Data Transfer Node (DTN) Workflows

Joseph Hill, Gerben van Malenstein, Cees de Laat, Paola Grosso, Leon Gommans

Why Data Transfer Nodes (DTNs)

- DTNs can act as an interface to a high performance link

- Configured to maximize performance for a given workflow

- Simplifies configuration of client systems

- Multiple clients may share a DTN

- DTNs strategically placed to best benefit clients

- DTNs can be compared to specialized high speed transport systems of the past

Pneumatic Tube Messaging System, 1943

United States Library of Congress's Prints and Photographs division (Digital ID fsa.8d30861)

Example: Entry Point for High Speed Transport

A typical use case for DTNs is as a high speed file transfer service. A computer system's configuration may allow for the utilization of all available bandwidth in a LAN environment. However, it is often the case that in a WAN environment with high latency or packet loss the same system performs poorly. A DTN could be tuned to maximize performance on a high latency path. It could also use specialized transfer protocols to mitigate high packet loss. The DTN may also have access to an optimized path such as a light path. Files destined for a distant receiver would be first sent to a DTN located on the same LAN as the sender. That DTN would then forward it at high speed to a DTN near the receiver. That DTN would then forward it to the final destination.
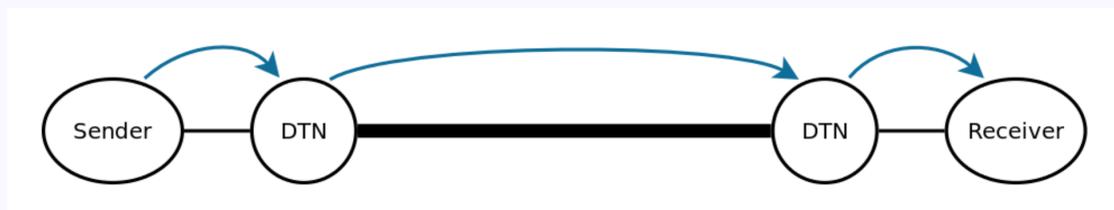
Example: Storage Access Point

Another possible use case for DTNs is to be used to access distributed data from remote locations. In this scenario a system located at a compute facility requests the data from the local DTN as it is required. That DTN would then transparently retrieve the data from multiple remote sites as needed. In contrast to the first example here block level access is provided by the DTNs. To the system performing the computations the nearby DTN appears to be the actual and only storage system. This hides both the remote and distributed nature of the data. While the compute side DTN may perform some caching, there need not be permanent storage of data at the compute facility.

UNIVERSITY OF AMSTERDAM   SURF NET   AIR FRANCE KLM