Learning Lessons from Accidents and Incidents Involving Safety-Critical Software Systems

A Thesis

Presented to

the Faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Master of Science, Computer Science

by

William S. Greenwell

May 2003

APPROVAL SHEET

The thesis is submitted in partial fulfillment of the

requirements for the degree of

Master of Science, Computer Science

_____

Author

This thesis has been read and approved by the Examining Committee:

_____

Thesis Advisor

_____

Committee Chairman

_____

Committee Member

Accepted for the School of Engineering and Applied Science:

_____

Dean, School of Engineering and Applied Science

May 2003

Abstract

Accidents and incidents involving safety-critical software systems often provide lessons to the systems' users and designers, to industry, to government regulators, and to the software engineering community at large. Proper identification and documentation of these lessons is critical in order to prevent the recurrence of an untoward event. This is especially important in the domain of commercial aviation. Safety-critical software systems both onboard commercial aircraft and on the ground at airports and air traffic control facilities perform essential functions that pilots and air traffic controllers must rely upon in order to operate and direct aircraft safely. Accidents and incidents involving these systems jeopardize the safety of the national airspace system and with it the safety of passengers, crew members, and society. When an undesired event occurs, a rigorous investigation must be conducted to identify as many lessons as possible so as to prevent a recurrence of the event in the future.

Unfortunately, incidents involving safety-critical software systems are not being investigated and documented with sufficient rigor to identify and disseminate important lessons arising from them, and consequently the aviation community is missing opportunities to correct defects that could lead to future incident recurrences. This phenomenon is due in part to the manner in which software is developed and maintained and to the widespread practice of allocating investigative resources to incidents based on the extent of their associated losses.

To illustrate this point, this thesis presents a case study of two commercial aviation incidents involving safety-critical software systems. In one case, the incident lead to an accident with extensive casualties. In the other, the incident lead to a near-collision between two Boeing 747s that threatened the lives of over 400 passengers and crew members. The incidents and the investigations that ensued are analyzed, and crucial lessons are identified that

were not documented in the official incident reports or acted upon by regulatory authorities. After examining each incident, common lessons are presented to the aviation industry relating to the design of safety-critical software systems and to investigators regarding the manner in which incidents involving such systems are investigated.

In addition to reviewing the incidents themselves, the investigations into each incident are also examined and compared to illustrate the disparity in attention given to each. Despite the fact that both incidents had equally important lessons to be learned, one was investigated in much greater detail solely because it resulted in casualties and the other did not. After arguing that this loss-based incident classification scheme should be rejected, this thesis proposes an alternative scheme based on the risk associated with future incident recurrence that will more closely match investigative resources to incidents whose recurrence would likely have catastrophic consequences.

"What is the cause of most aviation accidents: usually it is because someone does something too soon, followed very quickly by too little too late."

—Steve Wilson, NTSB

"There are no new types of air crashes—only people with short memories. Every accident has its own forerunners, and every one happens either because somebody did not know where to draw the vital dividing line between the unforeseen and the unforeseeable or because well-meaning people deemed the risk acceptable."

—Stephen Barlay

"Those who cannot remember the past are condemned to repeat it."

—George Santayana

Acknowledgements

It is an honor to thank my advisor, John Knight, for sparking my interest in safety-critical software systems, for opening research opportunities to me that otherwise would have been unavailable, and for his countless contributions to my work. I also thank Elisabeth Strunk, my friend and colleague, for her contributions to this research and for being a sounding board for my ideas. I am also grateful to Michael Holloway, my co-advisor from the NASA Langley Research Center, for introducing me to the field of incident investigation and reporting and for his suggestion that important lessons are sometimes not learned in incident investigations. This work would not have been possible without the efforts of these individuals, and I indebted to them for the assistance they provided me.

I would also like to thank Peter Ladkin of the University of Bielefeld, Chris Johnson of the University of Glasgow, and others for their insightful reviews of portions of my work. Peer review is an essential step in preparing any research document, and their feedback has helped to improve the accuracy and comprehensiveness of this thesis.

Finally, I would like to thank my parents. My interest in aviation is undoubtedly due to the trips in the Bellanca my father took me on when I was a child. He also lent his experience as a private pilot to assist my research by answering my aviation-related questions. My mother and father have given me their unwavering support throughout my education, and I will always be grateful to them for everything.

This work was funded in part by NASA Langley Research Center under grants numbered NAG-1-2290 and NAG-1-02103.

Table of Contents

Table of Illustrations

Table of Abbreviations

ANSI	American National Standards Institute
AOS	Operational Support Directorate
ARTS	Automated Radar Terminal System
ASRS	Aviation Safety Reporting System
ATC	air traffic control
ATO	Air Traffic Operations
ATSB	Australian Transport Safety Bureau
CAA	Civil Aviation Authority
CAST	Commercial Aviation Safety Team
CFIT	controlled flight into terrain
CFR	Code of Federal Regulations
CRC	cyclic redundancy check
CVR	cockpit voice recorder
DOT	Department of Transportation
FAA	Federal Aviation Administration
GPS	global positioning system
GPWS	ground proximity warning system
IFR	instrument flight rules
ILS	instrument landing system
MDA	minimum descent altitude
MM	middle marker
MSA	minimum safe altitude
MSAW	minimum safe altitude warning
NAS	national airspace system
NASA	National Aeronautics and Space Administration
NFSD	National Field Support Division
nm	nautical miles
NOTAM	Notice to Airmen
NTSB	National Transportation Safety Board
OM	outer marker
RA	resolution advisory
RTCA	Radio Technical Commission for Aeronautics
STARS	Standard Terminal Automation Replacement System
TA	traffic advisory
TCAS	Traffic Alert and Collision Avoidance System
VOR	very high frequency omnidirectional radio range

1 Introduction

By their very nature, commercial aviation accidents demand our attention. Major accidents can create spectacular scenes of carnage and destruction that threaten public confidence in commercial air travel. At the very least, accidents remind us that, while very safe, there is still some risk in commercial air travel, and they often force engineers and regulators to rethink their safety analyses and add additional safeguards to the air transit system. It is out of a desire to improve safety and prevent the recurrence of tragedy that society demands investigations into accidents in an attempt to learn as many lessons from them as possible.

Learning lessons from accidents and incidents is particularly important when they involve safety-critical software systems, which are becoming increasingly ubiquitous in our society. In the realm of commercial aviation, these systems are used both onboard aircraft and at air traffic control facilities to assist pilots in operating their aircraft safely and air traffic controllers in managing the national airspace system (NAS) in a safe and efficient manner. Either in their capacity to control potentially hazardous operations or to advise human controllers via warnings and guidance when danger is present, we rely on these systems to function in a dependable fashion and not threaten our safety. If such a system falls short of its dependability requirements, the consequences could be catastrophic, and lives or property could be put at risk. Therefore, incidents involving safety-critical systems are serious occurrences. Whether or not an incident results in a catastrophe, it indicates a weakness in the systems involved and underscores the need for improving the affected systems to prevent future occurrences that could have more severe consequences. How we investigate incidents in which a safety-critical system failed to function as intended may determine whether lives or property are affected in the future by the same system behavior.

1.1 Case Study Overview

This thesis describes two commercial aviation incidents involving safety-critical software systems. The first incident involved the failure of a ground-based warning system that contributed to an accident with over 200 fatalities. The second resulted from the failure of an onboard collision avoidance system that caused two aircraft to nearly collide, jeopardizing the lives of over 400 passengers and crew members. In both incidents, the failure of the system involved caused operators to receive false or misleading information and make incorrect inferences, which either caused them to assume that all was safe when in fact something was awry or to take actions they believed were necessary to maintain safety that actually endangered the lives of passengers. When such a failure occurs, it is imperative that investigation boards conduct a thorough review of the systems involved so that aviation authorities can enact appropriate remedies to prevent the incident from recurring.

The official investigations of the incidents presented in this thesis did not examine the software systems involved with sufficient rigor, and consequently crucial lessons in software engineering were not documented or acted upon. As subsequent chapters explain, the nature of software systems makes it especially important for investigators and engineers to detect and repair defects as quickly as possible in order to prevent recurring manifestations of faults that may be present in those systems. To this end, this thesis reviews both of the incidents mentioned earlier focusing on the software failures that contributed to the incidents and then presents new lessons for the aviation community relating to the design of safety-critical software systems and the manner in which incident investigations should be conducted when software systems are involved.

After examining the incidents themselves, the investigations into each of the incidents are then reviewed and the question of why one incident received a much more rigorous investigation is posed. This discrepancy appears to have occurred simply because the first incident resulted in extensive casualties and the second did not. It is well known that incidents in which no loss is incurred are as valuable as accidents in their ability to teach lessons [19]. Even so, incidents rarely command the attention that accidents do, which creates a serious imbalance with potentially serious consequences. Indeed, the second incident described in this thesis could have easily developed into a catastrophe with almost twice the casualties as the first. In order to remedy this problem, current loss-based incident classification schemes should be rejected in favor of alternative schemes based on risk that will more appropriately match investigative resources to events whose recurrence would likely have catastrophic consequences. The fundamentals for such a scheme are proposed in this thesis.

Both of the events that are discussed could have been prevented in many ways. However, the need for change in incident classification is illustrated very clearly by the fact that both events were *preceded* by similar incidents that indicated the possibility of a systemic problem.

1.2 Thesis Organization

The remainder of this thesis is organized as follows. Chapter 2 presents background information on how aviation accidents are investigated and the role of software in commercial aviation. Chapters 3 and 4 present each of the incidents selected as case studies for this research and review each incident separately. Chapter 3 recounts the crash of a Boeing 747 jumbo jet into Guam in August 1997 and examines the failure of a ground-based minimum

altitude warning system (MSAW). Chapter 4 describes a near-collision between two Boeing 747s over Chinese airspace in June 1999 that was caused by a failure of a collision avoidance system on board one of the aircraft. Chapter 5 identifies a common lesson to the aviation community concerning the design and maintenance of safety-critical software systems and contains recommendations to investigators for improving incident investigations. It also notes the problems with investigating incidents and accidents differently simply because the latter involve loss and the former do not. Chapter 6 explores this topic in greater detail by contrasting the investigations into each of the incidents described in this thesis and illustrating the disparity in attention given to the second. Chapter 7 then proposes a new scheme for allocating investigative resources to incidents and accidents based on risk of future recurrence instead of loss. Lastly, chapter 8 contains the conclusions of this thesis.

2 Background Information

To understand why investigators are missing important lessons from accidents and incidents involving safety-critical software systems, one must be familiar with the nature of safety in commercial aviation as well as the reliance on safety-critical systems in the air transit system. This chapter begins with an overview of the organizations that regulate aviation in the United States and investigate domestic civil aviation accidents. It then discusses how safety-critical software systems are developed and employed in the air and on the ground to enhance the safety and efficiency of air travel as well as the issues surrounding them.

2.1 Safety and Commercial Aviation

The notion of safety has been prevalent in air travel since the first powered flight on December 17, 1903. Prior to their demonstration of the "Wright Flying Machine" in Kitty Hawk, North Carolina, Orville and Wilbur Wright conducted thousands of glider tests, performed wind tunnel experiments, and constructed a simulator to train pilots in the operation of rudder controls. Their airplane even featured a flight data recorder that was the predecessor of the first data recorders to be mandated for use on airliners [24].

Today, commercial air travel is one of the safest modes of transit available, and the commercial aviation industry enjoys one of the finest safety records compared to other potentially high-risk industries. According to National Transportation Safety Board (NTSB) statistics, accidents involving scheduled U.S. commercial air carriers (14 CFR 121) during 2002 were at a rate of about one accident per 300,000 departures [20]. Hull-loss accident rates among North American commercial carriers are less than half the worldwide rate [27]. Much of the industry's success is due to safety improvements implemented in response to major

accidents. These improvements include long range air traffic control service, wind shear alert systems, ground proximity warning systems, altitude alert systems, and collision avoidance equipment [25]. Indeed, the Federal Aviation Administration (FAA) itself was originally created by Congress in response to a series of midair collisions in the 1940s and 50s [26].

Despite these improvements and the rarity of commercial aviation accidents, difficult challenges lie ahead for the industry in preserving its safety record [27]. With annual domestic flights expected to increase from 66 million in 2001 to 81 million in 2013, the increase in departures will result in a corresponding increase in accidents unless steps are taken to further reduce accident rates. Airlines and government regulators will have to tackle the often conflicting goals of improving safety and efficiency simultaneously. To achieve the needed safety improvements, the aviation community will have to take a more proactive stance in identifying hazards and implementing measures to avoid them.

2.1.1 Investigation of Aviation Accidents

International civil aviation regulations require aviation authorities to appoint independent boards to investigate aviation accidents. In the United States, this role is fulfilled by the NTSB. In accordance with its charter issued by Congress, the NTSB "investigates every civil aviation accident in the United States" [28]. As defined by the NTSB and FAA, an accident is "an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and until such time as all persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage" [19]. In 2002, there were 1,820 such accidents, 360 of which were fatal, mostly involving general aviation (non-commercial) aircraft.

Because the NTSB cannot fully investigate every civil aviation accident, it must prioritize accident reports and investigate them accordingly. In order to allocate its investigative resources efficiently, the NTSB classifies each accident report it receives into one of the four categories defined in Table 1. In the table, the term "Part 121 aircraft" refers to aircraft with

Category	Definition
Major	An accident in which any of three conditions is met: • a Part 121 aircraft was destroyed; • there were multiple fatalities; or • there was one fatality and a Part 121 aircraft was substantially damaged.
Serious	An accident in which at least one of two conditions is met: • there was one fatality without substantial damage to a Part 121 aircraft; or • there was at least one serious injury and a Part 121 aircraft was substantially damaged.
Injury	A nonfatal accident with at least one serious injury and without substantial damage to a Part 121 aircraft.
Damage	An accident in which no person was killed or seriously injured, but in which any aircraft was substantially damaged.

Table 1: NTSB Accident Categories and Criteria [21]

10 or more seats. Figure 1 shows annual U.S. commercial aviation accidents by NTSB classification from 1983-2002. While rates of major and serious accidents have remained fairly constant, those of less severe accidents have risen dramatically as domestic flights have increased.

An accident's classification determines the extent to which it is investigated by the NTSB. The Board prepares a synopsis of every accident it investigates containing a factual account of the accident sequence and a brief analysis describing the probable cause. If safety

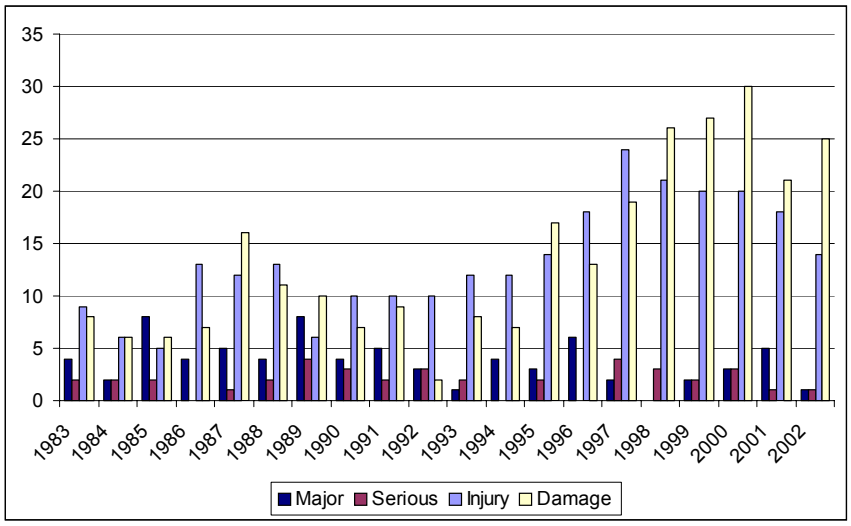

Figure 1: U.S. commercial (14 CFR 121) aviation accidents by NTSB classification, 1983-2002 [21].

deficiencies are identified, recommendations to the FAA, equipment manufacturers, operators, or other entities involved in the accident may also appear in the synopsis. For major accidents, the board prepares a much more extensive report, often spanning hundreds of pages, and occasionally holds public hearings to gather testimony from witnesses and experts. Accident synopses are stored in a public database, and reports for major accidents are made publicly available as well.

2.1.2 Investigation of Aviation Incidents

The FAA and NTSB draw a distinction between accidents and incidents. Specifically, an incident is defined as "an occurrence other than an accident associated with the operation of an aircraft, which affects or could affect the safety of operations" [19]. Except under limited circumstances, the NTSB does not investigate incidents. The FAA collects reports of civil aviation incidents and stores them in a database and may investigate selected incidents at its discretion. Informal incident reporting mechanisms also exist, such as the Aviation Safety Reporting System (ASRS) maintained by the FAA and NASA. ASRS is a voluntary program

in which aviation professionals including pilots, air traffic controllers, flight attendants, mechanics, and others submit reports of incidents where safety was compromised [29]. ASRS staff use these reports to prepare monthly safety bulletins and to identify immediate safety hazards that should be reported to the FAA. As Johnson notes, however, incident reporting systems such as ASRS have limitations and tend to focus on short-term fixes to safety problems rather than addressing underlying issues [30].

2.2 Safety-Critical Software in Aviation

Software systems are relied upon extensively in the aviation industry to advise pilots and air traffic controllers and to control certain aspects of flight. Over time, they have become an integral part of the national airspace system (NAS), and as the FAA proceeds with its modernization plans for the NAS, these systems will become ever more ubiquitous. As they take on more complicated tasks, the likelihood that such a system will contribute to an accident or incident will rise. The FAA treats these systems differently depending on whether they will be used aboard aircraft or on the ground at airports and air traffic facilities.

2.2.1 Airborne Systems

A typical commercial flight today is flown almost entirely by autopilot, with only the most delicate tasks of takeoff and landing performed manually. The pilot's instrument panel on a jumbo jet, once occupied by numerous analog gauges and dials, is now dominated by two computer displays, one showing electronic versions of essential flight instruments and the other depicting the aircraft's heading, course, and nearby air traffic. With the advent of global positioning system (GPS) navigation, autopilots can fly all but the last few miles of a transcontinental flight with little intervention from the pilot. The duty of the pilot has shifted from

keeping his aircraft aloft and on course to programming and monitoring the onboard avionics systems and interacting with air traffic control (ATC). Even in this new role, the pilot receives a great deal of help from software systems. Ground proximity warning systems (GPWS) track the aircraft's position and alert the pilot of threatening terrain. Collision avoidance systems monitor nearby aircraft and issue advisories when they detect traffic conflicts. Other systems assist the pilot in computing cruise and approach speeds and course corrections due to weather or turbulence.

Until the 1980s, large commercial aircraft, despite their reliance on computer systems for many of their functions, still provided pilots with direct, analog interfaces to flight control surfaces and engines. Even if the computer systems were to fail, pilots could still fly using the manual controls and backup analog instruments. Then, in 1984, Airbus Industrie launched the A320 commercial passenger jet, the first of its class to employ a "fly-by-wire" cockpit in which analog flight controls were replaced with digital electronic ones. Pilot commands were no longer sent to control surfaces via direct mechanical, pneumatic, or hydraulic links but were instead transmitted as signals to a computer, which would stimulate the necessary control surface actuators to execute the commands. Moreover, in order to prevent pilots from exceeding the operational envelope of the A320, Airbus added constraints to the computer logic that would cause the computer to refuse pilot commands that could compromise safety. This flight envelope protection system is suspected of contributing to fatal accidents, although Airbus insists these accidents were due to human error [31]. Since the deployment of the A320, Airbus has used its fly-by-wire design on each of its subsequent models. Boeing has developed its own fly-by-wire design, albeit with a more lenient flight envelope protection system, that debuted on its 777 jetliner.

Each of the systems described above relies on software for at least some of its function, and each could contribute to a hazardous situation if it were to fail. To reduce the chance of failure, the FAA requires all software intended for use onboard aircraft to comply with a set of guidelines known as DO-178B. These guidelines specify the manner in which avionics software may be developed and require developers to pass certain testing standards to show that their software will perform correctly and not interfere with other systems. Compliance with DO-178B is not a proof that a piece of software is fault-free, however, nor does it imply that the software will function as intended when it is put into operation.

2.2.2 Ground-based Systems

Software systems on the ground at air traffic facilities assist controllers in tracking and managing aircraft. The most prominent of these is the Automated Radar Terminal System (ARTS), which is installed at each of the FAA's en route and approach control centers. ARTS was originally deployed during the late 1960s and early 1970s and has since undergone numerous upgrades and enhancements. The most recent version, Common ARTS, was written in ANSI C and runs on commercial off-the-shelf hardware. The system can support 10,000 simultaneous tracks across 200 controller displays [32]. In addition to its basic function of processing and displaying radar returns, the system includes conflict alert and minimum safe altitude warning (MSAW) functions that alert controllers when it detect traffic conflicts between tracks or when a tracked aircraft is flying too low, respectively.

In use for over 30 years, ARTS has become antiquated and expensive to maintain. The FAA currently manages five different versions of ARTS, three of which still use monochrome displays. In 2002, the FAA began replacing ARTS installations with the Standard Terminal Automation Replacement System (STARS). According to the FAA, this system features color

displays, a windowing operating environment, and fault tolerance mechanisms including an Emergency Service Level that "will get critical information to the controllers, even in the event of a total hardware or software failure in the primary system" [33]. By 2008, the FAA expects to have deployed STARS at over 300 facilities.

As of 2002, STARS was four years behind schedule and $760 million over budget. To compensate for the delay, the FAA announced it would defer independent testing of STARS until after it had already been deployed at major air traffic centers. This announcement drew criticism from many, including others within the Department of Transportation (DOT). In a memorandum to the FAA administrator dated June 3, 2002, the Inspector General of the DOT wrote, "Independent testing provides the final assurance that the product is safe, effective and suitable for full-time use in the real world. We have serious reservations about declaring STARS 'operational' before FAA satisfactorily completes its standard Independent Operational Test and Evaluation" [34]. The memorandum identified 221 open "critical" trouble reports concerning STARS, which it described as "those that would prevent or preclude the performance of a mission, jeopardize safety or security, or adversely affect a mission-essential capability." According to the memo, when the DOT approached FAA officials with these concerns, they were told that the FAA would focus on the "truly critical" reports and that in doing so would be able to commence operations with a system that was "not perfect but acceptable." This response did not satisfy the Inspector General, who described the FAA's distinction between "critical" and "truly critical" and the notion of "not perfect but acceptable" as vague, particularly for an air traffic control system. Nevertheless, the FAA has proceeded with its deployment schedule for STARS.

The FAA currently does not have standard certification guidelines for ground-based software systems. DO-178B applies only to onboard systems, and no companion document exists yet for software used at airports and air traffic control centers. RTCA, the organization that developed DO-178B, is working to fill this void with DO-278, "Guidelines for Communication, Navigation, Surveillance, and Air Traffic Management (CNS/ATM) Systems Software Integrity Assurance," which mirrors DO-178B in its classification of software assurance levels based on the likelihood of an occurrence and the severity of the consequences. DO-278 is still in development, however, and has not yet been adopted as official guidance by the FAA.

2.3 Summary

Modern commercial aviation is extremely safe, thanks in large measure to the safety improvements put in place in response to past accidents. Accident investigations have been crucial in effecting technological and regulatory improvements in order to enhance safety aboard aircraft, at airports, and throughout the air traffic control system. Many of the technological improvements in these areas involve the use of software. As the reliance on software to ensure safety increases, so does the probability that software failure will contribute to an accident or incident. With software systems shifting from an advisory role to one of actually controlling operations, such as with the fly-by-wire design employed by Airbus, the consequences of failure could be catastrophic. The following chapters describe two cases in which such failures have already occurred and illustrate the need to alter the manner in which incidents involving software systems are investigated. By doing so, the aviation community will be able to learn more from these incidents and correct software problems before they contribute to additional occurrences.

3 Korean Air Flight 801

On August 6, 1997 at about 1:42 am Guam local time, Korean Air flight 801, a Boeing 747-300, crashed into Nimitz Hill, Guam while attempting a nonprecision approach to runway 6L at A.B. Won Guam International Airport. Of the 254 persons on board, 237 of which were passengers, only 23 passengers and 3 flight attendants survived. The National Transportation Safety Board (NTSB) investigated the accident and classified the crash as a controlled-flight-into-terrain, or CFIT, accident. During its investigation, the NTSB found that a ground-based minimum safe altitude warning system (MSAW), designed to alert air traffic controllers of aircraft flying too low, had been inhibited. In its final report, the NTSB concluded that the crash was largely due to pilot error, but also noted:

> Contributing to the accident was the Federal Aviation Administration's (FAA) intentional inhibition of the minimum safe altitude warning system (MSAW) at Guam and the agency's failure to adequately manage the system [1].

Despite its finding that the inhibition of the MSAW system at Guam was a contributory factor, the NTSB did not issue any safety recommendations to the FAA pertaining to the MSAW system in response to this accident.

3.1 Background Information

3.1.1 Guam Approach Procedure

Korean Air flight 801 crashed during its final approach to runway 6L at Guam International Airport while operating under instrument flight rules (IFR). Most scheduled commercial flights operating under IFR in U.S. airspace use a collection of navigation aids known as the instrument landing system, or ILS, in making their landing approaches. ILS was designed to allow aircraft to land under poor visibility conditions and consists of a localizer, glideslope,

marker beacons, and special runway lighting. The localizer and glideslope are radio navigation aids that help the pilot maintain the proper descent path. The localizer assists in maintaining horizontal alignment with the runway while the glideslope aids in controlling the aircraft's altitude and descent rate. In addition, marker beacons are typically placed at key points on the approach path to alert the pilot as the aircraft passes over them. Most ILS approaches have an outer marker (OM), which is usually placed where the aircraft is expected to intercept the glideslope descent path, and a middle marker (MM), which is typically positioned where the aircraft is expected to be 200 feet above runway elevation. Some ILS approaches also employ an inner marker that is placed where the aircraft is expected to be at *decision height*—the altitude at which the pilot must be able to see the runway to continue the approach. Flying over each marker causes distinct visual and aural indications in the cockpit, allowing the pilot to verify the aircraft's position on the approach path. This approach sequence, in which the pilot uses the localizer, glideslope, and marker beacons to perform the approach, is known as a full ILS or precision approach [8].

Approach procedures are published in aeronautical charts called *approach plates*. Each approach plate depicts procedures for a particular approach to a particular runway, for example the ILS approach to runway 6L at Guam. Approach plates are organized into a top-down *plan view*, which assists pilots in intercepting the approach path, and a *profile view* that guides pilots along the approach path once they are established on it. Figure 2 contains the

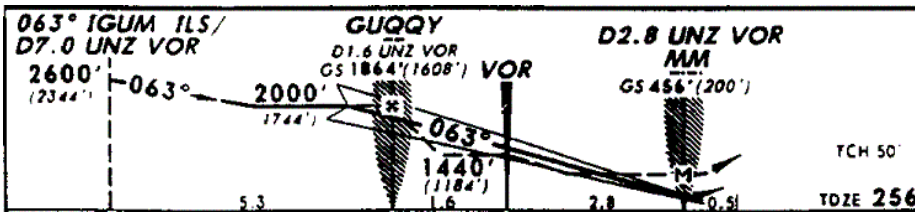

Figure 2: Profile view of the Guam ILS runway 6L approach plate as of 8/2/1997. NOT TO BE USED FOR NAVIGATION.

profile view for the Guam ILS runway 6L approach plate that was in effect at the time of the accident. Reading left-to-right, the procedure directs pilots flying the approach to maintain at least 2,600 feet until seven nautical miles (nm) from the point labelled "VOR." Pilots may then descend to and maintain 2,000 feet until becoming established on the localizer and glideslope, which must occur before crossing "GUQQY"—the outer marker for this approach. Assuming they have intercepted the localizer and glideslope by this point, pilots would then follow the glideslope descent path down toward the runway, crossing the middle marker as they approach decision height. Upon reaching decision height, pilots would either fly the remaining portion of the approach visually if they could see the runway or abort the approach, climb to a safe altitude, and either attempt the approach again or land at an alternate airport.

At the time of the accident, the FAA had issued a Notice To Airmen (NOTAM) for the Guam airport stating that the runway 6L glideslope was out of service, meaning that pilots were not to rely on the glideslope signal when landing at Guam. The flight crew of KA 801 received this notice both prior to departure and again from air traffic control during their approach to Guam. When the glideslope is unavailable, it is still possible to perform a non-precision or localizer-only ILS approach. The nonprecision approach procedure is published alongside the precision approach as a sequence of step-down altitudes. In lieu of a glideslope, pilots make a series of intermediate descents using these step-down altitude fixes. The non-precision approach procedure for the Guam ILS runway 6L approach is represented as a dashed line in Figure 2. The procedure is identical to the precision approach until crossing GUQQY. Instead of following the glideslope descent path, pilots flying the nonprecision approach would step down to 1,440 feet after crossing GUQQY and maintain this altitude until passing the point labeled "VOR." After crossing the VOR, pilots could descend to the

minimum descent altitude (MDA), which is the minimum height to which pilots may descend without having sight of the runway. MDAs vary according to airspeed and are printed in a table elsewhere on the approach plate. If the runway was not in sight after crossing the Missed Approach Point (denoted by a capital "M" on the approach plate), which for the 6L approach is the middle marker, pilots would be required to abort the approach. Otherwise, the remainder of the approach would be flown visually as with the ILS approach.

3.1.2 KA 801's Final Approach

Postaccident analysis of radar data indicates that flight 801 began a premature descent on its nonprecision approach and violated the 2,000 step-down clearance approximately 1.9 nm short of GUQQY. The aircraft proceeded on a steady descent, violating the 1,440 step-down clearance before impacting terrain adjacent to the VOR approximately 3.3 nm short of the runway threshold. Although it is impossible to fully assess the captain's state-of-mind while flying the approach since he and the other members of the flight crew perished in the accident, the NTSB theorized based on its analysis that he might have believed the VOR was collocated with the runway, even though the approach plate clearly places it 3.3 nm short of the runway threshold. Some dispute this theory, however, noting that it would have been necessary for the captain to ignore an over 400-foot discrepancy between the altitude of the VOR and that of the runway to make this error [9]. Cockpit voice recorder (CVR) transcripts also indicate confusion on the part of the captain as to the status of the glideslope. In the transcript, the captain repeatedly asks whether the glideslope is functioning, even though the NOTAM should have prompted him to disregard it entirely.

In its report, the NTSB concluded, "the captain lost awareness of flight 801's position on the [ILS] localizer-only approach to runway 6L at Guam International Airport and improp-

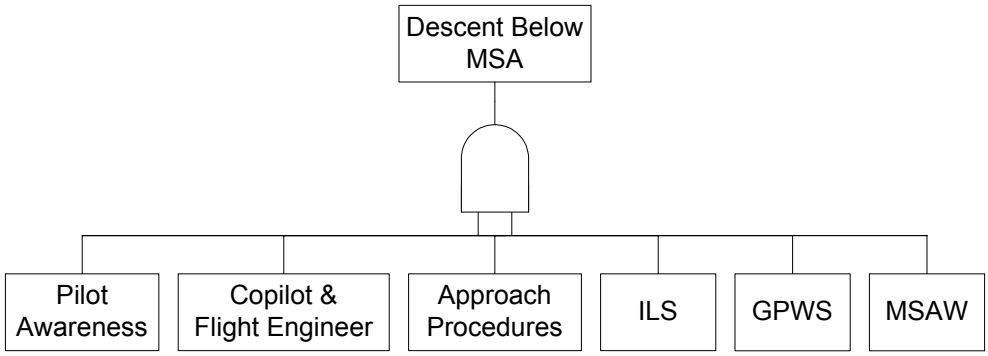

Figure 3: Barriers to controlled flight into terrain.

erly descended below the intermediate approach altitudes...which was causal to the accident." The NTSB classified the accident as a controlled flight into terrain (CFIT).

3.1.3 Controlled Flight Into Terrain

In a statement released March 26, 2001, the FAA's Commercial Aviation Safety Team (CAST) cited controlled flight into terrain as "the leading cause of fatal commercial air accidents worldwide" [4]. It defines a CFIT accident as one in which "a fully qualified and certified crew flies a properly working airplane into the ground, water, or obstacles with no apparent awareness by the pilots." Most CFIT accidents occur during nonprecision approaches, and the risk of CFIT is heightened when flying "black hole approaches"—ones in which unlit terrain makes it extremely difficult to distinguish between surface features and the sky above when flying at night. Guam presented a particularly notorious black hole to pilots, and even on clear nights Nimitz Hill, which lies directly on the approach path, could disappear completely into the darkness.

Under its own initiatives and in response to safety recommendations from the NTSB, the FAA has adopted numerous systems and procedures designed to reduce the frequency of CFIT-induced accidents. These barriers are shown in Figure 3. In the cockpit, the Instrument

Landing System (ILS), comprised of the localizer and glideslope, marker beacons, and special runway lighting, provides precision guidance to the flight crew as the aircraft makes its final approach. Approach plates specify procedures for becoming established on the ILS approach as well as backup approach procedures in case the ILS approach is unavailable. In addition, an onboard ground proximity warning system (GPWS) gives aural altitude callouts as the aircraft descends and features a special callout when the aircraft reaches its decision height or MDA. Lastly, the other members of the flight crew, typically the copilot and possibly the flight engineer, monitor the pilot's approach and may challenge it if they sense trouble. On the ground, the MSAW system alerts air traffic controllers to low-flying aircraft so that they can contact the flight crews and advise them accordingly.

Under normal circumstances, each of these measures—the ILS components, the GPWS, the flight crew following approach-plate procedures with onboard instruments, and the MSAW system—serves as a *barrier* against CFIT. While individual systems might fail occasionally, an accident will be prevented if just one of the systems above functions as intended. For a CFIT-induced accident to occur, *all* of these barriers must fail to prevent the accident, and typically the probability of such a catastrophic failure is extremely small provided the systems fail randomly and independently of each other. By deploying numerous systems to serve as barriers and keeping them as independent from one another as possible, the FAA hopes to achieve greater safety in the national airspace system (NAS) in precisely this manner.

3.1.4 The Minimum Safe Altitude Warning (MSAW) System

The MSAW system is a ground-based system that alerts controllers visually and aurally when an IFR-tracked flight descends below, or is predicted to descend below, a prede-

termined minimum safe altitude (MSA). The system is software-intensive and relies upon existing sensors to provide data. The MSAW system was incorporated into the FAA's Automated Terminal Radar System (ARTS) in 1977 in response to a NTSB Safety Recommendation resulting from a December 1972 Learjet accident and was designed to address the scenario in which a flight crew had become disoriented and improperly descended below its MSA, particularly while on landing approach.

Each MSAW installation operates with a terrain database and configuration information that are tailored to the airport at which the installation is running. The system identifies low-flying aircraft by employing two monitoring techniques. General monitoring tracks all aircraft operating within the MSAW service area. For each aircraft, the system reads the maximum terrain elevation for the region in which the aircraft is operating from the terrain database and applies a 500 foot safety margin to determine the MSA for that region (although the safety margin value can be adjusted). If the aircraft has violated its MSA, the system raises an alert to air traffic controllers. Approach path monitoring, the second technique, tracks aircraft operating within specially designated rectangular regions called capture boxes where aircraft typically perform final approach maneuvers. Inside each capture box, the MSAW system simulates a glideslope descent path and can determine whether an aircraft on final approach has descended, or is projected to descend, below the desired path.

The FAA's Operational Support Directorate (AOS) is responsible for overseeing the MSAW system. During an NTSB public hearing on the KA 801 accident, the acting manager of the FAA's National Field Support Division (NFSD), AOS-600, testified that proper functioning of the MSAW system was "very important" to his organization and classified the system as a "safety-critical item" [3].

Some MSAW-equipped sites were plagued by frequent nuisance warnings generated by the system, typically triggered by aircraft that had just taken off or were about to land. In order to reduce the frequency of nuisance warnings, site adaptation managers could request that inhibit zones be added to the configuration information for their airports. All aircraft operating within these inhibit zones would be excluded from MSAW processing. These requests were approved and implemented by the FAA's Air Traffic Operations office (ATO), which would send back a rebuilt system to the site that submitted the request. FAA Order 7210.3 describes the conditions under which the MSAW system may be inhibited. At the time of the accident, the order stated:

> When their continued use would adversely impact operational priorities, facility [Air Traffic] managers may temporarily inhibit the MSAW, the Approach Path Monitor portion of the MSAW, and/or the [Collision Avoidance] functions. Except when equipment or site adaptation problems preclude these functions being used, a brief written report shall be sent to the ATD whenever they are inhibited. A copy of the report shall be sent to ATO-100 [2].

3.1.5 MSAW at Guam

According to the NTSB report, the Guam MSAW system was installed in 1990, and was originally configured to monitor an area within a 55-nm radius of the Guam radar. In March 1993, Guam air traffic managers, in conjunction with the FAA's Western-Pacific Region office and the FAA Technical Center, prepared new site adaptation parameters for the Guam MSAW system that included a 54-nm inhibit zone centered at the Guam radar site as illustrated in Figure 4. According to NTSB investigators, this change was "neither a fluke nor a malfunction but rather was an intentional adaptation change for the purpose of eliminating numerous nuisance low altitude alerts," and was put in place "for temporary use until a better solution to the problem of nuisance alarms could be found" [5].

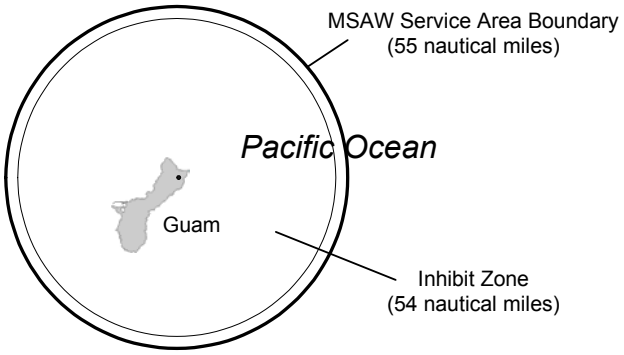

Figure 4: The Guam MSAW inhibit zone. Figure not drawn to scale.

According to testimony from the acting manager of the National Field Support Division (NFSD) of the FAA Technical Center at the time of the accident, this change effectively reduced MSAW processing to a 1-nm ring located between 54 and 55 nm from the radar facility as shown in Figure 4. No general or approach path warnings would be generated by the system for aircraft operating within the inhibit zone [3].

The new system became operational in February 1995. In July of the same year an FAA inspector conducted a biannual evaluation of the Guam facility and noted the inhibition of the MSAW system as an "informational" item, but did not recommend any follow-up action be taken. In April 1996 the FAA installed new MSAW software at Guam to update the terrain database, however this software also contained the 54-nm inhibit zone. This software remained in operation through the time of the accident. The FAA conducted another facility evaluation of Guam in May 1997, but this inspection failed to note the inhibition of the MSAW system entirely.

3.2 Postaccident Actions

After the accident, the FAA and NTSB investigators conducted a simulation of Korean Air flight 801's final approach with the MSAW inhibit zone removed. The simulation indi-

cated that, without the inhibit zone, the MSAW system would have generated visual and aural low altitude warnings for KA 801 64 seconds prior to impact. The NTSB concluded in its report that this time interval would have been sufficient for air traffic controllers to notify KA 801 and for the flight crew to take remedial action.

On August 15, 1997, the FAA announced that it had begun a review of MSAW systems nationwide as a precautionary measure. Of the 192 in-service systems, the FAA found two that were configured improperly, and reported that these systems were corrected and recertified. In addition, FAA flight inspections of 23 ATC facilities uncovered a previously unknown inhibit zone at Florence, SC. In response to these findings, the FAA developed policy to "require that MSAW be flight checked and ground certified as part of commissioning process for a new radar and periodically thereafter."

The FAA also conducted a fact-finding review of 10 ATC towers to assess controllers' knowledge of the MSAW system. The review found that most controllers possessed only a cursory knowledge of the system and gave inconsistent answers when asked about who had the authority to adapt MSAW parameters and how daily MSAW testing should be conducted if the system had been inhibited.

The review recommended that, among other things, (a) uniform site adaptation parameters should be established for MSAW operation, (b) periodic evaluations of MSAW systems should be conducted "to ensure system integrity and reliability," and (c) configuration management of MSAW software should be appropriately documented and centrally controlled. In an October 1997 briefing to the NTSB the FAA also presented new guidelines for certifying and maintaining MSAW systems to establish "strict oversight and control" over MSAW operations.

3.3 Analysis

The MSAW system was developed to address scenarios in which the onboard barriers designed to prevent CFIT accidents fail. This is precisely what happened on August 6, 1997 over Guam. The glideslope for runway 6L was out of service, and the captain lost awareness of the aircraft's position on final approach. Although the onboard GPWS gave aural altitude callouts to the flight crew as the aircraft descended and an additional callout when the aircraft reached its MDA, CVR transcripts indicate these callouts were largely ignored by the flight crew, possibly because traditional GPWS systems were known to generate nuisance messages over Guam. The "black hole" surrounding Nimitz Hill made it difficult for the captain to verify his approach visually. Lastly, the copilot and flight engineer failed to challenge the captain's approach in time to save the aircraft.

As noted earlier, the FAA had been aware of the inhibition of the MSAW system at Guam since July 1995 when an FAA inspector cited it in his evaluation of the Guam radar facility. During the NTSB hearing, the manager of the FAA's NFSD testified that no national policy was in place to prevent such a change, and that changes to MSAW configurations in general were left to the discretion of individual sites. The FAA did not provide site adaptation managers with any guidance or standards for adapting their MSAW configurations beyond that contained in Order 7210.3. Thus, sites were free to adapt their MSAW configurations as desired to reduce nuisance warnings without reassessing the system's ability to detect and report low-flying aircraft. As illustrated by the Guam case, this lack of oversight removed any guarantee that the system would function as intended, thereby degrading its ability to serve as a barrier against CFIT-induced accidents.

In its final report, the NTSB concluded that "the FAA's quality assurance for the minimum safe altitude warning system was inadequate, and the agency's intentional inhibition of that system contributed to the flight 801 accident." The NTSB did not, however, identify the underlying problems in the FAA's quality assurance process or recommend changes to the FAA's maintenance programs for MSAW or its other software systems.

Clearly the FAA's quality assurance of MSAW was inadequate. They had taken a trial-and-error approach to adapting MSAW site parameters, allowing sites to make configuration changes at their discretion to reduce or eliminate nuisance warnings with little oversight from the AOS or the ATO. Moreover, the FAA provided individual sites with no instructions for making configuration changes or guidance for reducing nuisance warnings while minimizing the extent of inhibit zones. These are merely symptoms of a deeper problem, however, and as Leveson notes, "If we only patch the symptoms and ignore the deeper underlying cause of one accident, we are unlikely to have much effect on future accidents" [6].

The underlying problem with the manner in which the FAA maintained the MSAW installations at its 193 ARTS IIA and ARTS III facilities is that it allowed changes to be made to the system without examining the effect those changes would have on the safety case the system was designed to address. The MSAW system was designed to address scenarios in which onboard altitude warning systems malfunctioned or failed to convince flight crews that they were operating at dangerously low altitudes. It stands as the only ground-based barrier against CFIT-induced accidents aside from the vigilance of air traffic controllers and enhances the overall level of safety in the NAS. While the FAA does not, in general, regard ground-based software systems as safety-critical, in his testimony at the NTSB hearing, the acting manager of the NFSD classified the MSAW system as a "safety-critical item" [3]. By allow-

ing unchecked configuration changes to be made to this system, the FAA jeopardized the level of safety the system was able to provide at each of its 193 ARTS AT facilities. In the case of Guam, these changes effectively disabled the system, removing it as a barrier against CFIT-induced accidents.

3.4 Related Incidents

The Korean Air flight 801 crash at Guam was not the only accident in which an improperly configured MSAW installation was found to be a contributory factor. Two other notable incidents, one at Dulles International Airport and another at Houston Intercontinental Airport, underscore the extent to which the FAA's oversight of the MSAW program was insufficient. These incidents are discussed below.

3.4.1 Dulles International Airport, 1994

On June 18, 1994, a Transportes Aereos Ejecutivos, S.A. Learjet crashed on final approach to runway 1R at Dulles International Airport approximately 0.8 nm short of the runway threshold. During its investigation of the accident, the NTSB found two discrepancies in the site adaptation variables used by the Dulles MSAW installation. These discrepancies caused the system to incorrectly model the location of the threshold for runway 1R and to apply the wrong MDA for aircraft subject to approach path monitoring. While the NTSB did not find these discrepancies to be causal to the accident, on November 21, 1994 the NTSB issued the following Safety Recommendation to the FAA:

> Conduct a complete national review of all environments using MSAW systems. This review should address all user-defined site variables for the MSAW programs that control general terrain warnings, as well as runway capture boxes, to ensure compliance with prescribed procedures [7].

The FAA responded that it would undertake such a review, and on January 26, 1996 reported that the review had been completed. It is noteworthy that the 1995 facility inspection of Guam in which the MSAW inhibition was cited as an "informational" item was undertaken during the review period, even though no corrections were made. Perhaps the inhibition was not brought to the attention of the FAA officials conducting the national MSAW review, or perhaps the FAA did not consider the inhibition to be within the scope of the NTSB's recommendation.

3.4.2 Houston Intercontinental Airport, 1998

Four years later, on January 13, 1998, a Learjet crashed 2.3 nm short of the runway threshold while on final approach to runway 26 at Houston Intercontinental Airport. Investigators from the AOS determined that the MDA specified in the site adaptation parameters for the Houston MSAW installation was incorrect. The MSAW system was configured to use an MDA of 100 feet above ground when the actual MDA was 402 feet above ground. As a result, the Houston MSAW system failed to alert air traffic controllers when the Learjet violated the 402 foot MDA.

The configuration error in the Houston MSAW installation was the same error that had been made at Dulles four years earlier and should have been detected and fixed during the FAA's national MSAW review campaign. Moreover, the Houston accident occurred after the FAA had implemented the recertification programs and uniform site adaptation standards it had proposed in response to the Korean Air flight 801 accident. This accident underscores the view that the FAA's remedies to the MSAW program in response to the flight 801 accident were insufficient, as they failed to detect and repair a configuration error much like one the FAA had seen before at Dulles.

3.5 Lessons Learned

Accidents occur because of complex sequences of events and intricate combinations of circumstances. It is clear that many things could have prevented this accident. The NTSB report blames three factors—the flight crew, the lack of operation of the glideslope, and the FAA's inhibition of the MSAW system's service area. Presumably, changes were made based on the first two and the changes that were made as a result of the third were discussed earlier.

After examining all the evidence about this accident that is available, however, it is clear that the lessons learned from this accident were far short of what they should have been. Two additional prominent problems should have been identified and additional significant corrective actions taken.

3.5.1 Lesson 1—Configuration Management

Korean Air flight 801 crashed into Nimitz Hill, Guam on the morning of August 6, 1997 not only because of errors made by the flight crew, but also because of the manner in which the FAA made software changes to the MSAW system. By allowing each of the 193 MSAW-equipped air traffic control facilities to modify their MSAW installations at their discretion without guidance or review, the FAA allowed the system to be modified without regard to how the modifications might affect the system's ability to detect low-flying aircraft as well as the overall effect this policy would have on the safety of the NAS.

The MSAW system at Guam was a barrier designed to help prevent CFIT accidents. As such, it was a component of an overall system that included all of the barriers designed to prevent the hazard of flying below a safe altitude. Prevention of the hazard could have been achieved by any one of the barriers provided that particular one was perfect in its operation. But none of them were. The goal of preventing the hazard was to be achieved by accepting

that no barrier was perfect and providing several. Thus, the MSAW system's functionality was an integral part of the analysis of the overall system's safety. This does not mean that the system itself has to be ultra-dependable. It means that the system's dependability when coupled with that of the other barriers reduces the probability of an accident to acceptable levels.

The MSAW system's functionality was changed after its initial deployment by modifying the software, and thereby the safety analysis was invalidated. The crucial lesson here is that software configuration management is an essential part of maintaining system safety, and that any changes to a software system must be undertaken only in concert with a comprehensive safety analysis. The large inhibit zone that was in place in the MSAW system's service area was clearly a factor, but determining how and why this situation arose and deciding what to do about it is a complex undertaking. The official investigation of the accident did not lead to the vital changes in software practice that were clearly indicated. The MSAW systems at Guam and other locations were modified independently of the associated safety analysis, a procedure that proved fatal. It is essential that software maintenance procedures for safety-critical systems be conducted in the context of the safety requirements and that they be carried out without human error.

It is unlikely that this example of defective configuration management is an isolated incident. Correcting the software configurations of the various instances of the MSAW system is necessary but not sufficient. The overall configuration management of software in safety-related systems in fields such as aviation must be undertaken correctly, and it must be treated as a critical component of the engineering process.

3.5.2 Lesson 2—Human Error

The second lesson that should have been learned from this accident concerns human error. Software in a safety-critical system is a integral component that cannot be changed without suitable analysis of the impact of the change. The simplistic approach taken by the FAA to both changing the software in the Guam MSAW system before the accident and to checking the software in other MSAW systems after the accident indicates a poor understanding of the crucial role that software plays in safety system management.

Human error in the maintenance of software in safety-related systems is likely just as it is in the operation of those systems. Thus, complementing the first lesson noted above, the software research community needs to examine the complex circumstances that are present in widely deployed safety-related systems and develop techniques to verify properties that are crucial to safety. For example, requiring that the software (including all data and configuration files) for some particular system not be changed in the field unless the change is accompanied by suitable verification activities, re-establishment of safety properties, and compatibility checks with other software components might be a reasonable goal. Automating the process and its enforcement is also a reasonable goal. Many other ideas suggest themselves.

A critical problem with software in applications like aviation is the notion that software can be changed easily. Other engineering disciplines fail to appreciate that just because one can change software does not mean that one should. If inhibiting the MSAW system's service area had required extensive modifications to hardware, it almost certainly would not have happened.

4 British Airways Flight 027

On June 28, 1999, British Airways flight 027, a Boeing 747 carrying 419 passengers and crew members en route to Hong Kong, China, and another Boeing 747 operated by Korean Air Cargo nearly collided over a remote region of Chinese airspace. At their closest point of approach, the two aircraft passed within 600 feet of each other, and the British Airways copilot later recounted that his windshield was consumed by the fuselage of the other jet. No injuries resulted from the incident and both aircraft arrived at their destinations; however it is likely that if the two aircraft had collided none of the persons aboard either aircraft would have survived [10].

4.1 Background Information

Prior to the incident, the two aircraft were cruising in opposite directions along the same airway safely separated by 2,000 feet of vertical distance. The British Airways passenger flight was cruising at 33,500 feet and the Korean Air Cargo jet at 31,500 feet. The Korean Air jet was flying in a cloud, preventing the pilots from visually identifying each other's aircraft. Both aircraft were equipped with a collision avoidance system known as the Traffic Alert and Collision Avoidance System, or TCAS. The TCAS unit installed on the Korean Air jet indicated traffic 400 feet below and approaching head on and shortly thereafter instructed the pilot to climb to avoid the oncoming traffic. In reality, there were no other aircraft in the vicinity of the Korean Air jet except for the British Airways flight 2,000 feet above, and the TCAS unit's indication and climb instruction were erroneous. The pilot had no way of knowing this fact, however, as he was operating in a region of airspace without air traffic control service and the cloud layer severely limited his visibility, and thus he followed the climb

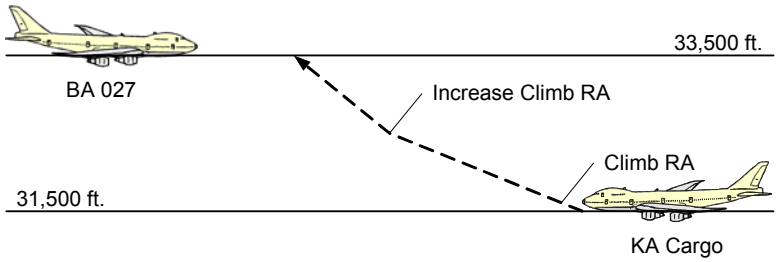

Figure 5: British Airways flight 027 incident sequence.

instruction issued by TCAS. The Korean Air pilot reported that the vertical separation between his aircraft and the phantom aircraft indicated by TCAS decreased to zero before increasing, and before reaching zero TCAS instructed him to increase his rate of climb. The pilot complied and pitched his aircraft further, unknowingly placing it on a collision course with British Airways flight 027, which was now closing in rapidly from above as shown in Figure 5 [10, 11].

As the Korean Air Cargo jet was making its climb, the crew of the British Airways passenger flight reported nothing unusual in their cockpit. Their TCAS display indicated traffic approaching head-on but still flying safely 2,000 feet below their own aircraft. Then, the TCAS unit suddenly issued a descend instruction and showed the traffic now approaching from only a couple hundred feet below. The flight crew began to comply with the instruction and pitched the nose down just before seeing the Korean Air jet emerge from the cloud layer below right in front of them. The two aircraft darted past one another separated only by 600 feet, well below the minimum separation limits for commercial aircraft. The entire incident sequence played out over an approximately 35-second period [13].

4.1.1 TCAS Overview

The Traffic Alert and Collision Avoidance System, or TCAS, is an onboard system designed to alert pilots of approaching traffic and provide guidance to avoid traffic conflicts

and maintain proper aircraft separation. The TCAS program was launched by the FAA in the 1980s, and TCAS units were deployed on commercial aircraft in 1993. TCAS is intended to serve as a backup to air traffic control, which retains the primary responsibility for maintaining proper aircraft separation for commercial flights. As such, TCAS is an entirely airborne system and does not require any ground support. All international flights are required to use TCAS as are all commercial flights operating in the United States [12].

TCAS comes in two variants, TCAS I and TCAS II, which offer different levels of advisory capability. Both systems track aircraft operating within four nautical miles and issue advisories to pilots when a conflict with another aircraft is projected to occur within approximately 45 seconds. These advisories are called traffic advisories (TAs) and consist of a visual indication on the pilot and copilot's primary flight displays along with an aural "TRAFFIC, TRAFFIC" annunciation. Pilots are not required to take any specific action in response to TAs but should attempt to visually identify the intruder aircraft if conditions permit. TCAS II provides a second level of advisory called a resolution advisory (RA). If a traffic conflict is projected to occur within approximately 30 seconds, TCAS II systems will compute an escape strategy in the form of a vertical maneuver and advise pilots to climb or descend to avoid the conflict. RAs take the form of visual indications on the flight displays that show the attitudes and vertical speeds the pilot must maintain to avoid the conflict as well as aural annunciations such as "CLIMB, CLIMB" or "DESCEND, DESCEND." If the intruder aircraft changes its course after the original RA was issued, TCAS may have to revise or reserve its RA accordingly. A revision is an instruction to alter the rate of the climb or descent directed by the original RA, while a reversal is a change in the direction of the RA. If the intruder aircraft is also equipped with a TCAS II unit, the two systems will coordinate to ensure that they issue com-

plementary RAs. TCAS notifies the flight crew accordingly when the aircraft has cleared a traffic conflict.

4.1.2 Aircraft Tracking

TCAS detects and tracks surrounding aircraft using technology similar to that used by air traffic control radars to track aircraft from the ground. As Figure 6 depicts, the system uses a radio transceiver to broadcast an interrogation signal via a directional antenna. Any nearby aircraft that are equipped with transponders will detect the signal and "squawk" back a reply containing information such as the aircraft's altitude, heading, airspeed, and vertical rate of climb or descent. The transceiver receives these replies and uses the directional antenna to determine the position of each aircraft that responded. The transponder replies along with the associated directional information are forwarded to a logic unit for processing. After filtering out returns from ground entities and aircraft farther than four nautical miles away, the logic unit uses the transponder returns to determine the position and velocity of each of the surrounding aircraft as well as information from local air data sources to determine the velocity of the aircraft on which the system is operating. It then predicts the course of each aircraft it is

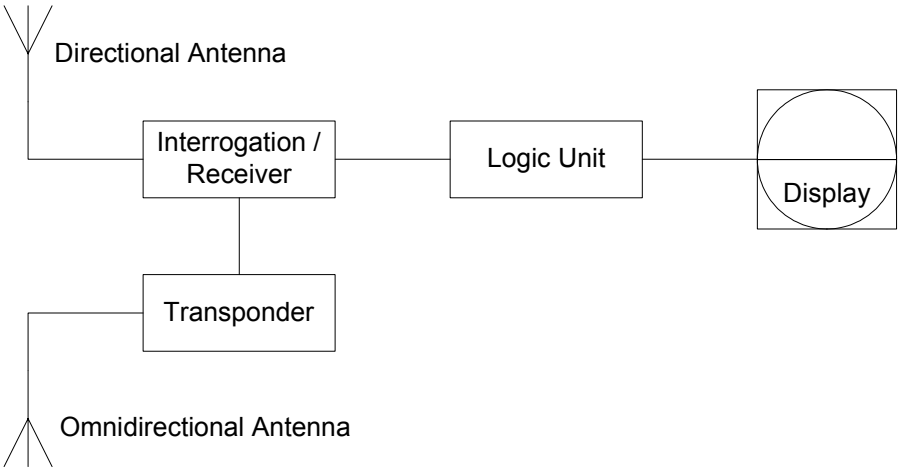

Figure 6: TCAS schematic [12].

monitoring to determine whether the aircraft poses a threat. If it predicts that a conflict will occur within about 45 seconds, it issues a TA. For TCAS II systems, if the logic unit predicts that a conflict will occur within about 30 seconds, it computes an escape strategy and issues an RA. The escape strategy is typically based on the relative altitude of the intruder aircraft. Generally, if the intruder is below, a climb RA is issued; if it is above, a descend RA is issued. The logic unit displays traffic information to the pilots including the threat classification of each tracked aircraft on the primary flight displays.

TCAS receives flight data for the aircraft on which it is operating from two independent air data sources. These data are passed into a comparator where they are averaged before being sent to the TCAS logic unit as illustrated in Figure 7. If the comparator detects that the variance in the inputs from the air data sources is too large, it raises an error signal that prompts TCAS to shut down and print an error message on the primary flight display. This design allows the system to detect but not tolerate disagreement between the air data sources or a failure of one of the sources. The comparator on the TCAS unit installed in the Korean Air Cargo jet featured an Enable line that if set to one would cause the comparator to function normally and if set to zero would cause it to function as a pass-through between one of the air data sources and TCAS [10].

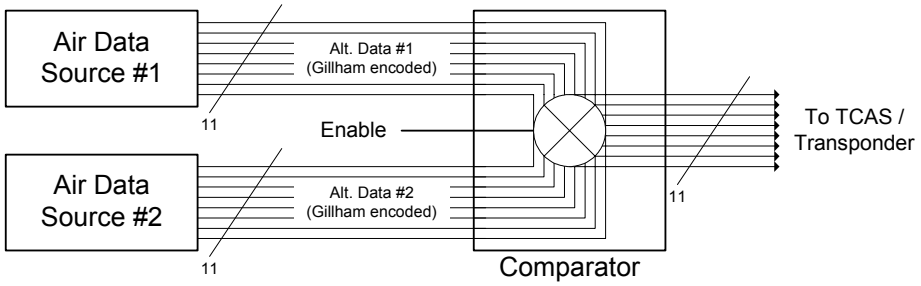

Figure 7: Simplified schematic of the air data comparator [10].

The air data sources report altitudes using an 11-bit binary encoding scheme known as Gillham code. Altitude data are sent using this encoding both to the transponder where they are transmitted as part of the transponder's interrogation reply and to the comparator where the data are averaged and forwarded to TCAS. Neither Gillham code nor the transponder protocol employ any error detection or correction mechanisms to verify the integrity of the data. TCAS compensates for this possibility by maintaining histories of the transponder returns from each tracked aircraft that it compares with new returns as they arrive. If a new return contains a fluctuation that is atypical of the performance capabilities of jet aircraft, such as a sudden change in altitude, TCAS assumes that the return is faulty and discards it. TCAS will continue to discard faulty returns for up to one minute from the detection of the original faulty return, at which point it resumes processing the returns normally irrespective of whether the fluctuation has disappeared.

4.1.3 Maintaining Aircraft Separation

TCAS is one of three mechanisms in the air traffic system designed to help maintain proper aircraft separation and prevent midair collisions. Air traffic control (ATC) is the primary line of defense on this front, and air traffic controllers can resolve traffic conflicts long before pilots or TCAS are even aware of them. TCAS is the secondary system and only reports conflicts when they are projected to occur within one minute in order to give ATC time to resolve the conflict first. Visual identification is the last defense mechanism and involves pilots actually looking through their windshields and attempting to spot aircraft operating in their vicinity. The latter technique is unreliable, however, as visibility conditions may be too poor to see outside the cockpit clearly, such as when an aircraft is flying in a cloud. Although ATC is the primary means of maintaining separation, followed by TCAS and then visual iden-

tification, these three barriers supersede one another in the reverse order. FAA and international regulations assign pilots the ultimate responsibility for maintaining the safety of their aircraft and grant pilots blanket authority to deviate from flight rules or ATC clearances in order to meet an emergency [14, 15]. Similarly, TCAS RAs take precedence over ATC instructions. An FAA advisory circular discussing the operation of TCAS II states, "For TCAS to work as designed, immediate and correct crew response to TCAS advisories is essential. Delayed crew response or reluctance of a flight crew to adjust the aircraft's flight path as advised by TCAS due to Air Traffic Control (ATC) clearance provisions, fear of later FAA scrutiny, or other factors could significantly decrease or negate the protection afforded by TCAS." The advisory later states, "If a TCAS RA requires maneuvering contrary to 'right-of-way' rules, 'cloud clearance' rules for visual flight rules (VFR), instrument flight rules (IFR), or other such criteria, pilots are expected to follow the TCAS RA to resolve the immediate traffic conflict" [16]. Thus, pilots are encouraged to follow TCAS RAs in spite of ATC instructions or clearances unless they have visually acquired the intruder aircraft, and must do so quickly since an RA indicates that a conflict is projected to occur within 30 seconds.

4.2 Postaccident Actions

This incident was investigated separately by the U.K. Civil Aviation Authority (CAA) and British Airways to explain the behavior of the Korean Air jet and determine why the TCAS unit onboard the British Airways flight failed to issue an advisory to the flight crew until moments before the two aircraft reached their closest point of approach. An inspection of the TCAS unit installed on the British Airways jet did not detect any problems. When the TCAS unit aboard the Korean Air jet was inspected, however, investigators found that cir-

cuitry related to TCAS function had been damaged in two locations. Part of the data line used to send the barometric altitude reading from one of the air data sources had been damaged, resulting in a bit-stuck-at-one error on the line. This error would have caused incorrect altimeter readings to be sent to the transponder, which would transmit the readings to other aircraft and to the comparator used by TCAS. The comparator should have detected this anomaly and raised the error signal that would have caused TCAS to shut down; however a pin on the Enable line to the comparator had been pushed back, causing it to short open, thereby disabling the comparator. As a result, faulty altitude values were allowed to pass through to the TCAS logic unit unchecked.

Although the air data source was sending the correct altitude value, the bit-stuck-at-one error on the data line caused TCAS to receive a value containing a one-bit discrepancy that corresponded to a 2,400-foot difference in altitude according to Gillham code [10]. Thus, the TCAS unit aboard the Korean Air jet believed it was flying at 33,900 feet, placing it 400 feet above the British Airways jet. According to the separation rules used by TCAS, this situation created a conflict between the two aircraft, and since TCAS believed it was the one flying above, it issued a climb RA to the pilot. As the pilot executed the instruction and the aircraft's altitude began to increase, the altitude value reported to TCAS also changed; however the one-bit error caused TCAS to think the aircraft was actually descending, decreasing the separation between it and the intruder. Consequently, the system revised its RA and instructed the pilot to climb faster, placing the two aircraft on what was actually a near-collision course.

Just as the error on the data line was causing incorrect altitude values to be sent to TCAS, it was sending the same incorrect readings to the transponder, which was transmitting

them to the British Airways jet in response to its interrogation signals. One might think that this behavior would have caused the TCAS unit onboard the British Airways flight to also believe that the Korean Air Cargo jet was flying at 33,900 feet instead of its true altitude of 31,500 feet. Instead, when the British Airways jet's transceiver began receiving the erroneous readings, the fault tolerance mechanisms discussed earlier detected the sudden altitude jump and began discarding the erroneous returns. This response prevented the TCAS unit aboard the British Airways jet from issuing a false RA, but it also meant that the crew of flight 027 was unaware that the Korean Air jet below was climbing toward them since the TCAS traffic display continued to plot the jet at 2,000 feet below. This erroneous indication continued until moments before the closest point of approach when the TCAS unit finally started processing the returns again and issued a descend RA to the pilot of the British Airways flight.

With the assistance of Korean Air, the CAA determined that the damage to the Korean Air Cargo jet's TCAS unit occurred during maintenance to the aircraft's avionics systems. Upon concluding its investigation, the CAA issued an airworthiness directive requiring air carriers using Gillham code to check the altitude values being transmitted by the transponder throughout the operational envelope of the aircraft and to periodically inspect the comparator unit to ensure that it is functioning properly. The CAA also notified other European aviation regulatory agencies and the FAA of the problems it found as well as manufacturers of transponder and TCAS equipment, and it issued a recommendation to aircraft operators urging them to consider using other encoding schemes for transmitting altitude data instead of Gillham code. At the end of its report on the incident the CAA noted, "This incident shows the effects that secondary failures can have on primary systems fitted to aircraft today. Regardless of the integrity of the collision avoidance system, it shows that relatively minor faults in the

interfacing system can still contribute to a serious safety risk" [10]. Indeed, safety is a systems issue, and the fact that one subsystem has high "integrity" does not imply that the resulting system will as well.

4.3 Analysis

The near-collision involving British Airways flight 027 revealed several design issues concerning TCAS. The CAA's investigation into the incident failed to document these issues, however, even though they could contribute to future recurrences if not corrected. This section discusses the design issues in TCAS as well as deeper problems with the manner in which agencies tend to investigate incidents.

4.3.1 TCAS Design Issues

The follow-up actions taken by the CAA focused on the maintenance issues that caused the damage to the TCAS system aboard the Korean air jet and those that allowed it to operate in such a state. While these issues are important, serious design issues also exist in TCAS at least in the models aboard the incident aircraft. As the CAA report suggested, the first relates to the manner in which altitude values and possibly other flight data are sent from the air data sources to the transponder and TCAS and then transmitted by the transponder to nearby aircraft. Although TCAS relies heavily on the altitude data it receives from transponder returns in order to make its traffic assessments, the transponder protocol does not employ any error detection or correction mechanisms to verify the integrity of the data, and even simple transmission faults that could be detected by employing parity checking or cyclic redundancy checks (CRCs) can pass through to TCAS unnoticed by the transceiver hardware. TCAS attempts to compensate for this possibility by examining the history of each tracked

aircraft to check for erroneous returns, but this is by no means a complete solution. Transponders that have permanently failed and are babbling erroneous data will not be detected as faulty by TCAS after the one-minute grace period nor can TCAS identify faulty returns when operating in areas of persistent radio interference sufficient to disrupt transponder communications. This problem could be remedied by adding a CRC or similar field to transponder messages; however achieving backward compatibility with older transponders and ground radars could be a difficult process, and replacing obsolete transponder and radar equipment would be very expensive.

The second issue pertains to the design of the comparator used to verify the data gathered from the two air data sources. The purpose of receiving data from two separate sources is to stabilize the air data TCAS receives and reduce the likelihood that faulty data is allowed to pass into TCAS undetected. The comparator performs both of these functions and signals TCAS if it detects a problem with the data, which prompts TCAS to shut down and notify the pilot. This mechanism prevents TCAS from acting on the faulty data and displaying inaccurate traffic information to the pilot or, more importantly, detecting a false traffic conflict and issuing an erroneous RA. The extra reliability achieved through this fault detection mechanism is defeated, however, if the comparator itself introduces vulnerabilities to the system. In the design used by the TCAS unit aboard the Korean Air jet, the Enable line to the comparator presented such a vulnerability. A simple electrical fault on the comparator line, such as a bent pin or a short to ground, would completely disable the fault detection mechanism intended to prevent TCAS from acting on faulty air data with no indication to the flight crew. While the Enable line may have provided some convenience for testing and troubleshooting the system,

its presence while the aircraft was in operation merely weakened the integrity of the comparator and with it the reliability of the air data that passed into TCAS.

The underlying issue concerning the design of the TCAS unit installed on the Korean Air jet is that vulnerabilities existed in the design that jeopardized the system's ability to satisfy its dependability properties. TCAS is relied upon to enhance the safety of the air traffic system by supplementing air traffic control in maintaining proper separation between aircraft and preventing midair collisions. Thus, its two key dependability properties as defined by Laprie are safety, the absence of catastrophic consequences for its users (the flight crew and passengers) and the environment, and reliability, defined as the probability that the system will provide service without interruption for a given period of time [17]. An informal minimum safety requirement would be that the system provide at least the level of safety present in its absence, or in other words that it not make matters worse. In the context of TCAS, this requirement means that the system must not issue bad guidance when it detects a real conflict and that it not detect false conflicts and issue erroneous RAs to pilots. Issuing bad guidance could cause two aircraft in conflict with one another to both evade in the same direction, exacerbating the conflict. Issuing RAs in response to false conflicts causes unnecessary disruptions to the air traffic system since aircraft might have to disobey ATC clearances to follow the RAs, which could lead to real conflicts with other aircraft that otherwise would not have occurred. Either scenario is dangerous because pilots would not know that they were following bad guidance or responding to a false conflict until they visually acquired the intruder aircraft or unless air traffic control detected the problem and intervened. As discussed earlier, visual acquisition is an unreliable collision avoidance technique, and air traffic controllers would have to act quickly to intervene and resolve the conflict since TCAS issues RAs

approximately 30 seconds before the aircraft are predicted to reach their closest point of approach. Even if controllers did intervene, the guidelines strongly encouraging pilots to follow RAs in spite of ATC instructions make it unclear as to what course of action a pilot would take when faced with this dilemma. Moreover, in uncontrolled airspace such as the area in which the incident occurred, TCAS is essentially the primary authority in maintaining aircraft separation because visual acquisition is the only other means of identifying traffic conflicts. Thus, a scenario in which TCAS issued improper or erroneous guidance could have serious consequences for the safety of the aircraft on which it is operating as well as for aircraft in the surrounding airspace. Since TCAS makes its conflict assessments and issues advisories based on the data it receives both from the local air data sources and the transponder returns from other aircraft, it is crucial that these data be free of errors or at least that TCAS be able to detect errors when they are present.

The design issues described earlier reveal two ways in which faulty data can reach TCAS undetected and distort its view of reality, thereby jeopardizing its ability to ensure safety. This observation does not mean that TCAS is an unsafe system, however, as it has been credited with avoiding numerous loss-of-separation incidents. Nevertheless, it is an imperfect system and perhaps needs to be improved further.

4.3.2 Incident Investigation

The result of the CAA's investigation into the flight 027 incident was a three-page report briefly describing the incident and the investigation, a paragraph documenting the analysis, and summary lists of the actions taken, conclusions, and recommendations [10]. The investigation was fairly informal and conducted with the assistance of British Airways and Korean Air officials. This degree of analysis pales in comparison to the formal investigations

that are launched in response to accidents involving loss of life, injury, or substantial damage to property and the voluminous reports they produce. This is not to single out the UK's Civil Aviation Authority, however, as the practice is shared by investigative agencies worldwide. British Airways conducted a more detailed investigation of the incident, but has not officially released the report of its investigation or findings to the public. They did note, however, that the two aircraft missed each other only by luck, and that had the aircraft been using the more precise Global Positioning Systems (GPS) navigation systems that are now widely used for navigation aboard commercial aircraft, they would have likely collided head-on [11].

The observation that only luck prevented a collision in this incident brings into question the practice of allocating investigative resources based on the severity of the loss associated with an accident or incident. An accident might be the result of a simple error involving a well-understood and accepted risk, such as forgetting to deploy the landing gear during final approach. On the other hand, an incident might involve a fault in a critical avionics system that goes unnoticed either because the fault was transient or because the pilot was able to continue flying the plane despite the manifestation of the fault. Most accidents involving commercial aircraft do involve complex sequences of events because of the numerous safeguards in place to ensure the safety of commercial flight, and these accidents warrant thorough investigations to determine how the safeguards can be improved. This fact does not make incidents any less important, however, because they are often precursors to accidents.

An incident can be thought of as a failure of a network of barriers designed to prevent a particular untoward event, and an accident is simply an incident followed by a loss event, such as a loss of life, injury, or destruction of property. For example, an incident has occurred when a pilot forgets to deploy the landing gear before reaching the runway. Whether this inci-

dent results in a loss event is dependent on several factors, such as the impact forces when the fuselage makes contact with the runway, whether sparks from the aircraft scraping along the pavement ignite the fuel, and possibly whether the pilot remembers the gear at the last second and aborts the landing. None of these factors can be relied upon to save the aircraft once the incident has occurred, however, and it is essentially luck that determines the fate of the pilot, the passengers, and the aircraft. While steps can be taken to help mitigate the extent of the loss, for example by reinforcing fuel tanks to help them better withstand impact and sheering forces, for the most part these measures acknowledge that some degree of loss is inevitable given the occurrence of an incident. Therefore, preventing the incident, not the loss event, is more important from an investigative standpoint.

Incidents provide opportunities for investigators to find ways of improving safeguards without the tragedy associated with accidents, and investigating incidents with the same rigor as accidents can save lives if those investigations uncover ways of preventing the incidents from recurring and possibly becoming accidents. The midair collision that occurred between two TCAS-equipped aircraft over southern Germany on July 1, 2002 is believed to have occurred in part because one of the pilots received contradictory advisories from TCAS and air traffic control and opted to follow the ATC instruction instead of the TCAS RA. A similar incident occurred on January 31, 2001 in airspace near Tokyo, Japan in which TCAS and an air traffic controller detected a traffic conflict at approximately the same time but issued contradictory instructions to one of the pilots. The pilot followed the ATC instruction and nearly collided with the other aircraft. Japanese authorities investigated the incident and determined the cause to be human error on the part of the pilot for disobeying the TCAS RA in favor of the ATC instruction; however the collision over Germany the following year raises doubts as

to whether there are deeper problems in the network of barriers intended to prevent midair collisions that need to be addressed [18].

Incidents in which design is suspected of being a factor in the failure of a safety-critical system deserve special attention. Unlike degradation faults that result from damage or normal wear and tear, design faults are present throughout the lifetime of a system. Moreover, whereas a particular degradation fault might only affect a handful of the population of a system, if a design fault is present in one member of the population, it is present in all of them. Either kind of fault can lead to system failure if the system is unable to tolerate the fault; however design faults tend to be more difficult to predict and tolerate because of the difficulty in understanding their failure semantics and the behavior of the system after encountering a design fault. The CAA's investigation into the British Airways incident over China focused on the sources of the degradation faults that contributed to the incident—the damage to the air data line and the bent pin that disabled the comparator. The investigators treated TCAS, the transponder, and the comparator largely as black boxes that could not be changed, which prevented them from discovering the design faults plaguing TCAS, the comparator, and the transponder protocol that allowed TCAS to process faulty traffic data and issue erroneous advisories.

4.4 Related Incident

The British Airways incident over China followed a similar incident that occurred between two aircraft in January 1998 over Hawaii in which one aircraft's TCAS unit issued a false traffic advisory because an air data computer had malfunctioned and was reporting the aircraft's altitude as 1,500 feet higher than its actual position. Fortunately, air traffic control-

lers happened to notice a discrepancy between the aircraft's altitude as reported by its transponder and that reported by the flight crew and were able to defuse the situation before it escalated further.

Because the aircraft with the malfunctioning equipment was operated by an Australian carrier, the Australian Transport Safety Bureau (ATSB) launched an investigation into the incident. When the British Airways incident occurred 15 months later, ATSB investigators saw the similarities between the two incidents and obtained a copy of British Airways' findings. The findings went beyond recommending better maintenance and addressed the design issues highlighted in this thesis along with other issues such as human factors. British Airways investigators recommended changes to the TCAS design and displays, which included adding a display of the local altitude to the TCAS traffic information so that flight crews could cross-check the altitude against other altimeter readings on the instrument panel. The findings also included recommendations for adding a special designation to traffic for which TCAS has received suspect transponder returns and indicating to the flight crew when TCAS is coordinating an RA with an intruder aircraft so they will know that the intruder is also likely to maneuver in response to the RA. Aware of the dangers of allowing TCAS to operate on faulty air data, British Airways also advised that Gillham code should be abandoned in favor of more robust solutions for transmitting altitude data.

The investigation conducted by British Airways addressed the design issues missed by the CAA and recommended changes to TCAS that would eliminate some of the faults and make it easier for flight crews to detect others when they appeared. As a result, British Airways produced a stronger set of recommendations that not only reduced the probability that damage-induced degradation faults would occur in TCAS and the systems that supply it data,

but also made it less likely that the occurrence of such faults in the future would cause the system to fail in an undetectable manner. British Airways should be commended for the completeness of its investigation and findings, and the ATSB adopted its recommendations largely verbatim in its report on the Hawaii incident [13]. Had such an investigation been performed in a timely fashion following the January 1998 incident, however, the near-collision over China might have been prevented and the lives of 400 passengers not put at risk.

4.5 Lessons Learned

British Airways and the CAA presented lessons and recommendations for the improvement of TCAS, transponder systems, and policies for maintaining and inspecting these systems. Once again, the lessons learned from this accident were far short of what they should have been. Two additional prominent problems should have been identified and additional significant corrective actions taken.

4.5.1 Lesson 1—Incident Classification

The first lesson is that classification schemes in which investigative resources are allocated to accidents and incidents based on their associated losses de-emphasize the importance of incidents with no losses even though these incidents may still have important lessons to be learned. Incidents provide opportunities to improve the affected systems without the consequences associated with accidents, and investigators should seize upon these opportunities to prevent similar sequences of events from occurring in the future, possibly with more dire consequences. Many accidents have been preceded by similar incidents, and lives and property could have been spared if the problems contributing to those accidents had been addressed when they manifested themselves earlier.

4.5.2 Lesson 2—Criticality of Design Faults

The second lesson is that when an accident or incident occurs involving a safety-critical computing system, investigators must pay particular attention to possible design faults that might be present in the system. While degradation faults are important, these are fairly well-understood and models exist to predict when and where degradation faults are likely to occur in a system throughout its lifetime. Design faults pose a vastly greater challenge to system engineers—their nature, when they will occur, what parts of the system will be most susceptible to them, and what effects they might have on system behavior and output are extremely difficult to predict and mitigate. Moreover, design faults do not vary across the population of a system; if one member of the population contains a design fault, the others do as well. Therefore, it is critical that investigators find design faults that contribute to a system failure because every other installation of the system will be susceptible to the same failure under similar conditions. Because design faults are difficult to understand, attempting to compensate for the design faults in a system by trying to prevent the conditions that trigger the faults from recurring rather than correcting the faults themselves is a strategy that is unlikely to succeed.

5 Common Lessons & Observations

While the Korean Air flight 801 and British Airways flight 027 incidents were distinct in most aspects, they each involved failures of safety-critical software systems. These failures were the result of problems in the manner in which the systems were developed or maintained, but they were not examined or documented sufficiently by the agencies that investigated the incidents. Moreover, the differences between the two investigations illustrate the disparity in the level of attention given to accidents versus that given to incidents. This chapter presents a common lesson to system designers, managers, and regulators arising from the incidents and makes observations on investigations involving safety-critical software systems.

5.1 Common Lesson

The systems examined in this thesis are each part of much larger systems designed to enhance the safety of commercial air travel. The MSAW system is part of the FAA's program to prevent CFIT accidents, and TCAS plays a significant role in reducing the likelihood of mid-air collisions. In both of the incidents described earlier, the systems involved were viewed as if they were isolated; ample consideration was not given to the roles these systems played in the overall systems they were part of. When the FAA inhibited the MSAW system at Guam, it did not consider how the inhibition would affect the safety of Guam airspace, especially in light of the other CFIT prevention systems that were also disabled or being ignored. Likewise, the designers of TCAS did not adequately address the effects faulty data could have on the system's functionality. Although TCAS is merely an advisory system, its advisories trump air traffic control instructions according to FAA and international regulations, and in some scenarios it is the only system capable of detecting and resolving conflicts in time to

avoid a collision. If the system issues false or erroneous advisories, it could actually decrease the safety of the air traffic system by creating conflicts where none would have otherwise existed.

When changes are made to a safety-critical system, the original safety analysis of the system is invalidated and must be performed again to ensure the system is still compliant with its original safety requirements. Moreover, when a new safety system is to be added to an existing network of barriers, that system must be examined to ensure that it will not adversely affect safety through faulty operation. Although these lessons are not new to the software engineering community, they are worth restating here because the incidents described in this thesis indicate that the community must work harder to disseminate them throughout the avionics industry.

5.2 On Incident Investigations

The two incidents described in this thesis illustrate the need for more comprehensive investigations of incidents involving safety-critical software systems. Both the NTSB and the CAA successfully determined the sequence of events leading to the Korean Air and British Airways incidents; however investigators missed important lessons when they interpreted this information, and as a result their recommendations were inadequate. In the case of MSAW, the FAA's changes to its MSAW program arising from the Guam accident failed to prevent the configuration error that contributed to the crash at Houston Intercontinental Airport in 1998. In the case of TCAS, the CAA's official recommendations focused only on maintenance and did not include the design changes recommended by the British Airways internal report.

These inadequacies occurred because investigators failed to examine how subtle changes in the event sequences could lead to the same failures. The crash at Dulles Airport in 1994 prompted the FAA to conduct a review of MSAW installations nationwide; however this review focused on fixing the specific configuration error that prevented the system from alerting air traffic controllers in advance. It did not consider that other configuration errors, such as improperly defined inhibit zones, could also prevent the system from functioning as intended. Similarly, the CAA's recommendations arising from the British Airways incident focused on the specific circumstances of that incident and did not address the broader issue of how a TCAS unit might act on faulty data received either from local air data sources or from transponder returns. Investigators must ensure that the lessons they extract from incidents are comprehensive enough to encompass slight differences in the event sequences that could lead to similar outcomes.

The need to consider subtle differences in event sequences reiterates the problems associated with investigating accidents and incidents differently. The Korean Air flight 801 accident received much more investigative and public attention than did the British Airways flight 027 incident, even though the latter could have easily developed into a tragedy with twice the number of casualties. All accidents begin as incidents, and luck could be the only factor determining whether an incident develops into a catastrophe. From an investigative perspective, the lessons to be learned from an incident and its related accident are equally important since these lessons usually focus on preventing the incident rather than mitigating the extent of the loss. This point is especially true in the context of safety-critical software systems, where design faults are shared by all instances of a particular system. If an incident reveals the presence of a design fault in a particular system, investigators have an opportunity

to develop recommendations to prevent the fault from manifesting itself again in other installations of the same system, possibly with more severe consequences.

6 Loss-Based Incident Classification

The practice of classifying incidents according to the extent of their associated losses might seem intuitive. Heavier losses result in greater tragedies, and the travelling public demands inquiries into major accidents in order to prevent their recurrence. The preceding chapters have shown, however, that the lessons learned from incidents involving safety-critical software systems can be just as important as those learned from accidents involving such systems. Despite this observation, many investigative agencies worldwide employ loss-based classification schemes to allocate resources to accident and incident investigations. This chapter examines the correspondence between loss and investigative rigor and the extent to which loss is indicative of the potential to learn new lessons and prevent future tragedies.

6.1 Investigative Resources

The need to prioritize accident and incident investigations arises from the limited resources available to investigative agencies. Most investigative agencies simply do not have the resources to fully investigate every aviation-related incident or accident that occurs within their jurisdiction. Agencies typically prioritize accidents according to the severity of their associated losses. For example, the National Transportation Safety Board (NTSB) classifies an accident as "major" if the accident results in the destruction of a commercial aircraft, multiple fatalities, or one fatality and substantial damage to a commercial aircraft. According to NTSB statistics, 74 major accidents occurred between 1983-2002 compared to 581 accidents receiving less severe designations [21]. NTSB investigators use a special operating manual when investigating major accidents that guides them in collecting evidence, holding public

hearings, and preparing final reports [22]. Reports are typically reserved for major accidents; synopses are prepared for less severe accidents and then stored in a database.

The NTSB is not unique in employing these procedures. Although many other agencies worldwide do not limit their investigations to accidents as the NTSB does, most distinguish between accidents and "serious incidents," including the U.K. Air Accidents Investigation Branch (AAIB), the French Bureau d'Enquêtes et d'Analyses (BEA), the German Federal Bureau of Aircraft Accidents Investigation (BFU), the Accident Investigation Board of Finland (AIB), the Australian Transport Safety Bureau (ATSB), the Taiwanese Aviation Safety Council (ASC), and others. While the definition of "accident" is typically clear, the term "serious incident" is often not well-defined. The AAIB and BFU offer guidelines that give examples of serious incidents but admit that these guidelines are not comprehensive. The ATSB uses a five-category system to classify accidents and incidents, but the criteria for categorizing an occurrence are subjective. The Canadian Transport Safety Board (TSB) does not actually distinguish between accidents and incidents but labels both types of events as "occurrences." They classify and investigate occurrences based on "whether the investigation is likely to lead to reduced risk to persons, properly, or the environment" [23]. This method is similar to the scheme proposed later in this thesis; however their criteria are still quite subjective.

The effect of allocating resources to accident and incident investigations based on the severity of their associated losses is that less severe accidents might receive only a small amount of attention from investigators and incidents might not be investigated at all. However, many major *accidents* are preceded by similar *incidents* in which it was only by coincidence that a loss did not occur. This observation is particularly important in the context of

safety-critical software systems, because design faults present in such systems can manifest themselves with unpredictable consequences. If the systems control hazardous operations, they might bring direct harm to passengers or crew. Alternatively, if the systems provide advice or warnings to pilots, they might raise false alerts or issue erroneous guidance to pilots, who could inadvertently jeopardize safety by acting on this information.

6.2 Incident Comparison

To illustrate the distinction between the way in which accidents and incidents are investigated, this section examines the investigations conducted following the Korean Air flight 801 and British Airways flight 027 incidents. They are then compared using a variety of metrics to illustrate difference in casualties and subsequent difference in effort devoted to investigating each incident.

6.2.1 Korean Air Flight 801

The NTSB began its investigation into the Korean Air flight 801 accident immediately after the crash. The Board adopted its final report, a 212-page document, on January 13, 2000. The report contains 134 pages of factual information pertaining to the accident and 37 pages of analysis. The investigation yielded 36 findings and a set of 15 recommendations mostly addressed to the FAA. During the investigation, the NTSB held a three-day public hearing into the accident in which officials from the FAA, Korean Air, the government of Guam, and other organizations gave testimony. The transcript from the hearing spans approximately 430 pages [23].

6.2.2 British Airways Flight 027

The U.K. Civil Aviation Authority (CAA) and British Airways each conducted their own investigations into the British Airways flight 027 incident. The CAA's report does not indicate when its investigation into the incident began; however the report is dated October 28, 1999, suggesting that the investigation lasted at most four months. The report is three pages long and includes eight paragraphs of factual information spanning two pages and a single paragraph of analysis. It contains a single conclusion and three recommendations directed at operators and equipment manufacturers. No public hearing was held in response to this incident. British Airways prepared a more detailed report on the incident, but that report has not been officially released to the public.

6.2.3 Comparison

Clearly, the Korean Air flight 801 accident received a more rigorous investigation than did the British Airways flight 027 incident. In order to help quantify the extent of the difference, data from the events and their investigations are summarized in Table 2 below.

	Korean Air 801	British Airways 027
Classification	Accident	Incident
Persons On Board	254	419
Fatalities	228	0
Injuries, Serious	26	0
Injuries, Minor	0	0
Total Casualties	254	0
Aircraft Damage	Destroyed	None
Investigation Length (months)	30	4
Final Report Length (pages)	212	3

Table 2: Comparison of Korean Air Flight 801 and British Airways Flight 027

	Korean Air 801	British Airways 027
Factual Information (pages)	134	2
Analysis (pages)	37	1
Findings / Conclusions	36	1
Recommendations	15	3

Table 2: Comparison of Korean Air Flight 801 and British Airways Flight 027

On the basis of loss, the near-collision involving British Airways 027 had no casualties compared to a 90% fatality rate in the Korean Air 801 accident. In addition, neither of the Boeing 747s involved in the near-collision sustained any damage from the incident, whereas the 747 involved in the Guam accident was destroyed. Examining loss alone, the Korean Air accident over Guam appears far more important than the near-collision over China, and thus the large discrepancy in output from the two investigations might seem warranted.

Comparing these events solely on the basis of loss is deceiving, however, as the British Airways incident could have easily developed into an accident with almost twice the number of fatalities as the Korean Air flight 801 crash in Guam. As British Airways officials noted, it was entirely by luck that the British Airways passenger flight and the Korean Air Cargo flight did not collide. By the time the Korean Air pilot inadvertently placed his aircraft on a collision course with British Airways flight 027, all of the barriers designed to prevent midair collisions had been defeated. If the incident sequence were to recur with similar aircraft, a variation in wind direction or in navigational precision could lead to a much more dire outcome and almost certainly would have if the incident aircraft had been using the Global Positioning System (GPS) navigation systems in widespread use today. Under the accident classification schemes employed by most investigative agencies, this catastrophic outcome would be necessary for a major investigation to be undertaken, even though the findings and recommenda-

tions would likely be the same as if an equally rigorous investigation had been conducted into the incident alone. This should not be the case. New classification schemes are necessary in order to better allocate investigative resources to incidents whose recurrence could have more severe consequences.

In reviewing this comparison, one might argue that the vast difference between the Korean Air and British Airways events was not necessarily because of their associated losses but rather due to the fact that different agencies investigated each event. Had both events been investigated by the NTSB or CAA, the figures might have matched more closely. Because the NTSB does not investigate incidents, however, had the near-collision involving British Airways flight 027 occurred in U.S. airspace, it is unlikely the NTSB would have issued any report on it. The FAA might have chosen to investigate the incident, but the extent of the investigation, if any, would have been at the administration's discretion. Similarly, if the Korean Air flight 801 accident had occurred in British airspace, it would have been investigated not by the CAA but by the AAIB, whose formal reports are similar to the NTSB's final reports in structure and length.

7 Risk-Based Incident Classification

This chapter proposes a new incident classification scheme based on risk. Instead of prioritizing incidents according to the extent of their subsequent losses, investigators should allocate resources to incident investigations based on the risks of the incidents' recurrence. To facilitate this approach, a new classification metric, Total Risk, is introduced in order to assess the risk associated with an incident as well as its importance relative to other incidents. A process called Iterative Reclassification is also developed to assist investigators in making initial Total Risk assessments, refining the assessments as investigations proceed, determining which which leads in an investigation to pursue next, and deciding when to defer or conclude an investigation.

7.1 Motivation

The term "incident" can be defined in a variety of ways but typically involves the failure of a network of barriers designed to protect a system from one or more hazards. An incident becomes an accident when it is coupled with a loss event such as a crash or collision in which damage or casualties are incurred. It is often the case that luck determines whether an incident develops into an accident and, if so, what the extent of the loss will be.

When investigating accidents, investigators can issue recommendations aimed at preventing the associated incident or at mitigating the severity of the loss, and they usually do both. While attempting to mitigate loss given the occurrence of an incident can help to reduce the severity of accidents, some degree of loss is almost always inevitable. On the other hand, if the incident itself is prevented, it cannot develop into an accident and thus no loss will occur. Therefore, recommendations aimed at preventing incident recurrences are likely to be

more effective in preventing future losses. Indeed, 13 of the 15 recommendations issued by the NTSB in response to the Korean Air flight 801 accident were aimed at preventing the recurrence of incidents in which aircraft descend below safe altitudes during final approach. Only two focused on mitigating losses by suggesting improvements to Guam's emergency response units.

Given that accidents begin as incidents and that incident prevention should be the focus of investigations, incidents are opportunities for investigators to identify problems and suggest safety improvements without the losses associated with accidents. Accident classification schemes based on loss alone place a low priority on incidents even though those incidents might be indicative of safety problems that could lead to more catastrophic outcomes should they recur. By itself, loss is a poor indicator of an incident's potential for learning new lessons and preventing future incidents. Therefore, classification schemes based on loss should be de-emphasized in favor of new schemes in which resources are allocated to incident investigations based on the risk associated with the incidents' recurrence. To this end, the fundamentals for such a scheme are presented below.

7.2 Risk as a Classification Metric

Risk is defined as the probability that an event will occur multiplied by the anticipated cost derived from the occurrence of the event. When an incident occurs, it suggests the presence of a deficiency in the safety systems involved that, if not corrected, could lead to recurrences of the incident. A useful measure of the importance of an incident, therefore, is the total

risk that society faces if nothing is done to prevent recurrences. The total risk of such a recurrence is given in Equation 1 below.

$$\begin{aligned} \text{Total Risk} &= E[\#\text{ Recurrences}] \times E[\text{Cost}] \\ &= P[\text{Incident Recurrence}] \times \text{Exposure} \times E[\text{Cost}] \end{aligned} \qquad \text{Eq. 1}$$

The term E[# Recurrences] represents the expected number of recurrences of the incident if nothing is done to reduce the likelihood of recurrence and is the product of P[Incident Recurrence], the probability that the incident will happen again, and Exposure, the number of opportunities for the incident to recur. The term E[Cost] is the expected cost of the incident given that it has occurred and is defined in Equation 2 below.

$$E[\text{Cost}] = \sum_{i \in S} Cost(i) \cdot P[i] \qquad \text{Eq. 2}$$

Equation 2 is simply the expectation of the random variable Cost associated with a particular incident. S represents the set of all possible outcomes that might result from the occurrence of the incident. For each possible outcome i, the cost of i, namely the loss, is multiplied by the probability that i occurs. The summation of these products yields the expected value of the random variable Cost, which is the expected cost of the incident.

As defined earlier, Exposure is the number of chances for an incident to occur. If a particular system has a chance of contributing to an incident each time it is operated, then the exposure from the system is the number of times the system is operated multiplied by the number of such systems in existence. When the system in question is used widely and frequently, this number can become quite large. For example, consider the in-flight breakup of

TWA flight 800 over the Atlantic Ocean in 1996. The NTSB concluded that the probable cause of the accident was an explosion of the aircraft's center wing fuel tank, and the Board identified design issues affecting all Boeing 747 airplanes [35]. Exposure in this case would be the number of Boeing 747s in operation multiplied by the number of flights each aircraft would be expected to make in its lifetime. Given the popularity of the 747 and the near impossibility of surviving a commercial aircraft breakup at cruise altitude, the Exposure and *E*[Cost] terms of the Total Risk equation would be very large, stressing the importance of implementing the Board's recommendations and reducing *P*[Incident Recurrence] in order to reduce the risk to an acceptable level.

The terms *P*[Incident Occurrence], Exposure, and *E*[Cost] follow one's intuition in prioritizing incidents. Clearly, an incident with a high probability of recurrence with high expected costs warrants significant investigation, particularly if numerous systems are already deployed that might also be susceptible to the incident. Likewise, an incident with a small probability of recurrence, a low expected cost, or for which there are only a handful of susceptible systems that are rarely used might warrant only a minor investigation. Thus, Total Risk can be used as a metric to prioritize incident investigations and determine where investigative resources would be best spent and which areas regulators, aircraft operators, and equipment manufacturers should focus on first when following up on investigators' recommendations.

As a second example of the use of Total Risk, consider the incident involving British Airways flight 027. It is very difficult to estimate the probability of recurrence but not impossible. The rates of failure of the relevant hardware components are probably known as is the rate of undetected damage occurring during maintenance. The cost of such an incident were it to result in an accident would be very high since there would be considerable loss of life and

equipment. Exposure is also likely to be very high because of the prevalent use of TCAS. Thus, a rough estimate of the total risk could be calculated quickly and used as an indicator of the significance of the incident.

Estimating the terms of the Total Risk equation above requires a degree of familiarity with the incident under consideration. Unfortunately, very little information is typically available immediately following an incident, and so some terms could be difficult to estimate. For example, to determine *P*[Incident Recurrence] and *E*[Cost] one must first know what kind of incident has occurred, and to determine Exposure one must know which systems were involved. The following sections provide possible guidelines for making an initial estimate of Total Risk and refining the estimate as the investigation progresses.

7.2.1 Estimating Recurrence and Cost

An initial assessment of Total Risk should not be performed until investigators have categorized an incident (e.g. rejected takeoff, loss of separation, descent below MSA, gear-up landing, etc.). Both *P*[Incident Recurrence] and *E*[Cost] depend on the type of incident, and a good estimate of Total Risk cannot be made without knowing these terms. Fortunately, incidents can usually be assigned to one of these general categories within a few days of their occurrence. Until enough information is available to make an initial Total Risk assessment, investigators should treat the incident with high priority.

Once an incident has been categorized, *P*[Incident Recurrence] and *E*[Cost] may be estimated using statistics for similar incidents. Since casualty figures are among the first details to emerge from an investigation, they may be compared against loss statistics for similar incidents to estimate *E*[Cost]. Likewise, *P*[Incident Recurrence] may be estimated using the rate of occurrence for the incident's general category. As investigators learn more about

the event sequence leading to the incident, such as what failures occurred or where faults might be present, they can refine the estimates to achieve a more accurate Total Risk assessment.

For example, consider a loss of separation incident involving two aircraft. *P*[Incident Recurrence] and *E*[Cost] might initially be estimated using the general rate for loss of separation incidents and the expected outcome of such an incident. If during the investigation it is discovered that a design fault in TCAS contributed to the incident, *P*[Incident Recurrence] could be refined using the conditional probability of an incident given the existence of a design fault in TCAS. This refinement would raise the probability of recurrence, thus increasing Total Risk and the priority of the investigation. On the other hand, if TCAS manufacturers later announce that they will correct the fault, *P*[Incident Recurrence] could decrease based on this new information, lowering Total Risk.

7.2.2 Estimating Exposure

Even after an incident has been categorized, it is still unlikely that investigators will know which systems contributed to it until later in the investigation, and Exposure will remain unknown. Until then, this term might be estimated using the aggregate number of flights expected to by made by same-model aircraft or, if a ground-based system is suspected of contributing to the incident, the number of facilities using the same system. If interactions between systems are suspected of contributing to the incident, Exposure can be taken as the number of instances in which the systems are used together.

To illustrate this point, consider the crash of Korean Air flight 801 into Guam. After categorizing the incident as a descent below MSA, investigators might initially estimate Exposure using the entire fleet because they are unsure which systems actually contributed to

the incident. This estimate results in a high initial Total Risk value, but also means that the investigation will be given higher priority until more is known. When investigators discover how the MSAW inhibition contributed to the incident, they could then reevaluate *P*[Incident Recurrence] as the probability that an aircraft descends below its MSA given that the MSAW system is inhibited, *E*[Cost] as the expected outcome of such an occurrence, and Exposure as the number of MSAW installations throughout the NAS multiplied by the number of landings at a typical airport where MSAW is used.

7.2.3 Multiple Systems

Complications arise when multiple systems contribute to an incident independently of one another. In this scenario, separate Total Risk assessments might be necessary for each contributory system, and the overall Total Risk for the incident could be taken as the sum of the individual assessments. Again using Korean Air flight 801 as an example, the unavailability of the glideslope and the MSAW inhibition were independent of each other, and each contributed to the accident. By themselves, each of these problems raised the risk associated with landing at Guam, but by occurring simultaneously this effect was exacerbated. To account for this fact, separate Total Risk assessments could be performed for the glideslope and the MSAW system to compute Total Risk for the incident. The separate assessments could also be used to determine which system to investigate first; systems with greater risk could be given higher priority.

7.2.4 Confidence

The Total Risk metric defined in Equation 1 is an estimate, and like any estimate it has some degree of error associated with it. The error associated with a Total Risk estimate deter-

mines the bounds for a confidence interval on the estimate. Initially, the confidence interval will be large because the Total Risk estimate for an incident will be based on preliminary information and general statistics for the category to which the incident belongs. As investigators discern more about the event sequence leading to the incident and begin to identify the systems involved, they will be able to estimate the terms of Equation 1 with greater precision, which will narrow the confidence interval for Total Risk. Narrowing the confidence interval is important because investigators will rely on the Total Risk estimate to decide where to focus their efforts. If the estimate does not accurately reflect the true risk of recurrence for an incident, investigators might waste time investigating relatively minor incidents instead of other potentially more important ones.

7.3 Follow-up Actions

A second important use of the concept of Total Risk is to guide the actions taken following an investigation. If Total Risk is high, then the follow-up actions should have a high probability of reducing it to an acceptable level. Many options are available to investigative and regulatory agencies and they need to be used carefully. At one extreme is the option of grounding the fleet and at the other there is the option of no action. In between, there are a variety of possibilities including required inspections, required equipment replacement, required equipment redesign, and so on. There are also options about how quickly any action should occur. Selection among options is a difficult activity if there is no effective mechanism for rating the seriousness of an incident.

Using British Airways flight 027 as an example once more, the actions taken following the incident were insufficient and fragmented despite the fact that Total Risk by the esti-

mation above was very high. Upon concluding its investigation, the CAA issued an airworthiness directive requiring air carriers using similar equipment to check and periodically inspect the equipment to ensure that it is functioning properly and notified other aviation regulatory agencies as well as equipment manufacturers of the problems it found. It also issued a recommendation to aircraft operators urging them to consider using other encoding schemes for transmitting altitude data since that was part of the problem. The CAA's recommendations did not require mandatory changes and the probability that they would reduce total risk to an acceptable level was small. More importantly, the report by British Airways contains useful insights about the incident yet it has not been made public nor led to appropriate general recommendations.

7.4 Iterative Reclassification

As an incident investigation proceeds, new details will emerge that affect the risk of future recurrence. The terms comprising the Total Risk equation will change as the breadth of possible event sequences is narrowed, faults are identified, and remedies are enacted. Consequently, new Total Risk assessments will periodically need to be made, and an investigation's priority relative to others will rise and fall as it is reclassified. After developing an initial set of recommendations, investigators might find that the risk associated with an incident has been reduced to the extent that their efforts would be better spent investigating other incidents with higher Total Risk assessments. Moreover, each reassessment will presumably lower the error in the estimate and thus narrow the bounds of the confidence interval. Relying only on the initial Total Risk estimate is insufficient because this estimate is based on preliminary information and probably will not have a high degree of confidence associated with it. Therefore, in

addition to the Total Risk metric for classifying incidents, a process is necessary to reassess incidents periodically in order to improve the confidence associated with Total Risk estimates.

Until an incident has been categorized, the initial Total Risk assessment cannot be performed, and the investigation into the incident should be given a high priority. Once assessed, the incident can be investigated according to its relative priority among other incidents. Investigators might then choose to reassess the incident on a strictly periodic basis (i.e. monthly or quarterly) or in light of major revelations concerning the investigation that might affect Total Risk, such as when a significant piece of evidence is discovered, when a defect is revealed, when a public inquiry is concluded, when recommendations are issued, or when remedies are implemented. Each reassessment will improve the confidence interval on Total Risk. If reassessing an incident causes its Total Risk to increase, the investigation should be intensified until the risk is mitigated; if Total Risk decreases, resources can be diverted to more urgent investigations. The investigation may be concluded when investigators are confident that Total Risk has fallen below a predetermined acceptable level, which may depend on the incident's categorization, the type of operation (commercial vs. general aviation, scheduled vs. unscheduled), the flight rules in effect, the type of aircraft, and possibly other factors.

The goal of investigating incidents is to learn lessons that help to prevent the incidents from recurring. Some incidents might be symptomatic of severe defects that could lead to future casualties if not corrected; others could be fairly straightforward and involve accepted risks. By employing the risk-based metric and process proposed above, investigators might be able to determine more accurately which incidents have greater potential for extracting important lessons. Doing so would enable them to allocate resources first to those investigations that would likely have the greatest impact on safety. As a result, investigative agencies could begin

to shift from a reactionary role in which loss motivates change to a proactive one focused on risk reduction.

7.5 Remaining Work

The notion of total risk is a starting point for a metric that will allow investigators to assess the importance of incidents more accurately and allocate investigative resources accordingly. By assessing incidents based on the risks of future losses from their recurrence rather than their immediate losses, investigators can be more proactive in detecting safety problems before they contribute to accidents involving casualties or damage to aircraft.

Much work remains to be done before this metric can be put into practice. Because incidents are rare occurrences, estimating their probabilities is difficult. A model of cost will be needed to assess the expected loss associated with an incident that takes into account fatalities, serious and minor injures, and damage to aircraft and other property. Moreover, the estimation techniques and reassessment process presented in this chapter are intended to serve as examples and are quite preliminary. Before they can be applied to any investigation, they must first be developed more fully and tested on sample incidents to determine their precision. Statistics concerning incident rates and casualties decomposed according to incident type must be computed in order to estimate the parameters comprising the Total Risk equation. While similar statistics already exist, it is unclear whether they are in a form suitable for this purpose. Perhaps most importantly, investigators will need to set acceptable risk levels and establish criteria for determining which level would apply to a given incident.

Once these challenges are overcome, the estimation and assessment procedures would need to be refined so that they could be employed in the field quickly. Total Risk assessment is

an overhead exercise and should not significantly detract from investigators' tasks of analyzing incidents and developing recommendations. While high precision cannot be expected from early estimates, they must be accurate enough to provide a rough indication of the worth of investigating an incident. Likewise, later assessments should help guide investigators in determining which aspects of the investigation to pursue next or whether to table the investigation and turn their attention elsewhere.

7.6 Summary

In spite of the unresolved issues, risk-based incident classification shows promise as an alternative to the loss-based schemes employed today. While loss might seem to be the more intuitive metric for classifying an occurrence, mitigating the risk of recurrence is what society hopes to accomplish by investigating incidents. By focusing on risk rather than loss in prioritizing incident investigations, investigators will more quickly and more comprehensively identify and address problems that pose significant threats to the safety of air travel. This approach will lead to fewer incident recurrences before a safety deficiency is corrected, reducing the opportunity for loss to occur.

8 Conclusions

Commercial air travel is one of the safety modes of transit available today due to the prompt attention given to accidents and the remedies put in place to prevent their recurrence. Although extremely safe, the system is still imperfect, and a handful of major accidents along with several less severe accidents and incidents happen each year. When these occurrences involve safety-critical software systems, they provide opportunities for investigators to identify system faults that, if not eliminated, could manifest themselves again and lead to recurrence of the incidents. To eliminate these faults, investigators must properly interpret the incidents and develop a set of recommendations that are comprehensive enough to cover the breadth of ways in which the faults could be triggered and disseminate those lessons so that appropriate remedies may be implemented.

The case studies of the Korean Air flight 801 and British Airways flight 027 incidents presented in this thesis indicate that this level of analysis is not being undertaken with regard to safety-critical software systems, and as a result important lessons are being overlooked. Both occurrences were preceded by similar incidents that warned of serious problems with the manner in which the MSAW and TCAS systems were developed and maintained. The Korean Air accident followed previous MSAW-related accidents such as the one at Dulles Airport in 1994, and the near-collision involving BA 027 followed a similar incident over Hawaii in 1995. Either the investigations into these incidents failed to uncover the underlying problems with the systems involved and develop suitable recommendations to prevent their recurrence or the remedies implemented were inadequate to resolve the problems.

8.1 The Systems Context

One problem was that the systems involved were treated as if they operated in isolated environments. In fact, they were part of much larger systems designed to enhance the safety of commercial air travel, and by viewing the systems in an isolated context, the designers and managers neglected to consider how their decisions could affect the larger systems and the overall level of safety. To overcome these issues, investigators must consider how subtle differences in the event sequence leading to an incident could lead to similar outcomes and develop recommendations that encompass these slight variations. Investigators and the avionics industry at large must remember to appreciate the larger role each safety system plays in enhancing the safety of commercial aviation and must consider how design and configuration management decisions regarding individual systems can affect safety overall.

8.2 Incident Classification

Current accident classification schemes used by investigative agencies to allocate resources to investigations place too great an emphasis on the immediate loss from an accident and as a result undervalue the importance of incidents with no losses. Consequently, incidents suggesting the presence of serious safety problems in onboard and ground-based systems are often ignored or not investigated with sufficient rigor to uncover these problems, which if left uncorrected could contribute to future incidents with more tragic outcomes. This dilemma was illustrated by the vast difference in the investigations conducted into the Korean Air flight 801 and British Airways flight 027 incidents, despite the observation that the latter could have easily developed into a major accident with almost twice the number of casualties as the Korean Air crash into Guam.

To mitigate this problem, investigators should reconsider the practice of classifying incidents based on their losses, and instead classify them based on the risk of future losses. Adopting risk-based schemes will allow investigators to be more proactive and address safety problems before they contribute to accidents with extensive casualties. For risk-based classification schemes to be useful, techniques will have to be developed for investigators to quickly assess the risk level of incidents early in the investigative process so that they can allocate resources accordingly.

Incidents are opportunities to correct safety problems without the losses associated with accidents. Many accidents are preceded by similar incidents, and by neglecting to treat incidents and accidents equally with respect to their potential for learning lessons, investigators are condemning themselves to correcting problems only when they result in casualties. Neither of the incidents presented in this thesis were unique; each had been preceded by previous occurrences. If investigators and regulators had addressed the configuration management and design issues surrounding the Korean Air flight 801 and British Airways flight 027 incidents when they manifested themselves earlier, these incidents probably would not have taken place.

9 References

1. National Transportation Safety Board. *Controlled Flight Into Terrain, Korean Air Flight 801, Boeing 747-300, HL7486, Nimitz Hill, Guam, August 6, 1997*. Aircraft Accident Report NTSB/AAR-00/01. Washington, DC.

2. Federal Aviation Administration. "Facility Operation and Administration." FAA Order 7210.3M. 29 February 1996. Washington, DC.

3. National Transportation Safety Board. *Public Hearing in Connection With the Investigation of Aircraft Accident, Korean Air Flight 801, B-747-300, Agana, Guam, August 6, 1997*. 24 March 1998. Honolulu, Hawaii.

4. Federal Aviation Administration. "Fact Sheet: CAST Accomplishments: Civil Aviation Controlled Flight Into Terrain (CFIT)." 26 March 2001. Washington, DC.

5. National Transportation Safety Board. "Guam ARTS-11A MSAW Chronology." *Korean Air Flight 801, B-747-300, Agana, Guam, August 6, 1997, Public Hearing Exhibit List*. docket no. SA-517, exhibit no. 3-U. 17 October 1997. Washington, DC.

6. Leveson, N.G. *Safeware: System Safety and Computers*. Reading, MA: Addison Wesley. 1995.

7. National Transportation Safety Board. *Controlled Collision with Terrain Transportes Aereos Ejecutivos, S.A. (TAESA) Learjet 25D, XA-BBA Dulles International Airport Chantilly, Virginia June 18, 1994*. Aircraft Accident Report NTSB/AAR-95/02. Washington, DC.

8. Federal Aviation Administration. *Aeronautical Information Manual*. Ch. 1, §1-1-9. 21 February 2002. Washington, D.C.

9. Ladkin, Peter B. "The Crash of Flight KE801, a Boeing 747-300, Guam, Wednesday 6 August, 1997: What We Know So Far." Article RVS-J-97-09. 11 September 1997.

10. U.K. Civil Aviation Authority. "Hazardous Loss of Separation Between Two Aircraft Over Chinese Airspace." Doc Ref KMH/Pap/059, issue 1. 28 October 1999. London, United Kingdom.

11. Carley, William M. "Wires Crossed: Flawed Safety Device In Jets Gets Blamed For a Near Catastrophe." *Wall Street Journal.* 12 October 1999, eastern ed.: A1.

12. MITRE Corp. "TCAS: Traffic Alert and Collision Avoidance System." 31 December 1998. <http://www.mitre.org/pubs/showcase/tcas/tcas.html>

13. Australian Transport Safety Bureau. "Safety Deficiencies: Errors in Traffic Alert and Col-

lision Avoidance Systems." Output no. R19990156. 9 September 1999. Canberra City, Australia.

14. Federal Aviation Administration. *Federal Aviation Regulations*. 14 CFR §91.3(b). Washington, D.C.

15. International Civil Aviation Organization. "Operation of ACAS Equipment." PANS-OPS (Document 8168), vol. 1, part VIII, ch. 3.

16. Federal Aviation Administration. "Advisory Circular: Air Carrier Operational Approval and Use of TCAS II." Advisory Circular 120-55B. 22 October 2002. Washington, D.C.

17. Laprie, J. C. *Dependability: Basic Concepts and Terminology.* Springer-Verlag. 1992.

18. Aviation Safety Network. "Aircraft incident description 31 JAN 2001 Boeing 747-446D." 31 July 2002. <http://aviation-safety.net/database/incidents/20010131-0.htm>

19. Federal Aviation Administration. "Aircraft Accident and Incident Notification, Investigation, and Reporting." FAA Order 8020.11B. 16 August 2000. Washington, D.C.

20. National Transportation Safety Board. "Accidents, Fatalities, and Rates, 2002 Preliminary Statistics, U.S. Aviation." <http://www.ntsb.gov/aviation/Table1.htm>

21. National Transportation Safety Board. "Accidents and Accident Rates by NTSB Classification, 1983 through 2002, for U.S. Air Carriers Operating Under 14 CFR 121." <http://www.ntsb.gov/aviation/Table2.htm>

22. National Transportation Safety Board. *Aviation Investigation Manual: Major Team Investigations*. Washington, D.C.

23. Transportation Safety Board of Canada. "Investigation Process." (18 September 2002). <http://www.tsb.gc.ca/en/investigation_process/what_we_do.asp>

24. Perrow, Charles. *Normal Accidents: Living with High-Risk Technologies*. Princeton: University Press. 1999.

25. National Transportation Safety Board. "We Are All Safer: NTSB-Inspired Improvements in Transportation Safety," 2nd ed. July 1998. Washington, D.C.

26. Federal Aviation Administration. *Blueprint for NAS Modernization.* October 2002. Washington, D.C.

27. Commercial Aviation Safety Team. "Western-Built Transport Hull Loss Accidents, by accident site, 1988 through 1997." 10 November 1998. <http://www.aia-aerospace.org/departments/civil/cast/charts.html>

28. National Transportation Safety Board. “NTSB History and Mission.” <http://www.ntsb.gov/Abt_NTSB/history.htm>

29. Aviation Safety Reporting System. “Program Overview.” <http://asrs.arc.nasa.gov/overview.htm>

30. Johnson, Chris. “The Limitations of Aviation Incident Reporting.” <http://www.dcs.gla.ac.uk/~johnson/papers/reminders/>

31. Main Commission Aircraft Accident Investigation Warsaw. *Report on the Accident to Airbus A320-11 Aircraft in Warsaw on September 14, 1993*. March 1994.

32. Federal Aviation Administration. “Common ARTS Description.” <http://www1.faa.gov/ats/atb/Sectors/Automation/CommonArts/description.htm>

33. Federal Aviation Administration. “STARS Facts.” <http://www1.faa.gov/ats/atb/Sectors/Automation/STARS/starsfacts.htm>

34. Department of Transportation, Office of Inspector General. “Follow-up on Federal Aviation Administration’s Acquisition of Standard Terminal Automated Replacement System.” Memorandum to Federal Aviation Administrator. 3 June 2002.

35. National Transportation Safety Board. *In-flight Breakup Over The Atlantic Ocean, Trans World Airlines Flight 800, Boeing 747-131, N93119, Near East Moriches, New York, July 17, 1996*. Aircraft Accident Report NTSB/AAR-00/03. Washington, D.C.