A First Course in Stochastic Network Calculus

Dipl.-Math. Michael Alexander Beck

January 20, 2014

Contents

1 Basics of Probability **4**

2 Stochastic Arrivals and Service **5**

3 Stochastic Performance Bounds **6**

3.1 Backlog Bound 6

3.2 Delay Bound 6

3.3 Output Bound 12

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

The delay bound is derived similarly to the backlog bound and we follow the same scheme as presented in the previous section. First a definition of delay is needed.

Definition 3.1. The *virtual delay* d at time n is defined as:

$$d(n) := \min\{m : A(n) < D(n+m)\}$$

The virtual delay can be interpreted as the time D needs to catch up with the amount of arrivals.

Theorem 3.2. *For all $N \in \mathbb{N}_0$ it holds that:*

$$\mathbb{P}(d(n) > N) \leq e^{\theta\rho_S(\theta)N + \theta(\sigma_A(\theta)+\sigma_S(\theta))} \sum_{k=0}^{n+N} e^{\theta(n-k)(\rho_A(\theta)+\rho_S(\theta))}$$

Proof. Assume it holds $d(n) > N$, i.e. $A(n) - D(n+N) > 0$. Then we have:

$$\begin{aligned} 0 < A(n) - D(n+N) &\leq A(n) - A \otimes S(0, n+N) \\ &= \max_{0 \leq k \leq n+N} \{A(n) - A(k) - S(k, n+N)\} \end{aligned}$$

Hence the implication

$$d(n) > N \quad \Rightarrow \quad \max_{0 \leq k \leq n+N} \{A(k,n) - S(k, n+N\} > 0$$

holds and therefore:

$$
\begin{aligned}
\mathbb{P}(d(n) > N) &\leq \mathbb{P}(\max_{0 \leq k \leq n+N} \{A(k,n) - S(k,n+N)\} > 0\} \\
&\leq \mathbb{E}(e^{\theta \max_{0\leq k \leq n+N}\{A(k,n)-S(k,n+N)\}}) \\
&\leq \mathbb{E}(\max_{0 \leq k \leq n+N} e^{\theta(A(k,n)-S(k,n+N))}) \leq \sum_{k=0}^{n+N} \mathbb{E}(e^{\theta(A(k,n)-S(k,n+N))}) \\
&= \sum_{k=0}^{n+N} \mathbb{E}(e^{\theta A(k,n)}) \mathbb{E}(e^{-\theta S(k,n+N)}) \\
&\leq \sum_{k=0}^{n+N} e^{\theta\rho_A(\theta)(n-k)+\theta\sigma_A(\theta)} e^{\theta\rho_S(\theta)(N+n-k)+\theta\sigma_S(\theta)} \\
&= e^{\theta\rho_S(\theta)N+\theta(\sigma_A(\theta)+\sigma_S(\theta))} \sum_{k=0}^{n+N} e^{\theta(n-k)(\rho_A(\theta)+\rho_S(\theta))}
\end{aligned}
$$

□

Analogous to the backlog bound, the above bound can be derived in a slightly different way. From $\mathbb{P}(\max_{0\leq k\leq n}\{A(k,n) - S(k,n+N) > 0\})$ one might continue with:

$$
\begin{aligned}
\mathbb{P}(d(n) > N) &\leq \mathbb{P}\left(\bigcup_{k=0}^{n+N} A(k,n) - S(k,n+N) > 0\right) \\
&\leq \sum_{k=0}^{n+N} \mathbb{P}(A(k,n) - S(k,n+N) > 0) \\
&\leq \sum_{k=0}^{n+N} \mathbb{E}(e^{\theta(A(k,n)-S(k,n+N))})
\end{aligned}
$$

and then proceed as before. Also, we can state the same corollary concerning the parameter θ.

Corollary 3.3. *Assume A is $(\sigma_A(\theta), \rho_A(\theta))$-bounded for all $\theta \in [0,b]$ and S is $(\sigma_S(\theta), \rho_S(\theta))$-bounded for all $\theta \in [0,b]$ and some $b \in \mathbb{R}^+ \cup \{\infty\}$. Then:*

$$
\mathbb{P}(d(n) > N) \leq \inf_{\theta\in[0,b]} e^{\theta\rho_S(\theta)N+\theta(\sigma_A(\theta)+\sigma_S(\theta))} \sum_{k=0}^{n+N} e^{\theta(n-k)(\rho_A(\theta)+\rho_S(\theta))}
$$

Making a further (but often reasonable) assumption on S the above bound can be improved.

Corollary 3.4. *Assume the situation as in the previous corollary. Further let $S(k, n+N) \geq 0$ for all $k \geq n+1$. Then:*

$$
\mathbb{P}(d(n) > N) \leq \inf_{\theta\in[0,b]} e^{\theta\rho_S(\theta)N+\theta(\sigma_A(\theta)+\sigma_S(\theta))} \sum_{k=0}^{n} e^{\theta k(\rho_A(\theta)+\rho_S(\theta))}
$$

Proof. Using $A(k,n) - S(k,n+N) \leq 0$ for all $k \geq n+1$ we have:

$$\begin{aligned} 0 &< A(n) - D(n+N) \leq A(n) - A \otimes S(0,n+N) \\ &= \max_{0\leq k \leq n+N} \{A(n) - A(k) - S(k,n+N)\} \\ &= \max_{0\leq k\leq n} \{A(n) - A(k) - S(k,n+N)\} \end{aligned}$$

The remainder of the proof follows the same lines as above. □

For the tail-bounded version we need a stability condition on the service.

Definition 3.5. We say S is *delay-stable* (for ε), if

$$S(m,n) - \alpha_S(n-m,\varepsilon) >\geq 0 \qquad \forall\, m,n \geq 0$$

and

$$S(m,n) - \alpha_S(n-m,\varepsilon) \xrightarrow{n\to\infty} \infty \qquad \forall\, m \geq 0$$

holds.

Theorem 3.6. *Assume S is delay stable for some ε. Then we have for all $n \in \mathbb{N}_0$:*

$$\begin{aligned} &\mathbb{P}(d(n) > \min\{N \geq 0 : (\alpha_S(\varepsilon) - S) \oslash \alpha_A(\varepsilon)(n, n+N') < 0\,\forall\, N' \geq N) \\ &\leq \sum_{k=0}^{n} \eta_A(k,\varepsilon) + \sum_{k=n+1}^{\infty} \eta_S(k,\varepsilon) \end{aligned}$$

Proof. Fix an arbitrary n and assume for a while

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

for all $l \geq 0$ and

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

for all $k \leq n$. Choose now some arbitrary $l < d(n)$, then we have by the definition of virtual delay and (3.1):

$$A(n) > D(n+l) \geq \min_{0\leq k\leq n+l} \{A(k) + S(k,n+l) - \alpha_S(n+l-k,\varepsilon)\}$$

From this and (3.2) we can derive

$$\begin{aligned} 0 &< \max_{0\leq k\leq n+l} \{A(n) - A(k) - S(k,n+l) + \alpha_S(n+l-k,\varepsilon)\} \\ &= \max_{0\leq k\leq n} \{A(n) - A(k) - S(k,n+l) + \alpha_S(n+l-k,\varepsilon)\} \\ &\quad \vee \max_{n+1\leq k\leq n+l} \{A(n) - A(k) - S(k,n+l) + \alpha_S(n+l-k,\varepsilon)\} \\ &\leq \max_{0\leq k\leq n} \{A(n) - A(k) - S(k,n+l) + \alpha_S(n+l-k,\varepsilon)\} \vee 0 \\ &\leq \max_{0\leq k\leq n} \{\alpha_A(n-k,\varepsilon) - S(k,n+l) + \alpha_S(n+l-k,\varepsilon)\} \vee 0 \end{aligned} \tag{3.3}$$

where we have used that S is delay-stable in the transition to the fourth line. Writing it shorter, we eventually have derived $0 < (\alpha_S(\varepsilon) - S) \oslash \alpha_A(\varepsilon)(n, n+l)$ for all $l < d(n)$. By the assumption of delay-stability, we know there exists for each k an N_k, such that

$$\alpha_A(n-k,\varepsilon) - S(k, n+N') + \alpha_S(n+N'-k,\varepsilon) \leq 0$$

for all $N' \geq N_k$. Define $N := \max_{0 \leq k \leq n} N_k$, then we have

$$(\alpha_S(\varepsilon) - S) \oslash \alpha_A(\varepsilon)(n, n+N') \leq 0$$

for all $N' \geq N$. Because of (3.3) we must have that $l < N$ for each $l < d(n)$ and hence arrive at:

$$d(n) \leq N = \min\{N \geq 0 : (\alpha_S(\varepsilon) - S) \oslash \alpha_A(\varepsilon)(n, n+N') \leq 0 \, \forall \, N' \geq N\}$$

Moving to probabilities yields:

$$\begin{aligned}
&\mathbb{P}(d(n) \leq N) \\
\leq &\mathbb{P}\left(\bigcup_{k=0}^{n} A(k,n) > \alpha_A(n-k,\varepsilon) \cup \bigcup_{l \geq 0} D(n+l) < A \oslash (S - \alpha_S(\varepsilon))(0, n+l)\right) \\
\leq &\sum_{k=0}^{n} \eta_A(k,\varepsilon) + \sum_{k=n+1}^{\infty} \eta_S(k,\varepsilon)
\end{aligned}$$

□

Again, we can make similar comments, to the ones made for the backlog bound: Solving the equation

$$x = \min\{N \geq 0 : (\alpha_S(\varepsilon) - S) \oslash \alpha_A(\varepsilon)(n, n+N') \leq 0 \, \forall \, N' \geq N\}$$

is quite involved and the existence of a solution is not guaranteed. Further for the generalized case uniqueness of a solution (if existence) is in general not given. We close with the tail-bound version of a delay bound, which allows optimization over ε and ε':

Corollary 3.7. *Assume S is delay-stable for all $\varepsilon \in I$ (I being some interval). Then we have for all $n \in \mathbb{N}_0$ and $\varepsilon' > 0$ and $\varepsilon \in I$:*

$$\begin{aligned}
&\mathbb{P}(d(n) > \min\{N \geq 0 : (\alpha_S(\varepsilon') - S) \oslash \alpha_A(\varepsilon)(n, n+N') < 0 \, \forall \, N' \geq N) \\
\leq &\sum_{k=0}^{n} \eta_A(k,\varepsilon) + \sum_{k=n+1}^{\infty} \eta_S(k,\varepsilon')
\end{aligned}$$

Exercises

Exercise 3.8. One can also define a backwards oriented delay by:

$$\tilde{d}(n) := \min\{m : D(n) - A(n-m) > 0\}$$

Show that

$$\mathbb{P}(\tilde{d}(n) > N) \leq e^{\theta\rho_S(\theta)N + \theta(\sigma_A(\theta)+\sigma_S(\theta))} \sum_{k=0}^{n-N} e^{\theta k(\rho_A(\theta)+\rho_S(\theta))}$$

for all $n > N \in \mathbb{N}_0$.

Exercise 3.9. If we compare the backlog bound with the delay bound we notice that they only differ in the leading terms $e^{-\theta x}$ and $e^{\theta\rho_S(\theta)N}$. In fact one can interpret the delay bound as a backlog bound, if there is some guarantee on the service.

Assume that $D(n) - D(m) \geq f(m,n)$ if $[m,n]$ is a backlog period. (i.e. the node offers a deterministic strict service curve). Show that the delay bound, can be expressed as a backlog bound:

$$\mathbb{P}(d(n) > N) \leq \mathbb{P}(q(n) > f(n, n+N))$$

(Hint: Construct first the dynamic S-Server. Then show that from $d(n) > N$ follows $q(n) > f(n, n+N)$)

Exercise 3.10. Assume now that the arrivals have a strict bound, i.e. $A(m,n) \leq f(m,n)$, where f is monotone decreasing in the first variable. Show that the backlog bound can be expressed by the backward delay bound from exercise 3.8:

$$\mathbb{P}(q(n) > x) \leq \mathbb{P}(\tilde{d}(n) > m)$$

with m maximal such that

$$f(m,n) < x$$

is still fulfilled.

Exercise 3.11. Let us investigate the bound from 3.2 in more detail, in expression the term $\sum_{k'=0}^{n} e^{\theta k'(\rho_A(\theta)+\rho_S(\theta))}$. You might remember the geometric series:

$$\sum_{k=0}^{\infty} q^k$$

which converges to the value $\frac{1}{1-q}$ if $|q| < 1$ holds. How behaves the delay bound for $n \to \infty$, in the case $\rho_A(\theta) \geq -\rho_S(\theta)$?

The above shows, that one can think of $\rho_A(\theta) < -\rho_S(\theta)$ as a stability condition. The following counterexample will however show, that it is not a stability condition in the intuitive sense, that the utilization of a node is below 100%. Assume for this a constant rate server with rate c and an arrival with exponentially distributed increments (as in example **??**). The average rate of arrivals per time step is equal to: $\mathbb{E}(a(n)) = \frac{1}{\lambda}$, where

λ is the parameter of the exponential distribution. This gives for the node an utilization of

$$\frac{\frac{1}{\lambda}}{c} = \frac{1}{\lambda c}$$

Show that there exists a choice of c, λ and θ, sucht that $\frac{1}{\lambda c} < 1$ holds, but also:

$$\rho_A(\theta) \geq -\rho_S(\theta)$$

This means our system is stable by construction (the node's utilization is below 100%), but our delay bound diverges for $n \to \infty$, no matter how large we choose N in 3.2. This makes it blatantly obvious, that the derived delay-bound is really only a *bound*. A bound, which in some situations is far from the real systems behaviour. Hence the right choice of θ is crucial[1].

Exercise 3.12. In this exercise we investigate how different timescales influence the quality of the delay bound. In expression, if one has two descriptions of a system differing only in different definitions of what one timestep is, which of the two descriptions leads to a better delay bound?

To do this assume a flow A which is $(0, \rho_A(\theta))$-bounded and a node, which offers a constant rate c. Introduce a parameter $m \in \mathbb{N}$ called granularity and a corresponding flow A_m such that:

$$a_m(k) := \sum_{l=1}^{m} a(m \cdot (k-1) + l)$$

Give a bound for $\mathbb{E}(e^{\theta A_m(k,n)})$.

Next we define the virtual delay with granularity m. Let T be a multiple of m:

$$d_m(T) := m \cdot \min\{k \,:\, A(T) < D(T + m \cdot k)\}$$

Show that $d(T) = d_m(T)$.

Show that for all N being a multiple of m holds:

$$\mathbb{P}(d_m(T) > N) \leq e^{-\theta c N} \sum_{l=0}^{\frac{T}{m}} (e^{\theta(\rho_A(\theta) - c) \cdot m})^l$$

Show now that for an even N and T the bound for $d_2(T)$ is strictly smaller than the bound for $d(T)$. Using this result, what is the best bound one can achieve for the event, that the delay at (an arbitrary) time $T \in \mathbb{N}$ is smaller than (an arbitrary) $N \in \mathbb{N}$?

---

[1] To see this even more obvious you can prove the following: For every choice of c and λ with $1/\lambda < c$ exists a θ such that $\rho_A(\theta) \geq -\rho_S(\theta)$. You may proceed like follows: Reformulate $\rho_A(\theta) \geq -\rho_S(\theta)$ as $(1 - \theta/\lambda)e^{\theta c} < 1$. Now interpret the last expression as a function in θ, denoted by f. Calculate the values $f(0)$ and $f(\lambda)$ and apply the intermediate value theorem.

3.3 Output Bound

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

Theorem 3.13. *For $\rho_S(\theta) < -\rho_A(\theta)$ the output is bounded by $(\sigma_A(\theta)+\sigma_S(\theta)+B(\rho_A(\theta), \rho_S(\theta)), \rho_A(\theta))$. With $B(\rho_A(\theta), \rho_S(\theta)) = -\frac{1}{\theta}\log(1 - e^{(\theta\rho_A(\theta)+\rho_S(\theta))})$*

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:

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

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 $0 < |q| < 1$). □

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 to 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 **??**, 3.2 and 3.13. 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 gets us bounds, which can be significally better (compare exercise 3.11 to see how much one can gain here).

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 3.14. We say S is output-stable (for ε) if for each $m \geq 0$ exists an N_m such that

$$S(k, k+n) > \alpha_S(n, \varepsilon) + \alpha_A(n+m, \varepsilon)$$

holds for all $k \geq 0$ and all $n \geq N_m$.

The above definition states, that the service can catch up on $\alpha_S(n,\varepsilon) + \alpha_A(n+m,\varepsilon)$ after at most N_m time steps and keeps being larger after that point, no matter at which time k we start observing the system. Further it states that such an N_m exists for every choice of m.

Theorem 3.15. *Assume S is output-stable for A. Define the following envelope*

$$\begin{aligned}\alpha_D(m,\varepsilon) &:= \alpha_A(\varepsilon) \oslash (S - \alpha_S(\varepsilon))(n_0, n_0+m) \\ &= \max_{0\le k\le n_0} \{\alpha_A(n_0+m-k,\varepsilon) - S(k,n_0) + \alpha_S(n_0-k,\varepsilon)\} \\ &:= \max_{n\ge 0} \max_{0\le k\le n} \{\alpha_A(n+m-k,\varepsilon) - S(k,n) + \alpha_S(n-k,\varepsilon)\}\end{aligned}$$

where n_0 is the index, which maximizes the expression on the last line.

We have 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

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

holds, as well as:

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

We have then:

$$\begin{aligned}D(n+m) - D(n) &\le D(n+m) - \min_{0\le k\le n} \{A(k) + S(k,n) - \alpha_S(n-k,\varepsilon)\} \\ &\le A(n+m) - \min_{0\le k\le n} \{A(k) + S(k,n) - \alpha_S(n-k,\varepsilon)\} \\ &= \max_{0\le k\le n} \{A(n+m) - A(k) - S(k,n) + \alpha_S(n-k,\varepsilon)\} \\ &\le \max_{0\le k\le n} \{\alpha_A(n+m-k) - S(k,n) + \alpha_S(n-k,\varepsilon)\} \\ &\le \max_{n\ge 0} \max_{0\le k\le n} \{\alpha_A(n+m-k) - S(k,n) + \alpha_S(n-k,\varepsilon)\} \\ &= \max_{0\le k\le n_0} \{\alpha_A(n_0+m-k) - S(k,n_0) + \alpha_S(n_0-k,\varepsilon)\} \end{aligned} \tag{3.3}$$

Here we must give a reason, why the second to last line is well defined. For this we show the set $\{\max_{n\ge 0} \max_{0\le k\le n}\{\alpha_A(n+m-k) - S(k,n) + \alpha_S(n-k,\varepsilon)\}\}$ is bounded from above, in such a way that a maximizing index n_0 exists. Of course, our stability condition is the striking argument:

$$\begin{aligned}
&\max_{n\geq 0}\max_{0\leq k\leq n}\{\alpha_A(n+m-k)-S(k,n)+\alpha_S(n-k,\varepsilon)\}\\
=&\max_{k\geq 0}\max_{n\geq k}\{\alpha_A(n+m-k)-S(k,n)+\alpha_S(n-k,\varepsilon)\}\\
\overset{n'=n-k}{=}&\max_{k\geq 0}\max_{n'\geq 0}\{\alpha_A(n'+m)-S(k,k+n')+\alpha_S(n',\varepsilon)\}\\
=&\max_{k\geq 0}\max_{0\leq n'\leq N_m}\{\alpha_A(n'+m)-S(k,k+n')+\alpha_S(n',\varepsilon)\}\\
=&\max_{0\leq k\leq \bar{n}}\max_{0\leq n'\leq N_m}\{\alpha_A(n'+m)-S(k,k+n')+\alpha_S(n',\varepsilon)\}
\end{aligned}$$

Here we have used the stability condition in the fourth line. The last line follows from the simple observation that the set in $\max_{k\leq n'\leq N_m}\{\alpha_A(n+m-k)-S(k,n)+\alpha_S(n-k,\varepsilon)\}$ is empty for $k>N_m$ and hence can not contribute to the maximum. The index set on the last line is finite and hence the maximum is finite and more importantly a maximizing index n_0 can be found as needed. From our initializing assumption we derived henceforth inequality (3.3). 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 to 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 3.16. S is *stable* (for some ε) if

$$S(m,m+n')-\alpha_S(n',\varepsilon)-\alpha_A(N+n',\varepsilon)\xrightarrow{n'\to\infty}\infty$$

holds for all $N\geq 0$ and $m\geq 0$ and

$$S(m,m+n')-\alpha_S(n',\varepsilon)>0$$

holds for all $m\geq 0$ and $n'\geq 0$.

It is easy to see, that S being stable implies that S is delay-stable and output-stable.

Bibliography

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