IDA DOCUMENT D-1244

FY92 CONTRIBUTIONS TO THE DATABASE MANAGEMENT SYSTEM INTERFACE STANDARDS WORKING GROUP OF THE NAVY NEXT GENERATION COMPUTER RESOURCES PROGRAM

Karen D. Gordon
Terry Mayfield, *Task Leader*

October 1992

Prepared for
Space and Naval Warfare Systems Command
Ada Joint Program Office

INSTITUTE FOR DEFENSE ANALYSES
1801 N. Beauregard Street, Alexandria, Virginia 22311-1772

IDA Log No. HQ 92-042853

DEFINITIONS

IDA publishes the following documents to report the results of its work.

Reports

Reports are the most authoritative and most carefully considered products IDA publishes. They normally embody results of major projects which (a) have a direct bearing on decisions affecting major programs, (b) address issues of significant concern to the Executive Branch, the Congress and/or the public, or (c) address issues that have significant economic implications. IDA Reports are reviewed by outside panels of experts to ensure their high quality and relevance to the problems studied, and they are released by the President of IDA.

Group Reports

Group Reports record the findings and results of IDA established working groups and panels composed of senior individuals addressing major issues which otherwise would be the subject of an IDA Report. IDA Group Reports are reviewed by the senior individuals responsible for the project and others as selected by IDA to ensure their high quality and relevance to the problems studied, and are released by the President of IDA.

Papers

Papers, also authoritative and carefully considered products of IDA, address studies that are narrower in scope than those covered in Reports. IDA Papers are reviewed to ensure that they meet the high standards expected of refereed papers in professional journals or formal Agency reports.

Documents

IDA Documents are used for the convenience of the sponsors or the analysts (a) to record substantive work done in quick reaction studies, (b) to record the proceedings of conferences and meetings, (c) to make available preliminary and tentative results of analyses, (d) to record data developed in the course of an investigation, or (e) to forward information that is essentially unanalyzed and unevaluated. The review of IDA Documents is suited to their content and intended use.

The work reported in this document was conducted under contract MDA 903 89 C 0003 for the Department of Defense. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1992	Final

4. TITLE AND SUBTITLE
FY92 Contributions to the Database Management System Interface Standards Working Group of the Navy Next Generation Computer Resources Program

5. FUNDING NUMBERS
MDA 903 89 C 0003

Task T-D5-1090

6. AUTHOR(S)
Karen D. Gordon

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses (IDA)
1801 N. Beauregard St.
Alexandria, VA 22311-1772

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA Document D-1244

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Space and Naval Warfare
Systems Command
SPAWAR 231-2B2
Washington, D.C. 20363-5100

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, unlimited distribution: December 31, 1992.

12b. DISTRIBUTION CODE
2A

13. ABSTRACT (Maximum 200 words)

The Next Generation Computer Resources (NGCR) Program is a Navy standardization effort designed to fulfill the Navy's needs for standard computer resources while at the same time allowing it to take advantage of commercial products and investments and to field new technology more quickly and effectively. The program is centered around the selection and adoption of open system standards in several areas, including backplanes, local area networks, operating systems, project support environments, graphics, and database management systems. IDA is providing support through participation in a recently formed group known as the NGCR Database Management System (DBMS) Interface Standards Working Group (DISWG). The objective of the NGCR DISWG is to identify and help define nonproprietary industry-based database management system interface standards for use in the development and maintenance of future Navy mission-critical computing systems. The document is a compilation of IDA written contributions to the NGCR DISWG from April 1992 through September 1992.

14. SUBJECT TERMS	15. NUMBER OF PAGES
NGCR; Database Management System (DBMS); Interface Standards; Open System; Navy Mission Critical Computer Resource (MCCR).	34
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

IDA DOCUMENT D-1244

FY92 CONTRIBUTIONS TO THE DATABASE MANAGEMENT SYSTEM INTERFACE STANDARDS WORKING GROUP OF THE NAVY NEXT GENERATION COMPUTER RESOURCES PROGRAM

Karen D. Gordon
Terry Mayfield, *Task Leader*

October 1992

INSTITUTE FOR DEFENSE ANALYSES
Contract MDA 903 89 C 0003
Task T-D5-725

PREFACE

The Space and Naval Warfare Systems Command (SPAWAR) and the Ada Joint Program Office (AJPO), Office of the Deputy Director of Defense Research and Engineering (Science and Technology), tasked the Institute for Defense Analyses (IDA) to support the Next Generation Computer Resources (NGCR) Database Management System (DBMS) Interface Standards Working Group (DISWG) in identifying and helping to define DBMS interface standards that can meet Navy mission-critical computing requirements. This document is a compilation of written contributions made by IDA under the task, and covers the time period from April 1992 through September 1992. It is submitted in partial fulfillment of the NGCR Database Management System Standards task.

This document was reviewed by the following members of IDA management: Dr. Richard J. Ivanetich and Mr. Terry Mayfield.

CONTENTS

INTRODUCTION

The U.S. Navy is conducting a standardization effort known as the Next Generation Computer Resources (NGCR) Program. The NGCR Program is designed to fulfill the Navy's needs for standard computer resources while at the same time allowing it to take advantage of commercial products and investments and to field new technology more quickly and effectively. The program is centered around the selection and adoption of open system standards in several areas, including backplanes, local area networks, operating systems, project support environments, graphics, and database management systems.

In April 1992, IDA began providing support to the NGCR Program in the area of database management systems. At that time, the U.S. Navy Space and Naval Warfare Systems Command formed the NGCR Database Management System (DBMS) Interface Standards Working Group (DISWG), of which IDA is a member. The DISWG is chartered to identify and help define nonproprietary industry-based database management system interface standards for use in the development and maintenance of future Navy mission-critical computing systems.

During the April 1992-October 1992 time period, the DISWG held a series of small preliminary meetings, in preparation for its first full-scale, joint industry/Navy meeting to be held 8–9 December 1992. At the preliminary meetings, the DISWG took initial steps toward identifying requirements, reviewing available technology and standards, and defining its approach to standardization. This document contains copies of IDA's contributions toward identifying issues, formulating a statement of requirements, and refining the DISWG Reference Model.

Background Information and Preliminary List of Issues

Each DISWG member was asked to provide this information prior to the first DISWG meeting, which was held 14 April 1992.

Submitted by Karen Gordon (IDA), 7 April 1992

Karen D. Gordon
IDA/CSED
1801 N. Beauregard St.
Alexandria, VA 22311
phone: (703) 845-6630
FAX: (703) 845-6848
email:: gordon@ida.org

Database Experience

- 1984–1985: Very Large Database (VLDB) IR&D Project at the MITRE Corporation. Investigated methods for improving the performance of relational DBMSs in VLDB environments, via a case study involving the commercial relational DBMS ORACLE and the Navy Sea Watch Database.
- 1985–1986: SDI Large Scale System Technology Study Panel. Served as a member of this nine-person panel, assembled by the U.S. Army Strategic Defense Command to assist the Army in formulating its SDI BM/C^3 research program. Represented the Data Management technology area. Identified four issues as being the most critical data management issues for BM/C^3: data allocation, data consistency, performance, and security. Investigated these issues and presented findings and recommendations in the panel's final report.

Potential Discussion Items/Issues

- Scope
  - Include/exclude data exchange standards?
  - Include/exclude transaction processing standards?
- Approach
  - Use NIST APP as baseline?
  - How can Navy influence evolution of commercially based standards?

- Shortfalls of current standards
  - What are they? (See NOSC White Paper)
  - Can they be overcome?
  - If they can be overcome, how?
  - If they can't be, what should Navy do?
- Requirements
  - Requirements vs. long-term goals
  - Consider formulating a strategy to incrementally meet long-term goals

Draft Data Management Requirements for Shore Systems

At its second meeting (held 5 June 1992), the DISWG broke up into *ad hoc* subgroups to focus on the data management requirements of different types of mission platforms. The four types of platforms considered were (1) air, (2) surface, (3) subsurface, and (4) shore. Karen Gordon led the shore subgroup. Below is a copy of the requirements list formulated by the shore subgroup during the meeting.

Submitted by Shore Subgroup, 9 June 1992

Applications:

a. Command and control

b. Sensors

c. Intelligence

d. Management information systems (logistics, medical, finance, accounting)

e. Simulation and training

f. Science (environmental)

g. Question: scope with respect to MCCR

Characteristics of shore-based information management:

a. Size (amount of data)

   1. Range is from megabytes to gigabytes and even terabytes

   2. Range of accessibility is from on-line to near-line to archival

b. Distribution

   1. World-wide

   2. Global interconnection of shore systems; interfaces to air, surface, undersea platforms

c. Heterogeneity

   1. Database management systems

      - ORACLE, Sybase, Informix

   2. Underlying hardware and software

      - Range is from PCs to supercomputers

3. Data models
    - Flat files
    - Hierarchical
    - Network
    - Relational
    - Object-oriented
    - Knowledge-based/deductive/rule-based
4. Types of data
    - Text (information retrieval)
    - Images
    - Voice
    - Spatial data (maps, environmental information systems)
    - Binary large objects, in general
    - Time dimension
    - Uncertainty

Requirements:

a. Reliability

b. Availability
    1. Media, system, communication failures
    2. Network partitions

c. Safety

d. Security
    1. Confidentiality
        - Inference
        - Aggregation
        - Data and traffic
    2. Integrity
    3. Availability
    4. Identification and authentication

e. Real-time performance

f. Flexibility

    1. Tradeoffs among reliability, availability, safety, security, real-time performance

g. Interoperability

    1. New systems

    2. Legacy systems

h. Portability

    1. Applications

    2. Data

    3. Users

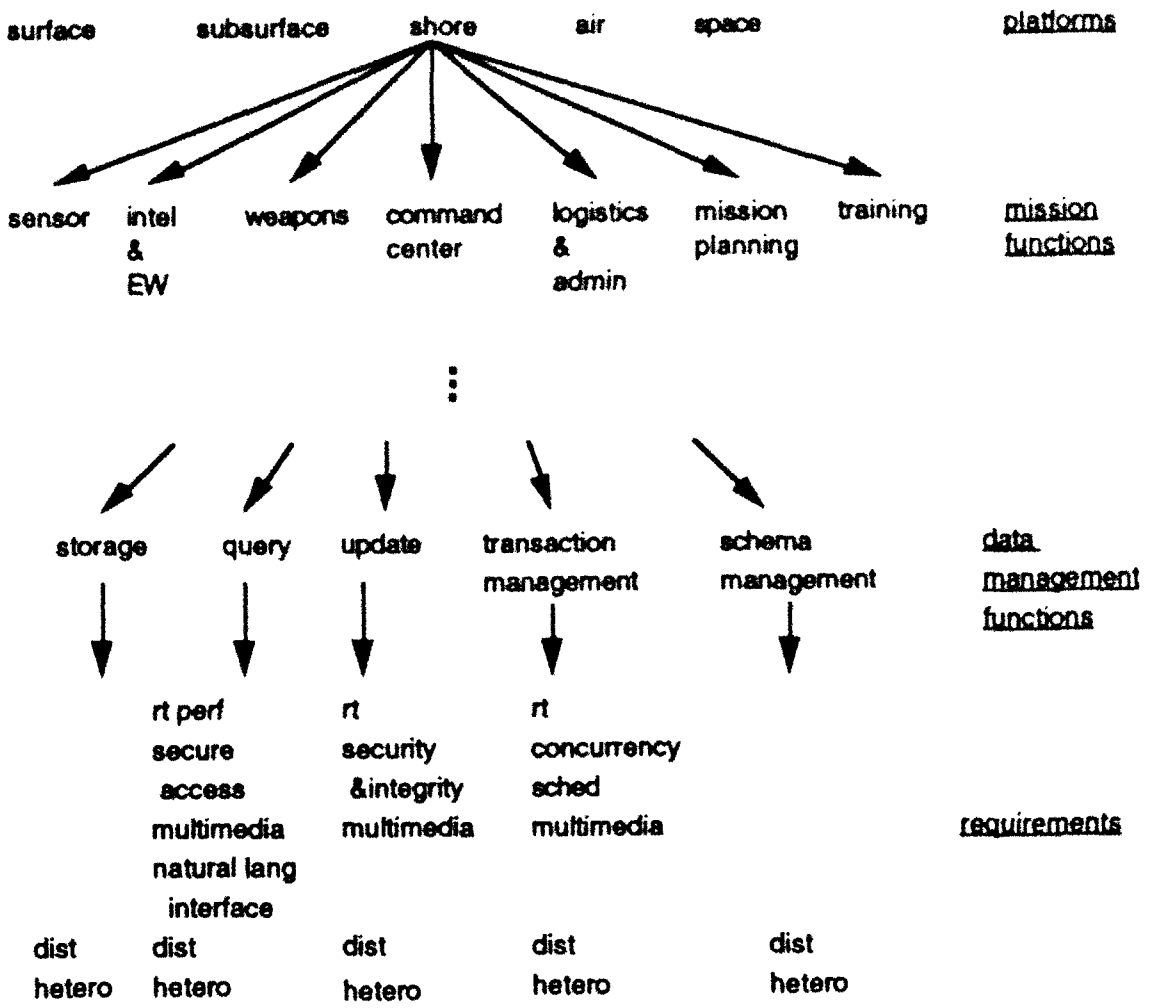
Platform-Independent View of Requirements

The platform subgroups (air, surface, subsurface, and shore) met on 11 August to discuss their requirements further. The shore subgroup, led by Karen Gordon, decided to look at requirements from a higher level. The figure on the next page captures some of their thoughts. After reviewing the lists of requirements being developed by the different platform subgroups, the shore subgroup found that the requirements tended to be the same for all platforms. The reason is that *all* platforms have to perform *all* of a wide variety of mission functions. That is, on shore, as well as on all other platforms, sensor functions, intelligence and electronic warfare functions, weapons functions, etc., have to be performed. This is what the top part of the figure is meant to convey. (Seven arrows could have been drawn from each platform, but then the figure would have been unreadable.)

The bottom part of the figure represents an attempt to place some structure on the data management requirements that were being identified by the various subgroups. As shown, there are certain data management functions (e.g., query, update, schema management) that need to be performed, and there are certain requirements (real-time, secure, distributed, heterogeneous) on the execution of those functions. It is possible to start with functions and then consider requirements (as done in the figure), or to start with requirements (real-time processing, secure processing) and then consider what demands are placed on the various data management functions.

As it turns out, the data management functions are independent of mission function and of platform. The requirements, however, are not independent; they are derived from the needs of mission functions. For example, sensors and weapons require *hard real-time* data management, while intelligence and command center functions can relax real-time constraints to some extent, but do require *secure* data management. (The list of requirements in the figure is incomplete. Also, the list was developed through consideration of all mission functions. Separate mission-unique lists could be developed, and they could possibly be used as the basis for profiles later.)

Submitted by Shore Subgroup, 11 August 1992


surface
subsurface
shore
air
space
platforms
sensor
intel & EW
weapons
command center
logistics & admin
mission planning
training
mission functions
storage
query
update
transaction management
schema management
data management functions
rt perf
secure access
multimedia
natural lang interface
rt security &integrity
multimedia
rt concurrency sched
multimedia
requirements
dist hetero
dist hetero
dist hetero
dist hetero
dist hetero

Comments on Draft DISWG Reference Model

A draft reference model was presented by Paul Fortier of the Naval Undersea Warfare Center at the 11–12 August meeting of the DISWG. Comments on the reference model were requested from DISWG members. Below is a copy of an electronic mail message that provides some comments.

Submitted by Karen Gordon (IDA), 17 September 1992

Paul,

In trying to formulate a structure for stating NGCR data management requirements, I have spent some time reviewing the reference model concepts you presented at the last DISWG meeting. I believe that the reference model, the requirements, and the evaluation criteria should be closely tied together. In order to be able to talk about certain data management requirements, I think it might be helpful to extend the reference model in the following ways:

1. The current model focuses on data manipulation (query and update). It could be extended to encompass data definition as well. We need to have interfaces for database administrators, as well as for application programs and interactive users.

2. The current model doesn't seem to reflect diversity in the kinds of data that need to be managed. I tend to think that we won't be able to have "universal" schemas and interfaces. Instead, we will probably need distinct kinds of data servers for distinct kinds of data (e.g., conventional ADP data, text, geographical data, images, video, etc.). The distinct kinds of data servers will require distinct interfaces. For example, the API for a conventional RDBMS will presumably differ from the API for a text retrieval system or from the API for a database of images.

   The model could be extended to show that we expect distinct sets of interfaces for distinct kinds of data.

3. The current model seems to illustrate a single data server. It could be extended to show multiple, heterogeneous data servers implemented on multiple, heterogeneous, distributed hw/sw platforms.

4. The current model illustrates local transaction management (within a single data server or DBMS). It could be extended to show global transaction management (across multiple, heterogeneous data servers), as discussed in X/Open's Distributed Transaction Processing literature.

5. The model could be extended to reflect the fact that data can be partially replicated.

6. In regard to interoperability, we may need protocols that define common formats for certain types of data (e.g., maps, video, etc.), so that servers of those types of data can interoperate and exchange data. (Is this within the scope of the DISWG? Can we show this in the reference model?)

I think that it would be a challenge to incorporate all these extensions into the reference model, but I believe that we should at least consider them.

Karen

Draft Data Management Requirements Outline

At the 11–12 August 1992 meeting of the DISWG, three official subgroups were formed: (1) Approach Subgroup, (2) Current Standards and Available Technology Subgroup, and (3) Requirements Subgroup. Karen Gordon is the chair of the Requirements Subgroup. The Requirements Subgroup is responsible for preparing the Requirements Document. Below is a draft outline that was distributed to members of the Requirements Subgroup.

Submitted by Karen Gordon (IDA), 29 September 1992

Requirements Subgroup Members:

First, let me thank those of you who submitted draft outlines. Below is a proposed outline of DISWG requirements based on your input, as well as on (1) the requirements lists developed by the platform subgroups, (2) the Requirements Integration briefing given by Paul Fortier, (3) the OSSWG Operational Concepts Document (OCD), (4) the White Paper on the DBMS Interface Standard for NGCR, and (5) the reference model being developed by Paul.

Now, let me explain some of the motivation behind the proposed outline. Based on your input and on the other sources mentioned above, it seems to me that requirements can be viewed as falling into three general categories (and eleven classes) as follows:

(1)	General requirements (e.g., scalability, extensibility, configurability)	[Requirements Class 1]
(2)	Basic functional requirements	
	Data manipulation	[Class 2]
	Data definition	[Class 3]
(3)	Extended requirements	
	Extended data types	[Class 4]
	Distribution	[Class 5]
	Heterogeneity	[Class 6]
	Real-time performance	[Class 7]
	Fault tolerance	[Class 8]
	Security	[Class 9]
	Extended data manipulation	[Class 10]
	Extended data definition	[Class 11]

The GENERAL REQUIREMENTS correspond to the general requirements of the OSSWG OCD. They tend to specify general goals of the interface standards. They do not really state specific functions (e.g., query, update, conceptual schema definition) or specific requirements on function execution (e.g., real-time, fault-tolerant, secure).

The division of the remaining requirements into BASIC FUNCTIONAL REQUIREMENTS and EXTENDED REQUIREMENTS reflects the notion apparent in some of your outlines (and in Paul's new version of the reference model, I think) that there are certain basic interfaces that must be provided in data management interface standards, and there are also some extensions that are important to Navy MCCR missions.

The BASIC FUNCTIONAL REQUIREMENTS correspond most closely to the following classes of OSSWG requirements: file interfaces, generalized I/O interfaces, process management interfaces, resource management interfaces, and time services interfaces. With respect to the OSSWG, these requirements specify what functions should be performed by various components of operating systems. In the proposed DISWG requirements outline, the basic functional requirements are grouped into two classes--data manipulation and data definition.

[I think it might be possible to come up with a better way of dividing basic functional requirements into classes. For example, perhaps classes based on the levels in the reference model being developed by Paul could be used. Your input on this is requested.]

The EXTENDED REQUIREMENTS reflect various directions in which basic data management functionality must be extended in order to meet Navy MCCR requirements. I believe that devoting a class to each direction would help us in formulating requirements and in adopting and influencing data management interface standards.

The specific classes of extended requirements listed below are based on the OSSWG "Big 6" requirements: distributed, heterogeneous, real-time, fault-tolerant, secure, and Ada-supportive. (I have initially left out Ada support as a separate class of requirements, but could include it if there prove to be Ada-specific interface requirements. I have also added extended data types, extended data manipulation, and extended data definition as separate (DBMS-specific) classes). These classes are discussed in the OSSWG OCD (Sections 3.1.5-3.1.10), and have been important throughout the OSSWG's history. In the OSSWG OCD, separate requirements classes were specified for some, but not all, of the BIG 6 requirements. For example, the OSSWG has separate requirements classes for security and Ada. Requirements for real-time performance, fault tolerance, and to some extent

distribution and heterogeneity, are scattered throughout the various requirements classes. On several occasions, the OSSWG has noted that it probably would have been helpful to pull real-time performance requirements into a single class, and to pull fault tolerance requirements into another single class.

I think that most of the individual requirements that have been suggested up to this time could be placed under one (or possibly more than one) of the eleven classes listed in the proposed outline given below. I have, in fact, tried to list most of the individual requirements that have been identified. Please let me of any suggestions you have for additional classes.

One more note: I am not sure what we should do with requirements having to do with tools. We could add a twelfth "Tools" class of requirements, but I'm not sure what data management interface requirements we would put into the class. With respect to interfaces to tools, I'm not sure whether they're in scope.

Based on your feedback on this proposed outline, I'll try to send out a message to the entire working group before the Orlando meeting. I'll also prepare a briefing to give at the meeting. I think it is most important to settle on an outline giving the classes (such as the eleven identified in the outline below) that we will cover in the requirements document. Then, once we've agreed on the classes, we can break up into groups to work through detailed requirements in each service class.

Thank you,

Karen

Proposed Outline of NGCR Data Management Interface Requirements

1. General

   Scope with respect to platforms: surface, subsurface, air, space, shore

   Scope with respect to mission functions: sensor, intelligence and electronic warfare, weapons, command center, logistics and administration, mission planning, and training

   Scalability: with respect to volume of data, number of users, transaction rate

   Modularity

   Extensibility

   Uniformity

   Completeness

   Configurability

   Interoperability with other NGCR standards Others (see, for example, OSSWG OCD)

2. Data Manipulation—Basic

   Centralized DBMS/data server

   Interactive and batch processing

   Multi-user

   Conventional alphanumeric data types, including integer, real, character string

   Consistent alphanumeric sorting order

   Static and dynamic data

   Embedded and ad hoc queries and updates

   Embedded and ad hoc transactions (with conventional ACID properties)

   Discretionary access control

   Independent of hardware/software platform, operating system, network, DBMS/data server (vendor)

   Independent of data model (flat files, indexed files, hierarchical, network, relational, object-oriented)

   Language bindings

3. Data Definition—Basic

Centralized DBMS/data server

Interactive and batch processing

Multi-user

Conventional alphanumeric data types, including integer, real, character string

Data model?

Conceptual schema definition, independent of hardware/software platform, operating system, network, DBMS/data server (vendor)

External schema definition, independent of hardware/software platform, operating system, network, DBMS/data server (vendor)

Mapping to internal schema and database

Language bindings

Supportability

Configuration management

4. Extended Data Types

Text

Image

Voice

Video

Multimedia

Spatial

Graphics

Binary

Time-dependent

Knowledge-based

Uncertain

5. Distribution

- Data distributed across heterogeneous hw/sw systems (transparent to user)
- Homogeneous distributed data servers (transparent)
- Universal accessibility of data, subject to security constraints
- Partial replication of data (transparent)

6. Heterogeneity

- Heterogeneous distributed data servers (transparent)
- Global transactions (across heterogeneous data servers)
- Universal accessibility of data, subject to security constraints
- Data interchange
- Data representation

7. Real-Time Performance

- Real-time (various levels)
- Means of conveying timeliness requirements
- Consistency applied selectively (wrt concurrency control and failures)
- Recovery applied selectively
- Means of conveying accessibility requirements (real-time, on-line, archival)
- Transient data
- Precision applied selectively

8. Fault Tolerance

- Data replication (transparent to user)
- Data server replication (transparent)
- Checkpointing
- Versioning
- Monitoring and logging of faults, errors, failures
- Fault detection, isolation, diagnosis, prediction
- Restarts: cold, warm, hot
- Integrity/correctness/completeness
- Consistency applied selectively (wrt concurrency control and failures)

Recovery applied selectively

Precision applied selectively

Survivability wrt media, system, communication failures

Graceful degradation

Availability

Configuration management

Reconfiguration

Maintainability

9. Security

Security auditing

Least privilege

Mandatory access control

Information labeling

Confidentiality

Inference

Aggregation

Data and traffic

Integrity/correctness/completeness

Availability

Identification and authentication

Multilevel security

Physical separation

10. Data Manipulation—Extended

Nested transactions

Global transactions (across heterogeneous data servers)

Open-ended transactions

Long-duration transactions

Validity of data

Natural language

Voice

Multimedia

Command aids

Training

11. Data Definition—Extended

On-line updates

Resource management

LIST OF ACRONYMS

ACID	Atomicity, Consistency, Isolation, Durability
ADP	Automated Data Processing
APP	Application Portability Profile
BM/C^3	Battle Management/Command, Control and Communications
CSED	Computer and Software Engineering Division
DBMS	Database Management System
DISWG	Database Management System Interface Standards Working Group
IDA	Institute for Defense Analyses
IR&D	Internal Research and Development
MCCR	Mission Critical Computer Resource
NGCR	Next Generation Computer Resources
NIST	National Institute for Standards and Technology
NOSC	Naval Ocean Systems Center
OCD	Operational Concepts Document
PC	Personal Computer
RDBMS	Relational Database Management System
SDI	Strategic Defense Initiative
VLDB	Very Large Database
wrt	with respect to