Requirements for a High Performance Transport Protocol Summary
Version 2 - 07/16/90

ABSTRACT

This paper analyzes Transport protocol requirements collected from various sources which go beyond the capabilities of present standards. Requirements considered are those of the "Transport Service User" avoiding any direct use of the services provided by existing or developing Transport Protocols. Four papers, evaluating the requirements in leading edge applications of networks, are summarized, categorized, evaluated and assessed. The four papers addressed twenty eight requirements of challenging applications in the areas of Technical Workstations, Process Control, Real-time Military Systems and Scientific Supercomputing. The paper identifies six significant requirements which appear to drive further Transport protocol development. This effort is supporting the ANSI X3S3.3 committee's study project investigating the Transport and Network standards abilities to support Very High Speed Networks. Further work is planned by this committee to analyze the significant requirements against the existing OSI protocols. The study effort is expected to end with recommendations for modifications to existing OSI standards and/or new OSI protocols. This is the second version of this paper provided for consideration by the ANSI X3S3.3 committee.

1. Introduction

The ANSI X3S3.3 High Speed Networking project has been performing a study to ensure the viability of OSI Network and Transport protocols in the very high speed networking environment. This paper addresses the requirements as seen by Transport protocol users independent of the capabilities of existing or developing Transport protocols. The applications identified included the areas of Technical Workstations, Process Control, Real-time Military Systems and Scientific Supercomputing. The emphasis is to identify the requirements which go beyond the capabilities of existing standards (i.e. TP4 and ISO Connectionless Transport). The paper provides some observations brought out by discussions carried out by the ANSI X3S3.3, reviews four papers on requirements, categorizes the requirements, evaluates their impact to using existing Transport protocols, assesses the significant requirements and then provides conclusions.

2. Observations from Committee Discussions

This sections provides summary points on requirements for Transport protocols that have been brought out by discussions within the ANSI X3S3.3 High Speed Networking project.

2.1. Paradigm Shift

The paradigm assumed for computer communications has undergone major changes in the last few years. In the past the paradigm involved large centralized computers working at a site interconnecting with other sites separated by long distances communicating with noisy 56KBit/second lines.

The environments considered by all of the applications discussed in this paper involve many computational elements (i.e. workstations) located in close physical proximity. The availability of inexpensive microprocessors has revolutionized the computing industry over the last ten years and now the new computing environment is changing the nature of data communications. While today clusters of these processors are located within a floor or throughout a building, in the near future it is expected that such clusters may be spread across a metropolitan region and will soon span large distances. Thus while today LANs are typically used within such applications soon both metropolitan area networks and wide area networks will need to support these applications.

The paradigm shift encompasses all of the OSI Basic Reference Model layers from the Transport layer down. At the Data Link layer the use of very high speed LANs (Local Area Networks, i.e. ISO FDDI) as well as developing MANs (Metropolitan Area Networks, i.e. IEEE 802.6) and ATM are providing new capabilities. At the Network layer interconnection of multiple high speed LANs, MANs and WANs will be required with performance that matches that achieved by these Data Link networks. The Transport layer requirements is covered by this paper.

2.2. Bandwidth Metric

Initial discussions within the subgroup focused on bandwidth capabilities of the existing TP4 protocol. With relatively minor modifications (i.e. moving the checksum to the end of packets and utilizing one set of options) TP4 has been run at very high speed when using very large packets. Similar results have been reported by the TCP community. The committee agreed that for point-to-point transfer between two devices with very large amounts of data to transfer, TP4 appears adequate given the minor modifications (which could possibly be handled through the ISO 8073 incremental improvement effort already in-process).

2.3. New Functional Capabilities

The new requirements identified for Transport protocols came from very challenging applications having many processor based devices communicating among themselves (e.g. via a high performance LAN interconnect). A need for multicast and reliable multicast communications was identified. A need to control maximum latency for some messages at the cost of increased latency for others was identified as well as the ability to transfer small high priority messages in the middle of on-going large file transfers. Selectable error control, efficient datagrams and support for time synchronization were also identified.

New functional capabilities are by their very nature beyond the scope of the present protocols. Further study is required to determine if individual capabilities can be added to existing protocols or if there is a paradigm conflict which precludes its addition.

2.4. Implementation Efficiency

During the meetings protracted discussions concerning implementation efficiency were carried on. The issue presented was reduced protocol complexity in order to achieve higher performance (i.e. faster execution achieving lower transfer latencies and performing more operations per second). A number of "measures" of complexity were proposed (i.e. number of C language instructions to carry out a critical function, die size of an integrated circuit and electronic board "real estate" consumed), however all were deemed arbitrary and none were generally accepted by the committee. It is proposed here to utilize "number of decision points" to carry out a critical function. While not perfect this provides a measure which is applicable to any implementation method.

A number of members of the committee expressed doubts as to the importance of implementation efficiency at a time when processors are getting faster and memory is getting cheaper; however, data rates by high speed networks are also growing at least a similar rate and for the application areas considered in this paper it appears justifiable that 10% or greater potential efficiency improvements are worth pursuing.

The means to achieve such efficiencies is by cutting the number of options within the "critical path" of the protocol to a minimum and by fixing such options to a point within the header or trailer. The placement of a field within a protocol data unit (PDU) can be very important; for example, placing the checksum in the trailer of a PDU has a major advantage when calculating a checksum on a PDU received.

A disadvantage in optimizing the protocol for efficiency is that it takes away the tremendous flexibility in calling out (and later adding) options that protocols such as TP4 have where an "ASN.1" approach is used.

2.5. Use of "Hard" Numbers

In the beginning of the requirements study it was anticipated that "hard" numbers would play a major role in defining the requirements. Examples of such numbers are a one millisecond maximum latency requirement for a complete transaction or a requirement to set up a connection in ten milliseconds. In pursuing the requirements study it was soon determined that numbers such as these exist but are typically events at the application layer which do not directly translate into "hard" numbers at the Transport layer. Numbers considered are typically from existing systems which may have been designed a number of years ago before the technology assumed by this effort became available. Worst case "hard" numbers which will hold over the next twenty years are not obtainable.

The approach proposed is to bring out numbers wherever they can be found as "ballpark" estimates of what is foreseen, but assume that as technology improves such numbers will be stressed more. Thus as LANs go to 100 Megabit/second and beyond and point-to-point communications go to 800 Megabit/second and beyond that very capable applications will be developed which will stress these technologies and continue to require more.

3. Documentation Summary

This section lists the requirements described in the primary reference papers. Details concerning these requirements are described in the papers. The requirements listed in the referenced papers are those that the authors felt were not being met by the existing standards (i.e. TP4, Connectionless Transport, TCP and UDP).

3.1. Technical Workstations

The "Notes on Technical Workstations Requirements" [1] paper identifies two primary requirements of the technical workstations industry. These requirements reflect the very nature of the very competitive workstation market where success depends on delivering the maximum performance at the lowest possible cost across the board. This paper describes the nature of the data traffic, the real drivers within the workstation market and describes six R&D issues for workstation Transport protocols. The six R&D issues are new functional capabilities which are being studied but no definitive trend can be identified at this time to consider these as requirements.

The two requirements listed are:

TW1) Deliver excellent performance on a wide range of traffic types with the minimum number of Transport protocols

TW2) Deliver maximum performance at the lowest possible cost

3.2. Process Control

The "User Requirements for Communications in Time Critical Applications [3] paper identifies ten requirements for the process control and manufacturing environment's command and control networks which are not being met by MAP 3.0. "Command and control networks interconnect devices such as PLCs, robots, CNCs vision systems and supervisory computers. Traffic on this type of network is essentially real-time, i.e. submitted to tight deadlines".[3]

The ten requirements listed are:

PC1) Prioritization of Messages at the Application Interface

PC2) Predictable Request/Response Times Observable at the Application Interface

PC3) Selection of Error Recovery by the User

PC4) Support for Multicast Communication

PC5) Support for Redundancy

PC6) Existence of Security Mechanisms

PC7) Interworking with Non-Real-Time MAP 3.0

PC8) Low Attachment Cost

PC9) Support for Time Synchronization

PC10) Quality of Service for Congestion Recovery

3.3. Real-Time Military Systems

The "Requirements for a High Performance Transport Protocol for Use on Naval Platforms"[2] paper identifies nine requirements from the real-time military environment. This paper describes a Tactical Console Display scenario in which many Display Workstations communicate among themselves, with a Common Data Base Control computer, with various file servers and with controlled subsystems. It is postulated that such requirements may exist for some non-military systems such as air traffic control.

The nine requirements listed are:

MIL1) Multicast Transfer

MIL2) Reliable Multicast Transfer

MIL3) Real-Time Scheduling

MIL4) Very Fast Connection Build-up and Tear Down

MIL5) Limited Routing Capability

MIL6) Special Services (i.e. Time Synchronization and Distributed Transaction Support)

MIL7) Reliable Datagrams

MIL8) Selectable Error Control

MIL9) Support for the Conservation of Local Resources

3.4. Scientific Supercomputing

The "Delta-t Transport Protocol: Features and Experience"[4] paper identifies seven requirements from the scientific supercomputing environment. A scientific supercomputing environment is found at a research facility which utilizes a few supercomputers along with many computers of all sizes all of which may communicate. The Transport protocol is viewed as a component to support a distributed operating system across such a facility.

The seven requirements listed are:

SC1) Minimum Packet Exchange for Request/Response Transactions

SC2) High Throughput Bulk Data Transport and other Stream services

SC3) Flow Control Without Polling for Reliable Zero Window Opening

SC4) Error Control of Lost, Damaged, Duplicated, and Out-of-sequence Packets

SC5) Large and Flexible Name Space for Transport End Points

SC6) Message Boundary Preservation

SC7) Secure Communications

3.5. Additional Requirements Identified

This section contains additional requirements identified outside of the referenced papers that form the primary basis for the requirements presented. Each requirement listed in this section lists where it was identified and provides additional descriptions (due to the lack of more detailed reference documents).

AD1) Capability of Working in a Global Heterogeneous Internet

This requirement was identified at the September 1989 X3S3.3 HSP meeting in Durham, N.C.. A consensus of the group assembled agreed to this as a requirement. Thus for any protocol recommendation that comes out of this SD3 Study a requirement exists that it must be capable of working along with existing and developing international standards.

AD2) Congestion Management

This requirement was identified at the June 1990 X3S3.3 HSP meeting in Boston, MA.. This is a major factor being considered in the development of new networking technologies in the telecommunications arena. The three primary issues are congestion avoidance, detection and recovery. It was pointed out that a network experiencing congestion can not be considered high speed by its users at that time and thus a question arose as to whether this was not a concern "by definition". The consensus of the group appeared to be that this was an area of concern and should be addressed by this Study effort.

4. Requirement Categorization

The section breaks into categories the requirements listed in the previous section. Requirements are combined wherever possible. It is the goal of this section to provide an organized listing of the Transport level requirements.

4.1. New Functional Capabilities

This section combines requirements which are not part of the present Transport protocols (i.e. TP4 and Connectionless Transport). Further analysis is required to determine if the capabilities listed here are outside of the present Transport protocol paradigm where a change in paradigm necessitates a new protocol.

NF1) Efficient Transactions. Incorporates MIL4, MIL7 and SC1.

NF2) Multicast Transfer (Both Unreliable and Reliable Options). Incorporates PC4, MIL1 and MIL2.

NF3) Selectable Error Control. Incorporates PC3, PC10 and MIL8.

NF4) Latency Control Facilities. Incorporates PC1, PC2 and MIL3.

NF5) Support for Time Synchronization. Incorporates PC9 and MIL6.

NF6) Large and Flexible Name Space for Transport End Points. Incorporates SC5.

NF7) Support for Redundancy. Incorporates PC5.

NF8) Distributed Transaction Support. Incorporates MIL6.

NF9) Support for Limited Switching and Relaying at the Sub-net Level. Incorporates MIL5.

NF10) Congestion Management. Incorporates AD2.

4.2. Efficient Protocol Implementations

The requirements listed in this section can be viewed as a need to ensure that the performance being provided by new standards at lower layers (e.g. FDDI and HSC) are being carried through the Transport layer. As technology at these lower layers progress, performance must be carried up.

EP1) Maximize Efficiency (over a wide range of traffic types, lowest possible cost and concern for conserving local resources). Incorporates TW1, TW2, MIL9

EP2) High Throughput Bulk Data Transport and other Stream Services. Incorporates SC2.

4.3. Design Features

The requirements listed in this section are important features needed in developing a Transport protocol. These features may already be incorporated or could be incorporated (through the on-going standard maintenance efforts) within the existing Transport standards.

DF1) User Selectable Security Options. Incorporates PC6 and SC7.

DF2) Flow Control Without Polling for Reliable Zero Window Opening. Incorporates SC3.

DF3) Error Control of Lost, Damaged, Duplicated, and Out-of-sequence Packets. Incorporates SC4.

DF4) Message Boundary Preservation. Incorporates SC6.

4.4. Interoperability Capabilities

The requirements listed in this section provide a context for any standard considered by this effort to exist within.

IC1) Capability of Working in a Global Heterogeneous Internet. Incorporates AD1.

IC2) Interworking with non real-time MAP3.0. Incorporates PC7.

4.5. Other Requirements

This section lists any requirement which does not fit in one of the above categories.

OT1) Low Attachment Cost. Incorporates PC8.

5. Requirement Evaluation

This section takes the requirements that have been described earlier and evaluates them according to their impact on the present study effort. The primary issue addressed here are to identify the requirements that significantly impact the present standard Transport protocols. The classification for the requirements are (by class number):

5.1. Class 1

These requirements are considered to be significant enough to drive a new SD3 development proposal for modification to existing OSI standards and/or new OSI protocols for the support of OSI higher levels.

NF1) Efficient Transactions

NF2) Multicast Transfer (Both Unreliable and Reliable Options)

NF3) Selectable Error Control

NF4) Latency Control Facilities

NF5) Support for Time Synchronization

EP1) Efficient Implementation

5.2. Class 2

These requirements are an important part of any Transport protocol considered by the SD3 study task but are not expected to drive towards a new SD3 proposal.

DF1) User Selectable Security Options

DF2) Flow Control Without Polling for Reliable Zero Window Opening

DF3) Error Control of Lost, Damaged, Duplicated, and Out-of-sequence Packets

DF4) Message Boundary Preservation

EP2) High Throughput Bulk Data Transport and other Stream Services

IC1) Capability of Working in a Global Heterogeneous Internet

5.3. Class 3

Those requirements needing further investigation.

NF6) Large and Flexible Name Space for Transport End Points

NF7) Support for Redundancy

NF8) Distributed Transaction Support

NF9) Support for Limited Switching and Relaying at the Sub-net Level

NF10) Congestion Management

5.4. Class 4

Those requirements viewed as being out of scope for the present requirement study.

IC2) Interworking with non real-time MAP3.0

OT1) Low Attachment Cost.

6. Significant Requirement Assessment

This section further discusses the requirements evaluated to be significant (i.e. Class 1) as to their importance to each of the application areas. Table 1 identifies the requirements by application area.

Table 1
Significant Requirement Assessment

Requirement	Workstation	Process Control	Military	Supercomputing
NF1) Efficient Transactions	X		X	X
NF2) Multicast Transfer		X	X	
NF3) Selectable Error Control		X	X	
NF4) Latency Control Facilities		X	X	
EP1) Efficient Implementation	X	X	X	X
NF5) Time Synchronization		X	X	

Table 1 can be used to cross-reference the significant requirements to the application areas that they are needed in. During the ANSI X3S3.3 committee discussions a request was made to prioritize the requirements. Such a prioritization does not appear meaningful, what does appear useful is to consider the significant requirements by application area. Using Table 1, it can be determined which significant requirement is important to each application area. Within the application areas the requirements are inter-related and thus prioritizing is not an issue.

The study effort found a similar set of requirements for both the Process Control and Real-Time Military application areas. The time critical nature of these environments have brought out a similar need for new Transport level functionality of Multicast Transfer, Selectable Error Control and Latency

Control features. The Time Synchronization requirement which is needed by both of these application areas is outside of the normal Transport protocol issues.

6.1. Efficient Transactions

The Efficient Transaction is a new function proposed to handle a function occurring increasingly in the application areas considered here. A request by a service user causes an indication to a user(s) at a remote processor, upon completing an operation this remote user issues a response to its service provider which ends with a confirm to the initially requesting service user. This four legged exchange is used for Remote Procedure Calls (RPCs) and other Client-Server exchanges.

This exchange can be accomplished using TP4; however, a number of operations are required such as a four legged exchange to first set up a connection, a data transfer phase and a disconnect phase. Using this technique results in many communication actions over the network which can greatly increase latency. Latency and number of transactions per unit of time are the primary metrics for this type of exchange and thus efficiency in carrying out this function is important. In addition if implementations use multiple service requests between service user and provider than an increase in latency can result which is detrimental to time critical systems. An additional requirement discussed in the Workstation application area was that the data sent via the transaction not be restricted in length. Thus the data may need to be passed via many PDUs at the lower layers.

6.2. Multicast Transfer

The Multicast Transfer is the simultaneous sending of the same PDU to a number of peer entities. It can be assumed for the application areas discussed here that the underlying layer services support multicast communication down to the communication media used.

The simultaneous transfer allows critical information to be distributed to all very quickly which is important in time critical applications. The application area of interest is anticipated to be larger than just Process Control and Real-Time Military applications but encompass embedded systems of all kinds that involve many processing elements. Other examples of such systems are discrete manufacturing industry, air traffic control complexes and flight simulation systems.

6.3. Selectable Error Control

Selectable Error Control provides a means for the service user to specify the amount of error control to apply to a specific Transport Service Access Point (TSAP). The two extremes are complete assurance and no error control. This feature has great utility when applied to Multicast Transfers; however utility for this feature has been seen when applied to point-to-point transfers in time critical applications. Along with the two extremes of selecting error control, intermediate levels are needed for Multicast Transfers. Another feature to provide is "hole" preservation in large transfers even when no error control is applied, this has value in transferring video images in time critical applications. The application areas of interest are the same as those listed under the Multicast Transfer section.

6.4. Latency Control Facilities

Latency Control Facilities provide a means for selecting data for transfer from one TSAP at the expense of others. In order to make the facility useful it needs to be implemented consistently (i.e. of global scope) across a system in order to achieve the desired results. A major need in time critical applications is to transfer small high priority (e.g. Efficient Transactions) messages in the middle of large on-going file transfers. Control of latency is not found in any present Transport standard. The application areas of interest are the same as those listed under the Multicast Transfer section.

6.5. Efficient Implementation

Efficient Implementation as used here refers to a protocol design discipline. As described earlier in this paper, minimizing the number of options within the "critical path" of the protocol, fixing such options to a point within the header or trailer and considering the placement of fields can be very important to the bandwidth/latency potentials of the protocol.

An important point is that a number of factors need to be considered when assessing efficiency including bandwidth, latency, amount of local resources required (i.e. memory for control structures and buffers). The only measurements that are meaningful are those made within the environment an application runs in. Thus for considering Technical Workstation efficiency, timings of two Workstations transferring a large file may be meaningless if it is assumed that such workstation's run in an X-Windows environment.

6.6. Time Synchronization Support

The Time Synchronization Support requirement can be treated separately from the other requirements since this is not a need of a basic Transport service but a requirement for hooks within this and lower layers. Time critical systems have a need for maintaining a consistent value of time and supplying it to the applications that run in all of the computing elements. The end result of this requirement may be a separate SD3 project or an assurance that the hooks needed to support such a facility are available through a Transport standard.

The type of support needed is a means to accurately determine or predict the local time that a PDU carrying synchronization information is placed on a network as well as when such a PDU is received from a remote time service.

7. Conclusions

The combination of new functional requirements along with the requirement for efficient implementation provides a strong justification for further Transport standard development. The significant requirements (i.e. Class 1) described within the previous section appear to be the drivers for further Transport standardization efforts. The next step is to analyze these significant requirements against the existing OSI protocols as described in this study effort's SD3. Any further efforts must also consider those requirements classified as important (i.e. Class 2). These further efforts are expected to result in one or more SD3 proposals for modification to existing OSI standards and/or new OSI standards.

There were no requirements identified for the telecommunication or low end system (i.e. personal computers) areas. It is assumed that the present standards are adequate for these areas for the foreseeable future. Telecommunications considerations of speed-of-light timing limitations (i.e. delay due to long distances) and a means for charging for the communication paths used are presently not issues in the applications considered here.

8. References

1. Chesson, G. Notes on Technical Workstation Requirements. X3S3.3/90-65.
2. Marlow, D. T. Requirements for a High Performance Transport Protocol for Use on Naval Platforms - Revision 1. 1989 July. HSP-8. X3S3.3/89-107(Revised).
3. Pleinevaux, P. User Requirements for Communications in Time Critical Applications. 1989 February. HSP-46.
4. Watson, R. W. The Delta-t Transport Protocol: Features and Experience. 1989 October. IEEE 14th Conference on Local Computer Networks. X3S3.3/89-349.