

불안정한 네트워크 환경에서 대용량 데이터의 전송 효율화를 위한 링 버퍼에 관한 연구

송민규* · 김효령**

A Study on Ring Buffer for Efficiency of Mass Data Transmission in Unstable Network Environment

Min-Gyu Song* · Hyo-Ryoung Kim**

요 약

본 논문에서는 불안정한 네트워크 환경에서 안정적으로 대용량 데이터 스트림을 전송할 수 있는 TCP/IP 기반 링 버퍼 시스템을 설계하였다. 우리가 제안한 방식에서 각 전파천문대의 백엔드 시스템으로부터 UDP 프레임으로 생성 및 출력되는 관측 데이터 스트림은 클라이언트 시스템 내의 소켓 버퍼를 경유해 대용량 링 버퍼에 UDP 패킷으로 저장된다. 이후 목적지에 해당하는 상관센터로의 안정적인 전송을 위해 해당 패킷은 TCP로 전송되고 상관센터 내의 서버 시스템은 소켓 버퍼에 수신된 패킷에 이상이 없으면 대용량 링 버퍼에 저장한다. 패킷 손실, 중복 및 순서 역전 등의 이상이 발생할 경우 TCP의 흐름 제어를 통해 패킷은 재전송되며 상관센터에 도착한 데이터는 신뢰성을 보장받게 된다. 또한 네트워크 성능 불안정으로 인한 혼잡 회피 발생 시 병렬 스트림 적용을 통해 성능 저하가 최소화될 수 있도록 하였다.

ABSTRACT

In this paper, we designed a TCP/IP based ring buffer system that can stably transfer bulk data streams in the unstable network environments. In the scheme we proposed, The observation data stream generated and output by each radio observatory's backend system as a UDP frame is stored as a UDP packet in a large capacity ring buffer via a socket buffer in the client system. Thereafter, for stable transmission to the remote destination, the packets are processed in TCP and transmitted to the socket buffer of server system in the correlation center, which packets are stored in a large capacity ring buffer if there is no problem with the packets. In case of errors such as loss, duplication, and out of order delivery, the packets are retransmitted through TCP flow control, and we guaranteed that the reliability of data arriving at the correlation center. When congestion avoidance occurs due to network performance instability, we also suggest that performance degradation can be minimized by applying parallel streams.

키워드

Advanced Network, Parallel Stream, Ring Buffer, TCP/IP Socket, VLBI Data,
첨단망, 병렬 스트림, 링 버퍼, TCP/IP 소켓, VLBI 데이터,

* 교신저자 : 한국천문연구원 천문우주기술센터

** 한국천문연구원 전파천문본부(hrkim@kasi.re.kr)

• 접수일 : 2020. 10. 31

• 수정완료일 : 2020. 11. 23

• 게재확정일 : 2020. 12. 15

• Received : Oct. 31, 2020, Revised : Nov. 23, 2020, Accepted : Dec. 15, 2020

• Corresponding Author : Min-Gyu Song

Technology Center for Astronomy and Space Science, Korea Astronomy and
Space Science Institute,

Email : mksong@kasi.re.kr

I. 서론

초고속 네트워크의 저변 확대 및 IT 기술의 발전으로 인해 다량의 정보를 온라인으로 송수신하는 것이 일상화되었다. 과학연구 분야에서도 네트워크를 이용한 데이터 송수신 및 공유가 활발히 이뤄지고 있는 가운데 대용량 데이터의 안정적인 전송의 중요성은 더욱 높아지고 있다[1,2]. 그중에서도 VLBI는 수 ~ 수십 Gbps 수준의 방대한 관측 데이터를 네트워크를 경유해 전송하고 그에 대한 분석 결과를 얻는 분야로서 시스템 측면에서 다수의 전파천문대와 중앙에 위치한 상관센터로 구성된다. 전자가 관측 데이터의 획득, 전달 기능을 수행한다면 후자는 이를 수신 받아 분석하는 역할을 하며 데이터 처리의 효율성, 생산성을 높이기 위해 두 양단은 초고속 네트워크로 연결된다[3].

기존에 각 전파천문대에서 획득한 VLBI 데이터는 내부에 설치된 시스템에 먼저 저장된 후 FTP 등의 프로토콜을 이용해 파일 형태로 상관센터 내의 스토리지로 전송되었다[4]. 하지만 전파천문대와 상관센터 간 초고속 네트워크 환경이 구축되어 있는데 송수신 양단에서 데이터를 중복 저장하는 것은 비효율적이고, 이에 따른 연구 생산성 저하는 극복해야 할 과제 중 하나로 인식되고 있다[5].

이를 개선하기 위해 각 전파천문대에 저장된 관측 데이터를 네트워크 상에서 스트림 형태로 상관센터에 직접 전달하려는 시도가 지속적으로 이뤄졌다. 하지만 이를 구현함에 있어 몇 가지 문제점이 있다고 할 수 있는데 아무리 안정적인 연구망이라 할지라도 성능이 저하될 수 있고 이로 인해 언제라도 패킷이 손실될 가능성이 존재하기 때문이다. 더욱이 VLBI 관측 과정에서 저장되는 데이터는 단일 스캔 기준으로 길게는 수 시간에 걸쳐 기록되는 장시간 특성을 갖기 때문에 네트워크 불안정으로 인해 발생하는 패킷 손실은 관측 데이터의 신뢰성 저하와도 직결된다

이러한 VLBI 데이터 스트림의 입출력을 위해 현재 PF_RING이라 불리는 링 버퍼 라이브러리가 널리 활용되고 있다[6]. 하지만 이 방식 하에서 시스템이 처리해야 할 IP/TCP/UDP 등의 헤더는 MAC 프레임에 포함되어 사용자에게 직접 전달됨에 따라 그 자체만으로는 커널 수준에서 흐름 제어를 통한 신뢰성 있는

패킷 처리가 불가능하다. 이로 인해 기존 방식의 링 버퍼는 상당수 네트워크 안정성이 보장된 LAN(Local Area Network) 영역에서 사용되는 실정이고, 페이로드 획득에 있어서도 사용자가 각 프레임마다 헤더를 제거하는 작업이 필요하다[7]. 네트워크 상에서 신뢰성 있는 데이터 전송을 구현하는데 문제가 있고 적용 범위를 WAN 구간으로 확장시키기 어려운 이러한 한계로 인해 현재 KVN의 경우 각 전파천문대와 상관센터 간 초고속 네트워크가 구축되어 있음에도 전파천문대에서 획득한 스트림을 상관센터로 바로 전달하지 못하고 있다.

이에 대한 해결을 위해 본 논문에서는 각 전파천문대에서 생성된 VLBI 관측 데이터 스트림을 데이터 분석이 이뤄지는 상관센터에 직접 송신, 저장하는 것은 물론, 데이터 전송 중 네트워크의 성능이 변화하더라도 관측 데이터의 손실이 최소화되도록 하는 시스템을 설계하였다. 이를 위한 본 논문의 구성은 다음과 같다. 본 서론에 이어 2장에서는 KVN 세 사이트와 상관센터 간 원격 기록을 위한 시스템 구성 및 실험 결과에 대해 알아보고 이를 로컬 기록 성능과 비교 분석한다. 3장에서는 불안정한 네트워크 상에서 UDP 데이터의 안정적 전송을 위한 시스템 설계 및 데이터 처리 절차에 대해 기술하고 4장에서 결론을 맺고자 한다.

II. KVN 세 사이트 원격 기록 실험

2.1 네트워크 및 시스템 환경 구축

기존 VLBI 관측에서 백엔드 시스템으로부터 출력된 관측 데이터 스트림은 각 관측소의 기록 시스템에 먼저 기록된 후, 분석 처리를 위해 상관센터로 재전송되었다[5]. 하지만 각 관측소에 이미 저장된 데이터를 상관센터에 추가로 전송하고 반복 저장되는 과정에서 시간 지연이 발생하였고 효율성은 저하되었다. 이러한 처리 방식은 이전의 Mark5B를 이용한 1Gbps 관측에서는 큰 문제가 되지 않았다. 하지만 대역폭과 채널 수가 각각 2, 4배에 해당하는 광대역 VLBI 관측에서 늘어난 데이터 용량에 비례하여 시간 지연이 급증하고 상관처리 생산성이 저하되는 일이 빈번하게 나타났다. 이를 극복하기 위한 방법은 네트워크 상에서 관

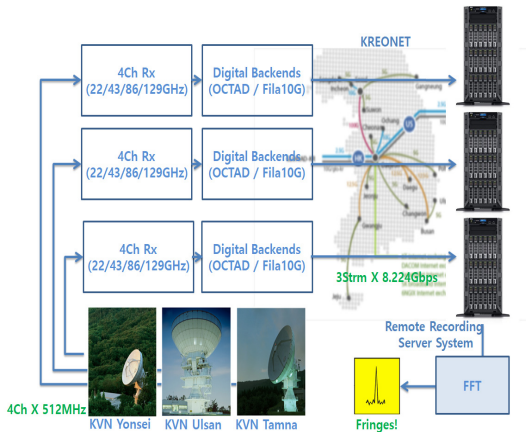


그림 1. 대용량 관측 데이터의 원격 기록을 위한 KVN 세 관측소, 대전상관센터 간 VLBI 시스템 및 네트워크 구성

Fig. 1 Configuration of VLBI System and Network between three KVN observatories and Daejeon Correlation Center for the remote recording of Massive Astronomical Data

측 데이터의 기록, 분석 처리가 이뤄지는 위치에 따라 크게 원격 상관과 원격 기록 두 가지로 분류된다. 먼저 원격 상관의 경우 관측 데이터는 관측이 이뤄지는 각 관측소에 저장되고 해당 데이터의 상관처리는 네트워크를 통해 원격으로 수행된다. 이에 대해서는 이미 2000년대 중반부터 관련 연구와 실험이 활발히 진행 중이고 미국의 NRAO(National Radio Astronomy Observatory)에서 개발된 프로그램이 제공되어 활용되고 있다[8]. 하지만 KVN의 경우 현재 여러 시스템 난제 및 효율 저하에 따라 이를 활발히 사용하지 못하고 있다. 반면, 원격 기록은 네트워크 안정성과 스토리지의 성능이 뒷받침된다면 관측 데이터의 생성 - 전송 - 기록 - 분석 과정에서 생산성과 효율성을 극대화시킬 수 있다는 장점이 있다. 이에 따라 본 논문에서는 후자에 초점을 맞춰 불안정한 네트워크 환경에서 관측 데이터 처리의 효율성을 개선하는 방안에 대해 논하고자 하며 이를 위한 실험 환경을 그림 1과 같이 구축하였다. 그림 1에서 하단에 위치한 세 안테나는 KVN의 각 관측소에 설치된 전파망원경을 나타낸 것으로 이를 통해 수집된 천체 신호는 4채널 수신기를 통해 각 주파수 대역별로 분리 수신된다. 각 수신기로부터 OCTAD 등의 디지털 백엔드에 각

채널별로 입력되는 신호의 대역폭은 512MHz이며 이는 나이퀴스트 샘플링과 2비트 양자화 적용을 통해 2048Mbps의 디지털 신호로 변환된다.

본 논문에서는 OCTAD의 4채널로부터 8Gbps급 데이터를 합성하고, 이에 대한 원격 기록 수행 및 로컬 기록과의 검증에 위해 각 채널별로 쓰레드를 개별 할당하였다. 각 쓰레드 당 VDIF 프레임의 속도는 헤더 정보를 제외하고 2048Mbps이고, 4개 쓰레드로부터 획득한 각 데이터 프레임을 하나로 병합해 처리할 경우 전체 프레임 속도는 8192Mbps가 된다. 해당 관측 데이터는 네트워크를 경유해 원격의 대전상관센터에 기록되며, 패킷 손실 및 무결성 점검을 위해 로컬 상의 Mark6 시스템에 동시에 기록될 필요가 있다. 이를 위해 상관센터의 원격 스토리지에 대한 전송 및 저장용으로 1, 2, 3, 4 쓰레드를 그리고 관측소 내부의 로컬 스토리지 Mark6 측으로 5, 6, 7, 8 쓰레드를 설정하였다. 또한 원격의 상관센터와 내부의 관측소 목적지로 복제된 두 관측 데이터 스트림이 KREONET과 사설망에서 동시에 병렬 스트림으로 전달될 수 있도록 10GbE 네트워크 스위치를 VLAN 기반으로 포트 분할하였다[9]. KVN 각 관측소에서 상관센터로 전송되는 관측 데이터는 최소 10GbE 회선을 통해 KREONET 백본에 입력된다. 이를 수신 및 저장하는 상관센터는 100GbE 회선으로 KREONET 대전지역망센터와 연결되며 네트워킹을 위해 현재 Dell 사의

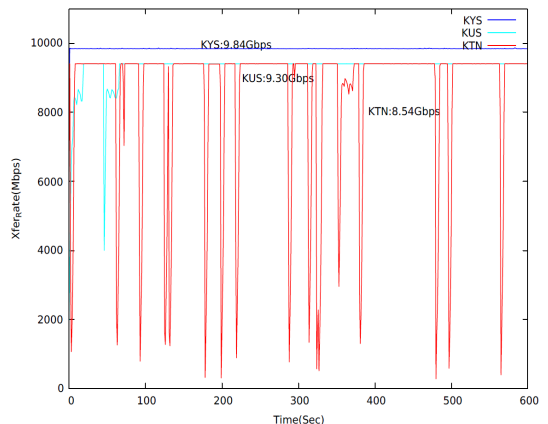


그림 2. KVN 세 사이트와 대전상관센터 간 네트워크 전송 성능

Fig. 2 Network Transmission Performance between Three KVN Sites and Daejeon Correlation Center

Z9100-ON이 설치되어 운용 중이다. 하부단 서버와의 통신을 위해 해당 스위치의 각 포트에 40GbE SR4 트랜시버를 장착하였고 이로부터 4가닥으로 분기되는 10G 회선 중 세 개를 각 관측소의 데이터가 저장되는 스토리지에 연결하였다.

2.2 실험 결과 및 성능 평가

상술한 네트워크 구성 하에서 관측 데이터 스트림의 원격 기록 실험을 수행하였다. 원격 기록은 네트워크를 경유해 상관센터의 스토리지에 바로 데이터를 저장하기 때문에 네트워크 성능이 패킷 전달의 안정성에 비례한다. 이에 따라 실험에 앞서 iperf3를 이용해 TCP 프로토콜 기반으로 KVN 세 사이트와 대전 상관센터 간 네트워크 성능을 측정하였다. 그림 2는 KVN 각 사이트에 대해 600초에 걸쳐 측정한 결과로서 10GbE NIC 기준으로 연세, 울산, 탐라 각각 평균 9.84, 9.30, 8.54Gbps에 달하는 전송 성능을 확인할 수 있다. 그래프 상에서 40GbE 백본으로 연결된 연세 사이트는 9.8Gbps 내외로 한치의 흔들림 없이 일정한 전송 성능이 유지되고 있다. 반면, 울산과 탐라의 경우는 급격한 성능 변화가 감지되는데 울산 사이트가 실험 개시 후 100초 동안 출렁거리는 패턴을 보이다가 안정화된 반면 탐라 사이트는 실험 전반에 걸쳐

| | KVN Yonsei | KVN Ulsan | KVN Tamna |
|-----------------------------------|---------------|---------------|---------------|
| Data Recording Time(second) | 15,540 | 15,556 | 15,304 |
| Number of Received Frames | 1,942,538,186 | 1,944,523,771 | 1,912,910,050 |
| Number of Packet Loss Occurrences | 10 | 20 | 29 |
| Number of Lost Frames | 765 | 536 | 1196 |

표 1. KVN 세 사이트와 상관센터 간 원격 기록 실험 결과

Table. 1 Results of Remote Recording Experiments between the Three KVN Sites and the Correlation Center

7Gbps 내외로 등락을 반복하며 성능이 일정치 않을 수 있다.

네트워크 성능 검증 후 KVN 세 사이트에서 획득한 관측 데이터 스트림을 원격지의 상관센터에 직접 전송, 기록하는 실험을 실시하였다. 데이터가 저장되는 스토리지는 물리적으로 16개의 1TB 디스크로 구성되며 RAID0에 기반해 14.546TB 크기의 단일 볼륨으로 가상화하였다. KVN의 세 관측소에서 8.224Gbps 속도로 출력되는 데이터 스트림은 장거리 네트워크를 경유해 해당 스토리지에 직접 기록되며, 해당 실험의 결과를 표 1에 요약하였다. 원격 기록 과정에서 초 단

| | Time | Frame_Num/Sec | Total_Frame | Buf_Stream_Size | Data_Size | Packet_Loss | Total_LFrame | Buf_Zone |
|----|----------|---------------|-------------|-----------------|----------------|-------------|--------------|----------|
| 1 | 14273348 | 26283/125000 | 36246 | 298087104 | 289772640 | 98717 | 213754 | 6307808 |
| 2 | 14273590 | 124874/125000 | 30286120 | 249073050880 | 249064580160 | 126 | 213880 | 7183104 |
| 3 | 14273593 | 124891/125000 | 30661111 | 252156976964 | 252142165440 | 9 | 213889 | 11636960 |
| 4 | 14273597 | 124995/125000 | 31101106 | 256268936744 | 256258953360 | 5 | 213894 | 8939488 |
| 5 | 14273598 | 124991/125000 | 31286097 | 257296861728 | 257288127840 | 9 | 213903 | 6962176 |
| 6 | 14273607 | 124996/125000 | 32411093 | 266548828832 | 266540868000 | 4 | 213907 | 7171328 |
| 7 | 14273811 | 124861/125000 | 57911054 | 476260508096 | 476246329620 | 39 | 213946 | 11275104 |
| 8 | 14273865 | 124857/125000 | 64661011 | 531772154466 | 531762770880 | 43 | 213989 | 7962608 |
| 9 | 14273869 | 124995/125000 | 66161006 | 535884113344 | 536869548640 | 5 | 213994 | 10364016 |
| 10 | 14274525 | 124940/125000 | 147160946 | 1210251619604 | 1210240427040 | 60 | 214054 | 9194432 |
| 11 | 14278237 | 124964/125000 | 611160910 | 5026187323640 | 5026136417280 | 36 | 214090 | 47814336 |
| 12 | 14278242 | 124995/125000 | 611785905 | 5031327282720 | 5031312366160 | 5 | 214095 | 10561392 |
| 13 | 14278246 | 124995/125000 | 612285900 | 5035439241600 | 5035429126080 | 5 | 214100 | 8075968 |
| 14 | 14278661 | 124951/125000 | 664160851 | 5462058838624 | 5462044397280 | 49 | 214149 | 11949472 |
| 15 | 14278962 | 124915/125000 | 751785766 | 6182686139584 | 6182678976480 | 85 | 214234 | 6727232 |
| 16 | 14279652 | 124928/125000 | 788035694 | 6480805547456 | 6480795070080 | 72 | 214306 | 9054624 |
| 17 | 14279666 | 124857/125000 | 789785551 | 6495196371424 | 6495183780480 | 143 | 214449 | 11924800 |
| 18 | 14280927 | 124880/125000 | 947410431 | 7791503384544 | 7791486681600 | 120 | 214569 | 11867232 |
| 19 | 14280966 | 124965/125000 | 952285396 | 7831596096704 | 7831585219680 | 35 | 214604 | 8758960 |
| 20 | 14280983 | 124995/125000 | 954410392 | 7849071063808 | 7849061507520 | 4 | 214608 | 7977280 |
| 21 | 14280985 | 124999/125000 | 954660391 | 7851127056584 | 7851099908160 | 1 | 214609 | 15939424 |
| 22 | 14282027 | 124862/125000 | 1084910383 | 892230289792 | 8922289436864 | 8 | 214617 | 11661632 |
| 23 | 14282654 | 124961/125000 | 1163285344 | 9566859669056 | 9566843709600 | 39 | 214650 | 11159968 |
| 24 | 14282909 | 124971/125000 | 1201536315 | 9881426430660 | 9881406896960 | 29 | 214685 | 13479136 |
| 25 | 14283329 | 124998/125000 | 1247660213 | 10260757591712 | 10260749260800 | 102 | 214787 | 8248672 |
| 26 | 14286107 | 124924/125000 | 1984910137 | 13116640966880 | 13116528573120 | 76 | 214883 | 11702720 |
| 27 | 14286891 | 124926/125000 | 1982910063 | 13922492368112 | 13922476214400 | 74 | 214937 | 13462688 |
| 28 | 14287496 | 124988/125000 | 1768536051 | 14544432259424 | 14544418229280 | 12 | 214949 | 10937920 |
| 29 | 14287756 | 124999/125000 | 1800910050 | 14810684261200 | 14810669324640 | 1 | 214950 | 11743872 |

표 2. 원격 기록 실험에서 가장 많은 오류가 발생한 탐라 사이트의 패킷 유실 내역
Table. 2 Packet Loss List of the KVN Tamna Where the Most Errors Occurred in the Remote Recording Experiment

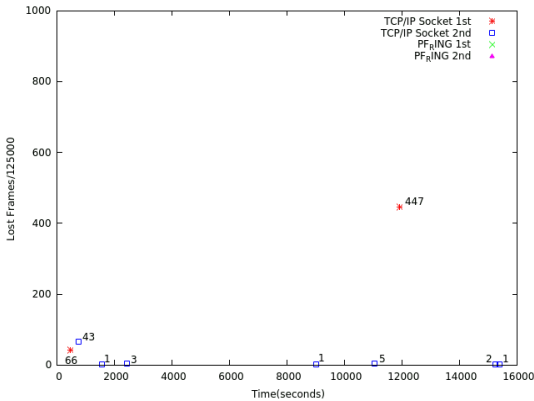


그림 3. PF_RING 링 버퍼와 TCP/IP 기반 링 버퍼를 이용한 프레임 수신, 저장 성능 비교

Fig. 3 Performance Comparison of Frame Reception and Storage using PF_RING ring buffer and TCP/IP based Ring Buffer

위로 KVN 연세, 울산, 탐라 각각 10, 20, 29번의 프레임 유실이 발생하였다. 이를 프레임 단위로 환산하면 각각 765, 536, 1196개의 프레임이 전송 및 기록 과정에서 유실되었다. 표 2는 그 중 가장 많은 에러가 발생한 탐라 사이트의 프레임 유실을 발췌한 내역이다. 원격 기록이 정상적으로 수행될 경우 매 초당 기록된 프레임 개수는 125000개이다. 하지만 실험 시작 후 29번의 에러 발생 과정에서 98717, 126, 9, 5개 등의 형태로 프레임이 매 초당 유실되었다. 그림 2에 도시한 네트워크 성능과 이에 비례하는 표 1의 프레임 유실 수치를 통해 성공적인 원격 기록을 위해서는 네트워크 안정성이 전제되어야 함을 알 수 있다.

2.3. 로컬 기록 성능과의 비교 분석

원격 기록 과정에서 발생한 프레임 손실과 네트워크 불안정의 연관성을 판별하기 위해 백엔드와 기록 시스템이 로컬 네트워크에 구축된 환경에서의 기록 성능을 추가로 검증하였다. 그림 1에 도식화된 KVN 세 사이트와 대전상관센터 간 원격기록 실험에서도 활용되었지만 백엔드에서 출력되는 8.224Gbps 데이터 스트림이 손실없이 원격에 전송, 저장되기 위해서는 수신 시스템 내부에 대용량 링 버퍼가 작동되어야 한다. 환경의 메모리에서 중단없는 데이터 입출력을 구현하기 위해 입출력 포인터는 개별적으로 관리된다. 쓰기 포인터가 외부에서 네트워크 인터페이스로 입력

되어 메모리에 쓰여지는 데이터의 주소를 추적한다면, 읽기 포인터는 최종 목적지인 디스크 저장을 위해 메모리에서 읽혀지는 데이터의 주소를 추적한다. 링 버퍼 상에서 벌어지는 상기 데이터 입출력 과정에서 각 포인터는 점진적으로 증가하며 두 포인터 간의 차이는 링 버퍼 상에서 실제 버퍼링되는 데이터 용량에 해당한다[10]. 본 논문에서는 근거리 환경의 원격 기록 성능 평가와 관련하여 2017년 전자통신학회논문지 12권 1호에 기술된 것과 동일한 환경을 실험실 내에 구축하였다. 객관적 기록 성능 검증을 위해 외부에서 입력되어 내부 메모리 상에 버퍼링되는 프레임을 TCP/IP 소켓과 PF_RING 기반으로 처리하는 두 가지 버전의 프로그램을 제작하였고 각 경우에 대한 성능을 측정하였다. PF_RING 방식의 경우 PF_RING 활성화, 함수 호출에 필요한 구조체 핸들 획득을 비롯 애플리케이션 지정, PF_RING 소켓과 결합된 디바이스 연결, 해제 등의 작업이 필요하다. 반면 TCP/IP 소켓 인터페이스에서는 일반적인 데이터 송수신과 동일하게 소켓 개설, 소켓 주소와의 결합 이후 recv, recvfrom 등의 함수 호출을 통해 외부에서 입력되는 데이터를 수신하는 방식이다. 로컬 네트워크 환경에서 위와 명시한 함수 호출을 기반으로 OCTAD로부터 전송된 VDIF 프레임이 스토리지에 저장하여 성능 평가를 수행하였고 그에 관한 결과를 그림 3과 같이 수치화하였다.

위 성능 평가는 OCTAD에서 8.224Gbps 속도로 출력되는 UDP 프레임을 1TB 용량 16개 디스크로 구성된 스토리지에 15,500초 동안 수신 및 기록하는 방식으로 진행되었다. 정밀한 결과 도출을 위해 로컬 기록 실험은 PF_RING, TCP/IP 소켓 각 방식에서 두 번에 걸쳐 수행되었다. 먼저 PF_RING 기반 링 버퍼 방식의 경우 두 번 모두 프레임 유실이 발생하지 않았다. 하지만 TCP/IP 소켓 기반 링 버퍼 방식은 두 번의 실험에서 각각 2번, 7번의 유실이 발견되었다. 해당 프레임 개수는 490, 79개로서 스토리지에 저장되는 VDIF 프레임이 8224 바이트 단위라는 점을 감안했을 때 실질적으로 손실된 데이터 용량은 각각 4MB, 0.6MB 내외에 해당한다. 이는 전체 스토리지 용량 면에서 미미한 수준으로, 본 실험을 통해 다음의 두 가지 사실을 알 수 있다. 먼저 10GbE 인터페이스의 경우 TCP/IP 소켓에 기반한 링 버퍼는 PF_RING에 못

지 않은 성능과 안정성을 보장하는 것으로 파악된다. 전자의 경우 다음 장의 시스템 설계에서 기술하겠지만 TCP에 기반한 신뢰성 있는 프레임 전달 구현이 용이하다는 점에서 활용 가치가 크다. 네트워크 특성 관련으로는 동일한 백엔드, 스토리지를 사용해 진행된 실험임에도 네트워크 거리가 짧아짐에 따라 프레임 송수신 및 저장 성능이 매우 안정화되었다. 이로부터 그림 1, 2의 프레임 손실이 백엔드, 스토리지보다는 상당부분 네트워크로부터 야기되었음을 알 수 있다. 이는 프로토콜 측면에서 네트워크 불안정 요소를 보완할 경우 원격에서 신뢰성 있는 프레임 송수신, 저장이 가능함을 의미한다.

III. 대용량 UDP 프레임의 안정적 전송

3.1 링 버퍼 기반 시스템 설계

KVN 세 사이트의 전파망원경과 수신기를 거쳐 유입되는 아날로그 형태의 데이터는 샘플러를 거쳐 디지털 변환되고 FILA10G/OCTAD 등의 이더넷 기반 광대역 전송장비를 통해 UDP 프레임으로 출력된다. 이를 각 관측소에서 상관센터로 네트워크 전송하는 과정에서 프레임 누락 발생 시, 스토리지 상의 해당 프레임 저장은 불가하고, 신뢰성 저하는 피할 수 없다. 하지만 각 관측소에서 출력되는 UDP 프레임을 TCP 기반으로 전송할 수 있다면 네트워크 상에서 패킷이 유실되더라도 재전송을 통해 복원할 수 있고 신뢰성 있는 원격 기록이 가능하다. 이를 구현하기 위해

본 논문은 기존 시스템에서 두 가지 부분을 개선하였고, 이를 반영한 시스템 설계를 그림 4와 같이 도식화하였다. 첫 번째 개선 사항은 링 버퍼로서 리눅스 커널에서 제공하는 흐름제어를 손쉽게 접목할 수 있도록 TCP/IP 소켓 인터페이스에 기반한 링 버퍼를 채택하였다. 또한 관측 데이터 전송에 있어서도 기존에는 백엔드에서 원격의 스토리지로 바로 UDP 프레임을 전송하는 방식이었다. 하지만 이 과정에서 네트워크 불안정에 따른 프레임 유실이 불가피하다. 이에 따라 본 논문에서는 이를 극복하기 위한 방안으로 UDP 프레임의 TCP 패킷 변환을 제시하였다. KVN 각 사이트의 디지털 백엔드에서 출력되는 UDP 프레임을 TCP 패킷으로 가공 후 송신할 수 있다면 상관센터는 보다 안정적으로 데이터를 수신할 수 있다. 이 알고리즘 하에서 클라이언트 내부의 링 버퍼는 UDP 프레임을 입력받아 TCP 패킷을 출력한다. 이를 위해 클라이언트에는 최소 두 개의 NIC가 장착된다. 클라이언트의 제2 NIC와 네트워크 스위치의 Public VLAN을 경유해 출력되는 관측 데이터는 네트워크 상에서 IP 기반으로 원격의 상관센터에 라우팅 전송되며 수신 및 저장에 따라 각각 스토리지 서버와 스토리지에서 처리된다. 스토리지 서버는 KVN 각 사이트로부터 전송되는 관측 데이터 스트림에 대한 단방향 수신을 위해 단일 NIC가 설치되어 있고 링 버퍼 상의 입출력을 통해 최종적으로 스토리지에 해당 데이터가 저장된다.

KVN 각 사이트와 상관센터 내의 시스템 작동 절차를 세부 기술하면 다음과 같다. 먼저 각 관측소의 경우 해당 원리를 그림 5에 도식화하였다. 관측 데이터는 각 관측소에 설치된 네트워크 스위치의 Private VLAN을 거쳐 제1 리눅스 서버 시스템의 제1 NIC(Network Interface Card)에 전달된다. 해당 관측 데이터 스트림은 OSI 7 계층 중 3, 4번째에 해당하는 IP/UDP 계층에 대한 데이터그램 소켓 API(Application Programming Interface) 호출을 통해 수신 소켓 버퍼에 저장된다. 수신 소켓 버퍼는 외부에서 실시간 입력되는 관측 데이터 프레임이 임시 저장되는 공간으로서 외부로부터 새로 입력되는 UDP 프레임이 연속적으로 처리될 수 있도록 기존에 저장되어 있던 UDP 프레임은 대용량 링 버퍼에 복사되어야 하고 이를 위해 memcpy 함수를 호출한다. 링 버

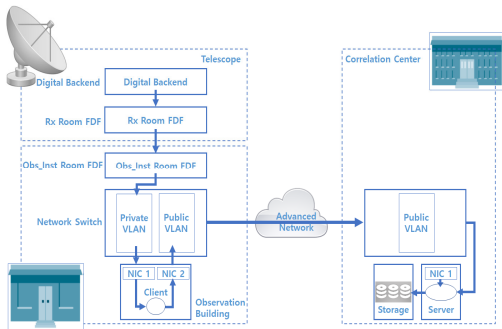


그림 4. TCP/IP 기반 링 버퍼 기반의 VLBI 데이터 스트림 전송 안정화를 위한 시스템 구성
 Fig. 4 System Configuration for Stabilization of VLBI data stream based on TCP/IP based Ring Buffer

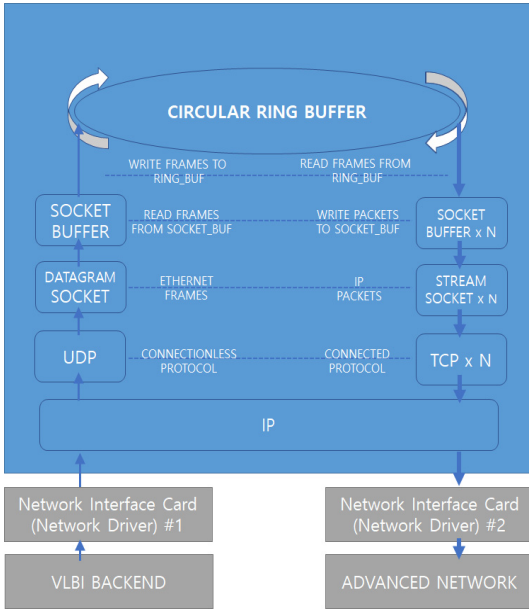


그림 5. 전파천문대에서 관측 데이터 스트림의 안정적 전송을 위한 시스템 설계

Fig. 5 System Design for Stable Transmission of Astronomical Data Streams at Radio Observatory

퍼에 저장된 프레임에서 사용자 데이터에 해당하는 페이로드가 추출하여 송신 소켓 버퍼에 저장되고 TCP/IP 계층에 대한 스트림 소켓 API 함수 호출을 기반으로 TCP 패킷으로 변환, 전송된다. 이 과정에서 전송 성능 향상을 위해 병렬 스트림이 지원되고 네트워크 성능 변화에 따라 가변적으로 운용될 수 있도록 시스템을 설계하였다.

초고속 네트워크에 연결되어 관측 데이터 스트림을 수신하는 상관센터 역시 데이터 처리 방식은 위와 유사하며 내부의 데이터 흐름을 그림 6에 나타내었다. 소켓 주소에 명시된 IP 주소, 포트 번호에 따라 상관센터의 제2 리눅스 서버로 전송된 TCP 패킷은 소켓 버퍼에 수신, 저장되고 이 과정에서 이상없음이 검증되면 대용량 링 버퍼로 복사된다. 그리고 읽기/쓰기 포인터의 점진적 증가 속에 헤더를 제외한 페이로드 누적량이 일정 용량을 초과하면 write 등의 파일 함수를 통해 스토리지에 최종 저장된다. 하지만 네트워크 전송 과정에서 패킷 손실, 중복 및 순서 역전 등의 이상이 발생한 경우에는 TCP의 흐름 제어를 통해 클라이언트에 재전송 요청이 전달되고 서버는 손상되지

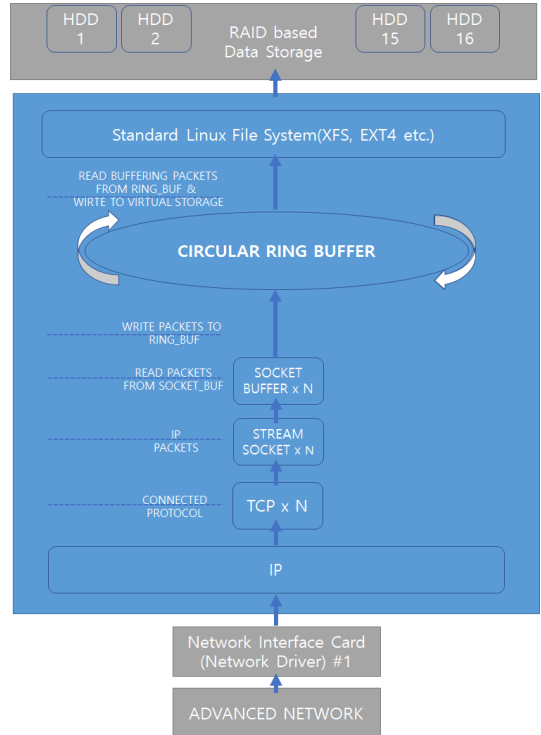


그림 6. 상관센터 내의 관측 데이터 스트림 수신 및 저장을 위한 시스템 설계

Fig. 6 System Design for Receiving and Storing Astronomical Data Streams in Correlation Center

않은 패킷을 수신받는다. 전파천문대와 상관센터에서 각각 클라이언트, 서버 시스템에 실시간으로 입출력되는 패킷은 상기 과정의 반복을 통해 링 버퍼를 경유해 송수신되며 네트워크에 연결된 어디에서라도 신뢰성이 보장된 패킷을 전송받을 수 있다. 다만, 이 과정에서 재전송 발생 수가 일정한 수치를 초과하게 되면 상관 센터에 설치된 제2 리눅스 서버는 전파천문대 내의 제1 리눅스 서버에 이를 통보하고 제1 리눅스 서버는 전송하는 스트림 개수를 증가시켜 성능 저하를 최소화한다.

3.2 제안된 데이터 흐름 및 처리 절차

지금까지 네트워크 상의 안정적인 데이터 전송을 위한 구조를 설계하였다. 그림 7은 이러한 구조 하에서 대용량 VLBI 데이터 스트림의 안정적 전송 방법을 나타낸 흐름도이며, 데이터 처리 절차를 서술하면

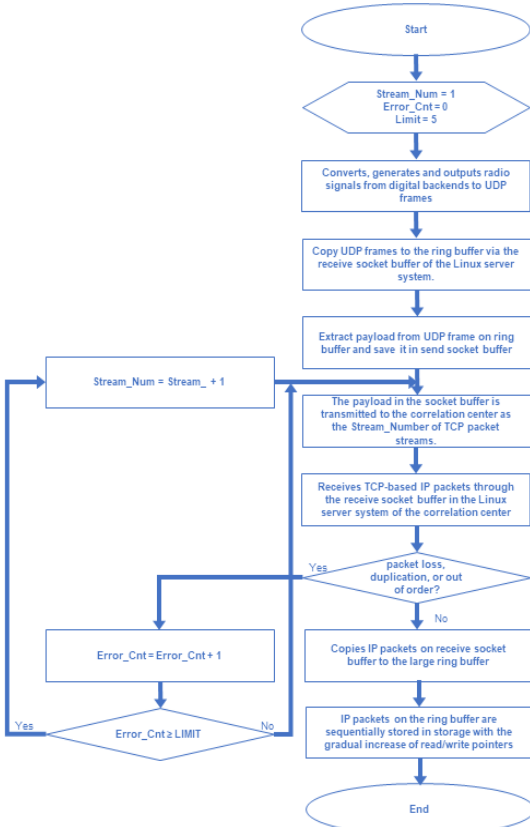


그림 7. 불안정한 네트워크 환경에서 대용량 VLBI 데이터 스트림의 안정적 전송을 위한 흐름도
 Fig. 7 Flow Chart for Stable Transmission of Large VLBI Data Streams in an Unstable Network Environment

다음과 같다. 대용량 VLBI 데이터 스트림의 안정적 전송을 위해 먼저, 각 전파천문대의 망원경으로 관측한 전파 신호를 디지털 백엔드에서 UDP 프레임으로 변환, 생성 및 출력한다. 상기 VLBI 관측 데이터 스트림은 관측동에 설치된 제1 리눅스 서버 시스템의 수신 소켓 버퍼를 경유해 링 버퍼에 UDP 프레임으로 저장된다. 해당 프레임에서 페이로드를 추출하여 송신 소켓 버퍼에 복사하고 커널은 이로부터 생성된 IP 패킷을 TCP에 기반해서 네트워크 성능 저하에 따라 가변 개수의 스트림으로 상관센터로 전송한다. 단일 또는 병렬 스트림으로 상관센터의 제2리눅스 서버에 전달되는 TCP 기반의 IP 패킷은 수신 소켓 버퍼 상에서 이상 없음이 검증되면 대용량 링 버퍼에 복사된다. 하

지만, 패킷 손실, 중복 및 순서 역전 등의 이상이 발생한 경우에는 TCP의 흐름 제어를 통해 해당 패킷은 제1 리눅스 서버 시스템에 재전송 요청된다. 링 버퍼 상에 입출력되는 패킷이 각 포인터의 업데이트 속에 기준 용량을 초과하게 되면 write 등의 함수를 통해 페이로드가 표준 리눅스 파일 시스템이 탑재된 스토리지에 저장된다. 위 순서에 따라 각 전파천문대로부터 출력되는 VLBI 관측 데이터 스트림은 네트워크 상에서 IP 기반으로 라우팅되고 소켓 주소에 명시된 IP 주소, 포트 번호에 따라 목적지의 스토리지로 직접 전달될 수 있다. 이는 네트워크에 연결된 어디라도 TCP의 흐름제어, 재전송을 통해 안정적인 패킷 송수신 구현이 가능함을 의미한다.

IV. 결론

VLBI에서 10Gbps 수준의 방대한 관측 데이터는 현재 각 관측소 내의 디지털 백엔드에서 UDP 프레임으로 출력되며 내부 네트워크까지는 안정적으로 송수신된다. 하지만 이를 수 백km 이상의 외부 네트워크로 확장할 경우, 네트워크가 불안정하면 전송 과정에서 패킷 손실이 발생하고 원하는 결과의 도출은 현재 불가능하다. 이러한 사실은 KVN 세 사이트를 이용한 UDP 기반 원격 기록 실험을 통해서도 확인할 수 있었는데, 내부 기록 실험 결과와의 비교 분석을 통해 데이터 송수신 시스템이 안정적이라도 네트워크 불안정에 따라 패킷 손실이 증가하는 현상을 입증하였다.

본 논문에서는 이러한 한계를 극복하기 위해 각 관측소에서 UDP 기반의 MAC 프레임을 단순 전달하던 기존의 방식에서 벗어나 TCP 기반 IP 패킷으로 가공, 변환 후 상관센터로 전송하는 시스템을 설계하였다. 제안된 시스템에서 각 관측소에서 출력되는 데이터 스트림은 네트워크 상에서 IP 기반으로 라우팅되고 소켓 주소에 명시된 IP 주소, 포트 번호에 따라 목적지의 스토리지로 직접 전달된다. 따라서 네트워크에 연결된 어디라도 TCP의 흐름제어, 재전송을 통해 안정적인 패킷 송수신 구현이 가능하고, 헤더 정보가 커널에서 처리됨에 따라 사용자는 실질적으로 필요로 하는 페이로드를 바로 획득할 수 있다.

또한 TCP를 이용한 데이터 전송에서 네트워크 성

능이 불안정해질 경우 병렬 스트림 개수를 증가시키면 전송 효율이 개선된다. 이에 따라 본 논문에서는 패킷 손실 발생 시 Congestion Avoidance/Slow Start로 인한 성능 저하가 최소화될 수 있도록 전송되는 스트림의 개수를 가변시키는 형태로 관측 데이터를 전파천문대에서 상관센터로 송수신하는 방법을 제시하였다. 이를 시스템에 적용할 경우 향후 네트워크 환경에 보다 최적화된 형태로 신뢰성있는 패킷 송수신이 가능할 것으로 기대된다.

Acknowledgement

본 논문은 등록 특허[발명의 명칭: TCP/IP 기반 링 버퍼를 이용한 대용량 VLBI 데이터 스트림의 안정적 전송 시스템 및 방법 / 특허 번호: 제 10-2129254호]를 기반으로 작성되었으며, 해당 연구는 “한국우주전파관측망 운영”과 “X-band 디지털 빔 포밍 배열 안테나 시스템 개발”사업 그리고, KISTI에서 운영하는 국가과학기술연구망 KREONET의 지원을 받아 수행되었습니다.

References

[1] A. Szomoru, “EXPREs and the e-EVN,” In *Proc. The 9th European VLBI Network Symp.*, Bologna, Italy, Sept. 2008, pp. 40-47.

[2] M. Lee and K. Yoo, “Implementation of a Platform for the Big Scientific Data Transfers,” *J. of the Korea Institute of Electronic Communication Sciences*, vol. 13, no. 4, 2018, pp. 881-886.

[3] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy*. Birkach: Wiley, 2004.

[4] F. Takahashi, T. Kondo, Y. Takahashi and Y. Koyama, *Very Long Baseline Interferometer*. Tokyo: IOS Press, 2000.

[5] M. Song, H. Kim, Y. Kang, D. Je, S. Wi and S. Lee, “Implementation of Ring Buffer based Massive VLBI Data Stream Input/Output over the Wide Area Network,” *J. of the Korea Institute of Electronic Communication Sciences*, vol. 14, no. 6, 2019, pp. 1109-1119.

[6] R. Cappallo, C. Ruszczyk, and A. Whitney, “Mark6: Design and Status,” *Reports of the Finnish Geodetic Institute*, vol. 5, no. 8, Mar. 2013, pp. 9-12.

[7] M. Lindqvist and A. Somoru, “Present status and technical directions of the EVN,” 12th European VLBI Network *Symp.* and Users Meeting, Cagliari, Italy, Oct. 2014, pp. 1-8.

[8] NASA, “*Int. VLBI Service for Geodesy and Astrometry*,” *Annual report 2008*, July. 2009.

[9] M. Song, H. Kim, D. Byun, T. Jung, Y. Kang, H. Kim, J. Kim, S. Wi, S. Lee, D. Roh, S. Oh, and J. Yeom, *The Design and Implementation of e-KVN Network for the Data Processing in Wideband VLBI Observation*. Daejeon: Korea Astronomy and Space Science Institute, 2019, pp. 3-5.

[10] M. Song, Y. Kang and H. Kim, “A Study on the Performance Monitoring and Optimization of a High Speed Network for the Transfer of Massive VLBI Data,” *J. of the Korea Institute of Electronic Communication Sciences*, vol. 14, no. 6, 2013, pp. 1097-1108

저자 소개

송민규(Gyu-Min Song)

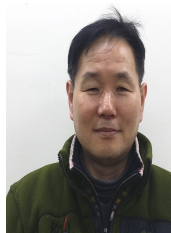


2001년 강원대학교 전기공학과 졸업(공학사)
2003년 강원대학교 대학원 전자공학과 졸업(공학석사)

2002년 ~ 현재 한국천문연구원 연구원

※ 관심분야 : 대용량 데이터 처리, 초고속 네트워크, 병렬 시스템

김효령(Hyo-Ryoung Kim)



1990년 서울대학교 천문학과 졸업(이학사)

1996년 부산대학교 대학원 천문학과 졸업(이학석사)

2003년 부산대학교 대학원 천문학과 졸업(이학박사)

1990년 ~ 현재 한국천문연구원 연구원

※ 관심분야 : 전파천문, 외부은하, 클러스터

