Chapter 1

Polling Networks

Each station on the network is polled in some predetermined order. When polled, a station uses the full data rate of the connecting channel to transmit its backlog of packets to the central computer. Between polls, stations accumulate messages in their queues, but do not transmit until they are polled.

Transmissions between stations take place through the central computer, which receives all incoming packets and transmits them to the appropriate locations.
1.1 Polling Mechanisms

Roll-Call Polling

Each station has to be polled in turn by the central computer (controller). After the station has transmitted its backlog of messages, it notifies the central controller with a suffix to its last packet. After receiving this suffix packet, the controller sends a poll to the next station in the polling sequence.

Hub Polling

In this case the go-ahead (suffix) packet contains the next station address. A monitoring channel must be provided to indicate to the appropriate station that it should start transmitting. Essentially the go-ahead is transmitted directly from one station to another.

1.2 Roll-Call Polling Operation

Station

- The central computer sends out a polling packet to station $i$ in the polling sequence.
• Station $i$ synchronises on bits and characters.

• Station $i$ reads and interprets the station address and the go-ahead contained in the polling packet.

• Station $i$ transmits all its backlogged messages to the central computer for distribution to the central computer and other stations.

• Station $i$ appends a go-ahead, and possibly a next-station address to its last packet.

**Controller**

• The central computer synchronises on bits and characters.

• The central computer reads and interprets the incoming packets, including the final go-ahead and next station address.

• The central computer sends out a poll to station $(i + 1)$.

These steps are repeated for each station. When all stations have been polled the whole sequence begins again.

The communications between the controller and stations may be half-duplex or duplex. We assume duplex.

**Transmission from Controller to Stations:** This can be asynchronous or continuous. If continuous then step 2 of station operation can be ignored.

**Transmission from Station to Controller:** This is asynchronous. If a station has no messages it either does not reply or sends back a go-ahead packet.
1.3 Polling Analysis Model

Messages arrive at random at each station and are stored in a buffer until the station is polled.
When polled, the station transmits all packets in its buffer and finishes by sending a go-ahead packet.
The next station is then polled on a roll-call or hub basis.
A walk time \( w \) is required to transfer the poll from one station to another.

1.3.1 Model Assumptions

In determining the performance, the following assumptions are made:

- Each station has the same Poisson arrival statistics with average arrival rate \( \lambda \) packets/second.
- The walk time \( w \) between stations is constant and the same for every consecutive station.
The channel propagation times between stations are equal and are included in the walk time.

Packet length distributions are the same for packets arriving at each station.

1.3.2 Average Cycle Time

Let:

\( N_m \) = Average number of packets stored at a station when poll arrives

\( X \) = Average length of packet (in bits)

\( R \) = Speed of transmission channel (in bits per second)

\( N_m X / R \) = Time needed to empty station buffer (in seconds)

The next stations starts transmitting after the walk time \( w \).

Length of average cycle time =

\[
T_c = M \left[ N_m X / R + w \right]
\]

where \( M \) is the number of stations.

From the input statistics:

\( N_m = \lambda T_c \)

Combining this with the previous equation to eliminate \( N_m \) gives:

\[
T_c = \frac{M w}{1 - M \lambda X / R}
\]

Define throughput \( S = M \lambda X / R \)

Therefore:

\[
T_c = \frac{M w}{1 - S} \text{ seconds}
\]

\( S \) must be < 1 in order to have finite buffer lengths at stations.
1.3.3 Delay Analysis

Let $W$ = average time an arriving packet must wait before reaching the head of the queue.

This delay is made up of two components:

1. The waiting delay $(W_1)$ in the station buffer while other stations are being served,

2. The waiting delay $(W_2)$ in the station buffer while that particular station is being served.

$$W = W_1 + W_2$$

![Division of Waiting Times for a Typical Packet](image)

**Figure 1.3: Division of Waiting Times for a Typical Packet**

**For $W_1$**

The average number of packets that are transmitted over channel from station $=$ $\lambda T_c$.

Average service time $= \frac{\lambda T_c \bar{X}}{R}$

Define $\rho = \frac{\lambda \bar{X}}{R} = \frac{S}{M}$

Service time per station $= \rho T_c$

Remaining part of average cycle during which the station is idle is:

$$T_c(1 - \rho)$$

Random arrivals occur in the interval $T_c(1 - \rho)$
On average the delay is:

\[
\frac{T_c(1 - \rho)}{2} \text{ seconds}
\]

Using the previous expression for \(T_c\):

\[
W_1 = \frac{Mw(1 - \rho)}{2(1 - M\rho)}
\]

For \(W_2\)

An approximation is used. Consider an equivalent network for which there is no walk time so that there is always some station being served if there are packets in the network.

Think of the whole system as having input \(M\lambda\) and only a single server. An \(M/G/1\) model will apply.

The delay in a buffer for an \(M/G/1\) system is:

\[
W' = \frac{\lambda'}{2(1 - \rho')} E[\tau^2]
\]

where \(E[\tau^2] = \sigma^2 + 1/\mu^2\) is the 2nd moment of the service time distribution.

In our case

\[
\begin{align*}
\lambda' &= M\lambda \\
\rho' &= M\rho = S \\
E[\tau^2] &= \frac{\left[ \frac{X}{R} \right]^2}{\frac{X^2}{R^2}}
\end{align*}
\]
Therefore
\[ W_2 = \frac{M \lambda (\bar{X}^2 / R^2)}{2(1 - S)} = \frac{M \lambda \bar{X}}{R} \cdot \frac{1}{2(1 - S)} \cdot \frac{1}{XR} \bar{X}^2 \]

As \( S = M \lambda \bar{X} / R \)
\[ W_2 = \frac{S \bar{X}^2}{2XR(1 - S)} \]

Total delay is \( W_1 + W_2 \) so
\[ W = \frac{Mw(1 - \rho)}{2(1 - M\rho)} + \frac{S \bar{X}^2}{2XR(1 - S)} \]

Rewriting, the average time a packet must wait before reaching the head of the queue in a polling network is:

\[
W = \frac{Mw(1 - S/M)}{2(1 - S)} + \frac{S \bar{X}^2}{2XR(1 - S)}
\]

Two distributions of packet lengths are of interest:

**Constant:** \( \bar{X}^2 = (\bar{X})^2 \)
\[
W = \frac{Mw(1 - S/M)}{2(1 - S)} + \frac{S \bar{X}}{2R(1 - S)}
\]

**Exponential:** \( \bar{X}^2 = 2(\bar{X})^2 \)
\[
W = \frac{Mw(1 - S/M)}{2(1 - S)} + \frac{SX}{R(1 - S)}
\]
1.4 Polling Network Example

Consider a metropolitan area network with a single central processor located at the headend of a broadband CATV system that has a tree topology. The following are specified:

- Maximum distance from headend to a subscriber station - 20 km
- Access technique - roll-call poling
- Length of polling packet - 8 bytes
- Length of go-ahead - 1 byte
- Data rate of channel - 56 kbps
- Number of subscribers - 1000
- Packet length distribution (subscriber to headend) - exponential
- Mean packet length - 200 bytes
- Propagation delay - 6 µs/km
- Mode of transmission - duplex
- Modem synchronisation time (at headend) - 10 ms

(a) Find the mean waiting time (delay) for arriving packets at the stations if each user generates an average of one packet per minute.

(b) If the channel data rate is reduced to 9,600 bps, what is the largest possible mean packet length that will not overload the system.

(c) For mean packet lengths of $2/3$ the result found for (b), determine the mean waiting delay.

First determine walk time ($w$). This is made up of:

- Tx time for go-ahead
- Propagation delay of go-ahead
• Tx delay for polling packet
• Propagation delay for polling packet
• Sync. delay of modem for received data at headend

Calculating these:
• \(8/(56 \times 10^3) = 0.14 \text{ ms}\)
• \(20 \times 6 = 120 \mu s = 0.12 \text{ ms}\)
• \(8 \times 8/(56 \times 10^3) = 1.14 \text{ ms}\)
• 0.12 ms
• 10 ms

Total = 11.52 ms.

(a) Throughput per station
\[
\rho = \frac{\lambda \bar{X}}{R} = \frac{1/60 \times 200 \times 8}{56 \times 10^3} = 4.76 \times 10^{-4}
\]

Total throughput
\[
S = M\rho = 1000 \times 4.76 \times 10^{-4} = 0.476
\]

Mean Delay
\[
W = \frac{Mw(1 - S/M)}{2(1 - S)} + \frac{S \bar{X}}{R(1 - S)}
\]
\[
W = \frac{1000(11.52 \times 10^{-3})(1 - 4.76 \times 10^{-4})}{2(1 - 0.476)} + \frac{0.476 \times 200 \times 8}{56 \times 10^3(1 - 0.476)}
\]
\[
= 10.99 + 0.026 = 11.02 \text{ sec}
\]

(b)
\[
S_{\text{MAX}} = \frac{1000(1/60 \times 8 \times \bar{X}_{\text{MAX}})}{9600} < 1
\]
\[
\bar{X}_{\text{MAX}} < 72 \text{ bytes}
\]
(c)

\[ \bar{X} = \frac{2}{3} \times 72 = 48 \text{ bytes} \]

throughput per station \( \rho = \frac{\lambda \bar{X}}{R} = \frac{1}{60} \times 48 \times 8 \times \frac{8}{9600} = 6.67 \times 10^{-4} \]

total throughput =

\[ S = M \rho = 1000 \times 6.67 \times 10^{-4} = 0.67 \]

Polling packet and go-ahead transmission times change because of the new channel bit rate. New walk time is given by:

\[ w = \frac{8 \times 8}{9.6} + \frac{8}{9.6} + 2 \times 0.12 + 10 = 17.74 \text{ ms} \]

Therefore \( W \) is given by:

\[ W = 26.62 + 0.01 = 26.63 \text{ sec} \]

Note that \( W \) in both cases is dominated by \( W_1 \) (\( W_2 \approx 0 \)).
A ring network is characterised by a sequence of point-to-point links between stations, forming a loop. All messages travel in one direction around the loop, passing through network interfaces at each station.

### 2.1 Operation of Ring

Bits from the ring enter in one direction in a serial fashion and, after a delay of several bits, are retransmitted over the ring either unchanged, or after some modification.
Access to the ring for transmissions is controlled by a token. A station wishing to transmit waits for an idle token to arrive. It then changes it into a busy token and follows it with data. An idle token is transmitted at the end of the holding time.

In the listen mode, each station passes on the packet received at its input after a delay referred to as the station latency. Define:

\[
\text{Ring Latency} = \text{Station Latency} + \text{Propagation Delay}
\]

The operational mode of a ring is determined by when idle tokens may be generated. The modes are:

- Multiple-Token
- Single-Token
- Single-Packet
Figure 2.2: Station / Ring Interface
### Figure 2.3: Operational Modes of Ring

<table>
<thead>
<tr>
<th>Discrete time</th>
<th>#1 Out</th>
<th>#2 Out</th>
<th>#3 Out</th>
<th>#4 Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>16</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>14</td>
<td>16</td>
<td>18</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>16</td>
<td>18</td>
<td>20</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>18</td>
<td>20</td>
<td>22</td>
<td>24</td>
<td>26</td>
</tr>
</tbody>
</table>

- **In**: Free token
- **Out**: Busy token
- **d**: Data bit

The figure illustrates the operational modes of a ring, showing the progression of data and tokens through the system at discrete time intervals.
2.2 Delay Analysis for Token Ring Networks

The following assumptions are made in the analysis:

- All stations behave the same.
- The arrival process at each station is Poisson with average arrival rate of $\lambda$.
- The average distance between the sending and receiving station is $\frac{1}{2}$ the distance around the ring.
- The stations are spaced so that the propagation delays between consecutively serviced stations are equal and given by $\frac{\tau}{M}$, where $\tau$ is the total ring propagation delay and $M$ is the number of stations.
- The packet length distribution is the same for each station with mean length $\bar{X}$ bits and second moment $\overline{X^2}$.
- All packets queued at a station are transmitted during a service period.

Let:

- $R =$ Channel bit rate (bits/sec)
- $B =$ Latency per station (bits)
- $\tau =$ Round trip propagation delay for the ring (sec)
- $\tau' =$ Ring Latency (sec)

A common channel is used for both a token ring network and a polling network. The master controller is not of importance to the operation of the system in so far as the analysis is concerned. Thus, the polling equations, developed in the previous chapter, can be used for token ring analysis.

The average waiting time in a polling network was derived to be:

$$W = \frac{Mw(1 - S/M)}{2(1 - S)} + \frac{S(X/R)^2}{2(X/R)(1 - S)}$$

The average transfer delay $T$ is given by

$$T = \frac{\bar{X}}{R} + \tau_{avg} + W$$
where

\[ \bar{X}/R = \text{Average time to transmit a packet} \]

\[ \tau_{avg} = \text{Average latency from Tx station to Rx station} \]

We need to state what \( \tau_{avg} \) is for a token ring LAN.

We assume that a packet travels on average half way around the ring before it reaches its destination station.

\[ \tau' = \tau + MB/R \]

and \( \tau_{avg} = \tau'/2 \).

The expression for \( W \) must be determined for token ring operation. Need to evaluate:

1. The walk time \( (w) \)
2. Moments of the service time \( \bar{X}/R \) and \( \bar{X}^2/R^2 \)
3. The network throughput

1. **Walk Time**

\[ w = \frac{\tau'}{M} = B/R + \frac{\tau}{M} \]

2. **Service Time**

For ring networks, the service time can include any time interval during which the ring is inactive in the course of completing one packet transmission and becoming available for another transmission. Call this time the **Effective Service Time** \( E \).

3. **Network Throughput**

define \( S' = M\lambda \bar{E} \)

while \( S = M\lambda \bar{X}/R \)
Using these expressions gives

\[ T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'(1 - S'M)}{2(1 - S')^2} + \frac{S'E^2}{2E(1 - S')} \]

This gives a general expression for the delay in a token ring network. The expressions for \( \bar{E} \), \( E^2 \) and \( S' \) will depend on the operational mode of the ring and the packet length distribution.

Note that \( T \) is normally referred to as the unnormalised delay. The normalised delay is defined as:

\[ \hat{T} = \frac{T}{(\bar{X}/R)} \]
2.3 Multiple Token Operation

In this case a new token is generated immediately after the last bit of the packet leaves the transmitting station. This is similar to the polling case as the ring is always transmitting between walk times.

The average effective service time is $\bar{X}/R$ and the second moment is $\bar{X}^2/R^2$.

Average transfer delay is therefore:

\[
T = \frac{\bar{X}}{R} + \frac{\tau'(1 - S/M)}{2(1 - S)} + \frac{S\bar{X}^2}{2XR(1 - S)}
\]

Maximum throughput is obtained when $S = 1$. However $S = 1 \Rightarrow T \to \infty$ which is not desirable.

For fixed length and exponentially distributed packet lengths we have:

**Fixed Length Packets** \[\bar{X}^2 = (\bar{X})^2\]

\[
T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'(1 - S/M)}{2(1 - S)} + \frac{S\bar{X}}{2XR(1 - S)}
\]

**Exponentially Distributed Packet Lengths** \[\bar{X}^2 = 2(\bar{X})^2\]

\[
T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'(1 - S/M)}{2(1 - S)} + \frac{S\bar{X}}{2XR(1 - S)}
\]

**Note:** In practice, maximum throughput is determined as the largest throughput for which the average transfer delay remains finite.

2.4 Single Token Operation

A new free token is not generated until the busy token from that station has been received again, i.e. has circled the ring.
Case 1

If the packet transmission time $X/R$ is greater than the ring latency $\tau'$, the station is still transmitting a packet when its busy token returns after circulating around the ring. In this case the operation is the same as the multiple token case.

Case 2

If $X/R < \tau'$, single token operation is different. In this case, the effective service time is $> X/R$ due to periods in which the ring is waiting for the busy token to return.

Define normalised ring latency delay:

$$a' = \frac{\tau'}{X/R}$$

Fixed and exponentially distributed packets must be treated differently.

Fixed Length Packets

$\bar{X} = X = \text{constant}$. If $a' < 1$ then the multiple token result applies. If $a' > 1$ then the result is different. In this case, the ring cannot be used to transmit another packet until the token has completed its "tour" of the ring.

$$\bar{E} = \tau'$$

$$S' = M\lambda\tau' = M\lambda(\bar{X}/R)a' = Sa'$$

Substituting into the general equation for $T$ gives:

$$T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'(1 - Sa'/M)}{2(1 - Sa')} + \frac{Sa'\tau'}{2(1 - Sa')}$$

If $a' < 1$ then the maximum achievable throughput on the ring is 1

If $a' > 1$ the throughput can reach a maximum of $1/a'$

(Noting that maximum throughput is found by setting $S' = 1$ and finding the corresponding value of $S$.)
Exponential Packet Distribution

In this case the packet length is random. A probability distribution must be assigned to the effective service time.

If $X/R < \tau'$ then $E = \tau'$

If $X/R > \tau'$ then $E = X/R$

$X$ has an exponential distribution function with mean $\bar{X}$ and since $R$ is a constant, the distribution function of $X/R$ is:

$$F_{X/R}(y) = \begin{cases} 
1 - e^{-Ry/\bar{X}}, & y \geq 0 \\
0, & y < 0 
\end{cases}$$

So the distribution function of the effective service time is:

$$F_E(y) = \begin{cases} 
0, & y < \tau' \\
1 - e^{-Ry/\bar{X}}, & y \geq \tau' 
\end{cases}$$
The first and second moments of $E$ are:

$$\bar{E} = (\bar{X}/R)e^{-a'} + \tau'$$
$$\bar{E}^2 = (\tau')^2 + 2(\bar{X}/R)^2 e^{-a'}(1 + a')$$

And also $S' = Se^{-a'} + Sa'$ thus:

$$T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'[1 - S(e^{-a'} + a')/M]}{2[1 - S(e^{-a'} + a')]} + \frac{\bar{X}S[(a')^2 + 2(1 + a')e^{-a'}]}{R \frac{2[1 - S(e^{-a'} + a')]}{2}}$$

The maximum throughput obtainable is given by $\frac{1}{e^{-a'} + a'}$
2.5 Single Packet Operation

In this mode of operation there is never more than one packet circulating on the ring.
There is always a delay of $\tau'$ seconds after the end of each packet. Therefore:

$$ E = X/R + \tau' $$

The 1st and 2nd moments are:

$$ \bar{E} = \bar{X}/R + \tau' $$
$$ \bar{E}^2 = (X/R)^2 + 2\tau'(\bar{X}/R) + (\tau')^2 $$

This results in the average delay expression:

$$ T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'[1 - (1 + a')S/M]}{2[1 - (1 + a')S]} + \frac{S[\bar{X}^2 + (R\tau')^2 + 2\bar{X}R\tau']}{2[1 - (1 + a')S]RX} $$

Fixed Length Packets

$$ T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'[1 - (1 + a')S/M]}{2[1 - (1 + a')S]} + \frac{S(\bar{X}/R)(1 + a')^2}{2[1 - (1 + a')S]} $$

Exponentially Distributed Packet Lengths

$$ T = \frac{\bar{X}}{R} + \frac{\tau'}{2} + \frac{\tau'[1 - (1 + a')S/M]}{2[1 - (1 + a')S]} + \frac{S(\bar{X}/R)(1 + a')^2 + 1}{2[1 - (1 + a')S]} $$

Upper bound to throughput is $1/(1 + a')$