Aalborg Universitet

# Worst case queue length estimation in networks of multiple token bus segments

Schiøler, Henrik; Nielsen, N.N.; Nielsen, Jens Frederik D.; Jørgensen, N.

*Publication date:*
1998

*Document Version*
Også kaldet Forlagets PDF

# WORST CASE QUEUE LENGTH ESTIMATION IN NETWORKS OF MULTIPLE TOKEN BUS SEGMENTS.

Henrik Schiøler M.Sc.EE, Ph.D, Niels Nørgaard Nielsen
M.Sc.EE, Jens Dalsgaard Nielsen M.Sc.EE, Ph.D and Niels Jørgensen M.Sc.EE
Dept. of Control Engineering
Inst. of Electronic Systems
Aalborg, Denmark

**Abstract**

The paper describes a a novel methodology for estimating worst case queue lengths in real time communication networks of connected segments each of token bus type. A general non probabilistic traffic model is suggested, which is believed to be applicable even where a probabilistic model traditionally is the first choice, e.g. alerting and user interface handling. Completion time theory is provided justifying the correctness of an iterative estimation scheme for worst case queue lengths, and numerical results illustrating the virtues of the algorithm are presented.

## 1 Introduction

A variety of local area network technologies are suggested to serve communication medium in hard real time environments such as distributed proces control systems, where hard bounds on communication delays are required to guarantee the functional correctness of e.g. closed loop control across the network and high level system diagnostics. Among the more recent suggestions is TTP [1] where a TDMA scheme quarantees a fixed bandwith to each node, a tight clock synchronization between nodes as well as direct information about liveliness and internal state. In [3] and [5] approximate network preemptivness as well as prioritization among nodes is assumed in order to facilitate the translation of results from scheduling theory ([?], [2] and [?]) to real time network communication. In this paper an existing network technology is analyzed w.r.t. real-time properties. We consider a network, where a number of token bus segments are connected by a number of gateways in an arbitrary configuration as depicted in fig. 1

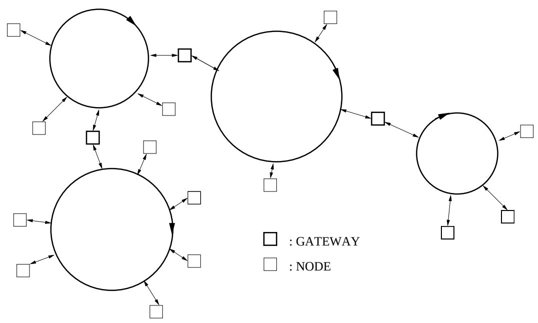


Figure 1: *Overall network topology*

Throughout the network different patterns of traffic are mixed, so that both periodic and aperiodic sources of traffic reside on the same segments. Periodic traffic is for example generated by sensors transmitting measured values for logging, supervision, control and presentation, whereas aperiodic traffic might be commands from the operator to machinery level or alerting in the opposite direction.

When a message is generated in some node it initially is placed in some queue of outgoing messages served according to a FIFO discipline, in turn it is transmitted to another node on the same segment. If the receiving node is a gateway to another segment the message is statically routed forwards and put into a queue of outgoing messages of the gateway. Thus to reach the end receiver the message has to travel across a number of network segments. Since outgoing messages have to advance through queues of outgoing messages transmission delays and queue lengths are closely linked. High transmission delays can be accomodated for w.r.t. alert messages or errors reported from low level error detection, by introducing priority queuing in nodes generating traffic of different priority as well as in gateways. Oppositely to facilitate timing related system diagnostics on a high level, it is of major importance that hard bounds (worst case estimates) on transmission delay can be established even for regular low priority traffic.

From control theory it is well known that in-loop delay might introduce a serious deterioration of performance and robustness to closed loop control systems or it may even destabilize the loop. If closed loop control is performed across the network it is therefore of vital importance that only a predictable and bounded delay is introduced by the network. In this paper a general methodology for establishing worst case bounds on transmission delays and queue sizes in a network of connected segments of token busses.

The paper is organized as follows. First a non probabilistic traffic model is sugggested, which is a generalization of a pure periodic traffic model. The validity of the model is briefly argued. Then a completion time theory for unprioritized token bus networks is developed identifying an extreme scenario regarding transmission instants and in turn queue lengths. The extreme scenario is subsequently utilized in an estimation scheme, where worst case queue lengths are minimized iteratively. Next a numerical example is given illustrating the main virtues of the method. Worst case estimation of transmission delays is discussed and the case of multiple segments is treated. Finally conclusive remarks on perspectives and future research are made.

## 2 Modelling traffic sources.

From the Poisson assumption traditionally made for aperiodic traffic, traffic of any size is possible within intervals of arbitrary shortness. Thus Poisson traffic does not allow for worst case analysis and is therefore not considered feasible for analysis of hard real time systems. Therefore traffic sources in hard real time systems are assumed to produce only a bounded traffic within time intervals below a certain minimum length. Traffic $R_i(I_T)$ received within a time interval $I_T$ in node $i$ is assumed to be bounded by

$$R_i(I_T) \leq \bar{R}_i(|I_T|) = C_i + \alpha_i \cdot (|I_T|) \tag{1}$$

Assume that a lowest interarrival time $T_i$ as well as a maximum message size $c_i$ can be identified for some source then (1) is fulfilled for $\alpha_i = \frac{c_i}{T_i}$ and $C_i = c_i$. Likewise if a number of such sources characterized by parameters $C_j$ and $\alpha_j$ $j \in J_i$ produce traffic to the same node $i$ the total received traffic is bounded by

$$R_i(I_T) \leq \bar{R}_i(|I_T|) = \sum_{j \in J_i} C_j + (\sum_{j \in J_i} \alpha_j) \cdot |I_T| \tag{2}$$

Thus the above bound is invariant to convergence of traffic sources.
Next consider some queue subject to incomming traffic characterized by $\alpha$ and $C$. If it is guarateed that the queue length will not exceed a value of $K$, the outgoing traffic is characterized by $\alpha$ and $C + K$. Thus the form (1) is invariant to queueing of bounded queue length.
The bound on received traffic by an afine function, even for aperiodic traffic, is illustrated by the following 2 examples

- When an alarm goes off it will at least take the minimum repair period before an alarm from the same source goes off again (a firealarm should not be transmitted twice before the fire has been put out). Thus a minimal period can be identified.
- Command messages should not be transmitted closer than the biological reaction time of the human initiating them, once again yielding a minimal period of transmission.

## 3 Completion time theory for token bus networks.

Based on the traffic model derived in the previous section a non probabilistic framework for estimating queue lengths and in turn transmission delays is developed in the sequel. The traffic model of the previous section can be viewed as a generalization of pure periodic traffic models and includes as well aperiodic traffic, where a minimum interarrival time can be identified. This work is inspired by the basic results from scheduling theory and especially the critical instant theorem of [?], stating that for a periodic task set, where execution times and periods are fixed, the worst case situation is when all tasks are reported ready simultaneously. In [5] the results from scheduling theory are applied to a shared communication channel of ring topology with periodic traffic sources. The basic assumptions regarding prioritization among tasks and preemptivness are translated into prioritization among nodes and a fixed small packet size respectively. Both are bound to introduce a significant overhead to the system and in turn decreasing the data efficience of the network. Instead we consider token bus networks without priorities and with packet lengths of arbitrary lengths.
For simplicity token passing overhead is assumed to be negligible in the following. Likewise nodes are assumed to transmit according to a bounded and gated unprioritized FIFO discipline, where the traffic residing at a queue at token capture time is transmitted up to a certain maximum network packet size. We consider an extreme situation where maximum traffic is received from a common critical instant. This situation is compared to a set of sitations where traffic is assumed to be received in packages of fixed size below the maximum mentioned above. First consider an instant $t_0$ where queue lengths $Q_i(t_0),\ i = 1, 2, .., N$ are not all zero, that is

$$SQ(t_0) = \sum_{i=1}^{N} Q_i(t_0) > 0 \tag{3}$$

Assume then for any $t > t_0$ that $SQ(t) > 0$. When only one queue is not empty, all time is spent transmitting so that

$$SQ(t) - SQ(t_0) = \sum_{i=1}^{N} R_i([t_0, t]) - (t - t_0) \tag{4}$$

$$\leq \sum_{i=1}^{N} c_i + (\sum_{i=1}^{N} \alpha_i - 1) \cdot (t - t_0) \tag{5}$$

Thus if $\sum_{i=1}^{N} \alpha_i < 1$ , $SQ(t)$ needs to cross zero for some $t > t_0$. More specifically, because all queue lengths initially are zero $SQ(t) \leq \sum_{i=1}^{N} c_i$ and the time for reaching zero obeys

$$t - t_0 \leq T_z = \frac{\sum_{i=1}^{N} c_i}{1 - \sum_{i=1}^{N} \alpha_i} \tag{6}$$

Thus all queue lengths can at most be away from zero for a time interval of length $T_z$.
Let $\mathcal{S}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$ denote the set of scenarios over $t \geq t_0$ where

- $Q_k(t_0) = 0$
- $Q_i(t_0) \leq \bar{Q}_i$ for $i \neq k$
- $R_i([t_0, t]) \leq \bar{R}_i(t - t_0),\ i = 1, 2, .., N$
- Node $h$ is the first node to transmit after $t_0$

and let $\bar{\mathcal{S}}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$ be the one scenario in $\mathcal{S}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$, where

- $Q_i(t_0) = \bar{Q}_i$ for $i \neq k$
- $R_i([t_0, t]) = \bar{R}_i(t - t_0),\ i = 1, 2, .., N$

Next we define the following symbols

- $succ : N \to N$ is the usual successor function according to the ring topology, i.e.

$$\begin{aligned} succ(i) &= i+1 \;\text{ for }\; i < N \\ succ(N) &= 1 \end{aligned} \tag{7}$$

- $pred : N \to N$ is the inverse of $succ$
- $SUCC : (1..N)^3 \to (1..N)^2$ is a three place successor function defined by

$$\begin{aligned} SUCC(h,i,j) = (succ(i), j) \;\text{ for }\; succ(i) \neq h \\ SUCC(h,i,j) = (succ(i), j+1) \;\text{ for }\; succ(i) = h \end{aligned} \tag{8}$$

  That is if $(i,j)$ denotes quantities associated to the $j$. th token capture of node $i$ in $\mathcal{S}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$ then $SUCC(h,i,j)$ relates to the token capture following immediately after.

- $S_{i,j}$: Amount of traffic transmitted from the $i$ th. node between its $j$ th. token- capture and -release after $t_0$
- $t_{i,j}$: Time for $j$ th. token capture, i.e.

$$t_{SUCC(h,i,j)} = t_{i,j} + S_{i,j} \tag{9}$$

- $T_{i,j}$: Total amount of traffic transmitted from the $i$ th. node after its $j$ th. token release after $t_0$, i.e.

$$T_{i,j} = T_{i,j-1} + S_{i,j} \tag{10}$$

The intiuitive locations of the above symbols are depicted in fig. 2. All quantities denoted $\bar{}$ are defined to relate

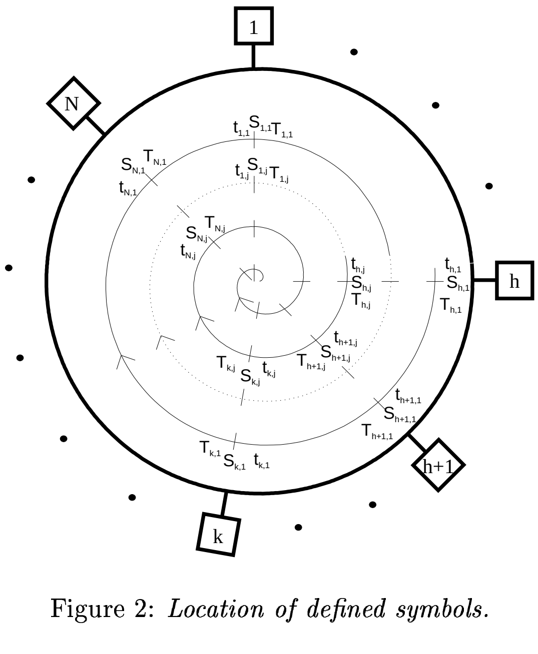


Figure 2: *Location of defined symbols.*

to $\bar{\mathcal{S}}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$.
Now $\bar{t}_{h,1} = t_{h,1} = t_0$ and because $\bar{Q}_h \geq Q_h(t_0)$, $\bar{S}_{h,1} \geq S_{h,1}$ and in turn $\bar{t}_{succ(h),1} \geq t_{succ(h),1}$ implying

$$\begin{aligned} \bar{R}_{succ(h)}(\bar{t}_{succ(h),1} - t_0) &\geq R_{succ(h)}([t_0, t_{succ(h),1}]) \text{ and} \\ \bar{S}_{succ(h),1} &\geq S_{succ(h),1} \end{aligned} \tag{11}$$

The reasoning follows the same line all trough the first token rotation. Now assume that for some token rotation $l$ and some node $n$ that

$$\bar{t}_{i,j} \geq t_{i,j},\ i = 1..N,\ j < l \tag{12}$$
$$\bar{t}_{i,j} \geq t_{i,j},\ i = h, .., n,\ j = l \tag{13}$$
$$\bar{T}_{i,j} \geq T_{i,j},\ i = 1..N,\ j < l \tag{14}$$
$$\bar{T}_{i,j} \geq T_{i,j},\ i = h, .., n,\ j = l \tag{15}$$
$$\tag{16}$$

and further that $\bar{T}_{SUCC(h,n,l)} < T_{SUCC(h,n,l)}$. This implyes that $\bar{S}_{SUCC(h,n,l)} < S_{SUCC(h,n,l)}$, which can only happen when the queue is exhausted in the former situation, i.e.

$$\bar{Q}_{succ(n)}(\bar{t}_{SUCC(h,n,l)}) - \bar{S}_{SUCC(h,n,l)} = 0 \tag{17}$$

Up until the $l$ th. token capture of node $succ(n)$ all nodes transmitted more in the first situation, that is $\bar{T}_{i,j} \geq T_{i,j}$ so that $\bar{t}_{SUCC(h,n,l)} \geq t_{SUCC(h,n,l)}$ and in turn $\bar{R}_{succ(n)}(\bar{t}_{SUCC(h,n,l)} - t_0) \geq R_{succ(n)}([t_0, t_{SUCC(h,n,l)}])$ so that

$$\begin{aligned} 0 &= \bar{Q}_{succ(n)}(\bar{t}_{SUCC(h,n,l)}) - \bar{S}_{SUCC(h,n,l)} \\ &= \bar{Q}_{succ(n)} + \bar{R}_{succ(n)}(\bar{t}_{SUCC(h,n,l)} - t_0) - \bar{T}_{SUCC(h,n,l)} \\ &> Q_{succ(n)}(t_0) + R_{succ(n)}([t_0, t_{SUCC(h,n,l)}]) - T_{SUCC(h,n,l)} \\ &= Q_{succ(n)}(t_{SUCC(h,n,l)}) - S_{SUCC(h,n,l)} \end{aligned} \tag{18}$$

which is not possible, because it would mean that node $succ(n)$ should transmit more than it has queued up at $t_{SUCC(h,n,l)}$. All together it is proved by the 2. principle of induction that

$$\bar{t}_{i,j} \geq t_{i,j},\ i = 1, .., N,\ \forall j \geq 1 \tag{19}$$
$$\bar{T}_{i,j} \geq T_{i,j},\ i = 1, .., N,\ \forall j \geq 1 \tag{20}$$

Thus $\bar{\mathcal{S}}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$ is extreme in the above sence within $\mathcal{S}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$. Next we show that $\bar{\mathcal{S}}_{succ(k),k}(\bar{Q}_1, ..\bar{Q}_N)$ is extreme for node $k$ among all possible values of $h$. In this case

$$\bar{t}_{k,j} \geq t_{k,j},\ \forall j \geq 1 \tag{21}$$
$$\bar{T}_{k,j} \geq T_{k,j},\ \forall j \geq 1 \tag{22}$$

We now redefine notation slightly so that $^-$ relates to $\bar{\mathcal{S}}_{h,k}(\bar{Q}_1, ..\bar{Q}_N)$ and quantities without $^-$ relates to $\bar{\mathcal{S}}_{succ(h),k}(\bar{Q}_1, ..\bar{Q}_N)$.
To easen up notation let $m = succ(h)$. Then $\bar{t}_{h,1} = t_0$ and $\bar{t}_{m,1} = \bar{t}_{h,1} + \bar{S}_{h,1} \geq t_0 = t_{m,1}$. Consequently $\bar{R}_{m,1}(\bar{t}_{m,1} - t_0) \geq R_{m,1}([t_0, t_{m,1}])$, so that $\bar{S}_{m,1} \geq S_{m,1}$ and in turn $\bar{t}_{succ(m),1} \geq t_{succ(m),1}$. The argument follows similarly all through the first token rotation, so that $\bar{t}_{pred(h),1} \geq t_{pred(h),1}$ and $\bar{S}_{pred(h),1} \geq S_{pred(h),1}$ so that $\bar{t}_{h,2} \geq t_{h,1}$.
Assume for the sake of contradiction that $\bar{T}_{h,2} < T_{h,1}$. Then $\bar{S}_{h,2} < S_{h,1}$, which can only happen if $\bar{Q}_h(\bar{t}_{h,2}) - \bar{S}_{h,2} = 0$. Now

$$\begin{aligned} 0 &= \bar{Q}_h(\bar{t}_{h,2}) - \bar{S}_{h,2} \\ &= \bar{Q}_h + \bar{R}_h(\bar{t}_{h,2} - t_0) - \bar{T}_{h,2} \\ &> Q_h(t_0) + R_h([t_0, t_{h,1}]) - T_{h,1} = \\ &= Q_h(t_{h,1}) - S_{h,1} \end{aligned} \tag{23}$$

which is a contradiction as stated previously. Now assume generally that

$$\bar{t}_{i,j} \ \geq \ t_{i,j},\ i = 1..N,\ i \neq h,\ j < l \tag{24}$$

$$\begin{aligned}
\bar{t}_{h,j+1} &\geq t_{h,j} \; j < l & (25)\\
\bar{t}_{i,j} &\geq t_{i,j}, \; i = succ(h), .., n, \; j = l & (26)\\
\bar{T}_{i,j} &\geq T_{i,j}, \; i = 1..N, \; i \neq h, \; j < l & (27)\\
\bar{T}_{h,j+1} &\geq T_{h,j} \; j < l & (28)\\
\bar{T}_{i,j} &\geq T_{i,j}, \; i = succ(h), .., n, \; j = l & (29)\\
& & (30)
\end{aligned}$$

and for the sake of contradiction that $\bar{T}_{SUCC(h,n,l)} < T_{SUCC(m,n,l)}$, then as before $\bar{S}_{SUCC(h,n,l)} < S_{SUCC(m,n,l)}$ which implyes that $\bar{Q}_{succ(n)}(\bar{t}_{SUCC(h,n,l)}) - \bar{S}_{SUCC(h,n,l)} = 0$ so that

$$\begin{aligned}
0 &= \bar{Q}_{succ(n)}(\bar{t}_{SUCC(h,n,l)}) - \bar{S}_{SUCC(h,n,l)}\\
&= \bar{Q}_{succ(n)} + \bar{R}_{succ(n)}(\bar{t}_{SUCC(h,n,l)} - t_0) - \bar{T}_{SUCC(h,n,l)}\\
&> Q_{succ(n)}(t_0) + R_{succ(n)}([t_0, t_{SUCC(m,n,l)}]) - T_{SUCC(m,n,l)}\\
&= Q_h(t_{SUCC(m,n,l)}) - S_{SUCC(m,n,l)}
\end{aligned} \tag{31}$$

Once again by the second principle of induction

$$\begin{aligned}
\bar{t}_{i,j} \geq t_{i,j}, \; i = 1..N, \;\; i \neq h \forall j \geq 1\\
\bar{T}_{i,j} \geq T_{i,j}, \; i = 1, .., N, \; i \neq h \forall j \geq 1
\end{aligned} \tag{32}$$

From the above it is concluded that when the starting node $h$ is rotated one location in token direction, then $t_{k,j}$ and $T_{k,j}$ becomes smaller for $k \neq h$. This proces can be repeated from $h = succ(k)$ until $h = k$ and all steps reduce values for $t_{k,j}$ and $T_{k,j}$. Thus $h = succ(k)$ is extreme.
Now it is assumed that traffic arrives at node $k$ as complete network packages, that is if $Q_k(t_{k,j}) > 0$ then $S_{k,j} = P$, where $P$ is the size of one network package from node $k$. Then from $t_0$ until some later instant $t_z$, where $Q_k(t_z) = 0$

$$\begin{aligned}
\bar{Q}_k(\bar{t}_{k,j}) &= \bar{R}_k(\bar{t}_{k,j} - t_0) - \bar{T}_{k,j-1} & (33)\\
&\geq \bar{R}_k(\bar{t}_{k,j} - t_0) - (j-1) \cdot P\\
&\geq R_k([t_0, t_{k,j}]) - (j-1) \cdot P = Q_k(t_{k,j})
\end{aligned}$$

Inequality (33) reveals that $Q_k(t_{k,j})$ will not exceed $\bar{Q}_k(\bar{t}_{k,j})$. Therefore $\bar{S}_{succ(k),k}(\bar{Q}_1, ..\bar{Q}_N)$ is extreme also w.r.t. queue lengths among all possible values of $h$.

## 4 Iterative method for estimating maximum queue lengths.

Since $\bar{Q}(\bar{t}_{k,j})$ bounds $Q(t_{k,j})$ for all numbers of token rotation $j \geq 1$ the maximum queue length $\bar{Q}_k$ in node $k$ can be computed by

$$\bar{Q}_k = \max_{\bar{t}_{j,k} \in [t_0 t_z]} \bar{Q}_k(\bar{t}_{k,j}) \tag{34}$$

Therefore if bounds $\bar{Q}_1, .., \bar{Q}_N$ on queue lengths are known along with maximum traffic bounds $c_1, .., c_N$ and $\alpha_1, .., \alpha_N$ it is possible to infer new and hopefully lower bounds on queue lengths by computing $\bar{Q}_k \; k = 1, .., N$ using the previous bounds. The following algorithm computes transmission times, transmitted amounts of traffic and queue lengths for all nodes in the situation
$\bar{S}_{succ(k),k}(\bar{Q}_1, ..\bar{Q}_N)$.

```
while(sum(queue_length) < PACKET_SIZE)
   node=k;
   for i=1:NUM_NODES
```

```
        node=NEXT(node);
        rec=c(node)+alfa(node)*time;
        queue_length(node)=Q(node)+rec-sent(node);
        tr=min(PACKET_SIZE,queue_length(node));
        sent(node)=sent(node)+tr;
        time=time+tr;
    end

    rec=c(k)+alfa(k)*time;
    queue_length(k)=rec-sent(k);
    tr=min(PACKET_SIZE,queue_length(k));
    sent(k)=sent(k)+tr;
    time=time+tr;
end
```

If the map $\Gamma$ is defined by

$$\Gamma : \bar{Q}_1, .., \bar{Q}_N, c_1, .., c_N, \alpha_1, .., \alpha_N \mapsto \bar{Q}_1, .., \bar{Q}_N \tag{35}$$

we can construct a non increasing sequence $\bar{Q}^1, \bar{Q}^2, ..$ by

$$\begin{aligned} & (\bar{Q}_1^{p+1}, .., \bar{Q}_N^{p+1}) \\ = \; & \min(\bar{Q}_1^p, .., \bar{Q}_N^p \\ , \; & \Gamma(\bar{Q}_1^p, .., \bar{Q}_N^p, c_1 .., c_N, \alpha_1, .., \alpha_N)) \end{aligned} \tag{36}$$

As a starting point of the iteration (36) the common value $\sum_{i=1}^N c_i$ can be chosen so that

$$\bar{Q}_k^0 = \sum_{i=1}^N c_i \; k = 1..N \tag{37}$$

since this value is known to bound the total amount of queued traffic it especially bounds traffic queued up on a single node.

## 5 Numerical example.

As a numerical example results are presented for a token bus connecting 7 nodes with periodic traffic. Traffic characteristics of the node set are shown in table 5

| Node | T | C | P | $\bar{Q}$ | $\underline{Q}$ | $Td^*$ |
|---|---|---|---|---|---|---|
| 1 | 1000 | 200 | 100 | 400 | 400 | 2500 |
| 2 | 2000 | 500 | 100 | 1000 | 700 | 5500 |
| 3 | 4000 | 1700 | 100 | 4200 | 2600 | 15800 |
| 4 | 8000 | 200 | 100 | 200 | 200 | 1300 |
| 5 | 10000 | 100 | 100 | 100 | 100 | 600 |
| 6 | 20000 | 1000 | 100 | 1000 | 1000 | 5500 |
| 7 | 40000 | 500 | 100 | 500 | 500 | 3000 |

From table 5 we get $\sum_{i=1}^N \alpha_i = 0.9725$, which guarantees finite time completion of the algorithm, and $\sum_{i=1}^N c_i = 4200$ which defines the starting point of the iteration (36). In the $\bar{Q}$ column the final worst case estimates of queue lengths. Every node $j$ for which

$$\frac{C_j}{T_j} < \frac{P_j}{\sum_{i=1}^N P_i} - 1 \tag{38}$$

can maximally build up a queue length of $c_j$, since even for maximum token turn around time each periodic traffic reception will be transmitted before the expiry of the associated period. This basic fact has certainly been captured by the algorithm as can be seen from table 5. The figures listed in column $\underline{Q}$ indicate maximum queue sizes for a situation, where all queues initially are zero and all nodes receive maximum traffic from the instant, where node $j$ just released token. This situation is intuitively critical (extreme), and is assumable, that is it may occur. The worst case results on the other hand may or may not be assumable, so that if they are not, the worst case estimates will be conservative. If some maximum queue length criterion $ML_j$ is defined then

$$\bar{Q}_j \leq ML_j \tag{39}$$

guarantees proper operation with respect to queue lengths. In other words the worst case estimates supply a sufficient condition. On the other hand if

$$\underline{Q}_j \geq ML_j \tag{40}$$

there exists a situation, where the queue length criterion is not fulfilled, so that the figures in the $\underline{Q}$ column define a necessary condition. Finally if $ML_j \in [\underline{Q}_j, \bar{Q}_j]$ no precise statement can be made.

## 6 Estimating worst case transmission delays.

Two different definitions for transmission delay of some message $M$ of size $L$ can be made; $Td^{**}$ is the time from receiving the first bit of $M$ in the transmitting node $i$ until the last bit of $M$ reaches the receiving node $j$, or alternatively $Td^*$ is the time from receiving the last bit of $M$ in $i$ until the last bit of $M$ reaches $j$. We consider only delays according to the latter definition. Now following the reasoning of the former section an extreme situation can be identified; the time from node $i$ receives the last bit of $M$ until the last bit is transmitted is largest if it is received exactly when $i$ releases token and in that instant all queue lengths are maximum and from that instant all queues receive maximum traffic. Computing the transmission delay in this situation yields a worst case estimate of $Td^*$ when ignoring cable delays. Worst case estimates of $Td^*$ associated with the example in the previous section are shown in table 5.

## 7 Multi segment analysis.

For an entire network consisting of a number of segments each of ring topology, the previous consideration only carry over to a certain extend, that is if the overall network topology is free of directed cycles. A directed cycle is, in this context, a sequence of segments $S_1, .., S_M$ so that traffic travels from $S_i$ to $S_{i+1}$ for $i < M$ and from $S_M$ to $S_1$, as depicted for $M = 2$ in fig. 3. When the network is free of such cycles there must be at least one segment that does not receive traffic from other segments. The analysis of the previous section is initially applied to such *leaf* segments. If a gateway $G$ in such a segment $S$ receives traffic from nodes $j \in J$ described by parameters $c_j$ and $\alpha_j$ $j \in J$ and the queue length analysis is carried out on $S$ producing estimates $\bar{Q}_j$ $j \in J$, then $G$ is described by parameters $\sum_{j \in J}(c_j + \bar{Q}_j)$ and $\sum_{j \in J} \alpha_j$. Subsequently there is at least on segment receiving data only from leaf segments. The queue length analysis is then performed on such segments using the parameters describing gateways allready obtained from the analysis of leaf segments. The analysis can be carried on in this way all the way to a number of top segments which do not transmit traffic to other segments.
When cycles appear, as shown in fig. 3, the following problem arises; the parameters describing gateway 1 effects the analysis and in turn the queue length estimates of segment $B$, which effects the parameters describing gateway 2, which effects the analysis of segment $A$ and in turn the parameters of gateway 1. One could imagine an iterative scheme of analysis alternating between segments $A$ and $B$. The conditions for such a scheme to be stable remain yet to be found.

Instead an alternative appoach is considered based on the following assumption concerning the gateways of the network; a gateway between segments $A$ and $B$ is assumed to consist of 2 queues as shown in fig 4. One queue $q_A$ is receiving traffic from segment $A$ and one queue $q_B$ is transmitting traffic to segment $B$. Traffic is transferred

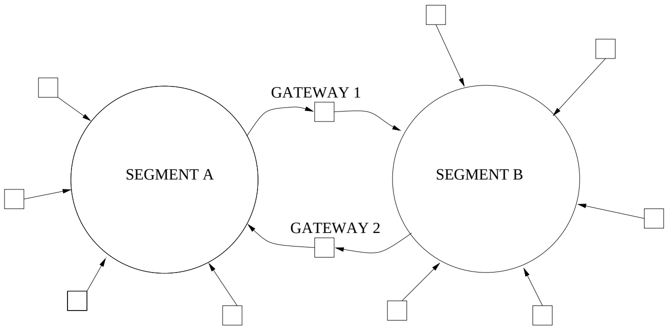


Figure 3: *Network with a directed cycle.*

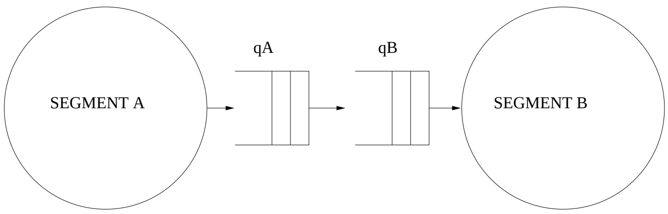


Figure 4: *Gateway divided into separate queues por decoupling segments.*

from $q_A$ to $q_B$ periodically so that a traffic amount of $c$ is transferred each $T$ time units. The parameters $c$ and $T$ fulfill the following relationships

$$\frac{c}{T} = \alpha \geq \sum_{j \in J} \alpha_j \tag{41}$$

where $J$ denotes the set of nodes in segment $A$ transmitting traffic through the gateway. In this way parameters describing each gateway remain fixed and the queue length analysis of each segment can be performed independently. Assume for some instant $t_0$ that $Q_A(t_0) < c$ and $Q_A(t) > c$ for $t > t_0$. During the interval $[t_0, t]$ the receive queue $q_A$ received maximally $\sum_{j \in J}(c_j + \bar{Q}_j + \alpha_j \cdot (t - t_0))$. Because $Q_A(t) > c$, after $t_0$, minimally $\alpha \cdot (t - t_0) - c$ is transferred to the transmit queue $q_B$ in $[t_0, t]$. Consequently $Q_A$ is bounded by

$$\begin{aligned} Q_A(t) &\leq \sum_{j \in J}(c_j + \bar{Q}_j + \alpha_j \cdot (t - t_0)) + c - \alpha \cdot (t - t_0) + c \\ &\leq \sum_{j \in J}(c_j + \bar{Q}_j) + 2 \cdot c \end{aligned} \tag{42}$$

## 8 Summary

A theoretical and operational framework for estimating worst case queue lengths and transmission delays in networks of connected token bus segments is presented. A non probabilistic traffic model is presented which bounds received traffic below an afine function. The validity of the model even for traffic sources traditionally modelled in a probabilistic framwork is argued by examples. On the basis of the assumed traffic bounds a critical or extreme scenario regarding transmission instants and queue lengths is identified. The extreme scenario is utilized in an iterative algorithm, which, on the basis of a known set of worst case queue length estimates, computes new and hopefully less conservative estimates. A valid starting point of the algorithm is found and numerical results illustrating the virtues of the algorithm are presented. Multiple segments are treated under a non restricitve assumption concerning the interior operation of the gateways connecting segments. The generality of the suggested traffic model yeilds hope for the entire methodology to by broadly applicable even outside the area of marine automation where it was originally developed.

## References


[1] Herman Kopetz and Gunther Grunsteidl, TTP - A Protocol for Fault-Tolerant Real-Time Systems.," *IEEE Computer*, pp. 14-23, 1994.

[2] J. P. Lehoczky, Fixed Priority Scheduling of Periodic Task Sets with Arbitrary Deadlines.," *Proceedings of the 11th IEEE Real-Time Systems Symposium*, pp. 201-209, 1990.

[3] J. P. Lehoczky and L. Sha, Performance of Real-Time Bus Scheduling Algorithms," *ACM Performance Evaluation Review, Special Issue 14(1)*, 1986.

[4] C. L. Liu and J. W. Layland, Scheduling Algorithms for Multiprogramming in a Hard Real Time Environment.," *JACM*, Vol. 20, pp. 46-61, 1973.

[5] J.K Strosnider and T. Marchok, Responsive Deterministic IEEE 802.5 Token Ring Scheduling," *The Journal of Real-Time Systems*, Vol. 1, pp. 133-158, 1989.