Alternative Architectures for Distributed Work in the National Airspace System

Philip J. Smith*, Charles E. Billings*, Roger Chapman*, Jodi Heintz Obradovich*,
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

Abstract

The architecture for the National Airspace System (NAS) in the United States 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. This paper characterizes a number of different subsystems that have been recently incorporated in the NAS. The goal of this discussion is to begin to identify the critical parameters defining the differences among alternative architectures in terms of the locus of control and in terms of access to relevant data and knowledge. At an abstract level, this analysis can be described as an effort to describe alternative "rules of the game" for the NAS.

Introduction

Previously, Smith, McCoy and Orasanu (1998) have discussed attempts to improve efficiency within the NAS by changing the locus of control, and consequently changing patterns of interaction and information transfer between air traffic service providers (FAA Air Traffic Management Units or TMUs) and air traffic service users (Airline Operations Centers or AOCs). This previous work introduced the idea that one of the more powerful ways to influence decision making within a highly distributed system like the NAS is to "change the rules" on the assumption that this will influence organizations and individuals to adapt their decision-making processes in desirable ways. In this context, two interesting ways in which the rules can be modified are by changing the locus or nature of the control process or by changing the refereeing process. Below, this past work is briefly reviewed to remind readers of the context. Then new examples of such changes are discussed.

1

Previous Examples

One of the initial approaches to changing Traffic Flow Management (TFM) procedures was a shift from management by direction, the standard air traffic control (ATC) paradigm, toward management by permission, wherein exceptions to preferred routes could be requested by users, considered by providers in the light of their greater knowledge of system constraints and capabilities, and then granted if appropriate. Under this new paradigm, referred to as the National Route Program (NRP), while control remained within the FAA's Air Traffic Management (ATM) system, important new information was shared between TMUs and AOCs. This paradigm shift induced greater information transfer between TMUs and AOCs, resulting in:

- Increased understanding by users of provider and system constraints, and

- Increased understanding by providers of user economic and operational needs.

The use of this approach was associated with substantial fuel and time savings for air carriers. The weaknesses of the new approach included:

- Increases in the time, and therefore personnel, required to accomplish one-on-one interactions between TMUs and AOCs concerning requests for exceptions to the preferred route structures; and

- The increased flexibility afforded to airspace users, while considerable, was still felt to be inadequate by air carriers because, they believed, of inherent conservatism on the part of the ATM system.

Experience with this program led to an important modification of the traffic management paradigm, which can be called management by exception. In this paradigm, there was an actual transfer of control from TMUs to AOCs; AOCs were now permitted to file their desired flight plans, which were automatically accepted by the ATM provider unless environmental conditions required more central control. The clearances granted remained subject to later tactical modification during flight if weather or traffic contingencies required such measures. This program is still in effect, and is called the expanded National Route Program (also abbreviated NRP). The benefits of the modified system include:

- Greater flexibility for users to accommodate economic and other business objectives, and

- Potentially, an ATM system that is more directly responsive to user requirements and needs, as the users now have a way to express their preferences for routes very clearly (by actually filing those routes).

The costs of this modified approach to air traffic management include:

- Significant additional information must be considered by dispatchers if AOCs are going to plan effectively around known or predicted ATC constraints;

- Less information and knowledge exchange occurs between TMUs and AOCs, because users do not have to interact with traffic managers prior to executing their plans. Thus, this new procedure tends to negate the benefit of AOCs having a routine process for sharing air traffic management's significantly greater knowledge of traffic patterns and constraints.

Problem Statement

As outlined above, observation of these two innovations in air traffic management revealed substantial benefits, but also significant shortcomings in the efficiency and flexibility of the resultant systems. In particular, the following issues had not been adequately addressed in the enhanced NRP:

- Use of some airspace was still inefficient due in part to inadequate distribution of relevant information and knowledge to accompany this new distribution of decision making authority.

- There are cases where there is no longer an independent decision-maker ("referee") to allocate finite resources when they are insufficient to meet the needs of all users. This lack of a neutral resource broker can lead to cases wherein certain air carriers may carry out operations that significantly reduce system capacity, thus impacting the ability of other airlines (and sometimes their own airline) to conduct operations efficiently.

Approaches to these Problems

A new FAA/industry initiative called the Collaborative Decision Making (CDM) program is one of several approaches that have been initiated in order to improve information flow and procedural management of remaining bottlenecks in the NAS. Additional procedures are being developed in an attempt to allocate system resources fairly while providing as much flexibility as possible. New technologies are also under study with the intent of providing better information management and decision support for providers and users.

Several of these initiatives appear to hold promise for improving:

- Understanding by both providers and users of each others' needs and priorities; and

- Ways of planning and executing operations that make better use of constrained resources, without unacceptable increases in system overhead, cognitive complexity or operator workload.

It should be noted that the design of the original NRP in 1991-92 and its enhancement in 1995 were ambitious attempts to improve system operation, motivated in large part by naturalistic observation of system deficiencies. The NRP "experiments" have given us a better understanding of the capabilities and limitations of both organizational and operator behavior in the changing context of this complex, rapidly-evolving real-world system. This increased understanding has been an important input into CDM activities aimed at further improvements in an ATM system operating under continually increasing pressure.

In particular, the studies summarized above illustrated how, if the system's architecture gives one person or group control of the situation, but that person or group:

3

- Does not have the data or knowledge to support an effective decision; or

- Does not initiate an interaction with the person or group that has this data or knowledge,

then significant inefficiencies or even safety hazards can result.

Several of the efforts underway as part of the CDM program are intended to increase transfers of relevant information and knowledge to improve decision making in situations where control has been shifted in the system. In addition, other efforts have taken a different approach, looking for a different way to modify the locus of control. A few examples follow.

Example 1. Increasing the Dissemination of Knowledge

Dissemination of knowledge from traffic managers to AOCs was markedly improved under the original coordinated NRP, because the procedures required them to interact with each other in order to obtain approval for non-preferred routes. The system overhead and personnel workload were high, however, and these factors led to modifications that eliminated most of those interactions.

Such interactions, however, need not be synchronous (a difficult problem for these extremely busy people) in order to support such a transfer of knowledge. A potential surrogate for such real-time interactions could be:

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

- Development of synchronous and asynchronous communication tools that can provide AOC staff with a rich environment in which they can interact off-line with traffic managers to learn more about the bottlenecks identified by the analysis tool, and by which they can explore potential solutions to these problems with traffic managers. Asynchronous tools will often be preferable if they offer sufficiently rich means for two-way communication.

Development of procedures by which the results of these post-operations analyses and interactions are disseminated to responsible dispatchers at airlines and to appropriate management staff at ATM facilities is also, of course, an essential part of such a scheme.

Tools to assist in acquiring such knowledge (such as the post-operations evaluation tool, POET) have been developed and linked to asynchronous communications tools (such as the Collaborative Slide Annotation Tool or C-SLANT) (Smith et al., 1999). The former is now being tested in limited operational use to support post-operations reviews involving the FAA's Air Traffic Control Systems Command Center, Enroute Centers and AOCs. Such technologies enhance the sharing of information and knowledge and can thus improve understanding by providers and users of each others' needs and priorities, potentially with less overhead cost than

was observed in the original NRP procedures. Figures 1-4 show examples of displays from these two software packages.
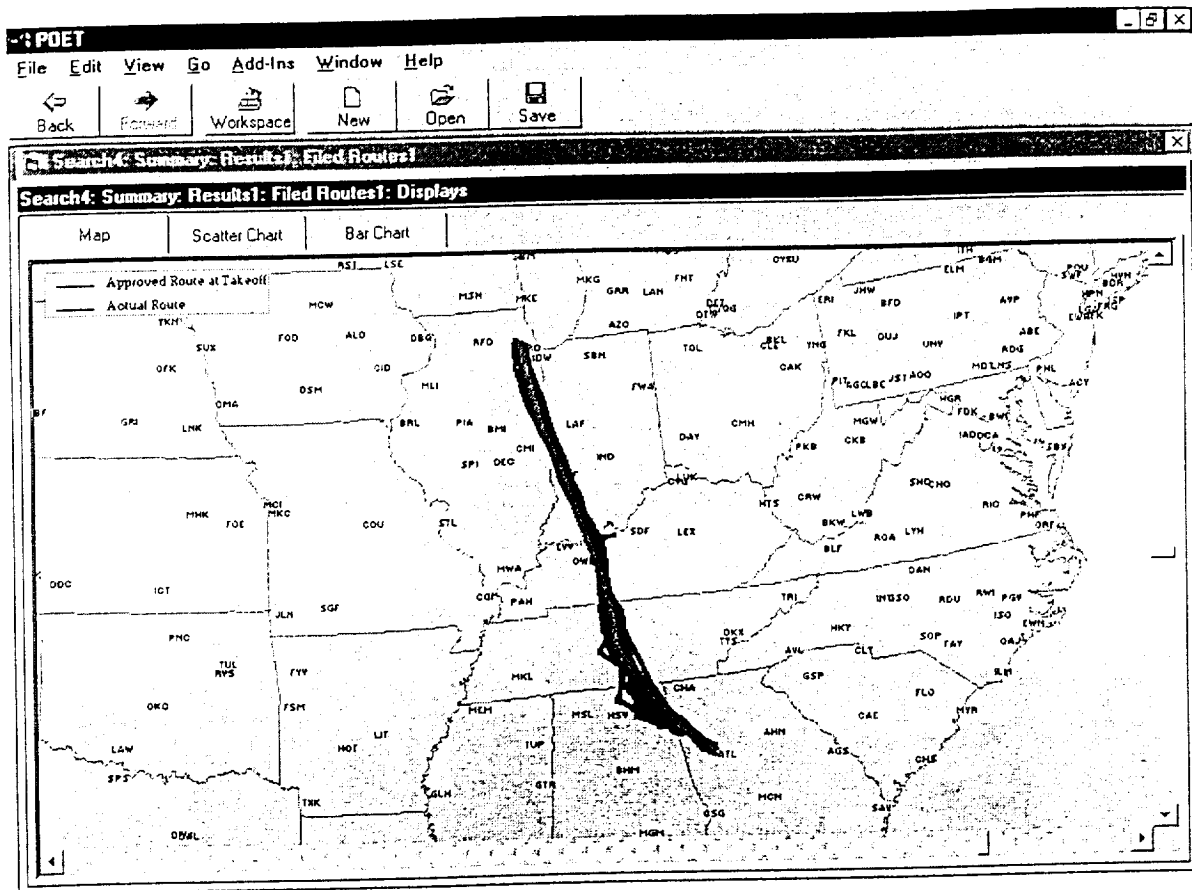

Figure 1. POET display of filed and actual routes for flights from ORD to ATL departing 1115Z.

File Edit View Go Add-Ins Window Help

⇦ Back | ➡ Forward | 🗂 Workspace | 🗋 New | 📂 Open | 💾 Save

Search1: Summary: Results1: Filed Routes1
Location: Search1: Summary: Results1: BBB ORD ATL 1115 SUPER80 Filed Routes

Search: MinInstsForFlt: 15; Dates: 4/1/98 To 4/30/98; DepartTime: 1115; DptApt: ORD; ArrApt: ATL

Explore Data Mining...

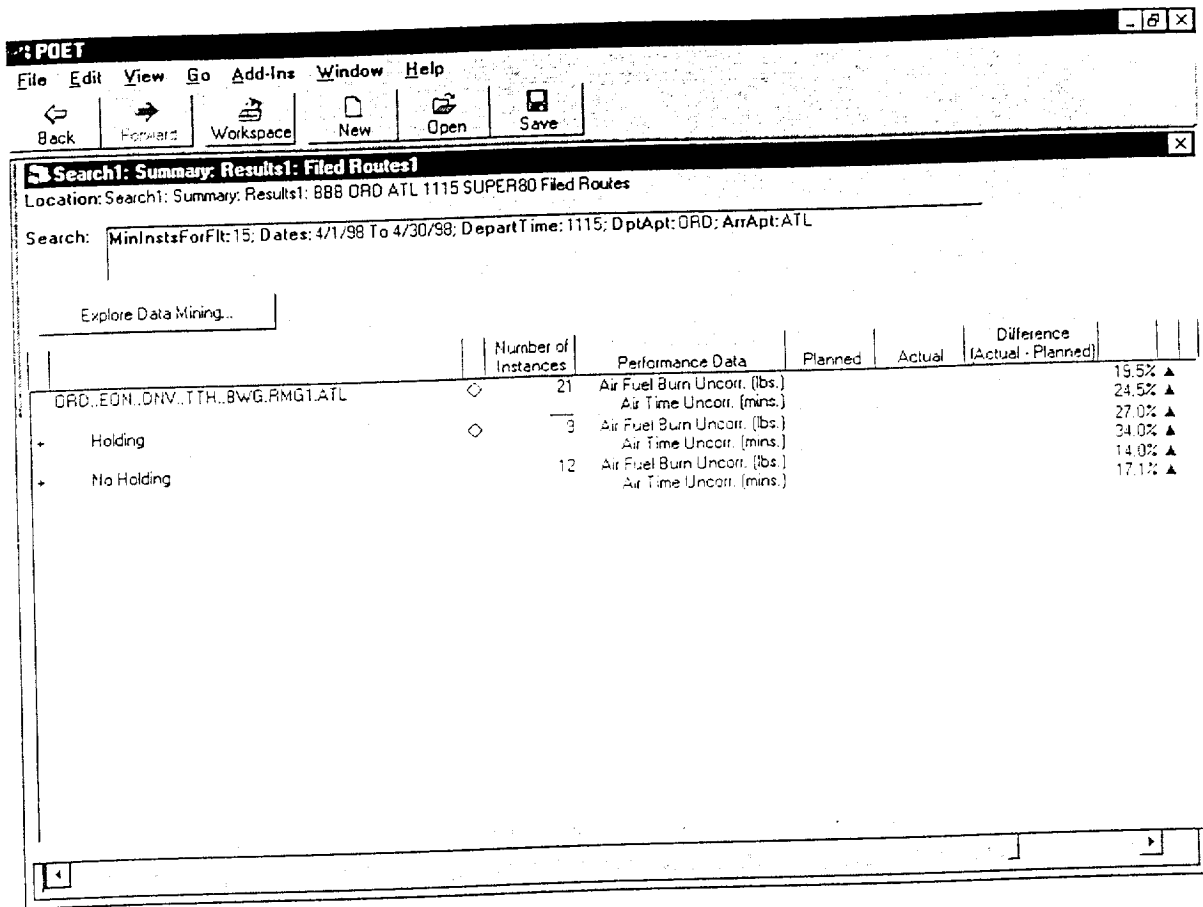|	Number of Instances	Performance Data	Planned	Actual	Difference (Actual - Planned)		
ORD..EON..DNV..TTH..BWG.RMG1.ATL	◇ 21	Air Fuel Burn Uncorr. (lbs.)				19.5% ▲	
		Air Time Uncorr. (mins.)				24.5% ▲	
Holding	◇ 9	Air Fuel Burn Uncorr. (lbs.)				27.0% ▲	
		Air Time Uncorr. (mins.)				34.0% ▲	
No Holding	12	Air Fuel Burn Uncorr. (lbs.)				14.0% ▲	
		Air Time Uncorr. (mins.)				17.1% ▲	

Figure 2. POET display of performance statistics for flights with and without holding from ORD to ATL departing 1115Z.
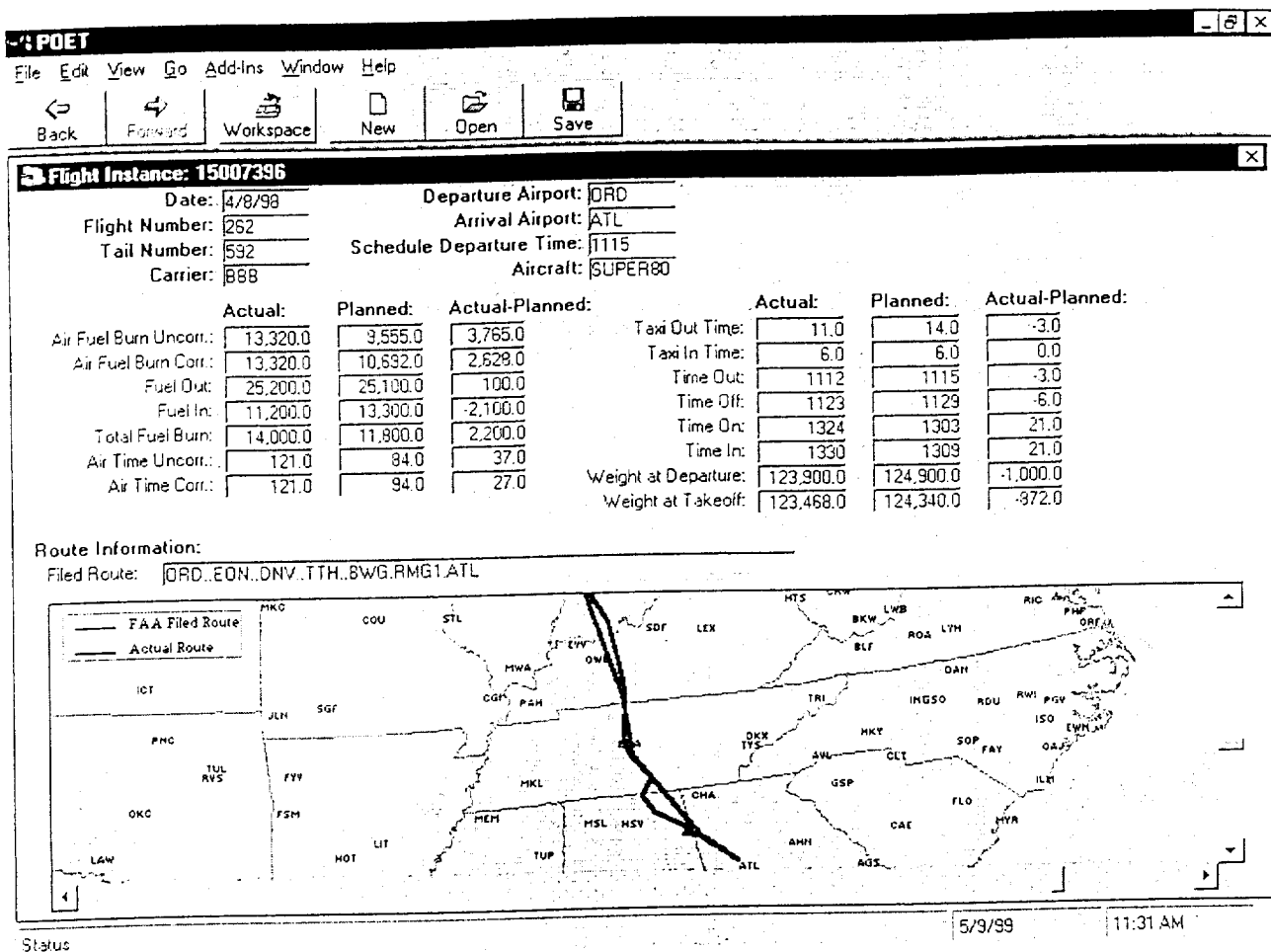
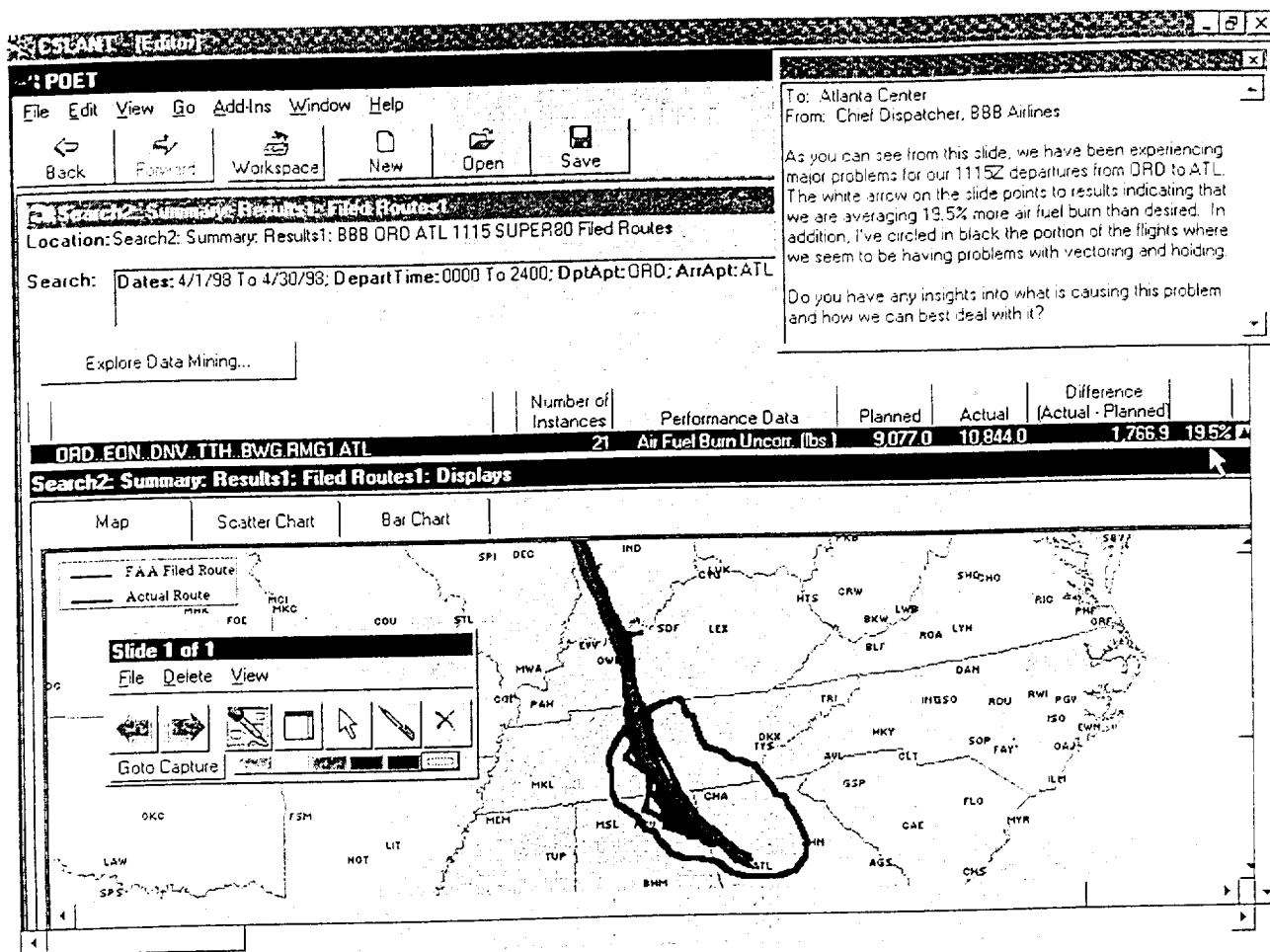Figure 3. Individual instance of the ORD-ATL 1115Z flights.

Figure 4. Sample C-SLANT slide. (Typically, annotations using CSLANT use voice and pointing rather than text.)

Example 2. Management by Control with Increased Flexibility

Another approach to improving traffic management when resources are constrained, while providing as much flexibility as possible, is for traffic managers to provide routing options to airlines, which then decide which options they prefer for specific flights. For example, if there is a 20 miles-in-trail restriction for southbound flights through central Florida, traffic managers inform AOCs that there are two options: to file flights along that route with a 20 miles-in-trail restriction, or to file those flights along the east coast of Florida with no dynamic capacity constraints. In this fashion, traffic managers are communicating their knowledge of the situation at an efficient, abstract level, leaving carriers free to adopt whichever strategy favors their business objectives. (This is primarily a one-way flow of information and knowledge, however.)

Example 3. Changing the Parameter of Control

Making better use of constrained resources while increasing operator flexibility has been approached in a number of ways. Historically, the air traffic management system has handled arrival restrictions at airports with Ground Delay Programs that held specific flights at their departure points, thus limiting the arrival rates at the destination airports. Thus, the parameter of control was at the level of specific flights.

Since the goal is only to limit the arrival rate, however, this procedure has been modified under the enhanced Ground Delay Program so that when there is a need to constrain arrivals (due to weather, runway closures, etc.), traffic managers now limit the number of arrival slots allocated to each airline during a specific time period. Each carrier is then allowed to use its slots for whichever flights it prefers. The parameter of control becomes the allocation of arrival slots, giving airlines more flexibility to meet their business objectives.

Example 4. Use of a Neutral Resource Broker

In the enhanced Ground Delay Program introduced in Example 3, traffic management controls the use of a constrained resource, ensuring that it is used in a fair and impartial manner, while giving the airlines maximum flexibility to meet their objectives. Another dimension of this program is the use of a procedure called "compression" that allows arrival slots to be exchanged between airlines when an arrival rate restriction has been imposed by traffic managers (Wambsganss, 1997). If a given airline has been assigned an arrival slot at an affected airport in some 15-minute window, but is unable to utilize the slot because of cancellations, delays for mechanical reasons, etc., a compression algorithm is used to ascertain whether some other airline has a later flight that could be moved up into the unfilled slot rather than waste that slot. If so, system efficiency is improved because capacity is used to the fullest extent possible; the flight moved into the unfilled slot benefits because its delay is reduced. The airline that gives up the slot could not have used it, but as an added incentive, that airline now receives the slot vacated by the aircraft moved up.

Example 5. Shifting the Locus of Control to Match the Locus of Data

For a number of reasons, there has been a significant increase in the demand for certain high altitude sectors, as well as an increase in the complexity of the traffic patterns within those

sectors. One contributing factor has been the changes in traffic patterns because of the use of the expanded NRP by the airlines. A second has been the transition of commuter and regional carriers from turbopropellor aircraft into high-performance small jets. The magnitude of this latter change is illustrated by transport aircraft sales during the last year, "when 534 regional jets were sold and only 25 turboprops," (Eccleston, 2000).

An example of this problem is associated with crossing aircraft traffic in New York Air Route Traffic Control enroute airspace from the Hancock VOR flying westbound across J(et airways) 95/36/223 departure routes and J-584/146 arrival routes into New York. The departing and arriving traffic is either climbing or descending, and the crossing enroute traffic adds increased complexity into air traffic management. One such flight on a crossing route introduced at an inappropriate time can require several controllers to make numerous decisions and take control actions that limit the ability to work the normal traffic in this high altitude sector. When such a crossing flight appears during a departure "push" at a major airport, that single flight can delay departures by 10-15%. Thus, although the route for the enroute crossing flight may be more fuel efficient, departure rates for all other aircraft may be significantly decreased.

One possible solution to such problems is to provide AOCs with more detailed knowledge about air traffic bottlenecks and to allow them to use the new knowledge to resolve the problems among themselves as much as possible (the approach discussed in Example 1). For cases where the situation is competitive, however (such as the case in which one carrier is filing the crossing routes while others are impacted by the departure delays), some sort of refereeing may be necessary. In these cases, the first step is to establish which type of traffic should have priority. Specific cases at any one airport can pose difficult decisions, though widespread application of prioritization may balance out specific inequities.

However, it is important not to become fixated on single solutions. To deal with departure delays due to crossing traffic from NRP flights, it might be possible to at least partially solve the crossing problem by making use of lower altitudes for departures and arrivals. An example of the use of this strategy is the use of a wider spectrum of altitudes to relieve ground delays for departing traffic in the New York area, through the application of the Low Altitude Arrival and Departure (LAADR) program. In this case, cooperation between FAA traffic managers and AOCs has led to shifting the locus of control from AOCs back to traffic management units, which have better real-time data to make decisions about which flights to assign to LAADRs. To better accommodate airline constraints, however, this is done as part of a collaborative process.

As an illustration, if the high altitude sector for North Gate departures is projected to be overloaded during the evening (usually because of specific wind patterns), New York Center traffic managers have in the past initiated departure stops on all northbound flights, often delaying departures through the North Gate for about 45 minutes until the situation is resolved. These blanket departure delays have been effective in preventing traffic overloads in the high altitude departure sector, but at a substantial cost to many airlines in terms of serious delays.

To deal with this situation more effectively and less intrusively, procedures have been developed to allow the Center traffic managers to work in collaboration with AOCs to dynamically adjust departure altitudes for specific flights. It is intended for use only as needed and typically

involves capping 2-4 departing aircraft at a lower than normal altitude (22,000 feet) to reduce peak congestion in the higher altitude air traffic control sector. This process makes it possible to avoid abrupt departure stops at New York airports during peak periods in the evenings. Flights eligible for involvement in the program would typically be short-haul flights to destinations such as Buffalo and Toronto; in general, the selected flights remain at the lower altitude to their destinations.

In brief, the use of LAADRs for New York Center (ZNY) North Gate departures involves:

- Making early predictions about conditions likely to cause excessive traffic delays in the high altitude departure sector. If ZNY expects wind conditions will lead to route 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 needed, ZNY will send out an advisory at least two hours before the time when LAADRing may become necessary. This advisory goes to the ATC System Command Center (ATCSCC), to affected surrounding Centers and TRACONs, and to the Airline AOCs. It is updated if conditions change.

- Airlines can inform the ATCSCC if one or more of their flights should not be requested to accept a LAADR clearance on that day, because of fuel requirements or other limitations. Those flights will not be considered by the TMU; if required, they will be given ground holds instead of low-altitude departures.

- Other flights of participating airlines departing New York during the time period specified in the advisory are fueled so that they can fly either at their preferred cruise altitude or at the lower LAADR altitude, and their pilots are informed that the flight may be asked by ATC to fly at the lower altitude. (Pilots are also asked not to request higher altitudes once enroute to avoid excessive radio frequency congestion and increased controller workload.)

- Based on traffic loads close to the departure time, traffic managers make a decision whether to leave each participating flight at its original filed 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.

An alternative solution for this problem would have been for airlines to voluntarily file some flights through the lower, less congested low altitude departure sector. This would have been inefficient, however, as in this case AOCs 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, ZNY had only one tool available to deal with this situation: delaying departures by very disruptive ground stops. Given airline priorities in 1999 (they were willing to expend slightly more fuel flying short flights at lower altitudes if this lessened departure delays), the LAADR procedure offered a way to decide dynamically which flights should be held at lower altitudes and thus increase needed capacity. It shifts the locus of control (selecting altitudes for certain flights) from AOCs back to traffic managers, as the latter are in the best position to make the real-time decisions. It does so, however, in a way that allows the

AOCs to place certain constraints on the process, by exempting flights from the process when this is necessary or desirable for economic or safety reasons.

This is a significant architectural change. As with the expanded NRP, which gave airlines more control over pre-flight planning because they had the best knowledge and data about the costs associated with various flight plans, in this case control is also being shifted, but from AOCs to traffic managers because the TMUs have the real-time data and knowledge to make appropriate tactical adjustments. 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.

General Discussion and Conclusions

The various FAA/industry initiatives illustrated here have led to more efficient use of airspace, to greater flexibility for system users, and to implementation of a "referee" function as necessary, after prior consultation with users. Such innovations in the ATM system during the past decade suggest that procedural approaches based on naturalistic observations and decision making principles can produce considerable further improvements in the functionality of today's air traffic management system despite its inherent complexity and the severity of the demands being placed on it by rapidly increasing traffic.

These observations support the hypothesis that shifting the locus of control in the system will be accompanied by a need for substantial changes in the management, distribution, and display of relevant information and knowledge to system participants (Smith, et al., 1999). They further suggest that there is no one best "architecture" to deal with all of the situations that arise in the air traffic management system. Rather, they suggest that there are various ways to distribute control, along with access to the relevant data and knowledge.

In short, these observations about the evolution of the NAS lead to the following hypotheses:

- That distributed decision-making can work if (and probably only if) we can limit the amount of information assimilation and knowledge required of each decision maker. Time pressures may require that decisions be partitioned or performed by teams to limit cognitive complexity for individuals. Conversely, we must also provide each decision maker with access to the data and knowledge necessary to adequately perform his or her function, and with feedback about the impacts of these decisions.

- In a system involving constrained resources and competitive pressures, there must be a referee to ensure that all users are treated as nearly equitably as possible. Stakeholders must be involved in planning to ensure that their interests and information sources are considered by the system. Though "committee decision making" usually will not work under temporal pressure, "committee consultation or collaboration" prior to decision making is often possible and useful, and may significantly improve and limit criticism of system decisions once made.

- The usefulness of asynchronous communications as an avenue for interactions deserves to be explored intensively because of its potential to increase knowledge flow within the system.

• At heart, the relationship between the distribution of information and the distribution of authority to use that information for decision making is a cognitive engineering and psychology problem, above and beyond any technical issues or technologies involved in its solution. Technology can certainly assist, once the system's parameters, information content, objectives and permissible modes of behavior are specified, but technologies are at best simply tools to assist human decision makers to order and direct system behavior.

Clearly there is a need for data to evaluate the impact these hypotheses in the context of these naturalistic "experiments" that are being conducted as the NAS evolves. However, one of the first steps is to clearly identify the parameters defining different cognitive architectures for such a complex system, so the contributing factors influencing individual and organizational performances can be studied more effectively.

Acknowledgements

References

Eccleston, B. (2000). In Thomas, G. and Mecham, M. Continental's 100-Aircraft Order Reflects Regional Market strength. Aviation Week & Space Technology, 152 (9), p. 35.

Smith, Philip J., Billings, Charles E., McCoy, C. Elaine and Orasanu, Judith. (1999). Alternative Architectures for Distributed Cooperative Problem-Solving in the National Airspace System. Technical Report CSEL-1999-24. Cognitive Systems Engineering Laboratory, Institute for Ergonomics, The Ohio State University, Columbus OH.

Smith, Philip J., McCoy, C. Elaine and Orasanu, Judith. (1998). Distributed Cooperative Problem-Solving in the Air Traffic Management System. Proceedings of NDM '98. Washington, D.C.

Wambsganss, M.C. (1997). Collaborative Traffic Flow Management. Washington, D.C.: Metron.