IVS Memorandum 2008-009v01

15 July 2008

“VLBI2010 Sensitivity and Data Storage Requirements”

Bill Petrachenko

Memo: VLBI2010 Sensitivity and Data Storage Requirements

July 15, 2008
Bill Petrachenko

1. Introduction

In the VLBI2010 quest for 1 mm position accuracy, two strategies under study are to increase the number of observations per session and to increase the precision of the delay observable.

Increasing the number of observations per session requires the use of both short slew times and short on-source times. VLBI2010 will use burst mode data acquisition (in which data are taken at a very high rate and streamed to media more slowly during slew periods) and high record rates to reduce on-source time, and has chosen a smaller minimum antenna diameter of 12 m to help ease the implementation of the fast slew speeds required to shorten slew times. The need for smaller antennas and shorter on-source periods both result in reduced sensitivity. This has the negative impacts of increasing the minimum detectable source flux density (hence reducing the number of available sources) and also making it more difficult to achieve the VLBI2010 delay precision target of 4 ps. Increasing the data volume per observation will be used to overcome the former problem and the broadband delay technique has been proposed to overcome the latter.

The purpose of this memo is to study:

- The relation between the source low flux density cut-off (i.e. weakest source that will be observed) and the number of available sources;
- The relation between the number of available sources and the total data volume required for each VLBI2010 session; and
- The operational consequences of the required data volume on shipping and media costs.

Although a source low flux density cut-off may need to be applied at the start of operations to help reduce total data volume, this will only apply to regularly (daily) observed sources. It is assumed that all higher quality sources (even if below the flux density limit) will continue to be monitored on a cyclic basis perhaps with the assistance of larger legacy dishes to enhance sensitivity.

2. Sensitivity Considerations

In VLBI, a commonly used measure of antenna sensitivity is the system equivalent flux density (SEFD), which can be written,

$$SEFD = \frac{2kT_S}{Ae} = \frac{8kT_S}{\pi D^2 \cdot eff} \qquad (1)$$

where k is Boltzman's constant

T_S is the system temperature

Ae is the effective area of the antenna

D is the diameter of the antenna, and

eff is the antenna/feed efficiency.

The VLBI2010 antenna is specified to have minimum D = 12 m, Ts = 50 and eff = 0.5. For these values, SEFD is about 2440.

VLBI sensitivity is governed by the equation

$$SNR = S \cdot \eta \cdot \sqrt{\frac{2 \cdot B \cdot T}{SEFD1 \cdot SEFD2}} \qquad (2)$$

where SNR is the signal-to-noise ratio of the correlation of signals from antennas 1 and 2.

S is the correlated flux density of the source (J)

η is the VLBI processing factor (~0.7 for 2-bit sampling)

B is the processed bandwidth, and

T is the integration time (s).

For a pair of VLBI2010 antennas with SEFD = 2440, a half filled 1 GHz IF bandwidth in two polarizations (i.e. B = 1.0 GHz), and 2-bit sampling (i.e. $\eta \sim 0.7$), the SNR per band can be rewritten

$$SNR = 12.8 \cdot S \cdot \sqrt{T} \qquad (3)$$

The use of a half filled band will allow better bandwidth factor optimization and RFI avoidance.

In IVS memo 2008-005v01, it was shown that, for over 50% of the broadband frequency sequences considered, broadband delay was resolved for SNR's per band less than 8. The question remains though, how much margin is required to account for real world factors such as RFI, source structure and the fact that VLBI2010 observations will be taken at frequencies as high as 18 GHz while current flux density lists are for S/X data only? In this memo, the cases of SNR targets of 10 and 14 per band will be considered. Equation 3 can be rearranged to give,

$$T = \frac{SNR^2}{164 \cdot S^2} \qquad (4)$$

3. Source availability

Source availability is based here on a list of 230 low structure sources generated specifically for geodetic observations by Leonid Petrov. The list is attached in Appendix A. The first entry in the list is the source name; the second is the minimum correlated flux density (mJ) (among all baseline lengths up to 13000 km and both S and X bands); and the last is the integration time T (s) to achieve SNR=10 per band (equation 4) assuming antennas and record bandwidths for VLBI2010, i.e. D=12 m, Ts=50 K, eff=0.5

and B=1GHz. Integration times are rounded up to the next largest second. As an example, a flux density (S) of 250 mJ will require an integration time (T) of about 10 s to achieve SNR=10 per band.

Based on this list and given a lower flux density limit, it is possible to determine a sub-list of sources capable of achieving that lower limit and then. For those, an average integration time to achieve a given SNR target can also be calculated. In Table 1 number of available sources and average integration times (for SNR=10) are listed for lower flux density limits from 0 to 1000 mJ in steps of 50 mJ. The same information is also displayed in Figure 1 but only for flux density lower limits in the range 250 to 500 mJ. In Figure 2, the relation, from Table 1, between average integration time and the number of available sources is displayed.

Flux limit (mJ)	Available No. of Sources	Mean T (s) for SNR=10
0	230	7.97
50	230	7.97
100	228	7.20
150	221	6.29
200	210	5.55
250	185	4.51
300	155	3.70
350	125	3.04
400	91	2.41
450	75	2.07
500	61	1.82
550	49	1.53
600	42	1.41
650	34	1.27
700	31	1.19
750	26	1.04
800	24	1.00
850	20	1.00
900	18	1.00
950	17	1.00
1000	15	1.00

Table 1. For low flux density limits from 0 to 1000 mJ in steps of 50 mJ, the available number of sources (out of a total of 230) are listed along with the average integration times T (s) required to achieve SNR=10 (assuming uniform observation rate for all sources).

Figure 1. Based on data from Table 1, the number of low structure sources (see list of 230 in Appendix A) available for observation vs minimum flux density limit.

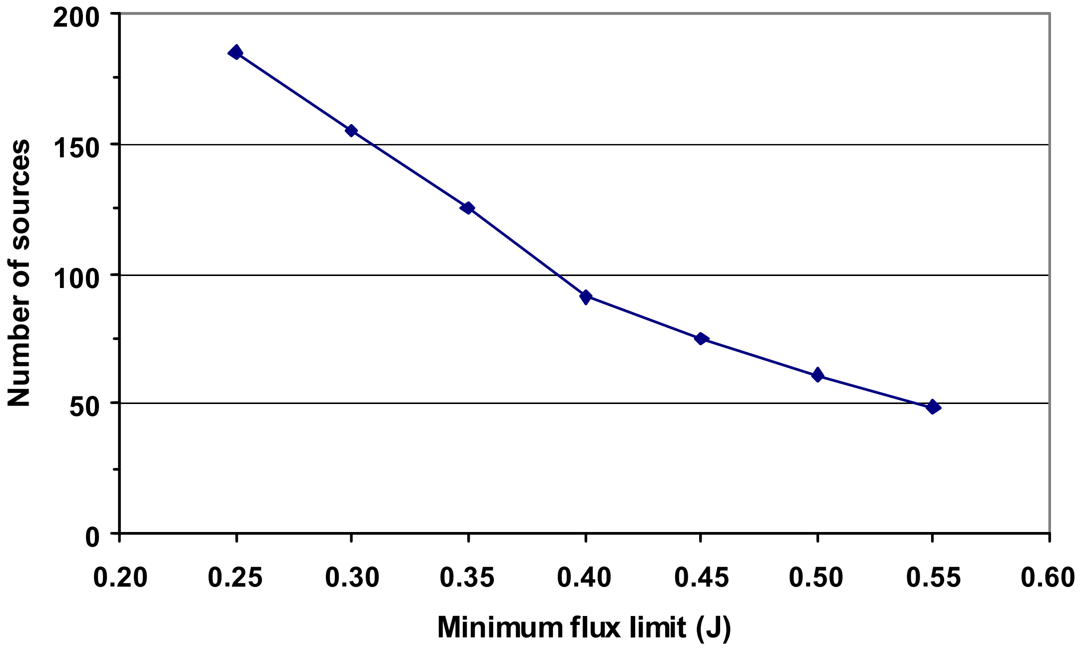
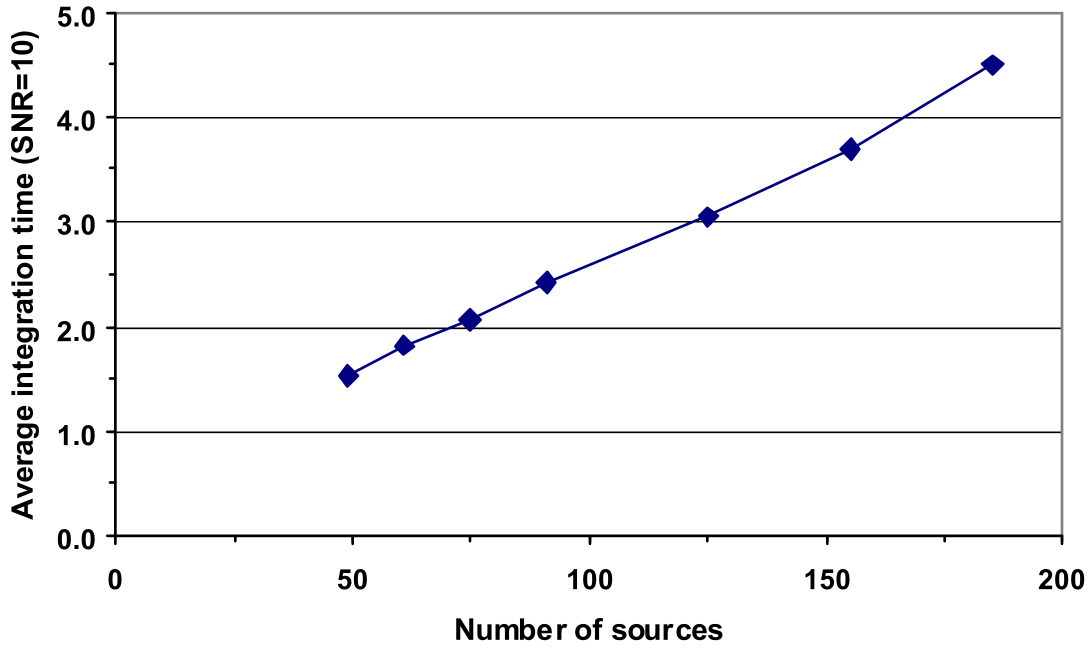


Figure 2. Based on data from Table 1, the average integration time (for data rate=16 Gbps and SNR=10) vs number of available sources



4. Media considerations

Data can be transported from stations to the correlator either by eVLBI or recorded on disk and physically shipped. In this memo, we will only consider issues related to disks.

The current established maximum capacity for 3.5” hard disks is 1 Tbyte. Recent advances in areal density in data recording have been achieved by using perpendicular recording, where magnetic domains are oriented perpendicular to the surface instead of longitudinally as in the past. The 1 Tbyte milestone has been achieved through the use of both multiple platters and perpendicular recording. The fact that data is now passing more quickly below the sensor heads means that higher data rates will also be available for the same rotational speeds.

Hitachi’s new “current perpendicular to the sensor plane giant magneto-resistive (CPP-GMR) heads are now a demonstrated technology that will make better use of perpendicular recording to achieve at least a factor of four increase in areal density, providing the technological basis for a 4 Tbyte 3.5” hard disk. However, since market demand has not yet caught up with the 1 Tbyte capacities, initial applications for the new heads will be in smaller disks such as the 2.5” form factor where a 1 Tbyte version is expected in 2009; 4 Tbyte 3.5” versions are not expected until 2011, which, incidentally, will maintain the historical factor of 2 per 2 years rate of capacity increase for hard disks. One worry with the 4 Tbyte disks is that it will take a major fraction of a day to back one up. However, historically, demand has always kept pace with capacity, so this is not expected to be a fundamental problem.

For the remainder of this memo, the standard unit of disk storage will be an 8-pack of 4 Tbyte disks having a total capacity of 32 Tbytes. These should be available prior to initial VLBI2010 operations, which are not expected before 2012. As will be seen, the fallback to 1 Tbyte 3.5” disks will be a significant limitation for continuous VLBI2010 operations. Another possible fallback might be disk arrays of 1 Tbyte 2.5” drives, which are likely to become available as early as 2009. A group of four 2.5” HDD’s is about 70% of the weight and volume of a single 3.5” HDD of equivalent total capacity, making this an attractive option with respect to shipping cost.

Another option for the future may be silicon storage based on flash memory. At present flash memory is about four times more costly than disk memory per Gbyte but has recently been doubling in capacity every year as opposed to every 2 years typical of disks. Alternate (but not yet fully proven) flash technologies (apparently independent of lithography limits) are promising 10’s of Gbytes per chip by 2010 and 1 Tbyte per chip by 2018.

5. VLBI2010 Media and Shipping Costs

The number of 32 Tbyte disk packs required per day (Dppd) can be calculated according to

$$Dppd = \frac{86400}{\Delta T_{SRC}} \cdot \frac{T_{av} \cdot br}{dpc} \qquad (5)$$

where $Dppd$ is the number of 8-packs required per day

T_{av} is the average integration time

br is the bit rate

dpc is the disk pack capacity

ΔT_{SRC} is the source switching interval

This relation is plotted in Figure 3 with respect to number of available sources assuming br =16e9, dpc=32e12, ΔT_{SRC}=45 or 60 s, SNR= 10 or 14 and the relation between T_{av} and the number of available sources is obtained from Figure 2.

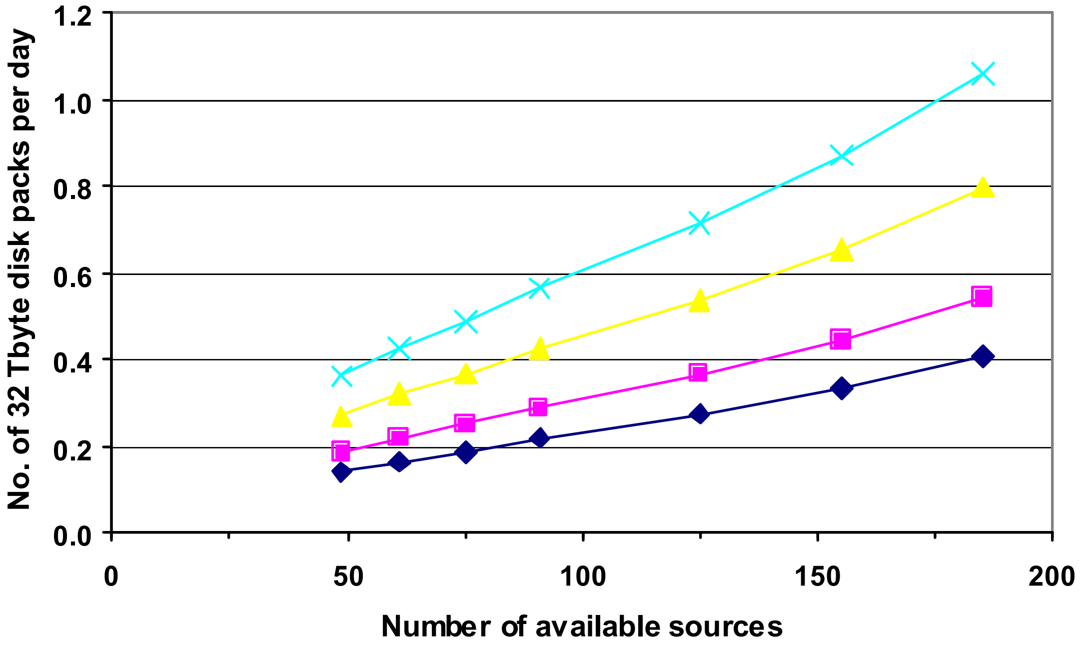


Figure 3. Number of 32 Tbyte disk packs needed per day is plotted relative to the number of available sources. The top curve assumes SNR=14 and ΔT_{SRC}=45 s, the next-to-top curve assumes SNR=14 and ΔT_{SRC}=60 s, the next-to-bottom curve assumes SNR=10 and ΔT_{SRC}=45 s and the bottom curve assumes SNR=10 and ΔT_{SRC}=60 s

To make most efficient use of the disk packs, they should be shipped only when completely full and it should be possible to switch disk packs in mid session. In the situations with lowest data volumes (e.g. SNR=10, ΔT_{SRC}=60 s, and number of sources=50) this has the negative impact of delaying the shipment of a session by as much as week, assuming daily observations. In this case, much lighter disk packs of 2.5" 1 Tbyte disks could be an answer. However, for more likely VLBI2010 operating modes (e.g. SNR=14, ΔT_{SRC}=45 s, and number of sources=125) the added shipping delay will typically not be longer than a day.

Shipping costs. If a typical VLBI2010 antenna is down 1 day per week for scheduled maintenance, there will be about 313 operations days per year for each station. Based on

Figure 3 and a shipping cost per 8-pack of \$200 each way, the annual 2-way shipping costs (\$k) per station is plotted in Figure 4.

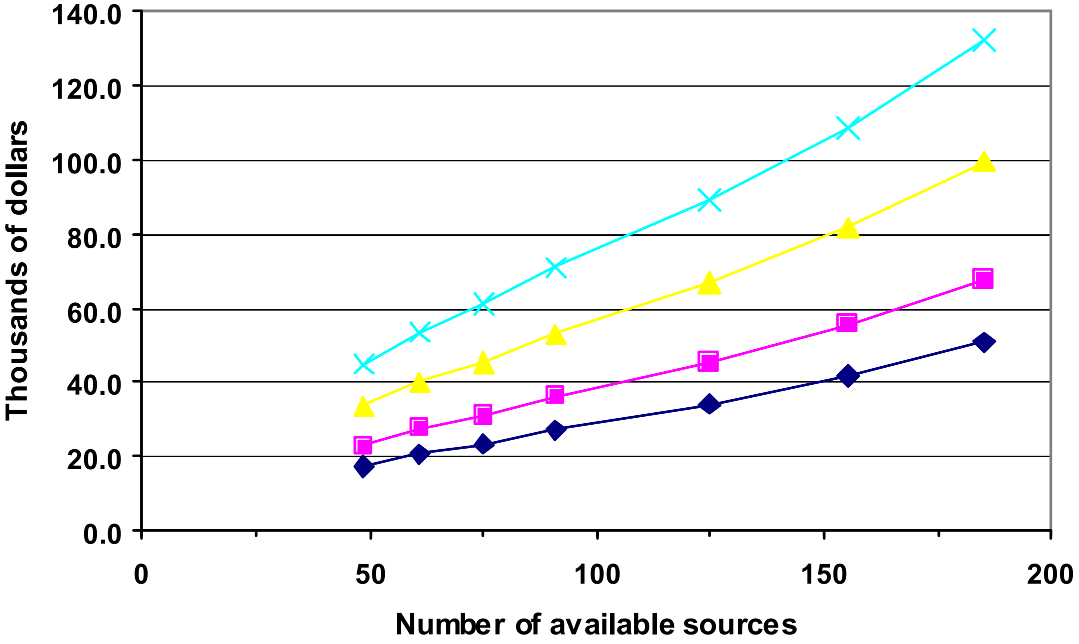


Figure 4. Annual 2-way shipping cost per station (\$k) is plotted relative to the number of available sources. The top curve assumes SNR=14 and ΔT_{SRC}=45 s, the next-to-top curve assumes SNR=14 and ΔT_{SRC}=60 s, the next-to-bottom curve assumes SNR=10 and ΔT_{SRC}=45 s and the bottom curve assumes SNR=10 and ΔT_{SRC}=60 s

Media pool. If 4 weeks of media are required per station (1 week to ship to the station, 1 week at the station, 1 week to ship to the correlator, and 1 week at the correlator) and the station observes for 6 days per week, the total number of disk packs per station in the pool can be found by multiplying the curves in Figure 3 by 24. See Figure 5.

6. Conclusions

The lower capital and maintenance costs resulting from the use of smaller fast slewing VLBI2010 antennas will be offset to some extent by higher operational costs for shipping larger volumes of recorded data. Assuming the existence of 4 Tbyte 3.5” HDD’s by 2011, shipping costs in 2012 will be significant but not insurmountable (see Figure 4), especially for some VLBI2010 operating modes. If 4 Tbyte 3.5” HDD’s don’t become available in time, fallbacks exist including the use of 1 Tbyte 2.5” HDD’s (which should be available as early as 2009). Typically, 2.5” HDD’s weigh about the same per Tbyte as 3.5” HDD’s. Assuming continued rates of improvement in HDD capacity, the situation can be expected improve with time. eVLBI remains an attractive option, but requires good fiber connections to all regularly observing sites. Perhaps this can be included as a site selection criterion, although it may be a significant limitation for approaching uniform global distribution of sites.

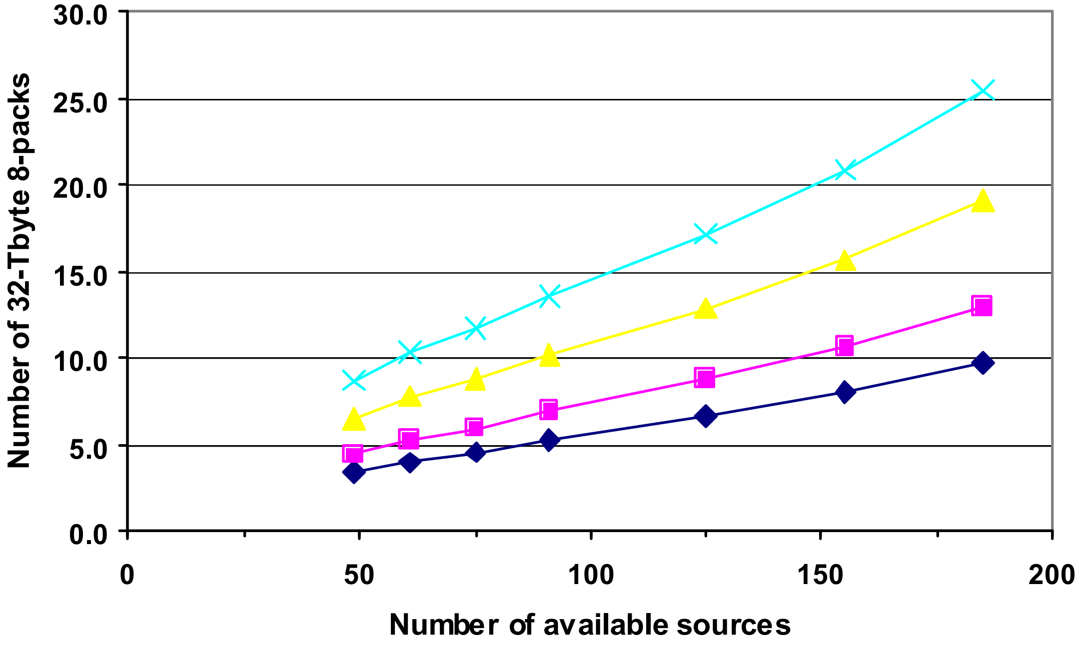


Figure 5. 24-day media pool per station is plotted relative to the number of available sources. The top curve assumes SNR=14 and ΔT_{SRC}=45 s, the next-to-top curve assumes SNR=14 and ΔT_{SRC}=60 s, the next-to-bottom curve assumes SNR=10 and ΔT_{SRC}=45 s and the bottom curve assumes SNR=10 and ΔT_{SRC}=60 s

Appendix A. Leonid Petrov's list of low structure geodetic sources

The first entry is the source name; the second is the minimum correlated flux density (mJ) for the source (among all baseline lengths and both S and X bands); and the last is the integration time T (s) to achieve SNR=10 per band assuming antennas and record bandwidths for VLBI2010, i.e. D=12 m, Ts=50 K, eff=0.5 and B=1GHz. Integration times are rounded up to the next largest second. For reference, a flux density (S) of 250 mJ will require an integration time (T) of about 10 s to achieve SNR=10 per band.

Src name	flux(mJ)	T(s)
2358+189	430	4
IIIZW2	170	22
0016+731	140	32
0019+058	370	5
0025+197	240	11
0035-252	280	8
0047+023	310	7
0048-097	460	3
0048-427	340	6
0054+161	260	9
0055-059	380	5
0059+581	1310	1
0104-408	580	2
0109+224	170	22
0115-214	330	6
0119+041	420	4
0119+115	1010	1
0134+311	190	17
0206+136	280	8
0208-512	600	2
0215+015	1130	1
0227-369	220	13
0229+131	680	2
0235+164	630	2
0237-027	290	8
0239+175	180	19
0256-005	390	4
0308-611	640	2
0307+380	450	4
0322+222	480	3
0332+078	340	6
0332-403	330	6
0338-214	320	6
0340+362	280	8
0346-279	810	1
0345+460	510	3
0347-211	300	7
0358+210	360	5
0400-319	380	5
0402-362	560	2
0405-385	810	1
0406-127	210	14
0414-189	480	3
0415+398	350	5

0420+022	730	2
0420-014	1360	1
0422+004	560	2
0422-380	480	3
0430+289	200	16
0436-129	280	8
0442+389	420	4
0446+112	900	1
0454-234	990	1
0458-020	380	5
0506-612	790	1
0506+101	360	5
0515+208	300	7
0524+034	360	5
0529+483	400	4
0534-340	340	6
0537-441	930	1
0536+145	830	1
0537-286	300	7
0544+273	240	11
0548+378	180	19
0556+238	370	5
0606-223	310	7
0607-157	1320	1
0613+570	430	4
0627-199	400	4
0632-235	340	6
0637-752	690	2
0641+392	350	5
0646-306	480	3
0648-165	1270	1
0657+172	430	4
0700-197	530	3
0714+457	370	5
0716+714	360	5
0723+219	210	14
0718+793	380	5
0727-115	2210	1
0729+259	310	7
0738+491	270	9
0743+277	220	13
0748+126	750	2
0747+185	190	17
0759+183	310	7
0800+618	370	5
0805+410	270	9
0808+019	500	3
0847-120	310	7
0854-108	490	3
0912+029	340	6
0920-397	410	4
0920+390	250	10
0925-203	200	16
0943+105	230	12
0951+268	230	12
0958+346	270	9
1004-217	310	7

1004-500	460	3
1013+127	320	6
1013+054	370	5
1015+057	390	4
1015+359	620	2
1027-186	300	7
1040+244	340	6
1053+704	260	9
1053+815	410	4
1057-797	360	5
1056+212	240	11
1059+282	290	8
1100+122	230	12
1101+384	150	28
1111+149	170	22
1125+366	200	16
1124-186	720	2
1128+385	610	2
1133-032	390	4
1144+402	500	3
1145+268	390	4
1144-379	1150	1
1149-084	970	1
1156+295	490	3
1212+171	220	13
1219+044	450	4
1226+373	170	22
1243-160	360	5
1244-255	620	2
1255-177	330	6
1300+580	320	6
1306+360	250	10
1308+328	480	3
1308+554	440	4
1324+224	770	2
1327+504	250	10
1339-287	420	4
1342+662	150	28
1348+308	290	8
1352-104	210	14
1357+769	310	7
1354-152	530	3
1406-267	540	3
1417+385	670	2
1424+366	370	5
1424-418	870	1
1428+370	200	16
1432+200	120	43
1441+252	360	5
1456+044	390	4
1502+036	390	4
1520+437	390	4
1520+319	280	8
1519-273	440	4
1519-294	410	4
1546+027	720	2
1550-242	250	10

1555+001	550	3
1601+112	220	13
1602-115	370	5
1608+243	220	13
1615+029	280	8
1616+063	300	7
1617+229	330	6
1623+578	240	11
1622-253	1090	1
1636+473	320	6
NRAO512	530	3
1639+230	250	10
1639-062	830	1
1651+391	220	13
1656-075	530	3
1657-261	150	28
1659+399	270	9
1705+018	390	4
1722+330	210	14
1718-649	550	3
1726+455	390	4
1725+123	460	3
1732+389	540	3
1736+324	290	8
1739+522	730	2
1746+470	160	24
1749+096	1460	1
1751+288	590	2
1754+155	270	9
1758+388	290	8
1800+440	270	9
1759-396	350	5
1806+456	100	61
1823+689	600	2
1846+322	470	3
1909+161	200	16
1920-211	1410	1
1929+226	290	8
1958-179	1580	1
2000+148	80	96
2000+472	1060	1
2008-159	540	3
2013+163	80	96
2052-474	140	32
2059+034	340	6
2121+053	1720	1
2127-096	440	4
2141+175	260	9
2142+110	240	11
2143-156	190	17
2144+092	300	7
2155+312	510	3
2205+166	270	9
2208-137	350	5
2214+350	400	4
2215+020	520	3
2215+150	360	5

2216+178	280	8
2227-088	1070	1
2229+695	170	22
2243+047	310	7
2250+194	200	16
2255-282	580	2
2300-683	280	8
2307+106	330	6
2309+454	410	4
2318+049	620	2
2319+317	280	8
2319+444	360	5
2335-027	240	11
2355-106	350	5
2357-318	480	3