

WIDE AREA TESTING OF CD-1.1 RELIABLE MULTICASTING

J. Roger Bowman, Ray Cordova, Warren Fox, and Dale Lovell

Science Applications International Corporation

Sponsored by Army Space and Missile Defense Command

Contract No. DASG60-03-C-0009

ABSTRACT

New software for reliable multicasting of sensor data has been successfully tested over a wide-area network between an International Monitoring System (IMS) infrasound station in Newport, Washington (I56US) and the Science Applications International Corporation (SAIC) testbeds in San Diego, California and Arlington, Virginia, and an Air Force Technical Applications (AFTAC) testbed in Satellite Beach, Florida. Multicast communication is a group communication mechanism that provides one-to-many (or many-to-many) communication. Multicasting is not an inherently reliable communication mechanism, so applications requiring reliable data transmission must provide their own reliability mechanism. The reliable multicasting application software uses the formats and protocols for continuous data (CD-1.1). CD-1.1 was designed for reliable and secure transmission of data from data providers, which are generally stations of the IMS or other sensor networks, to data consumers, which are generally data centers. The initial, unicast software implementation of the CD-1.1 formats and protocols was designed for point-to-point transmission and relied on TCP/IP (Transmission Control Protocol/Internet Protocol) for data transmission and data packet reliability, ordering and flow control. The unicast version of the Continuous Data Subsystem (CDS) CD-1.1 has proven to be reliable and secure for transmission of data. It is currently in use transmitting data from five data providers distributed globally to three data consumers.

IP multicasting can overcome two inherent drawbacks of unicast data transmission. Unicasting is inefficient for transmission of data to multiple data consumers, because data packets are duplicated for transmission to each consumer. In contrast, multicasting simultaneously transmits data packets to multiple data consumers using multicasting features of network routers to duplicate packets only when necessary to service all members of a multicast group. The reliability of the unicast implementation is limited by dependence on intermediate data forwarding sites, which are potential single points of failure. In contrast the CD-1.1 reliable multicasting system multicasts data packets to all data consumers simultaneously and relies on packet retransmission directly from the data provider to all data consumers to provide the principal reliability service. Thus, under both normal and somewhat degraded conditions, CD-1.1 with reliable multicasting provides a stream of data to all data consumers independent of any intermediate forwarding center. In addition, the CDS CD-1.1 with reliable multicasting provides an additional level of reliability with a unicast catch-up mechanism.

We implemented a multicast version of the CDS CD-1.1 application software. This version could not rely on the services of TCP/IP, which is strictly a point-to-point protocol, and therefore uses the Universal Datagram Protocol (UDP) over IP. Services such as data packet-level reliability, ordering and flow control were incorporated in the application software.

We tested the CD-1.1 reliable multicasting system in three phases progressing from a local testbed environment toward a configuration resembling the target environment. The first phase was conducted using a simulated data provider on SAIC's San Diego testbed, with data consumers on the San Diego testbed and a testbed in Arlington. In the second phase, the I56US infrasound station in Newport was used as the data provider with data consumers in San Diego and Arlington, while the third phase added a data consumer on a testbed at AFTAC. The *mrouted* application was used to tunnel multicast packets through the public Internet, which did not have routers configured for our multicast use. Some of the capabilities tested were: basic multicast transmission, recovery from multicast connectivity failure, acceptable multicast layer network usage, change of multicast group address, gap notification and unicast catch-up, and acceptable data timeliness. The CD-1.1 reliable multicasting system performed well during testing, providing a reliable, timely, and secure stream of data to data consumers. The wide-area tests demonstrated the feasibility of using this system in an operational setting to address US requirements for data reliability and security. Several minor anomalies identified during testing were resolved with software or configuration changes or were logged in a defect database for future action.

OBJECTIVE

The objective of this research is to develop and demonstrate prototype software for multicasting sensor data simultaneously from stations of the US Atomic Energy Detection System (AEDS) and the International Monitoring System (IMS) to more than one data consumer (center). An important application of this research would be to transmit data from stations to both the U.S. National Data Center (NDC) and the alternate NDC, which would simplify the transfer of the test-monitoring mission between the two data centers.

IP multicasting can overcome two inherent drawbacks of unicast data transmission. First, unicasting is inefficient for transmission of data to multiple data consumers, because packets are duplicated for transmission to each consumer. In contrast, multicasting simultaneously transmits data packets to multiple data consumers using multicasting features of network routers to duplicate packets only when necessary to service all members of a multicast group. Second, the reliability of the unicast implementation is limited by dependence on intermediate data forwarding sites, which are potential single points of failure. In contrast the CD-1.1 reliable multicasting system multicasts data packets to all data consumers simultaneously and relies on packet retransmission directly from the data provider via multicast to all data consumers to provide the principal reliability service. Thus, under both normal and somewhat degraded conditions, CD-1.1 with reliable multicasting provides a stream of data to all data consumers independent of any intermediate forwarding center.

RESEARCH ACCOMPLISHED

Introduction

Multicast communication is a group communication mechanism that provides one-to-many or many-to-many communication. For transmission of data from sensors to data centers, only the one-to-many capability of multicast is utilized. Figure 1 shows unicast and multicast data flows, contrasting the duplication of data for all paths between a data provider and data consumers in the unicast case, and the lack of duplication until the last multicast router before two members of a

multicast group in the multicast case. Multicasting is not

inherently a reliable

communication mechanism, so

applications requiring reliable

data transmission must provide

their own reliability mechanism.

SAIC has designed, developed

and tested prototype software

for reliably multicasting sensor

data from monitoring stations to

data centers. The reliable

multicasting application

software uses the formats and

protocols for continuous data

(CD-1.1) (SAIC, 2003b).

Multicasting required minor extensions to the CD-1.1 formats and protocols that were originally defined for unicasting (SAIC, 2001; IDC, 2002). The application software that implements the CD-1.1 formats and protocols is called the Continuous Data Subsystem (CDS) CD-1.1. The CDS CD-1.1 for multicasting is an extension of the CDS CD-1.1 for unicasting.

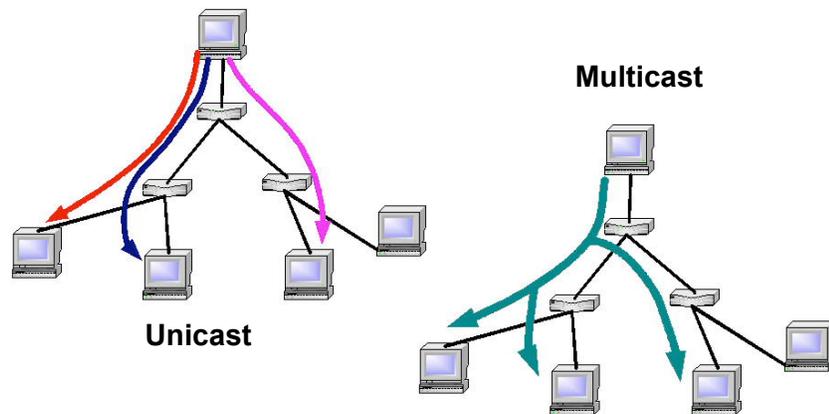


Figure 1. Comparison of unicast and multicast data flows.

The CD-1.1 formats and protocols were designed for reliable and secure transmission of data from data providers, which are generally stations of the US AEDS, IMS or other sensor networks, to data consumers, which are generally data centers. The initial, unicast software implementation of the CD-1.1 formats and protocols was designed for point-to-point transmission and relied on TCP/IP (Transmission Control Protocol/ Internet Protocol) for data transmission and data packet reliability, ordering and flow control. The unicast version of the CDS CD-1.1 has proven to be reliable and secure for transmission of data. It is currently in use for transmitting data from five data providers (e.g., AKASG, I56US, and I57US) and for receiving data at three data consumers (US NDC, SAIC-

Arlington, and the International Data Center). Independent and reliable delivery to multiple data consumers makes CD-1.1 reliable multicasting well suited for transmitting data to both the U.S. NDC and the alternate U.S. NDC.

Different multicast data transmission modes provide distinct levels of reliability, as illustrated in Table 1. Application-level Error Correction (AEC) is a critical reliability mechanism, determining which data have been sent and received, and retransmitting missing data as necessary (Saltzer et al., 1984). CD-1.1 reliable multicasting has Application-level Error Correction for both data packets, which are fragmented from CD-1.1 Data Frames, and for Data Frames.

Table 1: Services Provided by Various Modes of Multicast Data Transmission

Transmission Mode	Definition	Service
Multicast	Group delivery mechanism.	Best-effort delivery.
Reliable multicast	Group delivery with packet-level AEC.	Reliable delivery of data packets absent major failure modes (better best effort).
CD-1.1 reliable multicast	Group delivery with packet-level AEC and individual delivery with frame-level AEC.	Reliable delivery of Data Frames, even with major failure modes.

Software Design and Development

We designed and implemented a multicast version of the CDS CD-1.1 application software (SAIC, 2003a). This version of the CDS could not rely on services of TCP/IP, which is strictly a point-to-point protocol, and therefore uses the Universal Datagram Protocol (UDP) over IP. Services such as data packet-level reliability, ordering and flow control were incorporated in the application software. The design was constrained by several factors. First, we applied all CD-1.1 unicast requirements to reliable multicasting, most notably the requirement for 99.99% reliability of transport of Data Frames. Second, we minimized changes to the CD-1.1 formats and protocols (SAIC, 2003b) to preserve compatibility with unicast implementations. Third, we designed CDS CD-1.1 software components to operate in either multicast or unicast mode, according to parameter settings.

Multicasting functions used in normal operation have been separated from unicast functions used during catchup operation in two largely independent subsystems. Normal operation has at most short outages and modest network congestion. During normal operation, data are multicast to all data consumers, as illustrated in Figure 2. Data Frames are the minimum unit of data transmitted from a station and generally contain 10-30 s of data from all sites

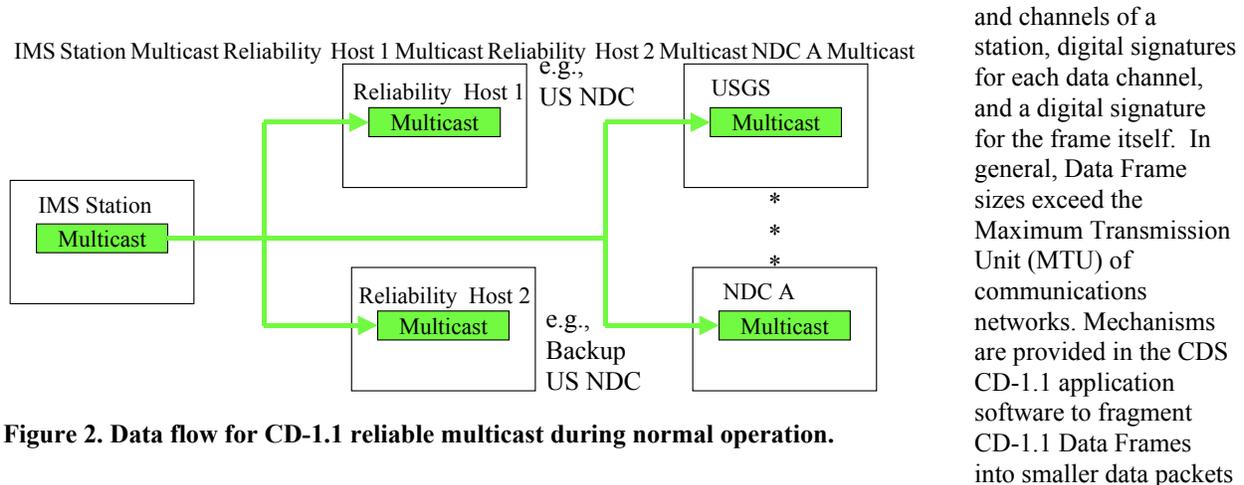


Figure 2. Data flow for CD-1.1 reliable multicast during normal operation.

that can be transmitted using UDP, to control the packet transmission rate, to re-assemble packets into Data Frames upon receipt, to identify and request re-transmission of missing data packets, and to re-multicast missing data packets.

Catch-up operation occurs after long outages or periods of high network congestion. Long outages are those exceeding the configurable, but relatively short, buffer of the multicast subsystem. Long outages might result from failure of hardware or communications circuits at a station or a data consumer. The design of the catchup subsystem employs reliability hosts, one of which is collocated with a station and the other(s) located at one or more locations of convenience. The reliability host provides re-transmission of Data Frames after those outages exceeding the multicast subsystem's buffer. During catchup operation, illustrated in Figure 3, point-to-point connections are first established between a data consumer and a reliability host. The data consumer requests Data Frames that were missed during the outage and the reliability host unicasts these Data Frames to the requesting data consumer.

Catchup processing uses the frame-level reliability mechanisms of CD-1.1 and provides a higher level of reliability than can be provided by the packet-level reliability mechanisms of reliable multicasting. Note that it is not practical for all data consumers to interact directly with the reliability host at a station. This approach would not scale well to multiple data consumers

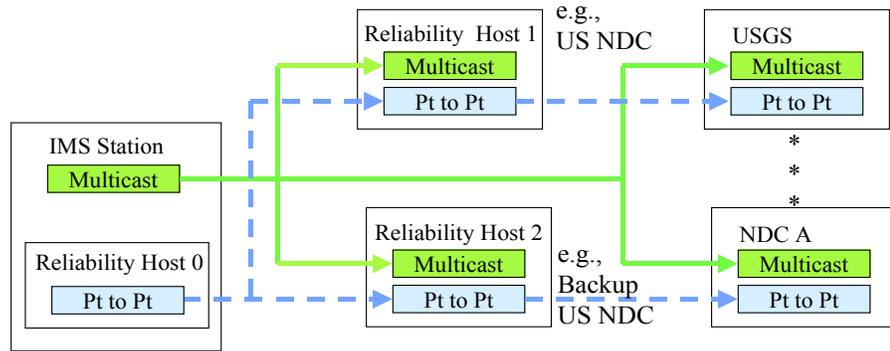


Figure 3. Data flow for CD-1.1 reliable multicast during catchup operation after an outage.

in circumstances in which the bandwidth of the communications link to the station is limited. In case of an outage of the station or the tail circuit to the station, which is often vulnerable to outages, all data consumers would request their missing data at a similar time, resulting in contention for the limited available bandwidth. Our design includes one or more reliability hosts remote from the station at sites with broad bandwidth communication links and optionally with communications and other infrastructure that is more reliable than that for an individual station. Key design features of the multicasting subsystem of the CD-1.1 reliable multicast software are (SAIC 2003a):

- Allows up to 20 data consumers in a single multicast group.
- A data provider is not limited in its sending by the absence or presence of any specific data consumer.
- A support increases and decreases in the size of the multicast group without the need to restart the sending activity of the data provider.
- A data consumer can enter or leave a multicast group without negatively impacting other group members.
- Transmission rates are constant and configurable to mitigate network congestion.
- The size of multicast data packets is configurable to support small MTU (Maximum Transmission Unit) networks (frames broken into packets).
- Missing multicast data packets of an active multicast session can be identified, requested, and resent, availability limited by configurable data buffer size.
- A data provider transmits frames in the order in which they were written to its CD-1.1 frame set, (except retransmissions). A multicast data provider can begin delivery at the current time or a configurable lookback time less than 10 minutes prior to the current time to prevent gaps in the stream caused by short data provider outages.

Key design features of the unicast catchup subsystem of the CD-1.1 reliable multicast software are (SAIC 2003a):

- Uses sequence numbers to detect Data Frames not received by a data consumer via multicast.
- Provides reliable delivery to a data consumer of frames not received via multicast.

The CD-1.1 reliable multicast software was written in C and C++. Maximum reuse was made of existing components of the Continuous Data Subsystem CD-1.1.

Testing Configurations

We followed the test plan for the CD-1.1 reliable multicasting system of Lovell and Bowman (2003). This plan defined configurations for hardware, communications, and application software, and defined test cases for four phases (Table 2). The first phase of testing was local area network testing in San Diego, California. Wide-area testing of the system consisted of three phases progressing toward a configuration resembling the target environment (Table 2). The first phase of wide-area testing (Phase 2) was conducted using a simulated data provider on SAIC’s San Diego testbed, with data consumers on the San Diego testbed and a testbed in Arlington, Virginia. In the second phase of wide-area testing (Phase 3), the I56US infrasound station in Newport was used as the data provider with data consumers in San Diego and Arlington, while the third phase of wide-area testing (Phase 4) added a data consumer on a testbed at AFTAC in Satellite Beach, Florida. The IP multicast routing daemon (*mrouted*) application (Deering, 1989, Xerox, 2003) was used to tunnel multicast packets through the public Internet, which does not have routers configured for our multicast use.

Table 2: Test Phases for CD-1.1 Reliable Multicasting

Phase	Title	Data Provider	Network	Data Consumers
1	Local Area Network Testing on the San Diego Testbed	San Diego testbed (synthetic source)	LAN	San Diego testbed (two Consumers)
2	San Diego to Arlington Wide-Area Testing	San Diego testbed (synthetic source)	Internet	Arlington testbed; San Diego testbed
3	I56US to Arlington and San Diego Wide-Area Testing	I56US	Internet	Arlington testbed; San Diego testbed
4	I56US to AFTAC, Arlington, and San Diego Wide-Area Testing	I56US	Internet	AFTAC, Arlington testbed; San Diego testbed

Local-Area Tests

We extensively tested the CD-1.1 reliable multicast application software during Phase 1 using a Local Area Network on a testbed in San Diego. Because this paper focuses on wide-area testing, here we only enumerate the test results achieved during Phase 1. The test cases demonstrated:

- The ability to establish and maintain a multicast group (a data provider and one or more data consumers).
- The ability to do an orderly, commanded shutdown of a multicast group via an Alert Frame sent by the data provider.
- The ability to receive and store CD-1.1 multicast data Frames at a data consumer.
- That the multicast software enforces the configured upper limit on network bandwidth usage.
- The ability to request re-transmission of a packet.
- That the addition of a new data consumer to a multicasting group does not alter the ability to provide a flow of data packets to existing group members.
- That the data Frame auditing process can determine when unicast catchup needs to take place.
- That data Frame auditing produces an accurate list of frame sequence number gaps.
- That a unicast catchup connection can be established.
- That during a unicast catchup connection, frames specified by a requesting data consumer are sent to that data consumer.
- That reliable multicast makes efficient use of host computer CPU and memory.
- That reliable multicast occurs within reasonable transport network utilization constraints.
- Timely delivery of data to data consumers.
- Complete delivery of data to data consumers.
- The ability to recover from a data provider reboot.
- The ability to recover from a data consumer reboot.

Wide-Area Tests

Figure 4 shows the topology of the wide-area tests of CD-1.1 reliable multicasting. Phase 2 used a simulated data provider on SAIC’s San Diego testbed and data consumers on SAIC’s Arlington testbed as well as on the San Diego testbed. Phase 3 used an IMS station in Newport, Washington as a data provider and used data consumers San Diego and Arlington. Phase 4 added a data consumer at AFTAC.

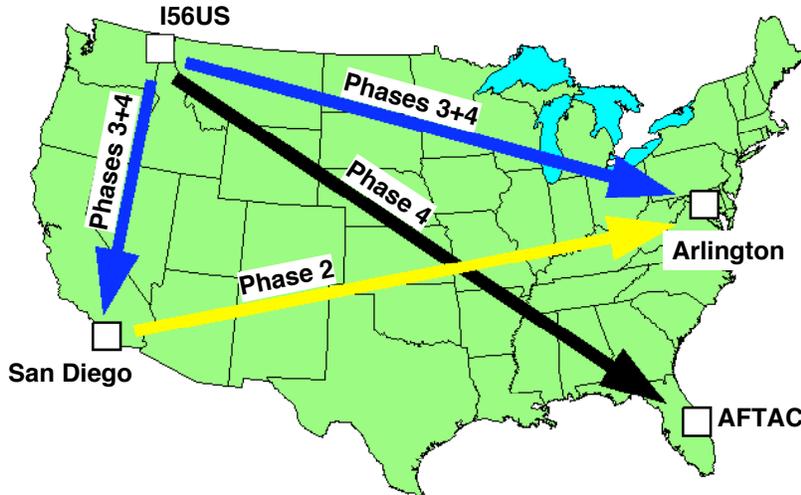


Figure 4. Map of data providers and data consumers used in wide-area testing.

Table 3 summarizes test results for Phases 2 and 3. Several minor anomalies identified during testing were resolved with software or configuration changes during testing or were logged in a defect database for future action. For example, heavy load on the link from I56US caused a high rate of packet loss, which was resolved by adjusting transmission control parameters. Initially, firewall and router network address translation at I56US prevented correct connection for unicast catchup. This was resolved by modifying the application software to specify the external IP address. Detailed test results are reported by Cordova and Lovell (2003).

Table 3: Results of Phase 2 and 3 testing.

Test Case	Purpose	Expected Result	Actual Result
Multicast data transmission	Demonstrate that Data Frames are successfully transmitted and stored.	Data Frames will be successfully transmitted to data consumer and stored in Frame Store.	As expected.
Gap notification and unicast catchup	Demonstrate that catchup connection is established if gaps are detected and that missing Data Frames are supplied.	Data Frames originally listed as missing will be stored in Frame Store.	As expected.
Change of multicast group address	Demonstrate that a data consumer can change group address at which it receives data following change of address by data provider.	Data will be successfully transmitted to data consumers using new multicast address and destination port.	As expected.
Multicast layer network usage	Determine bandwidth used in transmission of data from data providers to data consumers.	Average throughput will be less than configured value.	As expected.
Data timeliness	Demonstrate that average data latency is within specifications.	Maximum latency for test period will not exceed 3 minutes.	As expected.
Multicasting connectivity fails	Demonstrate proper behavior when multicast connectivity fails.	Multicast data transmission should stop and catchup processing should fill in gap when connectivity is restored.	As expected.

Figure 5 shows the configuration for Phase 4 testing. A testbed was established at AFTAC during June 2003. Hardware was designated for testing, and the firewall and application software were configured. Phase 4 testing will be performed during July and August using a subset of test procedures from Phase 3 testing. Therefore, we expect the testing to complete the successful demonstration of the capabilities of CD-1.1 reliable multicasting.

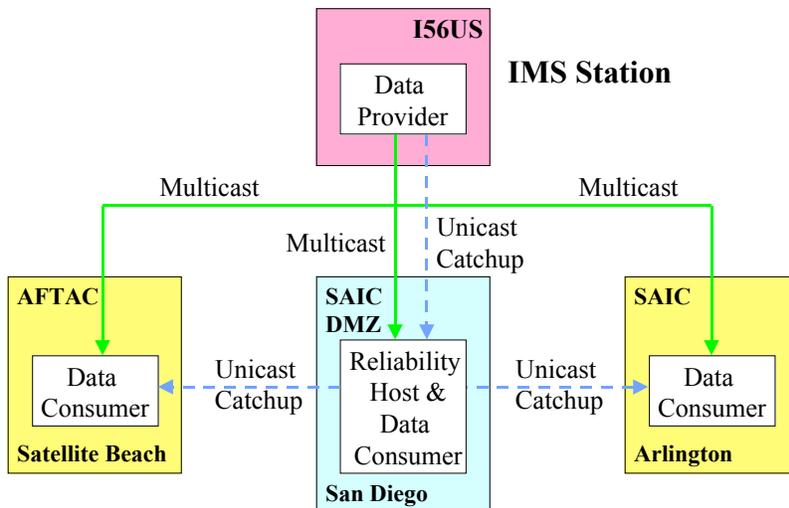


Figure 5. Configuration used in Phase 4 wide-area multicast tests.

CONCLUSIONS AND RECOMMENDATIONS

Wide-area tests of CD-1.1 reliable multicasting have demonstrated the feasibility of using a multicast system in an operational setting to address US requirements for data reliability and security. We recommend that CD-1.1 reliable multicasting be completed as an operational product and then be deployed at selected stations, at the US National Data Center, and at the alternate National Data Center.

ACKNOWLEDGEMENTS

We thank Jim Burns and Brian Coan of Telcordia Technologies for valuable suggestions on the design and testing of the CD-1.1 reliable multicast software, Jon Berger, Michael Hedlin and Clint Coon of UCSD for use of the I56US infrasound station as a multicast data provider, Victor Oancea of SAIC for supporting Arlington tests, Bruce Varnum, Dan Harvala and Richard Baumstark of AFTAC for coordinating wide area tests that included AFTAC, and Deb Agarwal of Lawrence Berkeley Laboratory for contributing to plans for wide-area testing.

REFERENCES

- Cordova, R. and D. Lovell (2003). Test Report for Continuous Data Subsystem CD-1.1 Reliable Multicasting, SAIC Technical Report SAIC-03/3014.
- Deering, S. (1989), IP multicast routing daemon, Stanford University.
- IDC (2002), IDC Software documentation Formats and Protocols for Continuous Data CD-1.1 Revision 0.3, IDC3.4.3Rev. 0.3.
- Lovell, D. and R. Bowman (2003), Plan for Wide Area Network Testing of Continuous Data Subsystem CD-1.1 Reliable Multicasting, SAIC Technical Report 03/3013.
- SAIC (Science Applications International Corporation) (2001), Formats and Protocols for Continuous Data CD-1.1 Revision 0.2.
- SAIC (Science Applications International Corporation) (2003a), Continuous Data Subsystem CD-1.1, Revision 1, SAIC-03/3000, 2003a.
- SAIC (Science Applications International Corporation) (2003b), Formats and Protocols for Continuous Data CD-1.1, Draft revision to SAIC-01/3027 Rev. 1.0.
- Saltzer, J.H., D.P. Reed, and D.D. Clark (1984), End-to-end arguments in system design, ACM Transaction on Computer Systems, 2, 277-288.
- Xerox Corp. (2003), <ftp://ftp.parc.xerox.com/pub/net-research/ipmulti/mrouted3.8-sparc-solaris2.tar.gz>.