

Storage Usage of Custody Transfer in Delay Tolerant Networks with Intermittent Connectivity

Matthew Seligman
Laboratory for Telecommunication Sciences
Army Research Lab, ALC, Bldg. 601
2800 Powder Mill Road
Adelphi, MD 20783
mattseligman@gmail.com

Abstract

Delay Tolerant Networks (DTNs) operate in a environment characterized by intermittently connected links and long, variable latency. End-to-end reliability schemes, such as TCP, perform poorly, or not at all, in this environment. Source node storage becomes a constrained resource when intermittent links have long downtimes or latency to receive acknowledgement from the destination is long. To alleviate this constraint, custody transfer is a proposed DTN reliability scheme using hop-by-hop reliability to enhance end-to-end reliability. A simulation study of storage usage is presented in this paper comparing end-to-end reliability and custody transfer with three different intermittent link connectivity schedules. The results presented show significantly lower storage usage when custody transfer is used as opposed to end-to-end reliability in intermittently connected networks for all schedules considered.

Keywords: Delay Tolerant Networks, Custody Transfer

1 Introduction

Delay Tolerant Networking (DTN) is a field of network research focused on architectures and protocols that are able to operate in challenged networking environments. A challenged networking environment is a network with extremely limited resources including CPU processing power, memory sizes, and network capacity. Specifically, a challenged networking environment is characterized by one or more of the following attributes.

- Intermittently connected network
- Long variable latency
- Asymmetric bandwidth

There are several scenarios in which DTN technologies will perform significantly better than current Internet technologies. This is because DTN solutions are designed

to perform well in challenged networking environments, whereas current networking solutions are designed to perform well in the Internet.

1.1 Characteristics of Challenged Networking Environment

Intermittent Connected Network : An intermittently connected network contains links that become available and unavailable during normal operation. This behavior is caused by mobility of nodes, lack of line-of-sight, physical disconnection, node failure, and transmission power among other factors. Link availability may be scheduled, probabilistic, or random based on the cause of disconnection. In addition, these networks may be sparsely connected due to long periods of disconnection and a large number of intermittent links. For example, in DARPA's Disruption Tolerant Networking BAA Industry Day presentation [1], 20% link availability or less is defined as a metric for a sparsely connected network to evaluate possible DTN solutions. End-to-end reliability, flow and congestion control, and routing protocols based on current connectivity will not be suitable since the fundamental assumption in all of these designs is end-to-end connectivity.

Long and Variable Latency : As the DTN name implies, long latencies are characteristic of a challenged networking environment. In the case of InterPlanetary Internet (IPN) research, latencies are in the range of tens of minutes to hours or days. Long latencies are caused by large distances between communicating nodes, low data rates, and congested resources in the network. In an end-to-end reliability scheme, long latencies require large source retransmission buffers when the source transmits a large amount of data. Variable latencies yield large variance for in values of adjacent packet or message latencies. Large variances in latency values cause difficulty in selecting reliability protocol parameters such as timeout values.

Asymmetric Bandwidth : Asymmetric bandwidth occurs on a bidirectional link that has a different bandwidth

in each direction, which commonly occurs in space and satellite communications. All reliability schemes require feedback, so asymmetric bandwidth effects their ability to receive timely feedback.

1.2 Lack of End-to-End Connectivity

The characteristics of a challenged networking environment are fundamentally different than those characteristics of networks that are commonly implemented using the TCP/IP protocol suite. The fundamental differences contribute to one major attribute of a challenged networking environment that must be addressed in any DTN solution, namely the lack of end-to-end connectivity between source and destination.

End-to-end connectivity is defined as the existence of one or more paths between the source and destination nodes through the duration of transmission. Lack of end-to-end connectivity occurs due to several of the characteristics of a challenged networking environment. This is a very challenging problem to consider with current technologies because commonly used Internet protocols, such as TCP, provide end-to-end reliability, flow control, and congestion control under the assumption that end-to-end connectivity exists. In addition, Internet routing protocols assume an end-to-end path exists to determine valid routes. Therefore, lack of end-to-end connectivity requires a different approach for reliability, flow and congestion control, and routing.

1.3 Node Storage

Reliability protocols require node storage to implement a retransmission buffer to store unacknowledged data, which is generally volatile storage. This buffer may contain the only copy of the data, so if data is lost when a node fails, it is lost from the network and cannot be retransmitted in case of transmission failure. To prevent data loss, DTNs use non-volatile, persistent node storage.

In addition, a challenged networking environment further constrains storage in the network because intermittent connectivity and long latency require data to be stored for long periods of time throughout the network causing competition for the persistent node storage. Therefore, storage usage is a key factor in DTN reliability protocol selection. This paper evaluates and compares the storage usage of traditional end-to-end reliability protocols and *custody transfer*, which is the DTN proposed reliability scheme.

2 Related Work

Reliable communications research and protocols in the past have focused on developing solutions for a given set of assumptions, which are stable network topology and end-to-end connectivity. Several protocols have been developed, of which TCP is one of the most commonly

used. Since the given assumptions do not hold true in a challenged networking environment, connection-oriented end-to-end reliability protocols, including TCP, are unsuitable for DTN applications. Variations of TCP have been developed to compensate for its shortcomings for DTN applications in space for the InterPlanetary Internet (IPN). These variations show improved performance, but only under a strict set of networking conditions. Other related fields of research, such as Mobile Ad-Hoc Networks (MANETs) and sensor networks, are related and address a subset of a challenged networking environment's characteristics, but no previous reliability solution is suitable when a lack of end-to-end connectivity exists.

2.1 TCP

From its original design [2] and standardization in RFC 793 [3], TCP was designed for end-to-end reliable data communications over stable networks with high bandwidth and low latency. Originally designed for DARPA for use in ARPANET and other DoD networks, it was designed for managing reliable data communications from a sender and receiver, retransmission of missing packets, addressing, and reassembly. TCP is a connection-oriented transport protocol and requires a connection to be established prior to data transfer between the source and destination. It is not possible to create a connection when no end-to-end path exists. Therefore, TCP, or any other end-to-end connection-oriented protocol, will not work for DTNs since the connection is necessary for reliable TCP data transfer.

For reliable communications, TCP uses an Automatic Repeat Request (ARQ) algorithm to notify the source of successful or unsuccessful delivery of data segments. ARQ is an end-to-end algorithm that allows the receiver to request the retransmission of specific data segments not received from the sender. End-to-end handshaking used for reliability, such as ARQ, will not operate without an end-to-end path. In addition, TCP uses source retransmission buffers to store data segments until an acknowledgment is received by the sender. The lack of end-to-end connectivity requires extremely large source retransmission buffers. Without an end-to-end path, TCP periodically attempt to retransmit the data segment until a maximum number of retransmission is met, and the source is required to store the data segment until the maximum number of retransmissions is achieved. The shortcomings of TCP for DTN applications are also well-documented in [4].

TCP variants were developed specifically for IPN applications. IPN applications operate in a challenged networking environment, similar to DTN applications. First, TCP-Peach+ was developed for satellite IP networks [5] and attempts to address throughput performance [6], but it does not consider reliability. TP-Planet was developed as a replacement for TCP in IPN applications [7]. It does describe a delayed SACK solution for bandwidth asymme-

try. Also, an improved congestion control solution is provided when a link is unavailable, called a Blackout State. During this state, the transmission rate is reduced to avoid a large number of retransmissions for lost packets. Again, the lack of end-to-end connectivity was not considered in this solution. All of these solutions are application-specific to IPN, and they do not generalize for DTN applications.

In wireless networking, TCP was considered as a solution due to its widespread use in the Internet, but TCP is not designed to operate efficiently in a wireless network. A wireless network contains many of the same challenges encountered in a DTN. Many papers, such as [8], discuss several issues pertaining to the use of TCP over wireless links. This paper highlights the key assumptions made by TCP designers and how wireless links contradict those assumptions such as low bit error rates and errors caused by congestion.

2.2 MANET

MANET is a field of research focusing on mobile wireless networking without an infrastructure of fixed access points. DTNs and MANETs are similar in their networking environment and assumptions, in fact, a DTN solution may be used as an overlay over a MANET network. Mobility in a MANET creates a lack of connectivity in the network and variable link parameters, such as latency and bandwidth, as distance between nodes varies. Similarly, DTNs consider lack of connectivity due to mobility among other scenarios. The ad-hoc nature of MANETs yield dynamic network topologies due to nodes entering and leaving the physical range of the network, just as DTNs consider dynamic networks. The one major difference between MANETs and DTNs is how end-to-end reliable communications work. MANETs assume the destination is reachable when the source sends data, whereas DTNs make no assumption about the connectivity to the destination in any part of the network. This key difference illustrates how MANET technologies cannot offer end-to-end reliability for DTNs.

2.3 Sensor Networks

Sensor networks are used to retrieve information from a large number of physically dispersed computing devices called sensors. Each sensor gathers information to send upstream for further processing because sensor devices are limited by electrical power budget, computing resources, and memory resources. Specifically, limited radio range, mobility, and power budget constraints cause intermittent connectivity. Therefore, sensor networks also suffer from lack of end-to-end connectivity.

Sensor network transport protocols, such as RMST [9] and PSFQ [10], were designed for reliable data delivery. RMST offers reliability using a selective-NACK commit protocol. PSFQ considers out-of-order fragments by re-

questing retransmission of these fragments and reordering them prior to forwarding. An evaluation of the performance of these protocols along with a TCP-like protocols using only end-to-end acknowledgments in sensor networks characterized by high round-trip delay, disconnections, unreliable nodes, large messages, and high mobility is discussed in [11]. The evaluation determined that none of the protocols perform well under all of the network conditions described above including lack of end-to-end connectivity. Based on the discussion presented, custody transfer is necessary to provide reliability in networks without end-to-end connectivity.

3 Custody Transfer

First described in [12], custody transfer is a mechanism to improve reliability by using hop-by-hop reliability, one or more hops, in the absence of an end-to-end connectivity by transferring the responsibility of reliable delivery to intermediate nodes along a path from source to destination. In addition, custody transfer improves storage usage for the retransmission buffer at the source node.

A DTN node maintaining custody of a message is called a *custodian*. A custodian is responsible for reliable delivery of the message from itself to another custodian or the destination. Acknowledgments and custody transfer complete messages are transmitted between two communicating DTN nodes using a distributed commit protocol described in [13]. The custodian of a bundle changes one or more times on its path from source to destination. The maximum number of custody transfers possible is equal to one less than the number of nodes traversed by the message.

Each message is persistently stored in a retransmission buffer at its custodian node. Persistent storage at some or all of nodes in a DTN is a feature of the architecture [14]. It is necessary to overcome node failure and intermittent connectivity. The custodian of a message may be the only node in the DTN containing a copy of the message since each previous custodian, including the source node, has removed the message from its retransmission buffer after relinquishing custody. All messages at a custodian node without persistent storage would be lost if node failure occurred.

Operationally, a message is sent from its current custodian to another node closer to the destination or to the destination. A node may or may not accept custody of a received message. If custody is accepted, a message is sent to the previous custodian containing the acknowledgment of receipt and custody acceptance by the receiving node, and the previous custodian removes the message from its storage.

Joint custody occurs when the acknowledgment and custody acceptance message is not received by the previous custodian. At this point, the previous custodian has not received notice of the custody transfer, and the current

custodian has already asserted custody. Therefore, both nodes believe each is the custodian of the message and attempts to forward the message as its custodian. Joint custody leads to message duplication. DTN destination nodes, or intermediate nodes, must be able to deal with duplication by recording the receipt of unique messages to identify duplicate messages and eliminate them from the DTN, which requires future work.

An example of end-to-end delivery from source Node 1 to destination Node 8 using custody transfer is shown in figure 1. All nodes along the path from source to destination are filled in gray, and the node storing the message in its retransmission buffer is filled in black. Node 1 sends a message to Node 5 along the path drawn in arrowed lines because no path currently exists to the destination, Node 8. Node 1 stores the message in its retransmission buffer in figure 1(a). Node 5 receives the message, stores the message in its retransmission buffer, and sends a positive acknowledgment (ACK) and custody acceptance back to Node 1. Next, Node 1 removes the message from its retransmission buffer in figure 1(b). The responsibility of reliable delivery has moved from Node 1 to Node 5 through custody transfer. No path exists from Node 5 to Node 8, so Node 5 persistently stores the message until a path is available or it is deleted by Node 5 explicitly according to retransmission buffer management. This is a subject for future research.

Figure 1(c) shows two network topology changes. The link from Node 3 to Node 5 is not available, and the link from Node 5 to Node 6 has become available. Node 5 initiates a message delivery transaction with Node 8. The same mechanism is followed as described between Node 1 and Node 5. The only difference in the message delivery from Node 5 to Node 8 is that Node 8 is the destination, and no further message forwarding is necessary. After Node 8 receives custody and Node 5 removes the message from its retransmission buffer, the end-to-end message delivery is complete as shown in figure 1(d). This simplified example of custody transfer shows successful delivery of a message without end-to-end connectivity. The simplification of this example has omitted several custody transfer issues, but it does illustrate the ability of custody transfer to reliably transmit a message through an intermittently connected network.

4 Reliability and Storage Issues with End-to-End Schemes

The use of end-to-end reliability schemes is common in today's networks due to the end-to-end principle. End-to-end reliability require the source to store data packets until the source receives an acknowledgment (ACK) of successful delivery. The lack of end-to-end connectivity and long latency requires the source node to store the data packets for a potentially long period of time, so source node storage becomes a bottleneck for reliability. Relia-

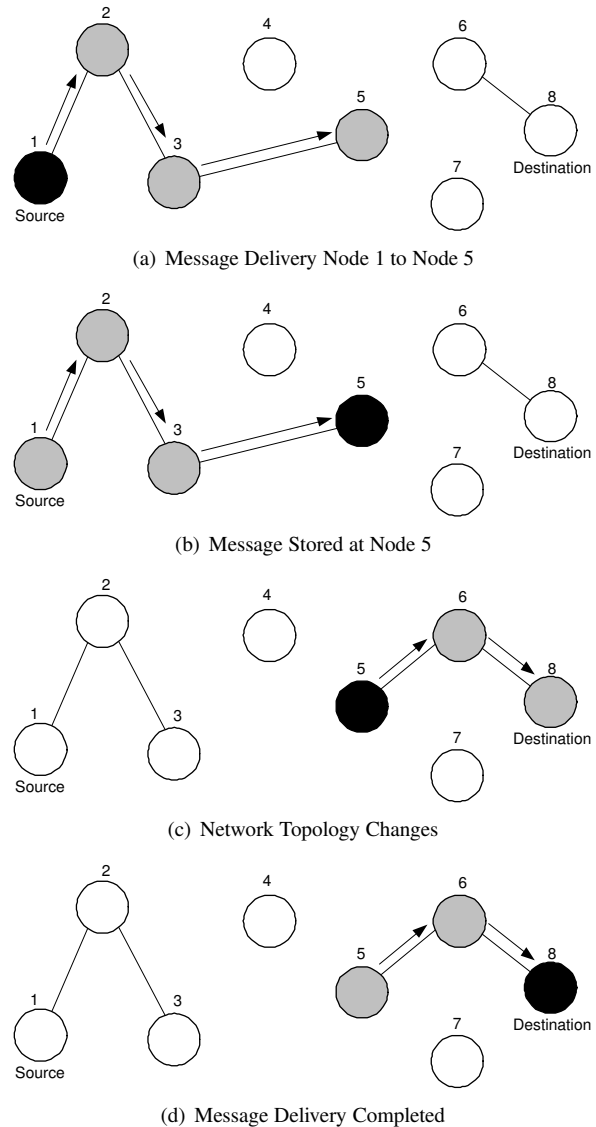


Figure 1: Custody Transfer Example

bility is not guaranteed when the source node storage is completely full. End-to-end reliability schemes in DTNs may require extremely large source node storage, but as previously discussed, DTN nodes lack abundant networking resources, such as storage. Therefore, a different reliability approach is necessary for reliable communications.

One reliability improvement to an intermittently connected network is persistent node storage. This allows data packets to be stored in the middle of the network while waiting for intermittent links to become available towards the packet's destination. This solution reduces re-transmissions because data packets are not dropped due to a node failure. When the intermittent link becomes available, the node removes the data packet from its persistent storage and forwards the data towards the destination.

Two issues are still present in this solution. First, this solution does not reduce the amount of storage necessary at the source node for reliable communications. Secondly,

all retransmissions are still requested from the source. Therefore, the delay in forwarding the retransmitted data packets is higher in most cases due to the intermittently connected links between the source and the location in the network where the data packet was dropped or lost. The custody transfer solution addresses both of these concerns by allowing the retransmission buffer for a data packet to reside at any capable node in the network rather than only at the source node.

Custody transfer reduces storage requirements at the source node by implementing a hop-by-hop ACK. In any reliability scheme, a node removes a data packet from its retransmission buffer when a corresponding ACK for the given packet is received. Any DTN node able to persistently store a data packet becomes the custodian node for the packet and sends an ACK to the previous custodian node. Upon receipt of the ACK, the previous custodian node removes the data packet from its storage. This mechanism reduces the storage requirement at a DTN source node. Therefore, a study of storage requirements for end-to-end reliability and custody transfer is presented.

5 Simulation Environment

To investigate the node storage improvement claims discussed in this paper, a simulation environment was developed using YACSIM, a discrete-event simulator, and a custom network simulation library. The custom network simulation library was developed to accurately model the dynamic behavior of a challenged networking environment. The details of the challenged networking environment and DTN mechanisms, such as custody transfer, contain detailed models. Other factors not associated with DTN are abstracted to increase the performance of the simulation environment and produce meaningful results in a shorter period of time. Other DTN simulation environments were investigated including the Java DTN simulator [15] and DTN simulator available in the DTN2 reference implementation [16], but neither of these simulation packages met the requirements of this particular DTN storage research.

Specifically, this custom simulator has the ability to allow for highly variable latency, intermittent connectivity, storage management, and traffic generation. The custody transfer mechanism is implemented as described in the bundle specification [17]. Nodes are modeled as storage, and links are modeled as queues using delay and capacity parameters. Therefore, network traffic stored and dropped due to intermittent connectivity is modeled with more detail than other network-based event-driven simulator that only add link latency time to transmitted packets without consideration of the FIFO property of the link.

For these simulations, intermittently connected links had randomly selected on-off times while maintaining the on-off percentage specified by the particular simulation run. This research investigated intermittently connected

links with on-off percentages of 25/75, 50/50, and 75/25, where the first number is on-time percentage and the second number is the off-time percentage. Bursty traffic generation models were used to simulate the use of file transfer applications in a DTN while, for this particular study, abstracting the details of the specific file transfer application. A 10-node moderately connected network was used for simulation. Each DTN node performs its own storage management.

6 End-to-End vs. Custody Transfer without Storage Constraints

Time-weighted mean network storage measurements were recorded from simulations of a network with intermittently connected links. These measurements are the time-weighted mean of the summation of the size of all data packets stored in the network. The results of the simulations are shown in figures 2, 3, 4. All of these results show significantly lower storage usage when custody transfer is used as the reliability mechanism as opposed to only an end-to-end reliability mechanism. The difference between the two curves in each figure increases as the number of links that are intermittent increase. In addition, the mean network storage using end-to-end reliability is approximately double the mean network storage using custody transfer when the number of intermittent links was greater than or equal to 5. These results show the storage usage benefits of using custody transfer as a reliability scheme.

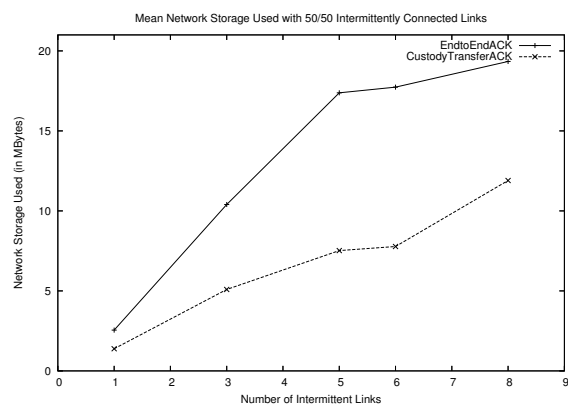


Figure 2: Network Storage Usage for 50/50 IC Links

The mean network storage results show a significantly reduced network storage requirement when using custody transfer. This result occurs because custody transfer enables the network to use available storage throughout the network as opposed to source node storage. End-to-end reliability protocols operating with intermittent connectivity require a large source node storage due to long round-trip times between source and destination. Custody transfer effectively shortens the round-trip time by reducing

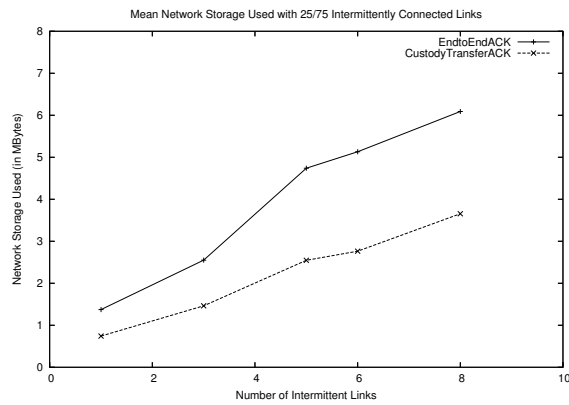


Figure 3: Network Storage Usage for 25/75 IC Links

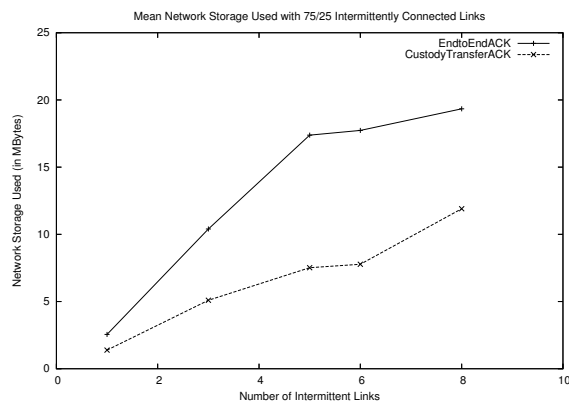


Figure 4: Network Storage Usage for 75/25 IC Links

the latency of the path between pairs of custodian nodes and transfers the responsibility of retransmission to other nodes in the DTN. On the other hand, the simulation results showed that non-source nodes may require more storage when using custody transfer because the storage requirements are distributed throughout the node in the network. Since nodes on the path from source to destination may become custodians, they require more storage to accept custody as opposed to only end-to-end reliability.

7 End-to-End vs. Custody Transfer without Storage Constraints

All previous simulation results assumed unlimited storage at each node. This set of simulations restricts the amount of storage at each node which shows the impact of node storage requirements on the reliability protocol selection. Simulation runs were executed with 1 MB storage at each node and 4 MB storage at each node for all three types of intermittent connectivity (50/50, 25/75, and 75/25). Results for each storage size and type parameter of intermittent connectivity were similar. Since the storage is constrained, it is more interesting to study the number of dropped packets due to the storage constraint. The results

presented in figure 5 used 1 MB node storage and 25/75 intermittent connectivity, which are the most constricting values for each simulation parameter.

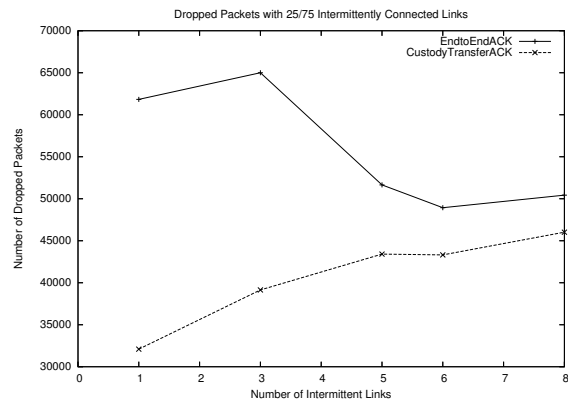


Figure 5: Dropped Packets for 25/75 IC Links

For a small number of intermittent links, the difference is dropped packets between end-to-end reliability and custody transfer is large, but this difference decreases as the number of intermittent links increase. When the number of intermittent links is low, end-to-end reliability drops packets at the source and the node (or nodes) connected to the few intermittent links. This is why the number of drops is significantly higher for end-to-end reliability as opposed to custody transfer. As the number of intermittent links increases, the number of drops at nodes between the source and destination increases dramatically and dominates the number of total drops. This is why the drop packet values are significantly closer when a large number of links are intermittent. Nevertheless, the results show clearly that custody transfer reduces the number of packets dropped when storage is constrained.

8 Future Work

The results given in this paper show the benefits of using custody transfer in networks with intermittent connectivity in terms of network storage and packet loss. The ultimate goal of this is to provide the best reliability possible in a DTN while minimizing the storage constraints put on DTN nodes to reduce size, cost, and power. Future work includes refining the storage management scheme for custody transfer to better utilize currently unused storage throughout the network.

References

- [1] P. Marshall, "Disruption Tolerant Networking Industry Day BAA 04-13." http://www.darpa.mil/ato/solicit/DTN/DTN_industryday.pdf, Jan 2004.

- [2] V. G. Cerf and R. E. Kahn, "A Protocol for Packet Network Intercommunication," *IEEE Transactions on Communications*, vol. 22, pp. 637–648, May 1974.
- [3] J. Postel, *Transmission Control Protocol - DARPA Internet Program Protocol Specification*, Sep 1981. RFC 793.
- [4] S. Burleigh, A. Hooke, L. Torgerson, and et al., "Delay Tolerant Networking : An Approach to Interplanetary Internet," *IEEE Communications Magazine*, vol. 41, pp. 128–136, Jun 2003.
- [5] I. F. Akyildiz, X. Zhang, and J. Fiang, "TCP-Peach+ : Enhancement of TCP-Peach for Satellite IP Networks," *IEEE Communications Letters*, vol. 6, pp. 303–305, Jul 2002.
- [6] O. B. Akan, J. Fang, and I. F. Akyildiz, "Performance of TCP Protocols in Deep Space Communication Networks," *IEEE Communications Letters*, vol. 6, pp. 478–480, Nov 2002.
- [7] O. B. Akan, J. Fang, and I. F. Akyildiz, "TP-Planet : A Reliable Transport Protocol for Interplanetary Internet," *IEEE Journal of Selected Areas in Communications (JSAC)*, vol. 22, pp. 348–361, Feb 2004.
- [8] G. Xylomenos, G. C. Polyzos, P. Mahonen, and M. Saaranen, "TCP Performance Issues over Wireless Links," *IEEE Communications Magazine*, vol. 39, pp. 52–58, Apr 2001.
- [9] F. Stann and J. Heidemann, "RMST : Reliable Data Transport in Sensor Networks," in *1st IEEE International Workshop on Sensor Net Protocols and Applications*, 2003.
- [10] C.-Y. Wan, A. T. Campbell, and L. Krishnamurthy, "Pump-Slowly, Fetch-Quickly (PSFQ) : A Reliable Transport Protocol for Sensor Networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, Apr 2005.
- [11] R. Patra and S. Nedeveschi, "DTNLite : A Reliable Data Transfer Architecture for Sensor Networks," *CS294-1 : Deeply Embedded Networks (Fall 2003)*, 2003. <http://www.dtnrg.org/>.
- [12] V. Cerf and et al., *Interplanetary Internet IPN : Architectural Definition*. IRTF, May 2001. <http://www.ipnsig.org/reports/memo-ipnrg-arch-00.pdf>.
- [13] K. Fall, W. Hong, and S. Madden, "Custody Transfer for Reliable Delivery in Delay Tolerant Networks," Tech. Rep. IRB-TR-03-030, Intel Research Berkeley, Jul 2003.
- [14] K. Fall, "A Delay-Tolerant Network Architecture for Challenged Internets," in *Proceedings of ACM SIGCOMM '03*, pp. 27–34, ACM Press, Aug 2003.
- [15] "Simulator for DTN." <http://www.dtnrg.org/>.
- [16] "DTN2 Reference Implementation." <http://www.dtnrg.org/wiki/Code>.
- [17] "Bundle Protocol Specification." <http://www.dtnrg.org/docs/specs/draft-irtf-dtnrg-bundle-spec-04.txt>.