A First Course in Stochastic Network Calculus

Dipl.-Math. Michael Alexander Beck

January 28, 2014

Contents

Foreword

This *First Course in Stochastic Network Calculus* is a “First” course in two different perspectives. One can see it as a introductory course to Stochastic Network Calculus. On the other side it was my first course I have given about SNC to a few students in June 2012. It builds on the lecture *Performance Modelling in Distributed Systems [1]* at the University of Kaiserslautern. The concepts of stochastic network calclulus parallels those of deterministic network calculus. This is why I reference on the lecture of 2011 at several points to stress these connections. This document however is thought of a stand-alone course and hence a deep study of [1] is not necessary (but recommended).

This course contains a rather large probability primer to ensure the students can really grasp the expressions, which appear in stochastic network calculus. A student familiar with probability theory might skip this first chapter and delve directly into the stochastic network calculus. For each topic exercises are given, which can (and should) be used to strengthen the understanding of the presented definitions and theory.

This document is still in process and hopefully will evolve at some day into a fully grown course about Stochastic Network Calculus, providing a good overview over this exciting theory. Hence, please provide feedback to `beck@cs.uni-kl.de`.

- Michael Beck

1 Basics of Probability

2 Stochastic Arrivals and Service

3 Stochastic Performance Bounds

Now, based on the models for stochastic arrivals and service guarantees, we ask for performance guarantees in this chapter. We are interested in bounds on the backlog and the delay of a node, as well as in bounds on the departures of a node. Again, the theory here parallels the deterministic approach, with the difference, that the bounds only hold with a high probability. Or, stated more positively: The achieved bounds are only violated with very small probabilities.

If not stated otherwise we assume in this chapter a flow A, which is bounded by $(\sigma_A(\theta), \rho_A(\theta))$ and a stochastically independent node S, which is $(\sigma_S(\theta), \rho_S(\theta))$-bounded for the same $\theta > 0$. Further, for the tail-bounded case, we assume A and S to be tail-bounded by envelopes α_A and α_S with errors η_A and η_S, respectively. The flow A is the input for node S and the corresponding output flow is denoted by D.

3.1 Backlog Bound

3.2 Delay Bound

3.3 Output Bound

Again we follow the same strategy as for the backlog bound.

Theorem 3.1. *For $\rho_S(\theta) < -\rho_A(\theta)$ the output is $(\sigma(\theta), \rho(\theta)$-bounded with*

$$\sigma(\theta) := \sigma_A(\theta) + \sigma_S(\theta) - \frac{1}{\theta} \log(1 - e^{\theta(\rho_A(\theta)+\rho_S(\theta))})$$

and

$$\rho(\theta) := \rho_A(\theta).$$

Proof. We have:

$$\begin{aligned} D(n) - D(m) &\leq D(n) - \min_{0 \leq k \leq m} \{A(k) + S(k, m)\} \\ &\leq A(n) - \min_{0 \leq k \leq m} \{A(k) + S(k, m)\} \\ &= \max_{0 \leq k \leq m} \{A(n) - A(k) - S(k, m)\} = A \oslash S(m, n) \end{aligned}$$

Hence:

$$
\begin{aligned}
\mathbb{E}(e^{\theta(D(n)-D(m))}) &\le \mathbb{E}(e^{\theta(A \oslash S(m,n))}) \le \mathbb{E}(e^{\theta A}) \circ \mathbb{E}(e^{-\theta S})(m,n) \\
&= \sum_{k=0}^{m} \mathbb{E}(e^{\theta A(k,n)})\mathbb{E}(e^{-\theta S(k,m)}) \\
&\le \sum_{k=0}^{m} e^{\theta\rho_A(\theta)(n-k)+\theta\sigma_A(\theta)} e^{\theta\rho_S(\theta)(m-k)+\theta\sigma_S(\theta)} \\
&= e^{\theta\rho_A(\theta)(n-m)+\theta\sigma_A(\theta)+\theta\sigma_S(\theta)} \sum_{k=0}^{m} e^{\theta(m-k)(\rho_A(\theta)+\rho_S(\theta))} \\
&\le e^{\theta\rho_A(\theta)(n-m)+\theta(\sigma_A(\theta)+\sigma_S(\theta)-\frac{1}{\theta}\log(1-\exp(\theta\rho_A(\theta)+\rho_S(\theta)))}
\end{aligned}
$$

Where the sum was bounded by its corresponding geometric series in the last line ($\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ if $0 < |q| < 1$). $\square$

Note that the last step in the above proof is not necessary to get a bound on the output. The only reason for performing the last inequality is land in the class of $(\sigma(\theta), \rho(\theta))$-bounded arrivals. Coming back to this framework allows analyzing whole networks, since the output bound can be used again as input bound in theorems about backlog or delay. In this course we sticked to the $(\sigma(\theta), \rho(\theta))$-bounds, since they make formulas more tractable and give us a clearer view on the overall picture.

However all of the presented theorems can be generalized to the usage of $f(\theta, n)$-bounds, avoiding the last step in the previous proof. Leaving it out, results in bounds, which can be significally better.

The last thing to do is the tail-bounded version of output bounds. For these we need again a stability condition, similar to the one for delay-bounds:

Definition 3.2. We say S *is output-stable for some* A *and* $\varepsilon > 0$, if for each $m \ge 0$ exists an n_m such that

$$\alpha_A(\varepsilon) \oslash (S - \alpha_S(\varepsilon))(n, n+m) \le \alpha_A(\varepsilon) \oslash (S - \alpha_S(\varepsilon))(n_m, n_m + m)$$

holds for all $n \ge 0$.

This definition allows to give a description of the output under tail-bounds.

Theorem 3.3. *Assume* S *is output-stable for* A*. Define the envelope*

$$\alpha_D(m, \varepsilon) :=:= \alpha_A(\varepsilon) \oslash (S - \alpha_S(\varepsilon))(n_m, n_m + m)$$

We have then, that D *is tailbounded by envelope* α_D *with error*

$$\eta_D(m, \varepsilon) := \eta_S(0, \varepsilon) + \sum_{k=0}^{\infty} \eta_A(m + k, \varepsilon)$$

Proof. Let $m, n \in \mathbb{N}_0$ and $\varepsilon > 0$ arbitrary. We consider the expression $D(n+m) - D(n)$. Assume for a while that

$$D(0,n) \geq A \otimes (S - \alpha_S(\varepsilon))(0,n) \tag{3.1}$$

holds, as well as:

$$A(k, n+m) \leq \alpha_A(n+m-k, \varepsilon) \qquad \forall\, k \leq n \tag{3.2}$$

We have then:

$$\begin{aligned}
D(n+m) - D(n) &\leq D(n+m) - \min_{0 \leq k \leq n} \{A(k) + S(k,n) - \alpha_S(n-k,\varepsilon)\} \\
&\leq A(n+m) - \min_{0 \leq k \leq n} \{A(k) + S(k,n) - \alpha_S(n-k,\varepsilon)\} \\
&= \max_{0 \leq k \leq n} \{A(n+m) - A(k) - S(k,n) + \alpha_S(n-k,\varepsilon)\} \\
&\leq \max_{0 \leq k \leq n} \{\alpha_A(n+m-k) - S(k,n) + \alpha_S(n-k,\varepsilon)\} \\
&= \alpha_A(\varepsilon) \oslash (S - \alpha_S(\varepsilon))(n, n+m) \\
&\leq \alpha_A(\varepsilon) \oslash (S - \alpha_S(\varepsilon))(n_m, n_m + m) = \alpha_D(m,\varepsilon)
\end{aligned}$$

Here we have used the stability condition in the last line. Again, moving to probabilities finishes the proof:

$$\begin{aligned}
&\mathbb{P}(D(n, n+m) > \alpha_D(m,\varepsilon)) \\
\leq\, &\mathbb{P}\left(D(0,n) < A \otimes (S - \alpha_S(\varepsilon))(0,n) \cup \bigcup_{k=0}^{n} A(k, n+m) > \alpha_A(n+m-k,\varepsilon) \right) \\
\leq\, &\eta_S(n,\varepsilon) + \sum_{k=0}^{n} \eta_A(n+m-k,\varepsilon) \\
\leq\, &\eta_S(0,\varepsilon) + \sum_{k'=0}^{n} \eta_A(m+k',\varepsilon) \\
\leq\, &\eta_S(0,\varepsilon) + \sum_{k'=0}^{\infty} \eta_A(m+k',\varepsilon)
\end{aligned}$$

□

We have seen two stability conditions needed to derive performance bounds, when we are in the tail-bounded case. To state the theorems as general as possible, we introduced delay-stability, as well as, output-stability. The following definition, brings these two together and allows computation of both bounds at the same time (while being a "too" strict of an assumption, if one is interested in either delay or output):

Definition 3.4. S is *stable for some arrival A and $\varepsilon > 0$*, if S is delay-stable and output stable for A and ε.

Bibliography

[1] *J. B. Schmitt. Lecture notes: Performance modeling of distributed systems. http://disco.informatik.uni-kl.de/content/PDS_ WS1112, October 2011.*