東京大学地震研究所技術研究報告，No. 19，1-6 頁，2013 年.
Technical Research Report, Earthquake Research Institute, the University of Tokyo, No. 19, pp. 1-6，2013.

WIN パケット表示解析ツールの開発

中川茂樹*・鶴岡　弘*

Development of a WIN packet analyzer

Shigeki NAKAGAWA* and Hiroshi TSURUOKA*

Abstract

The WIN system is the national standard for exchanging seismic data between Japanese universities, JMA, JAMSTEC and NIED using the JDXnet, which is a wide area layer-2 network. These seismic data are packetized in the WIN format and transmitted using the UDP/IP protocol. WIN packets are rarely lost in the network, but the determination of the reasons for those few losses is necessary to maintain stable data transmission. Typically in this case, the tcpdump command is used. However, since tcpdump is a general-purpose tool, the output of tcpdump is not suitable for WIN packets. In order to facilitate WIN packet analysis, we have developed winpdump, a WIN packet analyzer program.

Key words : WIN system, packet dump, JDXnet

はじめに

日本の地震や地殻変動観測データの集配信システムでは，多くの場合，観測データを WIN フォーマット（卜部，1994）化したパケットを UDP/IP プロトコルで送受信している．この集配信においては様々な種類の伝送経路が用いられている．観測点からデータセンターへのデータ収集では，衛星通信，ISDN，ADSL，光回線，無線 LAN，携帯電話回線などが観測点の観測項目や通信環境に応じて使い分けられている．一方，観測データの配信では，防災科学技術研究所，気象庁，海洋研究開発機構，大学等による地震観測データ等をリアルタイムに全国流通させるため，超高速広域ネットワークである JGN-X や SINET4 の広域 L2 網を用いた JDXnet（Japan Data eXchange network）が運用されている（鷹野ほか，2005）．また，JGN-X や SINET4 の広域 L2 網が到達しない拠点では，NTT のフレッツ網等を用いてデータの配信を行っている．

観測機器や送受信装置，ネットワークの故障・不具合等のため，データの集配信がしばしば滞ることがある．機器類の物理的な障害であれば復旧は容易であるが，ネットワークが関係する障害の場合はその原因究明が困難である事が多い．このような場合，tcpdump や snoop 等のコマンドを利用してネットワーク中を流れるパケットを収集し，分析を行う．Wireshark のように GUI で分析も行えるツールもあるが，一般的なプロトコル（例えば，telnet，arp 等）にしか対応していない．従って，これまでは，WIN パケット用のフィルタを自作するか，作業者が自分の目で内容を解読するしかなかった．そこで，本研究では，ネットワーク中を流れる WIN パケットを収集し効率的に分析するプログラムを開発した．

2013 年 8 月 19 日受付，2013 年 12 月 19 日受理.
† nakagawa@eri.u-tokyo.ac.jp
* 東京大学地震研究所地震火山情報センター
* The Earthquake and Volcano Information Center
Earthquake Research Institute, the University of Tokyo.

WIN パケットの通信

現在，日本の大学で行われているテレメータ地震観測では，多くの場合，次の経路でデータが転送されている．まず，地震観測データは観測点のテレメータ装置で WIN パケット化して送出され，JDXnet に設置されたデータ中継装置（卜部ほか，2013）や観測機関のデータセンターに衛星通信回線や NTT のフレッツ網，携帯電話回線等を用いて伝送される．その後，データは JDXnet で全国にブロードキャスト配信され，大学等の各機関で受信･収録される．

データは，ネットワークの境界に設置されたサーバを介して，バケツリレー式に転送される．データ転送には，WIN システムの recvt と sendt_raw の組合せ，または relay/relaym 単独といったコマンドが利用される（図 1）．いずれもサーバのあるポート番号に届いた WIN パケット

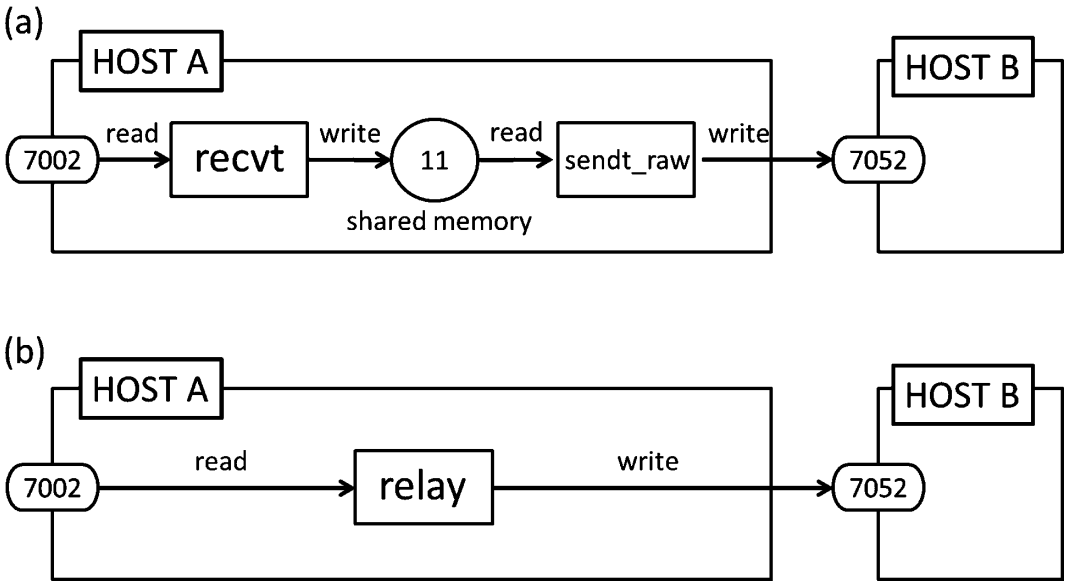

図 1. WIN パケット転送の模式図．7002 と 7052 はポート番号，11 は共有メモリ番号の例を表す．(a) recvt と sendt_raw を用いた場合．(b) relay を用いた場合．

```
[root@bbbb auto]# /usr/sbin/tcpdump -i eth0 -x -s 128 -c 5 src 172.17.15.115
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 128 bytes
18:02:03.059076 IP aaaa.eri.u-tokyo.ac.jp.55408 > bbbb.eri.u-tokyo.ac.jp.
7052: UDP, length 417
        0x0000:  4500 01bd 0000 4000 4011 c223 ac11 0f73  E.....@.@..#...s
        0x0010:  ac11 0f77 d870 1b8c 01a9 9d8c cbcb a001  ...w.p..........
        0x0020:  9e10 0116 1801 56d0 1020 c8ff f7e4 0edb  ......V.........
        0x0030:  2c46 8804 47fa 9404 73e7 e629 890f acba  ,F..G...s..)....
        0x0040:  2a10 5a30 1acc ea09 df28 f6e7 3e05 7e0c  *.Z0.....(..>.~.
        0x0050:  bad4 e8fa 944d 61e0 46de 523d 89db e4cc  .....Ma.F.R=....
        0x0060:  9d28 562a 15f2 6600 e400 16fd 9810 bbd8  .(V*..f.........
        0x0070:  3cfd                                     <.

●IP header
4:IP version(IPv4)
11:Protocol number (UDP)
ac110f73:source address (172.17.15.115)
ac110f77:destination address (172.17.15.119)

●UDP header
d870:source port (55408)
1b8c:destination port (7052)

●WIN header
cb:packet number (107)
a0:identification code
```

図 2. tcpdump を用いた WIN パケットの収集例．

を別のサーバのあるポート番号に向けて送り出す機能を提供する．relay/relaym はすべてのパケットを転送するのに対し，recvt と sendt_raw の組合せでは共有メモリを使用するため他ホストに送り出す WIN データのチャネルによる分割，併合が可能である．また，recvt で受信したデータは共有メモリに書き出されているため，データ処理も行うことができる．かつて，衛星通信を用いたデータ配信を行っていた（鷹野ほか，2001；卜部ほか，2001）時は，防災科学技術研究所の Hi-net，F-net（Okada et al., 2004; Obara et al., 2005）と産業技術総合研究所の一部（防災科学技術研究所経由で配信）のデータを raw_shift コマンドにより 2 ビット右シフトして各大学へ配信していた．その際，この共有メモリを用いたデータ転送の方式が使われていた．ちなみに，この 2 ビット右シフト操作が行われてい

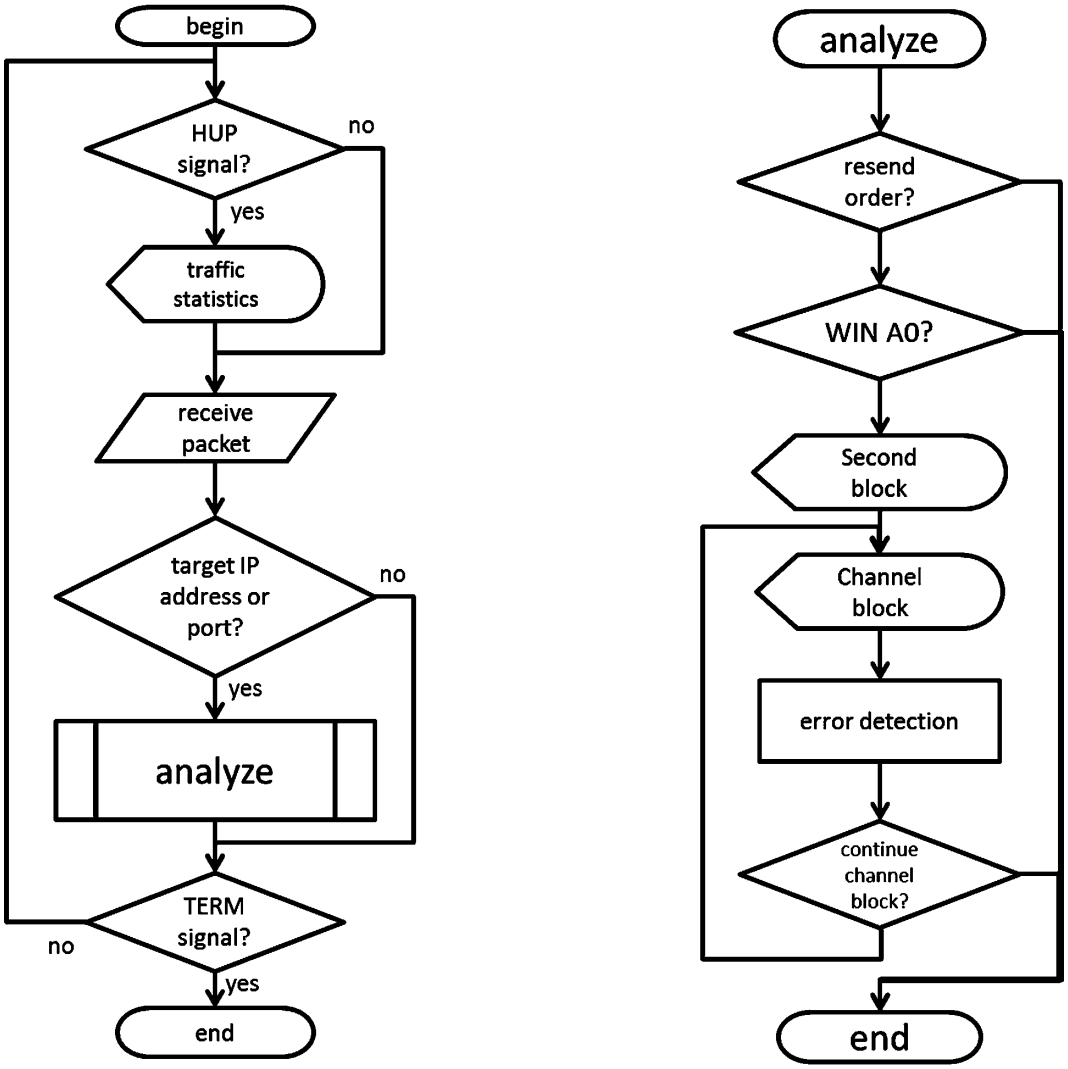

図 3. 開発したプログラムのフローチャート．左図はプログラム全体の流れを表す．右図は，左図のうち analyze の部分の流れを詳細に示している．

た理由は，当時の Hi-net 100Hz サンプリングデータはフルスケール 26.32 ビット長のデータであるが，利用していた AD 変換装置の実効ダイナミックレンジは約 135dB で 23.4 ビット相当のため，衛星配信系の帯域を節約するために 2 ビット右シフトして 24.32 ビット分を配信したからである（卜部，私信）．

さて，データ転送はネットワークを介して行われるため，データ再送要求の多発，パケット順の入れ替わり，パケットの消失，送信元が不明のデータ，WIN のヘッダやデータの異常など様々な障害が発生する．回線の輻輳が生じて全くデータが転送されなくなることもある．

障害の原因究明にあたっては，最初にデータ転送に関わるプロセスのログを確認し，ホスト間で照合することが行われる．共有メモリを介したデータ転送の場合，shmdump コマンドを用いて共有メモリ内のデータをダンプすることにより，送られているデータの中身を調べることができる．しかし，relay コマンドを使った転送を行っている場合やデータ転送を行うホストのネットワークインタフェースやスイッチに問題がある場合には，shmdump は使えず，プロセスのログからだけでは原因を特定することはできないことが多い．従って，このような場合は，tcpdump や snoop コマンドを用いてネットワークインタフェースを流れるパケットを収集し，その出力を目視により確認して原因の特定作業を行っていた．しかしながら tcpdump の出力は見やすいものではなく，かつ膨大な量のデータとなるので，この特定作業は根気と時間を要する作業であった（図 2）．

開発したプログラムの機能と特徴

開発したプログラムは，ネットワークインタフェースを

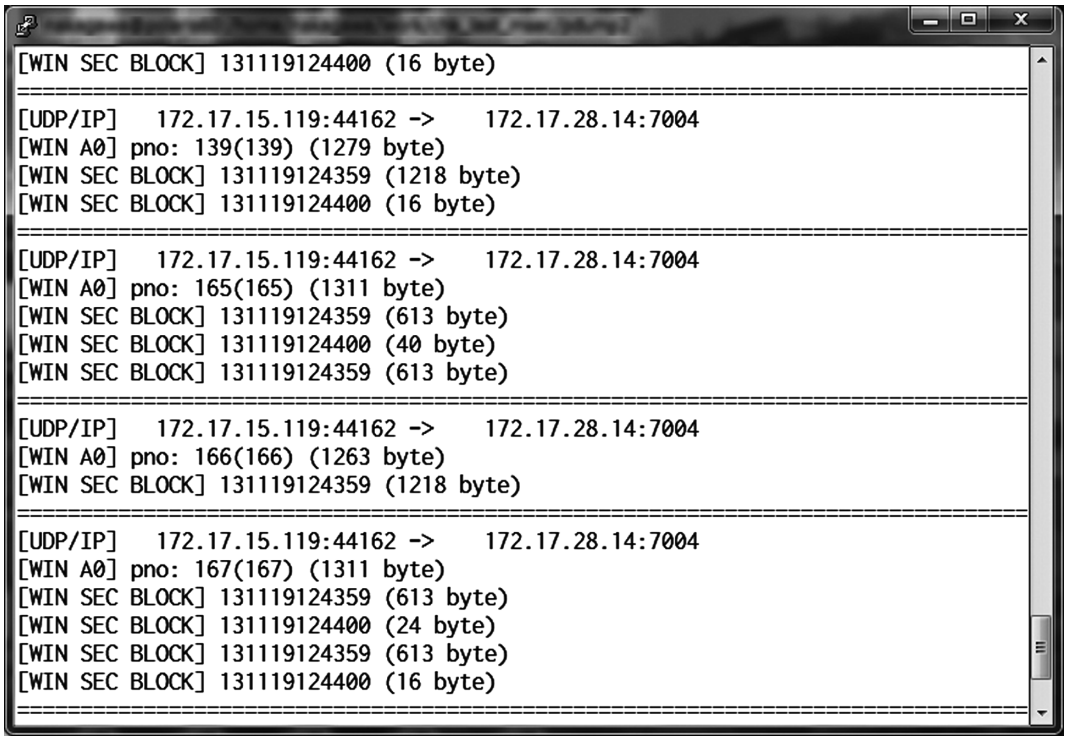

図 4.　開発したプログラムを実行した例.

無差別モードに設定して流れるすべてのパケットを収集し，その中から WIN の UDP パケットのみを抽出，分析する構造のプログラムとした．障害の分析を目的としたプログラムなので，送信元及び送信先の IP アドレスやポート番号でデータをフィルタすることも可能とした.

また，最近では，WIN システムは OS として FreeBSD と Linux を用いることがほとんどであるので，本プログラムはその 2 つの OS のみに対応することとした．これは，Solaris や Windows においてネットワークインタフェースを直接操作するプログラムは複雑になりがちで，今後の保守性等を考慮したことによる．libpcap といったパケット収集用のライブラリを利用する案も考えられたが，障害時に素早く検査できることを重視し，別プログラム等のインストールを要件としない方針としたことによる.

プログラム中でパケットを取り込む部分は，小俣(2011)，村 山（2004），Linux 日 和（2009） を 参 考 に し た．FreeBSD の BPF デバイスの使い方については，Rieck（2010）を参考にした.

本プログラム（winpdump と名付ける）の流れ図を図 3 に示す．winpdump は，まず収集したパケットを IP アドレスやポート番号によりフィルタする．その後，WIN のデータパケット（A0 パケット），制御パケット（例えば A8 パケット），再送要求パケットを自動判別する．データパケットについては，含まれる秒ブロック，チャネルブロックの中身を読み取り，秒ブロックの場合はその時刻，チャネルブロックの場合はチャネル番号とサンプリングレートを表示する．また，チャネルブロックに含まれる観測データも表示する事ができる．また，recvt や relay と同様に HUP シグナルを受け取ると，送信元と宛先のホスト，ポートの組合せ毎にそれまでに受信したパケット数や再送要求数などの流量情報を表示する.

さらに，WIN パケットの解析専用プログラムとしたことで，波形データの内容をパケット収集の時点で検査することができる（図 3 の error detection 部分）．このために，プログラム中に波形データを読むための関数 special_check( ) を用意した．この関数を任意に書き換えることにより，ユーザはその目的に応じた検査を実施することができる.

開発したプログラムの使い方

winpdump を実行した画面の例を図 4 に示す．図 4 からは，以下の事が読み取れる.

[UDP/IP] の行：ホスト 172.17.15.119 のポート 44162 番からホスト 172.17.28.14 のポート 7004 番へパケットが流れていることがわかる.

[WIN A0] の行:パケットの種類は WIN A0（データパケット）であり，パケット番号とサイズがわかる.

[WIN SEC BLOCK] の行：WIN パケットに含まれる秒ブ

ロックの時刻とサイズがわかる．1つのパケット内に複数の秒ブロックが含まれることもある．

図2に示すように tcpdump の出力を解析する場合は，パケットを構成する様々なヘッダ（IP ヘッダ, UDP ヘッダ）を先頭から順番に読み，それに続く WIN パケットの中身を調べる必要があった．複数の秒ブロックで構成されている場合は，パケットの途中に WIN の秒ヘッダが現れることになるので，作業時間を要する作業となる．

winpdump は，実行者が root であるか，またはプログラムが root に setuid されている必要がある．また，FreeBSD の場合 BPF を利用するので /dev/bpf* に読み込みの権限が設定されている必要がある．なお，FreeBSD の場合，パケットの取りこぼしを防ぐために BPF バッファを以下のように大きく設定したほうが良い（silvernetworks, 2010; Heyde, 2008; But and Bussiere, 2005）．

```
# sysctl -w net.bpf.bufsize=10485760
# sysctl -w net.bpf.maxbufsize=10485760
```

winpdump はオプションを付けずに起動すると，標準ではネットワークインタフェースの eth0 を通過する WIN パケットを表示する．しかし，障害の原因究明のためには，送信元，送信先ホストやポート番号を絞りこむことや表示する情報の種類を選択することが重要なので，次のオプションが用意されている．

-a

収集した WIN パケットをすべてダンプ表示する．-vvv と等価である．

-c detectlog

special_check( ) 関数に定義された検査を実行し，detectlog に検出結果を書き出す．detectlog が指定されない場合は，デフォルトで detectlog というファイルに書き出す．

ユーザは，ソース中の special_check( ) 関数を独自に定義することにより，様々な検査を行うことができる．

-d dstport

宛先ポート番号を指定し，フィルタする．

-D dstaddr

宛先ホストの IP アドレスを指定し，フィルタする．

-s srcport

送信元ポート番号を指定し，フィルタする．

-S srcport

送信元ホストの IP アドレスを指定し，フィルタする．

-i interface

パケットを収集するインターフェイスを指定する．指定されない場合は，eth0 が使われる．

-h

ヘルプメッセージを表示する．

-v

ダンプ表示のレベルを指定する．-v は詳しい情報，-vv はより詳しい情報，-vvv は全ての情報を表示する．-vvv と -a は等価である．

これらのオプションを適切に使って WIN パケットをダンプした出力は，障害の原因究明の一助となる．例えば，送信元が不明の WIN パケットがある場合は winpdump の出力のうち [UDP/IP] の行を確認すれば送信元を特定することができる．波形データにある障害が見られる場合には，それを検出するような関数 special_check( ) を作成した上で winpdump を実行することで，異常なデータを検出することができる．

まとめ

ネットワーク中を流れる WIN パケットを収集し分析するプログラム winpdump を開発した．これを用いることで従来 tcpdump 等の出力を分析し障害を検出する目視での作業が自動化され，障害の原因究明作業が効率よく行えるようになった．

開発したプログラムは，基本的には WIN の A0 パケットのみを対象としているが，今後はさらに改良を進め，最近一部のロガーで利用されている ACT プロトコル（森田ほか，2010）のパケットも分析，表示可能としたい．

謝辞：岩崎貴哉教授と酒井慎一准教授の査読意見は，本稿の内容を改善する上で大変有益でした．深く感謝します．

文　献

But, J. and J. Bussiere, 2005, Improving NetSniff Capture Performance on FreeBSD by Increasing the PCAP Capture Buffer Size, *CAIA Technical Report*, 051027A.

Heyde, A. A., 2008, Investigating the performance of Endace DAG monitoring hardware and Intel NICs in the context of Lawful Interception, *CAIA Technical Report*, 080222A.

Linux 日和 , 2009, RAW ソケットを使った受信 , http://linux-biyori.sakura.ne.jp/pr_rawsock.php,（参照 2013-05-06）.

森田裕一・酒井慎一・中川茂樹・笠原敬司・平田　直・鏡　弘道・加藤拓弥・佐藤峰司，2010，首都圏地震観測網（MeSO-net）のデータ伝送方式について－自律協調型データ送信手順（ACT protocol）の開発－，地震研究所彙報，**84**，89-105.

村山公保 , 2004, 基礎からわかる TCP/IP ネットワーク実験プログラミング第2版 , オーム社 , 378 頁 .

Obara, K., K. Kasahara, S. Hori and Y. Okada, 2005, A densely distributed high-sensitivity seismograph network in Japan: Hi-net by National Research Institute for Earth Science and Disaster Prevention, *Review of Scientific Instruments*, **76**, 021301-doi:10.1063/1.1854197.

Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara and A. Yamamoto, 2004, Recent progress of seismic observation networks in Japan -Hi-net, F-net, K-NET and

KiK-net-, *Earth, Planets and Space*, **56**, xv-xviii.
Rieck, B., 2010, Using FreeBSD's BPF device with C/C++, http://bastian.rieck.ru/howtos/bpf/, (参照 2013-05-10).
silvernetworks, 2010, BPF buffer, ネットワーク管理者の憂鬱な日常, http://blog.goo.ne.jp/silvernetworks/e/4aad4121b2cf6708034d8535dd3466fd, (参照 2013-05-28).
小俣光之, 2011, ルータ自作でわかるパケットの流れ, 技術評論社, 191 頁.
鷹野　澄・卜部　卓・平田　直・笠原敬司・小原一成・堀　貞喜・西出則武・若山晶彦・中澤博志・松森敏幸, 2001, 高感度地震波形データの全国リアルタイム流通システムの開発, 日本地震学会講演予稿集, No. 2, B57.
鷹野　澄・卜部　卓・鶴岡　弘・中川茂樹・三浦　哲・松澤　暢・岡田知己・中島淳一・中山貴史・平原　聡・伊藤武男・大見士朗・植平賢司・松島　健, 2005, 超高速ネットワーク JGNII によるリアルタイム地震波形データ交換システムの構築実験, 日本地震学会講演予稿集, No. 2, C098.
卜部　卓, 1994, 多チャンネル地震波形データのための共通フォーマットの提案, 日本地震学会講演予稿集, No. 2, P24.
卜部　卓・鷹野　澄・平田　直, 2001, 大学の衛星テレメータシステムにおける次の 10 年, 日本地震学会講演予稿集, No. 2, B58.
卜部　卓・鷹野　澄・鶴岡　弘・中川茂樹, 2013, JDXnet/SINET4 上に実現した観測データ中継システム, 日本地震学会講演予稿集, No. 2, D11-11.