DEVELOPING DECISION SUPPORT SYSTEMS:
A CONTAINER MANAGEMENT EXAMPLE

Matthias Jarke

August 1982

Center for Research on Information Systems
Computer Applications and Information Systems Area
Graduate School of Business Administration
New York University

Working Paper Series

CRIS #39

GBA #82-52(CR)

To appear in the International Journal of Policy Analysis and Information Systems, Special Issue on Decision Support Systems.

Abstract

The problem of managing an intercontinental container transportation system is used as an example of how knowledge from the areas of database design, management science, and human factors research can be combined in the design of a decision support system. Using a new representation of time-related database objects, we first present a logical data model of a container transportation system. A hierarchically distributed decision support system can be based on this model. A physical database structure is proposed and a survey of partial optimization models to be used in the decision support system is given.

1.0 INTRODUCTION

Several different research areas within computer science, management science, and human factors research have contributed to the emergence of decision support systems (DSS). The purpose of this paper is to outline a procedure for developing a DSS that clearly shows the contributions of: logical and physical database design; process modeling and computerized optimization; and user interface design. One justification for this segregation of activities is that it may be difficult to find systems analysts knowledgeable in all these areas.

The example used throughout this paper is based on the results of consulting and theoretical work on container management problems [6]. The management of intercontinental container transportation appears to be a suitable application because the system is fairly complex and geographically distributed. It requires a high degree of computer support to be efficient, yet many unforeseeable events make the human planner indispensible.

The introduction of container transportation systems some 20 years ago has been called the industrial revolution of general cargo transportation. Containers are large standardized steel or plywood boxes used for intermodal transportation of general cargo. The container offers a uniform interface between the cargo and the transport devices that allows mechanization of the turnaround in ports and inland terminals. However, container transportation systems require high investments in ships, containers, port facilities, and adequate inland transportation. Transportation planning and monitoring become more complex because a new process, container circulation, must be integrated with transport orders and vehicle scheduling.

Early in the container era, carriers expressed the need for better information systems and planning tools. Planning and tracking container movements requires a worldwide system of communicating databases with associated data entry stations and planning support tools. The problem is complicated by the fact that the information needs differ in actuality and precision and that data entry personnel vary widely in skills and working environment necessitating extensive error controls.

The strategy has largely been to build isolated systems for urgent tasks such as port terminal information systems [9], order processing, ship loading plans, or container movement tracking [13]. It has proved difficult to integrate such systems.

In this paper, we attack the problem from a broader perspective. In section 2, we analyze the specific structure of container systems and consider some elementary circulation strategies. Section 3 presents a logical data model of container transportation systems. It provides the background for a decision support system, discussed in section 4, that integrates various levels of container management. A possible physical database structure is outlined and contrasted with the logical structure. In section 5, we give a survey of partial optimization models which can be used to support container movement decisions and relate them to the database structure to determine the communication requirements. Section 6 discusses the role of the human user. The paper ends with the discussion of some general conclusions and open questions.

The structure outlined above can serve as a model for similar DSS development problems. The first step is a process-oriented analysis of the (technical) system in question, and of the major operational objectives of the organizations using it. In the second step, a logical database model is derived from the process model giving a unified description of the technical and organizational information used in decision making. The third step analyzes the decision structures in the system and determines the system boundaries for design.

Within those boundaries, step 4 requires to determine a physical database structure and data entry system that provide the necessary information without excessive communication costs. Steps 5 and 6 develop the optimization algorithms and interactive procedures for supporting the human decision-maker.

The structure of this development process is depicted in Figure 1. As in all "life cycle" concepts, the sequence of steps is not to be taken too seriously. Rather, the three design steps should be addressed in parallel and their interdependence will often require multiple cycling especially through the last steps. Note also that we do not discuss implementation, conversion, or later stages of the life cycle in this paper.

In the following sections, we shall concentrate on the container example; however, generalizable ideas will be stressed where appropriate and reviewed in the conclusions.

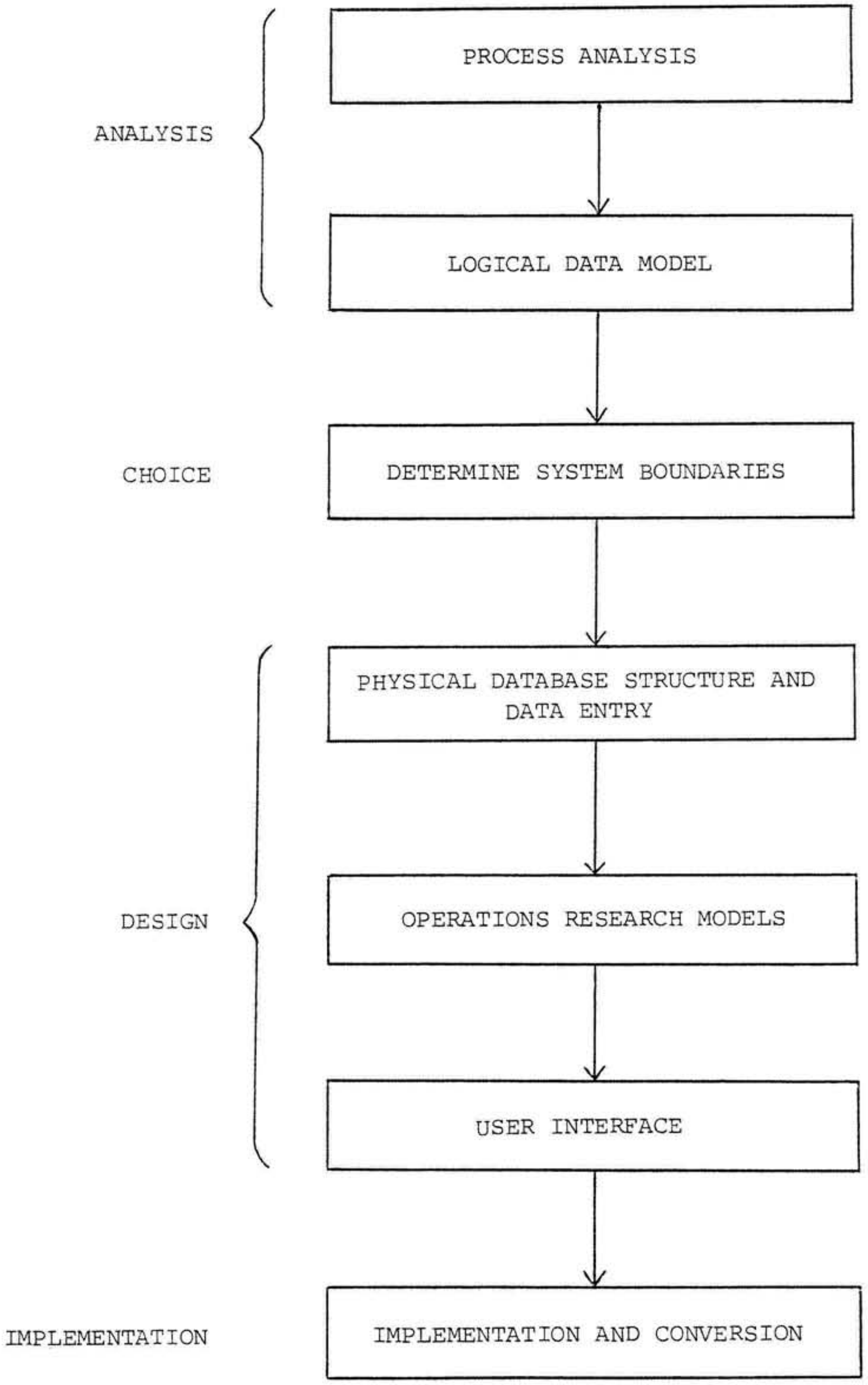

FIGURE 1: A Strategy for DSS Development

2.0 CONTAINER TRANSPORTATION SYSTEMS

This section describes the application of step 1 of our methodology. It examines the technical processes of container transportation and the organizational goals behind them. A brief analysis of some medium-term strategies provides a perspective for the objectives of short-term decision making.

2.1 Process Model

To formalize the structure of overseas transport, we shall adopt the following model of an intercontinental transportation system.

A transportation system connects two ranges. A range is a set of ports that share a common inland system and geographic orientation and therefore compete for the same cargo.

The inland of each range is divided in areas each having a regional center or gateway. The transportation process moves cargo from a sender in some area of one range, the exporter, to a receiver in some area of the other range, the importer.

Ships are used for overseas transportation. Transport between areas is mainly done by train via the gateways. Trucks collect and distribute cargo inside the areas.

Thus, the intercontinental transportation process must be coordinated with circulation processes of ships, trains, and trucks. Between any two modes of transportation, turnover and storage operations occur.

In a container transportation system, the cargo is packed into the container at the exporter and taken out at the importer ("house-house concept"). In between, all transport and turnover facilities can handle the container as a standard-type cargo. This reduces the handling costs and the time of ships in ports because it allows for mechanization of the turnover process. The price for this simplification is the newly introduced process of container circulation: the empty container must be taken away from an importer and stored or transported without cargo to an exporter to be reused.

A schema of the interacting processes described above is displayed in Figure 2. The figure illustrates one circulation of a container (black box when with cargo, white when empty) which is used for two intercontinental transportation processes (one down the left, the other up the right side of Figure 2). For simplicity, only one mode of inland transportation, trucking, is shown.

Figure 2 resembles a Petri net [8] and can be interpreted like one. For example, a container must arrive (on a truck) at the same place where the cargo (black half circle) is before the transportation process can begin.

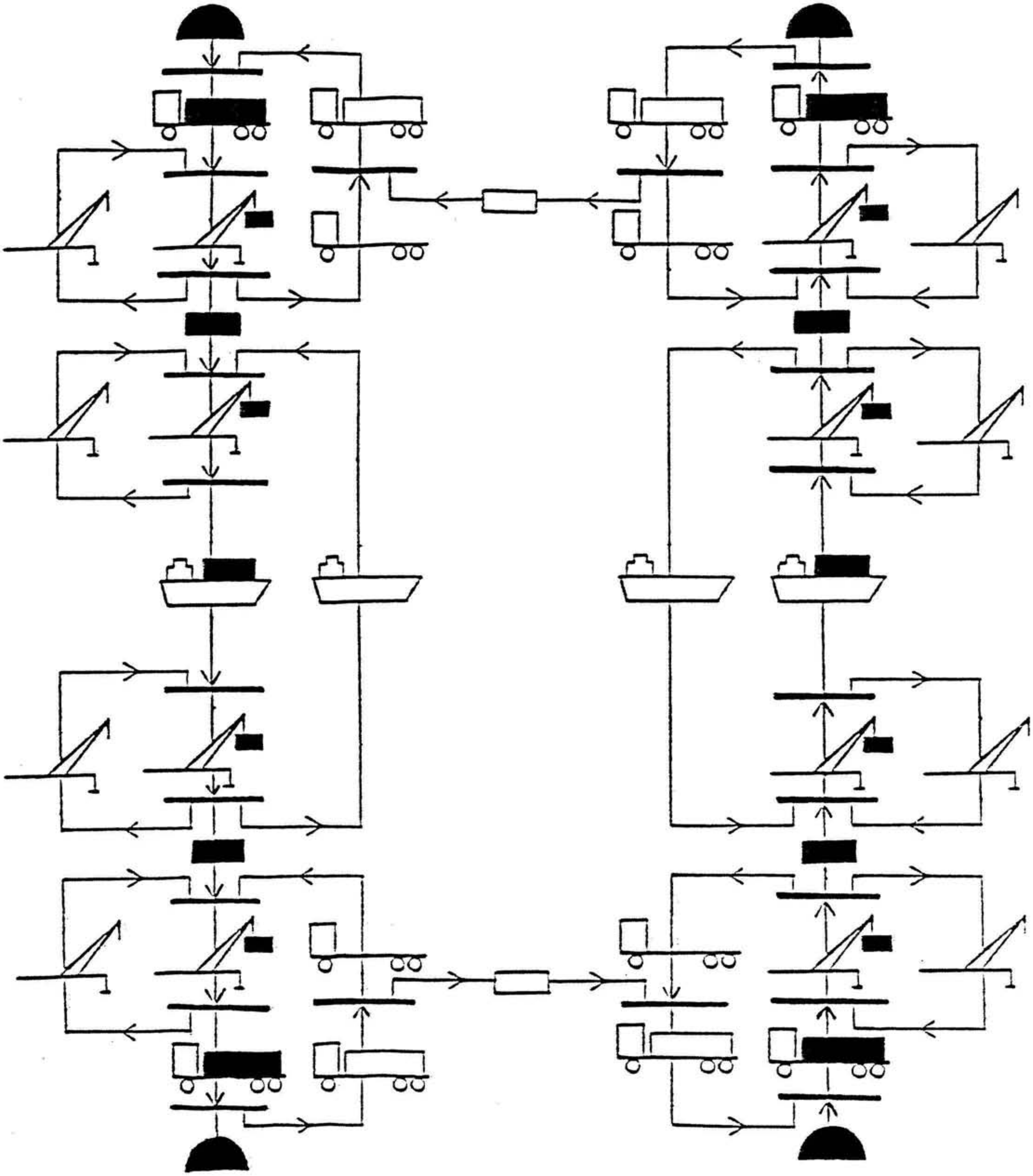

FIGURE 2: Intermodal Transport Using Containers

Note, that the figure distinguishes between time-consuming processes (the icons) and the events (bars) that begin or end a process. We shall return to this important distinction when discussing the logical data model.

The reader is cautioned that the figure gives only a schema of the interaction of different process types rather than showing particular instances like a real Petri net. Nevertheless, this seems to be a convenient notation to describe the technical processes and their interrelationship in a system. We now turn to analyzing the objectives and policies underlying this structure.

2.2 Medium-Term Policies For Container Circulation

After the initial investments in a container transportation system have been made and a service has been established (see [3,12]), planning can mainly influence operational profits and replacement policies of containers.

Medium-term planning concentrates on container circulation and treats the other processes as data. We shall return to the problem of simultaneous planning suggested by Figure 2 when describing short-term decision making.

The income depends on the flow of containerized cargo the service can attract. It is therefore determined by the overall cargo flows, the degree of containerization, and the market share of the carrier. These variables can be influenced to some degree by marketing though the flexibility of a single carrier is often limited by conference agreements. In this paper, we treat the marketing strategy as a datum and focus on cost management.

Important short- and medium-term cost components include the variable costs of container circulation and of replacing or leasing containers. To explain some elementary container circulation policies underlying the short-term decisions described below, we show briefly how cargo flows and container inventory are related.

Let a and b be independent random variables that denote the demand for export containers per day in the two ranges, A and B. Assume $E[a] >= E[b]$. Define the circulation time, T, as the time between two departures of the same container from A. T is a random variable which we assume is limited above, i.e. all containers are eventually returned. T will depend on the value of b, because empty containers can return faster than full containers. In a steady state of the system, the arrival of containers in A will have the same distribution as a and the expected value of the random variable, C, of containers used in the system will be $E[C] = E[a]\ E[T]$.

We conclude that the number of containers one needs depends on the size of the larger of the two cargo flows and on the circulation time; the size of the smaller cargo flow enters only indirectly via T. Thus, speeding up circulation is an important method to reduce container inventory. How can this be done?

In the medium term, the carrier can predetermine routes and modes of transportation for full and empty containers by selecting shortest paths in a cost network which includes both variable transportation and container capital costs. However, the carrier is not always free in its decisions. The shipper can impose time requirements for the transportation process. Further, in a system with relatively inflexible schedules of ships and trains, only inland transportation time and waiting time in harbors can be influenced by the carrier. In some cases, full containers are partially transported under control of the exporter ("merchant's haulage") and no optimization by the carrier is possible.

In short-term planning, unnecessary delays in the circulation process must be avoided. Most importantly, empty containers should be returned from the range, B, quickly. This reduces container inventory without additional variable costs, as may occur with the first method. There are often psychological barriers against such a policy, because transporting empty containers overseas costs money without giving immediate benefit. The consequences of too much hesitation, however, can be seen in many container terminals: export-oriented countries build new containers while empty containers amass in the rest of the world.

These rules are applicable mainly, if containers are a scarce resource in the system. However, an abundance of containers should be reduced in the long run, e.g. by not replacing damaged ones, to avoid unnecessary capital cost.

The above model is obviously very crude and was meant only to outline some important strategies. For instance, it assumes that the probability of b > a is small and therefore no empty containers are shipped from A to B. In addition, there must always be enough containers in reserve to cover peak demand. These conditions can usually be satisfied, because (1) demand averages out when there is an interdeparture time of ships of more than one day, (2) a safety stock of containers is held, and (3) additional containers can be leased.

A more complex model for determining the optimal container inventory was developed by Samuel and Ullman [10] for a mail container system serving two gateways. It was extended by Horn [4] to a network optimization model for the multi-gateway and multi-period case. By redefinition of the parameters, this model can easily be adapted to our problem. Its main drawback is that it is limited to solutions in which the container flow is determined at the beginning of the planning period and thus independent of the actual cargo flow. A more adaptive approach is needed and currently under investigation. For the short-term planning situation, we assume that time constraints which apply the above recommendations will be incorporated into the data.

In conclusion, the result of step 1 (process analysis) is an understanding of a container transportation system as a network of interacting processes guided by a cost minimization goal. The medium-term cost analysis suggests that container utilization over time must be taken into account when designing a DSS for short-term decision support.

3.0 LOGICAL DATA MODEL

Step 2 of the methodology outlined in Figure 1 requires the development of a logical data model from the process model. For planning purposes, one needs more than a simple description of the technical system in the database. The requirements include data about

1. the technical entities;
2. the observed behaviour of these entities (events);
3. organizational entities;
4. the exchange of orders among the organizations;
5. the owner or user relationship between technical and organizational entities;
6. constraints governing the planning process;
7. plans which organizational entities have developed for the usage of their technical entities.

Figure 3 shows a logical data model of a container transportation system the layers of which represent the above data types. We first discuss the modeling approach and then describe the specific model in more detail.

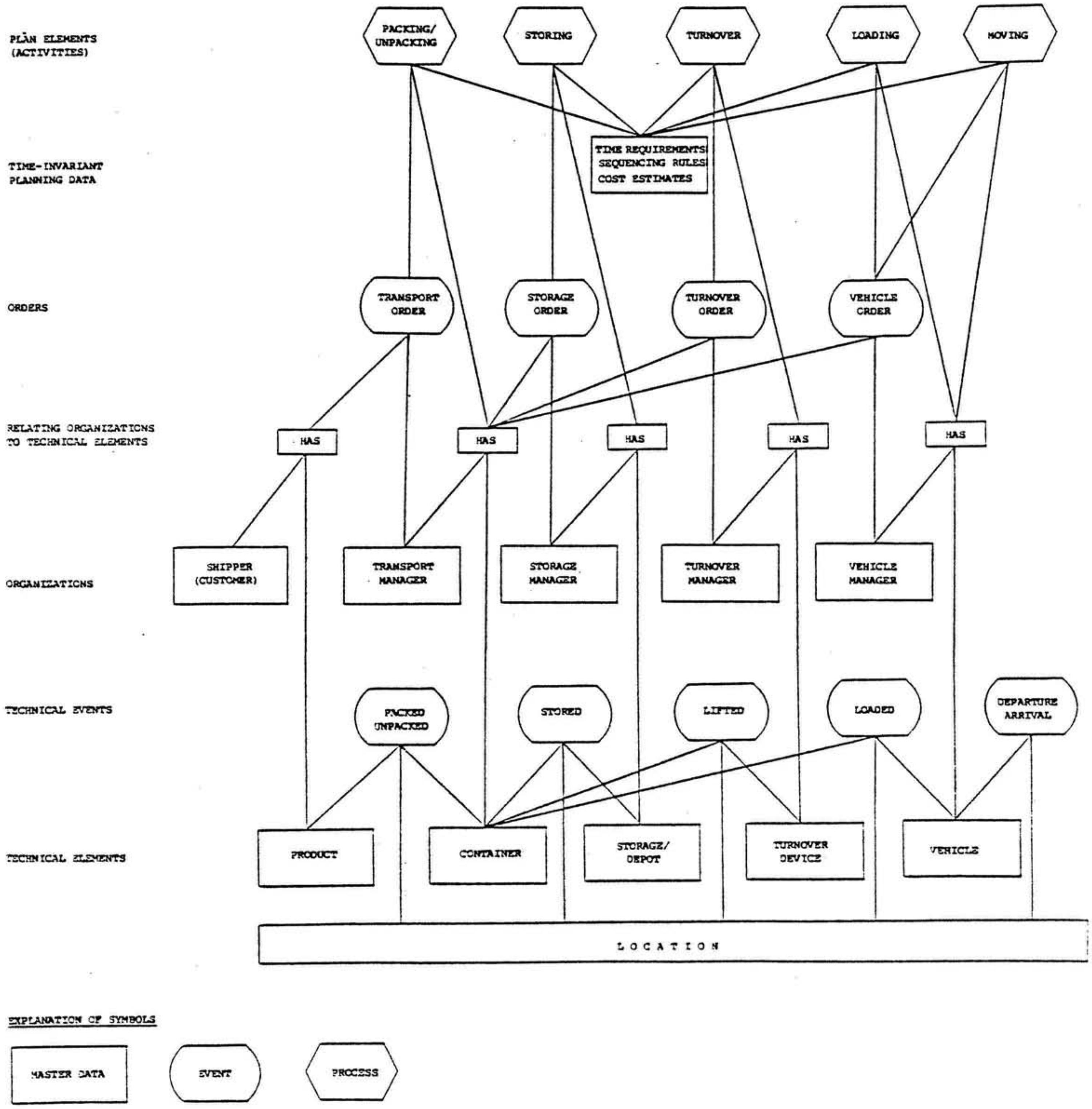

FIGURE 3: Logical Data Model

3.1 Modeling Approach

The design tool used in Figure 3 is an extended version of the aggregation hierarchy model [14] that maps the structure of the real world into a hierarchy of object classes.

Object classes are denoted by nodes of the model. Simple objects can be aggregated to complex objects; this is denoted by arcs connecting several lower-level object classes to one higher-level object class. For example, a LOCATION, a CONTAINER, and a transport VEHICLE, e.g., a truck, can be aggregated in the higher-level object, LOADED, meaning that a container is loaded on a truck in a specific location.

The answer to one important question has not been modeled so far: when did (or should) this happen and how much time did (will) it require? Our extension of the model aims at answering this by classifying objects according to their time dependence. We briefly discuss the shortcomings of existing models to motivate our approach.

Conventional data models map a state of the system, for example the fact that a container is currently aboard a truck. This is infeasible for container systems because they are geographically so distributed that it is impossible to check the state of the system at some moment.

Some approaches in the literature map historic developments by capturing events which change the state of a system [1], for example the event that a container has been loaded on a truck. This means capturing the time when a bar of Figure 2 was reached. It is a good solution for describing the past.

A planner, however, is more interested in future time behavior. For him, the duration of a process and the rules governing it are the important data while the starting time is his decision variable. As an example, assume that the planner wants to plan the utilization of a truck during the day. To determine, if and when a specific container should be transported by this truck, he needs information about the duration of the trip and about preconditions that must be satisfied before the trip can start (compare Figure 2).

To visualize the time dimension in the data model, different symbols are introduced for the different lifetimes of database objects. We distinguish master data with a very long lifetime (denoted by rectangles) from process data with a limited lifetime (hexagons) and from event data that map a "sudden" change in state (ovals). The further discussion will show how the introduction of time-related symbols supports the design of physical databases and data capture systems.

3.2 Application Analysis

The three lower levels of Figure 3 map the technical subsystem. The structure reveals the central role of the container between the transported good and the different types of transport facilities. As transportation means change of location, it is not surprising that a second object class of importance is LOCATION. The status of the technical subsystem at any moment can be described by the events that led to the current situation. Processes completed in the past can be described by two (or more) successive events involving the same combination of technical entities. For example, a trip of a vehicle can be described by ARRIVAL and DEPARTURE events; similarly, storage of a container at some location is mapped into a sequence, STORED - LIFTED.

The upper levels of Figure 3 map the organizations involved in container transport and their interaction via orders. Note, that the organizations are defined by their functional roles; in a specific system, one organization may carry out several of them. The importance of the concept, order, for the organization has been exploited in the BIAIT approach to business systems analysis [17]. BIAIT is, however, only interested in the role of the organization as a receiver of orders while an approach that tries to capture a whole system must include the issued orders, too.

The model defines an order as a three-way relationship (more precisely: event) among the issuing organization, the accepting organization, and the requested service which, in our example, is related to specific technical entities. This is why all orders are aggregations of HAS-type objects and an organization. It makes clear that the accpepting organization has freedom of choice how to excute the order.

An order must be related to an order execution process. An organization that receives and accepts an order has to divide order processing into elementary steps that can be executed either by the organization itself or by issuing suborders to other organizations.

To execute a step, technical facilities (and manpower) must be assigned to it for a certain time interval, i.e. the "lifetime" of the process associated with the step.

First, the planner must find out these times, by communication with a subcontractor, by estimation, or from a database as shown in Figure 3. A timetable or street network may be such a database. Next, the planner must aggregate the resource allocations into plans that assign starting points in real time to the steps. Plans such as order execution plans and resource utilization plans are different aggregations of these elementary assignments of resources to steps. To avoid overloading, they have been omitted from Figure 3 which shows only the plan el

In conclusion, the introduction of time-related symbols reveals an important distinction in data modeling for decision support systems. Plans for the future need information about both expected lifetime and position in real time whereas recording of past history requires only events to be stored. A decision support system can exploit this observation by combining a relatively simple message system for monitoring performance with more sophisticated databases and optimization strategies for planning. This is important especially in systems like container transportation where the state of the system cannot easily be determined directly.

In the remainder of this paper, describing steps 4 through 6 of our methodology, database structures and plannning procedures will be discussed. The _system boundaries_ are generally set as to apply to large integrated organizations that perform most of the activities themselves. At some points, however, alternative solutions for different organizational settings will be described.

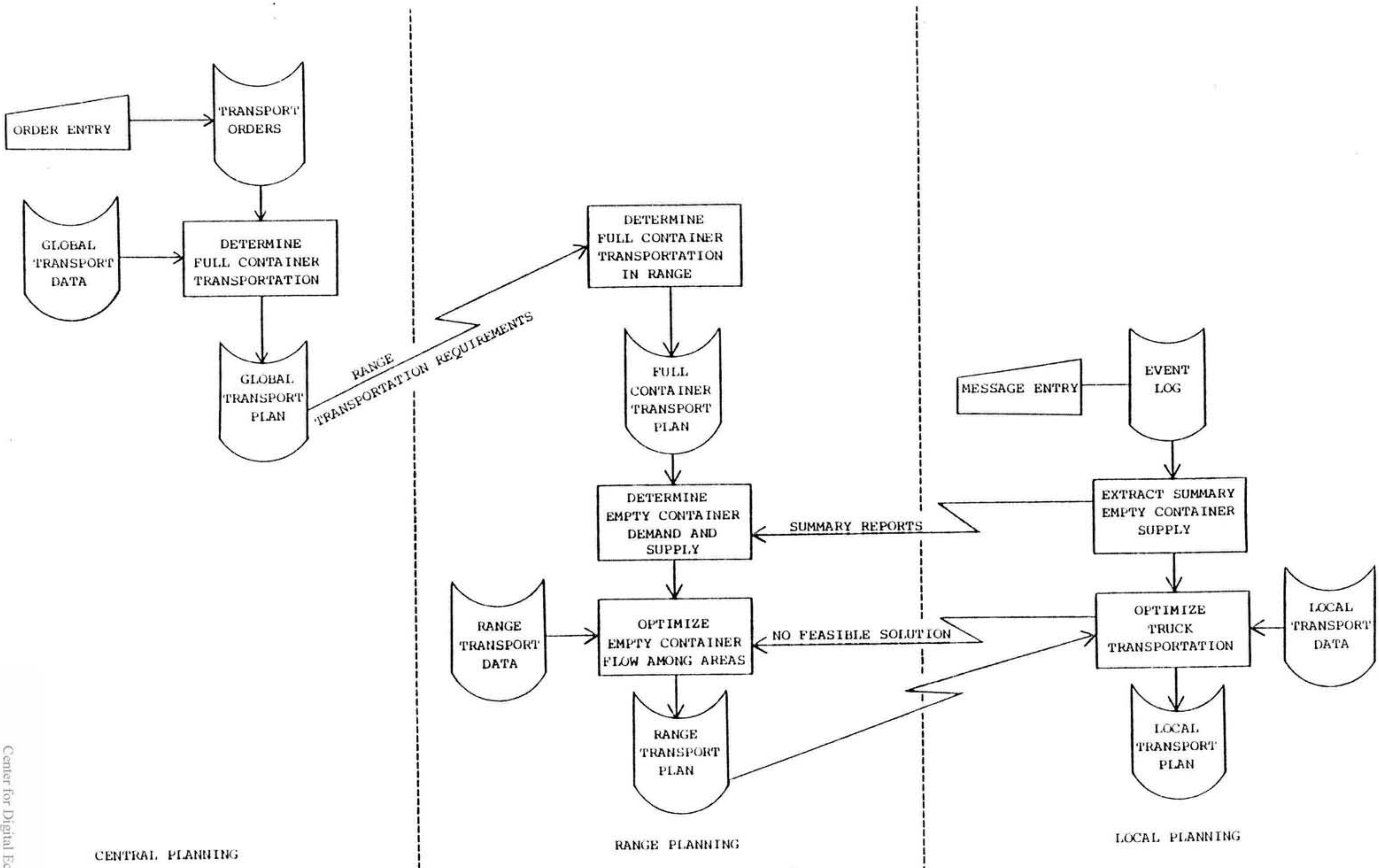

FIGURE 4: PLANNING PROCEDURE

4.0 DATABASE STRUCTURE AND DATA ENTRY

This section covers step 4 of Figure 1 and introduces the reader to the decision problems to be supported by the DSS. Although planning applications will be emphasized, the reader should bear in mind that only integration with order processing, transport documentation, and accounting justifies the data collection efforts.

4.1 Principle

The basic principle of planning container circulation is simple. If a container is <u>full</u>, it must be navigated individually following the transport order and the general policy of the carrier (throughout this section, carrier's haulage is assumed). <u>Empty</u> containers should be reused because of their high value, but the owner can exploit the fact that they can substitute for each other to minimize the costs of their reassignment to new jobs.

Because of the dispersion of containers over the world, container circulation planning is organized in a hierarchy that delegates regional decisions to the ranges and local decisions to the areas. The interdependence between the levels is resolved by "heuristic decomposition". The range level estimates the time requirements for area transportation. These estimates constitute data for the area. If they prove unr the range plan can be partially corrected using, e.g., sensitivity analysis techniques.

The decision hierarchy works on a hierarchy of databases. They must be structured in a way that allows for much local data capture, storage, and processing with limited communication between the levels. Our analysis of data capturing and usage for planning will show that a fairly clean separation of data is indeed possible.

Figure 4 gives an overview of the planning procedure In the following subsections, we discuss the important steps.

4.2 Getting Reliable Data

Planning information stems from two types of sources: incoming orders and messages reporting technical events in the system. Among the orders, the transport order is most important because it triggers the other orders. For planning container circulation, the following observations seem useful:

* Transport orders constitute demand for empty containers in the export area.
* Transport orders determine the movement of full containers.

- Transport orders determine future supply of empty containers in the import area.
- Messages like LIFTED, LOADED, or STORED report the actual movement of containers.
- Arrival and unpacking messages report the actual supply of empty containers.

We discuss these points in more detail, below.

Exporters need empty containers. Their transport orders define the <u>container demand</u> by time and location, as well as the (often predetermined) routes and modes of transportation. It is difficult to get advance information about intended transport orders. Interviews with European shipping companies showed that this is one of the major problems constraining the applicability of decision support systems [6]. Demand must usually be estimated, also to take into account the fact that business will go on after the end of the planning horizon.

Import and empty containers arriving in an area constitute <u>container supply</u>. Forecasting can use existing transport orders and is much simpler than export forecasting because the travel times of ships are relatively long. For example, many shipping companies give precise import information to the ports though they rarely offer realistic export forecasts.

The inland system shows many irregularities that influence the actual arrival times; they must be controlled by a sophisticated monitor system which compiles messages created by customers, truckers, gateway offices, railways, shipping companies, and ports. The basic message structure of a monitor system is defined by the event data of Figure 3. However, the following practical problems must be solved to form a useful system:

1. Messages can get lost and almost always contain errors. For important identification data like the container number, error-detecting codes [6] or scanning devices may help. Further, each message must include redundant information, in case a message is lost or a special event such as damage or theft of cargo is not communicated separately.

2. Different users need data of different actuality and detail. To plan, e.g., port terminal operations, messages about container movements must contain detailed locations and have to arrive within seconds while the carrier may need only daily information that the container is in the terminal, was loaded on a certain ship, etc. Typically, the local processor (in the above example the port) captures the detailed information and communicates appropriate summaries to other interested parties (the railways, the carrier, and the customer).

3. If an integrated information system does not exist, additional pre-information about major expected orders should be communicated among the organizations to avoid surprises that defeat planning efforts.

In existing monitoring systems, the main communication link for long-distance messages is the telex system [13]. Errors are often corrected manually, necessitating a batch-like procedure. The cost of on-line CRT input or automatic data capturing seems to be justified only at major terminals [9].

4.3 Database Structure

The integration of such message systems with order processing is in the initial stage at some advanced companies. An efficient implementation would use distributed database systems. Local databases on mini- or microcomputers in each area would contain street networks, detailed container positions, and current local orders. Other local databases (usually external to the carrier organization) would cover port and railway operations. They would communicate with a central mainframe database for the range which stores order data, routing information for full containers, and supply or demand of empty containers by area. Figure 5 outlines the structure of such a system.

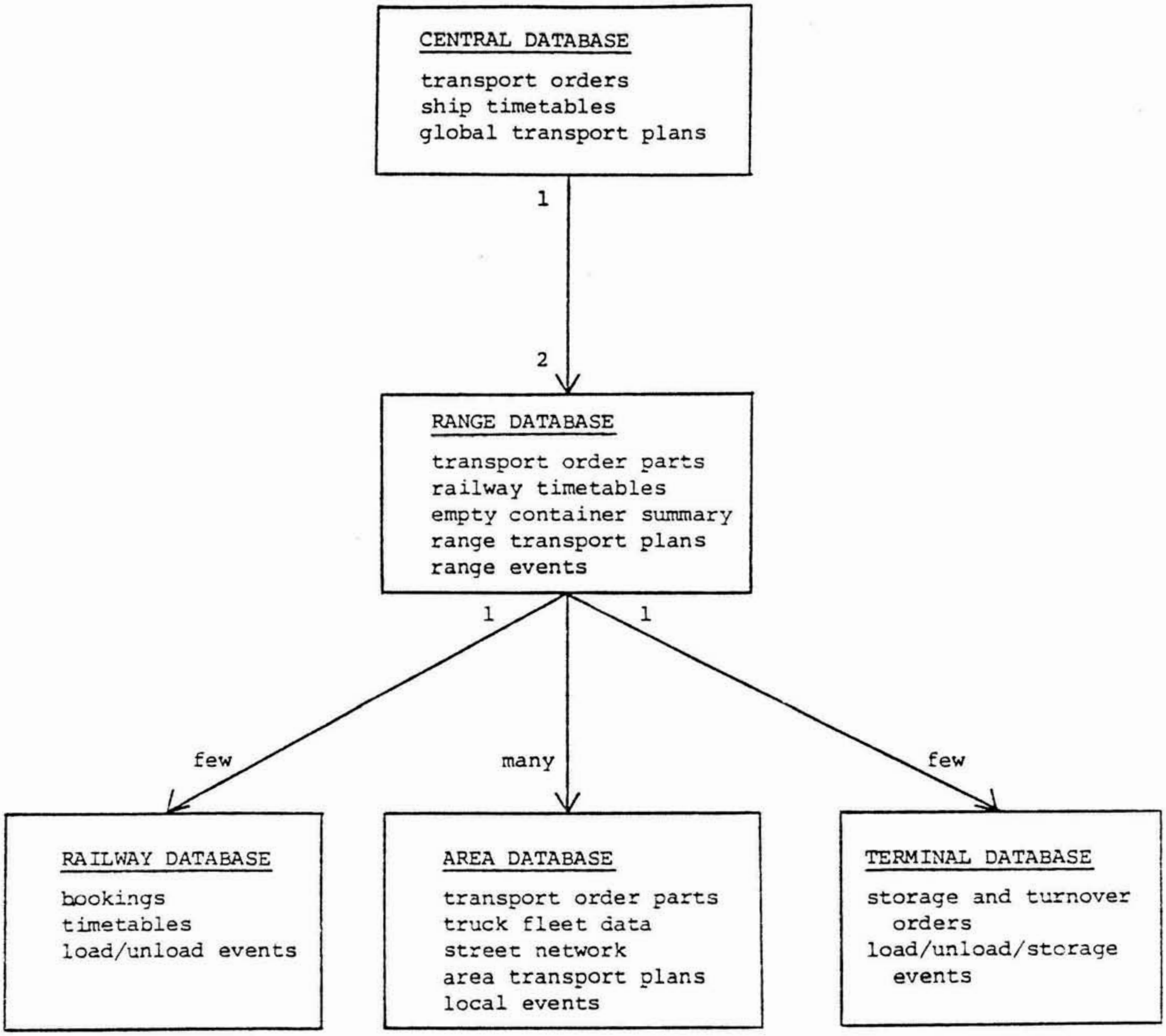

FIGURE 5: Physical Database Hierarchy

The physical database structure is influenced both by the logical structure given in Figure 3 (which data are needed) and by the geographically distributed planning procedure given in Figure 4 (where data are needed). To develop the physical data structure from the logical, the methods of vertical splitting, horizontal splitting, file duplication, and summarizing have been employed.

Vertical partitioning is based on the different roles of an object in the logical data model or, practically speaking, on the different application programs working on the data. As an example, consider a transport order. It is treated differently by the customer and the carrier, but there is also a split within the carrier organization. While the overall order data are kept centrally to determine the transportation process and to process financial data, the details of how to satisfy the overall plan are up to the export and import area, the port, and the inland transportation organization.

Horizontal partitioning is based on the fact that objects of the same type are geographically distributed and that almost all processing of each object is done on one site. For example, it rarely occurs that one area is interested in the street network or exact container locations of another area. It saves transmission costs to store and process this type of data locally.

Data duplication is useful when many reading accesses are made to the same data from several sites while there are relatively few updates. In the container system, this applies to some portions of the order data that are used both locally and at the center. However, the local copy will have a shorter lifetime than the central one because it is needed only during part of the transportation process. Data duplication is equivalent to the manual strategy of using multi-copy forms; this is also true in the sense that both techniques improve the reliability of the system if the consistency of all copies can be ensured.

As mentioned before, summarizing is one of the most important ways to reduce data transmission. Summary data are either selected detail data (e.g., the arrival message is chosen among all container movements messages kept in the port database to be communicated to the carrier) or functions applied to groupings of data (e.g., computing container demand by area). The advantage over duplication is the reduced cost of data transmission and storage. Again, however, summarizing is only feasible if the underlying detail data are relatively stable.

Figure 6 contrasts some important logical data objects from Figure 3 with the physical structure outlined in Figure 5 to give an indication of the degree of distribution and of the communications requirements of the database structure proposed here.

LOGICAL DATA ELEMENTS	PHYSICAL DATABASES				
	CENTRAL DB	RANGE DB	AREA DB	TERMINAL DB	RAILWAY DB
ORDERS					
transport	FO	VD	VD		
storage	FS	HS	HO	HO	
turnover	FS	HS		HO	
vehicle	FS	HS	HO		HO
PLANS					
ship	FO	VS		HS	
railway		VD	HS	HS	FO
truck			HO		
turnover		HS	VD	HO	VD
transport	FS				
EVENTS					
pack/unpack	FS		HO		
stored		FS	HO	HO	
lifted		FS	VD	HO	VD
loaded		FS	VD	HO	VD
departure/ arrival	FS	HS	HO	HO	HO

FIGURE 6: Mapping the Logical Data Model into the Physical Database Structure

EXPLANATION OF SYMBOLS:

F	all instances of object class are held
H	horizontal partition of object class is held
V	vertical partition of object class is held
O	original copy is held
D	duplicate is held
S	summary data are held

5.0 PLANNING PROCEDURES

We are now in a position to look in some detail at the computer-aided planning procedures and partial optimization algorithms that work on the database structure. The roles of computerized algorithms (step 5) and human planner (step 6) will be discussed for each level of the planning hierarchy and summarized in subsection 5.3.

The difference between the numbers of incoming export containers and the number of empty containers requested by exporters determines the net supply or demand of empty containers in each area and time period (usually a day). Long-term considerations may require the planner to assign artificial demand to certain areas.

The net position of empty containers is the interface between global and area models. The global model plans the container flow among the areas assuming that all containers are supplied or demanded at the gateway. The area models strive to satisfy this condition within the time limits set by the global model. We discuss both model types in turn.

5.1 Global Models

In the global models, the main problem is to minimize the cost of storing and moving empty containers. Full container transport enters the model only in that it reduces the capacity of some arcs of the transp network. For practical purposes, the problem can be partitioned into three subproblems, sea transport and the two inland systems.

Sea transportation models must provide a sufficient number of empty containers in both ranges taking into account temporary bottlenecks in overseas transportation capacity for both full and empty containers. As sea transport is relatively slow, the planning horizon should be at least three months.

If, as is common practice, the container owner also owns the ships, he can plan the schedule of the ships simultaneously with the assignment of containers to ships. A recently developed decision support system [18] allows the human operator to change the schedule interactively whereas the system optimizes the assignment problem with a few additional linear constraints.

The planning horizon of inland models can usually be smaller than in sea transport, but the inland network is much more complex. Therefore, a network optimization model that minimizes the cost of empty container transportation can be quite large. If a more or less time-invariant schedule exists, major parts of the model can be stored in the database as a time-location network [16]. In this case, only the time-varying data about supply, demand, and external transport capacity utilization (by full containers or containers of competitors) must be entered dail

While similar models have been successfully implemented for scheduling empty freight cars [11], their practicality for container planning depends heavily on improved message systems. Currently, planning is mainly done manually supported by batch reports from the existing message systems. Even with improved data that allow for computerized planning, however, many decisions will remain with the human planner. Probably the most important is the assignment of artificial demand for empty containers to areas in which orders are expected. Automatic forecasting methods can only support not replace this task.

5.2 Area Models

In each area, the planner must try to satisfy the requirements of the global plan. The planning horizon usually does not exceed one day; longer hauls are handled by restricting the number of available vehicles. Area transportation consists of three tasks:

- Distribute full import containers.
- Collect full export containers.
- Reassign empty containers; we assume that a depot near the gateway stores empty containers which are not currently needed.

Trucks, usually semi-trailers, perform these tasks. As an important planning element we use the concept of a roundtrip. A roundtrip is the sequence of trips between two successive stops of some given truck at the gateway. The sequence of roundtrips one truck is assigned during the planning period is called its tour.

The concept, roundtrip, has some nice properties for practical planning purposes. Many different area transportation policies can be transformed into restrictions to feasible roundtrips. To assign roundtrips to trucks, only their duration must be considered because all of them start and end at the gateway. Additional time constraints in a transport order, for example restrictions in office hours of the customer, can often be accommodated by interchanging the sequence of roundtrips within a tour or by exchanging roundtrips of similar length between trucks. Additional roundtrips required because of rush orders normally do not change other roundtrips.

Most trucks carry only one container at a time. Therefore, full container transportation determines routes that must be included in every feasible roundtrip plan. Transportation of empty containers is more flexible. The degrees of freedom - and the complexity of optimization models - depend on the following basic policy-making choices to be made by the carrier company:

1. The first choice is between using external organizations or a self-owned fleet of vehicles. If the company only issues orders to truck-owners covering single container movements at given tariffs, costs can only be reduced by optimizing the assignment of empty containers. If, on the other hand, own equipment is used, the carrier must simultaneously optimize vehicle scheduling.

2. A container truck consists conceptually of three parts: a tractor, a trailer or chassis, and the container. The second question to decide is whether to keep these parts together during roundtrips or to separate them at the customer's place.

Where to separate, is determined by technical constraints. If the customer has turnover equipment for containers, tractor and trailer can be used elsewhere while the container is packed or unpacked. Otherwise, at most the tractor can move freely. If it is not separable from the chassis, or if management determines that the driver should supervise packing for sea transport, no separation is possible at all.

Obviously, the more separation is feasible, the better the equipment utilization, resulting in lower investment costs. A policy of non-separation allows only for linear (to a customer and back) or triangular roundtrips including an empty container transp

3. Finally, it must be decided whether containers may be redistributed directly or only after some overhaul at the gateway. This decision determines again the possible structure of roundtrips. Redistribution via the gateway limits roundtrips to linear or at most triangular structures.

Each combination of the three policy decisions leads to a different optimization problem; the types of these problems are listed in table 1.

TABLE 1: Optimization Models for Area Transport

	redistribution	
	direct	via gateway only
foreign equipment	assignment problem (empty containers)	no optimization possible
own equipment		
- with separation	general routing problem (truck routes and empty containers)	assignment problem (truck routes)
- w/o separation	assignment problem (empty containers)	no optimization (except loading problem trucks[5])

Note that most of these models have a relatively simple structure. Even the general routing problem [7] can often be solved more efficiently than typical vehicle scheduling problems:

1. Full container trips (import and export) determine arcs which must be visited. All these arcs begin or end in the gateway and thus form one connected component of the network.

2. A subset of nodes that must be visited is given by the supply and demand points of empty containers. However, there are severe restrictions on the sequence of these points. Immediately after each supply point, a demand point must be visited, because a truck can carry only one container at a time. Thus, a fairly separated assignment problem is embedded in the general routing problem.

3. Time constraints for the daily use of the vehicles exclude very long roundtrips.

Orloff and Caprera [7] have shown that on the average general routing problems with a small number of connected components can be solved substantially faster than the usual vehicle scheduling problem where only nodes to visit are given. If the variable costs of transporting empty containers are high compared to those of driving an empty semi-trailer tractor without a trailer, the problem can be reduced to a rural postman problem [7] by solving the empty container assignment problem first. This approach will reduce the number of connected components especially in areas with unbalanced empty container positions, then most empty container transports will also start or end at the gateway. In [6], we devised a branch-and-bound algorithm suitable for smaller areas and a heuristic procedure based on the savings method [2] for larger areas.

5.3 The Human Decision-Making Input

We have shown that technically feasible models exist for each of the policies considered. Suppose for a given set of policy choices we have implemented the appropriate models together with the appropriate database as discussed in section 4. Nevertheless, the models cover only the standard situation where all relevant information exists at the beginning of the planning horizon. Additional short-term problems must be solved interactively by the human user and the model together with the underlying database. There are two essentially different ways in which the human planner can influence the results of the model.

Firstly, he can partially overwrite the contents of the database that provides the <u>input</u> for the model and do sensitivity analysis. Examples of situations where this may be useful include: changes in traffic conditions that influence travel times and costs; changes in the availability of trucks; and delays in packing or unpacking containers.

Secondly, the planner can change the output of the model. A typical example is the addition of simple roundtrips for rush orders. Other problems that can be solved in this way are the adaptation to additional time constraints and the assignment of personnel. Such tasks can be supported by the system in various ways. The system can check whether a solution proposed by the human operator is feasible. It can support the solution-finding process by graphics that show, for example, which trucks have enough slack time to assume additional tasks. The system can finally suggest solutions using human-controlled partial enumeration.

From the above discussion, it can be seen that a lot of work remains for the user of an area decision support system, even if partial optimization models take over standardized parts of it. It is therefore essential that the user interface of the system supports the tasks outlined above both by flexible input and report generation procedures and by suitable presentation methods for data and results. A number of recent approaches to the design and evaluation of user interfaces are described in [15].

6.0 CONCLUSIONS

The analysis of decision support systems for container transportation systems leads to some important findings about data modeling for DSS. Firstly, the relationship between planning and control (that is, between orders, plans, and events) can be modeled elegantly in a layered logical data model. Secondly, planning and control need information about both the past and the (expected) future. Often knowledge about the duration of processes exists independent of knowledge about their position in real time, or knowledge about events exists without knowledge about the underlying processes. Therefore, the logical data model should be capable to map both process and event data, that is, it should have a dynamic rather than a historical concept of time.

From the logical data model, database structures and message systems can be derived taking into account practical constraints like communication costs or imperfect information. In the container system, this means a hierarchical distribution of databases and planning responsibilities. It was shown that computationally feasible exact methods or good heuristics exist for partial optimization models. The ways in which the user can influence system behavior were categorized and the importance of flexible user interfaces was stressed.

More research is needed to establish the optimal division of work between the human planner and the decision support system. Another focus of current research in the database area is the design of interfaces for the wide variety of different users that perform data entry, database querying for details and summaries of time-related data, data analysis, and computer-aided optimization. Finally, research in heterogenous distributed databases has to provide the answer to open questions in the areas of data transmission between area and range databases and query processing that involves data on several different computers.

References

1. J.A.Bubenko: Conceptual Models and Their Semantics, *Proc. 62nd Infotech State of the Art Conferemce on Data Design*, London 1979, 58-87

2. G.Clarke and J.W.Wright: Scheduling of Vehicles from a Central Depot to a Number of Delivery Points, *Operations Research* 12 (1964), 568-581

3. S.Gilman and G.F.Williams: The Economics of Multi-Port Itineraries for Large Container Ships, *Journal of Transport Economics and Policy*, May 1976, 137-149

4. W.A.Horn: Determining Optimal Container Inventory and Routing, *Transportation Science* 5 (1971), 225-231

5. M.S.Hung and J.R.Brown: An Algorithm for a Class of Loading Problems, *Naval Reseach Logistics Quarterly* 25 (1978), 571-579

6. M.Jarke: *Ueberwachung und Steuerung von Container-Transportsystemen*, (Control and Logistics of Container Transportation Systems) Wiesbaden, West Germany, Gabler 1981

7. C.Orloff and D.Caprera: Reduction and Solution of Large Scale Vehicle Routing Problems, *Transportation Science* 10 (1976), 361-373

8. J.L.Peterson: Petri Nets, *acm computing surveys* 9(1977), 223-252

9. T.Sakai: Data Management Systems for Container Terminals - Design, Application, and Impacts, *Proc. 2nd Terminal Operation Conference*, Amsterdam 1980, 67-76

10. A.H.Samuel and Z.Ullman: Optimal Container Inventory for a Two-Gateway System, *Transportation Science* 5 (1971), 328-332

11. F.E.Schlaepfer: Kostenoptimale Verteilung leerer Gueterwagen, (Cost-Optimal Distribution of Empty Freightcars), in E.Nievergelt (ed.): *Praktische Studien zur Unternehmensforschung*, Berlin-Heidelberg-New York 1970, 147-192

12. H.Seelbach and J.Rohlffs: Ein Investitionsmodell fuer die Einrichtung eines Containerdienstes (An Investment Model for the Introduction of a Container Service). *Zeitschrift fuer Verkehrswissenschaft* 4 101-117

13. A.M.Smith: Telex Computer Tracks the Box, Containerisation International 12 (1978), 8, 71-73

14. J.M.Smith and D.C.P.Smith: Database Abstraction: Aggregation, Communications of the ACM 20 (1977), 405-413

15. Y.Vassiliou, M.Jarke, E.A.Stohr, and N.White (eds.): Proceedings NYU Symposium on User Interfaces, New York 1982

16. W.W.White: Dynamic Transshipment Network - An Algorithm and its Application to the Distribution of Empty Containers, IBM New York Scientific Center Report No. 320-296, Feb. 1969

17. J.A.Zachman: Business Systems Planning and Business Information Control Study: A Comparison, IBM Systems Journal 21, 1 (1982)

18. presented by H.-J.Zimmermann at the Hamburg OR Software Exhibition, 9th IFORS Congress, Hamburg 1981