ATNP/WG -WP/
07 March 1995

EUROCONTROL

AERONAUTICAL TELECOMMUNICATIONS NETWORK PANEL

Working Group 1,2,3

Toulouse, FR 13.3.95-24.3.95

QoS Management Concept for CNS/ATM-1 Package

Prepared by M.G. Adnams

SUMMARY

This document has been produced within the context of the WG1 deliverable WG1-10 QOS/Security Management Concept Package 1. It was felt that QoS and Security are separate topics, the deliverable has therefore been split and the are separate papers for these. The subject of this document is QoS; it covers Institutional, Application and ATN Internet related issues.

DOCUMENT CONTROL LOG

SECTION	DATE	REV. NO.	REASON FOR CHANGE OR REFERENCE TO CHANGE
	7-Mar-95	Issue 1.0	Document Creation

TABLE OF CONTENTS

1. Introduction

1.1 Document Purpose

This document has been produced within the context of the WG1 deliverable WG1-10 QOS/Security Management Concept Package 1.

It was felt that QoS and Security are separate topics, the deliverable has therefore been split and the are separate papers for these. The subject of this document is QoS; it covers Institutional, Application and ATN Internet related issues.

Where possible specific requirements, assumed requirements and constraints are identified.

1.2 Document Structure

Sections 2 and 3	Introduce the OSI QoS Concept/Model (mainly tutorial) but major assumptions and constraints are identified which should be reviewed by WG1/2/3.
Section 4	Documents Institutional requirements for QoS, of particular interest to WG1
Section 5	Documents User (application) requirements and assumptions for QoS, of particular interest to WG3.
Section 6	Documents the ATN Internet provisions in support of requirements. WG2 should check the correctness of the constraints/assumptions. WG3 should be aware of the constraints and assumtions.
Appendix A	Recalls Flimsy #2 of WG2 second meeting on the use of QoS and Priority.
Appendix B	The ADS Manual and QoS - QoS material from the ADS Manual.
Appendix C	QoS and MHS (Message Handling Systems) - Information with resepect to MHS systems.

2. QoS Management

Quality of Service is a way of expressing a network's capabilities as a set of independent performance metrics. The values of such metrics (e.g. transit delay, throughput, cost, residual error rate, etc.) derive from the technology used to implement the network and the network's topology. However, while the QoS that can be provided by a network can be expressed as a static quantity, the actual QoS provided can also vary dynamically with the applied load on the network. This is because as the volume of data handled by the network incises, queuing effects in the network can result in increased (e.g.) Transit Delay, and may affect the values of other QoS metrics. When parts of the network become congested, it is possible for there to be a catastrophic degradation in the QoS provided.

In order to counter the effects of congestion there is a need for QoS Management which combines QoS Maintenance and monitoring functions designed to avoid congestion occurring or to mitigate its effects. QoS Monitoring functions can schedule the maintenance functions when predefined conditions are met.

3. Quality of Service Model

3.1 The OSI Model

The Aeronautical Telecommunication Network is defined in the ATN Manual as an OSI communications system. As with other aspects of the system architecture, the ATN Quality of Service framework will be based on OSI architecture.

OSI specifications are based on a service model, where a particular set of functions is defined in terms of (i) the service provided by that set to other parts of the system (service-users) and (ii) the mechanisms employed within those functions to achieve the service. In the context of OSI, 'quality of service' refers to how well a service is provided, in terms of some set of user requirements. Thus, when the service is data transfer, QoS might be expressed in terms of throughput, transit delay and data integrity; cost and security might be relevant in some circumstances.

A quality of service framework for OSI is under development (ISO/IEC JTC1/SC21 N8871 Quality of Service – Basic Framework, working draft 4, July 1994). When complete, it will provide a uniform basis for the handling of QoS in OSI specifications.

The framework will cover:

- QoS categories, that define the typical needs for QoS in particular environments;
- QoS characteristics: the properties of OSI subsystems that are relevant to QoS and in terms of which QoS requirements may be stated;
- QoS mechanisms, that can be applied as part of QoS management functions to achieve specified QoS requirements;
- a model of QoS entities and information flows in OSI.

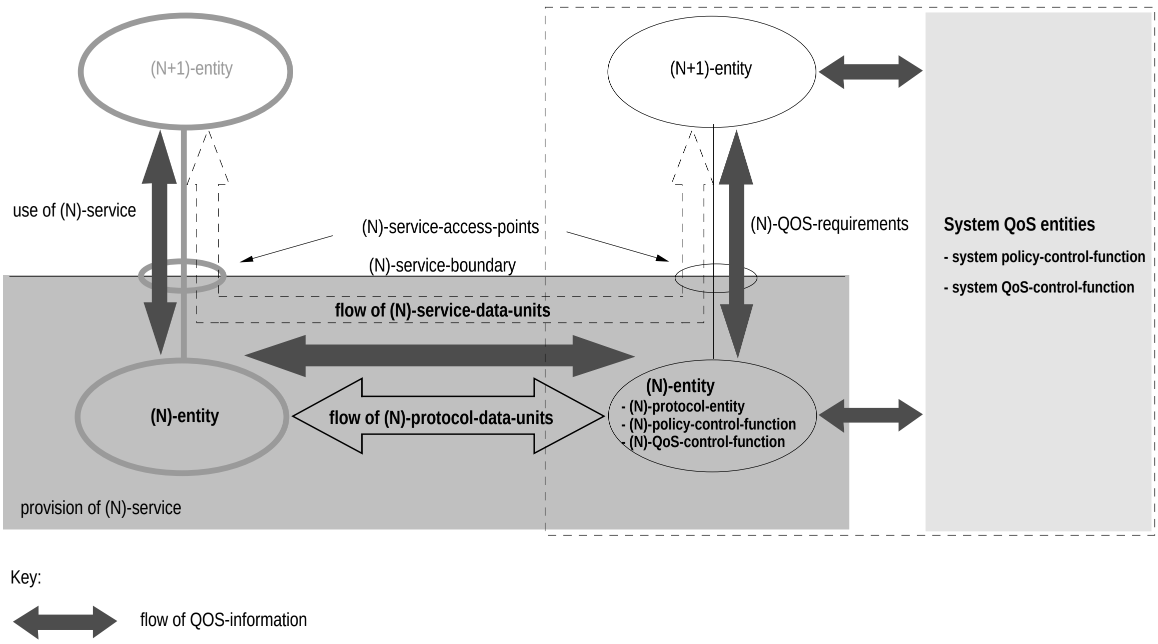


Figure 1 OSI model of QoS entities and information flow

Figure 1 shows the OSI QoS model of entities and information flows as it relates to a single layer service boundary. Within each layer and system-wide, two types of QoS-related entities are identified:

- policy-control-functions are the means by which policies are implemented concerning the acceptability of proposed or returned QoS requirements;
- QoS-control-functions are the means by which QoS requirements are interpreted and applied to a suitable configuration of protocol-entities.

The QoS framework identifies a variety of possible interactions between service-users, service-providers and third parties, such as system management entities. The (N)-layer aspects of these interactions are illustrated in Figure 2. In principle, the composition of such interactions across the OSI layer service boundaries provides the QoS capability of a particular OSI stack implementation.

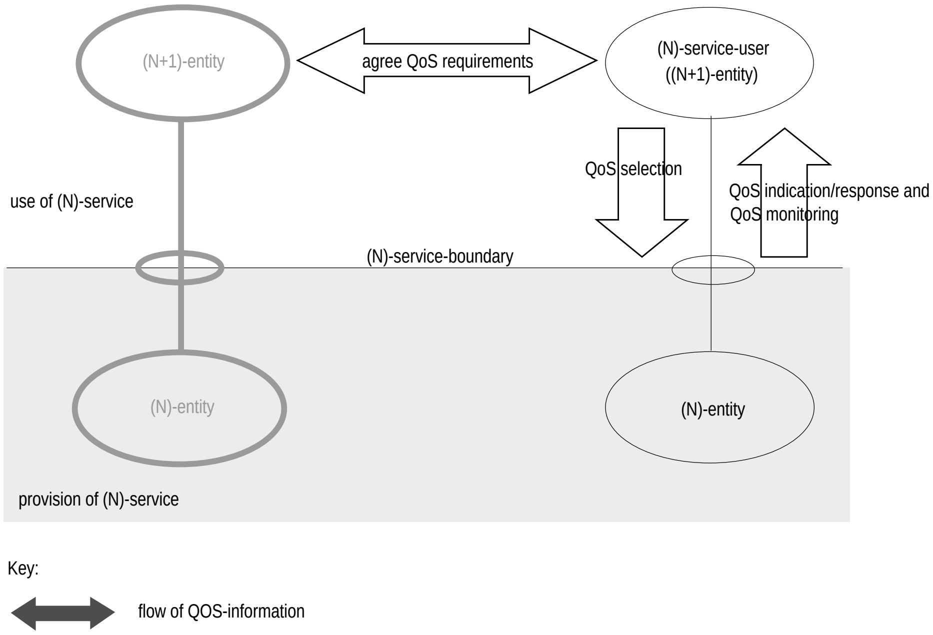


Figure 2 QoS interactions

3.2 Establishment and Maintenance

Target QoS levels may be established *a priori*, by layer management, by systems management, or by negotiation during connection establishment (in the case of connection-mode communications).

Where levels are negotiated, a range of values will be offered by one service-user that may be refined by the service-provider and the peer service-user.

The strength of QoS requirements needs to be considered. There are three types of requirement:

guaranteed QoS — the (N)-service-users and the service-provider agree a level of QoS and require the level to be maintained for the life of the connection, if necessary by degrading the service to users whose QoS is not guaranteed;

Note: Not provided in the ATN, practical only in connection oriented networks.

compulsory QoS — the (N)-service-users and the service-provider agree a level of QoS and require the connection to be cleared if the QoS level cannot be maintained;

Note: Not provided in the ATN, practical only in connection oriented networks.

best efforts QoS — the (N)-service-users and the service-provider agree a level of QoS but the level is not necessarily monitored and the connection will continue even if the QoS fails to meet the agreed target.

Note: This is what the ATN provides, practical in connectionless networks.

Constraint 1 - The ATN is a connectionless internetwork and only provides “best efforts” or “weak” QoS.

3.3 QoS Mapping

The QoS model implies that there is a mapping of QoS between communications layers shown in Figure 3.

In the **application layer**, a recursive application-service-object structure is shown, assuming the use of the Extended Application Layer Structure and the resulting modified upper layer protocols. In this case, there may be complex mappings within the application layer itself (in fact, as noted, these may involve session and presentation as well). Even if the 'traditional' application layer structure is adopted, there may still be several application-associations mapping on to corresponding transport connections.

Note: This reflects the current proposals from the WG3 Architecture subgroup.

In the **transport layer**, a set of T-data primitives from various transport connections may be transformed into a different set of N-unitdata primitives by TSDU segmentation and TPDU concatenation. The quantitative QoS that has been selected for each transport connection has to be implemented in each connectionless transmission.

In the **network layer**, NPDUs may be forwarded from the originating end system via by whatever routes are available, using the appropriate SNDCF based on qualitative QoS preferences. Subsequently, a multiplicity of routes may be taken by different NPDUs, depending on the routing choices of intermediate systems. In addition, there is a natural interaction with any other NPDU processing in each system, even if it does not relate to the same end-points.

At the **subnetwork level**, there is a further level of mapping to the QoS provided in a given subnetwork.

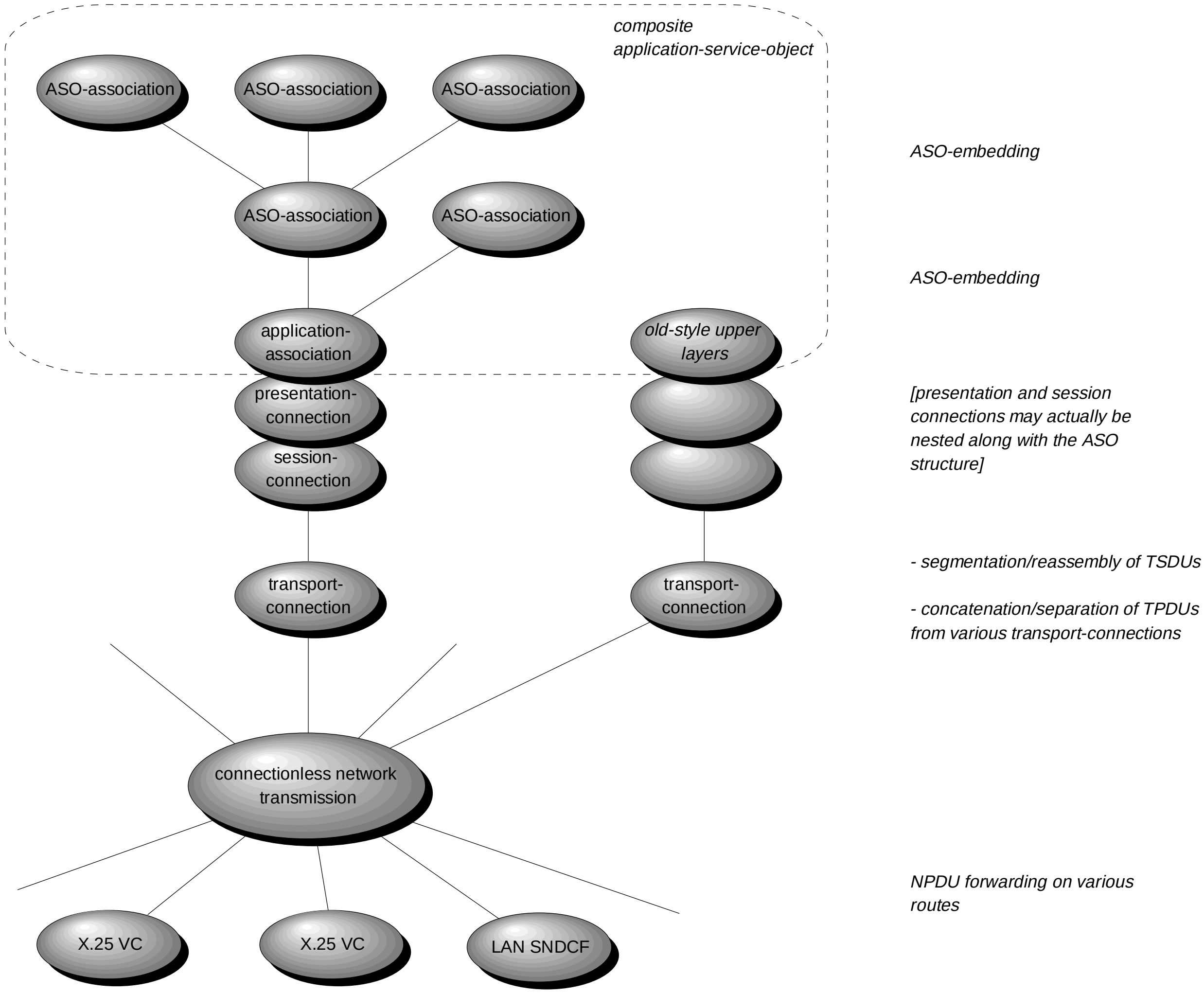


Figure 3 OSI communications stack

3.4 Problems with the Model

As stated previously a quality of service framework for OSI is under development (ISO/IEC JTC1/SC21 N8871 Quality of Service – Basic Framework, working draft 4, July 1994). When complete, it will provide a uniform basis for the handling of QoS in OSI specifications. The need for this is obvious since there are still significant problems with the current OSI specifications used for the ATN resulting from the lack of a coherent architectural approach to QoS handling and mapping in the OSI standards (details discussed below).

Assumption 1 - For coherent definitions of QoS characteristics to be agreed in the ATN layer protocols a fundamental problem of OSI standardisation must be solved.

Constraint 2 - A coherent architectural approach to QoS handling and mapping in the OSI standards will not be achieved within the Package 1 timescale, the solution would in any case result in the current ATN protocol versions being superseded.

The current OSI QoS “model” was designed with connection oriented networks in mind, there are therefore problems in adapting the model to connectionless networks.

3.5 QoS in Connectionless vs Connection Oriented Networks

One of the most useful QoS Maintenance tools is dynamic resource allocation. This is when network resources are allocated to a specific communications path in response to a user request. As long as the network is working below capacity, such resource allocations are little more than book-keeping, however, when all resources in a given network node have been allocated then new requests for resource allocation are rejected, thereby maintaining the QoS for the existing users. This avoids the impact of congestion by rationing the resources, typically on a first come first served basis. Although pre-emptive mechanisms may also be employed, allowing a higher priority user to claim resources previously allocated to a lower priority user, which then either deprives that user of their communications path, or reduces the QoS provided to the lower priority user.

Dynamic resource allocation is also useful for properly accounting for the use of network resources that may be scarcer than others. For example, a high cost network resource (e.g. a high speed data link) may be available that offers a significantly lower transit delay than the normal data links. Users that wish to pay for such the lower transit delay can signal this when the resources are allocated and be allocated this resource, instead of a normal data link.

However, this form of QoS Maintenance is only possible in connection mode networks, where communication paths (i.e. network connections) have a clearly defined establishment phase (when resources are allocated), data transfer phase (when those resources are used), and a connection termination phase (when the resources are returned to the network for re-allocation). On the other hand, dynamic resource pre-allocation is not practicable for connectionless networks.

In a connectionless network, each packet is sent as a discrete event; there is no opportunity for pre-allocation of resources. It therefore follows that with connectionless networks, the applied load may rise to the point that the network (or one or more nodes) becomes congested and no more data can be accepted. The normal response to such a situation is to discard packets in a congested node, in order to relieve the situation. Such packets may either be selected at random, or on a user priority basis, discarding lower priority data in preference to higher priority data. Users that see their data discarded, in consequence see a significant loss in the QoS provided.

Constraint 3 - QoS Maintenance based on dynamic resource pre-allocation is not practicable for connectionless networks.

The relative merits of connection mode and connectionless networks have been debated elsewhere. Neither is perfect when network resources become scarce, as the former simply prevents new users from gaining any service (which therefore have a zero QoS), while the latter, simply degrades its service as new users come along

3.6 QoS Maintenance and Monitoring in the ATN Internet

Connectionless networks can provide a QoS Maintenance function in a qualitative, rather than quantitative manner, it is possible for a network user to indicate on a per packet basis, the (e.g.) acceptable transit delay against cost required. The network may then forward packets on different routes according to such requests. Like priority, this mechanism can be used to differentiate users that are, for example, prepared to pay for a premium service, from those that are not, and hence allocate resources reserved for such a premium service to those users only. Premium resources are then not made available to other users, even when the resources they are using are congested.

Monitoring the QoS currently available in the ATN would require the dynamic collection of statistics from the nodes in the network by means such as distributed systems management. Intervening in routing decisions on this basis of gathered data would require the invention of new mechanisms and may cause route instability.

Constraint 4 - For the ATN the only adequate QoS Maintenance function is capacity planning and network design, and any network congestion is really the result of the failure of capacity planning.

Constraint 5 - When congestion occurs algorithms will be implemented to handle and report the problem, the problem can only be solved if the congestion is temporary.

Constraint 6 - The ATN offers a connection mode transport service over the connectionless network service. Although this would appear to be able to offer the potential for dynamic resource allocation, as the service is connection mode, the scope for such dynamic resource allocation is limited to the End Systems involved, and does not affect the network itself.

Constraint 7 - The use of distributed Systems Management for QoS Management is not defined and in any case not available in CNS/ATM-1 Package.

4. Institutional Requirements

Requirement 1 - Specifications of QoS requirements and capabilities shall be based on agreed operational requirements, such specifications must be stated in terms that allow for implementations to be tested for conformance with them.

Operational requirements relevant to QoS will generally be stated in terms of application performance. To derive QoS requirements for ATN applications from operational requirements, it will be necessary to apportion performance requirements between the application and the ATN communications subsystem. For instance, a response time operational requirement may lead to transit delay requirements on the ATN subsystem as well as processing time requirements on the ATN application.

Requirement 2 - ITU-R regulations on the use of aeronautical radio subnetworks shall be supported.

This requirement is met in the current ATN Manual by the specification of traffic type and the use of the security label.

Requirement 3 - In order to support the development of a safety case for an ATN application means of monitoring the actual QoS achieved during use (both in testing and operationally) are needed.

There is currently no support in the ATN Manual specification for monitoring actual QoS.

Requirement 4 - Commercial service providers and their customers need to be able to define service level agreements and subsequently to meet contractual guarantees, QoS specifications and mechanisms should support these needs.

There is currently no support in the ATN Manual specification for monitoring actual QoS (e.g. ATN QoS management capabilities to monitor the actual QoS provided against agreed service level).

Requirement 5 - Applications shall only be approved for operational use once the appropriate network capacity is in place.

The prime QoS Maintenance tool in the ATN is capacity planning and network design.

5. Application/Upper Layer Requirements

Constraint 8 - It should be realised that the operational concept and the SARPs for CNS/ATM-1 Package applications are at an early stage of development. All requirements, assumptions and constraints identified are therefore reflect the current state of knowledge and can not be considered complete or definitive.

5.1 The Generic Set of Application QoS Parameters

Assumption 2 - Application requirements for service can be expressed by a generic set of parameters; p1, p2, etc. WG3 will define such a generic set (TBD) and use them to express the technical requirements of applications as interpreted from the general requirements of users.

Constraint 9 - Applications will have requirements on the level of service to be provided and these requirements must be met by:

- design of the upper layer protocols, including message design
- transport service characteristics
- network design and implementation

Note: A requirement for Throughput or Transit delay can be met in a connectionless network but can not be controlled in any way by use of the Transport QoS.

5.1.1 The ADS Manual QoS Parameters

The ADS Manual specifies QoS Parameters relating to Absolute Performance:

- Transfer time,
- Response time.

The ADS Manual further states that these can be allocated on a message basis dependant on the operational requirement.

The ADS Manual specifies QoS Parameters relating to Relative Performance:

- Communications priority,
- Urgency.

The ADS Manual further states that these can be allocated on a message basis dependant on the operational requirement. See Appendix B for a summary of the ADS Manual material and definition of the above parameters.

Assumption 3 - WG3 will take into account the ADS Manual specified QoS parameters in identifying the generic set.

5.2 Application and Upper Layer Architecture

The WG3 architecture subgroup has identified for CNS/ATM-1 Package a 7 layer stack comprising:

- the Extended Application Layer Structure (XALS) supporting A^2CSE,

- the "fastbyte" session and presentation layers (effectively NULL layers for air/ground end systems but potentially FULL layers for ground/ground end systems).

Note: None of the above are yet ISO standards and no profiles for them have been proposed (at the time of writing). WG3 in Toulouse will discuss these issues.

Assumption 4 - Applications will specify a generic set of QoS requirements which will be mapped to those available at the ATN Internet where possible.

5.3 CNS/ATM-1 Package Air/Ground Applications

The connectionless network protocol used by the ATN can be used to indicate a qualitative ranking of QoS Metrics in order to select between premium and normal service levels. However, this is unlikely to impact operational air/ground applications as these should be designed within the limits of the QoS available over air/ground data networks, and be capable of operating over any that are available i.e. there is no potential for premium services.

Constraint 10 - air/ground applications shall be designed within the limits of the QoS available over air/ground data networks, and be capable of operating over any that are available i.e. there is no potential for premium services.

Assumption 5 - WG3 will express the service requirements for applications using the generic set of parameters defined for them. QoS requirements such as those from ADS Manual and ATS Data Link Applications Guidance Material will be interpreted to this form.

5.3.1 General QoS Requirements for Air/Ground Applications

5.3.1.1 QoS requirements related to reliability

Requirement 6 - The general performance requirements for applications as identified in the ADS Manual are:

1. The probability that a particular message will be delivered with one or more undetected errors shall be less than 10^{-7}.
2. The probability of non-receipt of a message shall be less than 10^{-6}.
3. The probability that a particular message will be misdirected shall be less than 10^{-7}.
4. The minimum availability of the end-to-end data communication system shall be not less than $1–10^{-6}$ to $1–10^{-4}$ over a given period - typically one month. This means an availability of at least 99.99%
5. Except in catastrophic situations, no single end-to-end outage should exceed 30 seconds. [End-to-end availability may be achieved through provision of alternate communications routings where feasible.]

5.3.2 QoS Requirements for Specific Air/Ground Applications

5.3.2.1 ADS requirements

Requirement 7 - The ADS Manual May 1994 has the following statements on ADS QoS parameters:

- The Transport Priority shall be 3 (It is not totally clear that this is consistent with the DLARD, which implies that 5 would be appropriate)
- The preferences shall be:

  - sequencing over transit delay
  - transit delay over cost
  - cost over residual error rate
- The end-systems priority handling shall be:
  - emergency
  - messages containing projected profile
  - other messages
- The information change...is to be afforded the maximum security...

Assumption 6 - WG3 will process the ADS Manual statements on ADS QoS parameters and express the service requirements using the generic set of parameters defined for them.

5.3.2.2 CPC requirements

Requirement 8 - The ADS Manual of May 1994 has the following QoS statements with respect to CPDLC

- The urgency of messages when placed in a queue shall be
  1. Distress
  2. Urgent
  3. Normal
  4. Low
- All CPC connections shall be set up using a "Safety of flight" priority.
- A security mechanism shall be provided which ensures that only authorised users are allowed access to the CPDLC application

The CPDLC application, in conjunction with the communications network, shall provide a security mechanism which ensures that only authorised CPDLC applications are allowed to access to other authorised CPDLC applications via the communications network.

Assumption 7 - WG3 will process the ADS Manual statements on CPDLC QoS parameters and express the service requirements using the generic set of parameters defined for them.

5.3.2.3 CM requirements

CM is an air-ground application used between an aircraft and its responsible ATC unit (and possibly neighbouring ATC units), associations being initiated by the ground end systems.

No additional performance requirements are specified, beyond the general data link requirements stated earlier.

5.3.2.4 FIS requirements

The flight information services application provides for distribution of information related to air traffic management (e.g., terminal status information, NOTAMs). It is a air-ground application, in which the airborne system is able to receive on request some set of data relating to (e.g.) conditions at a destination aerodrome. The ground end system may be in a routing domain remote from the those adjacent to the airborne domain.

The draft SARPs propose performance requirements for FIS in terms of availability, reliability, integrity, capacity, delay, and data loss.

No additional performance requirements are specified, beyond the general data link requirements stated earlier.

Assumption 8 - For FIS further work is required to analyse the message dialogues in addition to the QoS requirements of the Upper Layers and Transport services.

5.4 CNS/ATM-1 Package Ground/Ground Applications

5.4.1 General QoS Requirements for Ground/Ground Applications

None identified.

5.4.2 QoS Requirements for Specific Ground/Ground Applications

5.4.2.1 AIDC requirements

Requirement 9 - The ADS Manual identifies the following minimum requirements for the following Categories of AIDC Traffic:

Type	Transfer Time (Seconds)	Comm Failure Notification (Seconds)
Flight Planning Messages	60	60
Notification Messages	15	90
Co-ordination Messages	10	60
Transfer of Control Message	5	15
Surveillance Messages	5	15
Airspace Management	15	60

Assumption 9 - For AIDC, it should be noted that the Notification, Co-ordination, Transfer of Control and surveillance messages are exchanged between two specific controllers as part of the process for the acceptance and transfer of control procedure. It is therefore likely that communication will be carried out over a single transport connection.

Requirement 10 - The ADS Manual makes the following demands on the underlying communications service:

- All ATSU connections shall be established using an inter-network protocol priority value of 10 and, representative of flight safety communications.
- The transport priority value shall be 4.

For information handling:

- Emergency messages are critical communications"
- Surveillance messages are urgent communications
- All other messages are routine communications.

For security:

- These authentication exchanges, concerned with the control of aircraft in flight, are to be afforded the maximum security for authentication, access control and data integrity.

Assumption 10 - For AIDC work is required to analyse the message dialogues in addition to the QoS requirements of the Upper Layers and Transport services. The requirements for ensuring reliability of communication is achieved by a combination of a reliable communications service and the correct design of the applications protocols supporting the user application.

Assumption 11 - Definition of the application level messages and protocol in conjunction with the operational requirements will drive the requirements for Throughput, Transit Delay, Priority and Residual Error Rate. This in turn will impact the network design and use of the parameters negotiable on a per-connection basis.

Assumption 12 - WG3 will process the ADS Manual statements on AIDC QoS parameters and express the service requirements using the generic set of parameters defined for them.

5.4.2.2 ATS requirements (MHS) within AFTN Environment Type B

The ATS messaging service application is defined in [Draft Manual on Messaging Applications, Appendix A to the Report on Agenda Item 5, ATNP/1-WP/66, June 1994]. As described in this paper, ATS messaging is solely a ground-ground application used between ATC units. It is to be implemented using OSI MHS. In addition to the full MHS ATS stack, an AFTN migration stack is specified, to transfer AFTN messages directly in TSDUs between AFTN/ATN gateways.

Requirement 11 - In section 3.2.3, the Draft Manual on Messaging Applications specifies the use of five message priorities; these are to be mapped on to three transmission priorities (urgent, normal, non-urgent) that should be separately handled by MTAs and gateways.

It is understood that as a minimum MHS should provide the same level of service as the AFTN for which there are also no formally stated performance requirements.

Assumption 13 - For ATS further work is required to analyse the message dialogues in addition to the QoS requirements of the Upper Layers and Transport services.

Note: See Appendix C for further discussion on MHS and QoS.

5.5 Other System Application Requirements

Assumption 14 - There will be no standard systems and security management applications, nor use of the Directory using the ATN as a communications medium in CSN/ATM-1 Package.

6. Internet Provisions

The transport mechanism for CNS/ATM-1 Package is the Connection Oriented Transport Service Class 4 (TP4). The transport service user (TS_USER) is taken to be synonymous with the application using the service.

This section defines the QoS parameters available at each layer, their semantics in each layer and discusses the mapping of parameters between layers.

QoS, Priority and the Security Label are discussed separately for the network layer since they are treated quite differently.

Constraint 11 - It is not possible for an application to select by QoS or any other defined means the particular subnetwork to be used for data transfer.

6.1 The Transport Layer

The service definition of the Transport Service ISO 8072 foresees the following QoS parameters:

1. Throughput* (could be used for peer to peer info.exchange)
2. Transit delay* (could be used for transport timer setting, but this is not advisable)
3. Residual error rate* (maybe used to select checksum off/on)
4. Protection* (no use foreseen, use in ATN Manual was a defect)
5. Priority* (see section separate section below).
6. Transfer failure probability (no practical dynamic use)
7. Connection establishment failure probability (no practical dynamic use)
8. Connection release failure probability (no practical dynamic use)
9. Connection establishment delay (no practical dynamic use)
10. Connection release delay (no practical dynamic use)
11. Connection resilience (no practical dynamic use)

Note: The Security label is not a member of the above set of standard parameters

These QoS were defined with the connection oriented network layer (CONS) in mind where network and subnetwork resources could be requested and allocated on a per connection basis (e.g. Throughput could be mapped to X25 Throughput Class Negotiation Facility and Transit Delay to X25 Transit Delay Selection and Indication facility via N_CONNECT).

With a connectionless network layer many of these QoS parameters become less meaningful, unenforceable or useless (see below) to convey information to network layer. Since the route a packet takes through the network is not known a priori, it will not be possible or even desirable to configure network resources or subnetwork en route.

Assumption 15 - In the transport layer QoS parameters are set by the user at the service interface in the T-CONNECT primitive or may be set by local management for a particular connection.

Constraint 12 - QoS, Priority and the Security Label are set once per transport connection NOT per TSDU or TPDU and must be consistently applied for the lifetime of the connection in all NPDUs associated with it.

Note: It is has been suggested that certain users (applications) would like to specify priority per TSDU to in order to send distress messages ahead of normal data using the same *transport connection (setting up a separate connection being an unacceptable overhead). There is no defined mechanism that would allow this. Also this will not have the desired effect since, according to standard, TSDUs must be delivered in the order sent. If a pre-emptive message delivery service is required the connectionless transport service would provide the best means. Alternatively, in TP4 a user may send Expedited Data (max. 16 octects) messages. Expedited data interrupts normal TSDU peer to peer delivery on the connection until expedited delivery is acknowledged.*

Constraint 13 - Parameters 1-5 are negotiable at connection time, all others (6-11 above) are not, they are based purely on a priori information about the network. In the ATN TS_USERS with strict QoS requirements are themselves responsible for determining whether non-negotiable QoS parameter values can be satisfied, they can only do this by placing requirements on capacity planners and network designers.

Constraint 14 - TS_USERS may negotiate parameters 1-5 but the agreed values have end to end or transport layer internal significance only, they can not be passed to the network layer (using the standard interface) although they may be used by the transport layer (local issue) to rank network layer QoS in terms of preference.

Constraint 15 - The TS-USER may be informed of the non-conformance to QoS requirements of a peer transport entity during connection negotiation but not afterwards during the life of the connection. Catastrophic non-conformance will result in disconnection.

Constraint 16 - Negotiated values come from the peer transport entity NOT the network or subnetwork level. The proposed values in the negotiation are the result of a priori information “known” by the peer entities rather than being dynamically discovered by them.

Constraint 17 - Transport priority is only of significance between the end-users of the transport service; thus, there is no requirement in Package 1 that transport priority be used within the transport protocol layer for internal processing purposes or for internal resource allocation (connection and buffer management) purposes although this is not precluded

Constraint 18 - Only Priority is recommended of the 5 negotiable parameters in the ATN Manual, all others are optional.

Constraint 19 - Transport layer QoS (e.g. Transit delay) should not be used as a basis upon which to set Transport layer timers. These should be set by local management and maintained by the implementation on the basis of knowledge about the underlying network. Furthermore direct setting may interfere with flow and congestion control procedures which rely on timer settings.

Constraint 20 - The Transport QoS parameter, residual error rate, may be used to set checksums on/off on the basis that if the there is a highly reliable underlying service checksums will not be needed and will add to overhead. The standards Transport service interface offers no other mechanism for this, local management must otherwise be used to control the setting.

Constraint 21 - Transport QoS connection negotiation exchanges at connection time will obviously increase the data transfer overhead associated with setting up a connection and may not give real benefits. Negotiation should be limited to increase efficiency in general and particularly in the case of mobile subnetworks.

Constraint 22 - The Transport Service interface parameter set could be extended to support ATN specific parameters (e.g. the Security Label, checksum control etc.) but this will preclude the use of unmodified COTS products and is not recommended.

Note: When applications require independent behaviour of the underlying communication service, this can be provided by transport implementations allowing QoS, checksums and timers to be set on a per connection basis

6.2 ATN Network/Subnetwork Layer QoS

Note: Priority and the Security Label are discussed in separate sections below.

6.2.1 End Systems

QoS parameters in the connectionless network layer are:

1. Transit delay,
2. Residual error probability (no defined use in the ATN),
3. Protection (no defined use in the ATN),
4. Cost determinants.

Note: The planned use of cost determinants is not specified in the ATN Manual.

Constraint 23 - In a connectionless network there can be no peer to peer negotiation of QoS.

Constraint 24 - In CLNP QoS are not specified by VALUE from the NS_USER. They serve only as a RANKED SET which is configurable to represent an order of preference

Constraint 25 - The mandatory CLNP QoS maintenance function operates on the QUALITATIVE ranking (expressed in NPDUs by 2 bits in the globally unique format) of these parameters, a user can not specify a required QUANTITATIVE value for these parameters. Hence there can be no direct relation between these and the Transport QoS parameters of similar name.

Expected values of these parameters may be given to NS Users by NS Providers but there is no standard mechanism for this, management primitives may be used but there is NO SERVICE PROVISION in ISO 8348 nor any means defined to derive such values.

6.2.2 Intermediate Systems

The mandatory QoS maintenance function operates on the QUALITATIVE ranking (expressed in NPDUs by 2 bits in the globally unique format) of the QoS maintenance parameter.

When dynamic QoS based Routing in IDRP is implemented there is 1 QoS metric per route. IDRP selects the best route for each metric from its known list and makes an entry in the FIB offering the best fit for the metric. CLNP will then forward NPDUs appropriately according to the QoS bit settings (see ATN Manual A9.5.4.4.3) in the NPDU.

Assumption 16 - Values for IDRP QoS metrics (only Transit Delay and Expense are recommended for use in the ATN) in routing policy are set using a priori information about datalinks derived from Network Design by local management. IDRP QoS metric values will be set by local management.

Note: The planned use of Expense QoS metric is not specified in the ATN Manual.

6.2.3 Subnetwork Service

For ISO 8208 subnetworks service co-ordination between the SNDCF and the Subnetwork Provider comprises the following six service attributes:

1. Transit Delay;

2. Residual Error Probability;

3. Economic cost;
4. Protection against Unauthorised Access;
5. Throughput Class;
6. *Priority (discussed separately below).*

Constraint 26 - Subnetwork QoS are specified by quantitative value for use in the establishment of virtual circuits. It is not therefore possible to set them directly based on the qualitative CLNP parameters of similar name. The parameters may be set by local management from defaults based on a priori information for each subnetwork type or derived from CLNP parameters by a locally defined mapping for the subnetwork type.

The use of the Subnetwork Quality of Service is optional in the ATN (priority is treated separately below).

Note: Refer to Chapter and Appendix 10 of the ATN Manual for further guidance and standards.

6.2.4 The Setting of QoS Parameters

Assumption 17 - In the network layer the RANKED SET of QoS are set by local management using a priori information or by the transport layer in the N_UNITDATA request primitive using a locally defined mapping scheme (e.g. ATN Manual chapter 8.1.2.5.4) for a particular connection.

Requirement 12 - QoS is set in every NPDU and the CLNP QoS Maintenance function is mandatory for ATN implementations.

Requirement 13 - Every NPDU associated with a given transport connection must have the same QoS set.

Note 1: The ranked set must be set consistently in every NPDU associated with a given transport connection.

Note2: Local management in the above context includes the capacity of an application to set parameters directly by a locally defined mechanism.

Assumption 18 - In the subnetwork QoS parameters are set by local management using a priori information about the subnetworks available or by the use of the SN_UNITDATA request primitive using a locally defined mapping scheme (non defined at present).

6.3 ATN Network Layer Priority

Assumption 19 - There is no mandatory fixed relationship between transport and network layer priorities however local management or mechanisms may map TC priority to N-UNITDATA priority according to a locally defined scheme (e.g. ATN Manual Ch.8).

Requirement 14 - Priority is set in every NPDU and the CLNP priority function is mandatory for ATN implementations.

The priority function allows a PDU to be processed preferentially with respect to other PDUs. The resource of end and intermediate system Network entities, such as queues and buffers ban be used preferentially to process higher-priority PDUs ahead of lower-priority PDUs.

Requirement 15 - Every NPDU associated with a given transport connection must have the same priority.

Constraint 27 - In the network layer priority is used for queue and buffer management and is semantically unrelated to transport layer priority.

Note: The ATN Manual does not specify how the priority function should be implemented but mandates it. Priority based queuing and discard is not specified. This is a defect in the ATN Manual.

6.3.1 End Systems

There is nothing specific to end systems.

6.3.2 Intermediate Systems

Note: Somewhat confusingly the IDRP distinguishing path attribute priority is not specified for use in the ATN. The reason is historical, it had previously been thought to restrict access to routes based on this attribute, this functionality is now provided in the ATN by the security attribute.

6.3.3 Subnetwork Service

The following is taken from the ATN Manual A10.6.5.2.

The SNDCF shall have a priori information regarding the use of priority by the Subnetwork Provider.

If the Subnetwork Provider supports priority and specifies the mapping of Network Service to Subnetwork Service priorities:

1. The SNDCF shall convey the priority to the Subnetwork Provider by means of the ISO 8208 priority facility. If the subnetwork supports priority values ranging from **[0000 0000]** to **[0000 1110]** (0-14 sic), the subnetwork priority value shall be the same as the priority value in the ISO 8473 or ISO 9542 header, respectively to provide a one-to-one mapping between the network and subnetwork layer priorities. If fewer (or more) than fourteen levels of subnetwork priority are supported, the priority value in the ISO 8473 or ISO 9542 header shall be mapped onto the subnetwork priority value according to the requirements of the subnetwork provider.

2. The SNDCF shall establish a SN-connection for each requested priority. A SN-connection shall only be used for SNSDUs for which the indicated priority corresponds with the priority associated with the connection.

If the Subnetwork Provider does not require connections to be prioritised, then the SNDCF shall not convey priority information to the Subnetwork Provider.

Requirement 16 - The mapping between SNSDU priority (taken from the ISO 8473 NPDU header value 0-14) and subnetwork priority is specified separately for each subnetwork by a local manager based on a priori knowledge about the subnetwork as specified in ATN Manual A10.6.5.2

6.3.4 Mobile Subnetworks

Constraint 28 - The different mobile subnetworks have inconsistent specifications for and interpretation of priority. The reason for this is historical but gives rise to subnetwork specific behaviour based on priority. A user may set priority in NPDUs for a given transport connection but the meaning of this is conditioned by the subnetwork being used.

The mapping of ATN Internet priority to subnetwork priority may be one to one or many to one and even then interpreted by the subnetworks differently.

WG2/3 are invited to consider this problem and to give guidance as to the use and control of priority (sorry - input required).

The table below summarises the expected mapping of priority.

Category of Messages	Corresponding Communication Protocol Priority				
	Subnetwork				
	SN UNITDATA	CIDIN	Mode S	AMSS	VDL
Network/Systems Management	14	TBD	HIGH	14	N/A
Distress Communications	13	TBD	HIGH	14	N/A
Urgent Communications	12	TBD	HIGH	14	N/A
Communications relating to Direction Finding	11	TBD	HIGH	11	N/A
Flight Safety Messages	10	TBD	HIGH	11	N/A
Meteorological Communications	9	TBD	LOW	8	N/A
Flight Regularity Communications	8	TBD	LOW	7	N/A
Aeronautical Information Service Messages	7	TBD	*	6	N/A
Network/Systems Administration	6	TBD	*	5	N/A
Aeronautical Administrative Messages	5	TBD	*	5	N/A
<unassigned >	4	TBD	*	5	N/A
Urgent Priority Administrative and U.N. Charter Communications	3	TBD	*	3	N/A
High Priority Administrative and State/Government Communications	2	TBD	*	2	N/A
Normal Priority Administrative	1	TBD	*	1	N/A
Low Priority Administrative	0	TBD	*	0	N/A

Mapping of Communication Priorities in the ATN

Note 1. - The term "message(s)" in this table conforms to existing Annex 10 terminology.

Note 2. - "" indicates due to legal frequency allocation reasons no traffic of this priority category can pass over this subnetwork. Note also that it is security, not priority protocol mechanisms that can be used to restrict traffic in this way.*

Note 3. - TBD indicates these priority values are To Be Determined.

6.3.4.1 VHF

VHF is very simple, only specify one level of priority is specified.

6.3.4.2 Mode S

Mode S subnetworks offer 2 levels of priority LOW and HIGH.

The Mode S subnetwork processes 2 levels implementing a queue management scheme.

6.3.4.3 AMSS

AMSS offers 15 levels of “priority”(sic). The different levels are processed to control access to the subnetwork as well as to prioritise queuing.

6.3.5 The Setting of Priority

Assumption 20- In the network layer the priority may be set by local management using a priori information or by the transport layer in the N_UNITDATA request primitive using a locally defined mapping scheme for a particular connection.

Note: The mapping of transport to network layer priority was previously an ATN Manual requirement. This has now been recognised as a defect in WG2 although it may be the best solution for applications sharing a common transport entity and hence multiple connections.

Assumption 21 - In the subnetwork, priority is set by mapping SNSDU priority to the priority supported (if any) by a subnetwork service. The map is defined using a priori information about the subnetworks and set locally in End Systems and Intermediate Systems

6.4 ATN Network Layer Security Label Processing

It is understood that there is a clear user requirement to limit access to ITU restricted subnetworks in the ATN based on traffic type. The mechanism provided to control this access in the ATN is the Security Label and applies only to mobile subnetworks.

Note: This requirement was expressed in the Institutional Requirements section.

Constraint 29 - No known COTS software will support the specification or use of the security label.

Assumption 22 - The Security Label will be selected for CNS/ATM-1 Package implementation.

Note: An alternative to the Security Label is under consideration by WG2. The security label could be replaced by adopting addressing conventions for the same purpose. This section only looks at the security label.

Assumption 23 - The CLNP Security function is mandatory in all ATN End Systems and Intermediate Systems, the function must implement the ATN Security policy for access control based on traffic type.

Requirement 17 - Every NPDU associated with a given transport connection must have the same Security Label.

6.4.1 End Systems

Requirement 18 - A mechanism to set the security label consistently in each NPDU supporting a transport connection must be provided to a user communicating any traffic type other than General Communications. There is no standard mechanism at the network service interface.

6.4.2 Intermediate Systems

Requirement 19 - The security path attribute is mandatory in IDRP and the values that it supports must be configured by local management based on a priori knowledge of the supported subnetworks.

6.4.3 The Setting of the Security Label

A user wishing to communicate any traffic type other than General Communications must be provided with a mechanism to set the security label consistently in each NPDU supporting a transport connection.

Note: This means encoding the security label in the NPDU “Options part” as the Globally Unique Security option (see ATN Manual A9.5.4.).

There is no standard mechanism for achieving this but the possibilities are to:

- Extend the standard Transport service and Network Layer service parameter list to communicate the security label via the ATN stack to CLNP so that it may be encoded in NPDUs,

- Provide a local means for a user (application) to set the value of the security label directly by local management.

Note: The second possibility would work well in implementations where a single application uses only one transport connection. When this is not the case it will be difficult to relate the security label to be set for an NPDU to a particular transport connection.

7. Recommendation

This deliverable WG1-10 QOS/Security Management Concept Package 1 was produced in the context of the WG1 for delivery to WG2 and WG3 at their meetings in Toulouse on the 13-17 March. It has been proposed that there be a small side meeting combining members from both groups to discuss QOS related issues.

It is recommended that this combined group review this document and produce Flimsies from their respective WG2/3 perspectives.

Application requirements for service can be expressed by a generic set of parameters; p1, p2, etc. WG3 should define such a generic set and use them to express the technical requirements of applications as interpreted from the general requirements of users.

WG2 should check the Internet Provisions chapter for consistency with the CNS/ATM-1 Package Internet definition.

WG3 should review the Internet Provisions chapter to understand the constraints of the Internet provisions.

It is recommended that WG1 review the document together with the flimsies (WG1 meet after WG2 and WG3) and on the basis of the review recommend further action for Package 1 (if any) and the longer term deliverable, WG1-11 Overall QOS/Security Management Concept. In particular the scope of the latter should be defined.

APPENDIX A

Flimsy #2: Use of Transport Network and Subnetwork Priority/QoS in the ATN Internet

1. Background Information

1.1 Problem Statement

During the ATNP Working Group 2 discussion of the current draft ATN Internet SARPs and Guidance Material regarding transport, network and subnetwork priority/QOS provisions, it was observed that a number of areas of mis-specification or over-specification may exist:

1. in the area of semantic relationships and associated mapping among transport, network and subnetwork priorities; and,
2. in the area of semantic relationships and associated mapping among transport, network and subnetwork Quality of Service (QOS).

Further, certain ambiguities exist within the draft SARPs text regarding the proper invocation of these functionalities by the Transport Service User (TS-User).

1.2 Scope of this Flimsy

This flimsy presents the sense of the ATNP Working Group 2 discussion on these matters, and presents resulting recommendations for preparation of defect reports and coordination statements for other bodies.

1.3 Background on Priority Discussion

Regarding priority, the main observation of Working Group 2 (WG2) was that the ATN Manual is almost certainly defective in requiring not just a fixed relationship (independent of applications) between transport and network priority (Table A5-1 of the ATN SARPs and Guidance Material (V0.0)), but in also requiring that this mapping is explicitly implemented in the transport layer. It was observed that the specification of network and transport priority as separate parameters at the TS-User service boundary, in principle, should be allowed, since this approach would meet the requirements underlying Table A5-1, supporting transfer of TS-User PDUs in an appropriate manner through the ATN Internet. It was also observed that this approach would then fail to meet certain implementation requirements contained in the ATN SARPs and Guidance Material (V0.0) which, by themselves, have nothing to do with interoperability and appear to constitute over-specification.

This in turn led to a discussion on the semantics and proper use of transport and network priority. It was agreed that the purpose of network priority was to identify the relative priority of data for transmission through the network and to determine access to limited resources. On the other hand, the purpose of transport priority in the ATN was to signal the relative priority of the transport connection between peer applications.

Given such conclusions, there is no need to mandate the strict relationship between transport and network priority given in Table A5-1. It was agreed that while Table A5-1 is the preferred default, applications should be able to specify different relationships if desired.

1.4 Background on Quality of Service (QOS) Discussion

The discussion on Quality of Service (QOS) followed a similar path, and similar conclusions were reached. This problem is slightly different in its detail, since the semantics of QOS are dissimilar between the Transport and Network layers in the ATN, and thus no mapping algorithms are mandated within the ATN SARPs and Guidance Material (V0.0). However, the underlying principles to be applied are the same.

2. Recommended Use of Priority in the ATN Internet

2.1 Principles

The following principles were agreed with respect to Transport, Network and Subnetwork layer use of priority.

a) Transport priority is only of significance between the end-users of the transport service; thus, there is no requirement that transport priority be used within the transport protocol layer for internal processing purposes or for internal resource allocation purposes.

b) There is no mandatory fixed relationship between transport and network layer priorities. In addition, transport and network layer priorities are of different and possibly unrelated semantics.

c) The semantics of network layer priority must be specified and invoked in a manner to support correct mapping to the Subnetwork Service (SN-Service) priority. This mapping relationship is specified in the ATN SARPs and Guidance Material (V0.0), and thus no change is proposed at this time.

d) The syntax and semantics of the mapping of ATN SN-Service priority to the priority parameters provided by any given ATN subnetwork are specified by the SARPs appropriate to that subnetwork, and are not specified within the ATN SARPs and Guidance Material (V0.0).

2.2 Recommended Actions

The principles agreed above lead to the following recommended actions:

a) Based on the principles in 2.1, appropriate defect reports must be prepared to modify the ATN SARPs and Guidance Material (V0.0) to reflect the decoupling of Transport and Network layer priorities from each other, as is now specified.

b) Based on the principles in 2.1, ATN TS-Users (i.e. ATN applications and upper-layer protocol stacks) may need to be designed in a manner to support the separate conveyance of Network and Transport layer priority information to the ATN Network and Transport layer protocols, respectively.

3. Recommended Use of QOS in the ATN Internet

3.1 Principles

The following principles were agreed with respect to Transport, Network and Subnetwork layer use of QOS.

a) Transport QOS is only of significance between the end-users of the transport service; thus, there is no requirement that transport QOS be used within the transport protocol layer for internal processing purposes or for internal resource allocation purposes.

b) There is no mandatory fixed relationship between transport and network layer QOS. In addition, transport and network layer QOS contains different and possibly unrelated semantics.

c) The semantics of network layer QOS must be invoked in a manner to support correct mapping to the Subnetwork Service (SN-Service) QOS. This mapping relationship is specified in the ATN SARPs and Guidance Material (V0.0), and thus no change is proposed at this time.

d) The syntax and semantics of the mapping of ATN SN-Service QOS to the QOS parameters provided by any given ATN subnetwork are specified by the SARPs appropriate to that subnetwork, and are not specified within the ATN SARPs and Guidance Material (V0.0).

3.2 Recommended Actions

The principles agreed above lead to the following recommended actions:

a) Based on the principles in 3.1, appropriate defect reports must be prepared to modify the ATN SARPs and Guidance Material (V0.0) to reflect the decoupling of Transport and Network layer QOS from each other, as is now specified.

b) Based on the principles in 3.1, ATN TS-Users (i.e. ATN applications and upper-layer protocol stacks) may need to be designed in a manner to support the separate conveyance of Network and Transport layer QOS information to the ATN Network and Transport layer protocols, respectively.

4. Coordination Issues

The information contained in this flimsy should be communicated to ATNP WG1 and WG3 for comment and feedback during the upcoming joint meeting of the ATNP Working Groups in March 1995. Following this, a coordination statement should be prepared by the joint Working Group meeting to communicate relevant information to the ICAO ADS Panel, and to the other panels developing ATN subnetwork SARPs (i.e. SICAS Panel and AMC Panel).

APPENDIX B

The ADS Manual and QoS

1. QoS Parameters relating to absolute performance

1.1 Transfer time

Transfer time is the end-to-end time expected from the moment the triggering event is issued from the originator use process and the moment the message is received, validated and ready for further treatment at the destination user process. This duration includes the technical data extraction and composition of the data message together with the data transmission and processing of the message.

1.2 Response time

Response time is the end-to-end time expected from the moment the triggering event is issued from the originator user process and the moment the logical response is received by the originator user process. This duration includes :

- the time for technical data extraction and composition of the message;
- the data transmission and processing time
- the logical checks at the destination
- the transmission and receipt of a response

2. QoS Parameters relating to relative performance

2.1 Communication priority

Communication priority is envisaged to establish the priority of a connection between two end systems. In theory it would be useful to prioritise individual end-to-end message in an information flow, in relation to that of other messages in the same or in other end-to-end information flows. Unfortunately, the connection-oriented transport service does not give any freedom to vary the priority on a 'per message' basis, so this QoS requirement can only used to manage local (sending and response) queues.

2.2 Urgency

Urgency delineates the relative relationship among messages when placed in a queue for operator access. It relates to the handling of the information by the receiving system. It dictates order of display, processing, or other another action in accordance with the sequencing, route and time-expired data. Urgency does not influence communication processing, which is defined by communications priority; it applies to the end user processing application only. Valid urgency entries may include:

1. distress (indicating grave and imminent danger),
2. urgent (comprising movement and control messages),
3. normal (comprising routine operational messages, such as surveillance or navigation messages)
4. low (indicating any message with lesser urgency than the above).

Note: It is noted that the DLARD proforma only allows three levels of urgency.

3. QoS requirements related to reliability

The general performance requirements for applications as identified in the ADS Manual are:

1. The probability that a particular message will be delivered with one or more undetected errors shall be less than 10^{-7}.
2. The probability of non-receipt of a message shall be less than 10^{-6}.
3. The probability that a particular message will be misdirected shall be less than 10^{-7}.
4. The minimum availability of the end-to-end data communication system shall be not less than $1–10^{-6}$ to $1–10^{-4}$ over a given period - typically one month. This means an availability of at least 99.99%

Except in catastrophic situations, no single end-to-end outage should exceed 30 seconds. [End-to-end availability may be achieved through provision of alternate communications routings where feasible.]

4. DLARD Message Categories

Transfer time and response time can be allocated on a message basis dependant on the operational requirement. The following categories are distinguished in the report of the ATNP Meeting in the description of the Data Link Applications Requirements Document (DLARD). (It is recognised that the titles and descriptions below differ slightly from the entries in the DLARD Pro Forma):

1. Network/Systems Management (The reasoning is that such messages are imperative to the satisfactory operation of the network)
2. Distress, indicating grave and imminent danger;
3. Urgent, concerning the safety of the aircraft or persons on-board or within sight;
4. Communications relating to Direction Finding;
5. Flight safety, comprising movement and control messages, and meteorological or other advice of immediate concern to units involved in the operational control of an aircraft in flight or about to depart;
6. Meteorological messages, comprising forecasts, observations and other messages exchanged between meteorological offices;
7. Flight regularity communications;
8. Aeronautical Information Services Messages, comprising aircraft operator and other messages of concern to the aircraft in flight or about to depart;
9. Network/Systems Administration, including statistics reports;
10. Aeronautical Administrative Messages (NOTAM-Class I distribution?;
11. Urgent Priority Administrative and UN Charter Communications
12. High Priority Administrative and State/Government Communications
13. Normal Priority Administrative

14. Low, indicating any message with a lower priority than the above.

Each categorisation can be quantified based on maximum transfer or response time values, in 95% of the transfers, and on time within which communication failure is provided in 99,996 % of the transfers. For example, a given message category could be specified to have transfer/response times of messages within 16 seconds (for 95% of messages), and communication failure notification is provided within 30 seconds (in 99.996% of the messages).

APPENDIX C

QoS and MHS (Message Handling Systems)

There are currently no formal operational requirements defined by ICAO that demand the use of MHS. As a consequence there are no formal performance requirements. It is understood that as a minimum MHS should provide the same level of service as the AFTN for which there are also no formally stated performance requirements.

In order to address this situation an alternative set of requirements should be identified that can provide representative requirements. The Eurocontrol (CFMU) study [CFMU III Inter-Systems Communications - Final Report March 92, section 1.6.2] provides an example of performance requirements for an application that makes use of messaging technologies. This report identifies the performance requirements for the access mechanisms as described in the following text.

Performance: The highest performance requirements for communications over networks are summarised as follows:

- Interactive session: 10 seconds for 500 bytes Query plus a 200 Bytes Reply
- Message Transfer: 3 minutes for a 3 kilobyte message
- File Transfer: 4 hours for 100 Megabytes

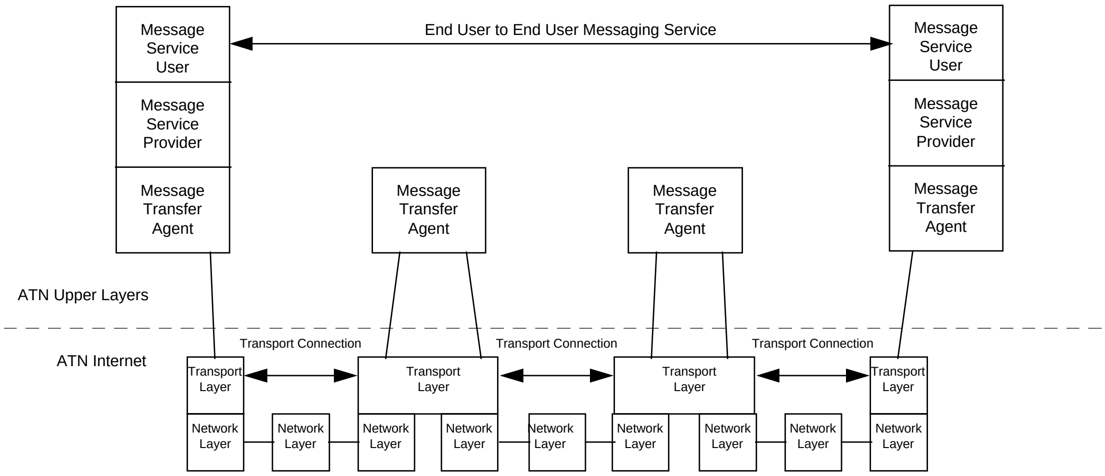
Traffic Load (hence performance) is also expressed in the following terms by the CFMU Study:

- Transit Time
- Number of Messages/ Unit Time
- Volume Transferred/ Unit Time (average figure)
- Bandwidth Required (KB/s) (peak figure)

These application level performance requirements do not necessarily translate directly or indirectly to parameters that are controllable at the transport service interface.

The diagram below shows the structure of a service based on message handling service. It is important to note that although there are elements of the structure that perform the role of relays as in the lower layer architecture for the ATN Internet that these relay elements described here as message transfer agents (MTA) do not operate in the same mode as an ATN router. The message transfer agent in a messaging services is required to receive all of the message before it will relay that message to next MTA. This differs from the router in the ATN Internet which will relay to the next router immediately each packet as it receives it.

Therefore in the diagram we see that two MTAs may communicate by establishing a transport connection between themselves. These MTAs view this transport connection as a single link, yet this may make use of a number of network layer routers to provide the end to end transport connection.



MHS has two significant characteristics that must be considered in the evaluation of performance of an MHS based message transport service.

MHS can provide a store and forward messaging service that can transfer a message from one end user to another via a series of MTAs. The relay between MTAs occurs sequentially thus a message must be completely received before it is relayed onwards.

MHS provides a reliable transfer service. This provides a mechanism for recovery so this if there is a network failure during the transfer of the message that the message transmission can be continued from the previous checkpoint and is not required to restart the transmission from the beginning. This makes MHS a mechanism by which a large amount of data e.g. a file can be transferred effectively.

Operational requirements that have been analysed in Europe show some applications that require the MHS to be able to transfer a large amount of data in a short time (100Mbytes in 4 hours). In order to achieve this sort of performance, it is usually necessary to implement the MHS using a single hop between the two high capacity end systems. The use of an intermediate store and forward relay would place heavy demands on the storage and communications bandwidth of the intermediate system, which is undesirable and may impact on the capability of the intermediate system to meet performance requirements for other users.