Internet Comunication in Real Time Systems

IONUT RESCEANU, MARIUS NICULESCU, CRISTINA PANA
Department of Mechatronics
University of Craiova
Address Bvd. Decebal Nr. 107
ROMANIA

Abstract: - There is a growing need for humans to perform complex remote operations and to extend the intelligence and experience of experts to distant applications. A blending of human intelligence, modern information technology, remote control, and intelligent autonomous systems is required. Traditionally, TCP has been considered unfriendly for real-time applications. The conventional wisdom is that TCP may be inappropriate for such applications because its congestion controlled reliable delivery may lead to excessive end-to-end delays that violate the real-time requirements of these applications. This has led to the design of alternative UDP-based transport protocols that favor timely data delivery over reliability while still providing mechanisms for congestion control. A new approach for solving one of the fundamental problems facing teleoperation and real time systems is discussed: the need to overcome time delays due to telemetry and signal propagation.

Key-Words: - **Web, Internet, Control, TCP/IP, UDP, Real-Time, Monitoring, Telerobotics**

1 Introduction

In order to overcome the limitations of net-bandwidth in real-time monitoring tasks, mobile agent technology can be used.

In a remote monitoring process, the signal data obtained from real machines are transferred to and processed at a remote host in the network through long distances. In a complex monitoring system, in order to make a precise prediction, large volumes of signals are measured and transferred. In reality, however, the network transportation speed is not satisfactory. Due to the latency, congestion, and instability of network transfers occurring either in critical real-time systems or if the machines that have encountered some emergencies are located far away from the monitoring system, commands from the monitoring system might not be transmitted to the remote machines within the required time periods. Our research reveals that real-time applications performance over TCP may not be as delay-unfriendly as is commonly believed. One reason is that the congestion control mechanism used by TCP regulates rate as a function of the number of packets sent by the application. Such a packet-based congestion control mechanism results in a significant performance bias *in favor* of flows with small packet sizes, such as VoIP.

Second, due to implementation artifacts, the average congestion window size can overestimate the actual load of a rate limited flow. This overestimation results in reduced sensitivity to timeouts and an improvement in the delay performance.

Teleoperation is the remote control of robot manipulators. Although commands can be sent from user to remote robot at different levels of abstraction, this section describes a kind of low-level teleoperation in which the human directly controls the motions and contact forces of the remote manipulator in real time. Perhaps the most common application of this technique is in construction equipment such as excavators in which the operator controls the velocity of the joints of the "robot" to accomplish the task. However construction equipment does not provide force feedback directly to the hand. When the user is located farther from the remote robot, considerable engineering effort must be applied to reproduce the sensory feedback information which allows accurate and efficient control. Both teleoperation and virtual environments require this rich and self-consistent sensory feedback. Haptic feedback devices were pioneered in teleoperation systems as far back as the 1940's. In both teleoperation and virtual environment applications of haptics, a loop is closed between the human operator's motion "inputs" and forces applied by the haptic device. In teleoperation this loop is closed via a communication link, robot manipulator, and the environment. In virtual environments, the loop is closed via a computer simulation.

Key issues for the advancement of teleoperation technology include:
Performance Evaluation: What quantitative measures can be developed with which to quantify the quality of a teleoperation system (including haptic displays)? Control: How can stable, high performance, control be obtained in spite of highly variable human operator and environment dynamics, time delays in communication channel, and kinematic effects such as singularities?
Scaling: What are the requirements for effective user interfaces when there is a large difference in scale (either up or down) between the master (human operator) and slave (remote robot)?
Mechanization: High quality teleoperation and haptic interaction depends critically on advanced mechanism designs for both master and slave sides. Key issues are light stiff structures and linkages, actuators with high torque/mass ratios and high linearity, compact, high resolution sensors for position velocity and force/torque.
Kinematics of Teleoperation: How can effective use be made of redundant degrees of freedom in teleoperation systems (i.e. when the number of slave DOF > number of master DOF)?
The field of telerobotics grew out of the need to perform operations where it is difficult to place a human being, due to constraints such as cost, safety or time. Telerobotic systems need to be able to perform tasks that a human would normally do. Due to limitations in robot autonomy, this often has to be achieved by using human operators to control the remote robot (via a communication link). Such a system is a telerobot. The human operator is responsible for high level control such as planning and perception, while the robot performs the low level instructions at the remote site.
An aspect of web telerobot systems that affects the choice of control scheme is time delay. Shared continuous control is less sensitive to these problems and has been demonstrated over the Internet, but only on short, high bandwidth Internet links. Discrete command control schemes and above are free of any time delay based instability problems as all closed loop control is performed locally. They are therefore the most appropriate choice for web telerobotic systems.
In order to overcome the limitations of net-bandwidth in real-time monitoring tasks, mobile agent technology can be used. Within the scope of the work described in current research on mobile agent's technology, a mobile agent can be defined as ''a software agent that is not bound to the system where it begins execution.'' A monitoring program can be dispatched as a mobile agent to the host that has sent the requirements to request for this monitoring program.
Operational failures of equipment may not only lead to a loss of production, but also, in some serious situations, may cause human casualties. Hence, equipment condition monitoring and fault diagnosis are often employed in maintenance to prevent operational failure of services and the fatal breakdown of manufacturing equipment.
The technology for equipment fault diagnosis has been developed in response to the demands of modern industry that has been concerned with such aspects as human safety, economic productivity and effectiveness. Machine fault diagnosis refers to the process of identifying a machine's operating condition and investigating its possible source of fault [1]. This is usually done in a way that is similar to medical diagnosis. Both of the diagnoses have to observe symptoms by sensing and analyzing signals collected either from a human body or a piece of equipment.
For machine-based fault diagnosis, the collected signals may be operating force, vibration, temperature, voltage, pressure, etc., any of which could be related to the inherent conditions of the machine [1]. Good diagnostic systems can provide managers and operators with the necessary information to determine the running condition of the equipment.
The concept of Web-based remote monitoring and a collaborative diagnostic system was initially proposed for medical care in 1988.
The system was designed to enable doctors or small clinics in rural areas to obtain instantaneous consultations from specialists in urban hospitals [2]. An ideal remote diagnostic system should include video-conferencing and remote measurements delivered as close to real-time as possible. It should also provide on-line fault diagnosis and support a multi-user kind of collaborative consultation. Hence, the system must install an operating system that supports a multi-tasking and multi-user operating environment. An eligible operating system must support distributive computing, cope with common servers and major communication protocols, and be easy to adapt to a variety of popular Web browser applets.
The Web allows universal access by having independent connectivity for different kinds of platforms using open standards for publishing (HTML, XML), messaging (HTTP), and networking (TCP/IP) [5]. Internet browser plug-ins should be able to handle new data types and allow different applets to be downloaded and run on any browser. A variety of software-based VIs should exist for

performing the required signal processing, features extraction, and analysis. These VIs are, preferably, to be compatible with ActiveX standards. Therefore, Web-enabled VIs can be operated as browser applets/ActiveX in a multi-user environment.

Since the Web enables multi-media support, both interactivity and extensibility [6], it can seamlessly include new forms of content and media [7]. Therefore, Web-based maintenance should employ multi-media to a large extent. Nowadays, broad bandwidth communication is available in many countries for use by Web-based multi-media. The developments in database and object technologies, such as CORBA, IIOP and component-ware concepts, enable users to connect to back-end databases and legacy applications via user-friendly Web interfaces [3]. All of these features make the development of Web-based maintenance and an interactive type of collaborative maintenance feasible.

To ensure future compatibility and allow for expansion of the sensors that will be used on the Internet, the standard for smart transducers, IEEE 1451.2, has been formulated for the design of next generation Web-ready sensors. The future smart transducer will have a built-in Ethernet module and support direct plug-and-play on the Internet without the need for a connection to a PC or having a separate Ethernet card, as is the case with today's systems. Sensor manufacturers, such as Hewlett Packard and Bruel & Kjaer, have already proclaimed that their new directions in designing sensors are based on the standard for smart transducers [9]. With the help of these Web-ready sensors, Web based maintenance becomes easier to implement, using fewer resources and involving less capital cost.

Even though the research on Web-based services and systems is advancing every day, progress in research on Web-based maintenance is lagging behind. As has already been mentioned, the three crucial functions for Web-based maintenance are, remote data sensing with the help of the mini-server, signal processing and fault diagnosis using Web-enabled VIs, and collaborative maintenance platforms for multiple users. Research has been conducted to target one or two of the three functions. However, there is a deficiency in developing a Web-based maintenance system that has all three important functions combined together.

2. Internet Delays and Real-time machine monitoring

Neural networks and expert systems techniques are finding increasing application in the field of machine tool diagnostics.

2.1 Transmission Delay

Transmission delay, sometimes also termed as latency, of a communication line is the time from the start of data packet transmission at the source to the start of data packet reception at the destination. The source of the latency can vary from the speed of the signal to how the signal is relayed among various gateways. In many instances this delay can become significant enough to become noticeable by a human.

Dedicated algorithms are specially designed for and tested with the Internet as the communication medium, although in principle they are applicable to any sort of transmission delay. The Internet is a complex network of servers and clients where data transmission is not direct but is forwarded over many links via many gateways. This can produce significant latency especially at certain times of the day with heavy network congestion and in areas with poor network infrastructure. The unpredictability of the Internet can result in variations in latency as well as lost data packets. To perform bilateral teleoperation over the Internet, the system must solve the problems posed by these pure delays in what is effectively a closed-loop system.

Propagation delay is unavoidable in a tele-operated system (due to the limitations of the speed of electricity, the speed of light). Preliminary studies have shown that a transmission delay of 200mS or more leads to problems with surgical accuracy and precision.

2.1.1 Fixed Delay

Propagation delay is unavoidable in a tele-operated system (due to the limitations of the speed of electricity, the speed of light). Preliminary studies have shown that a transmission delay of 200mS or more leads to problems with surgical accuracy and precision.

The International Telecommunication Union Telecommunications Standardization Sector (ITU-T) noted that with voice calls, most callers notice round trip delays when they exceed 250mS. As a result the ITU-T G.114 recommend the maximum desired one way latency to achieve high quality voice is 150mS. With Round Trip Time (RTT) delays of 500mS or more, a 'natural' phone conversation becomes very difficult. These figures for voice traffic will be used as an initial starting guide for our implementation of timely and realistic force feedback.[10]

ISBN: 978-960-6766-79-4

ISSN 1790-5117

2.1.2 Random Time Varying Delay

The Internet is a best effort service that offers no upper bound to response time or bandwidth guarantees. The result is a service that is time varying in a random nature. This fact introduces an extra level of complexity in the teleoperation of a system. A control engineer can deal with the problem of compensating of a constant delay with relative ease. However, a random time varying delay is very difficult to compensate for. Such a situation can often result in destabilizing the overall system. The key to timely and stable control of a closed loop system over the Internet is to effectively reduce the variance of the delay.

The problem of controlling a real time tele-system using the Internet as the link has been studied extensively over the past few years. Most researchers have tended to use TCP/IP (with its inherent short comings in ability to deliver data in a timely fashion), seemingly without firstly looking deeply into the IP protocols options available.

Generally this past research has concentrated on a variety of complex control methods in order help stabilize a telecontrol system in the presence of TCP/IP delay. Essentially the results have traded off a large amount of system response (delays of 5-6 seconds are not uncommon) in order to achieve stability.

2.2 TCP/IP Delays

TCP is window-based protocol that uses a feedback-based rate regulation scheme. The idea is to use a congestion window to regulate the amount of data that can be outstanding in the network at any time. TCP relies on losses as congestion feedback. It uses two mechanisms to detect packet losses: fast retransmit and timeout. If the sender receives three duplicate acknowledgements, it assumes that the data indicated by the acknowledgements is lost and immediately retransmits the missing data. This mechanism is called fast retransmit. After sending the missing data, TCP uses the fast recovery algorithm to govern the transmission of new data until a non-duplicate ACK arrives. Due to the roundtrip time (RTT) needed for the receipt of the loss indication, fast recovery typically takes on the order of the path's roundtrip-time [8].

The other loss recovery mechanism provided by TCP is the timeout mechanism. The TCP sender sets a timer for each transmitted packet. In case it receives less than three duplicate ACKs and the timer expires, the packet is retransmitted and the window size is reduced to one. If the first packet after a timeout is lost, TCP will double the length of the next timeout period.

Possible congestion due to TCP traffic flows is controlled by the congestion control mechanism that is native to TCP. This congestion control can inflict serious problems on real time applications. In addition to this, TCP has an error correction arrangement in the forms of:

- Ordered delivery
- Duplication detection
- Crash recovery
- Retransmission strategy

By TCP addressing these above issues, TCP offers a guarantee for the reliable transport of packets to destination, thus, shielding the data users from the unreliable nature of the underlying IP network. The downside is the fact that these flow and error control techniques employed by TCP present a major obstacle to achieving time guarantees over the Internet. For example, the TCP slow start mechanism is used to discover the channel throughput during the initial connection setup and for resumption of a broken connection. This is done by first sending a packet across the channel and waiting for a response. If a response is received, the next packet is sent a bit faster. This procedure is repeated until the speed of the link is discovered. With the half-second delay between responses, throughput is significantly slowed.

Since this process can take 7-15 Round Trip Times, for a link with a propagation delay of 500ms this can mean that for 3-7 seconds, the link is underutilized. (See figure 1)[7]

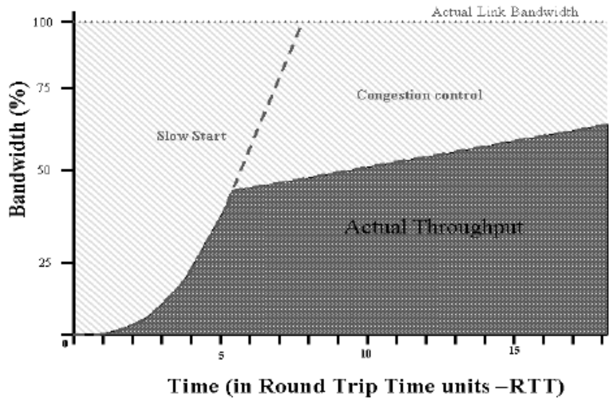

Fig 1. TCP/IP True Throughput

To further add to overhead, every TCP connection is established by a "3-Way Handshake" between the Receiver and Sender. On links with long propagation delays, this fixed overhead means that even very short data exchanges take at least a few seconds to be completed. Data links can be noisy, and this has profound effects on the performance of TCP/IP throughput because the slow start congestion control mechanism wrongly detects the noise as network congestion. Hence, from a real time viewpoint, TCP fails to provide an adequate solution, largely due to the enormous processing overhead it employs in order to provide a reliable path for data.[14]

2.3 UDP/IP and RTP/UDP/IP Delays

Seemingly, none of the present online teleoperated systems to date have used Real Time Protocol (RTP) running over User Datagram Protocol/ Internet Protocol (UDP/IP). This is probably largely due to the fact UPD/IP is seen an unreliable data medium, whereby data could arrive out of order or not at all. Even so, RTP/UDP/IP is fast becoming the popular protocol arrangement for streaming data in real time over the Internet. UDP is a connectionless protocol. This fact gives it very different characteristics to TCP. UDP is an unreliable service due to the fact that delivery and duplication of packets cannot be guaranteed. In addition it is likely that packets will arrive at the destination out of order. Even so, UDP with RTP is a far better option than TCP for real-time applications such as voice or video. Retransmission of a packet 1-2 seconds after it was sent when it contains a 20mS sample (as is the common case for voice) would produce disastrous implications to the real-time voice stream. In addition the cost in time for TCP to detect a packet loss, stop the data stream, request a resend from the point of loss and then finally receive the lost packet can be in the order of several seconds. As stated, packet loss is unavoidable with UDP/IP, but it can be compensated for in voice streaming by codec loss-concealment schemes. One such codec is G.723.1, which has the ability to interpolate a lost frame by simulating the vocal characteristics of the previous frame and slowly damping the signal. It has been shown that packet loss rates up to the order of 10 percent have little noticeable impact on the audible quality of the speech. It should be noted also that the connectionless quality of UDP/IP reduces the overhead of the protocol (from TCP/IP 40bytes to UDP/IP 28bytes) and this makes UDP a further preferred choice for constant flow applications such as multimedia and control sessions. Even though a UDP/IP implementation has a lesser header overhead than that of TCP/IP, the RTP/UDP/IP implementation returns the header overhead back to 40bytes since the RTP component adds an additional 12bytes to the header. Now 40-45 bytes of overhead would not be an issue if the data packet were in the order of 1500 bytes. The problem is that our implementation only involves packets with a data size in the order of 10-20bytes (due to the sampling rate). Hence a whopping total of 40-45bytes of overhead to transmit a 10-20byte payload. There are two possible solutions to this problem:

1. Increase packet size, at the expense of sample rate and potential delay jitter.
2. Use header compression. In the case of voice packets it has been shown that the increased delay incurred from increasing the packet size is unacceptable. For this reason a great amount of research is being undertaken into optimizing header compression.[17]

In summary, utilizing UDP/IP in place of TCP/IP will greatly increase network efficacy by:

- Removing the need for having a connection setup before data can start to flow.
- Removing the slow ramping up of
- Low rate packet loss does not halt transmission of the streaming data.[15]

In real time operations such as online gaming, some programmers would prefer to use user datagram protocol (UDP). This protocol eliminates the need for confirmation where the transmitting computer keeps sending the data packets with no regard as to whether the receiving computer has received the data. This means that all the data are sent in a timely fashion, an important feature for real time operations. But the lack of confirmation also means that it is less reliable. The transmission delay is also more stable than TCP.[16]

5 Conclusion

The Internet protocols do not guarantee a maximum delay for a message to be carried across a network link, which means that the control scheme must work under variable (and possibly large) time delays. Continuous control is not well suited, as it is prone to instability problems under time delay. In spite to all limitations, however, it is possible to realize reliable systems that in future will help in improving everyone's quality of life. In fact, remote diagnosis and rehabilitation, access to dangerous and/or remote sites will be more and more accessible and more applications are going to appear, all aimed at easing the interaction between distant worlds.

References:

[1]Allman, Hayes, Kruse and Ostermann, 1997 M. Allman, C. Hayes, H. Kruse, S. Ostermann, 'TCP Performance Over Satellite Links', '*The Fifth International Conference on*

Telecommunications Systems, Nashville,TN', March, 1997

[2]Anderson, R. J. and M. W. Spong, (1989) *"Bilateral control of teleoperators with time delay,"* IEEE Trans. On Automatic Control, vol.34, n.5, pp. 494-501

[3] Altintas Y, Yellowley I (1989) In process detection of tool failure in milling using cutting force models. ASME J Eng Ind 111:149–

[4]Munir, S. and Book, W. J., *Wave-Based Teleoperation with Prediction,* Proceedings of the American Control Conference, Vol. 6, pp. 4605-4611, June 2001.

[5]Niemeyer G. and J. J. E. Slotine, (1991*) "Stable adaptive teleoperation,"* IEEE Journal of Oceanic Engineering, vol.16, n.1, pp.152-162.

[6]Niemeyer, G. and J. J. E. Slotine, (1998), "*Towards force-reflection teleoperation over the Internet,"* in Proc. IEEE Conf. Robotics and Automation (ICRA), Leuven, Belgium, pp.1909-1915

[7]Oboe, R. (2003). *Force-reflecting teleoperation over the internet: The JBIT project*. Proceedings of the IEEE, 91(3), 449–462.

[8] 157K. Goldberg, \Introduction: The Unique Phenomenon of a Distance", in: *The Robot in the Garden. Telerobotics and Telepistemology in the Age of the Internet*, K. Goldberg (ed.), MIT Press 2000.

[9]Park J.H.,and H.C. Cho, (1999) *"Sliding-mode controller for bilateral teleoperation with varying time delay,*" in Proc. Advanced Intelligent Mechatronics Conference, Atlanta, USA, pp. 311- 316.

[10] Raibulet C, Demartini C (2000) Mobile agent technology for the management of distributed systems–a case study. Comput Netw 34:823–830

[11]Secchi, C., Stramigioli, S., & Fantuzzi, C. (2003).*Dealing with unreliabilities in digital passive geometric telemanipulation.* In Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (Vol.3, pp. 2823–2828).

[12]Secchi, C., Stramigioli, S., & Fantuzzi, C. (2003b). Digital passive geometric telemanipulation. In Proceedings of the IEEE international conference on robotics and automation (Vol. 3, pp. 3290–3295).

[13] Sheng J. and M. W. Spong,, (2004) *"Model predictive control for bilateral teleoperation systems with time delays",* Canadian Conference on Electrical and Computer Engineering, , v.4, p. 1877--1880, Tampa, Florida, U.S.A.

[14]Slama, T., D. Aubry, A. Trevisani, R. Oboe and F Kratz (2007a) *"Bilateral Teleoperation over theInternet: Experimental Validation of a Generalized Predictive Controller"*, in Proc. European Control Conference'07, Kos, Greece.

[15]Slama,T., D. Aubry, R. Oboe, and F. Kratz, (2007b*) "Robust bilateral generalized predictive control for teleoperation systems"*, in Proc 15th. Of IEEE Mediterraneanl Conference on Control and Automotion '07, Athens, Greece.

Lynn Conway, Richad Volz, Michael W. Walker

[16]Teleautonomous Systems: Projecting and Coordinating Intelligent Action at a Distance IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION. VOL 6. NO 2. APRIL 1990

[17] Venners B (1999) Under the hood: Solve real problems with aglets, a type of mobile agent. JavaWorld. http://www.javaworld.com/javaworld/jw-05-997/jw-05-hood.html