Delay Tolerant Networking (DTN) Protocols for Space Communications

Ruhai Wang†, Xuan Wu†, Tiaotiao Wang†, and Tarik Taleb§
†Phillip M. Drayer Department of Electrical Engineering, Lamar University, Beaumont, TX, USA
§NEC Europe Ltd., Network Research Division, Kurfürsten-Anlage 36 69115 Heidelberg, Germany

(REVISION)

Contact Authors

Ruhai Wang
Drayer Department of Electrical Engineering
Lamar University
Beaumont, TX 77710-1029 USA
Phone: +1-409-880-1829
Fax: +1-409-880-8121
Email: wang@ee.lamar.edu

Tarik Taleb
NEC Europe Ltd.
Network Research Division
Kurfürsten-Anlage 36 69115 Heidelberg, Germany
Phone: +49-6221-4342-179
Fax: +49-6221-4342-155
Email: Tarik.Taleb@neclab.eu

Abstract

Delay/disruption tolerant networking (DTN) technology offers a new solution to highly stressed communications in space environments, especially those with long link delay and frequent link disruptions in deep space missions. It is becoming a recognized research area in computer networks and space communications. Extensive research work has been done on DTN for space communications in the past several years including numerous publications (Internet Engineering Task Force (IETF) Requests for Comments (RFCs) among them), DTN protocol implementations, and some experimental work done over simulation testbeds and in space, both low-Earth orbit and deep space. In this chapter, we provide an overview of the emerging DTN protocols for space communications with focuses on DTN architecture, BP, convergence layer protocols and their application in space. An overview of recent research and experimental activities on DTN for space communications is also presented.
1. INTRODUCTION

The National Aeronautics and Space Administration (NASA) is moving towards packet-switched space communications using appropriate network architecture and protocols [1-3]. Transmission control protocol (TCP) [4] experiences severe performance degradation in space due to its inherent design and the vast differences between terrestrial and space environments [2, 3]. Delay/disruption tolerant networking (DTN) technology [5, 6] offers a new solution that use a store-and-forward mechanism to combat long link delay and frequent link disruptions generally characterizing space communications. DTN communications use a bundle protocol (BP) [7], to construct a store-and-forward overlay network. The BP, situated at the application layer of the Internet model, utilizes conventional Internet protocols to send and receive data bundles. The DTN architecture requires a convergence layer protocol (CLP) [7, 8] in order to run the bundle protocol over the underlying transport layer protocols. CLPs currently available or under development for space internetworking include the TCP-based CLP (i.e., TCPCL) [8], user datagram protocol (UDP) [9]-based CLP [10], Saratoga [11], Licklider transmission protocol (LTP, also called a long-haul transmission protocol) [12-14], and LTP-Transport (LTP-T) [15]. The TCPCL protocol for the DTN system works together with the well-known TCP to assure reliable communication services between DTN nodes. In other words, the DTN nodes are capable of invoking various types of recovery mechanisms if they experience data losses because of link interruption or a TCP connection outage. LTP is intended to operate over point-to-point, long-haul, deep-space radio frequency links or similar links characterized by an extremely long transmission delay and/or frequent interruptions in connectivity [12]. Unlike TCPCL, LTP does not include any flow or congestion control signaling.
Although originally targeting challenged interplanetary Internet (IPN) [16], the DTN techniques and concepts have also been considered for use in other operational environments that are subject to link disruption and long link delay, such as terrestrial wireless networks, sensor-based networks, and tactical/military ad-hoc networks [17-26]. NASA considers DTN the most suitable technology to be employed in space internetworking and hopes to fly with it on space missions soon [27].

In this chapter, we provide an overview of the emerging DTN protocols for space communications with focuses on DTN architecture, BP, convergence layer protocols and their application in space. We also review recent research and experimental activities on DTN for space communications.

The remainder of this chapter is organized as follows. In Section 2, we provide an overview of space communications, review the performance problems of TCP in space environment, and then discuss the recently developed data transport protocols for space and interplanetary Internet. In Section 3, we introduce general DTN, DTN for space, BP, and DTN convergence layer protocols. In Section 4, we review recent research and experimental activities on DTN for space communications. A summary is provided in Section 5. Finally, open research issues are discussed in Section 6.

2. SPACE COMMUNICATIONS AND DATA TRANSPORT PROTOCOLS FOR SPACE/INTERPLANETARY INTERNET

2.1 OVERVIEW OF SPACE COMMUNICATIONS

From early geo-stationary orbit (GEO) relay and broadcast communication satellites and the Advanced Communications Technology Satellite (ACTS) testing of the Internet technology,
through the advanced tracking and data relay satellite system (TDRSS) and the deep space networks (DSN), to the current space Internet and Interplanetary Internet, space communications has been changing significantly with respect to both communication architecture and operating model. This movement from the traditional model of managed, mission-specific pairwise links to an automated network involving multiple nodes using relay and the Internet technologies [28] is of historical significance.

Planetary exploration missions have historically accomplished communication between data sources on mission planets and stations on Earth with dedicated direct communication circuits. Considering the limitations of direct connection time between any two moving planets, missions have recently begun to use relaying in space communication for effective data transmission [29]. Over 95% of the data received from the Mars Exploration Rovers has been returned to Earth via the Mars Global Surveyor and Mars Odyssey spacecraft [30]. This experience has demonstrated that using orbiters to relay data from the Mars surface can greatly increase data return. Similar benefits are expected around the moon and at other planets [27, 29].

In a typical space communication scenario, this kind of relay can be achieved in two ways: (1) through the planet-orbiting spacecraft; and (2) through the Earth-orbiting spacecraft such as the TDRSS [31, 32]. The first way is the primary method, while the second can be considered the backup choice. In Fig. 1, we illustrate a typical architecture using relaying spacecraft to forward data from another planet to Earth. The data sources on the planetary surface can be roving robotics, crews, surface stations or ascent/decent vehicles. Although the relay operation shown in Fig. 1 is done through a single spacecraft, multiple spacecraft or even constellations of satellite can be involved in the realistic space relaying communications. As the number of relays and relay users increases, international cross-support for this relaying capability will increase mission
robustness and data return, giving planetary landers multiple opportunities to forward data to Earth.

It is important to bring internetworking technology to space communications to enable as much autonomous operation as possible. The more autonomous we can make operations, the less human controlling and scheduling is involved and the greater the savings in reduced cost of operations [28].

2.2 PERFORMANCE PROBLEMS OF TCP IN SPACE

As a reliable and widely-used transport protocol of the Internet protocol suites, TCP [4] performs very well in today’s Internet operations. However, its performance in space communications is severely degraded by the inherent conflict between its design and the challenging characteristics of the very different space environment [2, 3].

Several issues of the space internetworking are inappropriately addressed by TCP, limiting its performance. One major issue is that TCP is designed to use window-based transmission control algorithms. With the sliding window flow-control mechanisms, TCP regulates the amount of data a source can send by adjusting the window size in response to the acknowledgement information from the destination; the timeliness of that information is key to the effectiveness of the technique.

Another major issue is that TCP cannot distinguish between data losses caused by network congestion and link errors.

The impact on performance of TCP caused by these issues is not very obvious in the terrestrial Internet. However, it becomes obvious and even serious due to the challenging space communication environment. Long and variable propagation delays, frequent and length link
outages, strong channel noise, and high channel-rate asymmetry in space all conspire to adversely degrade the performance of TCP.

Round-trip delays involved in space communications, especially those in deep space, is much longer than the terrestrial Internet. They can be up to 40 minutes long for the Martian channel and even much longer for a channel from Jupiter to Earth. The problem with long end-to-end RTTs in space communications is that they hurt TCP interactions between the space node and the ground Internet node. As a result, this limits the usefulness of TCP acknowledgment information from the remote destination node and influences the effectiveness of TCP transmission control.

Planetary bodies, asteroids or spacecrafts moving in space all may periodically and randomly interrupt the communication link and cause frequent and long link outage between two communication endpoints [2]. This can cause a significantly large number of packet losses and retransmissions, resulting in severe performance degradation of TCP.

Space communications feature much noisier channels than terrestrial communications, resulting in very high bit error rates (BER) with $10^{-5}$ very common and even rates on the order of $10^{-1}$ in the deep space environment [2]. As well discussed in literatures, TCP makes an erroneous congestion decision which misinterprets any data loss as a congestion loss. In general, multiplicative-decrease in sending data rate is designed as a TCP response to data loss caused by network congestion. Consequently, the throughput of TCP is unnecessarily decreased when data loss is caused by bit error corruption instead of real network congestion. In fact, data loss due to a high channel BER in the space environment has a disproportionately negative effect on TCP performance.
Space communications channels are frequently asymmetric in terms of channel bandwidth: the bandwidth of the uplink, from the ground to the spacecraft, is generally much lower than the bandwidth of the downlink channel, from the spacecraft to the ground. The space channel asymmetry occurs mainly due to the scientific purpose of space missions and the cost that would be incurred by an increase in receiver power and antenna size as would be required for high bandwidth reception [3]. For most space missions, the data flowing from the space to the ground over the downlink is substantially larger than the data flowing from the ground to space over the uplink. For this reason, the downlink designed for bulk data transfer has a broad bandwidth, while the uplink designed to send acknowledgement (ACK) messages or telecommand messages for commanding the spacecraft generally has a much narrower bandwidth. Channel rate asymmetry in space can severely affect the performance of TCP because TCP is ACK-clocked, relying on the timely feedback of ACKs through the uplink to make steady progress of data transmission. However, the slow uplink rate cannot handle the transmission of returning ACKs effectively. This may result in frequent losses of ACKs and consequential performance degradation of the protocol. For a detailed discussion on how the network channel asymmetry affects the performance of TCP, see [33]. TCP generally can not work effectively when the channel asymmetric ratio (defined as the ratio of ACK channel rate over data channel rate) is lower than 1/50 [3, 34, 35]. The ratio of channel asymmetry can be as low as 1/1000 in Earth-orbit communications and even significantly lower in deep space environment.

The above performance issues of TCP in satellite and space communications have been fully investigated, and thus are not discussed in detail in this chapter. For a thorough literature review, see [2, 3, 34, 35].
2.3 Data Transport Protocols for Space/Interplanetary Internet

While the feasibility of operating space communications using the Internet-type protocol has been a source of contention within NASA [1], a large number of transmission control protocols have been developed for space and other environments with similar communication conditions [2, 3]. Based on the design and operating methodology, these protocols can be roughly classified into the following three categories:

(1) Category I: The protocols in this category involve changes only to the TCP protocol. The changes are generally on the congestion-control and error-control algorithms of TCP.

(2) Category II: The protocols generally involve modifications to the TCP protocol and/or network operation infrastructure and elements.

(3) Category III: This category involves new protocols that are designed to provide data transport functionality but operate at the application and link layers. DTN bundle BP [7] and LTP [12-14] produced DTN Research Group (DTNRG) of the Internet Research Task Force (IRTF) [36] are considered the representative protocols in this category.

For a comparative summary and classification of these protocols, their design techniques and performance evaluation, please refer to [37].

The current issues in network, link and physical layers for space and Interplanetary Internet are not discussed here. Refer to [2] for a detailed discussion.

3. DTN, BP and Convergence Layer Protocols

3.1 Overview of DTN

In comparison to conventional Internet architecture (i.e., TCP/IP architecture), DTN is a networking architecture designed to provide communications in highly stressed environments characterized by long or variable delays, intermittent loss of link connectivity, high error rates,
and asymmetric data rates [38]. Originally developed for IPN, DTN was extended to terrestrial wireless networks and wireless sensor networks (WSNs) as those types of networks also suffer from link disruption and delay.

Application data units in DTN are carried in variable-length protocol data units (PDUs) called bundles that are intended to minimize the round-trip exchanges needed to complete a protocol transaction. In other words, if all the information required to complete a transaction is bundled, the number of exchanges between the message sender and the receiver can be reduced. This is significant to the protocol performance if the link propagation delay is extremely long such as hours or longer (i.e., in deep space).

Extensive work has been recently done in DTN routing with most researchers studying routing algorithms. The DTN routing algorithms are mostly designed under the assumption that paths are not continuously connected end-to-end between message source and destination. Some recent studies have also integrated DTN concepts into mobile ad-hoc network (MANET) research which mainly focuses on routing in dense mobile ad-hoc networks where end-to-end connectivity is possible [39]. If DTN and MANET are combined, the routing of bundles will involve both simple forwarding and store-carry-and-forward operation as required.

DTN adopts a traditional custody transfer mechanism to deal with harsh communication environments [7]. This mechanism is optional. However, general description of DTN architecture assumes the “custody transfer” option is enabled. With this mechanism, DTN keeps track of a current custodian for each bundle. A DTN custodian retains in persistent memory a copy of each bundle it has sent, discarding that persistent copy only when an ACK is received from another node confirming that this node has received the bundle successfully and has accepted responsibility for forwarding it, as illustrated by the DTN operation architecture in Fig.
2. By this means, no bundles are lost even if contact with the designated receiving node is interrupted due to a transient change in network topology. The adoption of custody transfer and persistent storage at intermediate nodes makes DTN capable of delegating data retransmission responsibility to nodes other than the original source node; this feature is a significant departure from Internet architecture, in which all retransmission is end-to-end and the source node bears all responsibility for retransmission.

The DTNRG of IRTF [36] concerns itself with the whole range of potential uses for DTN, including “DTN for space”.

3.2 DTN for Space

As the Internet protocols (mainly TCP/IP) are not effective across frequently disrupted links, long and variable delay, and high rates of data loss due to corruption, DTN technology is being considered for space communications. It is believed that integrating DTN into current space communication architecture will enable Internet-like user interaction to be maintained even in highly stressed space exploration environments.

In the Internet, connectivity is generally continuous and signal propagation latency is very small. This means that changes in network topology can be discovered dynamically and communicated to routers in time to revise computed routes before much traffic is misdirected. To some degree this is also true of terrestrial DTNs, making it plausible to use routing protocols such as Probabilistic Routing Protocol using History of Encounters and Transitivity (PROPHET) [26]. But these conditions do not hold in space communications, especially in deep space, so the kinds of “routing protocols” that work well in the Internet or even in some terrestrial DTNs are of little utility in space networking: by the time a node learned of the start of a communication opportunity the opportunity might already have ended. On the other hand, most changes in space
network topology are intentional and scheduled rather than inadvertent, so a different sort of "routing protocol" may be used instead: a "contact plan" enumerating planned episodes of connectivity may be simply declared to routers well in advance of the scheduled changes in topology. Routes may be computed based on topology changes that are announced rather than discovered.

One other significant difference between Internet protocol implementations and DTN is the site of retention of data in transit. Because data may need to be retained for hours or even days in a relay satellite before it can be forwarded, bundles may best be stored in long term persistent mass memory \[38\] rather than dynamic memory. In contrast, an Internet router generally need not retain a received packet for more than a few milliseconds before transmitting it; transient retention in high-speed dynamic RAM is appropriate.

3.3 Bundle Protocol (BP)

BP forms a store-and-forward overlay network to provide custody-based, message-oriented transmission and retransmission. To provide the store-and-forward message switching service, BP \[7\] is designed to operate as an application layer on top of heterogeneous underlying "convergence-layer" protocol (CLP) stacks, among which may be the Internet protocol stack itself.

The major capabilities of BP are the ability to cope with connectivity interruption and the ability to take advantage of scheduled, predicted, and opportunistic connectivity. BP can also be configured for retransmission on any failure of custody transfer, and it supports optional requests for end-to-end acknowledgment at the application layer.

In order to utilize an underlying convergence-layer protocol stack such as TCP/IP or UDP/IP, BP needs a convergence layer adapter (CLA) \[8\] deployed between the bundle layer and
transport layer. (This implies that both the CLA and the BP function at the application layer in the Internet protocol stack.) The bundle protocol agent (BPA) [7] of a DTN node executes BP procedures, invoking the services of the CLA to do so. A CLA sends and receives bundles on behalf of the BPA, using the transport service of the underlying internetworking protocols.

3.4 DTN Convergence Layer Protocol (CLP)

A variety of protocols can act as the CLP, as long as they can provide the services the BPA needs. Fig. 3 illustrates how the DTN BP and CLP are layered in the Internet-protocol stack.

The TCPCL protocol adapter uses TCP to provide reliable communication services between DTN nodes. When a bundle node establishes a TCP connection, it establishes a TCPCL connection at the same time for bundle communication. The TCPCL connects DTN nodes via a TCP channel. Over the established TCPCL connection, the sender can send bundles to the destination through the next node. While TCP itself is a reliable protocol, the TCPCL adaptation is additionally capable of invoking various types of recovery mechanisms if the transfer of a bundle experiences interruption because of a TCP connection outage. For an idle connection, a one-byte keep-alive message may optionally be sent at a defined interval. The protocol uses this to detect loss of connection in the absence of bundle traffic.

A simple UDP-based CLP has been implemented in the DTN-2 reference implementation [10]. This UDP-based CLP is intended for use over dedicated private links where congestion control is not required. Its design is based on a presumption that a bundle will always fit into a single UDP datagram, which is limited to around 64 Kbytes. In other words, this CLP is not able to support segmentation of large DTN bundles across multiple UDP packets. Another simple UDP convergence layer has been defined for unidirectional transport over both unicast and multicast networks [40].
Compared to these basic UDP-based CLP implementations, Saratoga [11], LTP [12] and LTP-T [15] are relatively complicated CLPs that support segmentation of large DTN bundles across multiple UDP packets or, in the case of LTP, packet structures for other link-layer protocols. These CLPs inherit core design ideas from the Consultative Committee for Space Data Systems (CCSDS) file delivery protocol (CFDP) [41]. While Saratoga was originally developed for efficient transfer of image data onboard the IP-based satellites, it can now work in DTN as a CLP for exchanging bundles between DTN peer nodes. It is a lightweight UDP-based transfer protocol intended for use between point-to-point peers that have sporadic, intermittent connectivity using dedicated IP links [11]. In other words, Saratoga is not developed for use over shared paths, so it requires no congestion control mechanism. Saratoga is intended to run above UDP and UDP-Lite [42]. It implements checksums across each hop to protect data integrity.

Intended to serve as a reliable CLP underlying BP, the newly-developed LTP provides reliable communications between adjacent DTN nodes for interplanetary space. In contrast to Saratoga, LTP is designed to be able to run over both UDP and CCSDS link-layer protocols and to perform retransmission-based recovery of lost data with a selective repeat ARQ mechanism. LTP recognizes each block of data as having two parts: a "red-part", whose delivery must be assured by acknowledgment and retransmission, followed by a "green-part", whose delivery is attempted, but not assured. The length of either part may be zero; that is, any given block may be designated entirely red or entirely green. Thus, LTP can provide both TCP-like and UDP-like functionality concurrently in a single session. Unlike TCP, LTP includes no flow or congestion control signaling. Like CFDP, however, LTP implements optional mechanisms for accelerating retransmission at the cost of some additional signal traffic.
An extension of LTP, LTP-T is proposed as an end-to-end capable transport protocol for data transmissions over multi-hop space networks [15].

As an alternative transmission mechanism for deep-space data transfers, Deep-Space Transport Protocol (DS-TP) [43] has also been proposed. DS-TP is designed as a rate-based, reliable transport protocol whose advantage is its efficient and fast retransmission mechanisms.

For a detailed discussion of DTN, BP, CLPs and DS-TP, refer to [6–8, 13, 43].

4. RECENT RESEARCH AND EXPERIMENTAL ACTIVITIES ON DTN FOR SPACE COMMUNICATIONS

DTN has become a recognized research area in computer networks and space communications [39]. Extensive research work has been done on DTN for space communications in the past several years, leading to numerous publications including several Internet Engineering Task Force (IETF) Request for Comments (RFCs): RFC 5050 Bundle Protocol Specification [7], RFC 5325 Licklider Transmission Protocol – Motivation [12], RFC 5326 Licklider Transmission Protocol – Specification [13], and RFC 5327 Licklider Transmission Protocol - Security Extensions [14]. While RFC 5050 introduced the means by which a DTN store-and-forward overlay is formed for custody-based message transmission, RFCs 5325-5327 describe how retransmission-based reliability can be ensured and secured for DTN over deep space links.

The CCSDS is currently working on DTN standardization for space internetworking. Agreements are also being established on the standards that will enable interoperability and cross support operations. The Space Internetworking Strategy Group (SISG), which is composed of technical experts appointed by the Inter-agency Operations Advisory Group (IOAG) agencies, considers DTN to be the only mature candidate protocol available to handle long propagation
delays, frequent and lengthy network disruption inherent in space missions involving multiple spacecraft [27]. The SISG concludes that DTN should be provided as the main end-to-end routing service in the future space networks and it is a high priority to mature DTN to full flight readiness for a wide variety of space missions by 2012. For the protocol implementation, Ohio University’s BP and LTP reference implementations in Java [44], NASA JPL’s ION system for DTN in space [45] and the DTN reference implementation in C++, DTN-2 [10], are already openly available for evaluation.

Some experimental work has recently been done on DTN in the space environment.

Experiments with a United Kingdom disaster monitoring constellation (UK-DMC) satellite in low-Earth orbit have demonstrated data transfer from space using DTN BP [46, 47]. In this experiment, an implementation of a subset of the bundle protocol onboard the DMC satellite, utilizing Saratoga as the CLP, was the first successful use of bundles to transmit images from space to a ground data system in presence of link disruption [47]. The experiment showed that providing a bundle integrity checksum can improve the reliability of BP. It also shows that network time synchronization is an important network configuration consideration.

As the first of a series of NASA experiments to mature DTN to flight readiness for a wide variety of mission, the Deep Impact Network Experiment (DINET) was conducted by JPL in October and November of 2008, using the EPOXI spacecraft located about around 15 million miles away from Earth as the first DTN router in space [48]. The DINET project validated the use of DTN protocols (JPL’s ION implementation) in deep-space Internet. During this month-long experiment, dozens of space images (about 14.5 MB) were transmitted from DTN nodes at JPL to the spacecraft and then automatically forwarded from the spacecraft back to the JPL nodes, hours or days later, without data loss or corruption. The experiment exercised the
principal elements of DTN technology including automatic, contact-sensitive relay operations, custody transfer, rate control, and delay/disruption-tolerant retransmission [49] in the presence of significant light-time delay and both planned and unplanned network disruptions. The experimental results show that DTN BP and LTP work very effectively. They can tolerate signal propagation delay longer than 80 seconds and network outage on the order of days, and they effectively handle station handovers and transient failures in DSN uplink service [49]. ION had no failure in four weeks of continuous operation on VxWorks, Solaris, and Linux platforms, although several bugs were found in the contact graph routing implementation which resulted in some under-utilization of network capacity. It was also found that the EPOXI spacecraft clock drifted more rapidly (around 1 second per day) than expected. However, time synchronization errors and one-way-light-time (OWLT) estimation errors of several seconds had no significant impact on DTN network operation [49]. In summary, the DINET project demonstrated that DTN is ready for operational use in space missions [48] and that the ION implementation of DTN is of sufficient quality that future space missions can use it at low risk.

NASA also began experimenting with DTN on the International Space Station (ISS) in July 2009, in a deployment led by the University of Colorado, Boulder (CU-Boulder) [50]. With the ION implementation of DTN installed on a BioServe payload known as Commercial Generic Bioprocessing Apparatus 5 (CGBA5), DTN bundles are transmitted from ISS to Marshall Space Flight Center (MSFC) in Huntsville, AL, and then forwarded to the payload operation control center at CU-Boulder. NASA and CU-Boulder are currently working to extend the DTN experiments on ISS so that other space agencies such as European Space Agency (ESA) and the Japanese Aerospace Exploration Agency can participate [50].
Some other projects are also under way to investigate DTN operation in space. In close collaboration with NASA JPL, extensive work has been done at Lamar University, Texas, to study DTN operation and performance in space communications, including establishment of two PC-based testbeds and experimental performance evaluation of various DTN implementations. One testbed is built to simulate a slow-speed point-to-point space link [51-53] and another to simulate high-speed Solar System Internetwork (SSI). With these testbeds, extensive experiments have been done in evaluating the DTN protocols in cislunar and deep-space communication environments, including CFDP [52], DTN-2 [53], JAVA-BP/LTP Reference Implementations from Ohio University [54], and ION [55, 56]. The experimental results show that for both JAVA-BP/LTP and ION implementations, BP/LTP/UDP/IP transmissions have a performance advantage over BP/TCPCL/TCP/IP and BP/UDPCL/UDP/IP transmissions over cislunar communications links at a high BER around $10^{-5}$, and the advantage becomes larger as the link delay increases [54, 55]. The experiments at Lamar University also investigated the performance of hybrid of ION protocols, with LTPCL, TCPCL and UDPCL running over different hops of a typical three node, relay-type of interplanetary links. It is found that the hybrid of TCPCL and LTPCL has significant goodput advantage over absolute LTPCL transmissions (i.e., LTPCL over all hops) and any other protocol combinations, and the configuration involving LTPCL on a longer hop (e.g., the hop from orbital relay spacecraft to Earth over the primary relay path in Fig. 3) is better able to tolerate long propagation delay and high BER than other configurations [56].

The ESA has established a DTN/IP Space-Ground testbed and is developing protocol implementations for evaluation, mainly targeting Martian communications [57].
5. Summary

In this chapter, we provide an overview of the emerging DTN protocols for space communications. We discuss the operation of space communications, the performance problems of TCP in space, and available data transport protocols for space and Interplanetary Internet. The chapter focuses on a discussion of DTN technologies including the general DTN, DTN for space, BP, and convergence layer protocols. As an essential part of the chapter, the latest development and experimental activities on DTN for space communications are also reviewed. Based on the available simulation and experimental results, it can be concluded that DTN BP and LTP protocols work very effectively with space missions in the presence of significantly long link delay and frequent and lengthy network outage. While international organizations are working on DTN standardization, flight experiments have already demonstrated DTN readiness for operational use in space missions. DTN for space communications is gradually approaching maturity, despite some issues that need to be addressed.

6. Open Issues

Although some file transfer testing using DTN has been done successfully for space communications in both testbed [53-56] and real-world deployment [46-49], these exercises have all been experimental activities. While the DINET experiment demonstrated DTN readiness in space missions, DTN for space is still far from being incorporated into mission-critical long-running applications. At this stage, several issues must be addressed prior to pervasive deployment, especially in regard to routing, security, and clock synchronization [39, 46, 49].

Although contact graph routing was shown to be generally effective in the DINET experiment (despite some implementation bugs), no dynamic routing protocol has yet been standardized for DTN in space.
Security will soon become an important issue in space missions. A lot of effort has gone into providing encryption and authentication to support lunar missions, but the significant differences in bandwidth and propagation delay between lunar missions and deep-space missions make network security – in particular, key management – an area of continuing research.

Clock synchronization was shown as a problem in both LEO UK-DMC [46, 47] and deep-space DINET projects [48].

The DTN architecture remains a topic of some controversy. In [58], the authors assert that the DTN Bundle Protocol is not well-suited to many operational scenarios that it was intended to support, especially with respect to error detection and transmission reliability. This suggests that additional experimental and theoretical investigations need to be done to better understand the operation of space DTN, and practical remediation strategies may be needed for enhanced performance.

According to our experimental results, the TCP-based DTN-2 shows better performance over less error-prone links (such as terrestrial links) while the LTP protocol has performance advantages over lossy space links. With this conclusion, we suggest a hybrid of DTN-2 and LTP protocols should be tested to evaluate their performance in an integrated terrestrial and space networking scenario.

According to our research results [59], the traffic shaping mechanism of a rate-based transmission is much more effective than the bursty flow of window-based transmission over a long-delay, lossy space communication channel. There exist new protocols such as DS-TP and Saratoga for deep space communications to control data transmissions in a rate-based manner. We suggest these new protocols should be evaluated to investigate the performance advantage of a rate-based transmission mechanism in the deep-space environment.
7. Acknowledgment

The authors would like to thank the guest editors, Drs. Athanasios Vasilakos, Yan Zhang, Thrasyvoulos Spyropoulos, and the anonymous reviewers for their constructive comments and suggestions.

The authors would also like to thank Scott Burleigh, the globally well known expert and scholar in Interplanetary Internet/DTN and the primary developer of ION and CFDP, at NASA’s Jet Propulsion Laboratory (JPL), for significant discussion and help in improving the quality of the chapter. This chapter could not be completed without his help.
References


[45] Interplanetary Overlay Network (ION), November 2009 release, [Online]: https://ion.ocp.ohiou.edu/.


Fig. 1. Relay infrastructure in interplanetary communications.
Fig. 2. Custody transfer of DTN operation architecture.

Fig. 3. Layering of DTN BP and CLPs in the Internet-protocol stack.