BY THE COMPTROLLER GENERAL

Report To The Chairman, Subcommittee On Transportation House Committee On Appropriations

OF THE UNITED STATES

Federal Aviation Administration's Host Computer: More Realistic Performance Tests Needed Before Production Begins

At the request of the Subcommittee Chairman GAO reviewed the Federal Aviation Administration's (FAA's) $725-million Host Computer Program Host computers are intended to support the nation's air traffic control in the late 1980s and the 1990s

GAO found that (1) vendor testing prior to acquisition of the Host computers did not adequately simulate present or future operational requirements and (2) documentation of test plans and results and technical oversight of performance testing were inadequate

GAO also questioned the accuracy and reliability of the computer model upon which FAA s projections of air traffic delays are based FAA has not adequately substantiated that the urgent need for the Host computers precludes additional tests from being conducted prior to the production decision

GAO recommends that the Secretary of Transportation consider the merits of conducting realistic performance testing on the proposed Host computer systems before the production contract decision is made The Congress should consider directing the Secretary to defer the Host contract award if the Secretary decides to proceed without realistic performance testing and does not satisfactorily explain to the Congress the basis for such a decision

GAO IMTEC-85-10
JUNE 6 1985

COMPTROLLER GENERAL OF THE UNITED STATES
WASHINGTON D C 20548

B-206887

The Honorable William Lehman
Chairman, Subcommittee on
Transportation
Committee on Appropriations
House of Representatives

Dear Mr. Chairman:

The Federal Aviation Administration (FAA) plans to spend approximately $725 million to develop, acquire, and operate new computers for its 20 air traffic control centers. These so-called "Host" computers are intended to provide needed computer capacity to handle, without delay, air traffic anticipated in the late 1980s and the 1990s. The computers will also provide the needed capacity for software additions that are intended to improve air safety and efficiency. As such, it is important that the decision to acquire these capabilities be made on a sound basis. This report responds to your May 10, 1984, request that we determine whether FAA's testing of two vendors' proposed Host computer systems will provide adequate information upon which to base a production decision scheduled to be made in June 1985. Specifically, you asked us to

- examine the soundness of FAA's projections that capacity shortages in existing computers will significantly delay air traffic;
- evaluate the adequacy and completeness of the performance testing conducted on the proposed Host computers; and
- determine the status of FAA plans to install "functional enhancements" (new software) on the Host computers.

The House Committee on Appropriations has stressed the importance of testing systems under realistic conditions. In reports on the 1984 and 1985 appropriations,[1] the Committee directed FAA to fully test the Host computer's performance, before making a production decision, to ensure that operational requirements could be met or exceeded.

---

[1]Reports on Department of Transportation and Related Agencies Appropriation Bills, 1984, H.R. 98-246 and 1985, H.R. 98-859.

In February 1985[2] FAA responded to your concerns about Host computer issues. We have reviewed this response and have incorporated FAA's position in our evaluation.

Our findings are summarized below and are discussed in more detail in appendix I. Appendix II describes our objectives, scope, and methodology.

FAA HAS NOT ADEQUATELY SUPPORTED THE URGENCY OF MAKING ITS PRODUCTION AND VENDOR SELECTION DECISIONS

FAA has not adequately substantiated projections of air traffic delays due to computer capacity shortages--its primary reason for planning to make a production decision in June 1985 and a vendor selection decision in July 1985. FAA's projections of these delays are based on a computer model whose reliability and accuracy are uncertain.

FAA's model uses unverified assumptions about, for example, (1) air traffic distribution among air traffic control centers and (2) the point at which the current computers become overloaded and delay air traffic. While the model may be sufficiently accurate and reliable to support assertions that certain air traffic centers will eventually encounter computer capacity problems, it does not provide the precision and confidence to adequately support FAA's claimed urgency that it must make production and vendor selection decisions by July 1985. In discussions with FAA officials regarding the use of these unverified assumptions, we were told that FAA's projections can be viewed as accurate within plus or minus 2 years.

FAA's latest projections indicate that air traffic delays may not be as great as FAA originally reported to the Congress in 1982. At that time, FAA projected that up to eight centers could experience significant traffic delays by 1989. FAA now projects that four centers will experience delays by 1989. Of these centers, only two are projected to experience significant delays. Our discussions with automation officials at the two centers indicate that neither is actually experiencing the significant delays projected by the model. In its February 1985 response to the Congress, FAA projected that, by the end of 1987, 253 occurrences of air traffic delays could be incurred nationwide. Projections based on more recent aviation forecasts show 135 occurrences by the end of 1987, or about a 47-percent reduction. Projections through 1989 show about a 28-percent decrease--from 716 to 519.

---

[2]Response to Congressional Concerns Regarding the FAA's Advanced Automation Program for Air Traffic Control, DOD/FAA/AAP-85-1, Feb. 1985.

FAA TESTING DOES NOT CONVINCINGLY SHOW THAT PROPOSED SYSTEMS CAN PROVIDE NEEDED CAPABILITY

FAA's performance testing during the design competition phase was not in accordance with guidance established by the Office of Management and Budget and the Department of Transportation (DOT). Both the system performance test and the environment in which the test was conducted were significantly less demanding than the operational capabilities the Host computer must provide. According to FAA, the unavailability of an adequate workload test and limitations at its test center, among other reasons, prevented it from performing realistic system performance tests. Because of testing shortfalls, convincing evidence is lacking on whether either vendor's system can adequately provide the performance needed.

Office of Management and Budget Circular A-109 and DOT Order 4200.14B state that, before a production decision is made, realistic system performance tests should be conducted in the operational environment or in an environment which realistically represents the operational conditions in which the system is expected to perform. This guidance is key to ensuring that the selected system meets the government's needs. In addition, the House Committee on Appropriations has directed FAA to fully test system performance, before making a production decision, to ensure that operational requirements can be met or exceeded.

We found, however, that FAA used a performance test that represented the 1970s air traffic environment. For example, performance test results showed a peak aircraft tracking workload of about 450 aircraft per air traffic control center. However, FAA expects the Host computer to track as many as 600 aircraft at certain centers in the 1990s. Also, key automated operational functions, such as those which aid the controller in detecting potential aircraft collisions, were not used during the system performance test. FAA officials said that, because of its complexity, a representative test was not available during the design competition phase.

In addition, due to limited capabilities, FAA's Technical Center was unable to simulate an operationally realistic air traffic control center environment during design competition phase testing. While tests at the technical center were limited to 12 controller positions, the actual controller workload at some centers can include as many as 60 positions. FAA officials said that, in their view, the expense in time and money to create a more realistic test environment was not worth the perceived benefits.

Proceeding to production without sufficient information on system performance could lead to additional time and money being spent in order to provide a system that performs as required. Similar acquisitions of major computer systems that were based on inadequate system performance testing demonstrate these effects. The Air Force may have to spend an additional $200 million to

upgrade its Phase IV Base Level Computer System so that it will meet requirements. And the Army is experiencing difficulty meeting mission needs because its base operations system must use its contingency capacity to meet day-to-day requirements. In both cases, production decisions were made without sufficient information on system performance.

FAA recognizes that it did not perform operationally realistic system performance tests. However, program officials explained they have taken actions which will provide sufficient confidence that the chosen system can meet its operational needs. For example, FAA estimated the minimum computer processor speed required in millions of instructions per second and required vendors to demonstrate the capability to upgrade to the next size processor if this estimate is low. FAA also plans to conduct a more realistic system performance test after the production vendor is selected to ensure that the system fielded can handle the heavier workload anticipated in the 1990s. This performance test will be conducted as part of a planned 4-month software testing effort.

In addition, FAA may receive additional evidence of system performance when it evaluates vendor proposals in response to the acquisition phase request for proposals. However, this information will not be used to support the production decision.

Regarding FAA's actions, we have serious concerns about using computer processor speed as the basis for estimating the minimum performance needed in its Host computers. Computer processor speed is a measure of the capability of one central component of a computer system. We believe that response time is a better indicator of system performance. Response time--the elapsed time between the end of an inquiry or demand on a data processing system and the beginning of the response--is dependent on how well processors, peripheral devices, communications, and software work together as a system. We therefore question whether an upgrade in the speed of the computer processor is an adequate way to solve all potential system performance problems.

According to Office of Management and Budget and DOT guidelines, operationally realistic system performance testing should be conducted before the production decision. Such tests are an effective means of ensuring that the system selected will satisfy FAA's needs. It would appear that testing both vendors' systems before production would have provided better performance information for the production decision without greatly extending the program's overall schedule. This is particularly true in light of the potential plus or minus 2-year variation in FAA's traffic delay projections.

TESTING DOCUMENTATION AND TECHNICAL OVERSIGHT OF PERFORMANCE TESTING WERE INADEQUATE

Although required to do so, FAA has not adequately documented its actions in planning, monitoring, and appraising the testing of

vendor systems during the design competition phase. Technical oversight of performance testing for the Host computer was also inadequate. Consequently, the usefulness of test results was reduced.

Office of Management and Budget and DOT criteria emphasize management controls, such as documentation and program monitoring. Specifically, agencies should formally document their system testing and evaluation activities. DOT guidelines further state that the Deputy Secretary of Transportation and DOT's Transportation Systems Acquisition Review Council (TSARC) should monitor the Host computer acquisition to ensure that system performance risk is being adequately reduced. In addition, FAA orders require the Operational Test and Evaluation Staff to evaluate Host computer operational readiness before the production decision and report their results to the FAA Deputy Administrator.

We found that FAA (1) did not prepare a plan describing test objectives, scope, timing, and criteria; (2) did not document its plan for evaluating vendor tests reports; and (3) will not document its analysis of vendor tests until after the production vendor has been selected.

The Director of the Advanced Automation Program Office recognizes that formal documentation for some Host computer program activities does not exist. He explained that formal documentation, although important, would have placed a heavy burden on his staff given the ambitious acquisition schedule of the Host computer. Rather, he established procedures which he believed provided insight into the vendors' progress. For example, he required FAA to use the quarterly program review process to provide comments to the vendor, such as indicating whether vendor system design and corrective actions were responsive to the agency's engineering requirements. However, we believe that without documentation of the agreements reached between FAA and the two vendors, the effectiveness of FAA's management control cannot be fully evaluated.

We also found that independent technical oversight of system performance issues was inadequate. TSARC has provided limited monitoring of the program's technical aspects. FAA's Operational Test and Evaluation staff was limited by the Advanced Automation Program office's untimely release of technical information and test data used by the vendors to prepare their test reports. DOT officials told us that they believe TSARC has provided adequate overall management oversight and that all information pertaining to the Host computer program will be considered before determining if the computer is ready for production.

STATUS OF FAA PLANS TO INSTALL NEW SOFTWARE ON THE HOST COMPUTER

You asked us to determine the status of air traffic control software additions, or enhancements, which are intended to improve air safety and efficiency. We concentrated on the Mode C Intruder,

En Route Metering II, and Conflict Resolution Advisory functions because they are scheduled to be the first three enhancements to be installed on the Host computers. (See app. I, p. 12.) According to FAA Advanced Automation Program officials, these three enhancements were not ready when the design competition phase testing took place.

We found that the Mode C Intruder enhancement, contrary to FAA's February 1985 report, is scheduled for implementation on certain existing computers before, rather than after, the Host computer is installed. The status of En Route Metering II is uncertain. This software, described as a critical improvement with large benefits, may not be implemented because modifications to the current En Route Metering function on FAA's existing computer may make it more effective than the proposed features of En Route Metering II.

We verified that the implementation schedule of the Conflict Resolution Advisory enhancement is the same as the one provided in FAA's February 1985 response. However, FAA did not describe problems with the generation of false alerts produced by an existing function critical to the proper operation of Conflict Resolution Advisory--referred to by FAA as "conflict alert." These problems still exist.

CONCLUSIONS

In our opinion, a decision to acquire a system as complex and important to the Nation's air traffic control system as the Host computer must be made on the soundest possible basis to assure that the system will meet FAA's operational needs.

Our review did not disclose persuasive reasons why realistic performance testing should occur only after the production and selection decisions have been made. FAA's fundamental reason for the need to meet the July 1985 vendor selection milestone is to install the Host computers in time to avoid air traffic delays caused by the limited capacity of its current computers. However, important assumptions used by the model to forecast delays caused by computer capacity shortages have not been verified. Further, FAA's recent projections of significant aircraft delays caused by computer capacity shortfalls have shown that estimated delays have decreased--from a projection that they would occur at up to eight centers by 1989 down to two centers for that same year.

FAA's testing of proposed systems has not provided convincing evidence that either vendor's proposed system can meet operationally realistic workloads. Further, its actions to compensate for this less-than-realistic performance testing do not provide adequate assurance that the system to be selected will perform as needed. DOT's and FAA's technical oversight of the testing program has been limited by incomplete and untimely documentation and untimely release of technical information and test data to its test and evaluation staff.

It appears that completing the performance testing on both vendors' systems before production and providing technical oversight would yield better system performance information for the production decision without greatly extending the overall program schedule, particularly in light of the lack of precision in FAA's air traffic delay-projections and the availability of a more realistic workload tape in June 1985.

The House Committee on Appropriations directed FAA to fully test Host computer system performance before the production decision was made. FAA assured the Committee that sufficient testing would be performed. On the basis of our review of the performance testing in the design competition phase, we believe that FAA has not provided such assurance.

While the Congress expected FAA to conduct realistic performance testing of the proposed Host computers before the production decision, in our opinion, the FAA design competition testing did not satisfy this expectation. However, other factors may be relevant in determining whether to defer the production decision, extend the design competition phase, and perform more realistic performance testing on the proposed Host computer systems. Principal among these other factors is evidence offered in the vendors' responses to the acquisition phase request for proposals on how their systems would meet FAA's future operational needs. These factors were not included in the scope of our review because they are not currently considered as part of the basis for the production decision.

RECOMMENDATIONS TO THE SECRETARY OF TRANSPORTATION

To ensure that FAA's operational requirements for the Host computers are met, we recommend that the Secretary consider the merits of deferring the production and vendor selection decisions for the Host computers, extending the design competition phase, and performing more realistic performance tests on both vendor systems. The Secretary should consider (1) the uncertainties associated with the proposed Host computers' ability to support operationally realistic workloads, (2) the questionable precision and reliability of FAA's computer model to project near-term air traffic delays caused by current computer capacity shortages, and (3) the apparent lack of significant near-term air traffic delays associated with FAA's current computers. Our review did not disclose persuasive reasons why realistic performance testing should occur only after the production and selection decisions have been made. We believe that the Secretary may have to weigh other evidence together with our review findings, in arriving at the decision to proceed with or defer the production and selection decisions.

If the Secretary decides to proceed with these decisions without realistic performance testing, notwithstanding the above-mentioned uncertainties, then we recommend that the Secretary, in advance of proceeding, provide the appropriate congressional

committees with the Department's views and the support for these views, particularly on the performance uncertainties.

MATTERS FOR CONSIDERATION BY THE CONGRESS

We believe that if the Secretary decides to proceed without realistic performance testing and without adequate explanation of the urgency of proceeding, the Congress should consider directing the Secretary to defer the contract award.

- - - - -

As requested by your office, unless you publicly announce its contents earlier, we plan no further distribution of this report until 30 days from its date of issuance. We will then send copies to the Secretary of Transportation, the FAA Administrator, and other interested parties, and will make copies available to others upon request.

Sincerely yours,

Comptroller General
of the United States

RESULTS OF OUR REVIEW

BACKGROUND

Today, air traffic controllers use two types of computers made by the International Business Machines (IBM) Corporation--9020A and 9020D--to process radar surveillance and flight plan data and to help them manage air traffic at FAA's 20 air traffic control centers. IBM 9020A computers are located at 10 sites, and IBM 9020D computers are at the other 10 sites.

FAA has stated that these computers, acquired in the 1970s, will not be able to handle increased air traffic and planned software additions without delaying traffic. Thus, FAA is developing and acquiring Host computers to replace all IBM 9020s. FAA envisions that, compared with the IBM 9020, the Host computer will be larger and more reliable and will be able to carry out new air traffic control functions designed to enhance safety and efficiency.

FAA plans to modify IBM 9020 software to operate on the Host computer--a procedure known as "rehosting." Software redesign will take place in the 1990s when the Advanced Automation System is implemented. The Advanced Automation System will introduce new controller workstations, processors, and local communications networks. In addition, it will provide the vehicle for implementing new automated air traffic control functions.

FAA is introducing the Host computer in two phases: the design competition phase and the acquisition phase. During the first phase, two vendors--IBM Corporation and Sperry Corporation--were required to submit designs for rehosting the IBM 9020 software and essential support software. They then attempted to confirm performance claims by running the rehosted software on their respective computer systems at the FAA Technical Center in New Jersey. One of these two vendors will be selected for the production contract in July 1985 and will enter the acquisition phase. During this phase, the remaining IBM 9020 non-essential support software will be rehosted, and the new Host computer systems will be installed at the 20 centers, as well as 4 other locations. Each vendor is required to propose a system from a family of software-compatible products so that, should the need arise, the system can be upgraded without having to change the software.

DOT has designated the Host computer program as a major system acquisition. As such, the program progresses through several key milestones. At each milestone, the Deputy Secretary and TSARC[1] review FAA's basis for proceeding to the next stage. In June 1985,

---

[1]TSARC includes the Deputy Secretary, the Assistant Secretaries, and the General Counsel.

the Deputy Secretary and TSARC will determine if FAA has adequately justified the need for the Host computer and has sufficiently addressed system performance issues. If the Deputy Secretary believes FAA has addressed these issues satisfactorily, he will approve the Host computer for production. Then, in mid-July 1985, DOT's designated Source Selection Official (in this case the Deputy Secretary) can select the production vendor for the Host computer.

FAA'S PERFORMANCE TESTING OF HOST COMPUTERS WAS CONDUCTED UNDER UNREALISTIC CONDITIONS

FAA's testing of IBM's and Sperry's proposed systems was not done in accordance with established criteria for acquiring major systems. Both the system performance test and the environment in which the test was conducted were significantly less demanding than the operational capabilities which the Host computer must provide. Consequently, neither vendor's system may provide the performance needed. Experience has shown that proceeding to production without such assurance could lead to the installation of systems that do not provide the needed operational capability and to costly upgrades to achieve such capability.

Office of Management and Budget and DOT guidelines provide criteria for reducing performance risk

Office of Management and Budget Circular A-109 and DOT Order 4200.14B provide guidelines for acquiring major systems. These guidelines state that full production may be approved when an agency has (1) reaffirmed mission needs and program objectives and (2) addressed system performance issues adequately. Regarding system performance issues, the DOT order states that, before a production decision is made, realistic system performance tests should be conducted in the operational environment or in an environment which realistically represents the operational conditions in which the system is expected to perform. The guidance is intended to ensure that the decision official, in this case the Deputy Secretary, will have sound and objective information to determine whether to proceed to the production phase.

The importance of conducting realistic tests was reinforced during 1984 and 1985 appropriations process. The House Committee on Appropriations directed FAA to fully test Host computer system performance, before making a production decision, to ensure that operational requirements could be met or exceeded. In particular, the Committee was concerned (1) about whether FAA's tests would sufficiently explore potential software performance problems and (2) about the tests' impact on overall system performance, including reliability, especially under peak air traffic control

conditions of the 1990s. The Committee stated that it would consider deferring the Host production contract if its concerns were not addressed. A Transportation Systems Center study prepared for TSARC also pointed out the need for performance testing which would be representative of future heavy workloads.

FAA did not conduct a realistic performance test of the Host system

FAA used a system performance test which represented the air traffic environment in the early 1970s. The test did not represent the mix of controlled and uncontrolled aircraft, types of aircraft, or routing patterns in which the Host computer would operate today or in the future. For example, the performance test had a peak aircraft tracking workload of approximately 450 aircraft per air traffic center. FAA expects the Host computer to track as many as 600 aircraft at certain centers in the 1990s. The 600-track estimate was derived from FAA's annual aviation forecast. Key automated operational functions, such as those which aid the controller in detecting potential mid-air and ground collisions, were not active during the system performance test.

Because of limited capabilities, FAA's Technical Center could not simulate a realistic air traffic control center environment. For example, while tests at the technical center were limited to 12 controller positions, the actual controller workload at some centers can include as many as 60 positions. In addition, the technical center could not simulate a communications and input/output workload that included a realistic number of radar input channels and radar processing and a realistic amount of communication with other air traffic control centers. FAA officials explained that time constraints and costs prevented them from testing in a more realistic air traffic control environment. Advanced Automation Program officials believe this realism is not needed to adequately evaluate vendors' proposed systems before production. They said that, in their view, the expense in time and money to create a more realistic test environment was not worth the perceived benefits.

Consequences of inadequate information on system performance

Proceeding to production without complete and accurate information on system performance significantly increases risk. If FAA and DOT proceed to production without accurate information on system performance, FAA could acquire a system which may not handle the 1990s workload, and it would have to spend more time and money to field such a system.

We have seen the effects of inaccurate system performance data at other agencies. In 1984, we reported[2] that the Air Force made its production and selection decisions regarding the Phase IV Base Level Computer Replacement Program without adequately testing the workload. The Air Force is now assessing how much it will cost to upgrade this system to meet operational requirements. According to the Air Force, Phase IV upgrades may cost almost $200 million.

Similarly, the Army Audit Agency reported[3] that the Army did not conduct representative workload testing before awarding a contract to automate its base operations. Current computer usage is over twice that which was portrayed during testing. The Army indicated that this oversight has affected its ability to meet mission needs; computer capacity planned for mobilization contingencies must now be used to meet day-to-day requirements.

FAA'S PERSPECTIVE ON ITS PERFORMANCE TESTS AND OUR EVALUATION

FAA recognizes that it conducted its system performance in an unrealistic environment. Advanced Automation Program officials explained, however, that they have taken actions which will show that the chosen system can meet its operational needs. For example, FAA estimated the minimum computer processor speed required in millions of instructions per second and required vendors to demonstrate the capability to upgrade to the next size processor if this estimate is low. These requirements for minimum processor speed and processor upgrade were included in both requests for proposals for the design competition and acquisition phases. We found that FAA's approach does not adequately address response time, a primary indicator of system performance. FAA acknowledged that it cannot be certain that response time--the elapsed time between the end of an inquiry or demand on a data processing system and the beginning of the response[4]--at higher traffic loads can be determined without additional tests. FAA plans to conduct a more realistic performance test after the production vendor is selected to ensure that the system fielded can handle the 1990s workload.

---

[2]Air Force Progress in Implementing the Phase IV Base Level Computer Replacement Program (GAO/IMTEC-84-7, Jan. 18, 1984).

[3]Army Audit Agency Reports: HQ 80-204, Feb. 6, 1980; and SW 84-200, Feb. 10, 1984.

[4]Processor speed is a measure of the capability of one central component of a computer system. Response time is dependent on all hardware and software components working concurrently.

During the design competition phase, FAA required the vendors to perform their demonstrations on a computer having a minimum processor speed of 4.3 millions of instructions per second. FAA arrived at this estimate by analyzing computer processor use at its centers during 1980 and 1981. FAA concluded that the relationship between computer use and air traffic workload was linear, i.e., an increase in the number of aircraft tracked would be accompanied by a proportionate increase in such utilization. FAA used this linear relationship to project the processor speed needed to track a workload of 600 aircraft. It then adjusted this speed to account for planned functional enhancements and support functions. Finally, FAA doubled the total because it believes that real-time systems should not operate at greater than 50-percent processor utilization.

During the performance test, processor utilization on each proposed system was recorded at various aircraft tracking levels up to about 450 aircraft. The recorded processor utilization and other forms of analyses will form the basis of vendor responses to the acquisition phase's requests for proposals. The vendors' proposals must address how their processor can track 600 aircraft without exceeding a specified level of processor utilization. Thus, while FAA may receive additional evidence of system performance as part of its proposal evaluation process, such evidence will not be used to support the production decision. An official from the Advanced Automation Program Office acknowledged the difficulty of extrapolating response time to predict system performance at the 600-track level, because of the smaller workload used in the performance test. The results of this test cannot be disclosed because FAA considers the release of this information as compromising the vendor selection process.

In evaluating FAA's actions, we have serious concerns about using computer processor speed as the principal basis for estimating the minimum performance needed by Host computers. Computer processor speed is only one measure of a computer system's performance capabilities. In our opinion, response time is a more important measure of how the entire computer system--processors, peripheral devices, communications, and software--works to meet the demands placed upon it. Therefore, we question whether an upgrade in the speed of the computer processor can fully solve all potential system performance problems.

The Federal Computer Performance Evaluation and Simulation Center reported to FAA in 1981 that response time is one of the best measures of system performance. In its report, the center pointed out that when system performance problems, especially those involving response time, occur in actual heavy workload situations, resolution may require upgrading of more than the processor, e.g., input/output channels, memory, and software.

Similar concerns were expressed by contractor staff supporting FAA's Operational Test and Evaluation staff in its review of design

competition phase testing. In its study objectives, the contractor proposed to address the validity of using processor speed as an indicator of system performance. Specifically, the contractor stated that complex functional software interactions and software performance problems are more likely to occur under workload conditions of the 1990s and that correcting such problems could require more than increasing the processor speed. As a result, the contractor was concerned whether FAA's strategy of planning to upgrade to a faster processor was appropriate to solve potential performance problems.

In addition to our concern that response time at heavy loads was not determined, we question FAA's use of extrapolation to determine minimum processor speed. Extrapolating results from linear regression analysis to estimate computer performance based on processor utilization is risky. Knowledgeable sources also have reservations about using extrapolation to project computer utilization at higher tracking levels. For example, a technical assessment of FAA's Advanced Automation Program performed by the Transportation Systems Center for TSARC,[5] stated that, at higher aircraft tracking levels, computer processor utilization would most likely be nonlinear, i.e., processor utilization would rise disproportionately faster than the number of aircraft being tracked. In addition, an FAA study of computer utilization at air traffic control centers cautioned extending extrapolated results for a specific IBM 9020 computer to another computer. As discussed earlier, derivation of the minimum Host computer speed was performed by extrapolating IBM 9020 results.

FAA plans to conduct a realistic system performance test after the production vendor is selected. FAA stated that this test was not performed earlier because of the lack of a workload tape capable of representing the 1990s air traffic scenario. This new test will be conducted as part of a final software test performed at FAA's Technical Center from January to April 1986. An Advanced Automation Program official agrees with us that running this test before the production decision would have made it more useful, but that the Host computer's ambitious acquisition schedule did not permit it during the design competition phase. A workload tape needed to run this more realistic system performance test will be ready by June 1985 and will be used in the acquisition phase. This official also said that delaying the design competition phase to perform more realistic testing would affect other modernization activities, such as the Advanced Automation System. However, he could not provide documentation of this effect.

---

[5]Technical Assessment of the Federal Aviation Administration's Advanced Automation Program, Transportation Systems Center, Dec. 1982.

DOCUMENTATION OF TEST PLANS AND RESULTS AND TECHNICAL OVERSIGHT OF HOST PERFORMANCE TESTING WERE INADEQUATE

Although required to do so, FAA has not adequately documented its actions in planning, monitoring, and appraising the testing of vendor systems. In addition, we found some limitations in technical oversight of performance testing. Consequently, the usefulness and confidence in test results are reduced.

Office of Management and Budget and DOT criteria state that management controls, such as documentation and program monitoring, should be provided. Specifically, agencies should formally document their system testing and evaluation activities. DOT guidance further states that the Deputy Secretary of Transportation and TSARC should monitor the Host computer acquisition to ensure that system performance risk is being adequately reduced. In addition, FAA Order 1810.1 and 1810.2 require FAA's Operational Test and Evaluation staff to evaluate the operational readiness of the Host computer before the production decision and report their results to FAA's Deputy Administrator.

We found that FAA did not prepare a plan that describes the objectives, scope, timing, and criteria for tests to be conducted during the design competition phase. The Director of the Advanced Automation Program Office recognized that formal documentation for some Host computer program activities did not exist. He explained that such formal documentation, although important, would have placed a heavy burden on his staff, given the Host computer's ambitious acquisition schedule. Rather, he established procedures which he believed provided insight into the vendors' progress. For example, he required FAA to use the quarterly program review process to provide comments to the vendor, such as indicating whether vendor system design and corrective actions were responsive to the agency's engineering requirements. However, we believe that without documentation of the agreements reached between FAA and the two vendors, the effectiveness of FAA's management control cannot be fully evaluated.

Similarly, FAA did not document its plan for evaluating vendor test reports produced during the design competition phase. Nor does it plan to document its analysis of the vendor tests until after the production vendor has been selected. This analysis should have been available to both TSARC and the Source Selection Official before vendor selection. The Director of the Advanced Automation Program Office explained that these actions were not taken because of time constraints and the possibility of compromising the vendor selection process by disclosing FAA's analysis of the vendor tests.

We also found deficiencies in FAA's technical oversight of the performance testing. On the basis of discussions with staff supporting the TSARC and the Operational Test and Evaluation staff, we

are concerned that the depth and scope of their oversight were limited regarding evaluating technical system performance issues. For example, although TSARC had expressed concern in the past over FAA's testing plans, we found that TSARC was not monitoring whether the operational workload used in system performance tests was realistic and was not analyzing test results.

An independent assessment of system performance issues is important to the objectivity of TSARC's decision concerning FAA's request to proceed to the production phase. Without such an independent assessment, TSARC will depend primarily on FAA's documentation and analysis of system performance issues. Officials from the Office of the Secretary told us that they believe TSARC has provided adequate oversight and that all information pertaining to the the Host program will be considered before making a production decision.

We also found the review performed by the Operational Test and Evaluation staff to be limited in depth and scope. Technical review of issues was limited by FAA's (1) untimely release of technical information to the Operational Test and Evaluation staff, including software modifications made by each vendor to the existing IBM 9020 software to operate on the vendor's proposed Host computer system and (2) restricting the test data made available to the Operational Test and Evaluation staff to vendor-prepared test reports.

Operational Test and Evaluation staff told us that these restrictions hampered their efforts to do an in-depth technical review of system performance. FAA officials believe that the proprietary nature of the vendor equipment and software justified their denial of access. On the basis of procedures implemented by the Advanced Automation Program Office, we believe that the vendor selection process would not have been compromised by providing controlled access to the information requested.

FAA'S SCHEDULE URGENCY IS NOT ADEQUATELY SUPPORTED

In 1982, FAA projected that up to eight IBM 9020A computer centers may experience substantial capacity problems during the mid and late 1980s. FAA based these projections on its operational delay day model. On the basis of these projections, FAA said the Host computer production decision should be made in the summer of 1985 so that the computer could be operating at all 20 centers by November 1987. This would avoid potential computer capacity problems in existing IBM 9020As. In February 1985, FAA provided to the Congress an update of its projections and stated that the projections continued to support the urgency of the Host computer schedule and the need to make production and vendor selection decisions by July 1985.

We found that the model's results do not adequately support the urgency of this schedule. As a result, FAA may not have the

proper information on how soon it needs to make production and selection decisions. First, the model's accuracy and reliability are questionable. Second, the most recent results show that the projected aircraft flight delays--FAA's primary justification for the urgency in making production and vendor selection decisions by July 1985--are not as great as those projected in 1982. In its February 1985 response to the Congress, FAA projected that, by the end of 1987, 253 occurrences of air traffic delays could be incurred nationwide. Projections based on more recent aviation forecasts show 135 occurrences by the end of 1987, or about a 47-percent reduction. Projections through 1989 show about a 28-percent decrease--from 716 to 519.

Accuracy and reliability of assumptions in operational delay day model are uncertain

FAA's projections of operational delays days are based on a model whose reliability and accuracy are uncertain. Certain key assumptions in the model, which could affect the computer saturation levels, have not been verified. For example, the model assumes that air traffic delays will occur when computer processor utilization exceeds 80 percent for more than 1 hour after turning off all non-essential functions, such as data recording and controller training. However, the 80-percent utilization assumption has never been verified. Furthermore, in our review of the Houston Center and of analyses done by the staff responsible for the model, we found that several IBM 9020A centers had experienced utilization rates above 80 percent without a delay in air traffic.

The model also uses an assumed distribution of air traffic among traffic control centers which has not been verified. The assumed traffic distribution is based upon the estimated percentages of traffic in each center during a busy day. The percentages were based on 1980 traffic distribution. FAA's most recent forecasts still rely on the percentages used in 1980. Contractor staff responsible for preparing the projections stated that there may have been significant changes in daily traffic patterns which would change the delay day forecasts for those centers.

In discussions with FAA officials regarding the use of these unverified assumptions, we were told that FAA's projections can be viewed as accurate within plus or minus 2 years. Consequently, while the model does continue to support the assumption that certain IBM 9020A centers will eventually encounter capacity problems, it lacks the precision to adequately support FAA's position that it does not have the time to conduct further testing in the design competition phase.

Air traffic delays may not be as great as FAA projections indicate

Although we have concerns about the accuracy and reliability of FAA's model, we did examine FAA's February 1985 projections

provided to the Congress. We tried to determine the significance of potential air traffic delays due to computer capacity shortages. FAA's 1985 projections show that four centers will experience operational delay days by 1989. FAA also noted, that on the average, an air traffic control center will suffer adverse impacts--called operational impact days[6]--for a 5-year period before the onset of operational delay days.[7] On the basis of its analysis of such projected delays, FAA believes that the IBM 9020 computers must be replaced in the 1986-87 timeframe.

FAA's February 1985 response gave no evidence of the 5-year adverse impacts but referred to FAA's 1982 report to the Congress regarding FAA's air traffic control computer system. The 1982 report did define both operational impact days and operational delay days. However, the 1982 report did not treat the effects of operational impact days as adversely as did the 1985 response. FAA's 1982 report also did not emphasize the importance of the 5-year onset of operational impact days as claimed in the February 1985 report. Rather, the 1982 report cited operational delay days as the most serious adverse effect and the primary justification for the urgency of the Host computer acquisition. It is on operational delay days that slower response time may occur and the distance between aircraft may have to be increased to assure safety. In cases of sustained periods of computer overload beyond an hour, FAA would finally delay air traffic to assure safety. Each air traffic control center is responsible for determining when to delay aircraft traffic.

The following table summarizes FAA's projection of operational delay days based on its model. The data represents en route delays caused by shortages in computer capacity.

---

[6]An operational impact day is a day when the computer processor utilization exceeds 80 percent for a period greater than 7 minutes. At this point, computer recording and training simulation functions may be dropped to ensure that safety and controller response times are maintained.

[7]An operational delay day is defined as a day during which the processor utilization exceeds 80 percent for a period greater than 1 hour after dropping support processing functions not critical to safety.

APPENDIX I

FAA Operational Delay Day
Forecast For IBM 9020A Centers

Center	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Albuquerque												
1	24	48	104	141	178	205	223	230	233	236	238	240
2	-	-	-	-	-	-	2	2	2	2	14	28
3	-	-	-	-	-	-	-	-	2	8	24	47
4	-	-	-	-	-	-	-	-	2	9	29	59
Boston												
1	-	-	-	-	2	4	6	11	21	30	48	56
2	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	2	9	22	43	71	99	133	175	191	214
4	-	2	10	25	43	71	104	136	178	201	219	237
Denver												
1	152	218	264	291	313	331	342	351	355	359	365	265
2	-	-	2	5	19	39	85	139	198	235	263	289
3	-	-	-	-	-	-	-	-	1	2	2	6
4	-	-	-	-	-	-	-	-	1	2	2	6
Houston												
1	114	169	219	237	247	252	257	262	273	238	302	316
2	-	-	8	21	47	75	124	160	196	225	242	248
3	12	41	69	119	176	211	229	240	244	247	251	256
4	2	9	28	59	03	159	194	219	232	242	245	248
Salt Lake City												
1	-	-	-	-	-	2	5	14	29	46	61	89
2	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-
Miami												
1	78	126	186	230	279	311	338	349	357	365	365	365
2	-	-	-	4	11	24	60	90	127	171	214	261
3	-	-	-	-	-	-	-	-	-	-	-	3
4	-	-	-	-	-	-	-	-	-	-	-	-
Memphis												
1	-	2	16	45	99	152	194	211	228	237	242	245
2	-	-	-	-	-	-	-	-	2	12	25	57
3	-	-	-	-	-	-	-	-	-	1	4	12
4	-	-	-	-	-	-	-	-	1	4	11	16
Minneapolis												
1	2	8	26	51	90	147	192	218	240	250	258	267
2	-	-	-	-	-	-	2	2	8	26	48	81
3	-	-	-	-	1	6	6	37	55	80	100	132
4	-	-	-	-		4	12	32	47	76	89	120
Oakland												
1	100	140	173	93	211	227	241	243	246	249	252	254
2	-	-	-	-	2	2	21	29	43	69	96	119
3	-	-	-	-	1	3	9	15	27	45	56	81
4	-	-	-	-	-	3	8	15	24	38	54	77
Seattle												
1	-	-	7	30	64	105	142	164	186	217	234	245
2		-	-	-	-	-	-	-	2	7	24	54
3	-	-	-	-	-	-	-	-	-	-	-	1
4	-	-	-	-	-	-	-	-	-	-	-	1

Legend

1 Provided to subcommittee in January 1982 This projection is based on 1981 aviation forecast and does not consider effects of "buying back computer capacity" by off-loading certain functions

2 Provided to subcommittee in January 1982 and again in February 1984 This projection is based on 1981 aviation forecast and considers the estimated effects of FAA's planned initiatives to "buy back" computer capacity by off-loading certain functions

3 Provided to subcommittee in February 1985 This projection is based on 1983 aviation forecast and reflects the observed effects of FAA's initiatives to "buy back" computer capacity by off-loading certain functions

4 Projections made in February 1985 based on 1984 FAA forecast data and continues to reflect the observed effects of FAA's "buy back " These projections have not been provided to the subcommittee

Comparing FAA's January 1982 projections with projections made using the most recent aviation forecast (lines 1 and 4) for all centers shows that, through 1989, the number of centers with significant delay days (greater than 4, according to FAA's 1982 report to the Congress) has dropped from 8 to 2--Boston Center and Houston Center. In contrast, FAA's February 1985 response to the Congress assumed that any operational delay day is significant and thus stated that four centers (Boston, Houston, Oakland, and Minneapolis) would experience operational delay days by 1989.

Contractor staff responsible for preparing the delay day projections told us they do not place much confidence in the Boston data. According to FAA's January 1982 projections (line 2), the Boston Center will not experience any delay days before 1995. The most recent projection (line 4) shows that this center will begin incurring operational delay days in 1985. Contractor and FAA staff responsible for the projections told us the difference in these projections was unexplainable. We discussed potential computer capacity problems with an official responsible for automation at the Boston Center. Although the most recent projection indicated that the center should be experiencing operational delay days, this official told us that computer capacity was not yet a problem and did not cause air traffic delays.

FAA's operational delay day model shows that the Houston Center should be experiencing significant operational delay days. Therefore, we inquired at this center to ascertain (1) the impact, if any, these operational delay days have had on flight delays and (2) the urgency of increasing center computer capacity. We found that the Houston Center is experiencing periods when processor utilization exceeds 80 percent. An official responsible for automation at the Houston Center told us that several actions are taken when these periods of high computer utilization occur. For example, the center can curtail low-priority computer functions, reroute traffic, and use a processor which is normally in a stand-by mode. In contrast to FAA's most recent projection, the Houston Center official told us that the center is not experiencing serious traffic delays because of computer capacity problems.

PLANS FOR FUNCTIONAL ENHANCEMENTS HAVE CHANGED SINCE FAA'S FEBRUARY 1985 CONGRESSIONAL RESPONSE

In response to the Chairman's request, we also determined the status of FAA's plans to install air traffic control "functional enhancements" (new software). We found that current implementation plans differ from those described in FAA's February 1985 response.

Mode C Intruder, En Route Metering II, and Conflict Resolution Advisory are the first three functional enhancements to be installed on the Host computer. Mode C Intruder and Conflict Resolution Advisory both enhance the existing conflict alert function. Mode C Intruder will enable controllers to detect potential conflicts between aircraft flying under visual flight rules, if equipped with a Mode C transponder, and controlled aircraft flying under instrument flight rules. Without Mode C Intruder, only conflicts between aircraft using instrument flight rules can be detected. Conflict Resolution Advisory will generate and display possible resolution advice for detected potential conflicts between aircraft. The controller currently performs the function with no computer assistance. En Route Metering II will aid controllers in managing the traffic flow into congested terminals and increasing fuel efficiency.

According to FAA's February 1985 response, Mode C Intruder and En Route Metering II will be implemented after the Host computer is installed. FAA staff responsible for these functions provided a different status. Mode C Intruder will be installed in the spring of 1986 on those IBM 9020s which have sufficient capacity. According to the staff, the Mode C Intruder feature is too important to postpone until after the Host is installed. En Route Metering II, on the other hand, may not be implemented because modifications to En Route Metering I, already in place on the existing IBM 9020s, may have surpassed the proposed features of En Route Metering II. En Route Metering II was described in FAA's February 1985 response as a critical tool with large benefits.

In its response, FAA said the schedule for implementing the Conflict Resolution Advisory enhancement has not changed. However, FAA's response does not accurately reflect the status of problems associated with the conflict alert function. According to FAA's response, Conflict Resolution Advisory development has improved the conflict alert function by minimizing false alerts. FAA recognizes that the existing conflict alert function needs modifications to reduce false alerts. However, FAA staff responsible for Conflict Resolution Advisory said that this problem has not yet been resolved.

OBJECTIVES, SCOPE, AND METHODOLOGY

Our primary objective was to respond to the Chairman's request that we determine if FAA's Host computer test and evaluation effort will provide the basis necessary for making a production decision in June 1985. Specifically, the Chairman asked us to

--examine the soundness of FAA's projections that capacity shortages in existing computers will significantly delay air traffic;

--evaluate the adequacy and completeness of the performance testing conducted on the proposed Host computers; and

--determine the status of FAA plans to install "functional enhancements" (new software) on the Host computer.

We conducted our work primarily at DOT and FAA Headquarters in Washington, D.C., and the FAA Technical Center in Pomona, New Jersey. We reviewed DOT and FAA documents on planning, management, and operation of the Host computer program. We interviewed DOT and FAA officials and technically qualified persons from private industry. We also interviewed staff from the Office of the Secretary knowledgeable about TSARC's role and concerns. We evaluated testing efforts at FAA's Technical Center where vendor demonstration tests were being conducted. We were not permitted to witness any of these tests because of FAA concerns about vendor proprietary information.

We did not evaluate all sources of delays currently experienced by the air travelers. To respond to Committee concerns, we only analyzed FAA's projections of delays caused by computer capacity shortages. To determine the accuracy of FAA's latest projections of operational delay days, we analyzed FAA's operational delay day model. We reviewed the model's assumptions and data base and discussed their validity and accuracy with contractor staff responsible for monitoring and updating the model.

We did not solicit official air traffic control center views on delays caused by computer shortages. However, we interviewed center automation officials knowledgeable of their center's computer capacity shortages and their impact on operations. We interviewed staff from FAA Headquarters and the technical center to ascertain the status of the functional enhancements. We limited our analysis of existing computers to IBM 9020A centers because these centers are projected to incur computer capacity shortages much sooner than those centers equipped with IBM 9020Ds.

Due to time constraints, we were unable to independently corroborate some of the information provided by FAA officials. In those instances, we used this information with appropriate

attribution. We also were not given timely access to FAA's technical evaluation plan, which describes how vendor test results will be incorporated into the evaluation of responses to the acquisition phase's requests for proposals. Therefore, our ability to assess the impact of design competition phase testing on this evaluation plan was limited. We performed our work in accordance with generally accepted government auditing standards. During the course of our review, we sought the views of responsible officials and incorporated their comments in the report where appropriate. As the Subcommittee Chairman requested, we did not ask DOT or FAA to review and to comment officially on a draft of this report.

(510027)

31386

AN EQUAL OPPORTUNITY EMPLOYER

UNITED STATES
GENERAL ACCOUNTING OFFICE
WASHINGTON D C 20548

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

BULK RATE
POSTAGE & FEES PAID
GAO
PERMIT No G100