CONF-970157--1
ANL/DIS/CP--91437

SIMULATION AND REAL-TIME OPTIMAL SCHEDULING: A FRAMEWORK FOR INTEGRATION

Charles M. Macal and Michael R. Nevins
Decision and Information Sciences Division
Argonne National Laboratory
Argonne, IL 60439
email: macal@anl.gov

Michael K. Williams and Joseph C. Joines
Military Traffic Management Command
Transportation Engineering Agency
Newport News, VA 23606

KEY WORDS: object-oriented simulation, military logistics, scheduling

ABSTRACT

Traditional scheduling and simulation models of the same system differ in several fundamental respects. These include the definition of a schedule, the existence of an objective function which orders schedules and indicates the performance of a given schedule according to specific criteria, and the level of fidelity at which the items are represented and processed through the system. This paper presents a conceptual, object-oriented, architecture for combining a traditional, high-level, scheduling system with a detailed, process-level, discrete-event simulation. A multi-echelon planning framework is established in the context of modeling end-to-end military deployments with the focus on detailed seaport operations.

1 INTRODUCTION

This paper outlines a conceptual framework for combining coarse-grained scheduling and fine-grained process simulations together into a dynamic, real-time context of military logistics planning. The scheduler operates at a macro-level of detail in space and time, and because of the broad scope of the scheduling problem, cannot well represent the systems modeled at the detailed process level, which is the province of simulation. The problem domain consists of scheduling the global movement of military units, equipment, supplies, and transportation assets over an extended time period of weeks or months. Simulations of this deployment process include seaports, army installations, and the transportation processes connecting them, at the hourly or minute time step. The integrated scheduling/simulation also addresses real-time data acquisition, real-time scheduling, and continuous replanning (see Figure 1), which are challenges for future logistics systems of all types (Mentzer and Firman, 1994).

Scheduling Background

A specific scheduling problem is described by four categories of information (Conway, et al, 1967): (1) the *items* and the *activities* the items must undergo, (2) the *resources* that are required to perform the activities, (3) *constraints* and *priorities* that restrict the assignment of items to activities, and (4) the *criteria* by which a schedule will be evaluated and the schedules compared. A typical simulation model contains these categories of information, but lacks an objective to be minimized or maximized. The notion that control variables exist which can be manipulated to produce alternative schedules, that is, not every activity or resource is fixed or constrained in its operation or utilization through time, is implicit in the scheduling problem but is not inherently obvious in the formulation of a simulation.

Schedule Criteria. Typical scheduling objectives for port operations include the following:

- Total lateness or tardiness of all items processed (minimize),
- Time the last item completes processing (minimize),
- Total cost of the operation (minimize),
- Utilization of port resources (maximize), and
- Inventory or staging area required (minimize).

More formally, let T_{Ai} be the *arrival time* for item i, and let D_i be the *required due date* for item i. Let Δp_{ij} be the *processing time*, the amount of time that will be required to perform the j^{th} activity on item i. Let ΔW_{ij} be the *waiting time* preceding the jth activity of item i. Given a consistent set of processing times, two schedules are the same if and only if they have the same set of waiting times (Conway, et al., 1967).

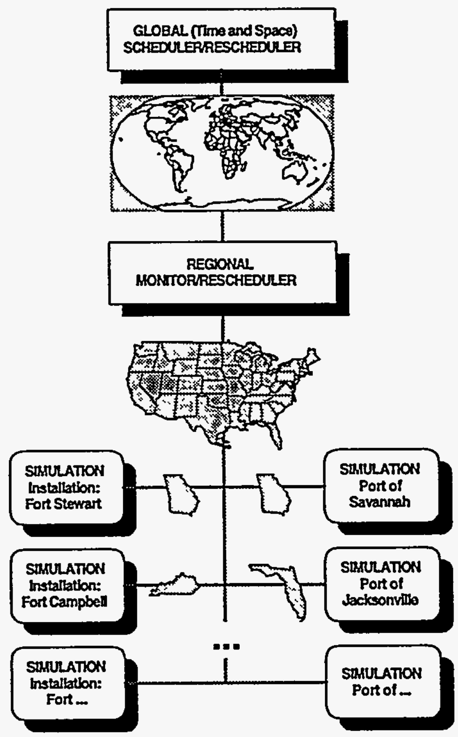

Figure 1 Multi-Echelon Nature of the Scheduling/Simulation Problem

Important measures derived from the ΔW_{ij} are: (1) times at which particular items leave the port, (2) length of time that particular items spend at the port, and (3) difference between the times when items leave the port and when they are supposed to leave. Theoretical work on scheduling has often employed simple measures of performance such as the average or maximum values over all items of flow time, completion time, lateness, or tardiness, where these measures are defined in terms of the above parameters: *total flow time* for item i is the total processing and wait times for the item in the j activities at the port: $F_i = \sum_j \Delta p_{ij} + \sum_j \Delta W_{ij}$; *completion time* for i is $C_i = T_{Ai} + F_i$; *lateness*, L_i, of i is $L_i = C_i - D_i$; and *tardiness*, T_i, of item i is $T_i = \max(0, L_i)$. Lateness for an item considers the difference between completion date and due date, regardless of the sign of the difference. Credit is given for being early under the lateness measure but not under the tardiness measure.

Simulated Schedules

Although simulations are often used for other purposes, a typical simulation produces information on all of the variables important to scheduling. The schedule produced by the simulation will be referred to as a *simulated schedule*. A simulated schedule is feasible in the sense that it is supportable by the processes represented in the simulation and the constraints imposed by the availability of the resources modeled. A simulated schedule may be one of many feasible schedules, and is not necessarily an optimal or even a good schedule. How close the simulated schedule is to an optimal schedule is an empirical question subject to investigation using the simulation model. Objectives, even if not formally optimized in simulation models, can still be used to measure the effectiveness of and to rank simulated schedules. The seaport simulation model, PORTSIM, illustrates a simulated schedule. Some background on PORTSIM is useful at this point.

Seaport Simulation (PORTSIM). PORTSIM is an object-oriented, discrete event simulation model of port operations focusing on the throughput of military forces (Nevins, et al., 1996). The key questions that PORTSIM addresses are (1) how long it will take to move specific forces through specific seaports under various assumptions on berth availability, ship availability, and resource constraints? and (2) what is the throughput capability of the particular seaport? PORTSIM includes detailed logic on processing items as they arrive at and move through seaports. The major processes modeled in PORTSIM consist of (1) reception and clearance of transportation items and unit equipment to and from the seaport, (2) staging of unit equipment at the port in preparation for loading/offloading, and (3) ship loading and offloading activities. Various port resources are required to support processing and movement through the port. The resources at the port for items arriving by highway include gates, open staging space, inspectors, drivers, berths, ships, cranes, roll-on/roll-off ramps, end ramps, and container handlers. For items arriving by rail, resources consist of docks, transit sheds, forklifts, interchange yard space, open staging spurs, apron spurs, rail end ramps, and locomotives. The simulation logic also includes constraints and priorities that restrict the assignment of items and resources to activities.

The limitations on port operations imposed by the availability of these resources are the key determinants of waiting time encountered by items being processed through the port. PORTSIM produces a simulated schedule (Figure 2) with the following measures:

- Flow Time for Item i = $\sum_j \Delta p_{ij} + \sum \Delta W_{ij} = 12 + 8$ units.
- Completion Time for Item i = $C_i = T_{iL} = T_{iA} + \sum_j \Delta p_{ij} + \sum_j \Delta W_{ij} = T_{iA} + 12 + 8$ units.
- Lateness of Item i = $L_i = C_i - D_i = (T_{iA} - D_i) + 20$ units
- Tardiness of Item i = $T_i = \max(0, L_i)$

These relationships imply that minimizing any of these measures is equivalent to minimizing the single term consisting of Wait Time = $\sum_j \Delta W_{ij}$. Wait time exists in the simulation due to a shortage of available resources for an activity to process an item. In principle, adding an unlimited amount of resources to the simulation will eliminate all wait time and result in an optimal schedule. Given that resources are limited, the scheduling problem with respect to the simulation consists of adding or shifting resources to minimize wait time.

2 SIMULATION FOR SCHEDULING

What capabilities does a simulation model need to have to be useful in the context of scheduling in a dynamic environment? The capabilities of traditional simulation models need to be expanded in order for a simulation to add value in the scheduling context. A simulation should have the capabilities to:

- Produce an optimal trajectory of the simulation state within the scope of the simulation's control variables,
- Update the simulation state based on dynamic, real-world data feeds and simulate from that point forward,
- Create alternate simulation scenarios and explore possible simulated futures, and
- Simulate to meet schedule requirements and understand what intermediate events must occur to meet required schedule goals (reverse planning).

Optimal Simulated Schedules

In the scheduling context a simulation must produce an optimal or near-optimal trajectory. The main questions are: (1) How well is the simulation performing relative to the optimal simulated schedule, if such a schedule even exists?, (2) What are the control variables in the simulation that can be manipulated to produce better schedules?, and (3) What are the constraints of the scheduling simulation that restrict attainment of the optimal schedule?

Control Variables. Control variables of the seaport simulation consist of the following: (1) ordering of items processed at the port, the queue ranking rules employed, and policies affecting port operations, such as whether activities in progress are preempted upon the arrival of a higher-priority item or they are allowed to continue to completion; (2) assignments of gates to staging areas, staging areas to berths, and routes connecting critical port elements; (3) matching of ships to berths; and (4) procedures for staging items (container stack height, etc.). Some factors considered as fixed parameters in the port simulation are

Port Activity for Item	Time / Duration	Resource(s) Used In Activity	Time: 0600 - 0700	Time: 0700 - 0800	Time: 0800 - 0900	Time: 0900 - 1000	Time: 1000 - 1100
Arrive at gate	T_{iA}	—	▲				
Wait at gate (queue)	ΔW_{i1}	—	□□				
Proceed through gate	Δp_{i1}	Gate	■■				
Transit to opening staging area	Δp_{i2}	—		■■			
Park in opening staging area	Δp_{i3}	Open staging space		■■			
Wait in open staging area (queue)	ΔW_{i2}	—			□□□□		
Inspect item in open staging area	Δp_{i4}	Inspector				■■	
Discharge item and transit to berth	Δp_{i5}	Driver				■■	
Wait in (roll-on/roll-off or container) call forward queue	ΔW_{i3}	—					□□
Load item onto ship at (RORO or container) berth	Δp_{i6}	Berth, ship, RO/RO ramp or crane					■■
Item loaded onto ship	T_{iL}						▲

■ Indicates time spent in an activity
□ Indicates time spent waiting for resource(s)
Note: Total Wait Time for Item i = $\sum_j \Delta W_{ij}$ = 8 units.

Figure 2 Simulated Port Schedule for Convoyed Vehicle Item Arriving via Commercial Highway

under the control of the scheduler, such as cranes, stevedores, terminal service battalions, and drivers. These resources could be shifted among ports, and make significant impacts on port performance. That some parameters may be fixed or variable depending on whether the perspective is that of the scheduler or the simulation leads to the notion of *hard and soft constraints* in the simulation.

Hard and Soft Constraints. A hard constraint is fixed under all circumstances of the simulation; a soft constraint could be relaxed in the simulation to investigate the impact of adding a resource, with the intention of communicating this information to the scheduler. For example, soft constraints consist of drivers, forklifts, and shifts; hard constraints consist of infrastructure elements such as berths, transit sheds, and gates. A new mode of running the simulation is needed to investigate the effects of relaxing the soft constraints. The key tasks are to identify the hard and soft constraints in the simulation processes and to structure the simulation control logic to operate the simulation in this "what-if" mode in which levels of available resources are varied.

Real-Time Simulation and Scheduling

Simulations in the dynamic scheduling context must deal with the present (real-world data feeds), the future (simulated time), and the past (simulated and real-world history). To address this requirement, the notion of *multiple worlds* is introduced.

Multiple Worlds. There are three worlds encompassed by the scheduling/simulation problem: (1) a simulated world, contingent on a state and a process representation that can project that state forward in time, (2) the real world, reflected in the data feeds that provide data on the locations and status of items in the port simulation at any time t, and (3) alternative future worlds that are contingent on external, uncontrollable, events and decisions not under the control of the simulation.

We make several assumptions regarding these worlds for the sake of making the problem tractable: (1) process representation in the simulated world is imperfect, and some discrepancies between the simulated and real worlds are going to occur, (2) real-world data feeds are error-free, accurate, representations of the real world and the information is known with complete certainty when it becomes available (the feeds express ground truth and are utilized instead of the simulation outputs if the simulated world is in conflict with the real world), (3) real-world data feeds are event-driven rather than reported in continuous time, and (4) there exist some finite, and as a practical matter, small, number of alternative future worlds that can be identified in terms of the key variables affecting the evolution of the real world toward these alternative futures. Of these assumptions, 2 is of the greatest concern, for to consider possibly erroneous data feeds would necessitate that ground truth not exist as a frame of reference.

Replanning in Simulation. The scheduling context requires the simulation to dynamically replan the simulated schedule as unanticipated, and unmodeled, events occur in the real world. The feeds may provide information which is in conflict with the locations and status of items as projected by the simulation. The simulation state must be reconciled with these real-world data, and the system must be resimulated from the updated state from that time forward. This process is termed *replanning*. We assume the data feeds take the following form: (1) the times at which an item begins or ends an activity, for example, an item enters the staging area or completes an inspection, and (2) the general status of an item, for example, "at time t item is in staging area waiting for inspection"

In the simulation, an item at time t is either (1) engaged in an activity for which there is a scheduled event time for completion of that activity in the simulation event queue, or (2) the item is waiting in an activity queue for resources to become available to begin the activity. The status of each item for which there is real-world data at time t must be checked in the simulation. Having found the item in the simulation, the item's attributes must be compared to the real-world data. For example, an item attribute could be "inspection status" with values "inspected" or "not inspected." If the simulated world is in conflict with the real world at time t, one or more of the following actions must be taken to recalibrate the simulated state to be consistent with the real-world data: (1) move items into or out of activity queues, (2) add, delete, or modify scheduled events in the simulation event queue, (3) change item attributes, (4) add or delete available resources, (5) add or delete resource assignments to activities affecting the item, and (6) modify the simulation history (that is, simulated events occurring before time t) to maintain consistency with the state at t. This last action is necessary only if it is important to maintain a simulation history that is totally consistent with real-world events. Once the simulated world has been made consistent with the real world at time t, the system is resimulated from time t forward to produce a new trajectory based on the updated state (see Figure 3). This is termed *exploratory simulation.*

Exploratory Simulation. Ideally one would have a set of simulation runs that would be able to answer any question that a decision-maker could ask about the effect of any possible event or any decision on the schedule, and this entire set of runs would be updated as the system state is updated according to real-world data. The basis

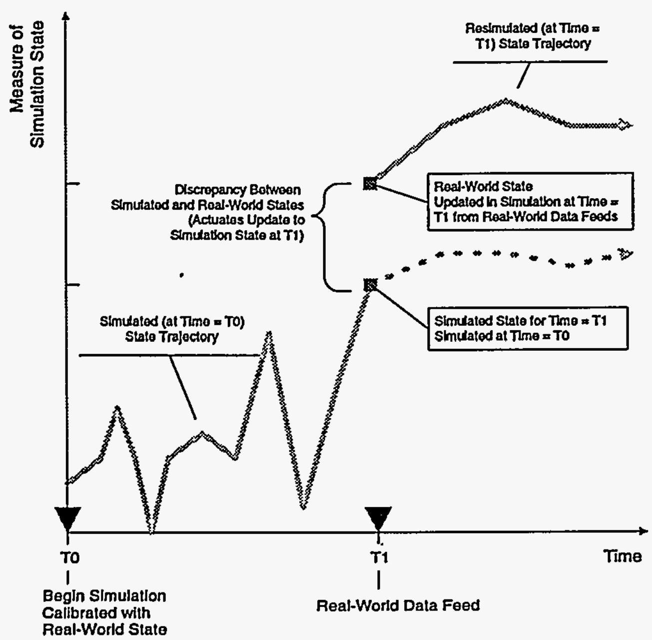

Figure 3 Replanning Initiated by Periodic Real-World Data

for these runs is as follows: (1) varying random number seeds that are the basis for the stochastic variables (for example, processing times and discharge times) in the port simulation, to understand the extent that variability inherent in the port simulation processes has on the overall port scheduling objective, (2) optimize over all port operational policies, to identify the best port schedule with respect to the variables under the control of the port, (3) sensitivity analysis, to understand the effects of changes in all model parameter values on the port schedule, and (4) exploration of the effects of key variables, such as increased or reduced resource levels. This information is used to communicate the effects of changes in the scheduler's control variables on port operations to the scheduler.

Reverse Simulation for Requirements Scheduling

The planning process sometimes begins with a requirement for items to be delivered to final destinations by specified times, as in required delivery dates for units to in-theater destinations as specified in operational plans. The scheduling problem then becomes one of determining the times by which intermediate steps in the deployment process must be completed for the operation to meet the final delivery date. This requirement is referred to as *reverse planning*. From the simulation perspective, the distinction between reverse and forward planning is important and requires a different approach from that adopted in the forward planning context. Some options for simulation to address reverse planning are:

1. Run the simulation backwards through time with the port processes restructured to operate in reverse time.

2. Develop a heuristic procedure to estimate start times given required completion times. This in effect would be a simplification of the detailed process simulation.

3. Take a schedule produced by a forward simulation run and shift the time line of that schedule so that the time for the completion of the simulation corresponds to the required time for the completion of the operations.

The problem with the reverse simulation approach is that the processes are not reversible in time. The process simulation would have to be rewritten with the same activities but with different process logic and decisions embedded in the processes. The second option is undesirable because it ignores the rich set of information on the port operations processes that is already embedded in the simulation model. The third option would be the most straightforward to implement. It also requires that a forward simulation be run and that the objective be the shortest time from beginning to end of the port operations, in effect, the most compressed time frame possible.

3 OBJECT-ORIENTED DESIGN FOR AN INTEGRATED SIMULATION/SCHEDULER

Research on simulation applications for real-time scheduling in manufacturing has recently appeared (see Coskun and Oren (1995) and Mebarki, Pierreval and Dussauchoy (1995)). Combining simulation and scheduling technologies leads to several technical issues that must be addressed by an integration framework. These include consistent aggregation/disaggregation of item and process representation, protocols for interaction between the simulation and the scheduling systems (for example, iteration schemes, convergence metrics), detection of differences between simulated schedules and real-world data on system state, resolution of discrepancies between simulated and real-world states, and model control schemes for replanning the simulation and the scheduler.

The simulation/scheduling problem is built up from a set of subproblems which naturally leads to an object-oriented framework, with objects representing the major components of the system and instructions that need to be passed among processes as the basis for defining object methods. An object-oriented structure lends itself to satisfying the requirements and interactions among the objects.

Framework for Integration

In the enhanced port simulation design, movement objects are represented hierarchically at different degrees of fidelity, with embedded methods checking the consistency of object attribute aggregation at all levels. Software agents monitor the status of the simulation and determine relevant information that should be passed between levels of the object hierarchy for purposes of reasoning about system status and for determining when communication between the scheduler and simulation is appropriate. We combine agents and monitors with objects in this design (see Figure 4). Major system components are described.

Master Scheduler: The master scheduler solves the aggregate-level scheduling problem over time and space. It determines a feasible set of dates for departures and arrivals of units, equipment, supplies, and assets, and in doing so globally allocates transportation and other resources to support the schedule. In reality, the master scheduler may be a collection of schedulers that operate at various levels of detail (scheduling global strategic movements, scheduling regional tactical movements) and that decompose the scheduling problem into parts such as modes (ship scheduler, air movement scheduler).

Resource Allocation and Schedule Object: This object stores data on the spatial and temporal allocation of global assets and resources as determined by the scheduler.

Resources Agent: This agent reasons over the spatial and temporal allocation of global assets and local resources. The agent updates the time line object with time-dependent resource data pertinent to the simulation.

Time Line Object: The time line object is the repository for time-subscripted information at all levels of detail, both at the high-level of detail that relates directly to aggregate date specifications in the plan and at the low level of detail that relates to individual activities at the ports and installations, produced by the scheduler. Besides identifying events and event times, the time line object also contains the relationships among the event times (see Figure 5 for example).

Time Line Agent: The time line agent reasons over the information that is posted to the time line object, the dates along the time line which

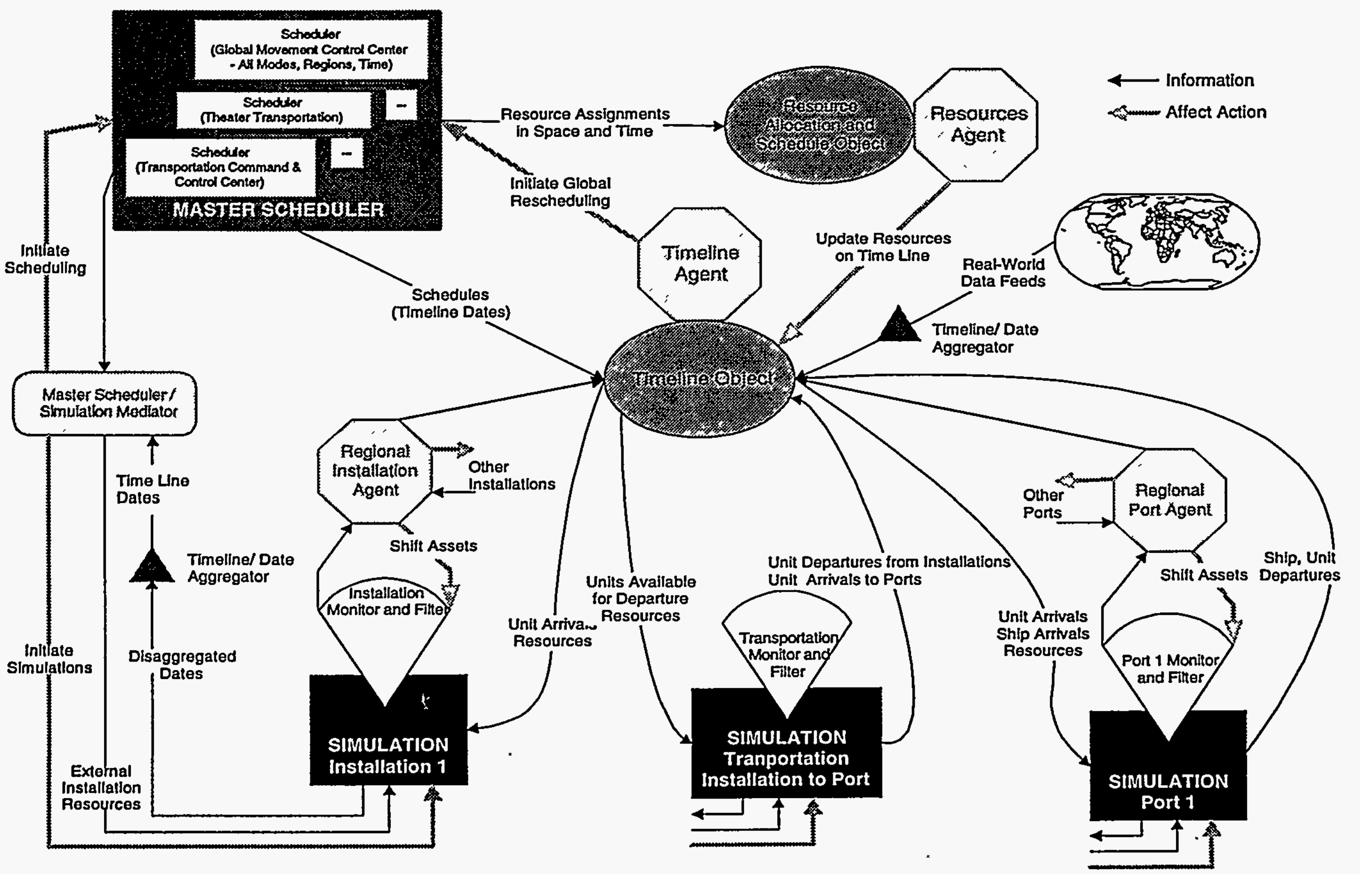

Figure 4 Integrated Simulation and Scheduling System

are posted by the scheduler, the simulations, and the real-world data feeds. The time line object monitors and identifies lateness conditions and initiates replanning of the master scheduler and the simulations.

Regional Port Agent: This agent allocates port-specific resources among ports (generally within a regional area that makes this action feasible) as shortages or surpluses of resources occur during port operations or that are foreseen to occur in future simulated time.

Port Monitor and Filter: These procedures are built into the port simulation code to continuously monitor conditions in the port simulation as it is running. Resource utilization levels at the port that cross a threshold are identified and communicated to the scheduler.

Port Simulation: This is the seaport simulation with the enhanced capabilities discussed in Section 2.

Master Scheduler / Simulation Mediator: This mediator coordinates the operation of the master scheduler and the simulations to ensure consistency between the results of these systems, ensuring that the schedulers and the simulations are producing consistent dates and resource allocations for future time periods.

Time Line/ Date Aggregator: The aggregator takes the detailed event times from the simulations and derives the dates implied by these event times that are consistent with the dates required by and produced by the schedulers so that a direct comparison between the simulation results and the scheduler results can be made. (The same functionality in the Time Line/ Date Aggregator is also embedded in the Time Line Agent.)

Modes of Simulation and Scheduler Interaction

There are two possible modes of interaction between the scheduler and the simulation that have been identified to achieve consistent schedules and plans: (1) using the simulation to determine coarse parameter values, such as port throughput capability that could be used directly in the scheduler, and (2) using the simulation to determine dates at which items will clear the port based on the arrival times from the scheduler and using the scheduler to determine dates that items will arrive at the port based on port processing times from the simulation. The first scheme is a one-way information flow from the simulation to the scheduler, with an aggregation step in-between. The second scheme would require a complex iteration scheme to achieve total consistency between the scheduler and the simulation.

CONCLUSIONS

The simulation and scheduling framework described here is extensible for coarse-grained schedulers used in conjunction with fine-grained process simulations. Computational performance issues are of concern because of the computational complexity of the scheduling problems, even at the coarse-grained level, and the need for exploratory simulations activated in response to real-world data feeds. Distributed, concurrent computing would be a natural area of investigation because of the weak, feed-forward linkages that exist at many points in the integrated simulation/scheduling system. This framework will be further developed as a basis for combining simulation with scheduling as part of the Advanced Logistics Program, of the Defense Advanced Research Project Agency (DARPA).

ACKNOWLEDGMENTS

This work was supported under military interdepartmental purchase request from U.S. Department of Defense, Military Traffic Management Command, Transportation Engineering Agency and the Defense Advanced Research Projects Agency, through U.S. Department of Energy contract W-31-109-Eng-38.

REFERENCES

Conway, R., W. Maxwell, and L. Miller. 1967. *Theory of Scheduling*, Addison-Wesley, Reading, MA.

Coskun, R. and T. Oren. 1995. "Architecture of a Real-Time, Simulation-Based and Intelligent Production Scheduling System," in *Proceedings of the 1995 Summer Computer Simulation Conference*, 367-372.

Mebarki, N., H. Pierreval and A. Dussauchoy. 1995. "Object Oriented Simulation-Optimization Software Devoted to the Real Time Scheduling of Manufacturing Systems," in *Proceedings of the 1995 Summer Computer Simulation Conference*, 585-590.

Mentzer, J. and J. Firman. 1994. "Logistics Control Systems in the 21st Century," *Journal of Business Logistics*, 15(1):215-227.

Nevins, M., C. Macal, J. Joines. 1996. "A Discrete Event Simulation Model for Seaport Operations," submitted to *Simulation*.

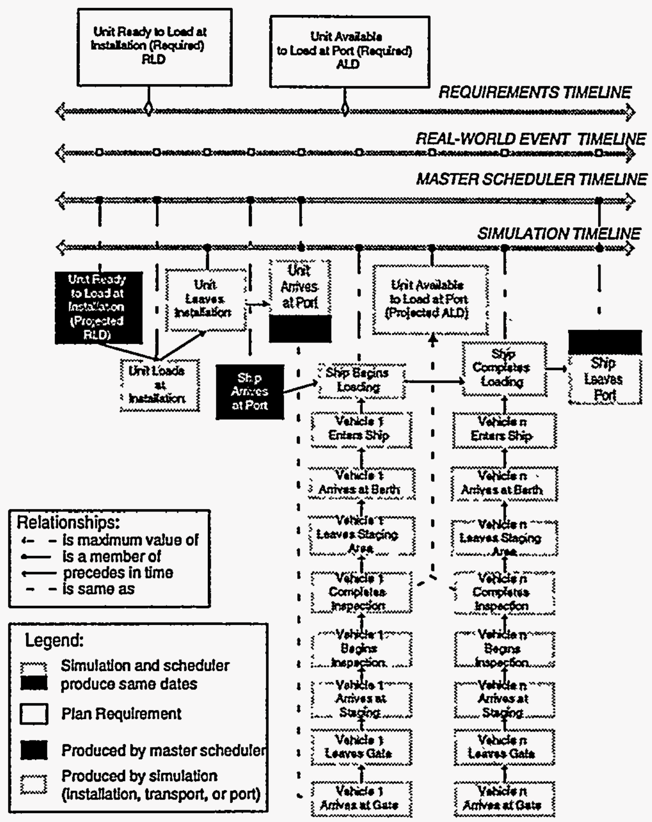

Figure 5 Port-Related Events and Relationships