N-DR Ames

15P.

IN-16842

Workload Characterization for the Space Station Data Communications System

Kenneth C. Sevcik

January 1986

Research Institute for Advanced Computer Science
NASA Ames Research Center

RIACS TR 86.4

(NASA-TM-89396) WORKLOAD CHARACTERIZATION FOR THE SPACE STATION DATA COMMUNICATIONS SYSTEM (NASA) 15 p CSCL 17B N86-29116

Unclas

G3/32 43339

RIACS

Research Institute for Advanced Computer Science

Workload Characterization For The Space Station Data Communications System

Kenneth C. Sevcik

Research Institute for Advanced Computer Science
NASA Ames Research Center
Moffett Field, California 94035

ABSTRACT

NASA plans to launch a permanent manned space station in the early 1990's. The station will be used to support a wide variety of activities involving earth and space observation, satellite maintenance, scientific experimentation, and commercial manufacturing. The control and monitoring of many of these activities will require extensive computer and communications system support.

In order to identify an appropriate computer and communication system for supporting the space station, an attempt to characterize the space station's data communications subsystem workload is currently underway. In this paper, we discuss some of the special aspects of the workload characterization problem in connection with the space station, and we present some possible approaches.

1. INTRODUCTION

The data communications system for the permanent manned space station that will be launched by NASA in 1992 is currently being designed. Choices of network structure, topology and protocols must be made by 1987 in order to allow sufficient time for implementation, experimentation in a testbed environment, and integration with the design of the rest of the space station. The workload that will be placed on the data communications system is an important factor in making these choices, so an attempt at workload characterization for the system is being made.

The space station project is unique in many ways, and these aspects seem to make workload characterization more difficult:

(1) Because this is the first permanent manned space station to be launched, there is no available knowledge of how space stations are "typically" used.

(2) The user community will be quite diverse, including commercial applications (materials processing, weather observations), scientific applications (crystal growth, space plasma physics), potential defense applications (some of which might be classified), along with control functions (navigation, environment maintenance).

(3) The elements of the system (space station, ground stations, shuttles, orbiting platforms, and satellites) have dynamic spatial relationships to one another, and higher quality communications services are required when elements are physically close to one another (such as when a shuttle docks at the space station).

The performance and reliability of alternative proposed configurations is being investigated using analytic and simulation models. These models can be helpful in making good system design choices only if they take into account the anticipated workload.

In this paper, we discuss two major issues. First, we describe some aspects of the problem of identifying what the components of the space station data system workload are likely to be, and classifying these components according to types of behavior. Second, we suggest a parameterized user profile by which, using various parameter settings, we can represent each of the types of anticipated usage. Goals in developing the user profile include (1) keeping the number of parameters small, and (2) allowing representation at varying levels of detail by providing reasonable default values for as many of the parameters as possible.

By extending or adapting the analytic and simulation models to accept the parameterized user profiles as definitions of the system workload, it will become possible to conveniently investigate the impact on performance of a variety of assumptions about the eventual composition of the data communications system workload.

2. SYSTEM ELEMENTS

In this section, we indicate how to view the space station system in such a way that a workload model can be formulated.

2.1. Users

Users of the computation and communication facilities on-board the space station will include personnel both on the ground and in space. A fundamental distinction between types of use is between *internal* users and *external* users [1]. Internal users include the critical functions of life-support environment maintenance, and guidance and navigation of the space station itself. Other uses that are also considered as internal but are less time-critical include mission planning and scheduling, crew training (through computer-aided instruction and simulation), and crew entertainment (games, electronic mail, and personal word processing).

The primary external uses can be categorized as commercial or scientific (with a possibility of some military applications as well). The commercial applications include crystal growth and materials manufacturing, each of which require a weightless environment. Also, observations of earth, ocean, and atmosphere will constitute commercial applications due to their utility in such applications as weather prediction. The long list of anticipated scientific applications includes astrophysics and planetary observation, space plasma and solar physics, and life sciences, among others [1,2].

2.2. Activities

Having some feeling for who the anticipated users of the space station are, it is possible to begin to identify various *activities* that the users will carry out and that will require the computation and communication facilities onboard the space station. Two major activities that relate to both commercial and scientific uses of the space station respectively are *process control* and *experimental control*. Automated process control will be required to manage crystal growth and other manufacturing operations. Similarly, many scientific experiments will require real-time monitoring and control. In both cases, *sensors* will be used to determine the status of the process or experiment, while *affectors* will be used to redirect or change the status [1].

Another class of activities is known as *proximity operations*. These include dockings with spacecraft (including the shuttle, the orbiting maneuvering vehicle and the orbital transfer vehicle). Proximity operations also include deployments and retrievals of tethers, and the extra-vehicular activity of crew members in external manned propulsion units.

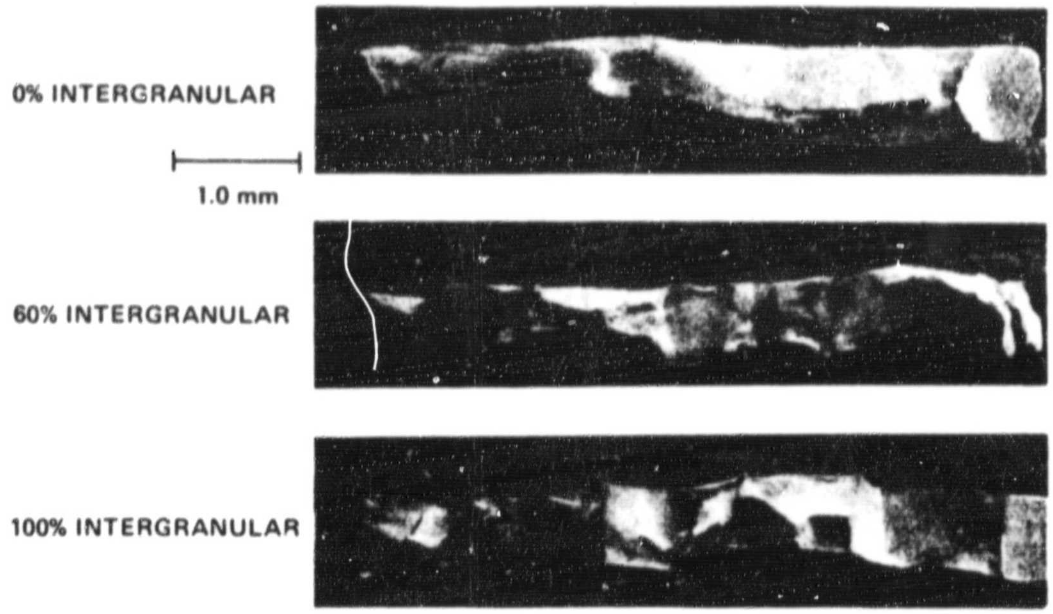

Figure 1 - SEM fractographs of typical fracture surfaces.

for some sets of specimens so that the maximum amount of intergranular fracture was somewhat less than 100%. By doing this, the change in fracture mode as a function of aging time can be used to investigate the kinetics of hydrogen segregation to grain boundaries.

Because the specimens were highly supersaturated with hydrogen at all times subsequent to charging, some, although relatively negligible, loss of hydrogen caused by outgassing occurred (except where noted). Because of this, the condition which maximizes embrittlement is not true equilibrium and will be referred to as "quasi-equilibrium." The aging time at the onset of maximum embrittlement (quasi-equilibrium condition) is denoted the "critical aging time."

The results of aging samples containing 95 ppm and 45 ppm hydrogen at 208 K are given in Fig. 2. The fracture mode of specimens containing 95 ppm hydrogen changes from zero to 100% intergranular fracture after aging for a period of approximately 10^6 seconds. Samples containing 45 ppm hydrogen reach a maximum percentage of intergranular fracture equal to approximately 25%. Similar results for specimens containing 275 ppm, 440 ppm, and 660 ppm of hydrogen which are aged at 253 K are shown in Fig. 3.

Figure 4 shows the effects of aging time at 318 K for specimens containing 765 ppm of hydrogen. As the aging time increases, the degree of intergranular fracture increases from about 30% at 10 seconds to 70% between 200 and 10^4 seconds. Upon further aging, the amount of intergranular fracture decreases because of the loss of hydrogen from outgassing. (This conclusion is based on comparison of the calculated RMS diffusion distance of hydrogen at 318 K during long aging periods with the width of the specimen.)

Interactions with co-orbiting platforms or occasional encounters with polar orbiting platforms or free-flyers are other activities that may involve the use of the orbital maneuvering vehicle.

Activities that correspond to the internal users of the space station include the critical functions of navigation and guidance of the station, and management of the communications link down to earth.

2.3. Types of Network Nodes

The space station data communication system will consist of a large number (roughly 300) nodes, all interconnected by a network. The nodes will have varying degrees of capability for the storage and processing of data. In order to deal with the large number of network nodes in characterizing the workload, it is desirable to identify classes of nodes that have similar functions. Below, we describe some such classes:

(1) Experiment nodes
interface to a user's experiment; may have varying degrees of internal processing power, but internal configuration is the responsibility of the user, so only its interaction with the SSDS is relevant to the workload characterization.

(2) Process control nodes
interface to a commercial production process, again with varying degrees of local capability.

(3) Crew workstations
used for many functions, including the monitoring and control of experiments and commercial processes, space station control and mission planning, crew training and education, etc.

(4) Data processing nodes
processing capacity for data analysis, reduction and compression.

(5) Data storage nodes
for storing the onboard data base (probably in distributed fashion) and buffering data for transmission to earth.

(6) Downlink management node
responsibility for scheduling and management of the TDRS satellite downlink, which is likely to be a critical resource due to its limited transmission capacity.

(7) **Life support nodes**
responsible for sensing the status of all aspects of the life support system and for initiating any required changes.

(8) **Space station control nodes**
certain nodes with specialized facilities for support of control functions.

2.4. Workload Characteristics

Each component of the space station workload may impact overall system performance in a different way depending on certain major characteristics. Thus, the workload characterization will have to associate with each workload component its character with respect to such attributes as:

- volume
  the amount of resource usage for computation, for database storage and retrieval operations, and for communication among system components over the data network
- intensity
  the density of resource usage when the component is active
- periodicity
  the manner in which the component cycles between activity and inactivity
- criticality
  the priority or importance of the component relative to other workload components
- constraints
  any constraints on the execution of the component, such as real-time deadlines

3. Problems and Approaches

In this section, we discuss several aspects of the workload characterization problem for the space station. While some of the problems are unique to the space station project, others are related to workload characterization problems in more general contexts. Thus, the approaches we suggest may also have wider relevance.

3.1. Uncertainty

Whenever an attempt is made to characterize the workload of a system that does not yet exist, there is a degree (probably large) of uncertainty of how the system will eventually be used. In the space station context, this problem is at least as severe as in any other environment. At the time of this work, the space station is still at least seven years away from being operational. Worse yet, because there has been no prior instance of a permanent manned space station, there are no existing systems that might be observed to form a starting point for predicting the eventual usage of the space station.

The closest things to precursors of the space station are probably the sky lab satellite, and the space shuttle, through which the spacelab experiments have been controlled. However, the control of experiments onboard the space station is expected to be significantly more interactive than was the control of earlier experiments in space. Thus, their data communications requirements could be quite different. The concept of *telescience* has been developed in the Space Station Users Group, which is composed of representatives from various scientific disciplines that may eventually benefit from use of the space station's facilities. Telescience is the act of carrying out experimental scientific research while in electronic rather than physical contact with the experimental equipment. That is, all observations and manipulations of the experiment are carried out remotely using television for viewing and robot manipulators for handling, where necessary.

Our approach to dealing with the uncertainty is to consider a broad range of scenarios based on a very high-level model with only a few parameters. The model distinguishes among several types of traffic, with the parameters reflecting the intensities of the various types. By varying the parameters a wide range of possible workload compositions can be examined.

3.2. Diversity

The anticipated user communities of the space station include many scientific disciplines and several commercial interests. The various disciplines and interests have not previously had to share research and production facilities, but in the space station, this will be necessary. The specific needs for computation and communication facilities are different for each of the groups, and the balance of the activity among the groups is not known currently. Furthermore, it is unlikely that the balance will be known any time before the space station becomes operational, and it probably will change continually during the lifetime of the space station.

The high-level model mentioned in the previous section and described in more detail in section 4 facilitates investigation of various balances of activity among the user groups. The objective in selecting a specific design for the SSDS data communications subsystem is to find one that performs well across a range of possible situations.

3.3. Time-Scale

Many of the activities that will cause high levels of data processing and data communication operations are specified as part of the internal activities (for example, shuttle dockings, and other proximity operations). However, these are specified on the time scales of days, weeks or even months. For example, proximity operations are specified as shown in Table 1 [2,3].

Proximity Operations	Frequency	Duration
Extra-vehicular Activity	1/day	6 hours
Shuttle Docking	4/year	24 hours
Orbital Maneuvering Vehicle	3/month	24 hours
Orbital Transfer Vehicle	1/month	8 hours
Tether Deployments	10/month	1 hour

Table 1. Frequencies and Durations of Various Proximity Operations.

Similarly, experiments and commercial process control activities are likely to have alternating periods of activity and quiescence.

Thus, to the extent that the schedule of activities aboard the space station is known, it is known with time granularities of days or more, while the operations in the data network occur at the seconds or milliseconds scale. Consequently, performance evaluation of the data communications network must be done on a time scale that is several orders of magnitude shorter than the time scale on which activities originate and cease.

In some cases, activities, experiments and processes can be scheduled so that not too many are coincident. In other cases, however, external events (e.g., sunspot activity) can trigger activity of a number of experiments simultaneously, resulting in a peak of activity on the network. (Unfortunately, most stochastic models used for performance evaluation do not reflect well the occurrence of simultaneous events such as those that could be caused by an external event.)

In evaluating a candidate network design, it is necessary to consider all potential activities, experiments and processes, to determine how they are initiated, and to identify what combinations are likely to be simultaneously active. Unfortunately, the number of possible combinations is very large.

3.4. Dynamic Configuration

The space station data network includes nodes onboard the space station plus nodes on other system components such as platforms, tethers, shuttles, maneuvering vehicles and transfer vehicles [3]. The positional relationships among these components are continually changing, and the pattern of data communications also changes accordingly. In particular, during a proximity operation such as a shuttle docking, the communications between the shuttle and the space station become much more intense and critical. Thus, the workload to which the data network is subjected is dependent on the relative locations of system components. Once again, this situation necessitates a case by case analysis of system performance, treating in turn each of many possible spatial configurations of the system components.

3.5. Mutual Dependence Problem

Neither the workload nor the space station is currently specified in detail. This leaves uncertainty in two directions. The system designers don't know the workload that their system will be required to support, and, similarly, the users do not know what facilities will be available. The users therefore don't know how ambitious they should be in identifying tasks and experiments that they would expect to carry out onboard the space station. Further, application software design should depend on the relative availability of various resources. A specific example of this type of problem arises in connection with the amount of local computation power built into each node.

The design of the space station data network interacts strongly with the decision about how much processing power to build into each component. For example, a typical experiment will generate a very large amount of data. Either all the data can be transmitted to earth for processing, or some pre-processing and/or data compression can be carried out by a processor onboard the space station in order to reduce the volume of data transmission over the downlink to earth.

This situation leads to the consideration of various assignments of processing power to space station system components. In order to evaluate the alternatives, it is necessary that the characterization of the workload components include the tradeoff between the amount of pre-processing or data compression carried out and the reduction in the amount of data that would be transmitted to earth. Only with this additional information is it possible to assess all potentially desirable configurations.

3.6. Evolving Specification of System and Workload

The basic design of the space station data network will have to be frozen in 1987 although the station will not be launched until 1992, at the earliest. Consequently, the knowledge of all aspects of the workload ... what components will be, what the balance will be among them, and what the resource usage characteristics of each are ... will evolve and generally increase. This situation motivates use of a hierarchical model, capable of representing information at varying levels of detail. At present, with only a very general knowledge of the workload, a hierarchical model would require only a few parameters to be specified, and further details would be based on default assumptions. Later, as knowledge of the workload becomes more detailed and refined, additional parameters can be set explicitly with confidence in order to increase the accuracy of the model.

4. Proposed Model for Workload Characterization

We now outline a strategy for formulating a hierarchical model that satisfies the requirements encountered in the earlier sections. (This strategy is an extension of an earlier proposal [6].) Some of the requirements that we will keep in mind are:

(1) There are a large number of potential user communities with differing characteristics and requirements.

(2) The system components will, at various times, be in many different spatial relationships with one another.

(3) There are many potential variations on the placement of computing power within the space station.

4.1. Model Entities

The fundamental entities and concepts in the model include *users*, *activities* (and their *variations*), and *situations*. In the paragraphs that follow, we indicate the basic parameters that describe each one.

4.1.1. Users

The various identifiable user groups would each be a separate "user" in the workload model. Initially, there might be as few as three users (commercial, scientific, and internal). Later, there would be at least twenty or so users, with the various scientific disciplines and various commercial enterprises being distinguished. Eventually, it might be desirable to distinguish even among individuals in a single discipline by associating each with a distinct user profile. The primary parameters indicating the overall behavior of each user would be the frequency and duration of their periods of usage (e.g., four experiments a year, each lasting three weeks on average).

4.1.2. Activities

A user would be associated with a set of activities, each one corresponding to one way in which the user exercises the facilities of the space station. Activities would be described by several attributes:

- how frequently they start (while the user is in a period of usage)
- how long they last
- how they are initiated (e.g., periodically, at random intervals, scheduled for periods of low activity, or triggered by external events)
- how they consume resources (rate of sending messages, average message length, computation required per message, etc.)
- how their level of resource usage dynamically varies (e.g., alternating between intense resource consumption and relatively low resource consumption)

Variations of activities are necessary to reflect such things as computation versus data transmission tradeoffs. For example, a particular activity that collects data might require very little computation in space if all the data is transmitted to earth in its raw form. On the other hand, if adequate computational power is available onboard the space station, then that power might be used to do data compression and/or data reduction, thus decreasing the volume of data transmitted to earth. These are two variations of the activity. In any single application of the workload model, only one variation of each activity would be "enabled".

4.1.3. Situations

Situations are used to distinguish such things as different spatial relationships among the system components (e.g., the shuttle being docked at the space station), and different environmental contexts (e.g., recent sunspot activity). Some activities are predicated on certain situations. For example, there might be two distinct activities representing the exchange of navigational control information between a shuttle and the space station. One, with heavy intensity, would be conditioned on situations in which a shuttle is currently docking at the space station, while the other, with much lower intensity, would be conditioned on situations in which no shuttle is in close proximity to the space station.

Thus, in using the workload model, by specifying a particular situation, the analyst would be able to filter out all activities of all users that are not appropriate to the situation under consideration.

4.2. Hierarchical Specification

With many users, many activities (with many variations), and many situations, there are a large number of parameter values required to specify the model, even at the simplest level. Many of these parameters are means of distributions of service times, interarrival times, message sizes, etc.

In the early stages of model and system development, information about these distributions beyond their means is not available. Consequently, simple defaults of exponential and geometric distributions can be adopted (since these distributions are completely specified by their means, and they have mathematical properties that facilitate analysis).

Later, when more information is available and more accuracy is desired, additional information about the distributional forms can be supplied. If both the mean and the variance are provided, then these can be used to specify a particular distribution among the families of hypoexponential and hyperexponential distributions. This family of distributions retains some of the same advantages of mathematical tractability possessed by the exponential and the geometric distributions.

In certain situations, still other distributional forms might be appropriate. For example, in many networks, a vast majority of messages are either of minimal length (the length of an acknowledgement, perhaps) or of maximal length (resulting from splitting a file into as few chunks as possible for transmission). In this case, a two-valued distribution with part of the mass at one point and the remainder at another is an appropriate representation.

Similarly, there are several ways of representing the degree of concurrency within an activity. The simplest case is with a single process corresponding to each activity. Slightly more complex situations can be specified by a rate of process initiations, or by an average number of processes in existence.

When messages are transmitted at some layer of the network, their availability can be indicated in several ways. Most simply, just the presence of each message can be signaled. If messages must be partitioned into packets for transmission, then the distribution of the number of packets per message should also be specified. Finally, if the processing overhead of packetization is thought to be significant, then the packets composing a single message can be thought of as becoming available for transmission at times separated by some short fixed interval.

4.3. Aspects of the Model

In this section, we briefly consider how the features of the workload characterization model proposed in this section contribute to alleviating the problems presented in section 3.

Diversity of the users and their activities is handled by directly reflecting users and activities as entities in the model. Distinctions among users and activities can be made to any desired extent by having the diligence to specify more and more "user" and "activity" entities in the model.

The knowledge of the workload will evolve over time, and there will be a great deal of uncertainty initially. To deal with this, the model is designed to be flexible and extendible. It requires few parameters initially, supplying appropriate defaults for anything not explicitly specified. On the other hand, as more and more knowledge of the workload becomes available, that knowledge can be incorporated into the model, by exploiting its hierarchical character.

The processor power allocation aspect of the mutual dependence problem is treated by specifying variations on activities, where one variation assumes that a significant amount of computation can be done in space, while another assumes that the data must be transmitted to earth in its raw form.

The fact that activities start and stop on a time scale several orders of magnitude slower than the rate of operations in the data communications system means that performance analysis must be carried out on each of a very large number of combinations of activities. Similarly, the dynamic spatial relationships among the system components also necessitate a combinatorial analysis of many possibilities. The model uses the concept of "situations" to distinguish these possibilities and to associate with each one the appropriate set of activities.

5. Conclusion

We have outlined a number of problems that make workload characterization for the permanent manned space station (presently under design) difficult. We have suggested approaches for handling each one, and proposed a hierarchical model for describing user profiles in a workload characterization.

Some of the problems encountered in the space station context are similar to workload characterization problems encountered in other contexts. Thus, some of the approaches that we suggest may also be applicable in other contexts.

6. Acknowledgements

My understanding of the space station project as a whole and also the development of the ideas in this paper were greatly aided by discussions with Terry Grant, Marjory Johnson, Don Dubois, and Silvano Columbano at the NASA Ames Research Center.

7. References

[1] Space Station Information System Final Study Report, JPL D-1737, Jet Propulsion Laboratory (August 1984).

[2] Space Station Data System Study, MDAC 3rd Quarterly Review (February 1985).

[3] SSDS Third Quarterly Review, TRW Systems (February 1985).

[4] Customer Data and Operations System Presentation to SSIS Workshop, NASA Goddard Space Flight Center (May 1985).

[5] Space Station Data Systems Study, presented at Johnson Space Center Workshop, MDAC (May 1985).

[6] K.C. Sevcik and M.J. Johnson, A Proposed Structure for Modelling Workload Classes, project note, Advanced Distributed System Networks Group, NASA Ames Research Center (May 1985).

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A163 415	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Hydrogen Embrittlement of Nickel: Modeling of the Effect of Diffusive Segregation of Hydrogen on Intergranular Fracture		5. TYPE OF REPORT & PERIOD COVERED Technical Report January 1986
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) D.H. Lassila and H.K. Birnbaum		8. CONTRACT OR GRANT NUMBER(s) N 00014-83-0468
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Illinois, Department of Metallurgy and Mining Eng.; Urbana; IL 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE January 1986
		13. NUMBER OF PAGES 13
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document is unclassified. Reproduction and distribution for any purpose of the U.S. government is permitted.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Hydrogen embrittlement
Nickel
Hydrogen

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Grain boundary fracture of hydrogen-charged nickel was studied under test conditions where hydrogen is essentially immobile. Prior to testing, hydrogen was allowed to diffuse during aging treatments. Experimental results show that the transition in the fracture mode from ductile rupture to intergranular is strongly dependent on aging temperature and time as well as on initial bulk hydrogen concentration. Analytical modeling of these dependencies using established thermodynamic and kinetic relationships, indicates that grain boundary

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) hydrogen concentration controls the fracture mode of hydrogen-charged nickel at low temperatures.