ABSTRACT

This paper presents a method to calculate the capacity, efficiency, and measurement update cycle in Supervisory control and data acquisition (SCADA) systems when these systems use token passing bus protocol as a MAC protocol in medium access control sub-layer. After deriving the proper formula for parameters, the simulations results will be discussed and compared with results in polling protocol. Finally, according to the most important criteria, a suitable cost function for optimal operation of such system will be derived and with the results of minimizing this cost function, an optimum number of remote in specific condition and parameter are obtained. These calculations will lead remotes to communicate with MTU without violating the channel efficiency.

Keywords: SCADA, MAC, telecontrol protocol, optimization.

INTRODUCTION

In SCADA systems, a central station called master terminal unit (MTU) communicates with several remote stations through a single media. So it needs a technique to manage access to the media, which is called medium access control protocol (MAC). Up to now, many protocols have been used in SCADA systems. For optimal operation of SCADA systems, many techniques can be used such as increasing channel capacity, MTBF... Moreover, MAC protocols can be improved by studying, analyzing, and comparing different protocols. In the previous papers, most of them have been studied widely such as polling, token ring, Aloha, and CSMA/CD [1]. In this paper, a token passing bus protocol will be analyzed thoroughly.

MATERIALS AND METHODS

Protocol
The token bus protocol uses the same basic principles as the polling protocol, seeing that both belong to the class of MAC protocols using selection techniques. This implies the absence of collisions. In this paper, the classic token bus protocol has been adapted to the peculiarities of the SCADA systems. In the following, this adaptation is described and analyzed. In the token
busprotocol the nodes are ordered following a circular sequence which begins at the control center, goes through all the RTUs, and then returns to the control center[2]. The cycle begins with the transmission of a message (either with or without information) from the control center to any RTU. This message serves as a token and after having been received by the first RTU of the cycle, makes it send a message (with or without information) to the control center or any other RTU. This second message serves as a new token which after been received by the second RTU in the cycle, starts up its corresponding transmission. This process is executed successively until again reaching the control center, at which point it goes back to start the cycle [2].

As can be seen, the described protocol is nothing more than a polling protocol in which the answer from an RTU is used as well as a question to the following RTU in the cycle. So the performance of this protocol can be calculated through simple adaptations of those presented in[3]. In any case, and without the need for any computation, it’s clear that the performance offered by this protocol is simply superior to that of the polling protocol, seeing that all the process of continuous questioning from the control center is eliminated. This gives it a level of efficiency of between 80 to 90 % [3].

**Generating the information**
For the purposes of our study, we will consider that an RTU transmits three types of information to the control center [3], two of which are done cyclically, and the third, sporadically. The first type of cyclic information(typically information about analogical and digital network measurements) we shall call "measurements", and it contains a perishable image of the electrical network; which is to say that once a certain amount of time has passed, if it has not been able to transmit, it will become invalid and will be substituted for other equivalent, but more recent, information. This type of information is by far the most common. In contrast, the second type of cyclic information (for example, commands to the motes, control information, etc.), which we shall call "general" messages, does not lose its validity over time and is not substituted for more recent information. Finally, we will consider that the remote is capable of generating sporadic and spontaneous messages- which we shall call "incidents"- with urgent information (usually alarms and incidents that have occurred in the network) that should be sent on to the control center as quickly as possible. Let us also consider that each RTU generates measurement messages in accordance with an exponential distribution with average Tm, or in other words, each RTU generates, on average, one measurement message every Tm seconds. Similarly, we will consider that general messages are generated exponentially in each remote with an average of Tg seconds. Finally we will assume that incident messages are very rare and do not affect normal channel traffic. Keeping all this in mind, we will state that each RTU generates messages exponentially with an average Ta, shown as[3].

\[
T_a = \frac{1}{\frac{1}{T_m} + \frac{1}{T_g}} \tag{1}
\]

**MATERIALS AND METHODS**

**Experimental**
The polling protocol under consideration alternates cycles of questions whose response will be a message with information, with other cycles whose response is a null message. In any case, in order to study the protocol's behavior, the following parameters must be considered:

N: number of remote stations. N=2 TO 100

Np: number of stations which has information to send. Np=0.1N TO 0.4N

P: Number of bits in the polling message and the null message. P=60 BIT
M: Number of bits in the messages within formation (we will assume that they all have the same length). M=380 BIT

Cp: Capacity in bits per second (bps) of the physical link used. C_p=1200 BIT

T_a: average time of generating information in a station. T_a=4s

Tc: Switching time required to change from receiving or idle state to transmission state (negligible).

Dp: Propagation delay caused by the physical medium (negligible).

In Fig. 1 these parameters can be seen in a typical example which has two cycles of null responses and three which contains information.

**Fig. 1, sample behavior of token passing bus protocol**

**Update cycle**

It is basically defined as time the station has to wait to get access to physical channel in order to send data and is equivalent to complete cycle time. We can strongly change the value of update cycle by manipulating the communication protocol, number of stations, link velocity and frame length[5].

In token passing bus protocol total update cycle is divided into two parts

1. Total time elapsed for sending token to all of the stations
2. Total time elapsed in transmission of data from i stations which has data to transmit.

\[
T_{cycle} = T_{token} + T_{data} \tag{2}
\]

For transmission of token between two stations, times that elapse is equal to:

\[
\frac{P}{C_p} \tag{3}
\]

And the total time of transmission of token between N remote stations is equal to

\[
T_{token} = N \cdot \frac{P}{C_p} \tag{4}
\]

The time it takes to send a data with length of M bits through the physical channel with capacity of C_p is equal to

\[
\frac{M}{C_p} \tag{5}
\]

Since the number of stations having data to send is N- Np, the total time for transmission of data from N-Np stations is given by equation

\[
T_{data} = (N-N_p) \cdot \frac{M}{C_p} \tag{6}
\]

Update cycle of token passing protocol is obtained from the equation
\[ T_{\text{cycle}} = \frac{N \cdot P + (N - N_p) M}{C_p} \] (7)

Protocol efficiency
Protocol efficiency is equal to the length of total real data which have been transmitted from stations to master divided by total data which transmitted through the stations.

\[ \eta = \frac{(N - N_p) M}{N \cdot P + (N - N_p) M} \] (8)

Throughput
If the layer N+1 send data at bit rate C_{n+1} and layer N transmit data at bit rate C_n then Throughput of layer N is defined as [5]

\[ E_n = \frac{C_n}{C_{n-1}} \] (9)

This parameter depends highly upon the telecontrol protocol and number of stations, the length of frames and link velocity.

Each remote generate message exponentially with an average time of \( T_a \), this is a typically Poisson Process.

From the Poisson process, the probability of generation of K message in time \( T \) is given by equation [7]

\[ P(K) = \frac{\left( \frac{K}{T_a} \right)^K}{K!} e^{-\frac{K}{T_a}} \] (10)

And the number of messages generated by a remote during a cycle \( T_{\text{cycle}} \) will be

\[ n_e = E(K) = \sum_{K=0}^{K=\infty} K \cdot P(K) \] (11)

This can be shown that

\[ n_e = \frac{T_{\text{cycle}}}{T_a} \] (12)

The average number of messages that collide and are therefore eliminated in a time period of \( T_{\text{Cycle}} \) is given by the following equation [7].

\[ n_f = \sum_{K=2}^{K=\infty} (K - 1) \cdot P(K) \] (13)

\[ n_f = e^{-\frac{T_{\text{coll}}}{T_a}} + \frac{T_{\text{cycle}}}{T_a} - 1 \] (14)

So the number of actual messages generate in interval \( T_{\text{cycle}} \) is

\[ n_{ac} = n_e - n_f = 1 - e^{-\frac{T_{\text{coll}}}{T_a}} \] (15)

This is equivalent to say that each remote generate an actual message every \( T_{ac} \) seconds, which can relate to \( T_a \) and \( T_{\text{cycle}} \) by the equation
The bit rate of sending real data (not token) through the physical channel is equal to the number of bits in a message with length (M) divided by the average time of generated message
\[
C_s(i) = \frac{M}{T_{ac}} \quad (17)
\]
And the rate of generation of data in N station is given by
\[
C_r(N) = \frac{N.M}{T_{ac}} \quad (18)
\]
Since the capacity of physical channel is Cp, the throughput of protocol is obtained by:
\[
S = \frac{C_r(N)}{C_p} = \frac{N.M}{C_p} = \frac{N.M}{C_pT_{ac}} \quad (19)
\]
By substituting from (16)
\[
S = \frac{N.M}{C_pT_{cycle}}(1-e^{-\frac{T_{cycle}}{T_a}}) \quad (20)
\]
Subject is the optimum number of remotes and intermediate stations that can share the same link which leads to the lowest cost (the maximum number of station) without violating the communication time requirement.

Minimum time that remote station needs to generate new data is Ta so it is reasonable to say that the optimal waiting time for each remote for token is Ta. This means that [8].
\[
T_{cycle} = T_a \rightarrow T_a = \frac{N.P + (N-N_p).M}{C_p} \quad (21)
\]
\[
IF \quad N_p = K.N \quad \rightarrow \quad N_{OPTIMAL} = \frac{T_aC_p}{P+M-KM} \quad (22)
\]

RESULTS AND DISCUSSION

The simulation results are shown for both the bus and pooling protocol. In figures 2,3 and 4 capacity, efficiency and update cycle of polling and bus protocol are showed .as you can see all of the parameter are improved in token passing bus in comparison with polling protocol in the same condition. Also in figure 5 the maximum numbers of stations that can be communicate with the central is determined.
In Scada system there are many techniques for improving the performance. Also by selecting the proper protocol for mac layer it can be achieved. In this paper by selecting the token passing bus instead of polling some important Parameter such as capacity, efficiency and update cycle was improved. Also by defining the proper cost function the maximum number of stations that can communicate with master without violating the channel efficiency.
REFERENCES

[1] Stuart A. Boyer, Supervisory Control and Data Acquisition, Instrumentation Systems, 4 edition (June 15, 2009)