Deep Space Optical Link ARQ Performance Analysis

Loren Clare
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109
818-354-1650
Loren.P.Clare@jpl.nasa.gov

Gregory Miles
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109
818-354-9191
Gregory.J.Miles@ jpl.nasa.gov

Abstract— **Substantial advancements have been made toward the use of optical communications for deep space exploration missions, promising a much higher volume of data to be communicated in comparison with present-day Radio Frequency (RF) based systems. One or more ground-based optical terminals are assumed to communicate with the spacecraft. Both short-term and long-term link outages will arise due to weather at the ground station(s), space platform pointing stability, and other effects. To mitigate these outages, an Automatic Repeat Query (ARQ) retransmission method is assumed, together with a reliable back channel for acknowledgement traffic. Specifically, the Licklider Transmission Protocol (LTP) is used, which is a component of the Disruption-Tolerant Networking (DTN) protocol suite that is well suited for high bandwidth-delay product links subject to disruptions. We provide an analysis of envisioned deep space mission scenarios and quantify buffering, latency and throughput performance, using a simulation in which long-term weather effects are modeled with a Gilbert-Elliot Markov chain, short-term outages occur as a Bernoulli process, and scheduled outages arising from geometric visibility or operational constraints are represented. We find that both short- and long-term effects impact throughput, but long-term weather effects dominate buffer sizing and overflow losses as well as latency performance.**

TABLE OF CONTENTS

1. SYSTEM OVERVIEW AND PROBLEM STATEMENT 1
2. SYSTEM DESCRIPTION 2
3. PERFORMANCE CHARACTERIZATION 5
4. PERFORMANCE RESULTS..... 5
5. SUMMARY 10
ACKNOWLEDGEMENTS 11
REFERENCES..... 11
BIOGRAPHY 11

1. SYSTEM OVERVIEW AND PROBLEM STATEMENT

Optical communications spanning deep space, such as between Mars and Earth, offers potentially much higher bandwidth than traditional RF-based communications. The optical link Earth terminal(s) might be ground-based or Earth-orbiting. We consider the ground-based case here, due to its much lower relative implementation cost and long-term upgradability. A ground-based Earth terminal is subject to Earth's atmospheric effects, in particular weather events such as a rainstorm that may persist for days during which a link outage occurs. These degrading effects are in addition to predictable outages arising from geometric visibility. Two key factors are considered in the following performance analysis of deep space mission resource requirements arising from optical communications:

- the resource capacity needed to accommodate the high data rates, and
- the impact of dynamic weather events.

Deterministically occurring link outages may arise due to mission operations constraints or over-subscription of resources. Commonly, predictable outages occur from geometric visibility. Such events may be mitigated by storing the data at the source (or intermediate relay point) whenever the link is known to be unavailable. Disruption-Tolerant Networking (DTN) [1, 2] provides protocols that automate this basic store-and-forward functionality. For example, a Mars orbiter will be in view of a single Earth ground station approximately 8 hours per day. During the two-thirds of each day when the ground station is unavailable, the DTN Bundle Protocol will retain the data until the next 8-hour pass becomes active, and will hold data remaining at the end of a pass until subsequent passes. Of course, data that is subject to extended storage must be "delay tolerant". This is often the case for deep space missions, which must operate in a high propagation delay environment in any case. Nevertheless, data delivery latency is an important metric for deep space missions. For example, a Mars rover requires return data delivery in time to plan and upload the next Martian day's operational commands. Also, envisioned deep space human space flight missions will impose more stringent timeliness requirements than typical robotic missions. Thus latency is a key metric of interest.

Large gaps in coverage may be mitigated with multiple ground stations spatially dispersed around the Earth; this is a fundamental feature of the NASA/JPL Deep Space Network (DSN). Some overlap occurs between ground station sites, allowing handover operations (including re-acquisition of links with the spacecraft). We focus on the single optical ground station case in the analysis below, so as to understand its benefits, with follow-on improvements arising as more optical ground stations are added in a budget-sustainable time sequence.

Individual transmissions may fail due to unpredictable channel errors, even if forward-error correction (FEC)

978-1-4673-7676-1/16/$31.00 ©2016 IEEE

techniques are used. We might simply accept such losses distributed randomly in the received data stream. Stochastic distribution of erasures in the received data may be determined by characterizing the physical and coding channel behavior.

However, missions generally demand high reliability on their data streams. This might be achieved with a large link margin, thereby reducing the chance of unpredictable error events. But such an approach is less effective as higher frequencies are employed. For example, research in deep space Ka-band communications has revealed very deep fades occurring during weather events, such that use of a sufficiently large link margin to guarantee low errors will destroy the advantages of Ka-band over X-band communications. Optical communications demonstrate even stronger effects, and therefore use of a higher link margin will not sufficiently solve this problem.

Instead, reliability is assured by means of retransmissions until the data is successfully received, that is, by using Automatic Repeat Query (ARQ) methods. ARQ requires a reverse channel in which each successfully received data unit is acknowledged. Each data unit transmitted is held at the transmitter in a queue until either the acknowledgment (ACK) is received or a timeout occurs, where the timeout accounts for the known round trip delay. If a timeout occurs, the data unit is retransmitted. The performance analysis below captures the operation of the optical link in which ARQ is used.

The throughput performance will depend on the type of ARQ used, discussed further below. ARQ will also impact the latency performance, as an individual data unit may require a random number of retransmissions until it is successfully received. These ARQ delays may also re-sequence the order of arrival of data units, which may impose additional delays when considering higher-layer information context, such as waiting until all data units forming a single large file are received (file-level latency).

Mission communications will very likely use a combination of optical and RF, particularly in early optical deployments. Also, traffic loading requirements for robotic missions are generally highly asymmetric, needing much higher rates on the return (spacecraft-to-Earth) link than in the forward direction. Therefore, we focus on an optical link as used for return data. We assume that the forward link (required for ARQ) is RF-based and is natively highly reliable, so that ACKs are not lost. For convenience we assume the forward link bandwidth is just sufficient to support ARQ ACKs, although in actuality the forward bandwidth will be higher and support multiplexing of forward user data. Lower-rate RF communications may also augment return data transfer, providing a highly reliable first-transmission capability for time-critical data, and/or backup for optical link outages to carry high-priority data. Such considerations require offered traffic to be separated into different classes, each with its own defined Quality of Service (QoS) requirement, in order to characterize performance; this extension will be deferred in the present analysis.

The next section provides a description of the system operation, factors that impact performance (particularly weather and how it's modeled), and measures of performance. Section 3 discusses the approach used to evaluate performance. Section 4 provides numerical results derived for a range parametric values. The final section provides a summary and conclusions of the work.

2. SYSTEM DESCRIPTION

ARQ Operation over the Optical Link

As indicated in the overview, ARQ is assumed used due to the unpredictable and unavoidable dropped transmissions that will arise with optical communications. Selective Repeat (SR) continuous ARQ is assumed, as simpler ARQ variants will suffer substantial performance degradation with the very high "delay-bandwidth product" environment of deep space communications. SR ARQ efficiently transmits each successful frame once provided ACKs are reliable. In addition, we assume disruption-tolerant ARQ such as the Licklider Transmission Protocol (LTP) [3] of the DTN suite. Specifically, the ARQ process is aware of the link schedule and correspondingly adjusts timeout timers. This provides some further gain in efficiency.

Frames await transmission in the Active Queue (AQ). Whenever AQ>0 and the link is active, frames are transmitted by the spacecraft. When the spacecraft transmits a frame, it moves it the frame from the AQ to the Pending Queue (PQ) and sets a timer equal to the roundtrip time (RTT) plus a small processing/guard time. The timer is adjusted for transmissions within a RTT of the end of pass, since their ACKs will be delayed by a known amount. For each successfully received frame, an ACK is transmitted to the spacecraft, which removes the frame from the PQ. If a frame times out, it is moved from the PQ back to the AQ. Such frames requiring retransmission are placed at the beginning of the AQ (receiving transmission priority) in order to minimize latency variance.

ACK transmission from the ground is assumed to be highly reliable, using an RF uplink capability. The ACK traffic load on the uplink will depend highly on the specific ARQ protocol as well as possible efficiencies from combining ACK information with spacecraft commands and other uplink user traffic. We defer developing a formal load analysis of ACK traffic in this paper.

We focus on a single optical ground station scenario, in which there is a single scheduled daily pass. If a second ground station is added, it will likely be positioned to provide additional geometric coverage with disjoint pass intervals. Scheduled service by multiple ground stations may be modeled as the union of their pass times and operating with ACK state transfers supporting handovers from one ground station to the next. Multiple ground

stations may alternatively be positioned with overlapping passes, providing weather diversity during dual coverage. In this case, the ARQ ground operation will ensure proper ACK generation to incorporate "besting" (successful reception if either station receives the frame correctly) and ACK radiation on the uplink.

Deep Space Mission Communications Dynamics

Deep space mission operations vary across a wide dynamical range. We focus on conditions that change over periods up to many days, such as a single extended weather event. Each mission can also vary at even slower rates, such as with the distance between the Earth and the spacecraft. For example, the distance to a Mars orbiter will vary between about 0.5AU to 2.5AU over the synodic cycle of approximately 780 days. Pass durations also vary as the view geometry changes slowly with time. Very substantial differences in optical link data rate capacity occur between these extremes. Also, ambient ground station weather patterns may slowly vary over the seasons. However we confine our model to be stochastically stationary, in which the defined system parameters are held to be constant. Extended performance characterization of mission-specific scenarios with these additional slow dynamics is briefly discussed in Section 3.

Input Variables and Performance Metrics

The general model is of a deep space spacecraft transmitting return data to Earth over an optical link. The optical link is subject to scheduled passes as well as unpredictable channel drops and outages. A forward channel from Earth to the spacecraft is modeled to carry ACK traffic to the spacecraft. The spacecraft maintains automated queueing mechanisms that implement an ARQ process with the ground.

<u>Link schedule</u>. For simplicity, we presume operations are based on a 24-hour cycle, recognizing that mission-dependent timing could also be a key schedule driver (for example, planning/execution on a Mars sol cycle). Typically we assume a single station for which the spacecraft endpoint is visible 8 hours per day, and that the link is scheduled for operation at all such times. An alternative daily "pass" value may be used and more precisely calibrated visibility duration or planned mission operation events that preclude communications. The duration might also be chosen as 16 hours to represent the 2 disjoint ground stations case.

<u>Transmission rate, short-term outages, and data unit size.</u>. The burst transmission rate of the optical channel will depend on the mission scenario, including distance to the spacecraft, day/night visibility constraints, sizes of spacecraft and ground station apertures, etc. The optical data transmission rate is given by r_T in bits per second. For example, the Mars trunkline rate supporting human exploration is envisioned to be r_T=250 Mb/s. The parameter r_T is the "burst" transmission rate that occurs during a pass (when the spacecraft is in geometric view of the spacecraft) and ignoring outages. Short-term outages may arise from multiple causes, as identified in [4], such as atmospheric dynamics or spacecraft platform instability that affects pointing. Analyses have shown that millisecond-scale outages have little effect on buffer performance, and may be mitigated by interleaving and FEC techniques. However, outages may persist for several seconds, and are included in our simulation model. Specifically, we model each transmission unit as requiring a 10 second transmission time, and refer to this unit as a "packet" of data. An actual implementation will use a data link layer protocol with frames typically having much shorter transmission times, and higher-layer offered service data unit sizes may also differ significantly, however the "packet" unit considered here captures the key effective characteristics of the physical optical channel for our purposes. Short-term outages are modeled as an independent identically distributed (i.i.d.) sequence in which a 10-second outage occurs with a given probability p_short, that is, the single-packet transmission short-term failure probability is p_short. During the course of a mission (e.g., a Mars spacecraft), the transmission rate r_T will vary slowly. Thus the size of a "packet" will vary, and appropriately accounted for in performance descriptions by conversion to units of bits. Note that the burst packet transmission rate is fixed at 10 packets/second, with each packet consisting of R_T*10 bits. For example, if R_T=250 Mb/s, the packet size is 2.5 Gb.

<u>Weather model</u>. In addition to scheduled link up/down times according to geometric line-of-sight as well as short-term outages, the optical channel is subject to long-term outages arising from weather events. A 2-state Gilbert-Elliot Markov Chain model is applied in which the channel moves between Good (clear sky) and Bad (cloudy) states at random discrete times according to defined transition probabilities p_GB and p_BG. These transition probabilities determine the mean time one remains in a state ("sojourn time") as E(G)=1/p_GB and E(B)=1/p_BG, and we will use E(G) and E(B) as input variables since they are easier to relate to empirical effects. These also determine the equilibrium proportion of time spent in each state: pi_G=E(G)/[E(G)+E(B)] and pi_B=E(B)/[E(G)+E(B)]. Short-term packet transmission failure probabilities r_G and r_B are in effect according to the current (G or B) weather state, and the overall mean transmission failure probability is r_ave=pi_G*r_G+pi_B*r_B. For example, if the weather time unit is one hour, then p_GB=1/24 implies G persists for a mean of 1 day, p_BG=1/24 implies B persists for an average of 1 day, and pi_G=1/2 and pi_B=1/2. If further r_G=.2 and r_B=1 (i.e., all frames are lost during Bad weather), then the average transmission failure rate is p_ave=0.6, that is, the weather allows transmissions at an overall 40% success rate while a link pass is active.

Figure 1 below depicts several different weather models based on parameter selection. These vary both in terms of the ratio of time spent in the Good versus Bad state (either 1 or 2 in these examples) and with regard to duration in a given state (the bottom two examples being 4 times the top two).

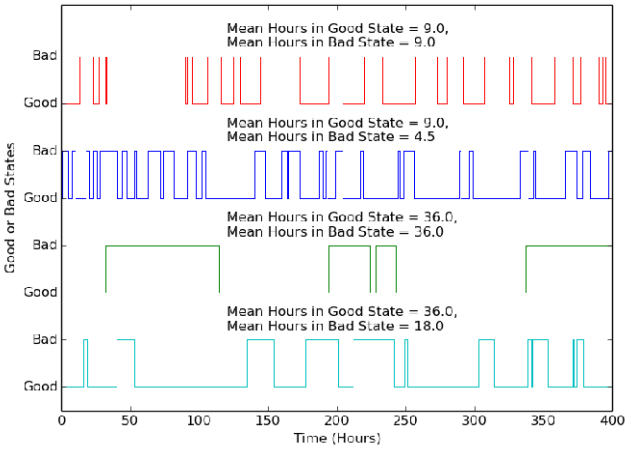

Figure 1. Example parametric weather model realizations.

In fitting empirical weather data to the model, only the weather state during the time interval when the link is active (nominally 8hr/day) is relevant. Thus, diurnal (day/night) variations of clear/cloudy conditions should not influence the parameter selection.

Offered traffic. While offered traffic may be allowed to vary as a stochastic process, the results presented below are restricted to cases of a constant-rate arrival of packets (inter-arrival times are fixed) without gaps. Greater entropy can be expected if packet arrivals are stochastic, leading to larger resource needs. The rate of arrivals is denoted r_A packets per second.

Buffer capacity and data loss from buffer overflow. Packets will be stored as arrivals occur while the link is inactive, and while they await their initial transmission behind earlier arrivals. These packets are said to be waiting in the Active Queue (AQ). Packets will also be stored as they await an ACK following transmission, possibly after multiple retransmissions. Packets await ACKs in the Pending Queue (PQ). Buffers may be allocated separately to the AQ and PQ, but we generally assume that a common packet buffer is shared among the AQ+PQ packets. We denote the sum of AQ and PQ as the Total Queue (TQ) size. The buffer holding the packets has a finite capacity denoted B, measured in packets; the capacity in bits is then B*r_T*10 since by definition each packet is the number of bits transmitted in ten seconds. It is noted that there will be packet overhead in addition to useful user data within the packet payload; we ignore that here as being insignificant for large-sized packets.

Packets that arrive when the packet buffer has reached maximum capacity are lost. We denote the total number of lost packets as L, and the mean number of packets lost per second is denoted r_L. This rate is independent of the burst transmission rate r_T, that is, r_L is the ratio of bits lost per bits transmitted.

It is noted that, because of the ARQ system, the only packet losses that occur are due to buffer overflow. An idealized system with an infinite buffer capacity could have an ever-increasing total queue size (TQ), but we generally focus on systems having steady state (equilibrium) performance. This will always be the case for a system with a finite-capacity buffer, and will also occur whenever the arrival rate is less than the channel capacity.

Throughput, capacity and utilization. The capacity of the system C is the theoretical maximum mean rate of packets per second that can be transferred over the optical link successfully in equilibrium, taking channel effects into account. There are three effects that prevent successful transmission: (1) unavailability whenever the link is not scheduled for an active pass, (2) packet transmission failure due to short-term outage, or (3) packet transmission failure due to weather outage. Assuming passes are scheduled for 8 hours per day, the channel capacity is given by C = (10)*(8/24)*(1-p_short)*[pi_G*(1-r_G)+pi_B*(1-r_B)] packets per second. The channel capacity in b/s is given by C*r_T/10 (recall a packet is defined as 10 seconds of transmitted data). Equilibrium will be reached with an infinite-capacity buffer if and only if r_A<C. The throughput is the mean rate of successfully received packets per second. The equilibrium throughput = r_A*(1-r_L). The utilization is the ratio of the throughput to the capacity.

We define the channel "pass availability" as the proportion of time when successful transmissions may occur during a pass, that is, we condition on transmissions during a pass. The channel capacity is the channel pass availability times the proportion of time scheduled passes are planned times the burst transmission rate r_T. Figure 2 presents the channel pass availability as a function of the ratio of the mean Good weather duration over the mean weather Bad duration, and assuming 8-hour passes, 0 and 1 short-term outage probabilities during the Good and Bad states respectively. The buffer sizing results presented below scale across the full range of burst transmission data rates r_T. For example, if the weather ratio E(G)/E(B)=1.0 then the channel pass availability is 50%, and so if r_T=10Mb/s and scheduled passes are 8hr per day,,then the channel capacity C=1.67Mb/s, and the maximum daily volume is 24hr*C=144Gb.

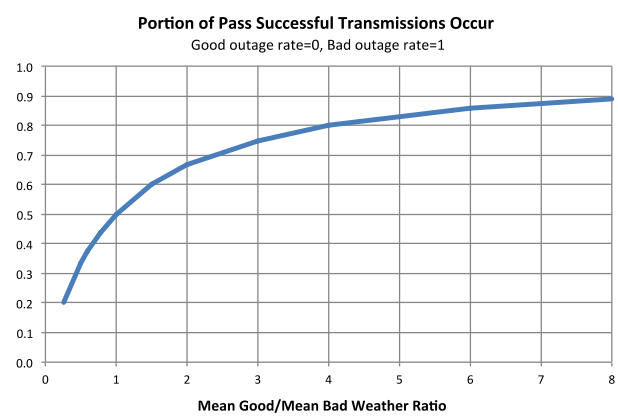

Figure 2. Channel Availability vs. Weather Ratio.

Latency. A packet remains within the communications system from its arrival time at the spacecraft until an ACK is received at the spacecraft (allowing the packet to be cleared from the PQ); this is called the packet latency. The ground will receive the packet a one-way light time earlier when it transmitted the ACK. Statistics are derived below for the distribution of latencies of the successful packets. Packets that are lost do not affect the latency statistics. All lost packets are dropped immediately upon arrival (retransmissions are given priority). Little's Law applies: E(TQ) = r_A*(1-r_L)*E(S) where E(TQ) denotes the mean equilibrium number of packets in the buffer and E(S) denotes the mean latency ("system time"). The sequence of packets successfully received is not guaranteed to be in the same order they were offered to the optical link system. If the user receiving the return data packets desires in-order delivery, a resequencing buffer system may process the stream of packets released by ARQ system output. The latency distribution of the resequenced stream will differ due to occasional added resequencing delays.

Busy Period. A standard unit considered in conventional queueing theory is the Busy Period, defined as the time interval between the time the queue-size becomes positive until it becomes zero. A Busy Cycle is the Busy Period plus the Idle Period, where an Idle Period persists until the queue-size becomes positive. The system stochastically regenerates each Busy Period, and therefore the Busy Period provides a measure of the "memory" that persists in the system. Our optical link ARQ system is not quite a conventional queueing system; we never expect the queue-size TQ to become zero again due to the constant arrivals and packets in the PQ awaiting ACKs (recall the PQ is contained within the TQ). We therefore construct a modified Busy Period corresponding to reaching the effective expected minimum queue-size. This is set as the number of arrivals occurring during a RTT (that is, r_A*RTT), plus the expected fraction of these that will require retransmission due to short-term or weather outages during Good weather. Thus a Busy Period begins/ends whenever the TQ reaches size r_A*RTT/[(1-r_short)*(1-r_G)].

3. PERFORMANCE CHARACTERIZATION

Performance is derived using a discrete-event Monte Carlo simulation based on the ns-3 simulation environment. A discrete-time Markov chain (DTMC) analysis as well as a fluid flow analysis method were also developed and used to validate the simulation model; these additional methods will not be further discussed. The simulation model captures the scheduled pass behavior, 10-second-scale short-term outages, long-term weather outages, ARQ operation, and buffer behavior including overflow losses, and provides full statistical characterization of the various metrics defined above. The smallest level of time resolution used in the simulation model is 10 seconds. We generally focus on equilibrium performance.

Many of the performance metrics derived are related to rare events, such as buffer loss. Rather than generating statistics over a large population of short-duration stochastic realizations, we take advantage of the ergodic property of the underlying stochastic system process, and derive statistics over very long simulated time periods. That is, by ergodicity, "the time average is the ensemble average." In the case where the deep space mission displays slow time variation of the system parameters (such as the Mars orbiter example that varies over a synodic period), time-dependent performance may be derived for each time point using static inputs and then this collection of results is "stitched together" over the long mission time range of interest.

The set of inputs are:

- One-way light time
- Link schedule; pass duration per day
- Burst transmission data rate r_T
- Buffer capacity (if choose infinite, require r_A<C)
- Data packet arrival rate r_A
- Short-term outage probability while in Good weather state r_G and short-term outage probability while in Bad weather state r_B
- Weather outage model parameters {E(G), E(B)} where E(G)=mean duration of continuous Good weather and E(B)=mean duration of continuous Bad weather
- Duration of simulated time (complete run)

The set of outputs are:

- Throughput, Channel Utilization
- Loss rate due to buffer overflow
- Latency statistics
- Queue-size statistics
- Busy Period statistics

4. PERFORMANCE RESULTS

Unless otherwise noted, the following values were assumed for the results presented below:

- One-way light time = 1342 seconds, corresponding to Earth and Mars at maximum distance of 2.7AU. Analyses reveal that performance at high loads is not sensitive to this parameter.

- The link is active 8 hours continuously per day. In actual operations, the duration will change slowly from day to day; for example, for a Mars 2022 mission this will vary between 5.17 hr to 11.00 hr; also, the pass start time will slowly slide.

- Short-term outage probability in Good state r_G = 0 (all transmissions succeed during Good weather) and in Bad state r_B=1 (no transmissions succeed during Bad weather).

- Packet arrival rate is taken to be 80% of the channel capacity: r_A=0.8*C where the capacity of the link C depends on the weather state durations ratio w=E(G)/E(B) and is given by C =

(10)*(8/24)*({w/[1+w]}*[1-r_G]) packets per second. Higher offered load values may be chosen, but will generate larger queue-sizes and require larger buffer capacities to meet acceptable buffer overflow performance.

Performance figures show queue-size occupancies and buffer capacities in bits, corresponding to a "baseline" case in which the burst transmission rate r_T = 10 Mb/s is chosen for concreteness. The results are easily scaled to account for other data rates. Since each packet is modeled to have a 10-second transmission time, this r_T corresponds to a packet size of 10Mb/s*10s = 100 Mb. For example, a buffer capacity of 1000 packets corresponds to 100 Gb in this case.

Queue-Size Process

Figure 3 below depicts two realizations of queue-size stochastic processes. The size (or "occupancy") of the queue grows higher during Bad weather, and decreases while Good weather predominates. The brown dashed line represents the infinite-capacity buffer case, while the green solid line case limits the buffer capacity, which stochastically causes lost data. These examples are generated from the same pseudo-random generator seed initiating the simulation runs, thus the overlap except for buffer-overflow-related events. Weather effects have high impact. For the examples of Figure 3, the mean duration of a Good state is 24 hours and the mean duration of a Bad state is also 24 hours, yielding a Good to Bad ratio of 1.

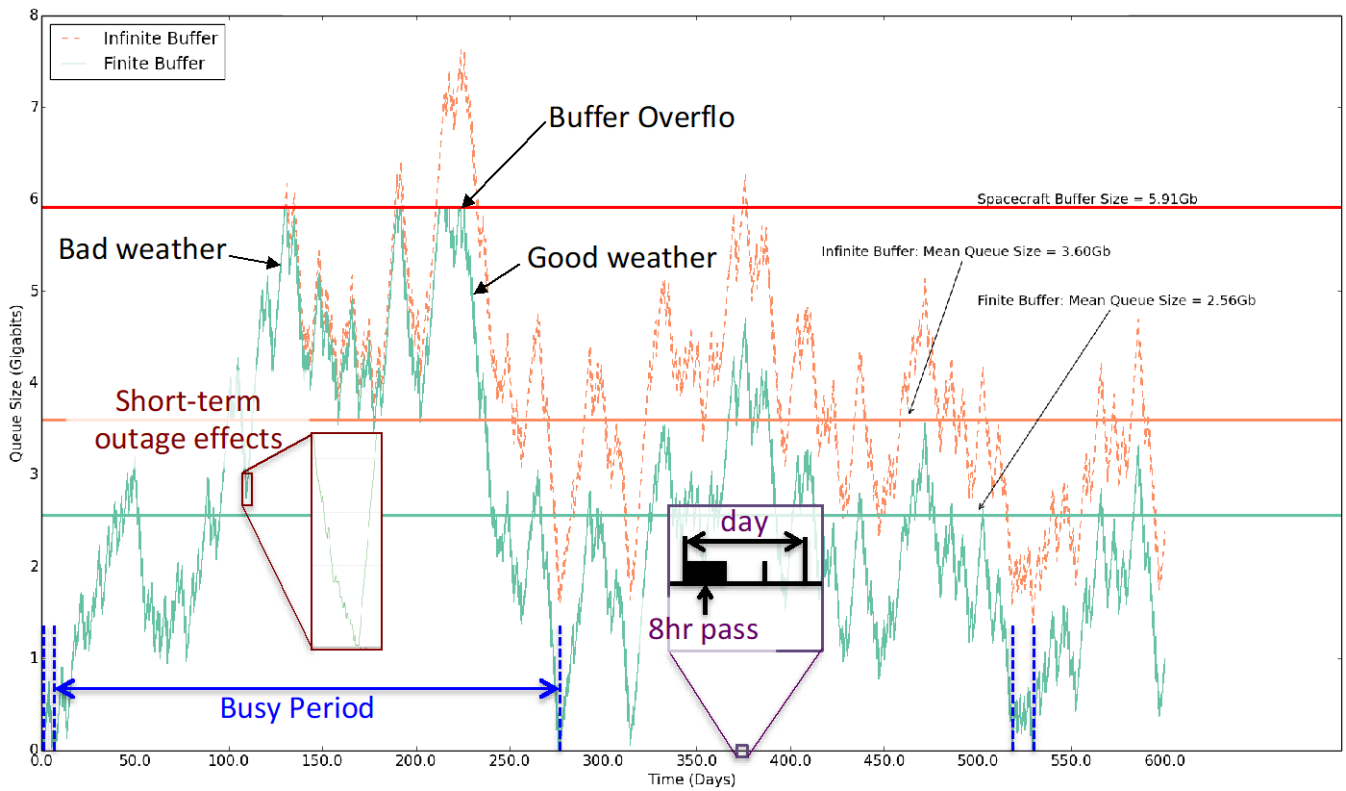

Figure 3. Queue-Size vs. time, Infinite and Finite Buffer Capacity Examples.

Queue-Size Performance

Figures 4 and 5 below depict queue-size performance over a range of the weather model parameters. The buffer capacity is unbounded in all cases. In both figures, the weather parameters are varied as (1) ratio of mean Good over Bad durations E(G)/E(B) and (2) the sum of the mean Good and Bad durations E(G)+E(B) (i.e., the mean "weather cycle" length). The ratio E(G)/E(B) impacts channel capacity C, and the offered load is taken as 80% of C. Smaller values of E(G)/E(B) generate larger buffer demands. Buffer needs also grow with the length of the weather cycle (E(G)+E(B)).

Figure 4 presents the mean queue size versus the mean weather cycle duration. From Little's Law, one may immediately generate the mean latency vs. weather cycle duration, using (Mean queue size)=(Mean arrival rate)*(Mean latency) where "arrivals" are not lost, always the case for an infinite capacity buffer. Included are the cases E(G)=30.6, E(B)=40.9 and E(G)=E(B)=9.5 which correspond to a fit of empirical data collected for the Table Mountain Facility (TMF) and for Edwards Air Force Base (EAFB) [5, 6, 7]. However, these datasets correspond to measurements taken 24/7 continuously, whereas the actual spacecraft buffering performance will depend on weather durations perceived only during pass time-windows. Further analysis of this issue is recommended. Nevertheless, we expect that these empirical data provide a rough indication of the range of values we may expect.

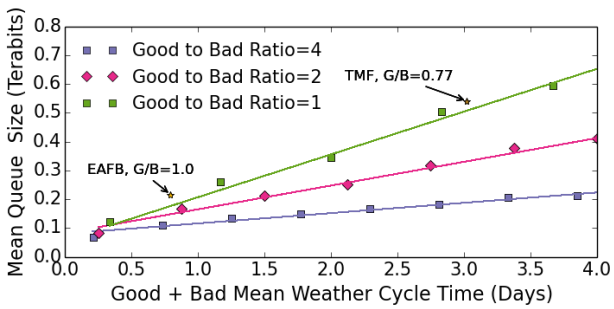

Figure 4. Mean Queue Size vs. Weather Cycle Time, 80% Offered Load, r_T=10Mb/s.

Figure 5 presents the 95th percentile of the queue size probability distribution versus the mean weather cycle duration. These measurements may suggest what capacity should be chosen for the implementation of the spacecraft buffer. Loosely, one might imagine that 5% of the data would be lost if one sets the buffer capacity to this 95th-percentile value. We found however that use of the infinite capacity 95th percentile queue-size actually provides loss rates closer to 1%. This measure can be used as an initial estimate to size the spacecraft buffer capacity.

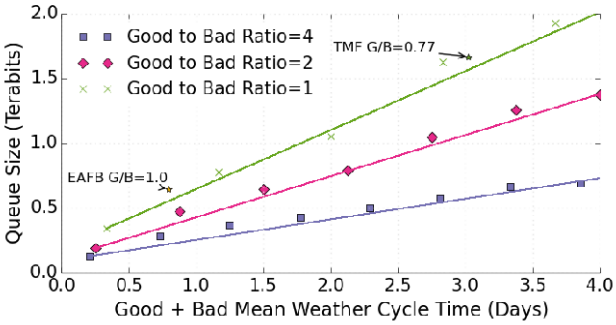

Figure 5. Queue Size 95th-Percentile vs. Weather Cycle, 80% Offered Load, r_T=10Mb/s.

We next consider the performance impact of using different finite buffer capacities. Mean queue-size results are combined with mean latency performance in Figure 15. Figure 6 presents the 95th percentile queue size versus buffer capacity for two different weather conditions, in which the mean Good duration and mean Bad duration parameters are {9.5, 9.5} hours or {36, 36} hours respectively. As the buffer capacity increases, the corresponding values converge to those shown in Figure 5 for these weather scenarios.

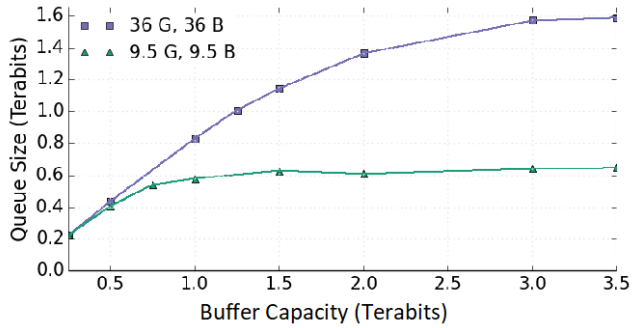

Figure 6. 95th Percentile Queue-Size versus Buffer Capacity, 80% Offered Load, r_T=10Mb/s.

Performances presented in Figures 4 through 6 all assume the offered load is 80% of channel capacity. We next consider system performance as the offered traffic load is varied. Mean queue-size results are combined with mean latency performance in Figure 17. The 95th percentile queue-size versus offered load is depicted in Figure 7 for the cases where the mean durations in the Good or Bad state are {9.5, 9.5} or {36, 36} hours.

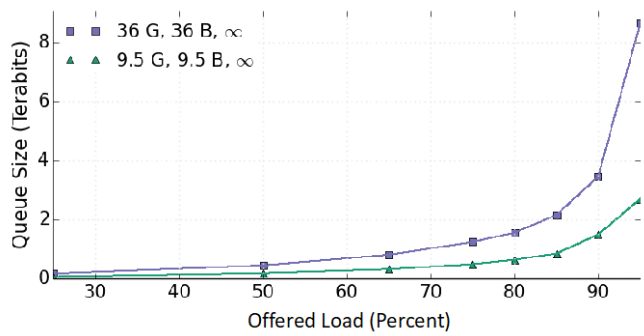

Figure 7. 95th Percentile Queue-Size versus Offered Load, r_T=10Mb/s.

Buffer Overflow (Data Loss) Performance

Because an ARQ technique is employed, the only losses that occur are due to new arrivals that find the buffer is full. In this case, we assume that the new arrival is dropped, as opposed to its replacement of an existing packet already in the queue (this impacts latency performance). In the following, we present the mean loss rate, which may also be viewed as the probability that a randomly selected packet is dropped due to buffer overflow. It is noted that losses will generally occur in bursts; this behavior can be characterized using higher-order statistical analysis.

Figure 8 provides an example realization of the cumulative packet loss process versus time. The mean Good weather duration is 8 hours and mean Bad weather duration is 4 hours. The buffer capacity is 10,000 packets, or 1Tb for the 10Mb/s baseline case. The offered load is 95%, and r_G=0.2. Note that time scale is 60,000 days, a little over

164 years. Of course, we are not concerned with a mission lasting such an enormous duration, rather, as noted earlier the ergodicity of the system ensures long time averages equate to large ensemble averages. Note that the standard deviation is more than 10 times the mean for the daily loss rate, that is, its coefficient of variation is greater than 10. The mean loss rate measured as the ratio of packets lost out of all packets arriving to the system is 0.0031. Notice that losses may occur in bursts.

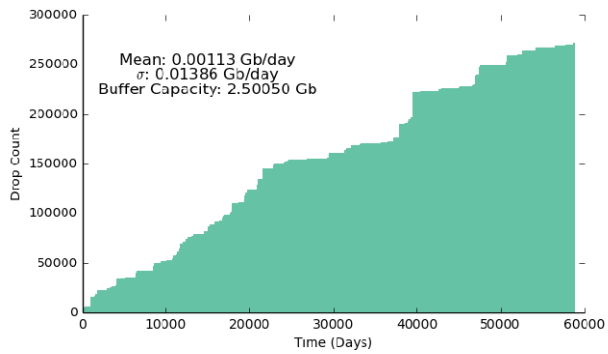

Figure 8. Example Cumulative Lost Packets vs. Time.

Figure 9 presents packet loss rate as a function of buffer capacity for four different weather scenarios, defined by (mean Good weather duration, mean Bad weather duration) as (9.5hr, 9.5hr) or (36hr, 36hr). In all cases the offered load was taken to be 80% of the channel capacity, and r_G=0. Note that for our 10Mb/s baseline scenario, a loss rate of 1% is achievable with 800Gb or 1.5Tb buffer capacities for the {9.5, 9.5} or {36, 36} weather cases respectively.

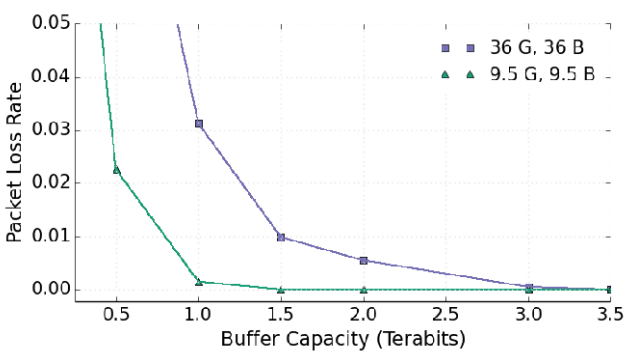

Figure 9. Packet Loss Rate vs. Buffer Capacity, 80% Offered Load, r_T=10Mb/s.

We also consider the loss rate performance across a range of offered load values. The same or lower loss rate can be achieved with a smaller buffer capacity at the cost of operating at a reduced offered load. Figure 10 presents the packet loss rate as a function of the offered load as it ranges between 25% to 95% of channel capacity when the weather is defined by mean Good and Bad durations of 12 hours each.

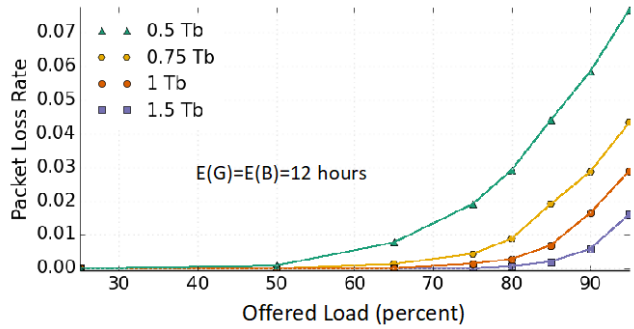

Figure 10. Packet Loss Rate vs. Offered Load, Good & Bad Weather Mean Duration = 12 hours, r_T=10Mb/s.

Latency Performance

Although we are nominally assuming deterministic periodic arrivals as well as pre-scheduled pass times (matching predictable geometric visibility), latency performance is stochastic due to the outages that arise. We derived the delivery latency performance for a randomly selected packet that is not lost due to overflow at its arrival instant. To convey the stochastic nature of latency, and to provide insight into the key variables of buffer capacity and weather effects, we present specific examples of empirical probability density functions in the next two figures. Figure 11 provides examples under two difference weather conditions (E(G)= E(B)=9.5 and E(G)=E(B)=36). One sees that is there is generally a substantial "spike" probability for near-zero latency values. Small deviations occur near the 1-day latency due to the manner in which packets are handled at pass boundaries. Larger latencies arise for larger mean weather cycle (E(G)+E(B)) values. The 95th-percentile values are shown for each case. The buffer capacity for both cases is infinite, and the offered load is 80%.

.

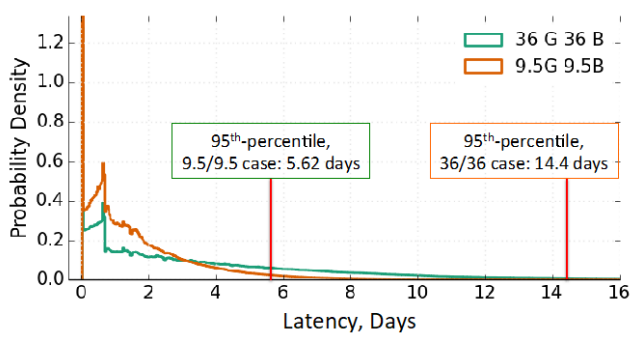

Figure 11. Latency Probability Densities, Short and Long Mean Weather Cycles, Infinite Buffer Capacities, 80% Load

Figure 12 presents the effect of buffer capacity on the latency distributions. Two cases are considered, both generated with the same parameters except for buffer capacity, infinite or 0.5Tb (where we are assuming the baseline burst transmission rate r_T=10Mb/s). The mean weather durations for Good or Bad states are E(G) = E(B) =

24 hr, and the offered load is 80%. It is seen that imposing a finite buffer can have a significant impact on latency performance. The 95th-percentile values are shown for each case. This finite capacity case resulted in a loss rate of 0.0697. This is a larger loss rate than one would typically tolerate, but illustrates the effect on latency performance. If a larger offered load were used (say 95%), buffer capacities associated with smaller loss rates would have a larger impact on latency. Thus it is important to consider whether a modest loss rate would be acceptable to not only reduce buffer costs but also to improve latency performance.

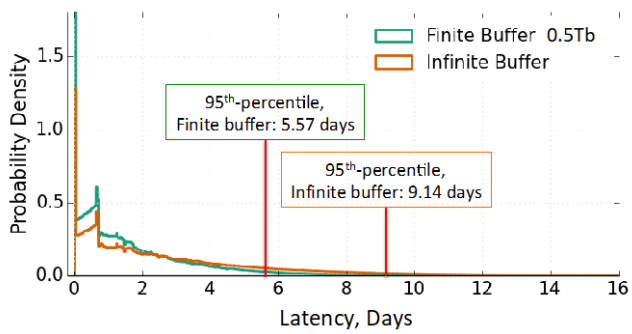

Figure 12. Latency Histograms, E(G)= E(B)=24, Infinite and Finite Buffer Capacities, 80% Load, r_T=10Mb/s.

Figure 13 presents the mean latency versus mean weather cycle for the same cases addressed for the mean queue-size in Figure 4.

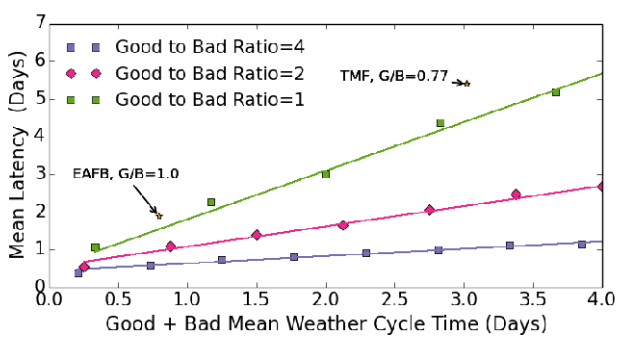

Figure 13. Mean Latencies versus Mean Weather Cycle Duration, 80% Offered Load.

Figure 14 depicts the 95th-percentile latencies versus mean weather cycle duration under the same two different weather conditions, again with 80% offered load. One finds that the 95th-percentile latency is roughly 3 times the mean value across the ranges considered.

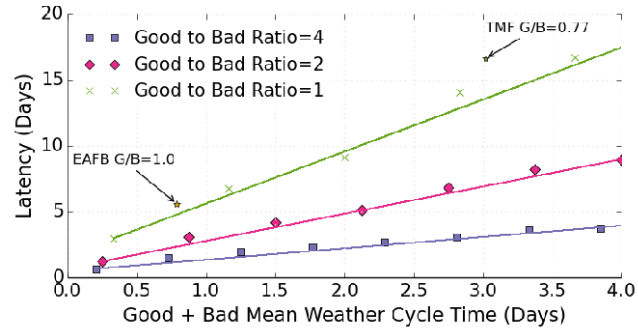

Figure 14. 95th-percentile Latencies versus Mean Weather Cycle Duration, 80% Offered Load.

Figure 15 provides both the mean queue-size and the mean latency versus buffer capacity for two different weather conditions (E(G), E(B))=(9.5,9.5) or (36, 36). The mean queue-size and mean latency follow Little's Result: (Mean queue size)=(Mean arrival rate)*(Mean latency); the divergence between the curves is due to arrivals that are lost from buffer overflow.

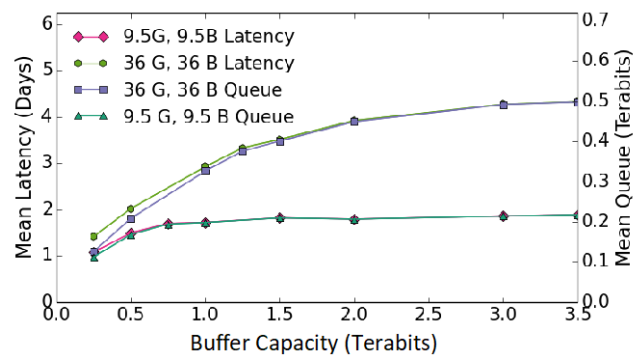

Figure 15. Mean Latency and Mean Queue-Size versus Buffer Capacity, 80% Offered Load, r_T=10Mb/s.

Figure 16 presents the 95th-percentile latencies versus buffer capacity for two different weather conditions (E(G), E(B))=(9.5,9.5) or (36, 36). Again, we found these are roughly three times the mean latency values.

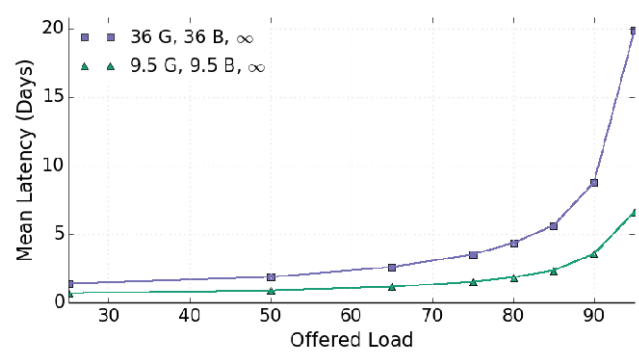

Figure 16: Mean Latency versus Offered Load, E(G)=E(B)=9.5 or 36hr, Infinite Buffer Capacities.

Figure 17 presents provides both the mean queue-size and the mean latency versus offered load capacity for two different weather conditions (E(G), E(B))=(9.5,9.5) or (36, 36) and an infinite buffer capacity.

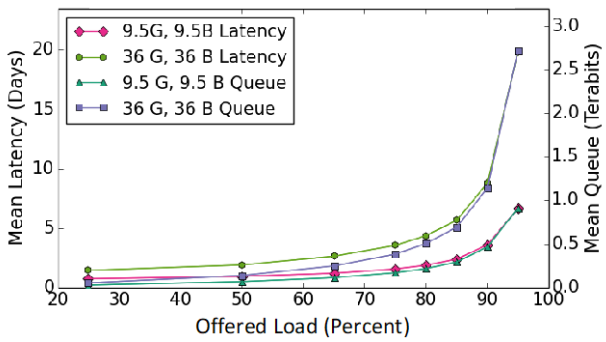

Figure 17: Mean Latency and Mean Queue-Size versus Offered Load, r_T=10Mb/s.

Figure 18 presents the 95th-percentile latency performance versus offered load for the same two weather conditions of (E(G), E(B))=(9.5,9.5) or (36, 36).

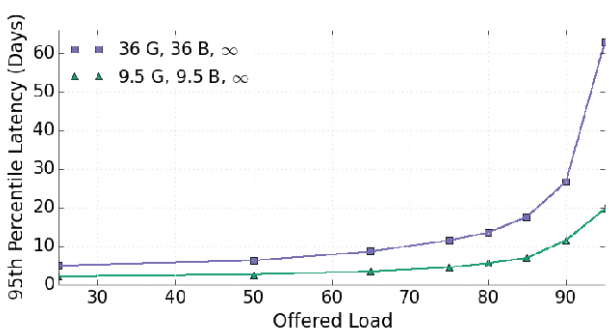

Figure 18: 95th-percentile Latency versus Offered Load, E(G)=E(B)=9.5 or 36hr, Infinite Buffer Capacities.

Busy Period Performance

The Busy Period was defined in Section 2. Figure 3 provides an example queue-size process realization showing Busy Periods that occur for the finite buffer case. As noted before, Busy Period instants correspond to times when the system "resets" and there is no dependency on prior history of the process. For heavy offered loads, the Busy Periods may persist for a long time. In the example in Figure 3 of the infinite capacity buffer case, the Busy Period persists longer than the 600-day run shown.

Figure 19 provides the empirical survivor functions (or complementary cumulative probability distribution functions) for the Busy Period processes for two different weather cases and where buffer capacity is either infinite or set to the 95th-percentile queue-size for the infinite capacity case and a 95% offered load. Very long simulation runs were made to obtain these statistics. These illustrate that the Busy Period process has a "long tail", in which the vast majority of the probability mass occurs for small values but very large Busy Periods are possible. These examples demonstrate that "memory" may persist in such systems for a substantial length of time, with some Busy Periods persisting for many hundreds of days.

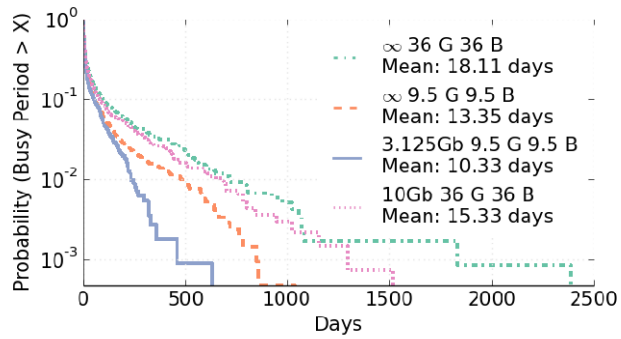

Figure 19. Example Busy Period Survivor Functions, 95% Offered Load.

5. SUMMARY

We have investigated the application of automated retransmission (ARQ) to a deep space optical link with a single Earth ground station. A discrete-event simulation tool has been developed. Outages are modeled for random short-term effects (~10sec outages), random long-term weather effects (persisting for many hours), and scheduled passes accounting for geometric visibility and operational constraints. These parameters generate the link "availability" or mean proportion of time when successful transmission can occur. The burst transmission rate combined with the availability yields the channel capacity. The traffic arrival rate divided by the channel capacity is the offered load. A crucial system parameter is the buffer capacity. Performance is also impacted by the link's propagation delay.

Performance was evaluated for several metrics, including data loss due to buffer overflow and data delivery latency. Weather has a strong impact, and defined by 2-state (Good or Bad) Markov chain parameterized by the mean continuous duration in each state, E(G) and E(B). Performance is found to depend on the "weather ratio" E(G)/E(B) and the "mean weather cycle" E(G)+E(B).

Published weather statistics were drawn on to best fit the Markov weather model. While limited (e.g., data used was from only two locations less than 50 miles apart), this provides a preliminary range to focus on. Presented results largely assume a system utilization with a 80% offered load, however results for different loads were also presented.

All performance results shown are for the case of a single optical ground station. Results presented generally assume an 8hr/day scheduled link pattern for simplicity, although passes in actual operations will vary according to mission- and time-dependent view geometry.

The analytical approach presented offers a mechanism to estimate the buffer capacity required to meet data loss (from buffer overflow) requirements. Moreover, latency performance must be recognized for deep space mission operations, having larger values largely due to the single ground station limitation as well as poor weather that may persist.

Adding more optical ground stations would very significantly improve the user performance, arising from both longer temporal coverage and weather diversity. In addition, RF systems are likely to be employed in parallel, particularly by early optical communications adopters. Novel methods for integrating optical and RF for improved Quality of Service (QoS) include using optical for "first-time" transmissions and RF for retransmission, thereby significantly reducing maximum delivery latencies. Such systems and operational techniques are under current investigation.

ACKNOWLEDGEMENTS

The research described in this paper was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology and the National Aeronautics and Space Administration. The authors acknowledge and thank Jay Breidenthal, Hua Xi, and Jay Gao for their valuable discussions and contributions to this paper.

REFERENCES

[1] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss, "Delay-Tolerant Networking Architecture," RFC 4838, April 2007.

[2] M. Ramadas, S. Burleigh, and S. Farrell, "Licklider Transmission Protocol - Specification," RFC 5326, September 2008.

[3]"Licklider Transmission Protocol (LTP) for CCSDS," Issue 1, Recommended Standard (Blue Book), CCSDS 734.1-B-1, Washington, D.C.: CCSDS, May 2015.

[4] Hamid Hemmati, Ed., *Deep Space Optical Communications.* New York: Wiley, April 2006.

[5] Paul W. Nugent, Joseph A. Shaw, and Sabino Piazzolla, "Infrared Cloud Imager Development for Atmospheric Optical Communication Characterization, and Measurements at the JPL Table Mountain Facility," IPN Progress Report 42-192, February 15, 2013.

[6] P. Elizabeth Amini, Stephen D. Slobin, and Sabino Piazzolla, "Cloud Coverage Statistics for Optical Communications at Table Mountain Observatory, California, Based Upon 1992-1998 Surface Weather Observations From Edwards Air Force Base, CA," JPL Publication 00-12, November 30, 2000.

[7] Sabino Piazzolla, Stephen D. Slobin, and P. Elizabeth Amini, "Cloud Coverage Diversity Statistics for Optical Communications in the Southwestern United States," JPL Publication 00-13, November 30, 2000.

BIOGRAPHY

Loren Clare *is the supervisor for the Communications Networks Group at the Jet Propulsion Laboratory. He obtained the Ph.D. in System Science from the University of California, Los Angeles in 1983. His interests include wireless communications protocols, self-organizing systems, network systems design, modeling and analysis, and distributed control systems. Prior to joining JPL in May 2000, he was a senior research scientist at the Rockwell Science Center, where he acquired experience in distributed sensor networks, satellite networking, and communications protocols for realtime networks supporting industrial automation.*

Gregory Miles *is a Computer Science senior at California State University, Los Angeles. He has been with the Jet Propulsion Laboratory since 2014.*