MITRE DATA COMMUNICATION LABORATORY INITIATIVE FOR DEPARTURE CLEARANCE (DCL)

John Gonda, Dongsong Zeng, Juliana Goh, Mike Bernock, and Jon Salisbury, The MITRE Corporation, McLean, Virginia

Abstract

In this paper, we present the preliminary results of a MITRE data communication initiative for FANS DCL procedures and requirements evaluation conducted in the MITRE Aviation Integration Demonstration and Experimentation for Aeronautics (IDEA) Lab and the Reconfigurable CNS/ATM Test (RCAT) Lab. Participants at the Integrated Communications Navigation and Surveillance (ICNS) Conference will find the data and observations of interest as the industry develops the avionics supporting data communications services and procedures.

The FAA Next Generation Air Transportation System (NextGen) and Single European Sky ATM Research (SESAR) Programs are becoming a reality. RTCA SC-214 and EUROCAE WG78 are jointly developing the data communication standards for safety, performance and Future Air Navigation System (FANS) interoperability requirements for Air Traffic Services (ATS). MITRE has been supporting the FAA Data Communication Program acquisition and implementation planning and is actively involved in SC-214 standardization. The FAA plans to implement a ground system, and operators are equipping with FANS avionics to support the Departure Clearance (DCL) service to be initially offered in the 2015 time frame. The work in this paper provides the first data towards SC-214/WG78 standards validation through laboratory experiments and data measurement for the Departure Clearance Service.

The IDEA lab provides cockpit simulators that can be used for human in the loop (HITL) exercises and pilot response time measurement. The MITRE RCAT lab has an operational Flight Management System (FMS), Communication Management Unit (CMU), VHF Digital Radio (VDR), access to Data Service Provider (DSP) networks, and is capable of measuring the actual communication technical performance for both VDL2 and VDL0 sub networks in the Boston metroplex. The measured pilot response time and the actual communication technical performance are then statistically combined to estimate the actual communication performance at Boston for comparison with the required communication performance (RCP) in the RTCA SC 214 draft standards.

1. Introduction

In support of NextGen and SESAR Programs, RTCA SC-214 and EUROCAE WG78 are jointly developing the data communication standards for safety, performance and Future Air Navigation System (FANS) interoperability requirements for Air Traffic Services (ATS). Meanwhile, the U. S. FAA Data Communication program plans to implement Tower Data Link Services (TDLS) with FANS avionics to support the Departure Clearance (DCL) in the 2015 time frame. MITRE has been supporting the FAA Data Communication Program acquisition and implementation planning and is actively involved in SC-214 standardization.

The data communications program is working toward a second Final Investment Decision (FID) with several key operational questions related to the interaction between pilots and controllers that must be addressed. Work is underway within the international community on the development and validation of the harmonized standards for FANS and ATN Safety, Performance, and Interoperability. MITRE's analytical, operational, and aircraft and ground station human in the loop (HITL) simulation capabilities afford essential support to these activities. In particular, empirical data obtained from simulation studies can be used as objective evidence on key issues that determine the success of the investment decision, but much more needs to be done beyond the limited scope of the preliminary work outlined in this report.

In this report, we present the preliminary results of a MITRE data communication initiative for FANS DCL procedures and requirements evaluation conducted in the MITRE Aviation Integration

Demonstration and Experimentation for Aeronautics (IDEA) Lab and the Reconfigurable CNS/ATM Test (RCAT) Lab. This work also has progressed in parallel with the Industry Data Comm Flight Trials using a prototype ground system, the DCL Trials Automation Platform (DTAP) and user aircraft equipped with operational FANS avionics. The work in this paper is of particular importance to SC-214/WG78 standards validation and provides the first operational FAA program decision support through laboratory experiments and data measurement for the Departure Clearance Service.

2. Human-in-the-Loop Simulation

2.1 Scope

A multitude of factors may affect pilot response time to departure clearances. The set of scenarios used in this HITL are representative of nominal flight operations. Non-normal events, such as resolving an emergent system failure, were not introduced in the HITL. In addition, the way in which flight crews manage their displays to input and verify route information can also affect their response time. For example, limited number of displays may require pilots to time-share the displays or print clearances that otherwise would not be required when more display space is available.

There are two types of pilot procedures in response to departure clearance: "process-then-accept" and "accept-then-process". In the "process-then-accept", pilots receive the message, read and enter the message elements into the avionics, and then accept the clearance when both pilots have verified the accuracy of the information that had been entered. In the "accept-then-process" approach, pilots read the message and then accept the message before entering the information into the avionics. This HITL used the "process-then-accept" set of procedures. It is expected that pilot response time for "process-then-accept" procedures will be longer than that for "accept-then-process" procedures.

Figure 1 shows a Boeing 737/757/767-like displays configuration, in which two Multi-Function Control and Display Units (MCDUs) are available to the flight crew and both the receipt and input of the departure clearance occurs via the MCDU.

Figure 1. IDEA Lab Cockpit Configuration

2.2 Test Design and Procedure

Thirteen flight crews participated in the HITL simulation. All participant pilots were current air transport pilots from various airlines. They had ratings on Embraer 145/190, Canadair Regional Jet, Boeing 737/757/767/777, and Airbus 319/320 aircraft. Their total flight hours ranged from 1060 hours to 27000 hours (mean = 12500 hours).

Four scenarios were developed for the HITL. Figure 2 provides an overview of the scenarios. This set of four scenarios was performed twice by flight crews, once by auto loading the clearances and once by manually loading them. Half the crews ran the scenarios using auto load first, then manual load while the other half of the crews did the opposite.

Each scenario began at the gate with the preflight complete and performance data entered in the MCDU. Flight crews started each scenario by requesting a departure clearance. This initial clearance used the UM80 message type. The first DCL revision, depending on the scenario, was either a simple or complex revision and received while the flight crew was away from the gate but in or close to the ramp area. If the first revision is complex, the flight crew received a revision to either the initial part of the route (UM79) or the latter part of the route (UM83). The second revision was received while the crew was taxiing to the departure runway. Similar to the first revision, this second revision was either simple or complex and if complex, either a change to the initial or latter part of the route.

At the end of each scenario test, the flight crews were asked to complete a post-scenario questionnaire.

At the end of the test day, flight crews were asked to complete the post-experiment questionnaire and debriefed on the goals of the study.

INITIAL CLEARANCE (at the gate)	1st REVISION (ramp/close to ramp area)	2ND REVISION (movement area)	
Initial DCL (UM80)	Simple	Complex (UM79)	S1
Initial DCL (UM80)	Complex (UM79)	Simple	S2
Initial DCL (UM80)	Simple	Complex (UM83)	S3
Initial DCL (UM80)	Complex (UM83)	Simple	S4

Figure 2. Overview of Scenarios

2.3 HITL Results

2.3.1 Post Scenario Questionnaire Results

2.3.1.1 Pilot Workload

Pilot workload was measured using a modified NASA-TLX on a seven point Liker Scale (1= very low/very successful, 7 = very high/not successful at all). Overall, the average rating for each of the six dimensions of workload was less than four which is the mid-point of the scale. Comparing the workload ratings on the individual dimensions, flight crews rated the Mental Demand involved in completing the datalink task to be higher with Non-Integrated Avionics (mean = 3.37) than with Integrated Avionics (mean = 2.54). This difference is statistically significant ($p<0.01$) as revealed by a t-test. This is also true for Physical Demand (Non-Integrated: mean = 2.35, Integrated: mean = 1.73, $p<.01$), Temporal Demand (Non-Integrated: mean = 2.98, Integrated: mean = 2.34, $p<0.01$), and how hard the crews had to work to complete the task (Non-Integrated: mean = 3.22, Integrated: mean = 2.54, $p<0.01$). While the average rating for how discouraged the flight crews felt tended to be higher with Non-Integrated Avionics (mean = 2.44) than with Integrated Avionics (mean =2.18), this difference was not statistically significant ($p>0.05$). Finally, regardless of type of avionics, flight crews rated themselves as equally successful at completing the datalink task (Non-Integrated: mean = 1.71, Integrated: mean = 1.68), $p >0.05$.

2.3.1.2 Content and Format of Datalink Message

Using a seven point Likert Scale (1: strongly agree, 4: neutral, 7: strong disagree), flight crews were asked to rate their level of agreement with statements regarding the acceptability of the presentation format and content of initial and revised departure clearance at the gate and while taxiing, from an operational/safety perspective.

Overall, average ratings on the acceptability of the format and content of the messages are better than neutral (rating of 4) regardless of type of avionics and location of receipt of the message. Regarding the acceptability of the format of the message received at the gate, there was no statistically significant difference ($p>0.05$) in the level of agreement between integrated avionics (mean = 2.13) and non-integrated avionics (mean = 1.86) as revealed by a t-test. This was also true for the acceptability of the format of the messages received while taxiing (integrated mean = 2.74 vs. non-integrated mean = 2.54), $p>.05$. Regarding the acceptability of the content of the message received at the gate, there was a statistically significantly ($p<0.05$) greater agreement that it was acceptable for non-integrated avionics (mean = 1.78) than with integrated avionics (mean = 2.50). There was however no statistically significant difference in the mean ratings regarding acceptability of the content of the messages received during taxi (integrated mean = 2.95 vs. non-integrated mean = 2.88), $p>0.05$.

2.3.1.3 Understanding the Datalink Message

Using a seven point Likert Scale (1: strongly agree, 4: neutral, 7: strong disagree), flight crews were asked to rate their level of agreement with the statement that they understood the datalink messages sufficiently to complete the task safely and effectively.

The average ratings show high levels of understanding of the messages regardless of type of avionics (Integrated: mean = 2.01, Non-Integrated: mean = 2.06, $p>0.05$).

2.3.1.4 Perceived Heads Down Time

Flight crews were asked to rate how much more or less time they perceived to have spent heads down when receiving a clearance revision in the movement area using datalink, compared to voice communications. Flight crews used a seven point Likert scale to provide their ratings (1: Much Less Time, 4: About the Same, 7: Much More Time).

Flight crews rated perceived heads down time when receiving a clearance revision via datalink in the movement area to be about the same as voice when using Non-Integrated Avionics (mean = 3.9), and slightly less than voice when using Integrated Avionics (mean = 3.06). This difference in rating is statistically significantly different ($p<0.01$) as revealed by a t-test.

2.3.2 Post-Experiment Questionnaire Results

2.3.2.1 Likelihood to Continue Taxiing

The pilots were asked to rate their agreement with the statement that they would continue taxiing in the movement area when they receive a route clearance that has to be manually (non-integrated avionics) or auto-loaded (integrated avionics) (1: Strongly Agree, 4: Neutral, 7: Strongly Disagree).

Pilots rated themselves relatively neutral to the likelihood that they would continue taxiing when receiving a clearance revision in the movement area, regardless of whether the clearance had to be auto-loaded (mean = 3.55) or manually loaded (mean = 4.38). This difference was statistically significant ($p<0.05$) as revealed by a t-test.

2.3.2.2 Adequacy of Procedures

The pilots were asked to rate their agreement with the statement that the procedures for handling auto-loadable or manually loaded departure clearances via datalink were adequate (1: Strongly Agree, 4: Neutral, 7: Strongly Disagree).

The pilots tended to agree that the procedures for handling departure clearances via datalink were adequate (Integrated: mean = 2.26, Non-Integrated: mean = 2.9). This difference in ratings is statistically significant ($p<0.05$) as revealed by a t-test.

2.3.2.3 Pilot Preference

The pilots were asked about their preference for mode of communications in the delivery of departure clearances in two ways. First, they were asked to rate their level of agreement with the statement that they prefer the use of datalink over voice communications when there are simple or complex revisions to the departure clearance. Second, they were asked to rank order their preference for auto loadable datalinked departure clearance, manually loaded departure clearance, and departure clearance via voice.

Pilots generally agreed that they preferred the use of datalink over voice for both simple (mean = 1.64) and complex (mean = 2.27) revisions to the departure clearance. This difference was not statistically significant, $p>0.05$ as revealed by a t-test

In terms of ranking their preferences, all pilots ranked the use of auto-loadable datalinked departure clearance as the most preferred method for delivering departure clearances. 19 of the 25 pilots ranked manually loaded departure clearances as the second most preferred method followed by departure clearances delivered by voice.

2.3.3 Qualitative Observations

This section describes qualitative observations that were made by the experimenter during the test sessions.

2.3.3.1 Display and Information Management

In general, the flight crews would use one MCDU to display the message and the other MCDU to input the information. In the HITL test, it was observed that the Pilot Not Taxiing (PNT) would reach across the aisle to configure the Pilot Taxiing's (PT) MCDU during complex revisions away from the gate. The PNT would look across the aisle to refer to the PT's MCDU while typing in the clearance in his/her own MCDU. The PT would periodically glance across at the PNT's MCDU to see what the PNT was typing.

If there is only one MCDU and no printer onboard, flight crews may need to develop adaptive strategies such as transcribing the clearance onto paper before entering the information or using personal electronic devices which have cameras (e.g., smartphones) to take a photograph of the clearance message and using that photograph as a reference.

2.3.3.2 Crew Coordination

Observations were made regarding how the flight crews defined their roles when performing the DCL procedure. For all crews, one pilot was designated the Captain and the other the First Officer and they were provided time before the test scenarios to discuss how they would allocate their roles and responsibilities.

At the gate, depending on the airline that the pilot was from, the Captain would sometimes request and load the flight plan and the First Officer would verify the Captain's entries. The majority of flight crews, however, had the First Officer performing this task at the gate. When loading clearances in the non-integrated avionics scenarios, it was often observed that one pilot would read each individual waypoint to the pilot who was entering the information in the MCDU.

While taxiing, the First Officer would read out and load the departure clearance message. The Captain would then verify the First Officer's entries. There was however, varying levels of verification from the Captain. Some Captains would glance across the flight deck periodically to observe the First Officer's entries, while others would pull over to check the entries him/herself.

2.3.3.3 Task Prioritization

Observations were made regarding how pilots prioritized their tasks when a datalink message was received on the flight deck. Since receiving and entering the initial departure clearance was the only task that pilots had to complete when the aircraft was parked at the gate, the issue of task prioritization was particularly important when a departure clearance revision was received while the crew was taxiing.

In general, regardless of the type of avionics interface that was being used, pilots may prioritize other tasks over dealing with the datalink messages. For example, if the flight crew was completing a checklist when a datalink message was received, the flight crew would complete the checklist before checking the datalink message. Similarly, if the Captain was making a turn onto a taxiway, the First Officer would ensure that the aircraft was clear of other aircraft before going heads down to deal with the datalink message.

4 Technical Performance Assessment

In a general DCL transaction scenario as shown in Figure 3, the DCL transaction starts when the controller composes a message. Then, the message is sent to the pilot. The time from message release by the controller to message arrival at the pilot is the uplink communication delay. The pilot receives the message, processes it, and then sends a response back to the controller (Process then Accept Procedure). The time from clearance message receipt to a response decision sent is called pilot response time. The time from the response leaving the aircraft to the response message arriving at the ground controller is the downlink communication delay. The time from controller receiving the response message to controller recognizing the response message is called recognition time. The uplink communication delay plus the downlink communication delay together is called Required Communication Technical Performance (RCTP). The time from the controller composing the clearance message to controller recognizing the response message is called the RCP time. In DCL scenarios, where the clearance is delivered by the automation in response to a pilot downlink request; the DCL messages are automatically generated by application software, so the message composition time and recognition time are negligible.

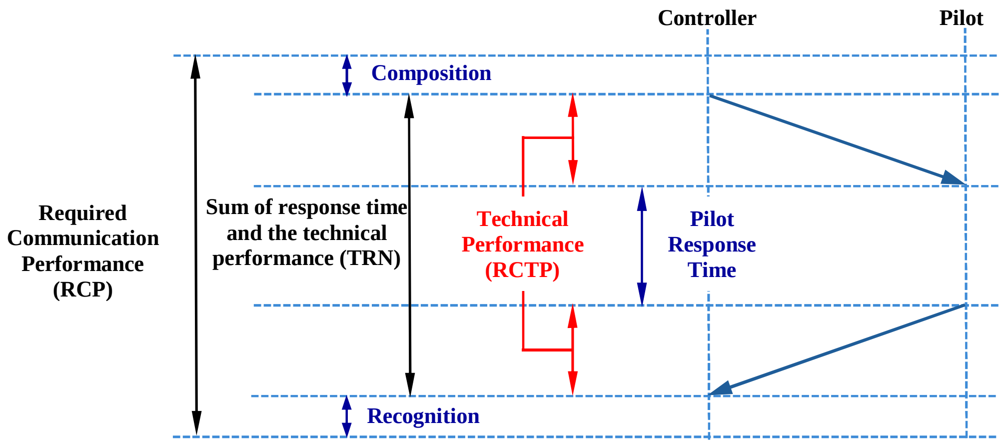

Figure 3. Exchange Transaction Diagram

In the current SC-214 data communication SPR [1], RCP400 is the performance requirement for the DCL Service. In the RCP400 requirements, the 99.9th percentile of the Actual Communication Performance (ACP) time is required to be less than 400 s, the 95th percentile of the Actual Communication Performance is required to be less than or equal to 174 s. The 99.9th and 95th percentiles of Pilot Operation Response Time (PORT) are required to be less than or equal to 371 s and 161 s, respectively. The 99.9th and 95th percentiles of Actual Communication Performance (ACTP) are required to be less than or equal to 32 s and 18 s, respectively.

4.1 ACTP Measurement Procedure

The ACTP measurement test was conducted using the MITRE RCAT Lab. Basically, the ACTP test set consists of FANS ATC workstation, live Datalink Service Provider (DSP) network (Aircraft Communications Addressing and Reporting System (ACARS)), and aircraft station with commercially compatible FANS avionics. The FANS ATC workstation generates a DCL message and sends it to an aircraft station through either ARINC or Societe Internationale de Telecommunication Aeronautiques (SITA) data network.

Figure 4 demonstrates the ACTP test setup diagram, in which the FANS ATC workstation measures the ACTP as the time from sending a DCL message out to receiving a response back from the aircraft station. Upon receiving the Current Data Authority (CDA) message, the aircraft avionics sends a response back to the ATC workstation instantaneously.

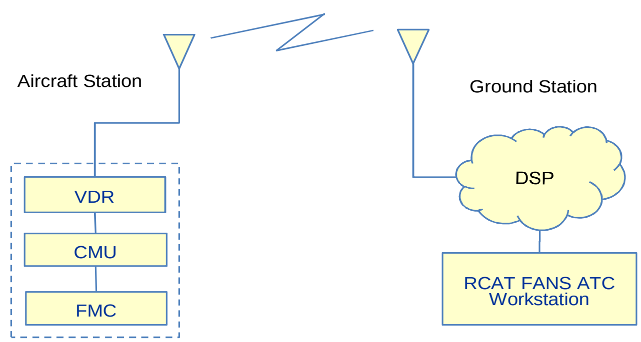

Figure 4. RCTP Measurement Setup Diagram

The aircraft avionics set consists of VHF Digital Radio (VDR), Communication Management Unit (CMU), and Flight Management Computer (FMC). In the RCAT avionics test set, the VDR is Honeywell RTA-44D (P/N: RTA44D-06728), the CMU is either Rockwell Collins CMU-900 (P/N: 822-1239-151) or Honeywell Mark II (P/N: 965-0758-001) CMU, and the FMC,") is a GE/Smiths 702A-4 (P/N: 2907A4), which has a FANS software load U10.5. All the modular units (VDR, CMU and FMC), with applicable software, reside "behind the glass". The displays shown in the RCAT cockpit provide the human computer interface associated with each.

The ACTP (Technical Performance) measurement procedure consists of the following 7 steps:

1. The RCAT ATC workstation establishes CDA as KESC (four-letter ICAO identifier) and Next Data Authority (NDA) as KRCT.
2. The aircraft station Airways Facilities Notification (AFN) application logs on to the ATC workstation.

3. ATC starts a CPDLC session with the aircraft station.
4. The ATC workstation sends departure clearance messages to aircraft from Next Data Authority (NDA) (KRCT) through a live sub-network using single-block (150 bytes) and double-block (300 bytes) messages that are generated in equal numbers.
5. The aircraft station receives the departure clearance and automatically replies with Not Current Data Authority (NCDA) message which has the same message size as WILCO.
6. At the RCAT ATC workstation, the time difference from UM80 sending to NCDA receipt is measured as the ACTP.
7. Repeat this procedure over 1000 times for each subnetwork.

All of the ACTP data for both VDL-2 and Plain Old ACARS (POA) sub-networks was collected at the RCAT lab in the Bedford area. Data is not identified as having traveled over SITA or ARINC networks or sub networks.

4.2 VDL-2 ACTP Measurement Results

In the post-measurement data analyses, we employed two methods: empirical method and curve fitting method. The empirical method calculates the physical probabilities from the measurement samples directly. The curve fitting method searches for the distribution function which is best fit for the measurement data first, and then calculates the probabilities based on the best fitted distribution function. The empirical method assumes zero probability of the delay time exceeding the maximum observed delay, i.e., truncates the probability at the tail. This tail truncation may cause increasing error in high percentile (e.g., 99.9th percentile) estimation. On the other hand, the curve fitting method does not have this tail truncation issue. Therefore, in our post-measurement data analyses, we compared the results from both methods, and chose the results of curve fitting method as the final results due to the fact that our measurement sample size is relatively small.

For the VDL-2 sub-network, we collected a total of 1009 ACTP samples. Then we fit more than 60 known distribution functions to the measurement data and calculated three goodness-of-fit statistics: Komogorov/Smirnov, Anderson/Darling, and Chi-squared, using a statistical tool EastFit developed by Mathwave Technologies. The goodness-of-fit results demonstrate that the Dagum distribution function is the best fit function for VDL-2 ACTP measurement data.

The survival function captures the probability that a random variable is longer than a specified time. Survival function is also called complementary CDF, as it is equal to 1-CDF. Figure 5 demonstrates the survival functions of both empirical VDL-2 ACTP data and fitted Dagum distribution in log scale. For the 99.9th percentile (where RCP is measured), the empirical estimation is 32 s while the fitted Dagum method estimation is 26 s. For the 95th percentile, the empirical estimation is 10 s and the fitted Dagum function estimation is 9 s.

From Figure 5, we can see that the empirical survival curve truncates at the observed maximal point. This is because the empirical method assumes that the probability above the observed maximal point is zero. Due to the limited number of observations, the empirical method may underestimate the probability above the observed maximal point. On the other hand, the tail of fitted Dagum distribution function tapers off smoothly. With limited sample size, the fitted Dagum curve looks more reasonable than the empirical curve. When the sample size is big enough, the empirical method may be more trustworthy than fitted functions.

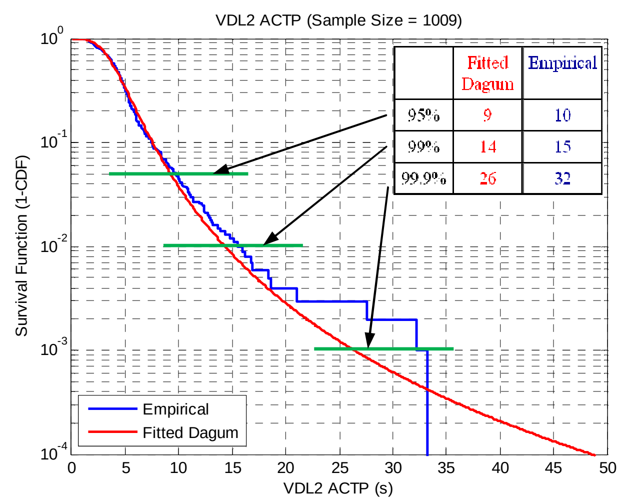

Figure 5. VDL-2 ACTP Estimated Percentiles

4.2 VDL 0 (POA) ACTP Measurement Results

For the POA sub-network, we collected a total of 997 ACTP samples. The goodness-of-fit results demonstrate that the Burr distribution function is the best fit function for POA ACTP.

Figure 6 plots the survival functions of both empirical POA ACTP data and the fitted Burr distribution in log scale. For the 99.9th percentile, the empirical estimation is 33 s while the fitted Burr method estimation is 29 s. For the 95th percentile, the empirical estimation is 12 s and the fitted Burr function estimation is 15 s.

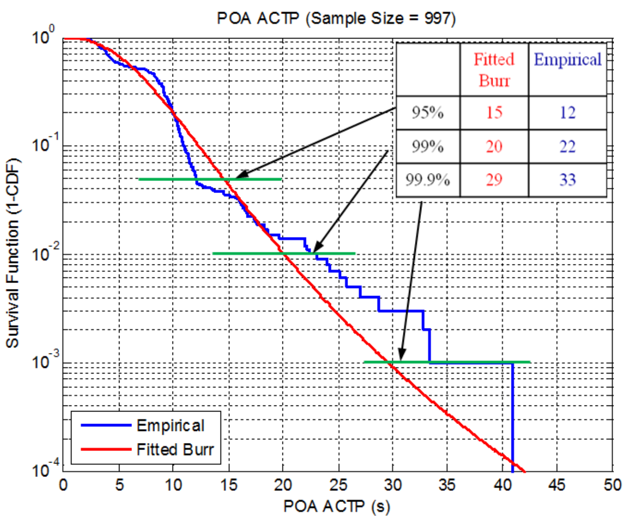

Figure 6. POA ACTP Estimated Percentiles

From the survival function picture, we can see that the empirical survival curve truncates at the observed maximal point. This is because the empirical method states that the probability above the observed maximal point is zero. Due to the limited number of observations, the empirical method may underestimate the probability above the observed maximal point. On the other hand, the tail of fitted Burr distribution function tapers off smoothly. With limited sample size, the fitted Burr curve looks more reasonable than the empirical curve. When the sample size is big enough, the empirical method may be more trustworthy than fitted functions.

4.3 DCL Pilot Response Time Measurement Results

4.3.1 DCL Pilot Response Time with Integrated Implementation

We collected a total of 223 PORT samples for integrated implementation. The goodness-of-fit results demonstrate that the Dagum distribution function is the best fit function for PORT data. Figure 7 plots the survival functions of both empirical PORT data and fitted Dagum distribution in log scale. For the 99.9th percentile, the empirical estimation is 276 s while the fitted Dagum method estimation is 360 s. For the 95th percentile, the empirical estimation is 152 s and the fitted Dagum function estimation is 156 s.

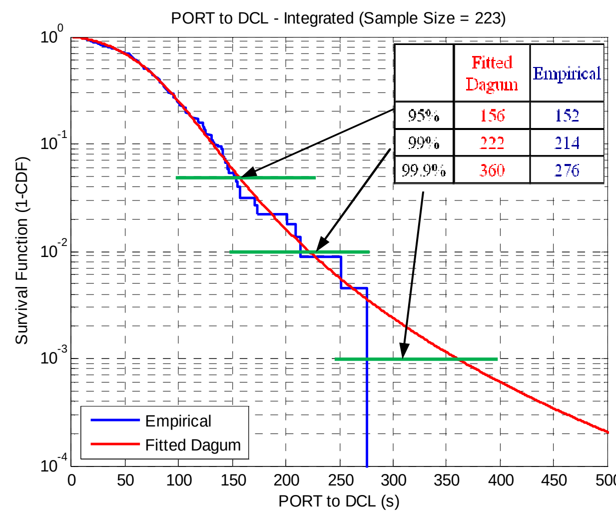

Figure 7. DCL PORT Estimated Percentiles with Integrated Implementation

From the survival function picture, we can see that the empirical survival curve truncates at the observed maximal point. This is because the empirical method states that the probability above the observed maximal point is zero. Due to a limited number of observations, the empirical method may underestimate the probability above the observed maximal point.

On the other hand, the tail of fitted Dagum distribution function tapers off smoothly. With limited sample size, the fitted Dagum curve looks more reasonable than the empirical curve. When the sample size is big enough, the empirical method may be more trustworthy than a fitted function.

4.3.2 DCL Pilot Response Time with Non-Integrated Implementation

We collected a total of 123 PORT samples for a non-integrated implementation. The goodness-of-fit results demonstrate that the Dagum distribution function is the best fit function for the PORT measurement data.

Figure 8 plots the survival functions of both empirical PORT data and fitted Dagum distribution. For the 99.9th percentile, the empirical estimation is 608 s while the fitted Dagum method estimation is 741 s. Similarly, for the 99th percentile, the empirical estimation is 393 s while the fitted Dagum function estimation is 448 s. For the 95th percentile, the empirical estimation is 310 s and the fitted Dagum function estimation is 310 s.

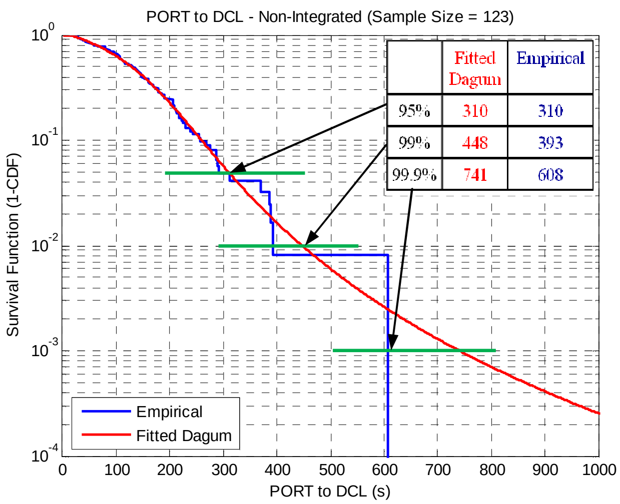

Figure 8. DCL PORT Estimated Percentiles with Non-Integrated Implementation

From the survival function picture, we can see that the empirical survival curve truncates at the observed maximal point. This is because the empirical method states that the probability above the observed maximal point is zero. Due to a limited number of observations, the empirical method may underestimate the probability above the observed maximal point.

On the other hand, the tail of fitted Dagum distribution function tapers off smoothly. With limited sample size, the fitted Dagum curve looks more reasonable than the empirical curve. When the sample size is big enough, the empirical method may be more trustworthy than fitted functions.

4.4 Required Communication Performance Assessment

4.4.1 Actual Communication Performance with Integrated Implementation

Table 1 compares the communication performance measurements with the corresponding requirements side by side, in which the red color represents measurement data and the green color denotes requirements.

For the 95th percentile, the required pilot response time is 161 s and the measured pilot response time is 156s. Our measurement data demonstrates that the pilot response time meets the requirements for the DCL Service. The required communication technical performance is 18 s. The measured VDL-2 communication technical performance is 9 s. The measured VDL-0 technical performance is 15 s. Our preliminary measurement data shows that both VDL-2 and POA can meet the required communication technical performance.

Table 1. Performance Requirement and Measurement Comparison—Integrated

		Pilot Response	Tech. Perf.	Sum of Pilot Resp. and Tech Perf.	Initiator	RCP
95%	Required	161 s	18 s	169 s	13	174 s
	VDL-2	156 s	9 s	(160,165) s [1]	Automated	(160,165) s [1]
	VDL-0		15 s	(164,171) s [1]	Automated	(164,171) s [1]
99.9%	Required	371 s	32 s	380 s	30	400 s
	VDL-2	360 s	27 s	(365,386) s [1]	Automated	(365,386) s [1]
	VDL-0		29 s	(368,389) s [1]	Automated	(368,389) s [1]
Note 1: (lower bound, upper bound), lower bound uses statistical sum while upper bound uses arithmetic sum.						

We used two methods to calculate the percentiles of the sum of measured pilot response time and measured ACTP. The lower bound method assumes that PORT and ACTP are independent random variables and uses a statistical sum to calculate the percentiles. On the other hand, the upper bound method assumes that PORT and ACTP are dependent and uses an arithmetic sum to calculate the percentiles. For the VDL-2 sub-network, the lower bound 95th percentile sum of PORT and ACTP is 160 s and upper bound 95th percentile sum of PORT and ACTP is 165 s. For VDL-0 sub-network, lower bound 95th percentile sum of PORT and ACTP is 164 s and upper bound 95th percentile sum of PORT and ACTP is 171 s. Our preliminary measurement data demonstrates that with an integrated implementation both VDL-2 and VDL-0 sub-networks could meet the required 95th percentile of communication performance, which is 174 s, for the DCL application. The lower bound 99.9th percentile sum of PORT and ACTP is 365 s and upper bound 99.9th percentile sum of PORT and ACTP is 386 s. For VDL-0 sub-network, lower bound 99.9th percentile sum of PORT and ACTP is 368 s and upper bound 99.9th percentile sum of PORT and ACTP is 389 s. Our measurement data demonstrates that with the integrated implementation both VDL-2 and VDL-0 sub-networks could meet the required 99.9th percentile of communication performance, which is 400 s, for DCL application.

4.4.2 Communication Performance with Non-Integrated Implementation

Table 2 compares the communication performance measurements with the corresponding requirements side by side, in which the red color represents measurement data and the green color denotes requirements.

For 95th percentile, the required pilot response time is 161 s. The measured pilot response time is 310 s. Our preliminary measurement data demonstrates that the pilot response time for DCL with a non-integrated implementation cannot meet the requirement of 161 s.

For the VDL-2 sub-network, the lower bound 95th percentile sum of PORT and ACTP is 315 s and upper bound 95th percentile sum of PORT and ACTP is 319 s. For the VDL-0 sub-network, lower bound 95th percentile sum of PORT and ACTP is 317 s and upper bound 95th percentile sum of PORT and ACTP is 325 s. Our preliminary measurement data demonstrates that a non-integrated implementation cannot meet the RCP400 95th percentile requirement of 174 s for the DCL Service. The lower bound 99.9th percentile sum of PORT and ACTP is 746 s and the upper bound 99.9th percentile sum of PORT and ACTP for the VDL-2 sub-network is 769 s. For the VDL-0 sub-network, the lower bound 99.9th percentile sum of PORT and ACTP is 748 s and the upper bound 99.9th percentile sum of PORT and ACTP is 771 s. Our preliminary measurement data demonstrates that a non-integrated implementation cannot meet the RCP400 99.9th percentile requirement of 400 s for the DCL Service.

Table 2. Performance Requirement and Measurement Comparison—Non-Integrated

		Pilot Response	Technical Performance	Sum of Pilot Resp. and Tech Perf.	Initiator	RCP
95%	Required	161 s	18 s	169 s	13	174 s
	VDL-2	310 s	9 s	(315,319) s [1]	Automated	(315,319) s [1]
	VDL-0		15 s	(317,325) s [1]	Automated	(317,325) s [1]
99.9%	Required	371 s	32 s	380 s	30	400 s
	VDL-2	741 s	27 s	(746,769) s [1]	Automated	(746,769) s [1]
	VDL-0		29 s	(748,771) s [1]	Automated	(748,771) s [1]
Note 1: (lower bound, upper bound), lower bound uses statistical sum while upper bound uses arithmetic sum.						

It is important to note that our sample sizes for PORT and ACTP are relatively small and the Radio Frequency (RF) channel is also lightly loaded in the Bedford Massachusetts area. So the results from our measurements are limited to our test conditions and may not be extensible to other conditions. In order to draw general conclusions, more measurement samples at various DCL operational airports are needed.

5. Summary and Next Steps

5.1 General Observations

The Data Comm Program has operational and performance requirements for their DCL implementation based on the evolving standards being developed in RTCA SC 214/EUROCAE WG 78. These standards will eventually form the baseline for a global set of ICAO guidance that will serve the NextGen, Single European Sky ATM Research (SESAR), and CARATS (Long-term Vision for the Future Air Traffic System in Japan) programs of the future.

The IDEA lab provides cockpit simulators that can be used for human in the loop (HITL) exercises and pilot response time measurement. The MITRE RCAT lab has an operational Flight Management System (FMS), Communication Management Unit (CMU), VHF Digital Radio (VDR), access to Data Service Provider (DSP) networks, and is capable of measuring the FANS DCL actual communication technical performance for both VDL2 and VDL0 sub networks in the Boston metroplex.

The measured pilot response time (IDEA lab) and the measured communication technical performance (RCAT lab) were statistically combined to estimate the actual communication performance at Boston for comparison with the required communication performance (RCP) in the RTCA SC 214 draft standards.

The work we are doing in our labs is helping to frame and answer relevant operational questions through operational experiments, which bring facts to inform decision making, including HITLs and Network performance measurements. These experiments:

- Reduce risk to the FAA Program by assessing Departure Clearance Service prior to Flight Trials
- Provide data to support development of cockpit procedures
- Provide input for Global Data Comm Standards and Data Comm Implementation

The DCL requirements we assessed are feasible for the part of the air/ground implementation being pursued by the FAA that we studied. Additional data must be collected to fully compare alternate avionics implementations should they become available.

Initial DCL performance requirements (RCP 300) were not attainable; therefore the Standards Committee revised the RCP upward to 400 based upon our earlier work. There still remains a question of whether RCP is needed at all for the DCL Service, since there is no separation requirement for surface operations. Significant safety questions remain, including the receipt of DCL revisions while taxiing. Operational trials will help to fully answer this question.

Compared to auto-loading clearances, manual loading of departure clearance takes longer time to respond. When using manual loading, pilots reported higher perceived heads down time and about the same heads down time as voice. Pilots also reported they were less likely to continue taxiing, were more likely to need additional "display location" to display the message and more likely to reach/read across the throttle. Given these observations, other workload factors, and wide variations within cockpit implementations, more work is needed to better understand the implications of auto-load and non-auto-load implementations. Additional HITLs and field trials should be conducted to examine the Pilot Response Time requirements of various different types of aircraft and installed avionics.

5.2 HITL

The findings from this HITL suggest the following when comparing the use of integrated (auto-loading) versus non-integrated avionics (manual loading) in the receipt of departure clearances.

When flight crews had to manually load departure clearances, they:

- Took longer to process and accept the message

• Reported higher mental, physical, and temporal workload

• Reported having to work harder at the task

• Reported perceived heads down time to be similar to voice communications

• Reported they were less likely to continue taxiing

• More likely to need additional “display location” (an actual physical display or a print out of the clearance) to display the message so that they may reference it while inputting the information

Indeed, the flight crews indicated unanimous preference for the use of integrated avionics so that the departure clearances may be auto loadable.

5.3 Technical Performance Assessment

On the basis of our limited measurement data and consequent data analyses, our preliminary observations are:

• The collected ACTP data of both VDL2 and VDL0 provide objective support for the RCTP allocations of the SC-214 data communication standards

• The ACTP measurements of both VDL2 and VDL0 sub networks provide objective data for FAA decision making in regards to Data Comm subnetwork requirements.

• For ACTP, the distribution functions with heavy tail capability such as Burr and Dagum functions demonstrate better fit for the measurement data than distribution functions without heavy tail capability such as exponential distribution.

5.4 Next Steps

Because the HITLs and Performance assessments we have completed are somewhat limited in scope, we believe that additional data collection and simulation can help to refine these preliminary results including:

• Collecting additional latency and human performance data at more highly congested airports and combining it with data gathered at the DCL Trials airports, and with different airlines and types of aircraft, such as at MEM, EWR, ATL, and IAD

• Conducting HITLs more focused on the ground side of the End to End System and using variations in the Human Computer Interface in the Tower.

• Collecting channel loading data in parallel with the latency data at various airports.

• Preparing position papers for standards activities which highlight the results of these analyses in the area of flight crew procedures.

• Extending the DCL HITLs to focus more on En Route services

References

[1] RTCA SC 214/EUROCAE WG 78 Document PU – 10 SPR-I-Annexes EFGHI-OPA-Feb 1_2012.doc.

Disclaimer

Email Addresses

John Gonda: jgonda@mitre.org

Dongsong Zeng: dzeng@mitre.org

2013 Integrated Communications Navigation and Surveillance (ICNS) Conference

April 23-25, 2013