Can Satellite Communications Support Time-Critical Applications

Frank Ball and Kashinath Basu

Frank Ball Consulting, Oxford UK .
School of Technology, Oxford Brookes University, Oxford UK.
frank_ball@ntlworld.com; kbasu@brookes.ac.uk

Abstract. Satellite networks are seen as having the potential to play an important role in machine-to-machine (M2M) communications and the Internet of Things, which in turn is seen as being important to a number of service sectors. However, certain M2M application have bounded latency requirements that in some cases may be quite stringent. Satellite networks general much higher latency that wired networks and therefore may not be able to meet the requirements of all M2M applications. This paper compares the latency requirements of certain time-critical applications with reported satellite network latency and address the problem of latency evaluation of networks support these types of application including those involving satellite links.

Keywords: Satellite Communications, IoT, Time critical application, monitoring and Control

1 Introduction

The global connective offered by the Internet and the number of heterogeneous, and in many cases smart, devices that are able to connect to it has led to the concept of the Internet of Things (IoT). Some see this concept as being a direction for the evolution of the current Internet, while others view the IoT as separate entity [1][2]. The important thing from the perspective of this paper, however, is that machine-to-machine (M2M) communication is quite different to that where human interaction is involved. In particular the traffic characteristic and QoS requirements of M2M applications are somewhat different from those of current Internet applications. For certain monitoring and control systems, strict latency requirements will need to be assured [2][3], and in some cases this assurance may need to be provided over large distances. However, given that propagation delay is a function of distance, then obviously, this sets a limit to what can be achieved, even with minimal switching delays and high bandwidth links.

Global (or very wide area) communication generally involves a combination of wired, wireless, and in many cases satellite networks. Satellite networks are seen as having the potential to play an important role in the IoT and M2M communication particularly in cases of remotely placed smart objects and group-based communications [1]. The IoT is seen as being important to a number of service sectors, including: Buildings; Energy, generation and distribution; Healthcare, e.g. remote patient monitoring, telesurgery; and transport etc.

The control of power grids is one area in which satellite communications already plays a role, and it is expected that they will also make a contribution to the smart grids of the future. However, certain smart grid applications that have been proposed for use in the near future are time-critical and have quite stringent latency requirements [3][4]. In these cases it seems unlikely that satellite communications could meet their particular requirements. Satellite communications may also be incorporated into Medical Grade networks (MGN) [5]. MGN also need to support time-critical applications, e.g. remote monitoring of a patient's vital signs, although for these applications, a much higher degree of latency is allowable. However, even though the latency requirements are less stringent, they must still be met for the application to function correctly. Also, in addition to bounded delay, both smart grid and MGN control and monitoring applications may also require a high level of reliability. There may also be many other examples of time-critical application for which satellite communication may be considered to be a suitable carrier, however, for economy of space, this paper will only focus on Smart grid control systems and MGN monitoring and diagnostic services.

To provide the necessary guarantees to time-critical applications, performance evaluation must be carried out *a priori* to deployment of the communications system. Therefore, in the case of supporting stringent real-time applications, performance evaluation will need to become an integral part of the system design process. Whilst performance evaluations of existing systems, based on either simulation or empirical testing, can provide useful information, they can only show whether or not a particular system may be able to meet its targets.

The aims of this paper are to address the potential for satellite communication to support time-critical monitoring, and control applications, and to consider how to evaluate the latency of a satellite link for individual cases.

The remainder of the paper is structured as follows: section 2, outlines the characteristics and requirements of time-critical applications; section 3, discuses smart grid applications and their relationship to satellite communications; section 4 presents MGNs and discusses the options for satellite interactions; sections 5. presents ongoing work and discusses future directions; and finally section 6. offers some concluding remarks.

2 Time-Critical Applications

Time-critical applications are a class of real-time application in where missed deadlines may result in serious consequences. The degree of seriousness will of course vary between applications, which in extreme cases could include, for example: significant loss of revenue: damage to equipment; and possibly personal injuries [3]. Less serious consequences could include a significant reduction in the efficiency of the control application. The degree of time criticality may also vary between applications depending on the flexibility of the control mechanism. Therefore, the type of real-time guarantee required will also vary with application. In general, there are three basis classes of real-time guarantee: Hard Real Time, in which information arriving after the deadline T is no longer of any use (i.e. equivalent to packet loss); Firm Real Time, that is similar to Hard Real Time, except that some losses may be tolerable; and Soft Real Time, in which information arriving before a deadline T_1 is fully useful, but will still have a degree of usefulness until a second deadline T_2 has been exceeded.

Proving a Hard-Real-Time service can be quite difficult in practice and the definition of Firm-Real-Time is not particularly explicit. Using a probabilistic interpretation as shown below in equations 1 and 2 offers a more pragmatic definition that can be matched to the tolerances of the control applications.

$$\text{Probabilistic Hard Real Time:} \quad P(t>T) <= 10^{-x} \qquad (1)$$

$$\text{Soft Real Time:} \quad P(t>T_1) <= 10^{-x} \text{ AND } P(t>T_2) <= 10^{-x} \qquad (2)$$
$$\text{with } f_u(t) = [1 \; t<=T_1;\; g(t) \; T_1 < t <= T_2;\; 0 \text{ otherwise}]$$

The setting of these parameters would need to be determined by the requirements of the application and the flexibility of the control algorithm. Certain smart grid and MGN control applications are examples of a Networked Control System (NCS) [6], i.e. a closed loop control system that operates over an open network that is shared with other classes of traffic. Designers of NCSs face two challenges: maintaining the appropriate QoS in the networks; and ensuring that the required Quality of Control (QoC) is provided. Therefore research into NCSs focuses on two objectives: Control of the network, to ensure a suitable QoS; and control over the network, that seeks to minimize adverse conditions in the network. Therefore, NCSs methodologies may prove useful for setting the real-time parameter.

The traffic patterns of a monitoring and control application may not only be different from that of other types of applications, but can also vary between different instances of monitoring activity. Communication between a monitoring device and the controller may take the form of: one-to-one, i.e. a single stream; many-to-one, i.e. a number streams fanning in to the controller; and many-to-one with synchronized sources. For the latter, it is possible that the controller may need to wait until a full set of synchronized messages arrive before processing them. In this case the delay bound will apply to the arrival time of the last message from the set. For the other two cases, however, the delay bound will apply to each individual message.

In general, average delays based on long periods of time do not provide a useful QoS metric for time-critical control applications. High percentiles of delay taken over relatively short periods are more useful for identifying adverse transitory conditions.

3 Smart Grid Applications

A communication network is an essential component of a smart grid system. Its role is to support a wide range of applications, many of which have very similar requirements to those of current Internet applications. In particular, they have the general requirements for security, resilience, reliability and wide area interconnectivity. However, a number of classes of smart grid applications, particularly those intended for controlling smart grids in the near future, have requirements that are significantly different from those of any existing Internet application. Time-critical smart grid applications are responsible for state estimation, control, protection, and ensuring the stability of power generation and distribution. Their domains of operation include both the local area for internal sub-station control, and the wide area for protection, control, and maintaining wide area awareness. Currently, when operating in the wide area, the role of these applications is generally limited to providing visualization for wide area awareness. Applications that provide visualization have near-real-time requirement and can tolerate latencies in the order of 100ms. Currently, automatic control, for which the more stringent delay requirements apply, is mainly limited to the local area [4].

De Sanctis et al [1] show that satellite communication based on low earth orbit (LEO) constellations can meet most of current power grid supervisory control and data acquisition (SCADA) requirements. But may not be able to meet the needs of emerging applications. In particular wide area situation awareness (WASA) has latency requirements of < 200ms, which is shown by Yang et al [2] to be just within the limits of a LEO constellation. It should be noted that the results of both [1] and [2] are based on the use of standard IP protocols.

WASA is based on a wide area measurement system (WAMS) that collects synchronised phase measurements of the power cycle, known as Synchrophasors from devices called Phase Measurement Units (PMUs) that are deployed throughout the grid. The result traffic pattern is that of many-to-one with synchronised sources[4]. Currently, in the wide area these measurements are used to provide visualisation for wide area awareness, and control decisions, and activations, are carried out manually. Automatic control is generally limited to internal control within sub-stations. However, in order to meet the smart grid objectives for a greater use of renewable energy and a more efficient of existing non-renewable source the aim is to extend automatic control to the wide area. The ultimate requirement for synchrophasor based wide area control applications is to carry out the measurement-to-decision process within one power cycle [3] [7] which in the case of a 60hz power cycle is 16.7ms. This is an application level target and it has been proposed that the communication latency should be in the order of 2 ms [3].

With such a stringent delay requirement it will not be possible for satellite communications to support this type of application. However, even in the case of a wired network that has high speed links and minimal switching delays, the size of a control domain will still be limited. For example, in the best case a circular domain with a centralized controller will be limited to a radius of less than 400k, based on a propagation delay of 5us/km. However, given a grid that comprises several control domains, inter-domain control will also be needed, thereby resulting in a hierarchical control system. If this does become the case, and the higher level control systems can operate with less stringent delay bounds, it is possible that satellite communication may be able to support the higher level control communications.

4 MGN Applications

Medical Grade Networks (MGNs) are required to deliver various e-Heath services to end users on an anywhere, anytime basis [8] [9]. These services include, for example, health information, electronic health recording, telemedicine and patient monitoring. In common with other communication networks MGNs need to provide reliability, resilience, various levels of security, and to support mobility for a number of applications. In addition, certain classes of these services, e.g. ECG monitoring, are delay intolerant and have firm or hard real-time communication requirements. Therefore, a MGNs needs to be an advanced multimedia network that provides the appropriate prioritisation to time critical sources.

Due to the requirements for mobility, local area wireless networks WLANs play an important part for patient monitoring. Generally, they provide a communications path between the monitoring devices and a wider area wired network. Numerous performance studies of IEEE 802.11 WLANs have addressed the problems of prioritization for time-critical applications at MAC and LLC levels, example of which include [9] and [10]. These studies emphasise the need for a thorough understanding of the performance issues in both the Link and MAC layers. In particular, for IEEE 802.11 WLANs, both data rates and PER can vary with channel conditions. This relationship offers a more challenging problem to performance evaluation than that experienced in the case of wired networks. Given the influence that channel condition can have on both data rates and packet error rates (PER), performance evaluations based on long term aggregation will most likely produce misleading results. Therefore, short term transitory conditions must also be considered .

Given that satellite communication performance can also be effect by channel conditions the comments made above relating to WLANs will also apply to the satellite link. If the satellite link is used to provide connection to a remote WLAN based monitoring system, then it could be the case that combined WLAN and satellite latency may exceed the delay target. However, with monitoring devices that are able to communicate directly with the satellite system the delays due to the WLAN can be eliminated.

5. Work in Progress and Future Directions

The delay requirements of time critical applications need to be addressed on an end-to-end basis, with the delays resulting from protocol stack operations also being taken into account. However, for practical reasons, some degree of decomposition in the evaluation process is usual needed.

In previous, work we have addressed the problem of post deployment evaluation of a power grid synchrophasor measurement system operating over a wide area wired network, and have carried out the first stage of developing of an evaluation approach based on generic parameterized models [4]. It was shown, that in the case of synchrophasor measurement and control systems, decomposition between the network and the application devices is quite straight forward due to their mode of operation. However, these initial models were based on high speed links, line rate strict priority switching and relatively deterministic real-time devices.

Currently we are extending these models to accommodate a higher degree of variability in both device output and switching mechanism performance. We have been experimenting with a convolution based approach for evaluating high percentiles of end-to-end delays in cases where decomposition has been applied. Initial results from this work are positive, however, they do show that even a low degree of variability leads to a much higher level of complexity in the final evaluation process.

The focus of our approach has been now widened to include the evaluation process for MGNs and to take into account the effects of wireless LAN communication. This line of the investigation is still in the early stages, however, it has indicated that the Link and MAC layers are appropriate boundaries for decomposition and that there may be a significant advantage in the co-evaluation of the Link and MAC layers due a level of interdependence as identified in [9] and [10].

The perceived importance of satellite communications for supporting, the IoT in general, and power grids and MGNs in particular has encouraged us to open our investigation to include satellite communication. Furthermore, as far as the evaluation process is concerned, we can see some commonality between the MGN Link/MAC evaluation and the evaluation of a satellite system. Although we do expect the latter to involve more stages of evaluation. In particular the effects of orbital configuration and different frequency bands will present a wider range of variability. Furthermore, unlike many other studies which focus mainly on average delays, our investigation is aimed mainly at evaluating the effects of short term worst case performance which is more relevant to requirements of time-critical applications.

6. Conclusion

This paper has addressed the potential for satellite communications to support time-critical monitoring and control applications, with particular focus on smart grid and MGN applications. It has noted that certain control applications proposed for the smart grids of the future have delay requirements that are too stringent to be met by satellite links. However, it has also noted that it may still be able to play a role in the smart grids of the future. It has also outlined the traffic characteristics and requirements of time-critical applications and discussed the need for post deployment evaluation. The paper has also outlined work in progress on the development of a post deployment evaluation technique based on generic and parameterized models. Finally, it has presented an outline plan for the progress of this work.

References

1. De Santus, M. Cianca, E. Araniti, G. biso, I. Prasad, R.: Satellite Communications Supporting Internet of Things. J. IEEE Intenet of Things, vol 3, No 1, 113-124. (2016)
2. Yang, Q. Laurenson, D.I and Barria J.A..: On the use of LEO Satellite Constellation for Active Management in Power Distribution Networks. IEEE Transactions on Smart Grid, Vol 3, 1371-1381, (2012)
3. Bakken, D.E.,Bose, A., Hauser, C.H.,Whitehead, D.E., Zweigle, G.C. :Smart Generation and Transmission With Coherent, Real-Time Data. Proceedings of the IEEE, Vol 99, No 6, June, (2011).
4. Ball, F and Basu, K,: Performance Evaluation of Time-Critical Smart Gris Applications: In: The Proceeding of the Eleventh International Network Conference (INC 2016) , Frankfurt, Germany. 13-18, July (2016)
6. Gupta, R. A. and Chow, M.-Y :Networked control system: Overview and research trends. IEEE Trans. Ind. Electron., vol. 57, no. 7, 2527–2535, Jul. (2010).
7. K. C. Budka,. J. G. Deshpande, T. L. Doumi,. M. Maddan and T. Mew, "Communication Network Architecture and Design Principles for Smart Grids", Bell Labs Technical Journal, vol. 15 No. 2, pp 205-228, (2010)
8. Skorin-Kapov, L. and Matijasevic, M.: Analysis of QoS Requirements for e-Health Services and Mapping to Evolved Systems Qos Classes. International Journal of Telemedicine and Aplications. Vol 2010. (2010)
9. Lin, C-F.: An Advanced Wireless Multimedia Communication Application: Mobile Telemedicine. WESEAS Transactions on Communications, Vol. 9, No. 3, 206-215, (2010)
10. Kang, K. Park, K-J. Song, J-J. Yoon, C-H. and Sha, L.; A Medical-Gread Wireless Architecture for Remote Electrocardiography. IEEE Trans Inf Technol Biomed. Vol. 2, 260-267. (2011)