On the influence of resequencing on the regularity of service

A. Jean-Marie, M. Tidball, M. Escalante, V. Leoni and H. Ponce de León

INRIA, Sophia Antipolis, France
LIRMM, Montpellier, France
Universidad Nacional de Rosario, Argentina

Performance’99, 16 October 1999
Plan

- Introduction
- The resequencing model
- General results
- Markovian analysis
- Performance comparisons
**Introduction**

Communication networks $\implies$ disordering $\implies$ resequencing.

Application Level Framing $\implies$ applications may cope with disordered elements of information.

Increased programming complexity $\implies$ is it worth the effort?

Intuition says “yes” when *jitter* is the measure. Jitter is caused by variations in the network delay *plus* resequencing. Diot and Gagnon confirm this using simulation.

Our contribution: propose a quantitative model to investigate the issue.
The resequencing model

The general resequencing model

Interesting quantities:

- \( r_n \) = total delay of customer \( n \), i.e. delay plus resequencing time.
- \( f_n \) = time at which the service starts.
- \( w_n \) = the waiting time in the queue.
- \( y_n \) = time spent by the server waiting customer number \( n \).
Jitter

Need one definition of jitter.

A function of the fidelity of the network:

\[ j_n = \left( (f_{n+1} + \sigma_{n+1}) - (f_n + \sigma_n) \right) - (t_{n+1} - t_n) \]
\[ = y_{n+1} + \sigma_{n+1} - (t_{n+1} - t_n). \]

Then, since \( \sigma_{n+1} \) is independent from the rest:

\[ \text{Var}(j_n) = \text{Var}(y_{n+1}) + \text{Var}(\sigma) + \text{Var}(t_{n+1} - t_n) \]
\[ - 2 \text{Cov}(y_{n+1}, t_{n+1} - t_n). \]

In the particular case where \( \sigma_n \equiv \sigma \) and \( t_{n+1} = t_n + \delta t \):

\[ \text{Var}(j_{n+1}) = \text{Var}(y_{n+1}). \]
Preliminaries


The queueing system is actually a $G/G/\infty \rightarrow ./G/1$ network.

The presence of resequencing modifies the arrival process to the second queue.

Results known in some cases

- $M/GI/\infty$ queue: resequencing time distribution (Harrus and Plateau, 1982)
- $M/M/\infty \rightarrow ./GI/1$ queue: Laplace transform of the end-to-end delay distribution (BGP, 1984).
Dynamics

Denote

\[ \xi_n := t_{n+1} - t_n \quad \beta_n := \xi_n - \sigma_n \quad z_n := w_n + r_n \]

Then:

\[ r_{n+1} = \max \{ \Delta_n, r_n - \xi_n \} \]

\[ \tau_n = [t_{n+1} - t_n + \delta_n - r_n]^+ = r_{n+1} - r_n + \xi_{n+1} \]

\[ z_{n+1} = \max \{ \Delta_n, r_n - \beta_n \} \]

\[ y_{n+1} = [w_n + \sigma_n - (r_{n+1} - r_n + \xi_n)]^- = z_{n+1} - z_n + \beta_{n+1} \]

\[ w_{n+1} = [w_n + \sigma_n - \tau_n]^+ \]

Using Loynes' scheme:

\[ z_n = \max \left\{ \max_{m=1}^n \{ \Delta_m - \sum_{j=m}^{n} \beta_j \}, z_0 - \sum_{j=0}^{n-1} \beta_j \right\} \]
Qualitative Results

Stability
Resequencing does not modify stability: if \( \{t_{n+1} - t_n\}_n \) and \( \{\sigma_n\}_n \) are jointly stationary and ergodic, then

\[
\mathbb{E}(t_{n+1} - t_n) > \mathbb{E}\sigma
\]

implies the existence of a stationary regime.

Insensitivity
For all \( \{t_{n+1} - t_n\}_n \) and \( \{\Delta_n\}_n \), in stationary regime:

\[
\mathbb{E}y_n = 1 - \mathbb{E}\sigma\mathbb{E}(t_{n+1} - t_n).
\]
Stochastic comparison

Internal monotonicity:
Assume that \( r_0 = 0 \) and \( z_0 = 0 \). Then for all \( n \geq 0 \):

\[
  r_{n+1} \geq_{icx} r_n \\
  z_{n+1} \geq_{icx} z_n .
\]

External monotonicity:
Assume that \( \Delta_n \leq_{icx} \tilde{\Delta}_n \) for all \( n \). Then:

\[
  r_n \leq_{icx} \tilde{r}_n \\
  w_n \leq_{icx} \tilde{w}_n
\]

Question: under which conditions is it true that

\[
  y_n \leq_{cx} \tilde{y}_n ?
\]
Quantitative results

Assume

– a disordered model of type $D/GI/\infty$ ($t_n = n$);
– $\Delta$ integrable;
– a queueing model of type $/D/1$ with $\sigma < 1$

Then:

\[
P\{r_n \leq x\} = P\{r_0 \leq x + n\} \prod_{m=0}^{n-1} \Delta(x + m)
\]

\[
P\{z_n \leq x\} = P\{z_0 \leq x + n\beta\} \prod_{m=0}^{n-1} \Delta(x + m\beta)
\]

\[
P\{R \leq x\} = \prod_{m=0}^{\infty} \Delta(x + m)
\]

\[
P\{Z \leq x\} = \prod_{m=0}^{\infty} \Delta(x + m\beta)
\]
Computing $\text{Var}(Y)$

In the stationary regime:

$$Y = [\Delta - Z + \beta]^+$$

Therefore for $y > 0$,

$$\mathbb{P}\{Y \geq y\} = \mathbb{P}\{Z \leq \Delta + \beta - y\}$$

and we know:

$$\mathbb{P}\{Z \leq x\} = \prod_{m=0}^{\infty} \Delta(x + m\beta).$$
Markov Chains

With resequencing
Assume that the sequences \( \{ \Delta_n, n \in \mathbb{N} \} \), \( \{ t_{n+1} - t_n, n \in \mathbb{N} \} \) and \( \{ \sigma_n, n \in \mathbb{N} \} \) are i.i.d. and mutually independent. Then, the following processes are Markov chains:

- \( \{ r_n, n \in \mathbb{N} \} \)
- \( \{ (r_n, w_n), n \in \mathbb{N} \} \)
- \( \{ (r_n, z_n), n \in \mathbb{N} \} \)

Assume that the delays \( \Delta_n \) are bounded by a constant \( \Delta \). Then \( w_n \) takes a finite number of different values and the Markov chains \( \{ (r_n, w_n), n \in \mathbb{N} \} \) and \( \{ (r_n, z_n), n \in \mathbb{N} \} \) are finite.
Without resequencing

Define: \( \tilde{r}_n \) as the list of the next relative times of arrival of the delayed customers, just after the \( n^{th} \) arrival to the queue.
\[
\tilde{r}_n = (\tilde{r}_n^1, \tilde{r}_n^2, ..., \tilde{r}_n^{m(n)}),
\]
where \( m(n) \) is the number of delayed customers at time \( n \).

For all \( \Delta \), \( \#E_\Delta = 2^{\Delta+1} - 2^{\Delta-1} = 3 \cdot 2^{\Delta-1} \).

The process \( \{(\tilde{r}_n, w_n)\} \) is a finite Markov chain.
**Distribution of inter-arrivals**

With resequencing

\[
\begin{align*}
\mathbb{P}(\tau = 0) &= \bar{\epsilon} (1 - \bar{\epsilon}^\Delta) \\
\mathbb{P}(\tau = 1) &= \bar{\epsilon}^{\Delta + 1} + \epsilon^2 \\
\mathbb{P}(\tau = k) &= \epsilon^2 \bar{\epsilon}^{k+1} \quad 2 \leq k \leq \Delta \\
\mathbb{P}(\tau = \Delta + 1) &= \epsilon \bar{\epsilon}^\Delta.
\end{align*}
\]

\[
\text{Var}(\tau) = \frac{\bar{\epsilon}}{\epsilon} (1 - \bar{\epsilon}^\Delta (1 + \Delta\epsilon)).
\]

Without resequencing

\[
\begin{align*}
\mathbb{P}(\tau = 0) &= \epsilon, \\
\mathbb{P}(\tau = k) &= \epsilon^2 \cdot \epsilon^{k-1} \quad \text{for } k \geq 1;
\end{align*}
\]

\[
\text{Var}(\tau) = \frac{2\epsilon\bar{\epsilon} \left[1 - (\epsilon\bar{\epsilon})^\Delta\right]}{1 - \epsilon\bar{\epsilon}}.
\]
Comparison of arrival processes

With resequencing: $\tau_R$. Without resequencing: $\tau_N$.
Graph of the difference between variances: $\text{Var}(\tau_N) - \text{Var}(\tau_R)$.

Increased variance: when $\Delta$ large and $\epsilon$ small.
Comparison of server waiting times

Values of $V_R(\sigma)$ (left) and $V_N(\sigma)$ (right) for different values of $\Delta$, $\epsilon = 0.1$
A bad performance measure

The idle period length distribution:

\[ \mathbb{E}I = \frac{1 - \sigma}{\epsilon + \epsilon^2 |1 - \sigma| + 1} \]

\[ \mathbb{E}I(\sigma) \]

Figure 1: \( \mathbb{E}I(\sigma) \) for \( \Delta = 20, \epsilon = 0.02 \)
Conclusion

- Intuition is not always founded: resequencing may decrease the variance of the arrival process, and jitter;
- Not resequencing is better if $\sigma \leq \sigma^*(\epsilon)$, and $\sigma^*(\epsilon) \simeq 1$ when $\epsilon \to 0$;
- More research with other tractable models: $D/Geom/\infty \to ./D/1$ with and without resequencing;