$$\to \infty \text{ (as } C^2 \to \infty)$$

THM: If $\rho_a < 1$ & $\rho_b < 1$ $\Rightarrow$ Convergent SITA
Trimodal Results

Lemma: As $C^2 \to \infty$, but $E[X]$: const,
  a $\to E[X]$
  b $\to \infty$, c $\to \infty$
  $p_a \to 1$

THM: If $m \leq 3$, SITA converges
If $m > 3$, SITA diverges

THM: LWL always converges for $\rho < 1$

$E[X^{\frac{3}{2}}] = p_a a^{\frac{3}{2}} + p_b b^{\frac{3}{2}} + p_c c^{\frac{3}{2}}$

$\rightarrow E[X]^{\frac{3}{2}} + 1 + 1$
Results
(2 server system)

Trimodal
\[ \rho < 1 \]
\[ c = b^m > 1 \]

Conv. SITA
\[ \text{Exp}(\mu_a) \]
\[ \text{Exp}(\mu_b) \]
\[ \text{Exp}(\mu_c) \]

Diverg. SITA
\[ \text{Exp}(\mu_a) \]
\[ \text{Exp}(\mu_b) \]

Conv. LWL

Diverg. LWL

Way more complex, because job types overlap!

"Separation in the limit"
Bounded Pareto (2 server system)

\[ X \sim \text{Bounded Pareto} (k, p, \alpha) \]

Lemma: As \( C^2 \rightarrow \infty \), but \( E[X] \), \( \alpha \): const, \( k \rightarrow \frac{(\alpha - 1)}{\alpha} \cdot E[X] \), \( p \rightarrow \infty \)

THM: SITA always diverges.

THM: If \( \alpha > 3/2 \) and \( \rho < 1 \), then LWL converges. Else LWL diverges.

Extends to \( n > 2 \) servers when \( \rho < n-1 \).
Bounded Pareto Results

Why was this not noticed?

$\alpha = 1.6$

$\alpha = 1.4$
Summary

Trimodal
\[ \rho < 1 \]
\[ c = b^n > 1 \]

Conv.
LWL

Conv.
SITA

Bimodal

Diverg.
LWL

Diverg.
SITA

or

Exp(\(\mu_a\))

Exp(\(\mu_b\))

Exp(\(\mu_c\))

H_3
\[ \rho < 1 \]

\(1-p\)

\[ 1 < \alpha < 2 \]

Bounded Pareto(\(\alpha\))

Exp(\(\mu_a\))

Exp(\(\mu_b\))

or

or
Old Nursery Rhyme

When SITA is good, it is very, very good  
But when it is bad, it is horrid.