

TECHNICAL REPORT

ISO
TR 12178

First edition
1994-04-01

Industrial automation — Time-critical communications architectures — User requirements

*Automatisation industrielle — Architectures des communications en
temps réel — Prescriptions des utilisateurs*



Reference number
ISO/TR 12178:1994(E)



Documents
Document Sharing Hub

ISO/TR 12178:1994(E)

Contents

	Page
Foreword	iii
Introduction	iv
Section 1: General	1
1.1 Scope	1
1.2 Normative references	1
1.3 Definitions	1
Section 2: User requirements for time-critical communications in discrete parts manufacturing	5
2.1 Introduction	5
2.2 Application models	7
2.3 Integration issues	9
2.4 Message differentiation and PDU attributes	12
2.5 Timing related issues	14
2.6 Resilience	20
2.7 Topology aspects	22
2.8 Management and sovereignty	22
Section 3: User requirements for metrics and benchmarks	24
3.1 Introduction	24
3.2 Standard metrics	25
3.3 Benchmarks for TCCA	29
3.4 Quality of Service Parameters	36

© ISO 1994

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

International Organization for Standardization
Case Postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The main task of technical committees is to prepare International Standards. In exceptional circumstances a technical committee may propose the publication of a Technical Report of one of the following types:

- type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;
- type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;
- type 3, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

ISO/TR 12178, which is a Technical Report of type 3, was prepared by Technical Committee ISO/TC 184, *Industrial automation systems and integration* Sub-Committee SC 5, *Architecture and communications*.

For an explanation of the relationship of this Technical Report to normative aspects of time-critical communications architectures that are dealt with in International Standards, see the Introduction.

Introduction

0.1 Summary

It is becoming apparent that in their present form the OSI standards cannot readily accommodate time-critical applications. The network architectures which have been standardized so far were primarily intended for general traffic and are not always capable of providing adequate performance and resilience for time-critical communications, especially where time-critical and non-time-critical traffic coexists. In particular in many CIM and control installations there appears to be a requirement for an intermediate network, between general enterprise-wide networks and the fieldbus-type networks. This intermediate network should carry both bulk data transfers and time-critical messages, and be able to operate over considerable distances and in hostile environments.

A task force was therefore set up under ISO/TC 184/SC 5/WG 2 to look at the requirements for a time-critical communications architecture, concentrating particularly on the requirements in intermediate networks and to prepare a technical report on that topic. This work could be extended to take into account requirements for time-critical communications in fieldbus.

The focus of this report is to identify the communication needs of tightly coupled control systems. This effort is intended to complement efforts which have focused on factory information networks (e.g. MAP) which are characterized by high throughput, large packet service between information processors (e.g. mini computers) and control processors (e.g. special purpose controllers). The effort is also intended to complement efforts to identify sensor networks (e.g. fieldbus) which concentrate on providing a low complexity interface to simple factory floor sensors, actuators and other devices.

There is no clear consensus on whether the time-critical communications architecture should be connection-oriented or connectionless. Either solution is acceptable providing that the requirements are met; however, some requirements may be easier to meet with one of these solutions than with the other. Since MMS uses a connection-oriented session layer then the overall service may appear connection-oriented even if the network and transport layers are connectionless.

This report summarizes the user requirements for time-critical communications systems that have been identified by the task force and provides a short explanation of each of the requirements. Consideration is also given to the requirement for standard metrics and benchmarks to allow users to assess the suitability of particular architectures for their applications.

There is still no all-embracing definitive quantification of the meaning of time-critical communication; however, communication is time-critical if the application process sending a message requires it to be received (or received and acted upon, or

received and acted upon and confirmed) within a certain time after it submits its 'send' request to the system. The network time constraints are determined by the requirements of the application, and there is no intention of defining absolute time constraints for any particular architecture. Indeed, many issues such as predictability are definable only in statistical terms. In all the following discussions it is assumed that the system under consideration is primarily concerned with dynamic sequencing, i.e. there is variable loading on the network because the applications are event driven. It is also assumed that practical applications will require the coexistence of time-critical and non-time-critical traffic on a single network.

By concentrating primarily on the requirements for event driven communications scenarios, the requirements for state driven scenarios and those where there is a mixture of state driven and event driven communications are also satisfied, since the state driven communications can be considered to be special, more predictable, cases of event driven communications.

A network which provides time-critical communications system functionality may, in certain applications, or in smaller systems, provide functionality of information and sensor networks as well. However, in no way should this possibility compromise its capability to serve as a control network.

The time-critical communications architecture shall provide services which allow for efficient interworking with systems which use MMS (ISO/IEC 9506). The time-critical communications architecture should be applicable in environments where the full 7 layer OSI stack is implemented but in certain applications reduced stack architectures may be more appropriate.

In certain applications requiring time-critical communications it may not be essential for the TCC architecture to meet all the identified requirements. Nevertheless, it is clear that all the requirements are essential for some applications.

0.2 Time windows

The term 'time-critical' is used to represent the presence of a time window, within which one or more specified actions is required to be completed with some defined level of certainty. Failure to complete specified actions within the time window risks failure of the applications requesting the actions, with attendant risk to equipment, plant and possibly human life.

While much agreement exists with respect to the presence of the window, the duration of the window continues to be an application-dependent matter. Time constants associated with batch and continuous manufacturing range from milliseconds to seconds, and sequencing times in discrete manufacturing range from microseconds to hours. It is not possible to define, in absolute terms, a universally acceptable time window for all manufacturing applications. Moreover, actual network performance is substantially perturbed in some implementations by dynamic

ISO/TR 12178:1994(E)

factors such as loading or collisions or token loss, and by static variables such as the number of participating nodes and message frame size. The user has to take all such factors into account when establishing the time window required for a particular implementation. Network performance changes resulting from such factors should not compromise the ability of the network implementation to meet the required time window.

However, it remains essential to provide a uniform means for determining whether a given configuration of networks and equipment will meet the time window requirements of a specific application. The metrics section of this document addresses such issues. Furthermore it is also not possible to define a networking protocol suite that is properly optimized for all conditions encountered in manufacturing applications. In general, there are three overlapping requirements that arise in varying degrees in manufacturing, and these requirements may lead to the definition of conflicting protocols.

a) A network is required that is optimized for robust interworking and guaranteed message delivery under harsh conditions.

b) A network is required that is optimized for cyclical or repetitive polling or data transmission. Possible bases for the repetition cycle include time and externally induced bases such as rotational position of associated mechanical equipment.

c) A network is required that is optimized for dynamic, unpredictable communications sequencing, both on the network and within the attached network time-critical communications entities.

This Technical Report focuses primarily on c).

It is recognized that all time-critical communications systems differ from other networks in their requirement for resilience. The tighter the time constraints are, the more important strategies for resilience become, since networks capable of meeting tight time constraints need to exhibit great resilience if the data flow is not to be perturbed by minor faults.

Industrial automation — Time-critical communications architectures — User requirements

Section 1: General

1.1 Scope

This Technical Report identifies user requirements for systems supporting time-critical communications. Such systems need networks which allow both time-critical and non-time-critical communications. It focuses on the requirements identified in the field of discrete parts manufacturing, but the requirements may also be applicable in other fields, including process control. Other work has concentrated on the requirements of state-driven systems, in which traffic flow patterns and network configuration are static, but this report concentrates primarily on the requirements of event-driven systems, in which traffic flow and configuration change dynamically.

It also discusses user requirements for metrics and benchmarks which can be used to compare and manage the performance of networks to ensure that the time constraints of applications can be met and proposes some simple benchmarks.

1.2 References

ISO 7498:1987, *Information processing systems - Open Systems Interconnection - Basic Reference Model*.

ISO/IEC 9506:1990, *Industrial automation systems - Manufacturing Message Specification (MMS) - Part 1: Service definition, Part 2: Protocol specification*.

ISO/IEC 9545:1989, *Information technology - Open Systems Interconnection - Application Layer Structure*.

1.3 Definitions

For the purposes of this Technical Report, the following definitions apply.

1.3.1 application dependent requirement: A requirement related to the usage of the TCC architecture which is independent of the implementation.

ISO/TR 12178:1994(E)

1.3.2 application entity: Part of the application process that deals with the communications system.

1.3.3 implementation-dependent requirement: A requirement which depends on the specific aspects of the implementation of the time-critical communications architecture being used.

1.3.4 interaction: Any causal relation induced by message transfer.

1.3.5 multipeer group: A group of peer entities which are mutually willing and able to be senders or receivers of multi-peer data transmissions with other members of the group.

1.3.6 multipeer data transmission: The transmission of a PDU to one or more destinations.

1.3.7 performance: The behaviour of the system with respect to time.

1.3.8 QoS attribute: An attribute of a managed object relating to QoS.

1.3.9 QoS category: A general class of QoS system elements relating to the user view of the system.

1.3.10 QoS characteristic: A quantifiable aspect of QoS system elements which is defined independently of the means by which it is represented or controlled.

1.3.11 QoS context: The set of information retained by one or more system elements, used for the purposes of QoS management.

1.3.12 QoS measure: A set of one or more observed values relating to a QoS characteristic.

1.3.13 QoS metrics: A set of measurement methods for quantifying QoS characteristics.

1.3.14 QoS parameter: A variable relating to a QoS characteristic conveyed between system elements as part of a mechanism for the management and provision of QoS.

1.3.15 spatial coherence: A property of duplicated lists of variables. It indicates whether or not all the copies are identical at a given time or within a given time window.

1.3.16 temporal coherence: A property of a list of variables. It indicates whether or not the value of each variable in the list has been produced and transmitted and/or received within a given time window.

1.3.17 time-critical communications: When one or more application process which sends a message requires it to be received (or received and acted upon, or received and acted upon

and confirmed) within a certain bounded time period "window" after its send request to the system. (This can be a time-critical communication transaction or a time-critical data transmission.)

1.3.18 time-critical communications architecture: An architecture, in the sense of the OSI reference model, which supports the identified requirements for time-critical communications. It is a structure which has layers, entities, service access points, protocols, connections etc. in accordance with the concepts introduced in ISO 7498.

1.3.19 time-critical data transmission: The transmission of a PDU within a given time window.

1.3.20 time-critical communications entity: An entity participating in time-critical communications (a time-critical communications entity may or may not map directly to an application entity since a single application entity could have time-critical and non-time-critical components).

1.3.21 time-critical communications group: A group of time-critical communications entities in a time-critical communications system.

1.3.22 time-critical communications system: A system in which there is time-critical communication; it is an implementation of a time-critical communications architecture.

1.3.23 time-critical multipeer data transmission: The transmission of a protocol data unit to more than one destination within a defined time-window or time windows.

1.3.24 time-critical communications transaction: An ordered set of message transfers (or exchanges) coming from different application entities which all have to be completed within a given time window. This may be a time-critical multipeer communications transaction or a time-critical peer to peer communications transaction. Such transactions are complex information processing operations involving multiple data transmissions for an application with time constraints.

1.3.25 time window: A bounded time interval which is characterized by starting time and delay or starting time and end time which are application dependent; the degree of resolution of times and delays is implementation dependent. Specific time windows are defined in 3.2.3.4.

1.4 Abbreviations

The following symbols and abbreviations are used in this Technical Report with the meanings indicated.

AE Application entity
AP Application process
ASE Application service element

ISO/TR 12178:1994(E)

CIM Computer Integrated Manufacturing
LAN Local Area Network
MAC Media Access Control
MAP Manufacturing Automation Protocol
MMS Manufacturing Message Specification
MTBF Mean time between failures
MTTR Mean time to repair
OSI Open Systems Interconnection
PDU Protocol Data Unit
QoS Quality of Service
SAP Service Access Point
TCC Time-critical communications
TCCA Time-critical communications architecture
TCCE Time-critical communications entity
TCCG Time-critical communications group
TCCS Time-critical communications system
TCCT Time-critical communications transaction
TCDT Time-critical data transmission
TW Time window

Section 2: User requirements for time-critical communications in discrete parts manufacturing

2.1 Introduction

Automated factories use networks for distributed control applications. These applications require communications systems which are designed to accommodate worst-case performance as opposed to average performance; in such systems timeliness is of paramount importance. The network should be able to cope with both bulk data transfers and time-critical messages and should support methods for achieving resilience.

If necessary, interfacing the TCC system with the main factory data communication network should be a simple procedure, using a linking device. Because of the requirement to operate in very harsh environments, highly reliable media and signalling methods are necessary so that a very low bit error rate results and a minimum number of retransmissions is necessary. The user needs to be able to define communications priorities and control error recovery mechanisms. Network management functions are necessary to allocate resources, control access to TCC groups, detect latent failures and actual failures etc.

The user requirements that are identified allow the concept of time windows to be realized in OSI. The interrelationship between TCC architecture, temporal coherence, time windows and applications requiring time-critical communications are set out in figure 1.

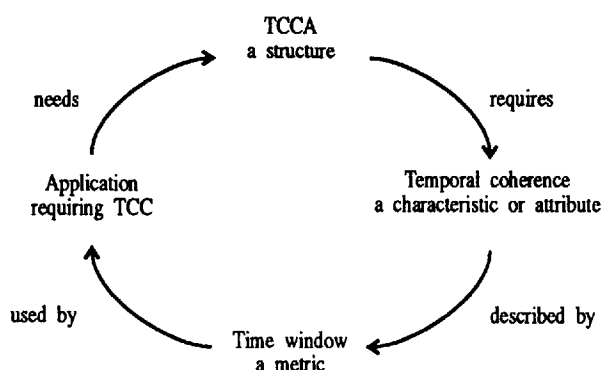


Figure 1

The following user requirements for time-critical communications systems (TCCSs) are grouped into six major categories: integration issues; message differentiation and PDU attributes; timing-related issues; resilience; topology aspects; and management and sovereignty issues. These requirements are discussed separately.

a) Integration issues

- 1) TCCSs supporting MMS functionalities in a timely fashion
- 2) TCCA requiring standardized user interfaces

ISO/TR 12178:1994(E)

- 3) TCCA defined within the context of OSI
 - 4) Time-critical messages short enough to be sent without segmentation
 - 5) Membership of a TCCS controlled
 - 6) Applications requiring time-critical communications not committing time-critical PDUs to the transmitting stack faster than it or the network or the receiving stack can handle the traffic
 - 7) Classes of conformance to individual requirement level
- b) Message differentiation and PDU attributes
- 1) Differentiation of messages
 - 2) Non-time-critical PDUs not prejudicing the delivery of time-critical PDUs
 - 3) Distinguishing the relative urgency of messages
 - 4) A TCCS coping with "dynamic sequencing"
- c) Timing-related issues
- 1) Necessity of Predictability of message transfer times and transaction times
 - 2) Necessity of agreed probability of interaction/transaction completion time
 - 3) Requirement for temporal coherence and possible requirement for spatial coherence in TCCA
 - 4) Reconfigurability of the application and the use of the TCCS
- d) Resilience
- 1) User control of error recovery
 - 2) Mechanisms to assist latent failure detection in TCCSs
 - 3) Availability
- e) Topology aspects
- 1) Off segment communication
- f) Management and sovereignty
- 1) Management issues
 - 2) Sovereignty issues

NOTE 1 The ordering of the above categories and points does not reflect their relative importance.

Some other sets of combinations of the identified user requirements may be made, for example: architectural requirements covering message differentiation, management and resilience issues; timeliness requirements; relationship to OSI standards and architectures. In practice users will probably order the requirements to reflect the importance of each in their particular application.

2.2 Application models

A TCC architecture needs to provide for a number of different communication scenarios. In general the user's overview is perceived from the application; an AE may map directly to the TCC entity or partly to the TCC entity because it also has non-TCC-entity elements in it (see figure 2).

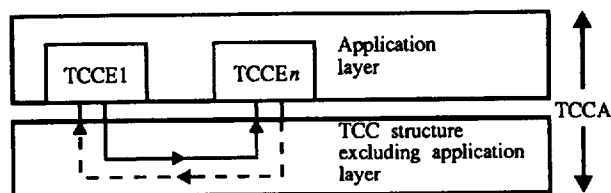


Figure 2 - Relationship of TCCEs and TCCA

A variety of application models have been identified and are referred to as client-server, initiator-sink, producer-consumer, producer-distributor-consumer, etc. The following types of transfers have been identified as essential for TCC architecture but the technical implications for TCC systems are not discussed in detail as these will depend on the particular TCC system chosen and its implementation.

- point to point, i.e. an exchange from one sender to one receiver.
- multicast, i.e. an exchange from one sender to many (n) receivers which form a known group of TCC entities.

NOTE 2 Where an exchange is from one sender to all receivers without negotiation of the communications context it may be referred to as broadcast.

In applications there will be the following types of interactions:

- Peer to peer interaction, where the application which initiates the interaction or requests the service (the initiator) is the sender and there may or may not be a requirement for a response [see figures 3 a) and 3 b)].
- Multipeer interaction, where the initiator may make one multipoint transfer or n point to point transfers and again there may, or may not, be a requirement for responses, which may be a confirmation by the group as a whole or n individual confirmations from each member of the group [see figures 4 a), 4 b), and 4 c)].
- Peer to distributing peer is sometimes recognized as a third category, but this can be equated to figure 3a or 3b plus figures 4 a) or 4 b) or 4 c) depending on what level of confirmation is required (see figure 5).