COMPARING PACKET FILL STRATEGIES IN ETHERNET-BASED DATA ACQUISITION SYSTEMS

Sérgio D. Penna, Engineer
EMBRAER Flight Test Instrumentation, Brazil

ABSTRACT

Ethernet-based data acquisition systems are becoming more and more common in the Flight Test Instrumentation environment. Digitized analog sensor output and various other types of digital data is captured and inserted into Ethernet packets using a "packet fill" strategy that in general is under control of the user. This paper discuss and compares two strategies "FILL-TO-TIME" and "FILL-TO-SIZE" for the acquisition of ARINC-429 digital data bus.

KEYWORDS

Flight Test Instrumentation, Ethernet Data Acquisition, Fill-to-Time, Fill-to-Size, ARINC-429.

INTRODUCTION

Today's Data Acquisition Systems used in Flight Test Instrumentation are capable of inserting instrumentation data into Ethernet/IP packets. These packets are transmitted commonly using one of two "packet fill" strategies:

FILL-TO-TIME: Packets are transmitted as result of timer expiration
FILL-TO-SIZE: Packets are transmitted as result of packet size limit exceedance

In either case, no packet is transmitted when its payload is empty.

The use of one or the other "packet fill" strategy is discussed here for the special case of the acquisition of ARINC-429 serial bus data.

SYNCHRONOUS AND ASYNCHRONOUS DATA TRANSMISSIONS

Data Acquisition Systems since the 80's are capable of acting as "bus monitors" for most digital buses used in modern avionics. ARINC-429, MIL-1553B, RS-485, RS-232, CAN, TTP, among others (more recently Ethernet), became important sources of data in flight testing complementing, sometimes replacing, analog or digital sensors traditionally deployed by Instrumentation Engineers.

Most of these digital buses are considered "asynchronous", in the sense that data transmissions from a source to a destination may begin at any given time without any previous agreement between participants in the bus. TTP can be considered "synchronous", in the sense that any data transmission starts and ends at a source at precise instants of time and repeats itself at precise time intervals, as dictated by a time-table shared by all participants in the bus.

PERIODIC AND NON-PERIODIC DATA SOURCES

In modern avionics, messages transmitted over digital buses are produced either by specialized hardware or by a combination of hardware and software executing in general purpose processors. In either case, data transmission can be initiated by either "periodic" or "non-periodic" tasks. A "periodic" task executes at precise time intervals, usually as a result of precise timer expiration. A "non-periodic" task usually executes as a result of an external event other than timer expiration (like flipping a switch or pressing a button).

Although ARINC-429, MIL-1553B, RS-485, RS-232 and CAN are in nature "asynchronous", "periodic" tasks transmitting data over any of these digital buses may generate a sequence of messages that appear to be as "periodic" as its source to an external observer. Under these special circumstances, messages are transmitted "asynchronously" (with no synchronization between source and destination), but "periodically" (equally spaced in time).

Periodic data sources are preferred, because it favors certain aspects of the design of data consumers. Fundamental to any analysis in the frequency domain, Nyquist's "Sampling Theorem" only works if signal samples are equally spaced in time. The discretization of any continuous model (as the transformation from the "s" domain to the "z" domain) also requires a fixed time step.

Modern avionic systems usually rely on "periodic task scheduling" and on "sampled signal acquisition", therefore one should expect "periodic" message flows through "asynchronous" digital buses.

PACKETIZED DIGITAL BUS DATA ACQUISITION

Today's flight test instrumentation hardware is commonly capable of performing similarly to commercial off-the-shelf "bus monitors" on the receiving side, as messages can be filtered and sorted out according to various criteria. Once received and processed, messages are inserted into Ethernet/IP packets and sent out to other devices, such as data analysis computers or instrumentation recorders. When no filtering or other processing is required, messages are accommodated into Ethernet/IP packets in the order they are received.

Ethernet/IP packets are transmitted according to one of two “packet fill” strategies:

FILL-TO-TIME (FTT): Packets are transmitted as result of timer expiration
FILL-TO-SIZE (FTS): Packets are transmitted as result of packet size limit exceedance

With FTT, packets are transmitted at regular time intervals with varying packet filling. With FTS, packets have the same size, but are transmitted at irregular time intervals.

In either case, no packet is transmitted when its payload is empty.

Choosing one or the other "packet fill" strategy is user's discretion and the hardware is in general programmable to accept one or the other strategy per digital data bus. Packet payload efficiency is vendor specific, as each message may be accompanied by other data, such as a "health" or "status" indication and a time-stamp generated at message reception.

FIELD OBSERVATIONS

In general, suppliers of modern avionics offer an "Interface Control Document" (ICD) which describes the contents of messages transmitted over a digital bus. Most of the time, the "periodic" nature of message transmissions is clearly expressed in terms of frequency (in Hertz) or period (in milliseconds). On the other hand, the order by which messages are transmitted is seldom expressed. However, fixed message groups are quite frequent.

A simple observation of message transmissions over a digital bus for a longer period of time can offer additional information about the actual profile of the message flow. Modern computer-based "bus monitors" are capable of recording and analyzing a message flow, so one can eventually determine its behavior over time.

Observations using such a "bus monitor" were performed in two ARINC-429 buses for a short period of time: one from a typical aircraft sensor (“A” bus) and one from a typical avionics processor module (“B” bus). The intention was to compare how these the two "packet fill" strategies perform and draw conclusions that may influence the choice of one or the other in packetized bus data acquisition.

EFFECT OF "FILL-TO-SIZE" IN THE OBSERVED ARINC-429 BUSES

In ARINC-429 buses operating at 100 kilobits per second, each message occupies a 360 microsecond slot, that is, 32 bit-times for the 32-bit message plus 4 bit-times of mandatory inter-message gap. A typical ARINC-429 bus monitor card uses a 12-byte block for each received message, so each Ethernet/IP packet formatted can accommodate as many as 120 messages without IP fragmentation. This 1,514-bytes packet is transmitted every 42.3 milliseconds, approximately 23 packets per second, if one assumes a regular flow of messages through the bus.

The choice of the maximum packet size depends on the application. Sometimes a higher rate of packet transmission is desired when, for instance, a short-period dynamic behavior has to be observed. In this case, the maximum packet size is reduced expecting the rate of packet transmission to be increased, if one assumes a regular flow of ARINC-429.

Using FTS packet fill strategy had dissimilar effects on A and B buses. In the A bus, a programmed maximum packet size of 1,008 bytes, enough to accommodate 84 ARINC-429 messages, resulted in packets being transmitted at intervals varying from a maximum of 68 milliseconds (apparently there has been a transmission interruption in the A bus) and a minimum of 34 milliseconds. Assuming a regular flow of messages through the bus, the packet transmission rate would have been 34 packets per second, roughly one packet every 30 milliseconds. In the B bus, packets were transmitted as a result of a 50 milliseconds, user-defined, "safeguard timer" expiration with payloads filled with a maximum of 26 and a minimum of 23 messages.

The A box seemed to produce a more irregular, faster message flow in its normal operation, while the B box seemed to produce a more regular, slower message flow during the short period of observation.

ZOOMING IN THE MESSAGE FLOW

By closely examining the A bus message flow, groups of messages that repeated themselves at harmonic rates could be identified, but there was still a significant count of messages for which the transmission intervals could not be exactly determined within the observation window. The message groups found are listed in Figure 1.

Transmission Group	Number of Messages	Transmission interval
A	9	10ms
B	10	20ms
C	6	20ms
D	4	40ms
E	3	40ms
F	32	>40ms

Figure 1 – ARINC-429 message transmission groups for the A bus

The B bus message flow was surprisingly regular, with two fixed groups of messages transmitted alternately every 40 milliseconds with very small jitter. The occurrence of ordered pairs of ARINC-429 labels indicated that the hardware designer wanted to "mark" the start and end of each message group.

Transmission Group	ARINC-429 Labels in Group				
A	0	1	204	75	374
	0	1	205	224	375
	0	1	145	231	376
	0	1	24	230	377
B	0	1	26	227	374
	0	1	74	237	375
	0	1	25	376	
	0	1	27	377	

Figure 2 – ARINC-429 message transmission groups for the B bus

In Figure 2, note that each group is a unique sequence of ARINC-429 labels accommodated in four rounds that start with a label pair "0" and "1" and ending with label "377" (the largest octal number in an 8-bit word) at the end of the fourth round.

None of this was clearly indicated in the supplied ICD for any of these buses. However, a simple "bus monitor" software tool listening to the physical bus could quickly identify message transmission rates and calculate its maximum, minimum, average and jitter (difference between maximum and minimum). If the message flow could be recorded on a hard-disk file for further investigation, the presence of fixed groups of messages could be more easily detected by a simple heuristic algorithm.

CONSIDERATIONS ABOUT "FILL-TO-SIZE" IN THE A BUS

In the A bus, it seemed waste of time finding any benefit for the little regularity observed in its message flow. However, a method adapted from the traditional "Task Schedulability Analysis" [1] by means of finding each task's "Response Time" can help finding a packet composition that better fits the observed message flow in which groups of messages repeat themselves at regular periods.

In the iterative formula (1), *C* ("Capacity") represents the task execution time (in general the worst-case), *T* ("Period") represents the task activation period, *HP()* represents the group of tasks that have scheduling priority higher than the task under analysis and *R* ("Response") represents the minimum time interval in which the task under analysis and all tasks with higher priority can execute to completion.

$$R_j^{n+1} = C_j + \sum_{i \in HP(j)} \frac{C_i}{T_i} R_j^n , R_j^0 = \sum_{i=1,j} C_i \qquad (1)$$

The adapted method uses the 360 microsecond ARINC-429 message inter-arrival time (32 plus 4 bit times at 100kbps) as the base unit for calculating *C* ("Capacity") for each one of the message groups. For instance, a group of 10 messages would have a *C* of 10*0.360 = 3.600 milliseconds. For *T* ("Period"), the observed average message transmission interval is used. The "Rate Monotonic

Priority Assignment" [1] is used as high-to-low priorities are assigned to high-to-low rates, that is, short-to-long periods.

Finally, an iterative process adds a complementary group of "low priority" messages arriving at the longest observed period until the total transmission time approaches but not exceeds the longest message period in a still schedulable set (task deadline equal task period). Figure 3 shows the result of the iterative process.

Count	**Times**	**Total Count**	**Delta T**
9	4	36	10
10	2	20	20
6	2	12	20
4	1	4	40
3	1	3	40
		75	

C	**T**	**Prio**	**Repeats**
3.24	10	4	n/a
5.76	20	3	n/a
2.52	40	2	n/a
12.96	>40	1	**36**

Sum C	24.48	**# of Msg**	
w0	24.48		
w1	36.72	102	
w2	39.96	111	
w3	39.96	**111**	
Pkt Time	30.24	84	
	40.32	112	
	43.2	120	**(max)**

Figure 3 – Iterative process for finding a suitable packet size

The resulting number of messages per packet (111) can be easily obtained by dividing the longest observed transmission interval (40 milliseconds) by 0.360 milliseconds (total duration of an ARINC-429 message). The iterative process has the benefit of finding how many additional slower rate messages (36) can be accommodated while maintaining a transmission period that is close to the smallest common multiple of all message periods.

However, in the A bus there are only so many messages (32) transmitted at slower rates (less than 1 in 40 milliseconds); therefore any number of messages (from 0 to 32) can arrive within the longest transmission interval (40 milliseconds). Even considering a message transmission burst, the resulting number of messages (75 plus 32) is still less than the total number of messages calculated by the iterative process (111).

Since there is no simple way of calculating an “optimal” packet size for this particular case, it would make more sense to use FTT and setting the packet transmission interval to the least common multiple of all harmonic message periods, plus some jitter.

CONSIDERATIONS ABOUT "FILL-TO-SIZE" IN THE B BUS

In the B bus, the repeating sequence of messages arranged in two fixed groups indicates that the correct packet size should fit an integer number of these two groups. Since the two groups added together contain 38 messages, it is possible to fit a total of 3 sets (114 messages) in a packet without fragmentation.

Although surprisingly regular in normal operation, there are no guarantees that the B box will always transmit messages without interruption, so it makes more sense to use FTT and setting the packet transmission interval to 1 to 3 times the period of one whole set of messages (80 milliseconds).

CONSIDERATIONS ABOUT "FILL-TO-TIME" IN THE OBSERVED ARINC-429 BUSES

It seems that FTT can be adequate for those ARINC-429 buses originating from a particular category of avionic boxes that execute a strictly periodic task scheduling. However, a question remains about at which point in time the timer should start. Let us examine two possibilities:

A) Timer starts as soon as one message arrives.

Assuming that no packet should be transmitted empty, the timer should start as soon as one message arrives and restart as it expires. In a regular, periodic message flow, packets should contain the same number of messages. However, even when messages are transmitted in groups, there is no guarantee that these groups will be entirely contained in one packet.

If the message flow is interrupted or slowed before timer expiration, the resulting packet size may be shorter than expected.

B) Timer starts upon arrival of a particular message

This can be very helpful when the message flow is "marked", that is, there is a message (or sequence of messages) that "marks" the beginning of a fixed periodic group of messages. To handle the situation whereby this "mark" is never transmitted, a "safeguard timer" should be set to a carefully chosen time interval, for instance, the least common multiple of all message group periods.

As above, if the message flow is interrupted or slowed before timer expiration, the resulting packet size may be shorter than expected. However, if the message flow is regular and message inter-arrival times show little jitter, there is a greater chance that the majority of packets will be transmitted with the same size.

CONCLUSION

It seems that "FILL-TO-SIZE" (FTS) or "FILL-TO-TIME" (FTT) packet fill strategies are not capable of improving characteristics such as regularity or predictability in the resulting packet flow coming out of a packetized ARINC-429 bus data acquisition, if used alone.

If one chooses FTS, there is still a need for a "safeguard timer" to prevent a situation whereby a partially filled packet may take an arbitrary long time interval to be transmitted. If one chooses FTT, there is a chance that an irregular message flow may cause partially filled packets to be transmitted degrading payload efficiency.

If there is enough knowledge about the ARINC-429 message flow indicating regularity and predictability, a combination of FTS and a FTT packet fill strategies can complement each other:

- Choose FTS as the "primary strategy" and FTT as the "safeguard strategy" for regular, periodic message flows;
- Choose FTT as the "primary strategy" and FTS as the "safeguard strategy" for irregular, non-periodic message flows.

It seems ironic to choose FTS as "primary strategy" for a regular, periodic message flow such as the one coming out of the B box. However, a fixed bit-rate source transmits always the same amount of bits per unit of time, so choosing a "size" indirectly fixes a time interval and makes the packet filling immune to a small transmission jitter, which may cause a packet to be transmitted with less than its expected size a tiny fraction of time ahead of its due time. The FTS as a "safeguard strategy" for irregular, non-periodic message flows is somewhat equivalent the usual "empty packets shall not be transmitted" rule. It means that no packet should be transmitted with less than the specified size, regardless of the time taken to fill it.

Taken to an extreme, a perfectly regular and predictable ARINC-429 message flow could allow switching from a "parser-aligned" to a "placed" iNet-X payload lay-out with all the benefits in payload efficiency and in the design of a data consumer application. In between, a periodic, "marked" ARINC-429 message flow can allow selecting the right size and right time for a better configured packet fill strategy.

ACKNOWLEDGEMENTS

I would like to thank Yuli Kolesnikov for his support in revising this paper from a hardware designer perspective.

REFERENCES

[1] Burns, A.; Wellings, A. *Real-time systems and programming languages*, 2nd edition, New York-NY, USA: Addison-Wesley, 1997.