Introduction of K5/VSSP software correlator and Fringe finding

Kensuke Kokado(1), Mamoru Sekido(2),
Yasuhiro Koyama(2), and Tetsuro Kondo(2)

1) Geospatial Information Authority of Japan
2) Kashima Space Research Center, NICT, Japan
(April 2010)

0. Introduction

NICT started the development of software correlator for VLBI data processing since 2002[1]. Two types of software correlators have been developed corresponding to the two types of K5 data acquisition systems(DAS). Please refer to a document [2] for details on the two DASs. Two types of software correlators are briefly introduced here. Especially software correlator of K5/VSSP will be discussed more detail, and an operation to find a fringe with it will be demonstrated in this class.

1. Software correlator for K5/VSI

K5/VSI is a DAS designed for wide-band high sampling rate observation. A software correlator for K5/VSI is named GICO-3. The main developer of GICO-3 is M.Kimura at NICT Kashima. Now DiFX correlator [3] developed by A.Deller is popular and going to be used in Australia and NRAO in the USA. The architecture of the core of the GICO-3 is quite similar with that of DiFX correltor. Although the implementation of GICO-3[4] was earlier than DiFX, its employment to large correlation system delayed, unfortunately.

The main algorithm of GICO-3 is FX correlator. VLBI data of a station is divided into a small chunk of data. Then it is converted to frequency domain by FFT, and distributed to available processors. Each one processor receives the chunk of data synchronized at the same epoch from all stations. Each one processor makes cross correlation for every combination of stations, and the results are sent back to managing server and merged for output file.

The GICO-3 software correlator is now employed for backup correlator for VERA project of NAOJ. Originally the GICO-3 was designed for processing of 32-bit stream defined in VSI-H specification and high speed data processing was realized with FX-type processing. Recently GICO-3 has been extended to accept the other types of data stream such as Makr5B and K5/VSSP. The data processing capability is getting over wrapped with the correlator of K5/VSSP.

Detailed usage of GICO-3 is not described here, because it is not formally used for geodetic purpose observations, but for astronomical correlation processing to generate data analyzed with AIPS.

2. Software correlator for K5/VSSP

Development of K5/VSSP data acquisition system (DAS) started from 2002 with intension of utilizing it multi-purpose scientific data acquiring with accurate time tag.

And also data transportation with Internet Protocol (IP) over the network has in the scope. That is why the data sampling board has been sometime called ‘IP-VLBI board’. Japanese geodetic VLBI observations are performed with K5/VSSP. Also the K5/VSSP system is regally used international IVS sessions, Wettzell-Tuskuba intensive-2 session, and VLBI observation at Showa station in Antarctica.

The minimum unit of K5/VSSP32 DAS[2] is composed of a PC and a sampling unit, which has video signal input up to 4 channels. Four set of K5/VSSP32 system compose a VLBI DAS for geodetic observation with 16 channels. Therefore the minimum unit of data processing by K5/VSSP software correlator is a data file which contains data up to 4 channels.

The main developer of the K5/VSSP software correlation system is T.Kondo of NICT. The software correlation system is composed of a set of command tools to manipulate the data. List of utility software is listed in Appendix A. This software correlator is freely available under a license agreement with NICT.

The correlation processing is performed with following steps

(1) **[A priori Delay]** A priori delay model computation is performed for the baseline to be processed. ‘apri_calc’ generates a priori delay model by using schedule file as input. Station coordinates, radio source coordinates, radio frequency information, and scan information is taken from schedule file. Clock offset/rate, polar motion and dut1 data are specified from command line. Also the paths to the data files for X,Y stations have to be specified in command line, and they are recorded in the output ‘apriori_file’. Therefore ‘apriori_file’ contains most of information necessary for correlation processing.

(2) **[Correlation]**Correlation process is performed by ‘cor’ or ‘fx_cor’ program with a priori delay file as an argument of command line. These two software are XF type and FX type respectively. Number of lag window, integration unit (parameter period), and additional clock offset can be specified from command line option.

The correlation software performs following tasks: (1) synchronization of two data stream, (2) delay and Doppler tracking in accordance with a priory delay model, (3) taking cross correlation and integration of data for a given integration period. (4) Extracting Pcal information (amplitude and phase) by multiplying sinusoidal signal to the data. (5) Generating ‘cor_file’ with storing of correlation output and pcal information.

(3) **[Fringe finding and clock adjustment]** Further integration is performed for correlation output data as the second step. ‘sdelay’ is the tool for the post processing and visualization of correlation result. Maximum of correlation coefficient is searched in delay and delay-rate domain. Fig.1 is an example of the plot generated by ‘sdelay’. The delay and delay rate value residual is thought to be a clock offset and rate. The possible clock parameters are used to generate the ‘apriori_file’ again and step 1-3 are repeated again to fix the clock parameters.

(4) **[Mass processing of correlation]** After clock parameters are fixed, mass correlation processing is performed for all the data. A command ‘cor_all’ can be use for processing more than one scans. Another mass processing scheme by using multiple processors has developed by M.Sekido and introduced in latter section 4. In any way, correlation output files are

generated corresponding to each scans. Since a set of 4 channels of data is minimum unit of the DAS, geodetic observation with 16 video channels is divided into 4 groups of data, where one group contains the data of 4 channels. Consequently the number of correlation output (cout) files is expected to be (#scans) X4.

(5) **[Bandwidth synthesis]** Bandwidth synthesis technique [5] is used to derive precise group delay observable. That algorithm is implemented in a program 'komb' [6]. A set of correlation output for one scan is composed of 4 cout files as described above. Thus komb reads 4 files at a time and computes bandwidth delay for S-band and X-band. Output of 'komb' called komb_file is generated for each scans as the processing result. The VLBI observation results are extracted from the group of komb_files and stored into mark3 database for baseline analysis in the next step.

3. Example of operation for fringe finding with software correlator K5/VSSP

For a demonstration of software correlation, an example of the command sequence to perform fringe detection is described here.

(1) A priori delay model generation is performed by:

```
apri_calc /home/vlbi/schedule/u8193a.skd –type 2 -apedir test ¥
        -baseid OT -wobbx 0.2250 -wobby 0.47030 -ut1_c -0.42932 ¥
        -coffset -25.2e-6 -g 1 -xdir /k55a/ad5/u8193a/kas34 ¥
        -ydir /sirius-a1/raid/u8193a/Ons-a
--------------------------- Comments ------------------------------
```

The descriptions of each command line options are as follows:

The first argument is schedule file name
-type 2: specifying naming rule type of k5 data file
-apedir test apri_file will be generated to 'test/' directory
-baseid OT: baseline id as combination of station id. Station id is A entry of $STATIONS section in schedule file.
-wobbx/-wobby: polar motion parameter in arc sec.
-ut1_c : ut1-utc in second.
-coffset: clock offset in second
-g 1: frequency channel group 1, which means channel 1-4 of the $CODE frequency information is used for processing.
-xdir/-ydir : Specify the directory path to the location of data file of X and Y station.

Since long command line options are needed, shell script is used to invoke this command. An example of apri_file is displayed at Appendix-B. The delay model is expressed with 3rd order polynomial as a legacy of hardware correlator. The polynomial epoch is selected around the center of scan and coefficient of Taylor series is indicated as entry of TAUx.

(2) Correlation processing in XF type can be performed by

cor test/ape1930000Ota.txt –lag 64 –cout cout_file1.txt
or FX type correlation by
fx_cor test/ape1930000Ota.txt –lag 1024
---------- Comments ----------
The first argument is apri_file name which contains most of necessary information for correlation processing.
-lag 64: Specifying lag window size by bit.
-cout <filename>: Specifying the correlation output filename. If it is not specified, conventional file naming is used.

(3) Post processing for fringe plot is done by
sdelay <correlation output file>
---------- Comments ----------
Following environment variables are used for data path and graphic output. 'sdelay env' command display the current values as follows:
K5SDELAY --- default directory for correlation data out
((null)), program deflt is (../cout/)
PGDISP --- default PGPLOT display device when selected so
(/xwin)
K5SDELAY value is preferable to be set '.' for processing the data at current directory. How to set the environment variables is
Case of csh/tcsh: setenv K5SDELAY <value>
Case of bsh: K5SDELAY=<value>; export K5SDELAY;

An example of fringe plot obtained by sdelay is indicated in Fig.1. Normally 4 channels of data will be displayed in one page.

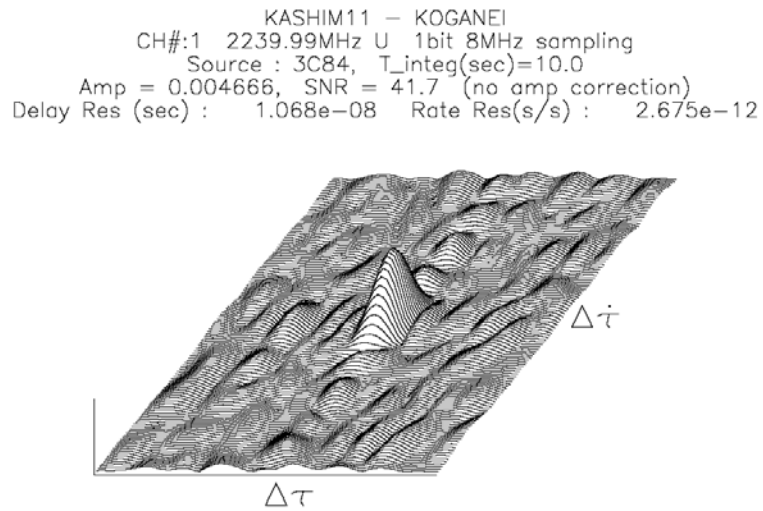

Fig. 1. A fringe plot generated by 'sdelay' command.

4. Mass processing of VLBI data with K5/VSSP software correlator via

distributed computing.

Mass processing programs by utilizing the K5/VSSP software correlator was developed with a set of tools written in Perl script. These programs are intended to utilize multiple processors in parallel to speed up the processing rate with minimum effort without modification of K5/VSSP software correlator. Since the K5/VSSP data is already divided into multiple data files containing 4 data channels, no further subdivision was applied. The K5/VSSP data files are shared by network file system (NFS) of Linux system. Major programs are following 2 tools.

- rapid_cor.pl (Management program(MP)) : Managing the correlation tasks
- Cor_srv2.pl (correlation program(CP)): Perform the correlation task given by the management server.

Fig.2 displays the overview of the distributed processing scheme.

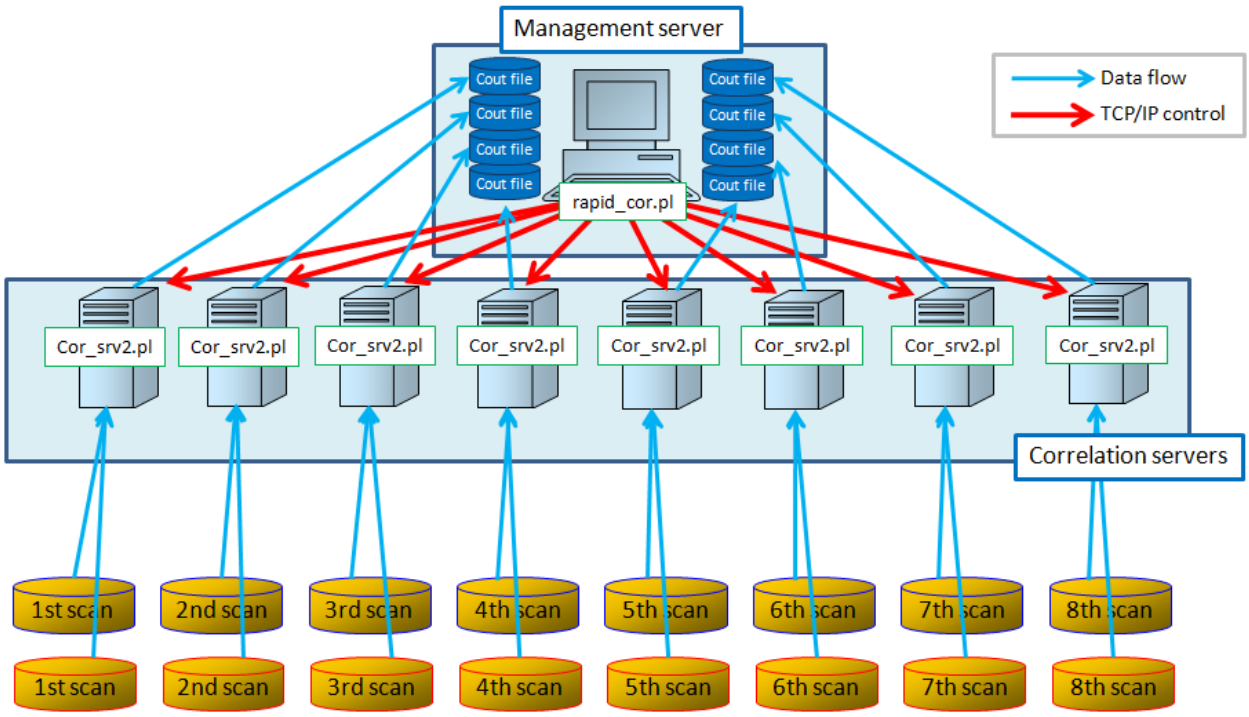

Fig.2 A scheme of distributed parallel data processing with K5/VSSP software correlator was realized with a set of agent program written in Perl script.

Brief overview of the behavior is as follows:

Environmental assumption: A priori files have to be generated in prior to the data processing. Multiple PCs or processors are supposed to be used for correlation processing and they can share the VLBI data (K5/VSSP) on data server via NFS. A priori files are also on the management server and they could be shared from all of the correlation servers.

Preparation: “Cor_srv2.pl” script has to be started at each processors. Cor_srv2.pl (CP) works as a server with waiting connection at port 9002(default). If multiple

CP scripts run on multi-CPU processor, the port has to be specified so that the server port will not conflict each other. The list of Correlation servers and its port number have to be registered on configuration file, which is read by rapid_cor.pl (MP).

Get starting: When MP is started at the management server with the configuration file, MP make connection to each CP via TCP/IP and send initialization message. After the initial connections to all the processors finished, MP wait for a task ordering message at TCP/IP port 9001(default). MP reads the schedule file of the session and searches for CP in idle state by scanning the CP list. The scanning is performed by question/answer communication via TCP/IP. If idle CP is found, MP stops the scanning and sends the task massage to CP. CP will performs the task and return to idle state when it the correlation task has accomplished. CP keeps the task record in memory until it is killed. Correlation tasks are distributed until all the CPs are occupied by a task. If the number of received task orders is larger than the number of processors, the remaining tasks are kept in a queue of MP, and wait until free CP appears. MP scans CPs registered in the configuration list periodically with searching for free CP, and a task in the queue is thrown when free CP is found.

The ‘task ordering message’ is ASCII character messages in the form

Scan: yyyy/ddd-HH:MM:SS <Baseline ID> <group id>

where <Baseline ID> is composed of two characters of two station IDs.

<group id> is one of ‘a’, ‘b’, ‘c’, ‘d’, which corresponds frequency channel 1-4, 5-8,9-12,13-16, respectively.

Since apri_file can be uniquely identified from ‘task ordering message’, CP can perform the correlation task only from that information.

Appendix A Utilities of K5/VSSP software correlator

#	Command	Function Description
1	apri_calc	A priori parameter calculation (both standard schedule file and VEX file are supported)
2	cor	Software correlator dedicated to 1 bit sampling data processing
3	cor_all	cor for two or more scan data
4	fx_cor	General purpose software correlator
5	fx_cor_all	fx_cor for two or more scan data
6	m5check	File format check of Mark5 data (automatic judging in 8|16|32|64 track mode), and display header block without sync check
7	m5time	Dispaly time label of Mark5 data
8	m5tok5	Conversion from Mark5 to K5 (all mode are supported)
9	m5tok5np	as same as "m5tok5 (Ver.2004-09-23)". No more updated.
10	m5tok5R	Conversion from Mark5 to K5 (for the purpose of checking program). Speed is slower than "m5tok5".
11	m5tok5Rnp	as same as "m5tok5R (Ver.2004-09-23)". No more updated.
12	m5vex_ana	Vex file analysis
13	k5tom5	Conversion from K5 to Mark5 (K5/VSSP32 is supported)
14	k5tom5H	Conversion from K5 to Mark5 (specialized for Huygens observation data)
15	sdelay	Coarse fringe search (2nd order search, fringe phase and amp plot, PCAL phase and amp plot were newly implemented)
16	skdchk	Check disk size required for K5/VSSP observation
17	datachk	K5 data check
18	datacut	shorten data period (size)
19	speana	display spectrum from K5 data file
20	speana2	display spectrum from K5 data file (more powerful than speana)
21	extdata	extract a certain channel data from K5 binary data and write out as an ascii data file
22	four2one	convert 4ch K5 data file to 1ch K5 data file
23	one2four	conbine four 1-ch K5 data file to one 4-ch K5 data file
24	data_half	reduce the sampling rate of K5 data file half by re-sampling the data every 2 samples
25	data_double	double the sampling rate by inserting a dummy sample as same as the sample just one sample before
26	adbitconv	change A/D converter resolution of K5 data file (already existed) artificially
27	data_recov	recover corrupted header of K5/VSSP and K5/VSSP32 data file
28	pcalcheck	monitor PCAL phase and amplitude in a K5/VSSP or K5/VSSP32 data file (recommended graphics is PGPLOT)

Appendix-B Example of aprif_file

```
** This is Apriori file made by apri_calc Ver. 2008-02-19
**      for cor, cor_all, fx_cor, and fx_cor_all
**
** Clock parameters at run are as follows,
**    Clock Offset (s) :    -2.520000000000000e-05
**    Clock Rate (s/s) :     0.000000000000000e+00
**    Clock Epoch        :      0000/000 00:00:00
**
**

$EXPCODE
U8193A

$OBS_NUMBER
1

$STATION1
KASHIM34 /k55a/ad5/u8193a/kas34/O193000000a.dat

$XYZ-STATION1
-3997649.201580 3276690.751330 3724278.805270

$STATION2
ONSALA60 /sirius-a1/raid/u8193a/Ons-a/T193000000a.dat

$XYZ-STATION2
3370606.045020 711917.494060 5349830.725710

$BASEID
OT

$FRQ_GRP(1-4)
1

$FREQUENCY
8212990000.0 U
8252990000.0 U
8352990000.0 U
8512990000.0 U

$PCAL_FREQ
10000.0
10000.0
10000.0
10000.0

$CLOCK
OFST=    -2.520000000000000e-05
RATE=     0.000000000000000e+00
XCOF=     0.000000000000000e+00
```

```
$SOURCE
0059+581

$RA
1 2 45.76238297

$DEC
58 24 11.13664300

$EPOCH
2000.0

$GHA
18 15 0.204000

$EOP
UT1-UTC= -0.429320
X_WOBB = 0.225000
Y_WOBB = 0.470300

$START
2008193000000

$STOP
2008193000120

$APRIORI
PRT=2008193000040
TAU0=   -9.898760399666164e-04
TAU1=   -9.566228916776988e-07
TAU2=   -1.944019701955070e-11
TAU3=    5.086822673153191e-15

$END
```

References

[1] *Hiro Osaki, Tetsuro Kondo, and Moritaka Kimura,* Development of Versatile Scientific Sampling Processor (VSSP) – A practical approach, IVS CRL-TDC News No. 20, pp.7-9, 2002.

[2] Mamoru Sekido, Yasuhiro Koyama and Tetsuro Kondo, introduction to the K5 System, TOW2009 class note, 2009.

[3] Deller, A. T., S.J. Tingay, M.Bailes, C. West, DiFX: A Software Correlator for Very Long Baseline Interferometry Using Multiprocessor Computing Environments, The Publications of the Astronomical Society of the Pacific, Volume 119, Issue 853, pp. 318-336, 2002

[4] Moritaka Kimura,and Junichi Nakajima, The implementation of the PC based Giga bit VLBI system, IVS NICT-TDC News No.21 pp.31-33, 2002.

[5] Rogers, E.E.A., Very long baseline interferometry with large effective bandwidth for phase-delay measurements, Radio Science, Vol. 5, pp.1239-1247, 1970.

[6] Takahashi Y., Hama S., Kondo T., Journal of Communications Research Laboratory, 38, 481,1991.