Alternative Architectures for Distributed Cooperative Problem-Solving in the National Airspace System

Philip J. Smith*, Charles Billings*, C. Elaine McCoy, ** and Judith Orasanu***

* Cognitive Systems Engineering Laboratory Institute for Ergonomics The Ohio State University Columbus OH

** Department of Aviation Ohio University Athens OH

*** NASA Ames Research Center Moffett Field CA

June 24, 1999
Table of Contents

Acknowledgements 3
Executive Summary 4
Abstract 7

Introduction 8
Alternative Architectures for Distributed Problem-Solving 9
A Model of Aviation System Management Interactions 10
A Problem-Driven Approach to Studying Alternative Architectures 10
Overview of Paper 11

Case Study 1: The Locus of Control vs. the Locus of Data 13
   Introductions: Task decomposition 13
   Details of the Scenario 13
   Important Features Illustrated by the Scenario 15
   Implications for System Design 16

Case Study 2: The Locus of Control vs. the Locus of Knowledge 18
   Introduction 18
   Alternative System Architectures 18
   Details of the Scenarios 19
   Follow-on Study 1: Predicted vs. Actual Fuel Consumption 22
   Follow-on Study 2: A Detailed Observational Study of LAX-DFW Flights 23
   Case Study 2: Discussion and Conclusions 24
      Alternative Solutions 25
      Dissemination of Knowledge from Traffic Managers to AOCs 26
      Allowing Traffic Managers to Play the Role of a Neutral Referee 30

Case Study 3: Use of a Referee for Resource Allocation 31
   Potential Solutions 32

Case Study 4: Use of a Neutral Resource Broker 33

Case Study 5: Shifting the Locus of Control & Knowledge to Match the Locus of Data 34

Overall Summary 36

References 38
Executive Summary

This paper deals with human performance in a distributed work environment. Specifically, it explores the impact of alternative system architectures on performance in the National Aviation System (NAS). These alternative architectures are discussed in terms of their impact on the performances of the people working within the system, and are defined in terms of the way the system is decomposed into subtasks, and in terms of the locus of control and access to relevant knowledge and data.

The fundamental premise is that the operation of the NAS requires some type of task decomposition in order to avoid excessive cognitive complexity for any one person. To deal with the realities of human limitations, such a decomposition strategy typically is designed to produce "good" rather than "optimal" performance, and is based on an independence assumption: If each person or subsystem performs its independent task well, the combined effects on overall system performance will be good. Since few systems are fully decomposable into a set of completely independent subtasks (and since the NAS is no exception to this limitation), there is also an assumption that, in those cases where interactions among subcomponents or people are necessary in order to ensure acceptable performance, they will either be mandated procedurally for these exceptions or will occur on an ad hoc basis at the initiative of some system participant. In this paper, five case studies are presented to explore these issues of system decomposition in the NAS.

Case Study 1 describes a system failure that arises because a flight crew and controller make a tactical decision that has adverse strategic implications. This situation arises because of a system architecture in which the locus of control (the flight crew and controller) does not match the locus of relevant data (the dispatcher). Two complementary solutions are suggested: Shifting the locus of data (providing flight crews with more global displays of weather) and designing intelligent agents (technology) to alert dispatchers when some "significant" change occurs in a flight, thus helping to ensure that their attention is focused on this change.

Three cautions are raised regarding such solutions, however. First, if the strategy of providing the person with control (the pilot in this case) with direct access to all relevant data is carried to its extreme for all possible scenarios, that person will be confronted with a task that has far too much complexity. Thus, we must carefully pick and choose the cases where this solution is applied. Second, if the flight crew is given more global weather data, they may be less inclined to contact dispatch, thus reducing the effectiveness of the safety net provided by the redundant evaluation of a flight by both pilots and dispatchers.

A third caution is that, if a technological "intelligent "agent" is provided to help ensure that important interactions occur between the flight crew and their dispatcher when a "significant" change occurs, so that the flight crew has the benefit of the data and knowledge available to the dispatcher, overreliance may result. The dispatcher may become too reliant on the technology and as a result, if the designers haven't anticipated or correctly dealt with all possible situations that can arise, a critical interaction may not occur. (This is especially of concern for situations where it is the lack of a response or action by the flight crew that is the critical event to be detected.)
Case Study 1 thus provides cautions for system designers to consider, rather than answers. Alternative solutions must be evaluated on a case-by-case basis in terms of the existing tradeoffs. Case Study 1 does suggest, however, that implementation of several complementary solutions, thus providing a kind of redundancy, may be a desirable design approach.

**Case Study 2** focuses on the locus of control in pre-flight planning. Over the past several years, the FAA has moved first from a system architecture based on management by control to one based on management by permission, and then to one based on management by exception. Each of these architectures has its own potential strengths and weaknesses. Management by permission under the original coordinated National Route Program, for instance, left the locus of control with FAA traffic managers, but induced interactions between traffic managers and Airline Operations Centers (AOCs) so that airline constraints and priorities were considered in the decision making, and so that knowledge about traffic constraints was disseminated from FAA traffic managers to airline dispatchers.

Management by exception, under the more recent North American Route Program, has resulted in some significant benefits to the airlines, but these are less than they could have been because the new locus of control (dispatch) has not been provided either with direct access to important, relevant knowledge and data about air traffic constraints or with indirect access though interactions with FAA traffic managers. In Case Study 2, it is suggested that the solution to this problem for pre-flight planning is not to revert back to a control by permission paradigm. Rather, there is a need to find ways to integrate the strengths of the management by permission paradigm into a management by exception paradigm. It is suggested that this can be achieved through the use of support technologies like the Post-Operations Evaluation Tool and the Collaborative Slide Annotation Tool to disseminate knowledge about air traffic constraints to AOCs and knowledge about airline constraints and priorities to traffic managers. Effective use of this knowledge may, however, require new technological supports for flight planning by dispatchers.

**Case Study 3** cautions that, while management by exception can be very beneficial if knowledge and data is distributed appropriately, in a competitive environment there are clearly cases where a neutral referee is needed in order to assure safe, efficient coordination of air traffic flows. Thus, Case Study 2 is a response to the request by airline AOCs to “give us the data and knowledge we need and let us try to solve the problems ourselves before having FAA traffic managers intervene.” Case Study 3 is a reminder that, given the airlines are in competition with each other, there will be situations in which “management by directive” by a neutral referee will still be necessary. In short, Case Study 3 suggests that we need to consider a hybrid environment in which, to deal with concerns with safety, overall system efficiency and equitable treatment across system users, different architectures must be interwoven to get the best of each for appropriate situations.

Case Study 3 also points out that the details of the implementation of “management by directive” by a referee are important. As an illustration, the principle of control at the least restrictive level possible (thus giving the airlines as much flexibility as possible) is discussed in the context of slot-substitution when there is a restricted arrival rate for an airport. By allowing each airline to determine which of its flights to use in filling a limited number of arrival slots (when FAA traffic managers have determined that the arrival rate for an airport needs to be restricted due to weather
or some other problem) both system capacity constraints and airline business concerns are accommodated.

Case Study 4 further explores this theme that a neutral referee is sometimes needed. In Case Study 4, though, the emphasis is on avoiding unnecessary waste of resources (such as unused arrival slots at an airport). The enhanced Ground Delay Program, developed cooperatively by FAA and airline staff as part of the Collaborative Decision Making program, is used to illustrate this point.

Case Study 5 ends the paper by suggesting that, although there are situations where a neutral referee is needed to deal with the competitive interests of the airlines, there are also many scenarios where FAA traffic managers or controllers are in a position to actively help an airline, because they have better access to the relevant real-time data or knowledge needed to deal effectively with the situation. Thus, Case Study 5 describes an architecture where the locus of control is shifted from the AOC to traffic managers or controllers, along with the communication of information about the relevant airline constraints and priorities.

Recommendations

The point of this paper is simple: In designing the architectures for different subsystems or functions in the NAS, we need to take a realistic view of human performance in an environment that at times is cooperative and at times is competitive. In part this means looking for those places where alternative control strategies such as management by directive, by permission or by exception are most effective, and carefully selecting the parameters of control used in the implementations of these strategies. In part this means that we must match the locus of control with effective access to the relevant knowledge and data, while considering the impact of the resultant task allocations on cognitive complexity. In part it means that we need effective ways to trigger and support interactions when the relevant knowledge is distributed across several people and organizations. And in part it means that we must be cognizant of the more detailed aspects of human performance that can result in errors, such as slips and mistakes, overreliance, cognitive biases, and the development of incorrect mental models and faulty assumptions and, as a result of these realities of human performance, we must include necessary redundancies in order to provide safety nets.
Abstract

The air traffic management system in the United States is an example of a distributed problem-solving system (Davis and Smith 1983; Durfee, Lesser and Corkill, 1989; Fleishman, and Zacarro, 1992; Layton, Smith and McCoy 1994; Orasanu and Salas 1993; Rasmussen, Brehmen, and Leplat 1991; Robertson, Zachery, and Black, 1990). It has elements of both cooperative and competitive problem-solving. This system includes complex organizations such as Airline Operations Centers (AOCs), the FAA Air Traffic Control Systems Command Center (ATCSCC), and traffic management units (TMUs) at enroute centers and TRACONs, all of which have a major focus on strategic decision-making. It also includes individuals concerned more with tactical decisions (such as air traffic controllers and pilots).

The architecture for this system has evolved over time to rely heavily on the distribution of tasks and control authority in order to keep cognitive complexity manageable for any one individual operator, and to provide redundancy (both human and technological) to serve as a safety net to catch the slips or mistakes that any one person or entity might make. Currently, major changes are being considered for this architecture, especially with respect to the locus of control, in an effort to improve efficiency and safety. This paper uses a series of case studies to help evaluate some of these changes from the perspective of system complexity, and to point out possible alternative approaches that might be taken to improve system performance. The paper illustrates the need to maintain a clear understanding of what is required to assure a high level of performance when alternative system architectures and decompositions are developed.
Introduction

Most complex system designs rely upon simplifications that allow the system to perform adequately, without trying to determine and implement "optimal" solutions. One common approach is to decompose the task of managing the overall system into subtasks, and to then assign those subtasks to separate individuals. The hope is that there is sufficient independence among these subtasks so that when each subtask alone is performed well, the combined effects produce good performance for the system as a whole. Furthermore, because few systems are actually decomposable into fully independent subtasks, it is also hoped that the operators responsible for particular subtasks will interact with one another as needed either because this interaction is procedurally mandated or because they decide that it is necessary to do so on an ad hoc basis in order to find an acceptable solution.

Alternative Architectures for Distributed Problem-Solving

This report uses a number of case studies from the current Air Traffic Management (ATM) system to discuss alternative system architectures, where an "architecture" is differentiated in terms of where control, knowledge and data reside within the system in order to support successful completion of particular tasks. To illustrate such alternative architectures, these case studies look at different subsystems within the ATM system (Kerns, Smith, McCoy and Orasanu, 1999). These examples are then used to discuss the strengths and weaknesses of each different architecture in dealing with specific types of situations.

The primary theme behind this series of analyses is that good decision-making requires effective access to appropriate knowledge and data by the person in control, as well as an appropriate strategy for distributing control. One of the principal factors determining whether a particular architecture is effective is the cognitive complexity of the task to be performed. If a task requires greater knowledge than one person can reasonably accumulate or integrate in order to perform that task, or if the task requires access to more data than that person can attend to or process effectively, then the work must somehow be distributed. This distribution may involve off-loading information processing, or it may take the form of decomposing some global decision into a set of sub-decisions that can be distributed to more than one person.

A second factor is the time stress associated with completing the task. Generally speaking, coordination and communication among people consume time. Thus, if some task or decision must be completed rapidly, and if a particular distribution of the work to complete that task or to arrive at a decision requires interactions that consume too much time or attention, then that architecture may not be acceptable.

A third factor is the mental model that each individual develops about the other participants in the system (Orasanu, 1991). This mental model influences what assumptions are made regarding "what the other guy is doing." This in turn affects decisions about when it is necessary to interact with "the other guy."

From a system design perspective, concerns over cognitive complexity, time stress and mental models suggest that at least the following questions must be addressed:
1. Is the complexity of a task sufficiently simple so that a single unaided person can be competent to complete it? (Note that the answer to this question requires a realistic view of the availability of resources to select and train a person to the required level of competence, and to retain that person for the job.)

2. Even if this single unaided person is competent, is he or she susceptible to slips? Or is the task being performed inherently susceptible to slips?

3. If the pool of individuals competent to perform this task is too small, can some of the work be assigned to designers and implementers (builders of technologies) who are competent to develop effective cognitive tools to aid human performance?

4. Even if they are competent, what slips are these designers and implementers likely to make?

5. As an alternative or complement to teamwork between an operator and a designer (as mediated by some piece of technology), can some of the work be distributed for other people to perform directly (rather than through some piece of technology that they have developed)?

6. If the work is distributed, how should control, knowledge and data be distributed in order to reduce the cognitive complexity for any one individual, provide "safety nets" to catch slips, mistakes or omissions, and achieve a uniformly desirable level of overall system performance?

7. If the work is distributed, has it been distributed in such a way that the necessary interactions among participants do in fact occur, but also do not interfere with the completion of tasks in a timely fashion?

**A Model of Aviation System Management Interactions**

In order to determine the architectures that might support problem-solving in this context, or to examine task decompositions in this setting, it is necessary to examine the agents that participate in ATM system management and control, and to lay out the relationships among them. In terms of the focus of this paper, six groups of people within airline and FAA organizations will be discussed.

From airlines these groups are:

- Pilots and other flight crew members, responsible for individual flights;
- Dispatchers (within Airline Operation Centers), who are jointly responsible with flight crews for flight safety (Airline Dispatchers Federation and Seagull Technology, Inc., 1995); Lacher and Klein, 1993)
- Airline Operations Centers (AOC) personnel responsible for air traffic control liaison on a daily operational level. This role is typically assigned an Air Traffic Control Coordinator (ATCC) or Chief Dispatcher within an AOC.

From the Federal Aviation Administration (FAA) Air Traffic Services:
• Air Traffic Control System Command Center (ATCSCC) personnel, who have primary responsibility for ensuring a safe and efficient flow of air traffic through the system. They have primary responsibility for national strategic air traffic management;
• Traffic Management Unit (TMU) personnel at enroute Air Route Traffic Control Centers (ARTCCs) and TRACONs;
• Air Traffic Controllers (ATC) in these enroute Centers and in Terminal Area Traffic Control and Airport Control facilities. These people are responsible for tactical air traffic control.

Until recently, air traffic management and control has been a generally hierarchical system involving the human operators listed above, aided by a variety of types of automation, most of which have been oriented toward transforming and moving data within the system. The aviation system is very broadly distributed, having command, coordination and communication nodes throughout the nation, and of course many aircraft that are also distributed throughout the national airspace. The classical model of the relationships among the persons and technologies involved in the management of aviation operations may be thought of as shown in Figure 1.

![Figure 1. A concept of the traditional air traffic management operational process.](image)

Recent developments in the air traffic management system and its policies have resulted in many changes in the strictly hierarchical relationships shown at the left in Figure 1 (Smith, Billings, et al., 1997). These developments, and their effects (both desired and unwanted) are the topics considered in this paper. We have used the construct shown here to illustrate some of the changes that have taken place.

A Problem-Driven Approach to Studying Alternative Architectures

To gain insights into the answers to the questions posed above, we have examined a collection of case studies representing different types of situations that have arisen in the ATM system (Wickens, Mavor and McGee, 1997). In each of these case studies, concerns about distributed problem-solving are identified. Then alternative approaches for overcoming these problems are presented and evaluated.
All of the alternatives considered here revolve around the classical approach to dealing with cognitive complexity: Decompose the overall task into a set of subtasks that are nearly independent, meaning that if each subtask is completed well by itself, then overall system performance will usually be good. Then add to the system some means for ensuring interactions among system components for those exceptional situations in which independent performance is not sufficient.

Abstractly, these alternative approaches for improving existing performance fall into several categories:

1. Decompose the task into a different set of subtasks, redefining the way responsibility and control is allocated to different individuals;

2. Shift the locus of knowledge and data to match the current locus of control (requiring the person with responsibility and control to personally have the necessary knowledge and direct access to the data);

3. Shift control to match the current locus of knowledge and data;

4. Leave the knowledge and data necessary for a particular decision distributed among several people, but ensure that effective interactions occur among these people when one of them needs certain knowledge or data from someone else in order to make an informed decision;

5. Provide automation aids for less critical subtasks, to facilitate concentration on the more critical sub-tasks. This may involve a redesign and simplification of the task architecture.

6. Hybrid solutions involving a combination of these alternatives.

Overview of Paper

In this paper, a number of case studies are presented and analyzed to understand the impact of alternative architectures on performance.

Case Study 1 involves a scenario dealing with problems that arose when a tactical decision was made by a flight crew and a controller, when neither of them had the necessary data to make an appropriate decision. In that scenario, a serious safety hazard arose because the people making the decision did not access the relevant data when deciding how to act. Two complementary solutions are suggested, one shifting the locus of data to give the flight crew direct access to additional data, the other using an intelligent alerting function to help ensure that the person with control (the pilot) interacts with the person who has access to the relevant data (the dispatcher).

Case Study 2 presents a scenario that is similar at a conceptual level, in that there is a mismatch between the locus of control and the locus of knowledge and data. Case Study 2 deals with strategic decisions, though, and the result of the mismatch is a decrease in efficiency rather than a potential safety concern. Partial solutions to this problem with strategic decisions include:
1. Providing better post-operations evaluation tools to determine where strategic plans have not been working, and then supporting synchronous or asynchronous information exchange between AOCs and traffic managers to discuss these problems. In this way, they can share their knowledge about what has been causing inefficiencies and how to reduce them (thus disseminating knowledge from the traffic managers to dispatchers and vice versa);

2. For relatively static constraints, providing access to constraints that are known to traffic managers so that dispatchers have access to this knowledge (and vice versa);

3. For dynamic situations, providing tools to allow more efficient and effective real-time collaboration between traffic managers and AOCs;

4. In those situations known to regularly produce inefficiencies due to air traffic constraints, giving the traffic manager the authority to help the airline with a flight by assessing the situation as it develops and choosing (with the concurrence of the flight crew) from a set of alternative options that the dispatcher has pre-approved for that flight;

5. Finding some way to increase capacity so that there is no constraint to deal with.

Thus, Case Study 2 emphasizes another important design concept: It is not always necessary to support real time exchange of knowledge or data in order to improve the performance of the person in control. In some cases, because of workload constraints in the real-time environment, a more effective (or complementary) solution may be to provide such information after the fact, in order to disseminate the relevant knowledge and improve performance in similar future situations. It also emphasizes that we need to look for long term solutions that increase capacity, thus reducing the constraints that need to be dealt with.

Case Study 3 involves the issue of resource constraints and how they may be fairly and impartially allocated in real time. In particular, it focuses on the reality that, in a competitive environment there are situations where a neutral referee is needed.

Case Study 4 deals with the issue of how to minimize the wasting of resources in a competitive environment. The FAA's enhanced Ground Delay Program is cited as an example of how to approach this.

Case Study 5 further explores the issues surrounding control and access to knowledge and data. The scenario discussed in Case Study 5, however, looks at another type of solution: Instead of shifting data to the person who has control in the current architecture, shift control to the person who already has access to the relevant data. The differentiating feature supporting this alternative solution or architecture is the timing involved. In this scenario, because the decision must be made close to the scheduled departure time for a flight, and because it isn't known until that time which flight(s) will be involved, it may make more sense to empower the traffic manager with the authority to make the decision (having informed the traffic manager of any relevant airline constraints and priorities before that point in time).

Each of these case studies is considered in detail in the following sections.
Case Study 1: The Locus of Control vs. the Locus of Data

Introduction: Task decomposition

As a specific example of task decomposition in the current National Airspace System (NAS), consider the following scenario. In order to reduce cognitive complexity, the overall task of selecting safe routes of flight and of operating these flights is currently decomposed in such a way that each of the participants (the flight crew, controllers, dispatchers, traffic managers, etc.) has only partial information. In particular, within the current ATM system, tactical decisions are made by flight crews and controllers without always having the information necessary to develop the same "big picture" about weather system developments that is available to dispatchers and traffic managers.

Because this distribution of data is not always adequate, controllers sometimes request reroutes for flights which do not have sufficient fuel for the proposed reroute. Similarly, flight crews sometimes fail to contact their dispatchers in situations where it would have been helpful to do so, because they have made incorrect assumptions about the importance of the knowledge or data which is available to the dispatcher. In short, although the current distribution of information and responsibilities generally permits an efficient and safe operation, it is susceptible to occasional errors due to false assumptions about "what the other person has already considered", or due to incorrect assessments of whether a particular change in route is "significant" enough to require interactions with someone who has the "bigger picture". This issue is illustrated strikingly in this case study.

Details of the Scenario

An incident arose involving a Boeing 727-200 flying from Dallas/Ft. Worth to Miami (Smith, Caisse, et al., 1998). As the airplane was traversing the Florida panhandle, there was a line of thunderstorms from the Tampa Bay area southeastward down to the Miami/Ft. Lauderdale area. The airline dispatcher, who was required to provide the pilot in command with information regarding any hazardous enroute weather, noted a line of thunderstorms that he felt potentially jeopardized the safety of the flight. He contacted the captain, briefed him on the enroute weather conditions, and recommended a revised route taking the aircraft direct to Ormond Beach and then down the east coast of Florida into the Miami airport from the northeast, ahead of the weather. The Captain concurred with the reroute and contacted the Jacksonville Air Route Traffic Control Center frequency to coordinate the reroute. The reroute was approved. The aircraft made a turn to the east and was proceeding direct to Ormond Beach on the Florida East Coast (Figure 2).
Figure 2. ASD display for Case Study 1
When the airplane was handed off from the controlling Center sector to the next sector, the receiving sector advised the Captain that, because of heavy traffic along the east coast of Florida, they would not be able to accommodate the reroute and that the aircraft would have to return to its originally filed route of flight along the west coast of Florida. The aircraft made a fairly abrupt turn back to the southwest, got offshore along the west coast of Florida and proceeded south toward the Ft. Myers area. Furthermore, the aircraft was slowed to 180 knots due to traffic, increasing its fuel burn.

At that time, the line of thunderstorms was sinking to the southeast, moving down toward Miami/Ft. Lauderdale/Sarasota/Ft. Myers. As the aircraft arrived in that vicinity and was preparing to turn to the east for its approach to Miami, landing to the east, the weather came across the Miami airport and shut down Miami's operations. As a result, the aircraft entered airborne holding and was given "expect further clearance" times from ATC that continued into the future. Thus, the crew was faced with uncertainty as to when the weather would clear and they would be released to proceed into Miami.

It was not until this time that the Captain contacted the aircraft's dispatcher and advised the dispatcher that the reroute they had agreed upon had been refused by an ATC sector, that the aircraft had ended up back on its original filed route of flight, and that they had encountered airborne holding. The dispatcher's attention had been diverted to another situation while this was happening and he had not noted the ATC-initiated reroute. Thus, at that point the aircraft was holding with thunderstorms between its position and the intended destination.

What complicated this scenario was that Sarasota, Ft. Myers, Ft. Lauderdale and West Palm Beach, which were all of the other usable alternate airports for this aircraft, were either unusable due to thunderstorms or were now north of the weather as well. The aircraft was trapped south of its intended destination and south of its usable alternates. (This airplane was not authorized to use the Key West airport.) Consequently, the crew was faced with a situation where they were now low on fuel with no desirable options in terms of available diversion airports. The flight crew finally had no choice but to break through the line of thunderstorms as the weather passed south of Miami, and then land at Miami. They encountered very heavy turbulence going through the line of weather. Fortunately, no one was seriously injured.

It is important to understand that the dispatcher working this particular flight was responsible for about 30 other flights at that time. He felt that this airplane's situation had been resolved by his previous rerouting action. He had therefore turned his attention to other situations that required his attention.

**Important Features Illustrated by the Scenario**

This scenario provides an example of one of the ways in which the air traffic management system has been decomposed into subtasks to reduce the cognitive complexity for individuals, as illustrated in Figure 3. As implied by this figure, it is assumed that the airline dispatchers and FAA traffic managers complete the strategic planning for a flight (resulting in a filed route that has been approved by the FAA). This flight plan is passed on to the flight crew and to a series of air traffic controllers, who then make tactical decisions to modify the plan if something unexpected arises. The dispatcher still has responsibility for monitoring the flight, and FAA
traffic managers still have responsibility for monitoring the traffic flow of which the flight is a part. Because of workload, however, this monitoring is often done at a more global level, with the dispatcher or traffic managers asking themselves whether something new has arisen that they need to respond to, rather than continuously monitoring each individual flight in detail in the fashion that a controller does.

In this particular situation, nothing new arose at a global level (the weather, in fact, did develop as predicted by the dispatcher), so it is easy to understand why the dispatcher did not monitor the flight more closely. (As far as the dispatcher knew, he had alerted the flight crew and a solution to the problem had been worked out, so there was no need to continue dealing with that flight unless there was some unexpected change in the weather.)

This particular scenario illustrates one potential weakness of such a decomposition: If the decomposition of the system in terms of the distribution of control differs from the decomposition in terms of the distribution of data or knowledge, interactions between components (people) will sometimes be required. If such an interaction is not triggered by some system or person, then a system failure can result. In this particular case, the locus of control resided with the flight crew and the controller, while the relevant data regarding movement of the storm system was available only to the dispatcher and FAA traffic managers.

### Implications for System Design

One response to such a system failure would be to try to change the distribution of data so that the locus of control matches the locus of relevant data. For example, in the current system, the flight crew (in cooperation with the controller) has the authority to make tactical decisions. These tactical decisions sometimes have significant strategic implications (as in this case), which implies that appropriate data regarding the strategic implications needs to somehow be considered.

**Solution 1: Changing the Locus of Data.** One solution, therefore, would be to provide the flight crew with access to all of the relevant data so that the locus of control and the locus of data coincided. In this specific case, the relevant data consists of a "big picture" view of the current and forecast weather. (In addition, knowledge of the legal alternates for the flight was needed. The flight crew had this information). Providing such knowledge to the flight crew might in fact
be a viable solution for this particular problem. Given access to such data, the flight crew would likely have reviewed the weather in south Florida themselves before accepting the second reroute from ATC.

As a solution to this specific scenario, this is an approach that likely would have improved performance. As a more general solution, however, this approach introduces some major concerns. If, for every scenario where there is a mismatch between the locus of control and the locus of data and knowledge, we simply shift the locus of data and knowledge to the person in control, we would likely produce a situation where the cognitive complexity becomes too great for that individual, who now would have to be capable of integrating all of the data and knowledge relevant to all of the possible scenarios that could arise. We could also create a situation where the protection against slips or mistakes, currently provided by redundancy (i.e., multiple persons monitoring a particular flight), would be reduced, as there would be no forcing function to assure that anyone other than this one operator was giving the flight careful consideration. (This latter concern is in fact why regulations state that a dispatcher and pilot share joint responsibility for each flight and require that the dispatcher be notified if a "significant" deviation from the flight plan is being considered.)

Thus, there are non-trivial design issues to consider when deciding whether it is appropriate to try to improve the current system by eliminating some decomposition so that one person now has all of the control that formerly resided across several people, and is given access to all of the data and knowledge relevant to this enlarged span of control. These issues will be discussed further after another type of solution is considered.

Solution 2. Using Intelligent Agents to Trigger Needed Interactions. An alternative to changing the current system decomposition in terms of the allocation of control, data and knowledge would be to ensure that interaction occurred between the flight crew and the dispatcher on those occasions where it was necessary (but without requiring constant interactions when they were not necessary). One approach to accomplishing this would be improved training (making the pilot the intelligent agent responsible for triggering an interaction). A second approach would be to make a designer the intelligent agent, (see Figure 4) developing technology to detect situations where an interaction between the flight crew and the dispatcher is needed. Both approaches merit serious consideration, since both the pilot and the designer are human and susceptible to error.
Evaluation of Alternative Solutions: Solution 1 above focuses on changing the system's architecture, shifting the locus of data to match the locus of control. Solution 2 maintains the current system decomposition, but provides a means for helping to ensure that interactions between different system components (such as pilots and dispatchers) occur on those occasions when they are needed. Other than the issue of cost, Solution 1 has significant merit for this particular problem, and for the general case where pilots are confronted with tactical decisions involving weather that may have strategic impact.

There are, however, some possible drawbacks to Solution 1. Given what we know of human performance, if the flight crews had such weather data, they might be less likely to confer with dispatchers when they decided to make some change, thus reducing the effectiveness of one of the safety nets. If Solution 2 were also implemented, this would help to ensure that dispatch became involved when a "significant" change arose, thus leaving only one “weak link” in the safety net: the designer. For if the designer did not adequately develop the algorithm for detecting a "significant" change and the dispatcher became reliant on such alerts to draw attention to specific flights, the assumed safety net wouldn't really exist (Guerlain, Smith, et al., 1999; Smith, McCoy, and Layton, 1993; Smith, McCoy, and Layton, 1993; Smith, McCoy and Layton, 1997). (This is especially of concern for cases where the flight crew has failed to take a necessary action, as it may be more difficult to design algorithms to detect the absence of a needed change in the flight plan.)

Finally, it is worth noting that we could do a similar analysis of this scenario from the perspective of FAA staff (traffic managers and controllers), asking what solutions might be available within the FAA’s ground support system and what their relative strengths and weaknesses are. In enroute flight, however, pilots are likely to be less heavily loaded than ATC personnel and therefore may represent a more productive approach to such problems.

Case Study 2: The Locus of Control vs. the Locus of Knowledge

Introduction

Historically, traffic flow management has been a function under the control of the FAA, with traffic managers at various facilities making decisions about what routes could be flown by flights scheduled by the airlines. In recent years, however, there has been increasing emphasis on giving the airlines greater flexibility, based on the assumption that the airlines have better information about the costs of alternative methods of operation and should therefore be in a position to make better decisions about the economics of alternative flight plans (Smith, McCoy, Orasanu, et al., 1995). In essence this shifts the task decompositions as, under such changes, airline dispatchers must consider a much larger set of factors if in fact they are to improve system performance. Issues surrounding such a shift are discussed here in terms of alternative system architectures that can be used to accomplish such major changes in the ATM system.

Alternative System Architectures

Alternative architectures for the air traffic management (ATM) system that change the decomposition of tasks for flight planning can be grouped into three categories (Sheridan, 1987; 1992; Smith, et al., 1997; Smith, McCoy, Orasanu, Billings et al., 1997):
1. **Management by directive** (where FAA traffic managers simply inform an airline regarding the route that can be used by a particular flight) (see Figure 5); this is the classical case shown in Figure 1;

2. **Management by permission** (where a default flight plan is assigned by the FAA, which can be revised if the Airline Operations Center requests an alternative and receives permission from FAA traffic management staff) (see Figure 6);

3. **Management by exception** (where an Airline Operations Center can simply file the flight plan that it desires for a given flight) (see Figure 7).

**Details of the Scenarios**

Over the past several years, the ATM system has been evolving from a system in which management by directive was the predominant form of interaction for flight planning toward a hybrid system including examples of all three forms of interactions (Federal Aviation Administration, 1992; 1995).

**Management by Permission:** The first major change arose in 1992, with a shift from management by directive to management by permission. FAA Advisory Circular 90-91 established a formal procedure allowing the airlines to request non-preferred routes (routes for flights that differed from the FAA-assigned preferred routes). Under this procedure, an airline could send a message via teletype to the FAA’s Air Traffic Control Systems Command Center (ATCSCC) requesting an alternative route for a particular flight. A specialist at ATCSCC would then evaluate this request, checking with traffic managers at the involved regional air route traffic control centers and, based on their input, would approve or disapprove the request.

This new paradigm was viewed very positively by both the airlines and the FAA. One airline, for example, reported that in one year, it submitted 15,279 requests for non-preferred routes and that 75% of these requests were approved. These approvals resulted in an estimated savings of 13,396,510 lbs. of fuel.

Thus, this shift toward management by permission gave the airlines a means to improve their efficiency by giving them a routine mechanism by which they could request more economical flight plans for their aircraft. It still left the locus of control with FAA traffic managers, however, as TMUs had to individually approve each requested alternative route. These approvals were given based on considerations of safety (avoiding excessive complexity and traffic bottlenecks), overall efficiency and equitable treatment of different airlines and system users in traffic flows. Thus, this shift left the basic task decomposition the same, but provided a procedure for increasing the relevance and frequency of interactions between traffic managers and dispatchers.

This new architecture, based on management by permission, caused routine interactions to occur between airline ATC Coordinators, ATCSCC traffic specialists and TMU staff (see Figures 5 and 6), giving each group a broader understanding of the factors considered by the other group, resulting in more effective and efficient interactions (which occurred only when the normal task
decomposition was inadequate and there was a need for interactions between the two groups in order to determine the best solution). Put in more general terms, there was no real shift in the locus of control, but there was a new mechanism which allowed airlines to better inform FAA traffic managers of their preferences (resulting in better decisions, from the airlines' viewpoints). A side effect was the dissemination of knowledge from FAA traffic managers to the airline ATC coordinators (allowing them to become more efficient because over time they increased their knowledge about what alternative routes were viable at particular times of day).

A side effect was the dissemination of knowledge from FAA traffic managers to the airline ATC coordinators (allowing them to become more efficient because over time they increased their knowledge about what alternative routes were viable at particular times of day).

\[\text{Figures 5 and 6: Management by Direction and Management by Permission}\]

Limitations of Management by Permission: The primary weakness of this paradigm was that it was manpower-intensive (requiring extra staffing to support the additional interactions), and was thought by the airlines to be excessively conservative at times in terms of the approval of requests for alternative routes. As a result, the traffic management system evolved further in 1995 to give the airlines additional flexibility using a different "architecture".

Management by Exception: Although the development and use of the "management by permission" architecture was viewed as a significant improvement, its perceived limitations were sufficient to motivate a revised program based on "management by exception". This new program, initially known as the expanded National Route Program (NRP) and now referred to as the North American Route Program (still abbreviated NRP), allowed the airlines, subject to certain constraints, to simply file the routes that they preferred for particular flights. FAA traffic managers would then monitor conditions, watching for situations (such as severe weather) when the program had to be canceled temporarily in particular portions of the country. Tactical changes could also be initiated by FAA air traffic controllers (as well as by pilots with the concurrence of the responsible air traffic controllers) after the flight was enroute. Unlike the earlier shift to "management by permission", this architectural change significantly altered the allocation of control, requiring dispatchers to consider factors (such as the prediction of air traffic bottlenecks) that in the past had been handled largely by FAA traffic managers, if the dispatchers wanted to file effective routes.
To evaluate the impact of this architectural change, two studies were conducted dealing with the impact of the expanded National Route Program on fuel consumption. The motivation for these studies came from two sources. First, dispatchers at a number of airlines, as well as traffic managers at enroute air traffic control centers, provided numerous examples of how flights filed under the NRP were sometimes given major amendments and suggested that some of these changes occurred on a regular basis. In some cases, the changes were clearly initiated by ATC to deal with traffic congestion.

Figure 7: Management by Exception

Along these lines, dispatchers made comments such as the following:

"Under the expanded NRP, it's like shooting ducks in the dark."

"The problem with the expanded NRP is that there's no feedback. Nobody's getting smarter. Someone has to be responsible for identifying and communicating constraints and bottlenecks."

"It used to be the weather that was the biggest source of uncertainty. Now it's the air traffic system."

In short, the dispatchers appeared to be indicating that the shift in their tasks gave them more flexibility, but did not give them the knowledge and data necessary to integrate considerations of air traffic (one of the major factors that had formerly been handled primarily by the FAA traffic managers) into their decision making.

As a specific example, one dispatcher indicated that NRP flights from Washington National to Cincinnati frequently had a problem because of the strategy used by ATC to deal with crossing traffic:

"It happens to us all the time. We file the flights at 35 or 39 [altitudes of 35,000 or 39,000 feet] and they're held at 23, 25 and 27. They don't tell us ahead of time that it's going to happen."

A second example of how traffic bottlenecks can affect NRP flights was provided by a traffic manager:

"Quite often ... 8-10 extra aircraft are on this northern route to DFW [from Southern California to Dallas flying north of White Sands (restricted airspace) into the northwest cornerpost at the Dallas-Fort Worth airport] during the noon [local time] arrival rush."
This causes a sector saturation problem in ZFW Sectors 93 and 47 [two Dallas-Fort Worth air traffic control center sectors]. To relieve this volume problem, the ZFW TMU [Fort Worth Center Traffic Management Unit] moves 5 aircraft back to the south route [south of White Sands] via CME.TQA.AQN.DFW [a sequence of navigational fixes into the southwest cornerpost of the Dallas-Fort Worth terminal area]. This longer route of flight, plus the fact that DFW is in a south flow (meaning these flights will spend more time flying below 10,000 feet), will reduce fuel savings or negate them altogether for this bank of flights.

Thus, anecdotal evidence suggested that traffic bottlenecks were arising which impacted the efficiency of NRP flights, raising questions about the effectiveness of this new decomposition of tasks. To gain further insights into this concern, two follow-up studies were conducted. These are described below.

**Follow-On Study 1: Analysis of Predicted vs. Actual Fuel Consumption.** To look for evidence of such inefficiencies, data was collected from a major airline on all of its flights filed over a 5-month period. These data were used to compare predicted fuel consumption on NRP routes with predicted fuel consumption on FAA preferred routes and also with actual fuel consumption. In the following discussion, a "flight" is defined to be a particular combination of an origin, destination, Ptime (scheduled departure time) and equipment type. Thus a given flight could have a new instance filed each day. Predicted and actual fuel consumptions were from takeoff to landing.

Predicted fuel consumptions were first analyzed, comparing performances on FAA preferred routes with the filed NRP routes. 21,334 flight instances were filed by this airline under the NRP during this time period. The average predicted fuel savings per day during this time period ranged from 2.3% to 6.0%. The total predicted savings was 17,723,329 lbs.

**Comparison of Predicted vs. Actual Fuel Consumption:** Given the anecdotal evidence outlined earlier, it seemed possible that these predictions overestimated actual fuel savings for some flights, since the computer's predictions did not take into account the new reroutings that might occur as a result of filing an NRP route and then encountering a traffic bottleneck while enroute. Consequently, we also compared predicted with actual fuel consumption.

To ensure adequate statistical power, only flights with at least 20 instances were considered. There were 267 such flights. A statistical analysis indicated that 94, or 35%, of these flights routinely burned more fuel than predicted (P<0.05). Of these 94, 21% routinely burned more extra fuel than was supposed to be saved by flying the NRP route instead of the FAA preferred route. The flight from DFW (Dallas-Fort Worth) to SNA (Orange County, CA) at 1645 UTC (flying an MD80), for instance, on average burned 1013 lbs. of fuel more than predicted. As a result that flight, which on average was supposed to save 759 lbs. of fuel compared to the FAA preferred route (a predicted 4% savings), actually burned 254 lbs. more than the prediction for the FAA preferred route (a 1.3% loss).

These data also indicated that the city pair that most often had flights with regular problems was LAX to DFW. Seventeen of those flights routinely burned more fuel than predicted.
Implications: At a minimum, these data indicate that there was some sort of a problem associated with 35% of the flights filed by this airline under the NRP during this time period. One possibility would be an underlying inaccuracy in the prediction model for one or more of these flights, over and above any new problems introduced by use of the expanded NRP. If, however, we assume that the prediction model provides unbiased estimates (after discounting any new problems introduced by use of the expanded NRP), then these data indicate that the actual benefits in terms of fuel consumption from the use of the NRP were less than predicted by this airline.

Follow-On Study 2: A Detailed Observational Study of LAX-DFW Flights. As mentioned above, the city pair that most often encountered problems was LAX-DFW. We therefore decided to study data from this city pair in detail in order to collect more refined data on the nature of the problems with NRP flights for this city pair, and to better quantify the impact of these problems.

Methods: Four students from the Aviation Department at Ohio University collected data from June 22, 1996 to August 23, 1996 on the performances of flights from LAX to DFW. Flights with five different scheduled departure times (Ptimes) were studied (1400, 1415, 1445, 1515, and 1810 Universal Coordinated Time or UTC). The students collected data on predicted and actual fuel consumptions and observed each flight instance on the Aircraft Situation Display to record any flight amendments.

Results: The resultant observations quickly made it clear that the underlying problem was the rerouting described earlier. Very briefly, what happens is:

1. A flight instance is filed under the expanded NRP along a route north of the White Sands special use airspace to the northwest cornerpost at DFW;

2. While that flight is enroute, the ATM system decides that there is likely to be a sector saturation problem in the Turkey or Falls high sectors when the flight reaches that point as it approaches the northwest cornerpost into DFW;

3. To deal with that problem, the flight or flights with the most southerly routes that are flying to the northwest cornerpost are rerouted south of White Sands to the FAA preferred route so that they will approach DFW via the southwest cornerpost.

The discussion below provides details on this problem.

<table>
<thead>
<tr>
<th>Ptime</th>
<th>Equipment Type</th>
<th>Number Observed</th>
<th>Pref Route</th>
<th>Route Flown NRP-No Swap</th>
<th>NRP-Swap</th>
</tr>
</thead>
<tbody>
<tr>
<td>1400</td>
<td>DC10</td>
<td>41</td>
<td>44%</td>
<td>17%</td>
<td>39%</td>
</tr>
<tr>
<td>1415</td>
<td>B767</td>
<td>42</td>
<td>48%</td>
<td>19%</td>
<td>33%</td>
</tr>
<tr>
<td>1445</td>
<td>MD80</td>
<td>36</td>
<td>50%</td>
<td>44%</td>
<td>6%</td>
</tr>
<tr>
<td>1515</td>
<td>MD80</td>
<td>41</td>
<td>51%</td>
<td>39%</td>
<td>10%</td>
</tr>
<tr>
<td>1810</td>
<td>DC10</td>
<td>29</td>
<td>38%</td>
<td>52%</td>
<td>10%</td>
</tr>
</tbody>
</table>
Cornerpost Swapping: Table 1 indicates the frequency with which the cornerpost swap occurred for the different flights that we observed. (Keep in mind that this swap usually occurs before White Sands, not as the flights are approaching the airport.) The results indicate that the flights that arrive at DFW for the noon rush (flights that are arriving into DFW around noon local time, and that have scheduled departure times or Ptimes of 1400 and 1415 UTC) are particularly affected. 33-39% of the flights during that time period fell into that category and were rerouted south of White Sands to the FAA preferred route.

<table>
<thead>
<tr>
<th>Ptime</th>
<th>Equipment Type</th>
<th>Number Observed</th>
<th>Expected Change</th>
<th>Actual Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1400</td>
<td>DC10</td>
<td>16</td>
<td>-3.5%</td>
<td>+0.4%</td>
</tr>
<tr>
<td>1415</td>
<td>B767</td>
<td>14</td>
<td>-4.5%</td>
<td>+0.3%</td>
</tr>
<tr>
<td>1445</td>
<td>MD80</td>
<td>2</td>
<td>-3.4%</td>
<td>+1.9%</td>
</tr>
<tr>
<td>1515</td>
<td>MD80</td>
<td>4</td>
<td>-2.3%</td>
<td>+0.1%</td>
</tr>
<tr>
<td>1810</td>
<td>DC10</td>
<td>3</td>
<td>-3.0%</td>
<td>+2.7%</td>
</tr>
</tbody>
</table>

Impact of Rerouting: Table 2 indicates the impact of this rerouting on overall savings for the NRP flights filed at particular Ptimes for those instances where an NRP flight was actually rerouted south of White Sands. All of these flights on average burned more fuel than was predicted if they had been filed on the FAA preferred route. On average, for example, it cost an additional 1502 lbs. of fuel each time the flight at 1400 UTC was rerouted to the southwest cornerpost. A statistical test comparing actual with predicted fuel consumptions for these flights was significant (p<.05) for the Ptimes of 1400, 1445 and 1810.

Case Study 2: Discussion and Conclusions

These discussions of the evolution of the National Route Program, discussing three different architectures (management by control, by permission or by exception), provide an important context in which to consider issues concerning the locus of control and access to the knowledge and data necessary to make the best use of this control. Under control by direction, where FAA traffic managers independently assigned routes for flights, the criticism was that the traffic managers had control, but didn't have the data and knowledge necessary to pick the best routes from the airlines' business perspectives (while still ensuring safety), and did not have the motivation or mechanism to routinely interact with the people who had this additional data and knowledge (airline dispatchers). With control by permission, interactions between the traffic managers and dispatchers were induced, thus helping to ensure that the person with control (the traffic manager) was given access to the relevant data and knowledge (by airline dispatchers). This occurred at a cost in terms of additional labor, however, and was also criticized as sometimes resulting in decisions that were too conservative.
Finally, Case Study 2 discussed control by exception. While in general this shift in architecture appears to have resulted in significant overall fuel savings, these were less than they might have been because the shift in control from FAA traffic managers to airline dispatchers was not matched by either a shift in the locus of data and knowledge about air traffic constraints (so that the dispatchers had direct access to such data and knowledge), nor was it accompanied by a shift in interactions or communication patterns so that the dispatchers could make use of the data and knowledge available to FAA traffic managers.

Thus, in both Case Study 1 and Case Study 2, a major theme is that if the system architecture gives one person or group control, but that person or group

- Does not have the necessary data or knowledge to make the best decision; or
- Does not initiate an interaction with the person or group that does have this data or knowledge;

then significant inefficiencies (Case Study 2) or even safety hazards (Case Study 1) can result.

*Alternative Solutions:* For the situation described in Case Study 1, two classes of solutions were discussed:

Solution 1. Changing the Locus of Data and Knowledge;
Solution 2. Using Intelligent Agents to Trigger Needed Interactions.

For the more strategic planning scenarios covered in Case Study 2, however, the necessary solutions are more complicated because the scenarios themselves are more complex and variable.

One solution would be to return to a management by permission paradigm. This, however, is expensive in that it requires significant extra staffing for both the FAA and the airlines when applied to routine pre-flight planning. Furthermore, there is no indication that either the FAA or the airlines would prefer such a change for that purpose. It is, however, an approach that is being explored for dealing with less frequent decisions such as selecting SWAP (Severe Weather Avoidance Program) routes during severe weather events, in the sense that ATCSCC teleconferences with AOC staff to discuss possible SWAP routes are essentially an opportunity for the airlines to request and recommend certain routes, with FAA making the ultimate decision.

A second solution would be to identify what was successful about management by permission in the original coordinated NRP, and to look for ways to incorporate those features into the current management by exception paradigm in place with NRP. In particular, those successful features included:

1. The dissemination of knowledge from traffic managers to airline AOCs;
2. The use of a single point of contact (ATC coordinators) at each airline to handle most of the interactions with ATCSCC specialists and traffic managers at ARTCC and TRACON TMUs;
3. Allowing FAA traffic managers to play the role of a neutral referee when there was some significant air traffic constraint requiring some type of management of flights through that airspace;
4. Streamlining the process over time, such that non-preferred routes that were routinely requested and approved were marked for automatic approval, without requiring lengthy consultation between ATCSCC specialists and traffic managers at the affected Centers.

Below, methods for incorporating some of these features into NRP are discussed.

The Dissemination of Knowledge from Traffic Managers to Airline AOCs. Under the original coordinated NRP, AOC staff began to learn about constraints in the system, or “tribal knowledge”, because the procedure caused them to routinely interact with traffic managers in an effort to get approval for non-preferred routes. This knowledge consisted of information about where and when bottlenecks in the system routinely arose, and in some cases about potential ways to avoid these bottlenecks while still getting approval for a much improved routing.

Two key notions are contained here. First, the knowledge that was being disseminated was about regularly occurring patterns. Second, there was a procedure that induced the interactions that led to this dissemination. This observation suggests that a surrogate for those real-time interactions would be:

1. Development some type of post-operations analysis tool that identifies routinely occurring patterns (constraints or bottlenecks) and displays them to AOC staff, thus disseminating this knowledge to AOCs and helping to ensure that the locus of control for preflight planning (dispatch) has access to the relevant knowledge;

2. Development of synchronous and asynchronous communications tools that provide AOC staff with a rich environment in which to interact with traffic managers in order to learn more about the bottlenecks identified by this analysis tool and about potential solutions. (Asynchronous tools may often be preferable in order to reduce the need for costly and sometimes difficult to arrange real-time discussions.);

3. Design of a procedure under which specific staff at AOCs and at the FAA have responsibility for reviewing the results of the analysis, for interacting with each other, and for communicating what they have learned to the responsible dispatch staff at the airlines. (As discussed above, experience from the original coordinated NRP suggests that specific individuals need to be assigned these roles as a routine part of their jobs.)
The first goal could be achieved with a tool such as POET (the Post-Operations Evaluation Tool), and the second with tools such as PicTel and Netmeeting for synchronous interactions and tools such as C-SLANT (the Collaborative SLide ANnotation Tool) for asynchronous interactions. As an example of this, consider the displays shown in Figures 8a and 8b from the Post-Operations Evaluation Tool (POET), focusing on flights from Chicago to Atlanta scheduled to depart at 1115 Z. For this particular flight, there were 21 instances in the month of April 1998. The routes flown are shown as the black lines on the map shown in Figure 8a, while the route filed is light. (For all 21 instances, the same route was filed by the airline.) On average, the actual air fuel burn during flight was 20% higher than the uncorrected estimate and the actual air time was on average 25% more than the uncorrected estimate.

A major factor contributing to this extra air fuel burn and air time was airborne holding. As Figure 8b indicates, this was happening 43% of the time.
Figure 8b. POET display of performance statistics for flights with and without holding from ORD to ATL departing 1115Z.

When airborne holding occurred on this route, on average 27% more fuel was consumed than the uncorrected estimate, while 14% more fuel was consumed for those flights where holding does not occur (meaning that something in addition to holding is accounting for a significant amount of the fuel burn). Similarly, average flight time is 34% higher than the uncorrected estimate for the flights with holding and 17% higher for the cases without holding. It is clear from Figure 8a that, in addition to holding, there was frequent vectoring to the southwest as the flights approach Atlanta.
This combination of vectoring and holding is shown more clearly with the individual instance above (see Figure 9). For this particular instance, the air fuel burn was 39% greater than the uncorrected estimate and the air time was 44% greater.

Thus, POET offers the potential to identify routine bottlenecks, and to provide AOC staff with that knowledge. To go beyond simply identifying bottlenecks, however, interaction between AOC staff and FAA traffic managers is needed. Tools like PicTel, Netmeeting and C-SLANT provide rich, efficient environments for such interactions. With C-SLANT, for instance, AOC staff can interact asynchronously with traffic managers by capturing a series of screens while viewing POET (or any other computer displays such as weather data). They can then annotate this with text, graphics and voice (similar to what can be done using powerpoint, but with a much simpler interface because the tool is tailored for this specific use), as shown in Figure 10. C-SLANT then combines with this a feature found in tools like LOTUS Screencam, allowing the user to move the mouse pointer around the screen to focus attention on the relevant details while a voice annotation is being produced, and having both the voice and pointing recorded for replay. The resultant series of annotated slides can then be emailed to another person. (Another feature of this particular application is that a file that might be 75-100 megs in size if produced using
LOTUS Screencam is only 1-5 megs if produced using C-SLANT, thus making email a practical vehicle for communication.

The intended use for this application, then, would be for an AOC staff member to create an annotated slideshow for a sequence of POET screens, raising questions that he would like a traffic manager to address in order to help him understand the problem better and to identify potential solutions. This slideshow would be emailed to the FAA contact, who could add additional annotations to the slide show, answering the questions posed. This response would then be emailed back to the AOC.

Figure 10. Sample C-SLANT slide.

In short, this process using POET and C-SLANT offers a mechanism for approximating (and possibly even improving upon) the dissemination of knowledge about air traffic constraints to AOCs that occurred under the old coordinated NRP. (Ultimately, to use this knowledge most effectively and to avoid excessive cognitive demands, the dispatchers may have to be given better flight planning tools that incorporate consideration of the air traffic constraints that have been identified.)
Allowing FAA Traffic Managers to Play the Role of a Neutral Referee. To the extent that, collectively, the airlines use this knowledge about air traffic constraints to find routings that avoid creating bottlenecks, there would be no need for intervention by traffic managers. Many airline dispatchers espouse this approach, saying: “Let us try to resolve the problems by ourselves before having traffic managers impose a solution.” On the one hand, there are recent examples, such as the collaborative process that is now part of the enhanced ground delay program, that indicate that the airlines can cooperatively react to situations and resolve certain air traffic congestion problems. On the other hand, it is clear that because this is a competitive environment, there will be times where a neutral referee is needed in order to maintain efficient traffic flows, assure safety and maintain equitable treatment of all airspace users. Thus, the first step may be to give dispatchers the knowledge they need in order to adjust their routings to be more effective. The second step would be for FAA traffic managers to intervene when the airlines can’t cooperatively produce acceptable traffic flows. This second approach is discussed more in Case Study 3.

Short-Term vs. Long-Term Solutions. In the discussion above, the discussion focuses on the use of tools like POET, C-SLANT and PicTel or Netmeeting to make better use of the existing capacity. It should be noted, however, that an equally important consideration for the long-term is to use such tools to identify and quantify the costs associated with existing constraints, and to then find ways to increase capacity so that these constraints are removed.

Case Study 3: Use of a Referee for Resource Allocation

Case Studies 1 and 2 emphasize the importance of ensuring that the person who must make a decision has access to the relevant data and knowledge, either directly or through interactions with other people. They also highlight concerns about cognitive complexity: If the knowledge, data or information processing requirements for a task are too great for one person to handle alone, some strategy must be applied to distribute the work. Finally, they caution us about the need to be realistic about human performance: People make slips and mistakes. Hence, it is important to design a system architecture with sufficient redundancies, providing safety nets to protect against human error.

Case Studies 1 and 2 thus emphasize cooperative aspects of the aviation system: FAA controllers and traffic managers working together with airline staff to improve efficiency and safety. In contrast, Case Study 3, which is discussed below, emphasizes the fact that this is also a competitive environment. For when decisions involve economics rather than safety, the airlines compete with one another.

A simple illustration of the issues associated with this competition is provided by surface movement at major airports during winter weather (Obradovich, Smith, et al., 1998). Without some sort of coordination or refereeing of the airlines, significant inefficiencies can result. Consider the case where, due to heavy snow, an airport is down to one runway during a departure push. If departures are handled on a first come, first serve basis, then each airline tends to feel compelled to deice its aircraft and get them into the queue for takeoff as quickly as possible. When all of the airlines do this, however, the net result can be long lines, which in turn may result in delaying some aircraft sufficiently long so that they need to be deiced a second time. The cost of deicing a second time is sizeable ($1200-$1500 per plane). Equally important, if it is
a priority flight for the airline that needs a second deicing, that flight now may be delayed much longer than desired.

An even more striking example of this type of situation arises when NRP flights are filed across the traffic flows for arrivals and departures into an airport. An illustration is provided by crossing traffic in ZNY airspace from HNK flying westbound across J-95/36/223 (departure routes) and J-584/146 (arrival routes). The traffic on these routes is either climbing or descending, and the additional crossing traffic introduced into these sectors adds an increased level of complexity. One such aircraft with a crossing route introduced at an inappropriate time can cause several controllers to make numerous decisions, and take a number of control actions that limit the ability to work the normal volume for this high altitude sector. When such a flight is filed during a departure push at a major airport, that single flight can delay departures by 10-15%. Thus, although the route for that one flight may be more fuel efficient, departure rates for other aircraft may be significantly reduced.

**Potential Solutions**

In the case of runway restrictions due to winter storms, the solution that is currently being explored is to use a neutral “referee” to allocate the limiting resource (departure “slots”) to the airlines involved, thus providing greater predictability and allowing each airline to plan better, reducing inefficiencies. At Boston, for instance, this allocation of departure slots during winter weather is done by a “referee” that is a computer system.

In the case of the NRP flights discussed above, alternative solutions are now being considered. One is similar to that discussed in Scenario 2, providing AOCs with more knowledge about air traffic bottlenecks and allowing them to use this knowledge to resolve the problems themselves as much as possible. For those cases where the situation is competitive (such as where one airline is filing the crossing traffic while others are experiencing the departure delays), however, some sort of refereeing may be necessary. In this case, it appears that the first step is to decide which should take priority. (For the example of crossing traffic through departure sectors, this means looking at the tradeoff between saving fuel for the NRP flight or increasing departures from the airport). At any one airport, this may be a difficult decision. However, given that it happens at many airports, it may be that at one airport one particular airline is benefiting from the NRP flights that are crossing arrival and departure flows, while at another its departures are being delayed by another airline’s NRP flights. If this is the case, it might be possible to establish an equitable solution if it is applied uniformly across many airports.

When considering solutions, however, it is important not to fixate on one solution. To deal with departure delays due to the crossing traffic from NRP flights, for instance, it might be possible to at least partially solve the problem through the use of lower altitude departures (as discussed later in Case Study 5). Furthermore, in the long run it might be possible to increase capacity by changes in sectorization procedures or some other modification of the way the airspace is used.

It should also be noted that refereeing can be imposed either statically (restricting the airlines from filing NRP flights through departure flows at busy times of day) or dynamically (allowing the airlines to file the NRP flights, but giving FAA traffic managers the authority to reroute NRP flights in those cases where it is clear they will actually interfere significantly with departures).
This latter solution is discussed further in Case Study 5 in the discussion of low altitude departures, where control is shifted from an AOC to a traffic manager in order to allow decisions to be made in a more timely, context-sensitive fashion. Along with that shift in control, however, is the communication to the traffic manager of the airline’s constraints and priorities for the traffic manager’s consideration in making the decision.

It should be noted that another important decision in terms of how to implement some type of refereeing is selection of the parameter of control (Wambsganss, 1997). For example, historically the FAA handled arrival restrictions at an airport with Ground Delay Programs that held specific flights on the ground before they departed for that airport, thus limiting the arrival rate. Under that procedure, the parameter of control was at the level of specific flights. Since the goal, however, is to limit the arrival rate rather than to hold specific flights on the ground, the procedure has evolved so that, in essence, when there is a need to limit arrivals due to weather, runway closures, etc., FAA traffic managers now limit the number of arrival slots allocated to each airline in a specific time period and each airline is then allowed to fill its slots with whatever flights it prefers. Thus, the parameter of control becomes the allocation of slots, giving the airlines more flexibility to meet their business objectives.

Finally, a similar approach to exerting control when there is a constrained resource while providing as much flexibility as possible is for the “referee” to provide options and for each airline to then decide which option it prefers. For example, if there is a 20 miles-in-trail restriction for southbound flights through central Florida, the traffic managers can inform AOCs that there are two options, to file a flight along that route with a 20 miles-in-trail restriction or to file it along the east coast of Florida with no restrictions. In this way, the traffic managers are communicating their knowledge of the situation at a more efficient, abstract level. (They are not giving the AOCs the details of why there is a 20 miles-in-trail restriction, nor do the AOCs require this information for effective decision making).

Case Study 4: Use of a Neutral Resource Broker

The FAA’s new enhanced Ground Delay Program not only illustrates the concept of selecting control parameters that provide as much flexibility as possible when allocating some constraining resource (as discussed above), it also illustrates the use of a “neutral broker” to exchange resources among the competing airlines (Wambsganss, 1997). In the enhanced Ground Delay Program developed cooperatively by the FAA and the airlines as part of the Collaborative Decision Making (CDM) Program, there is a procedure called “compression” that allows arrival slots to be exchanged between airlines when an arrival rate restriction has been imposed by FAA traffic managers.

In particular, even though a given airline may have been assigned an arrival slot at the affected airport in some 15 minute window, that airline may not be able to use that slot (because of cancellations, delays due to mechanicals, etc.). Under such circumstances, rather than waste that “resource” (since the slot will be lost once the time period is over), the CDM group has developed a computer algorithm called “compression” that checks to see whether some other airline has a delayed flight that could be moved up into the unfilled slot. When this occurs, the overall system is more efficient, since capacity is used to the fullest extent possible and the airline with the flight that is moved up to the unfilled slot benefits because the delay for that
flight is reduced. The airline that gives up the slot couldn’t have used it anyway, but as an extra incentive, that airline is now traded the slot that belonged to the flight that was moved up. Because that slot is later in time, it may be that the airline can in fact use its new, later slot to reduce delay for one of its other flights.
Case Study 5: Shifting the Locus of Control and Knowledge to Match the Locus of Data

Case Studies 3 and 4 focused on the role of the FAA as a neutral referee. In many cases, however, it is possible for traffic managers or controllers to work cooperatively with airline staff to assist them in better achieving their goals. This occurs, for instance, when a dispatcher calls a TMU and asks if the landing for an overseas arrival can be expedited as it otherwise may have to divert to another airport for refueling.

Case Study 5 explores this cooperative role of the FAA in more detail, looking at two recent cases where it has been expanded. Specifically, these two examples contrast with the solutions explored in Case Studies 1 and 2 in the sense that, rather than improving access to knowledge or data by the person in the system architecture who currently has control, thus letting that person do a better job, Case Study 5 looks at examples where the locus of control is shifted to the person in the best position to have real-time access to the appropriate data to make the needed decision.

Case Study 5a. Use of Low-Altitude Arrival and Departure Routes. The situation to be discussed in Case Study 5a deals with departure delays in the New York area. One example of such delays arises because the high altitude sector (Sector 134) for North Gate departures out of the New York area is frequently very busy in the evening. At times the traffic congestion and complexity threatens to exceed the capacity of the responsible controllers. When this occurs, ZNY traffic managers initiate a departure stop on all flights, often delaying all departures for about 45 minutes until the situation is resolved.

To deal with this situation, a new procedure has recently been developed by the New York Air Route Traffic Control Center, working in collaboration with air carriers and other operators. It is intended for use only as needed and would typically involve capping 2-4 departing aircraft at a lower altitude (FL220) to reduce peak congestion at higher altitudes. New York Center (ZNY) staff believe that this will help them avoid abrupt departure stops at the New York airports during peak periods in the evenings. Flights eligible for involvement in the program would be short-haul flights to destinations such as Buffalo, Toronto and Detroit. In general, such flights would be held at this lower altitude for their duration.

In brief, the use of Low-Altitude Arrival and Departure Routes (LAADRs) for ZNY North Gate departures involves:

1. Making early predictions about conditions likely to cause excessive traffic demands in Sector 134. If ZNY expects wind conditions on a given day will lead to filings that will significantly impact the North Gate departure sector between 6 and 9 pm, it will raise the issue on a mid-day telecon with the airline AOCs. If it is agreed that the procedure may be required, ZNY will send an advisory out at least 2 hours before the time when LAADRing may be necessary. This advisory goes to ATCSCC, to the affected surrounding Centers and TRACON, and to the Airline AOCs, and is updated if conditions change;

2. Upon receipt of the advisory, air carriers can inform ATCSCC that one or more of their flights should not be requested to use the LAADR procedure on that day (because of fuel requirements or other aircraft limitations). Based on this request, those flights are not be
considered by the TMU; if required, they are given ground holds instead of low-altitude departures.

3. All other flights for the participating airlines that are departing the New York area during the time specified in the LAADRing advisory are fueled so that they can fly at either their normal (preferred) cruise altitude or at the lower LAADR altitude (typically FL220). The flight planned is filed with the preferred cruise altitude, but the pilots are informed that the flight may be asked by the controllers to fly at the lower LAADR altitude. (The pilots are also asked not to request higher altitudes once enroute, as this can cause excessive radio frequency congestion and increase controller workload.)

4. In terms of the major point of Case Study 5, the participating flights are approved to fly at either altitude. Based on the traffic loads close to the departure, FAA traffic managers make the decision as to whether to leave each such flight at its filed (preferred) altitude or to change the flight plan to show the lower LAADR altitude. This change is normally communicated to the flight crew at taxi-out, asking for their concurrence.

As with Case Studies 1 and 2, this case study uses a specific example to illustrate another general architecture for cooperative problem-solving. In this case, the problem focuses on excess traffic due to the scheduling and filing practices of several airlines. In particular, the airlines scheduling North Gate departures out of the New York area, along with airlines scheduling overflights (North American Route Program (NRP) flights) through the high altitude sector of concern), are sometimes exceeding its capacity.

Prior to the introduction of LAADRing, the airlines were frequently overloading Sector 134 (a high altitude sector), resulting in departure delays and disruptive departure stops. One solution would have been for the airlines to voluntarily file some flights through the lower, less congested low altitude departure sector. However, in many cases this would be inefficient, as the airline dispatchers do not have the real-time data to decide which flights should be held at the lower altitude and which should not.

Thus, under past procedures (prior to the inception of the LAADRing procedure described above), ZNY traffic managers had only one principal tool available to deal with the situation: delaying departures and initiating very disruptive departure stops. Given current airline priorities (they are willing to spend a little more fuel flying short flights at lower altitudes if this improves departure times), The LAADRing procedure offers a way to dynamically decide which flights should be held at lower altitudes and thus increase needed capacity. Specifically, it shifts the locus of control (selecting the altitudes for certain flights) from the AOCs to traffic managers, as the traffic managers are in the best position to make the real-time decisions. It does so, however, in a manner that allows the AOCs to place certain constraints on the process (namely, they can exempt flights from the process when this is necessary or desirable for economic or safety reasons).

In summary, this is a significant architectural change: As with the expanded NRP, which gave the airlines more control over pre-flight planning because they had the knowledge and data about the costs associated with different flight plans, in this case control is also being shifted. It is being shifted from the AOCs to FAA traffic managers (see Figure 11), however, because the
TMUs have the real time data and knowledge to make the appropriate tactical adjustments (with the concurrence of flight crews and controllers). In essence, AOCs are giving the traffic managers a number of options that are acceptable for particular flights, and indicating their priorities for these options.

![Diagram showing Locus of data & knowledge and Information Flow among ATM Decision Makers](image)

**Figure 11.** Shifting Locus of Control to Person who has Access to Relevant Data

This architectural change contrasts with those discussed in Case Studies 1 and 2. In those two cases, proposed solutions focused on either shifting access to the needed knowledge and data to the person with control (leaving control where it already resided), or ensuring interaction between the person having control and the persons having the relevant knowledge or data (again leaving control where it already resided). In Case Study 5, we are looking at an architectural change in which knowledge and data are left with the person who currently has them (the traffic manager) and control is being shifted to that traffic manager because he is in the best position to make the effective decisions and act upon them. Along with this, however, is a concomitant communication of relevant information from AOCs to traffic managers, so that they can consider airline constraints when making the necessary decisions.

**Overall Summary**

We have made the point in this paper that solutions to the cooperative work problem will not come without costs to the various participants. Our fundamental premise is that the operation of the National Airspace System requires some sort of task decomposition that distributes the work in order to avoid excessive cognitive complexity for any one operator in the system. In this time-paced system, workload becomes an important variable. Excessive workload not only imposes costs upon operators, it also increases the likelihood of errors which can compromise system safety, and it potentially also decreases efficiency.
It should be pointed out, however, that insufficient workload may also pose problems by failing to keep operators sufficiently involved in the management task to be able to make good decisions when they are called upon to do so. As the NAS becomes more heavily automated, this may become a real danger. If controllers, in particular, become primarily monitors of a largely automated system, they are unlikely to be adequately involved in the management task and thus less likely to retain the acute awareness of the traffic situation required to detect problems when they are still easily manageable (Smith, Woods, Billings, et al., 1999; Smith, Woods, McCoy, et al., 1998; Wickens, Mavor and McGee, 1997). This is another cost, of a different type, but it can also compromise safety and decrease system efficiency (Billings, 1997, Carlson, Rhodes, and Cullen, 1966; Hopkin, 1995).

Another system cost will be in training the system participants to understand and work proficiently within the future ATM system. Training issues should be approached from the beginning of the design process and embedded in the design of system elements, to ensure that the needs of the various participants are worked out in advance of detailed system design. It is likely that whatever training is agreed upon should be conducted at least in part as a cooperative exercise among the groups that are to be involved in strategic and tactical management of the future system. Though this is mentioned as a cost, we believe that attention to this aspect of the future system will have very major payoffs for all concerned once the system is operational.

Case Study 2 discusses the implications of changes in management architecture in the evolving ATM system. Dispatchers found the lack of feedback in the expanded NRP disturbing; one commented that “it’s like shooting ducks in the dark.” Though controllers clearly need better information to help them maintain full tactical situation awareness, dispatchers have no less need for information and displays that can help them to achieve full strategic situation awareness, especially of the future flows of air traffic. They presently receive little or no help in visualizing the intent of system elements and participants. In earlier studies, we have discussed some of the economic benefits that have accrued from the NRP; here, we have discussed some of the economic costs, though it must be pointed out that thus far, the benefits probably exceed the overall costs.

Nonetheless, this case study makes it clear that we are not yet in a system posture that will permit us to take full advantage of the more flexible architecture. More studies of actual operational data are needed to provide insights that can be incorporated in the functional requirements for the architecture of the future ATM system and its components.

Another potential cost to participants in the system may be the requirement to work cooperatively in the interests of maximum system throughput. Case Studies 3 and 4 both illustrate the need either for cooperative effort by system users, or for an unbiased manager or “referee” to allocate limited resources in the interests of overall efficiency for all users. A critical question thus becomes how to establish rules and procedures to induce effective voluntary cooperation and coordination when it is needed to deal with either safety or system efficiency concerns, or how to provide refereeing to try to enforce necessary behaviors. Thus, while it is clear that airline users wish to have as much control as possible over NAS decisions and actions (RTCA, 1995), it is not a trivial matter to design and implement the necessary conditions to maximize such opportunities for the airlines.
Finally, these case studies indicate clearly the necessity of shifting access to data and knowledge from their present locations to the new loci of control in the system to minimize the real-time interactions that must take place, and to present that information in ways that support decision making under a variety of circumstances. Such new systems will not be cheap, but they will be vitally necessary adjuncts to the more distributed control that has been requested by users. A concomitant requirement is that the system be designed so that, in those cases where the data or knowledge necessary to make a decision is not directly available to the decision-maker, some mechanism is in place to initiate and support effective access to these data and knowledge.

References


