Delay- and Disruption-Tolerant Networking

Alex McMahon and Stephen Farrell • *Trinity College Dublin*

Delay- and disruption-tolerant networking (DTN) grew out of attempts to develop an Interplanetary Internet but has evolved into an active area of networking research, with applications in space networking, military tactical networking, and networking for various challenged communities. The DTN Research Group provides an open forum in which DTN researchers and developers can collaborate to further develop this experimental technology.

Many important Internet transport protocols fail in environments without contemporaneous end-to-end connectivity. Here, we review the *delay- and disruption-tolerant networking* (DTN) approach to dealing with this problem, with an emphasis on the "prestandards" work of the Internet Research Task Force's (IRTF) Delay-Tolerant Networking Research Group (DTNRG; www.dtnrg.org).

Background

In 1973, Vint Cerf and Robert Kahn wrote a pioneering paper on TCP,[1] and 30 years after sowing this seed, the Internet had become ubiquitous. In 1997, another seed was sown: Cerf thought that "an interplanetary backbone" was necessary for us to prepare for future needs (see www.wired.com/wired/archive/8.01/solar.html). He and scientists from NASA's Jet Propulsion Laboratory (JPL), who had been working since the early 1990s on adapting Internet protocols for space missions, shared a space-networking vision and, in 1998, started collaborating on developing an *Interplanetary Internet* (IPN).

To support the IPN, the IRTF formed an Interplanetary Internet Research Group (IPNRG). Figure 1 shows the group's vision from around this time. Initially, the IPNRG included contributors from Worldcom, JPL, Mitre, SPARTA, and a few universities. Completion of the first phase of the IPN project led to the publication of "Interplanetary Internet (IPN): Architectural Definition" (draft-irtf-ipnrg-arch-00.txt) in May 2001.

In the IPN scenario, transmission is subject to significant propagation delays (a minimum of roughly 4 minutes one-way light-trip time between Earth and Mars) and intermittent connectivity due to planetary movement and the occultation of spacecraft as they orbit a planet or as planets rotate. The extremely limited power available to many spacecraft dictates a particular need for efficiency at all protocol layers in the IPN. In addition to propagation delays and intermittent connectivity, low and highly asymmetric bandwidth as well as a relatively high bit-error rate also distinguish IPN communication from most of the terrestrial networking scenarios with which we're familiar.

The overall conclusion was that simply extending the Internet protocol suite to operate end-to-end over interplanetary distances wasn't feasible and that new techniques would be necessary. The IPNRG referred to their chosen approach as *bundling*, which builds a store-and-forward overlay network above the lower layers of underlying networks. Whereas the Earth's Internet was basically conceived as a "network of connected networks," the IPN was thought of as a "network of disconnected Internets" connected through a system of gateways forming a stable backbone across interplanetary space.

The Birth of DTN

During 2001 and 2002, IPN researchers investigated how they could apply the IPN architecture

 1089-7801/09/$26.00 © 2009 IEEE

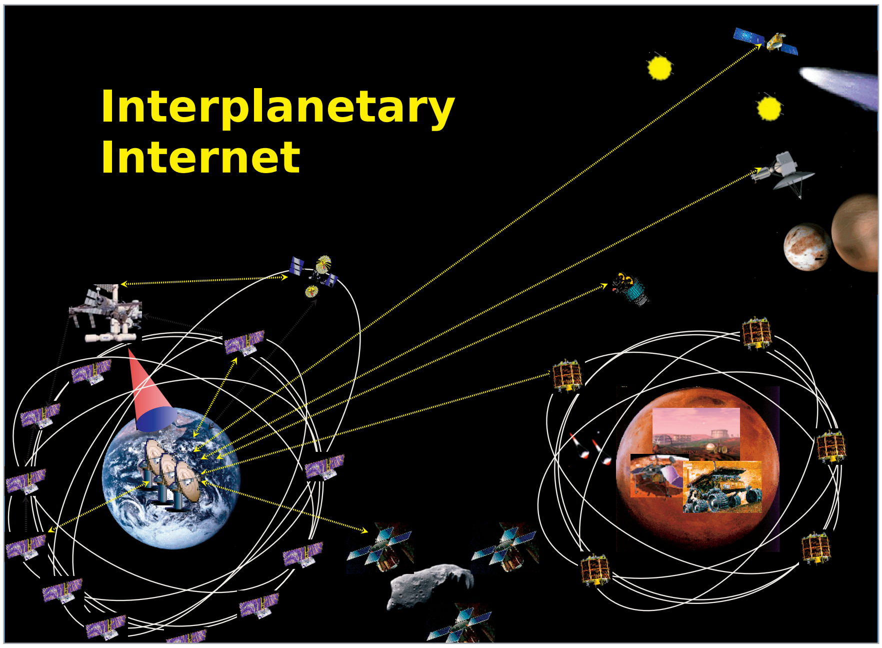

Figure 1. The early Interplanetary Internet vision. (Source: Adrian J. Hooke; used with permission.)

to other situations in which communications were subject to delays and disruptions. In August 2002, the IPNRG published an updated version of the draft as "Delay-Tolerant Network Architecture: The Evolving Interplanetary Internet" (see www.ietf.org/id/draft-irtf-ipnrg-arch-01.txt), which described a generalization of the architecture designed for IPN as an architecture for *delay-tolerant networks* (DTNs), a name coined by Kevin Fall of Intel Research. Among other updates, the IPNRG restructured the document to distinguish between an architecture for delay-tolerant networking and the application of that architecture to various extreme communications environments, including the IPN.

By this time, the IPNRG had realized that the different environments in which its architecture was applicable shared some essential characteristics, including the communication challenges introduced by long delays, intermittent connectivity, data rate asymmetry, packet loss, and errors. The updated draft provided examples of extreme environments and presented some problems inherent to using existing Internet protocols and applications. The authors also considered extreme terrestrial environments in which communications were subject to intermittent, probabilistic connectivity that would benefit from the architecture — these included military tactical networks, sensor networks deployed in oceanic environments, and communities living in extreme environments, such as the Sámi people of northern Scandinavia.

By May 2002, discussions were under way to recharter the IPNRG, and soon after, the DTNRG was formed. By 1 October 2002, a "dimming the lights" of IPNRG and its supporting interest groups had begun. Instead, the DTNRG began to address the architectural and protocol design principles arising from the need to provide interoperable communications with and among extreme and performance-challenged environments, where we can't assume continuous end-to-end connectivity.

The DTNRG worked first on further generalizing the IPNRG's architecture drafts and published its first draft in 2003 (see http://tools.ietf.org/html/draft-irtf-dtnrg-arch-00). Over the next few years, the research group refined the DTN architecture and eventually published RFC 4838.[2] Since then, the DTNRG published four more experimental RFCs, the first of which was the Bundle Protocol Specification (BP) describing the end-to-end protocol and abstract service description for exchanging messages (*bundles*) in DTN.[3] In 2008, three more RFCs[4–6] followed, describing the Licklider Transmission Protocol (LTP), a delay- and disruption-tolerant point-to-point protocol. LTP provides retransmission-based reliability over links characterized by extremely long round-trip times (RTTs).

The DTNRG is currently a very active research group with roughly 20 current Internet drafts on topics related to DTN security, routing, and various other BP extensions. DTNRG meets during most, but not all, IETF meetings and is an open research group, meaning that anyone interested can contribute simply by joining the mailing list and getting involved in the work. Aside from the DTNRG, an active research community is working on DTN-related topics. Researchers, for example, have conceived and implemented many routing schemes for DTN, including Delay-Tolerant Link State Routing (DTLOSR),[7] Contact Graph Routing (CGR),[8] and the Resource Allocation Protocol for Intentional DTN (RAPID),[9] although, so far, only one routing scheme has been fully documented as an Internet draft (see www.ietf.org/id/draft-irtf-dtnrg-prophet-02).

Further details about the DTN architecture's evolution are available in an architectural review paper published in the June 2008 special DTN edition of the *IEEE Journal on Selected Areas in Communications*.[10]

The DTN Architecture

RFC 4838 points out some fundamental assumptions built into the Internet architecture that are problematic in DTNs:

- An end-to-end path between the source and destination exists for the duration of a communication session.
- Retransmission based on timely and stable feedback from data receivers is an effective means for repairing errors (for reliable communication).
- End-to-end loss is relatively small.
- All routers and end stations support the TCP/IP protocol suite.
- Applications need not worry about communication performance.
- End-point-based security mechanisms are sufficient for meeting most security concerns.
- Packet switching is the most appropriate abstraction for interoperability and performance.
- Selecting a single route between sender and receiver is sufficient for achieving acceptable communication performance.

The DTN architecture relaxes most of these assumptions — it uses variable-length messages as the communication abstraction and a naming syntax that supports a wide range of naming and addressing conventions to enhance flexibility. It's designed to use storage within the network to support store-and-forward operation over multiple paths and potentially long timescales, and not to require but to support end-to-end reliability. The DTN architecture envisages security mechanisms that protect the infrastructure from unauthorized use by allowing for policy-based discarding of traffic as quickly as possible. The DTN architecture also assumes roughly synchronized clocks, an aspect currently provoking debate within the DTNRG given that various researchers have offered examples in which this synchronization is problematic.

The Bundle Protocol

BP was specifically developed conformant to the DTN architecture. It essentially runs at the application layer and generally follows the overlay-network approach. Although BP can run over TCP/IP, it can also run over other, lower-layer protocols (so-called convergence layers) — for example, proprietary protocols deployed in sensor networks or, for deep-space deployments, LTP. BP's key capabilities include custody-based retransmission; an ability to cope with intermittent connectivity; an ability to take advantage of scheduled, predicted, and opportunistic connectivity (in addition to continuous connectivity); and late binding of overlay network endpoint identifiers (EIDs) to convergence layer-specific addresses, such as IP addresses. Devices implementing BP are called *DTN nodes*.

BP forms an overlay that employs persistent storage to help combat network interruption and for its store and forward function. This overlay includes transfer of reliable delivery responsibility, optional end-to-end acknowledgment, and several diagnostic and management features. For naming, BP uses a flexible scheme (based on URIs[11]) that can encapsulate different naming and addressing schemes in the same overall naming syntax. BP is layer-agnostic and focuses on virtual message forwarding rather than packet switching.

DTN-Enabled Applications

A DTN-enabled application is modeled around sending and delivering *application data units* (ADUs), which can be of arbitrary length but which might be subject to implementation limitations — for example, in space applications, ADUs might be limited to being less than 64 Kbytes in size. The network might or might not preserve their relative order. Typically, the DTN sends or delivers ADUs to applications in complete units, although fragmentation can also occur within the network unless the ADUs are marked so as to prevent it. The BP transforms ADUs into one or more protocol data units called bundles that are then forwarded by DTN nodes. Applications send ADUs destined for an EID, and might arrange for delivery of ADUs sent to a particular EID, a so-called *registration*, which a DTN node can maintain persistently. This allows application registration information to survive application and operating system restarts.

Bundles and Fragments

Bundles contain a *primary block* and one or more other *blocks* of data. The primary block contains basic information, such as the destination EID, which is required for bundle routing and forwarding. Each block can contain either application data or other information used to deliver the containing bundle to its destination. Blocks hold information typically found in the header or payload portion of protocol data units in other protocol architectures. The term "block" is used instead of "header" because blocks might not appear at the beginning of a bundle due to particular processing requirements. Bundles can be fragmented into multiple constituent bundles during transmission, and these fragments can be further fragmented. Two or more fragments might in principle be reassembled anywhere in the network, again forming a new bundle.

End-Point Identifiers

As mentioned, bundle sources and destinations are identified via variable-length EIDs, which identify the original sender and bundles' final destinations, respectively. Bundles might also contain a "report-to"

EID, used when a DTN node requests special operations to direct diagnostic output to an arbitrary entity. A single EID might refer to one or more DTN nodes, which can be members of groups called *DTN end points*. A DTN end point is therefore a set of DTN nodes.

EIDs need not be related to routing or topological organization. An EID is a name, expressed using the general URI syntax that identifies a DTN end point. In URI terminology, each URI begins with a scheme name followed by a series of characters constrained by the syntax defined by the scheme, and called the scheme-specific part (SSP). The DTN architecture dictates that the scheme designer is responsible for defining how to interpret an EID's SSP so as to determine whether it refers to the equivalent of a unicast, multicast, or anycast set of nodes.

Binding

Binding means interpreting an EID to select a next hop to which a bundle can be forwarded toward its destination. Because the destination EID is potentially reinterpreted at each hop, binding might occur at the source, during transit, or possibly at the destination. The latter two scenarios are referred to as late binding.

Persistence

Bundles must wait in place in a queue until a communication opportunity is available. DTN nodes generally use some form of persistent storage, and stored bundles survive system restarts. *Persistence* assumes that storage is available, well-distributed throughout the network, and sufficiently robust to store bundles until forwarding can occur.

Bundle Routing

In DTNs, information for making scheduling and path selection decisions is based on the requested data transfers' size and performance requirements. To enable such decisions, bundles contain an originating time stamp, useful life indicator, and a class of service designator. We can consider a bundle successfully delivered to an EID when some minimum subset of the nodes, called the EID's *minimum reception group* (MRG), has received the bundle without error. An end point's MRG might refer to one node, one of a group of nodes, or the entire group of nodes, and a single node might be in multiple end points' MRGs. We assume a node can determine the MRG of the DTN end point named by an EID and that each node must have at least one EID that uniquely identifies it.

Routing schemes developed for DTN use various mechanisms, including packet replication, discovery of the meeting probabilities among nodes, and network coding. A DTN node might make forwarding decisions using measurements based on, for example, the known state of other DTN nodes, information on resource utilization, and the probability of an encounter. "Store, carry, and forward" routing schemes also use information on node contacts, location, and future movement. A DTN node must in general base such decisions on locally held information, and might constantly reassess forwarding decisions as contacts come and go — for example, to determine the best next-hop DTN node, time to forward, and highest delivery probabilities for each bundle, and to remove failed paths. The DTN architecture allows for the use of many different routing schemes, each of which might prove to be advantageous depending on circumstances. However, contact information must be known or discovered by a DTN node to form the basis for routing.

LTP

LTP[4–6] is designed to serve as a reliable convergence layer for BP over single-hop, high-latency (for example, deep-space) links. Long RTTs imply a substantial delay between the transmission of a block and the reception of an acknowledgment from the block's destination signaling its arrival. Unlike TCP, LTP sessions are unidirectional, so LTP peers can only achieve bidirectional data flow using two unidirectional links. To support scheduled communications, we might think of LTP as operating at a separate layer that knows the network state and uses lower-layer cues to tell each node when and how much to transmit. LTP does *A*utomatic *R*epeat Re*q*uest (ARQ) of data transmissions by soliciting selective-acknowledgment reception reports.

To avoid underutilizing expensive links, LTP doesn't postpone transmission until it receives acknowledgment that all prior blocks have arrived, but it allows multiple parallel data block transmission "sessions" to be in progress concurrently. Although LTP is principally aimed at supporting "long-haul" reliable transmission in interplanetary space as a convergence layer for BP, it's also been used in terrestrial environments such as in the Sensor Networking with Delay Tolerance project (SeNDT; http://down.dsg.cs.tcd.ie/sendt/).

DTN Adoption

Several wireless sensor networks have deployed DTN, and many other DTN deployments are described elsewhere,[12] including using DTN for underwater acoustic networking, meteorological and animal tracking, and various other sensor networks. Since 2003, DARPA has had a DTN program with the aim to develop and field network services that deliver critical information reliably even when no end-to-end path exists through the network. DARPA based phases one and two of its program on the Spindle (*S*urvivable *P*olicy-*In*fluenced *N*etworking: *D*isruption-*T*olerance through

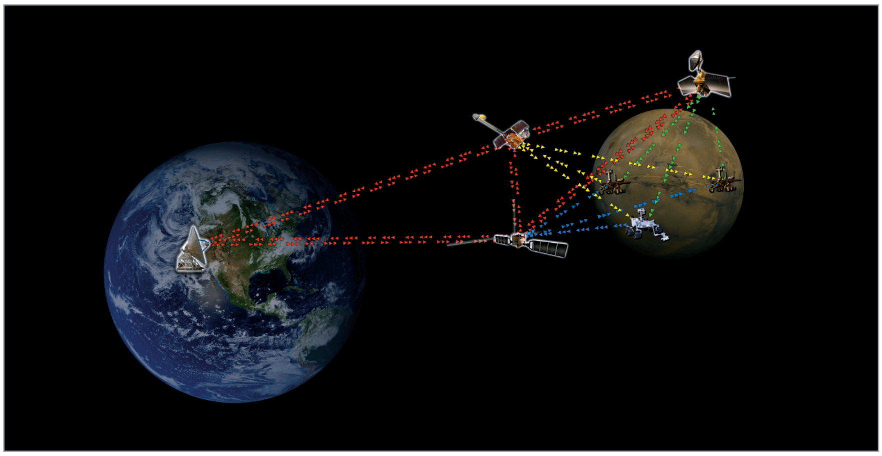

Figure 2. The Deep Impact Network (DINET) experiment. (Source: NASA/JPL-Caltech; used with permission.)

*L*earning and *E*volution) project led by BBN Technologies. The third phase of DARPA's DTN program aims to create the first "fieldable" equipment that uses DTN to access military tactical information.

The Technology and Infrastructure for Developing Regions (TIER; http://tier.cs.berkeley.edu/wiki/Home) project aims to address challenges in bringing the IT revolution to the masses in developing world regions. TIER, a research group at the University of California, Berkeley, has key projects in educational tools, healthcare, wireless (WiLDNet), distributed storage (TierStore), and speech technologies. DakNet was an early DTN project the MIT Media Lab developed, with, apparently, some commercialization from First Mile Solutions (see www.firstmilesolutions.com/documents/DakNet_IEEE_Computer.pdf). It was one of the first instances to use scheduled transport (data mules) to carry bundles between Wi-Fi equipped "kiosks" in villages on a regular basis. The EU's Seventh Framework Network for Communication Challenged Communities (N4C; www.n4c.eu) project also uses data mules and is based on previous work from the Sámi Network Connectivity (SNC; www.snc.sapmi.net) project. These projects use opportunistic encounters with data mules in the Swedish Arctic (for the most part helicopters, but also possibly snowmobiles or hikers) to transfer bundles between the temporary camps of the Sámi people, and the Internet.

Researchers are also investigating DTN in the context of disaster and emergency network support. The Multimedia and Mobile Communications Laboratory (MMLAB) at Seoul National University has been investigating its Architecture for Intelligent Emergency DTN using extensive temporary wireless communications.

And, not forgetting the origins of DTN, in 2008, NASA JPL conducted experiments simulating communications with rovers on the surface of Mars relayed through a DTN bundle agent installed on the Epoxi spacecraft, previously known as Deep Impact. As far as we know, this experiment set a distance record for RFC-compliant protocols with BP and LTP being used over 25 million km hops, for a total round trip of 50 million km! Figure 2 shows the set up for those DINET (*D*eep *I*mpact *Net*work) experiments.

Recently, the Consultative Committee on Space Data Systems (CCSDS; www.ccsds.org) has started a DTN working group that's examining the suitability of the BP and LTP experimental RFCs for use as CCSDS standards for future space missions. Both NASA and the European Space Agency (ESA) have several DTN-related activities under way that are feeding into this standardization process.

DTN Resources

As part of his PhD thesis work, Mike Demmer (then at UC Berkeley), developed an open source reference implementation of BP called "DTN2," which we at Trinity College Dublin currently help to maintain as part of our work on N4C. DTN2 provides a fairly complete DTN and BP software suite and has been used in many experimental DTN deployments.

The Interplanetary Overlay Network (ION) is another open source DTN implementation, developed by JPL and currently maintained by Ohio University (https://ion.ocp.ohiou.edu/index.php). ION consists of a completely separate BP implementation, contact graph routing, LTP, and NASA's Asynchronous Message Service (AMS), which provides an application-layer framework for using DTN. Various other implementations of the BP and LTP are linked from the DTNRG Web site at www.dtnrg.org/code/.

The DTNRG maintains an open mailing list (see http://mailman.dtnrg.org/mailman/listinfo/dtn-interest/) for general discussions, and a specific list exists for developers and users of the DTN2 reference implementation (http://mailman.dtnrg.org/mailman/listinfo/dtn-users/).

Recent discussions within DTNRG are considering whether aspects of the DTN technology have sufficient (commercial) backing and applicability to warrant creating an IETF working group to produce Internet standards for DTN. In the coming year or 18 months, we'll see whether this standardization aspect comes to fruition.

References

1. V. Cerf and R. Kahn, "A Protocol for Packet Network Intercommunication," *IEEE Trans. Communications*, vol. 22, no. 5, May 1974, pp. 637–648.
2. V. Cerf et al., *Delay-Tolerant Networking Architecture*, IETF RFC 4838, Apr. 2007; www.rfc-editor.org/rfc/rfc4838.txt.
3. K. Scott and S. Burleigh, *Bundle Protocol Specification*, IETF RFC 5050, Nov. 2007; www.rfc-editor.org/rfc/rfc5050.txt.
4. S. Burleigh, M. Ramadas, and S. Farrell, *Licklider Transmission Protocol — Motivation*, IETF RFC 5325, Sept. 2008; www.rfceditor.org/rfc/rfc5325.txt.
5. M. Ramadas, S. Burleigh, and S. Farrell, *Licklider Transmission Protocol — Specification*, IETF RFC 5326, Sept. 2008; www.rfceditor.org/rfc/rfc5326.txt.
6. S. Farrell, M. Ramadas, and S. Burleigh, *Licklider Transmission Protocol — Security Extensions*, IETF RFC 5327, Sept. 2008; www.rfc-editor.org/rfc/rfc5327.txt.
7. M. Demmer and K. Fall, "DTLSR: Delay Tolerant Routing for Developing Regions," *Proc. 2007 Workshop on Networked Systems for Developing Regions*, 2007, pp. 1–6; doi: 10.1145/1326571.1326579.
8. S. Burleigh, "Interplanetary Overlay Network (ION) Design and Operation v1.8," Feb. 2008, pp. 41 et seq.; https://ion.ocp.ohiou.edu/ION.pdf.
9. A. Balasubramanian, B. Levine, and A. Venkataramani, "DTN Routing as a Resource Allocation Problem," *SIGCOMM Computer Communications Rev.*, vol. 37, no. 4, 2007, pp. 373–384; http://doi.acm.org/10.1145/1282427.1282422.
10. K. Fall and S. Farrell, "DTN: An Architectural Retrospective," *IEEE J. Selected Areas in Communications*, vol. 26, no. 5, 2008, pp. 828–836; doi: 10.1109/JSAC.2008.080609.
11. T. Berners-Lee, R. Fielding, and L. Masinter, *Uniform Resource Identifier (URI): Generic Syntax*, IETF RFC 3986, Jan. 2005; www.rfc-editor.org/rfc/rfc3986.txt.
12. S. Farrell and V. Cahill, *Delay and Disruption Tolerant Networking,* Artech House Publishers, 2006.

Alex McMahon is undertaking PhD research within the Distributed Systems Group in the Department of Computer Science at Trinity College Dublin. His research is in the area of delay/disruption-tolerant networking. McMahon has an MPhil from the Department of Electronic & Electrical Engineering, Trinity College Dublin. Contact him at alex.mcmahon@cs.tcd.ie.

Stephen Farrell is a research fellow at Trinity College Dublin and chief technologist with NewBay Software. His research interests include security and delay/disruption-tolerant networking. Farrell has a PhD in computer science from Trinity College Dublin. He is a chair of the DTNRG, and an *IEEE Internet Computing* board member. Contact him at stephen.farrell@cs.tcd.ie.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.