Impact of Transport Layer Protocols on Reliable Information Access in Smart Grids

Kamal Shahid, Aamir Saeed, Thomas le Fevre Kristensen, Rasmus L. Olsen
Department of Electronic Systems, Aalborg University
Aalborg, Denmark
Email: {ksh, asa, tfk, rlo}@es.aau.dk

Abstract—Time is critical for certain types of dynamic information (e.g., frequency control) in a smart grid scenario. The usefulness of such information depends upon the arrival within a specific frame of time, which in other case may not serve the purpose and affect controllers performance. In this context, transport layer offers different levels of end-to-end communication services to the applications. For instance, TCP guarantees the transport of messages between two ends, however, at the cost of high end-to-end delays due to the retransmission mechanism. Whereas UDP offers minimum end-to-end delays at the cost of unreliable, best-effort data transportation service. The research question raised in this paper is thus, which is preferred for the delay-critical applications of smart grids, and to what degree of packet losses and round trip times, TCP is preferable to UDP and vice versa. The question is addressed by analyzing the performance of UDP and TCP over imperfect network conditions to show how the selection of transport layer protocol can dramatically affect controllers performance. This analysis is based on a quality metric called mismatch probability that considers occurrence of events at grid assets as well as the information update strategy in one single metric which otherwise is not very intuitive and difficult to allow a similar useful comparison. Further, the analysis is concluded by providing a clear guide on the selection of the transport protocol to meet application requirements.

Index Terms—Mismatch Probability (mmPr); Smart Grids; TCP; UDP

I. INTRODUCTION

Today, the governments of several countries are envisioned not only to upgrade the entire power grid system to smart grid but also to convert conventional fossil power plants into an entire renewable energy integrated system [1]. This goal will be attained by an active and reliable communication between various actors within the grid and a large scale installation of wind power and photo-voltaic (WP/PV) plants having a resilient communication infrastructure to coordinate their grid support services. This will bring new operational challenges for the system operators and will significantly change the control and operation of the existing distribution grids. For instance, with the high foreseen penetration of Renewable Generation (ReGen) plants, the electrical grid could face frequency and voltage problems [2]. In order to enable high penetration of ReGen plants in future and utilize them in a coordinated manner, the Distribution System Operator (DSO) would need to deploy private and/or use public communication networks.

Implementing new large-scale communication infrastructure is not economically feasible; therefore, existing communication infrastructure should be considered and further investigated for improved performance. Nowadays, cellular networks (EDGE, LTE etc.), fiber optics, cable internet and xDSL are already widely deployed by the telecom operators and have high geographical coverage [3], which could be used to connect the ReGen plants to the system operators. However, shared network solutions may not be able to provide quality-of-service required by the grid services and could bring additional risk, especially if they are exposed to internet access. Data exchange may suffer from stochastic non-controllable delays and packet drops. Therefore, a high consideration is required while designing systems like power grids that provide high dependability [3].

In smart grids, a system operator being in-charge of controlling multiple ReGen plants will depend exclusively on the information provided by these plants for sending correct set-points. It will, therefore, become very critical for the system operators to be well aware of the status of the connected plants. Especially, in case of delay-critical applications (e.g. protection and control related) where a delay of few milliseconds can cause the information to become outdated for the control center, which ultimately can become a big risk for the entire power system. In the worst-case, this may result in an unstable power grid and/or a blackout.

In order to ensure such control and monitoring, the International Electrotechnical Commission (IEC) has developed protocol standards for electric power systems and substations. For instance, IEC-61850 identifies the general as well as specific functional requirements for communications in a substation [4]. These requirements aid in the identification of the desirable services, data models, application protocol as well as all the underlying layers in the communication stack defined by the OSI reference model that will meet the overall requirements. However in the OSI model, it is the transport layer that is responsible for providing different levels of end-to-end data transportation service quality to the applications [5]. For instance, TCP provides a connection oriented [6] service that includes a mechanism to acknowledge the reception of data and a retransmission in case of lost data/acknowledgement. This allows a guaranteed transmission/reception of data packets in a causal order. TCP also provides congestion control, flow control and reliability by adding headers with the original
message. However, due to the retransmissions and congestion control mechanisms, TCP generally suffers with relatively higher delays in case of dropped packets or time-outs. UDP, on the other hand, provides connectionless, best effort service [7] with no guarantee of message delivery. It does not provide services like congestion control, flow control and reliability, therefore, faster than TCP. Due to the lack of such functionality in UDP, the application must accept that packets may vary well be lost in the network or arrive in different order than it was sent from the source. Typically, in practice TCP is used in industrial protocols such as MODBUS/TC and IEC-61850 for communicating over networks [8] [9]. Therefore, the research question raised in this paper is thus, which is preferred and to what degree of packet losses and round trip times, TCP is preferable to UDP and vice versa for various time critical applications in smart grids. The trade-off between using TCP and UDP, is in fact a trade-off between losing data in the network or accepting much higher delays in data reception.

A. State-of-the-Art

The performance analysis of transport layer protocols over communication networks in general has been addressed in several papers. For instance, [10] analyzes the performance of TCP, UDP and some improved protocols based on TCP in adhoc wireless networks based on throughput, packet loss, jitter, end-to-end delay and fairness. Reference [11] presents the same performance evaluation as [10] but on wired network environment. In [12], an analysis of both the transport layer protocols in a wireless LAN 802.11 test bed with different scenarios has been provided considering the flow fairness with a single access point and varying the number of mobile stations. These papers lack to address the performance of transport layer protocols in relation to the smart grid applications especially focusing the standards on the communication and control of electric power systems as, for instance, proposed by IEC. However, in [13] we analyzed information reliability over various imperfect communication network conditions with IEC-61850 MMS using the concept of mismatch probability. A resulting trade-off between quality of controller performance and mismatch probability has also been identified in [13]. Still, there is no solution proposed to decrease the probability of information mismatch and improve quality of the controller performance in [13]. Therefore, this paper is extending the previous analysis by providing a clear guideline to optimize communication performance in IEC standards (e.g. IEC 61850, IEC 60870) that currently use TCP as a transport layer protocol [8] [9].

In this paper, we analyze the trade-off between end-to-end delays and packet losses for the two transport layer protocols. This analysis is based on the information accuracy in the communication between plant controllers and the control center. It is hypothesized that a correct and timely reception of information leads to good/expected controller performance, while delayed information may cause information mismatch between the two ends, causing degraded controller performance. The information accuracy in a given scenario is measured using an information quality metric known as mismatch probability (mmPr). mmPr was first defined in 2010 in [14] and since then it has been applied to different scenarios, e.g. in [3], [15] and [16] to improve smart grid control. The benefit of using mmPr as a quality metric is that it considers both the occurrence of events and the update strategy in one single metric and put those in relation to the dynamics of the grid scenario [16]. This otherwise is not very intuitive and difficult to allow a similar useful analysis. Secondly, it has been ascertained in [17] that the simulation results on mmPr and voltage quality under the considered controller show same qualitative behavior. This implies that the mmPr as a quality metric can be used to identify relevant delay ranges as well as the update period intervals ranges that are expected to impact voltage quality performance. [17] also concludes that mmPr can be used to optimize communication network and information access configurations without the controller realization. Therefore, this paper only focuses on the network aspects of the communication without the realization of a specific controller. Finally, based on the outcome of the analysis, a solution is proposed that serves as a guide for the right selection of transport layer protocol, specifically in IEC 61850/60870 for various time critical applications in smart grids. The main contributions of the paper are:

- A procedure for estimating the information accuracy over congested communication networks.
- Based on the rate of occurrence of events, define a reference graph to serve as a guide to select appropriate transport layer protocol for a specific application.

The remainder of this paper is organized as follows: Section II defines and explain the quality metric selected to compare performance of the two transport layer protocols. Section III describes the case specific scenario adopted to get simulation results. Section IV provides evaluation of simulation and analytical results and finally Section V summarizes the conclusion drawn and directions for future work.

II. INFORMATION QUALITY METRIC

As described in section I, for delay-critical applications in a smart grid system, the reception of correct status information within a predefined frame of time is crucial for the control-center to take correct actions. The added delays due to, for example, poor network conditions can cause the information to become outdated for the control center. This is because the information age generally increases (approximately linear in case of periodic updates) as a function of the delay [17]. This implies that in order to process, for instance, 1 million information elements, each element should have the same priority to get through, which means to create a huge amount of high priority traffic — not good for the end-to-end delay. Secondly, since all information elements are potentially very different from each other, it is difficult to see the area where each element is sensitive to the delay — thereby giving a reason to priorities the data. Hence, both of these reasons contribute to a potentially erroneous prioritization of data packets in a
network. Eventually, this necessitates to have a quality metric that can measure the amount of information correctness (based on the information dynamics) and allows to see if it is worth to spend too much of resources to improve quality of the controller performance. Therefore, in order to evaluate the extent of information correctness, we make use of mmPr as a quality metric, defined in the following.

A. Defining Mismatch Probability (mmPr)

In order to model mmPr for a specific case in this paper, communication between a controller at some ReGen plant and a control-center is considered, which are located at different geographical and network locations. The control-center at certain control period accesses the dynamically changing controllers status information. This information access occurs over a shared network and thus offers stochastic end-to-end delays. Here, mmPr is defined as:

\[
mmPr = Pr(I_{cc}(t_c) \neq I_{ct}(t_c))
\]  

(1)

Here, \(I_{cc}\) and \(I_{ct}\) are the information available at control-center and the controller respectively, while \(t_c\) is the control time where the two sets of information are compared. This paper uses the proactive periodic access scheme for the controller to send its status updates [14]. In the periodic access, controller sends the state of the information (current status) to the control-center after every specified time interval (update rate), as shown in message sequence diagram in Fig. 1. This update rate is important as it can be used to determine the entire generated traffic. Notation used for the message sequence diagrams in Fig. 1 is as follows: \(D_i\) denotes the time at which \(i^{th}\) message is sent to the controller, while \(d_i\) is the delay experienced by this message. \(R_i\) is the time at which control-center (requester) needs the status information from the controller. \(E_i\) is the event detected at any time interval by the controller. The update process is assumed jointly independent to the event, delay, and request processes [14].

The case of information mismatch in TCP and UDP can be observed in Fig. 2(a) and 2(b), respectively. For TCP, in Fig. 2(a), it has been assumed that the communication connection is already established, i.e. there is no three-way handshake involved. In Fig. 2(a), \(R_1\) results in a mismatch from \(E_1\) because the time when control-center receives information, the status on ReGen controller had already changed through an event \(E_1\). Similarly, during message transmission, when a message is dropped due to, for instance, congestion in the network, it is retransmitted after a transmission timeout period. In case the congestion in network is too high and the message is dropped many times, it will be sent several times depending on the retransmission algorithm used. Although this mechanism ensures guarantees transmission of message at the control-center but at the cost of increased end-to-end delay which may become a cause of a mismatch of information between the two electric devices. This can be observed in Fig. 2(a) for information update between events \(E_3\) and \(E_4\). The suspected impact would be a wrong decision that leads to a wrong controller action. However, \(R_3\) in Fig. 2(a) succeeds in receiving correct information as no other event occurred during this period. In case of UDP, Fig. 2(b), a loss of information is not compensated with a retransmission. This may also become a cause of information mismatch depending on the events occurring in the controller side, as shown in Fig. 2(b).

To support our analysis, we make use of existing mmPr models, presented in [14], considering raw packet losses (UDP) versus the prolonged delay caused by packet losses (TCP) in the given model. The model for mmPr used in this paper is given in (2): [for detailed description of this model, see [14]]

\[
mmPr = \int_0^\infty exp(-\int_0^t \tau F_D(s)ds)A_E(dt)
\]  

(2)

Here, \(\tau\) is the status update rate, \(F_D\) is the CDF of delay and \(A_E\) is the CDF of backward recurrence time for an event process that is a stationary renewal process [12]. Packet losses for UDP in this model can be regarded as a thinning of the
update process, i.e. a reduction of the rate $\tau$ with the factor $(1 - P_{\text{LOSS}})$ such that $\tau_{\text{off}} = \tau (1 - P_{\text{LOSS}})$. Whereas, packet losses for TCP lead to a higher delay and in particular delay CDF. From the model shown, it is neither clear nor intuitive as to which change has the most severe impact on mmPr, and this is what we assess in section IV.

III. EVALUATION SETUP AND MEASUREMENTS

The assessment in this paper relies on measurements obtained from a network setup where, a control-center is connected to the ReGen plants, as shown in Fig. 1. The control-center and the controller within the ReGen plants are time-driven, where for each controller execution time $R_1(t_1), R_2(t_2), ..., R_N(t_N)$ the ReGen plants send updated flexibility values $D_1(t_1), D_2(t_2), ..., D_N(t_N)$. The controller execution times are equidistant $T_S = R_2(t_2)R_1(t_1) = R_N(t_N)R_{N-1}(t_{N-1})$, as well as the offset times $T_{\text{offset}} = R_1(t_1)D_1(t_1)$ denoting the time gap between sending the updates at the ReGen plants and the control-center execution times. The control step is $T_S = 1$ second, which corresponds to second-level scale necessary for load frequency control [18]. The offset value is considered to be constant and equal to $T_{\text{offset}} = 0.5$ seconds [8].

Communication between ReGen and the control-center is established first via TCP and then with UDP connection. With both TCP and UDP sockets, end-to-end delays at multiple packet loss rates are captured. The mmPr, defined in Section-II, is based on end-to-end delay measurements recorded from the time when ReGen plant sends a status update packet out to the time it is received at the control-center.

A. Simulation Setup

Analytical modeling of TCP throughput delay in [19] provides a mathematical model to compute delay for bulk TCP data. However, using the model in [19] and other such models for a small size of packets (in the order of few hundred bytes or less) and periodic interval of 10 seconds do not serve the purpose. This is because the sender congestion window for such settings does not exceed a certain limit. Normally, a packet would take half the duration of RTT to be received (at the receiver) if successful and would take duration of Timeout if packet is lost. Considering such settings, a network simulator can provide a mechanism to measure the delay that a packet of few bytes experiences to get across from sender to the receiver.

OMNeT++ is, therefore, used as a network simulator to obtain end-to-end delay traces. In order to obtain realistic traces of communication delays, a 3G network is realized offering gross data rate up to 200 kbps. The two communicating entities are placed for simplicity of setup at length of only 10 meter. Instead of attenuation and noise, we in this paper only focus on cross-traffic. By mimicking the information exchange shown in Fig. 1 and Fig. 2, end-to-end delay measurements have been collected from ReGen to the control-center.

The delay traces are obtained with different pairs of linearly distributed link propagation delays ($D_i$) and packet error rates (PER) for TCP as well as UDP, using identical network environment. For each pair of $D_L$ and PER, a set of 100 messages, each of 100 bytes in length at an exponentially distributed period of 10 seconds, was sent from controller to control-center to capture end-to-end delay traces. A set of 100 messages has been considered, specifically, to take the CDF of the end-to-end delays instead of one single delay for each group of network parameters. In case of TCP, this set of 100 messages are sent under a single three-way handshake. This assumption is made because the purpose here is not to get the exact model of TCP, but to get the potential solution of the impact that the additional delays have on the mismatch probability.

The end-to-end delay traces were then used to determine mmPr by comparing the time of reception of information with the exponentially distributed random events generated at a specific mean interval. Based on the results of mmPr, a comparison has been made to see which of the two protocols provide better performance in terms of information accuracy at different propagation delays and packet loss probabilities. Fig. 3 shows the complete simulation layout.

IV. PERFORMANCE STUDY OF TCP VERSUS UDP

A. Analysis by Simulation

This section presents the results of simulations, showing the impact of TCP and UDP on mismatch probability at two different packet loss probabilities. The results obtained via delay traces are also compared with those obtained through the analytical model (shown in green). A trade-off between the packet losses and added delays to the information will be presented in the end that would help selecting a protocol for the time critical message types in smart grids.

We estimate the mismatch probability by comparing at time instances of information access with the actual value. The average of mismatches yields the mmPr estimate:

$$\hat{\text{mmPr}} = \frac{1}{N} \sum_{i=0}^{N} I(I_{cc}(t_c) \neq I_{ct}(t_c))$$

Fig. 4 shows the impact of increasing propagation delay on mismatch probability (mmPr) considering perfect conditions for the network with no loss of information (i.e. PER = 0). It can be observed that under this condition TCP and UDP show the same performance approximately around 2.5 seconds of the propagation delay. However, as the delay continues to increase, the mmPr for TCP increases abruptly as compared to UDP. As the network is loss-less, this difference is because of the transmission timeout of TCP. Analytical results for UDPs mmPr in Fig. 4 clearly indicate that for low propagation
delays it gives the same results as obtained from the simulation model, but then deviates a little. This deviation is because the simulation results are gained from a combination of delay traces, and we expect the delay distribution to be slightly different from exponential distribution. This, nevertheless, matches the conclusions drawn in [14] i.e. more deterministic the distribution, higher the mismatch probability.

The difference of mmPr between TCP and UDP becomes more prominent as more and more packets are dropped i.e. higher PER. This is shown in Fig. 5 where the rate of packet loss in the network is increased to 10%. In case of TCP, each time a packet is lost, it is retransmitted causing delay in the information packet. The retransmitted information packet from the controller may become outdated for the control-center, causing mismatch of controller state information. The huge variation in the results is due to the variation in the mean end-to-end delays that depends upon the time a packet is lost during transmission as well as the cross traffic involved. However, in case of UDP, packet losses have no significant impact on mmPr. If any information packet is lost during transmission, the next request message can recompense the job of getting latest information, as observed in Fig. 5. It is also important to note that percentage of packet losses is higher for TCP than UDP simply because, TCP has more number of packets for request/responses due to the acknowledgement mechanism. It can, therefore, be concluded that higher packet loss probabilities cause TCP performance to degrade faster than UDP.

B. Trade-off between packet loss (PLOSS) and delay

Considering a simple case with delay and event, inter arrival processes are exponentially distributed with rates $\lambda$ (event) and $(delay)$, the following expression of the mismatch probability has been derived from (3) (for detailed derivations see [14]):

$$mmPr = \phi e^{\psi} \frac{\Gamma(\phi + \psi)}{\psi^{\phi+\psi}} \frac{F_{\Gamma(a,b)}}{\Gamma(a,b)}$$

(4)

With $\phi = \lambda/v$, $\psi = \tau_{eff}/v$ and $F_{\Gamma(a,b)}$ the CDF of a gamma distribution with parameters $a$ and $b$. Here, the value of deterministic delay is given by $D \equiv 1/v$, while $\tau_{eff}$ is given by $\tau \times P_{LOSS}$.

The important aspect to notice here is that the mmPr in reality is a complex function of ratios between the update rate, event rate and the delay rate, respectively. We, therefore, use this model to elaborate the trade-off between delay and packet losses reducing the effective update rate. This later is mapped into a comparison between UDP and TCP performance, since as (3) shows, the distribution of the delays (and events) are also important and for TCP these are certainly not exponentially distributed as we assume in the simple analysis.

The results shown in Fig. 6 illustrate the trade-off between packet loss probability and the delay it takes to achieve same level of mismatch probability. The point we make here is that for a given $P_{LOSS}$ probability, UDP leads to an effective reduction of update rate, which ultimately reduces the mmPr. If for the same $P_{LOSS}$ we use TCP (where the packet loss is reduced to zero through retransmissions), the plots in Fig. 6 show the mean delay that TCP should attain if the same mmPr should be achieved. Therefore, if a reliable protocol (e.g. TCP) can do this faster, then this protocol outperforms the UDP, and if it is slower, then UDP performs best.

Comparing these to the results with Fig. 4 and Fig. 5, it seems that in general the TCP is above the timely threshold shown in Fig. 6, indicating that the TCP assessed is in general performing poorly in the situation of sending dynamic data over network. However, as the plots in Fig. 6 also indicate, there is room for adjusting e.g. timeout values to accommodate for the losses for slow dynamic information (lambda = 0.1 events/sec), where a significant amount of time can be spent on retransmission before it no longer pays off. For faster information dynamics (here, 1 event/sec in average) there is so little time in overhead that it is very unlikely to be possible. The TCP timeouts in relation to this trade-off will be focused...
in our future research. On the other hand, other protocols may also be designed, e.g. multiple transmission of same information via UDP which reduces packet losses to nearly zero at the cost of a potential added delay, as long as the complete message transfer delay (end-to-end) is kept below the shown graphs, or else a simple UDP based protocol suffices.

Fig. 6. Trade-off between packet-loss and end-to-end delay for a given propagation delay at minimum and maximum event rates (lambda).

V. CONCLUSION

By using the concept of mmPr as a quality metric, this paper investigates how the selection of transport layer protocols effects the quality of information received and that the selection of UDP or TCP is basically a trade-off between loosing data in the network or accepting much higher delays in reception of data, respectively. It has also been shown that the trend in mmPr for UDP remains approximately the same for all cases of packet loss probabilities from which it can be concluded that UDP should be preferred for time critical message transmissions in smart grids compared to the standard TCP model assessed. However, the second aspect of this analysis shows that TCP is most suitable for information which changes slowly, and that there is room to adapt TCP e.g. by adjusting timeout values to achieve better performance. The analysis made in this paper, however, forms a basis to optimize the performance of IEC standards that defines requirements for communications in a substation.

The information access scheme used in this paper was proactive with periodic updates. This leads to the future direction of studies e.g. performance evaluation of transport layer protocols based on reactive (request-response based) and proactive access with event driven updates. Moreover, based on the verification of these simulation results in a Real-Time Hardware-in-the-Loop environment using IEC-61850, an adaptive algorithm will be developed by modifying the current communication protocols.

ACKNOWLEDGMENT

Energinet.dk is acknowledged for funding this work in contract number: PSO project 2015 no. 12347, Ancillary Services from Renewable Power Plants (RePlan), www.replanproject.dk

REFERENCES