Data Processing Architecture and Scaling for the ngVLA

Bhatnagar, S., Hiriart, R., Kern, J., Pokorny, M.

Science Requirements

The ngVLA telescope is being designed to improve the collecting area and angular resolution of the existing VLA by orders of magnitude. To achieve this, the ngVLA will have 244 18m-diameter antennas with the longest baseline of order 10^4 km. The antennas are arranged in four distinct arrays: the core-array consisting of 94 antennas with a maximum baseline of 1.3 km, the plains array with 74 antennas and 36.5 km maximum baseline, a mid-range array with 46 antennas and a maximum baseline of 10^3 km, and the rest of the antennas conforming the long baseline array, spread to radius of approximately 10^4 km. This increase in the number of antennas and maximum baseline length, while satisfying the scientific requirements, also increases the raw data rates from the telescope and the total data volume required to address the various **Key Science Goals** (KSG). Both higher sensitivity and resolution and the associated increase in data volume leads to significantly higher computing load – both in terms of the necessary raw computing and I/O load as well as computer resources (number compute-cores, RAM, data I/O bandwidth, etc.). Achieving noise-limited imaging with the ngVLA will require the use of sophisticated imaging and deconvolution algorithms which are inherently more complex (higher raw computing load and larger computer resources).

The ngVLA project defines a list of 24 science cases derived from the KSGs, with associated sensitivity, angular resolution and imaging performance requirements. These are shown in Table 1. For a viable end-to-end design of a telescope which can afford the desired scientific goals, it is also necessary to estimate the size of computing (**SofC**) associated with each case. For this, based on their technical requirements, we determine the observing parameters (integration time, total observing time, maximum baseline length, total bandwidth, number of frequency channels across the full bandwidth) and the imaging (W-Projection, A-Projection, AW-Projection) and deconvolution algorithms (Multi-Scale, Multi-Term, Multi-Scale Multi-Term) necessary to achieve the scientific goals. The SofC for the various algorithmic combinations vary significantly and the associated computing costs can vary from easily affordable to nearly unaffordable.

SC#	Name	Use Fraction	Field of View (arcsec)	PSF FWHM (mas)	Dynamic Range	Center Frequency	Bandwidth	Science Channel Width	Maximum Dump Time (sec)	Sub-array
1	*KSG1 Driving Cont Band 6 eg Taurus disk*	0.09	5.0	10	1.00E+03	100.0 GHz	20.0 GHz	120000000	1.0	Main
2	*KSG1 Driving Cont Band 4 eg Taurus disk*	0.04	5.0	10	1.00E+03	27.3 GHz	13.5 GHz	120000000	1.0	Main
3	*KSG2 Driving Line Band 5 eg Sgr B2(N)*	0.04	60.0	100	1.00E+03	40.5 GHz	4.0 GHz	13500	1.0	Main
4	*KSG2 Driving Line Band 4 eg Sgr B2(N)*	0.01	60.0	100	1.00E+03	27.3 GHz	4.0 GHz	9100	2.0	Main
5	*KSG2 Driving Line Band 3 eg Sgr B2(N)*	0.01	60.0	100	1.00E+03	16.4 GHz	4.0 GHz	5500	2.0	Main
6	*KSG3 Driving Line Band 5 eg COSMOS*	0.04	114.9	1000	1.00E+02	40.5 GHz	20.0 GHz	675000	1.0	Plains+Core
7	*KSG3 Driving Line Band 4 eg COSMOS*	0.01	170.5	1000	1.00E+02	27.3 GHz	13.5 GHz	455000	2.0	Plains+Core
8	*KSG3 Driving Line Band 3 eg COSMOS*	0.01	284.9	1000	1.00E+02	16.4 GHz	8.2 GHz	273300	2.0	Plains+Core
9	*KSG3 Driving Line Band 6 eg Spiderweb galaxy*	0.02	5.0	100	1.00E+03	72.0 GHz	240.0 MHz	7200000	1.0	Main
10	*KSG3 Driving Line Band 5 eg Spiderweb galaxy*	0.01	5.0	100	1.00E+03	36.0 GHz	120.0 MHz	3600000	1.0	Main
11	*KSG3 Driving Line Band 4 eg Spiderweb galaxy*	0.01	5.0	100	1.00E+03	27.7 GHz	92.3 MHz	2770000	2.0	Main
12	*KSG3 Driving Line Band 6 eg Virgo Cluster*	0.07	32.0	100	1.00E+03	112.5 GHz	6.0 GHz	375000	1.0	Plains+Core
13	*KSG3 Driving Line Band 1 eg M81 Group*	0.11	2473.8	1000	1.00E+03	1.4 GHz	7.0 MHz	4730	2.0	Plains+Core
14	*KSG3 Driving Line Band 1 eg M81 Group*	0.13	2473.8	60000	1.00E+03	1.4 GHz	7.0 MHz	47300	2.0	Core
15	*KSG5 Driving Cont Band 1 OTF Find LIGO event*	0.07	2920.1	1000	5.00E+03	2.4 GHz	2.3 GHz	2000000	0.5	Main
16	*KSG5 Driving Cont Band 4 OTF Find LISA event*	0.07	170.5	1000	5.00E+03	27.3 GHz	13.5 GHz	5000000	0.5	Plains+Core
17	*KSG5+4 Driving Cont Band 2 OTF Find BHs+PossiblePulsars*	0.04	1001.2	1000	5.00E+03	7.9 GHz	8.8 GHz	5000000	0.5	Plains+Core
18	*KSG5 Driving Cont eg Band 2 Followup from OTF*	0	1.0	10	5.00E+03	7.9 GHz	8.8 GHz	120000000	2.0	Main+LBA
19	*KSG5 Driving Cont Band 3 Gw170817@200Mpc*	0.24	1.0	1	1.00E+02	16.4 GHz	8.2 GHz	120000000	2.0	LBA
20	*KSG3 Supporting Cont Band 6 eg Virgo Cluster*	0	42.2	1000	5.00E+03	93.0 GHz	20.0 GHz	5000000	1.0	Main
21	*KSG3 Supporting Cont Band 5 eg Virgo Cluster*	0	114.9	1000	5.00E+03	40.5 GHz	20.0 GHz	5000000	1.0	Main
22	*KSG3 Supporting Cont Band 4 eg Virgo Cluster*	0	170.5	1000	5.00E+03	27.3 GHz	13.5 GHz	5000000	2.0	Main
23	*KSG3 Supporting Cont Band 3 eg Virgo Cluster*	0	284.9	1000	5.00E+03	16.4 GHz	8.2 GHz	5000000	2.0	Main
24	*KSG3 Supporting Cont Band 2 eg Virgo Cluster*	0	1001.2	1000	5.00E+03	7.9 GHz	8.8 GHz	5000000	2.0	Main

Table 1. Science requirements for each one of the science cases derived from the ngVLA Key Science Goals (KSGs).

Imaging Algorithms

For the purpose of SofC estimates, the process of converting raw data from the telescope to science quality image of the sky can be divided into two distinct steps: (1) data calibration and flagging, and (2) imaging and deconvolution. The cost of data calibration is relatively small and the overall cost of end-to end processing is dominated by the imaging and deconvolution step. We therefore focus on the SofC for imaging only.

The process of image-making can be divided into two steps:

1. Imaging: This step converts the raw calibrated data into a raw image. This involves re-sampling the raw data on to a regular grid (a.k.a. "gridding") and Fourier transforming the complex grid using the FFT algorithm to make a raw image of the sky (a.k.a. "Dirty Image"). The quality of this raw image is typically limited by the image-plane instrumental artifacts which are removed in the imaging modeling step.

Since the process of imaging is fundamentally iterative, the raw image is made using the residual data, after subtracting the current sky-model from the raw data. This requires the reverse process of converting a sky-model image to visibilities on a regular grid via the inverse Fourier transform, and then re-sampling this gridded data from a regular grid to the sampled points in the Fourier domain (a.k.a. "de-gridding"). In practice, the imaging step involves gridding and de-gridding and two FFT operations. The combined process is also often referred to as the "major cycle". Major cycle algorithms are typically data-domain algorithms and correct for various direction dependent (DD) effects. i.e. they work on complex-valued images and the complex valued irregularly sampled data.

2. Image modeling: This step (a.k.a. "deconvolution") involves making a wide-band model of the sky emission given the raw residual image and the telescope point spread function (PSF). This step itself is also iterative in nature and is often referred to as the "minor cycle". The minor cycle algorithms are image-domain algorithms, i.e., they work on standard real-valued images.

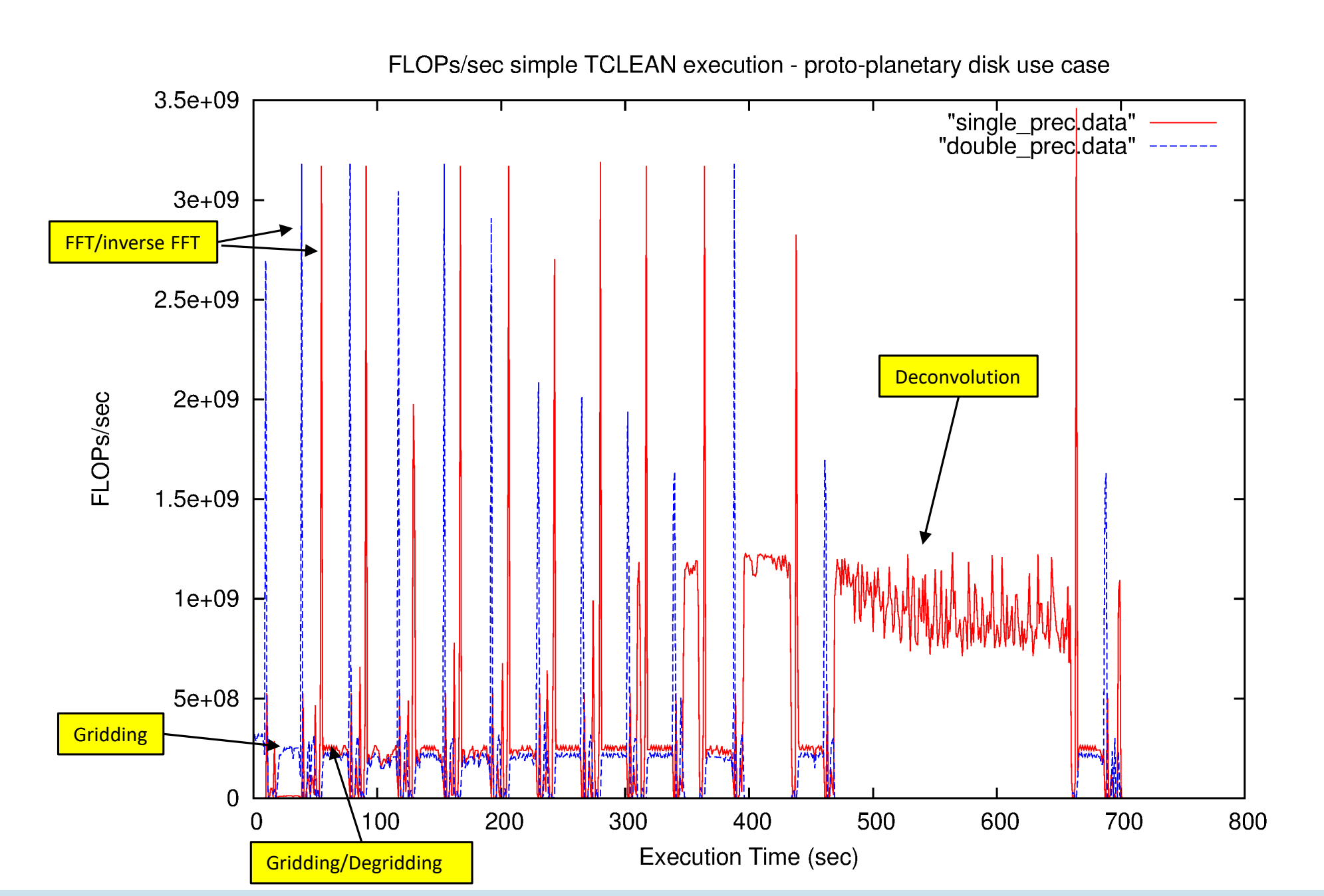

Figure 1. Number of floating point operations per second in a typical imaging execution over a small dataset. This plot exemplifies the computational costs involved by each operation. As deconvolution operates in the image domain, the number of FP operations doesn't scale with the number of the visibilities. The gridding/de-gridding cycles, on the other hand, scale linearly with the size of the dataset and become the major contributor to the computational cost for large datasets (the example shows a small dataset). The cost of the FFT and inverse FFT is relatively small and the implementation is highly parallel.

A typical end-to-end imaging requires a number of major cycles which, given the raw data and the current sky-model image re-calculates the current residual image; and a number of minor cycles which, given the current residual image iteratively improves the current model image. An example execution is shown in Figure 1. The final cost of imaging is a function of the number of major and minor cycle iterations, the algorithms used for major and minor cycles, the data volume and the image size. For most KSGs, the overall cost of imaging is dominated by the major cycle (gridding/de-gridding). We therefore developed parameterized scaling laws for the major cycle algorithms, parameterized by the data volume and the parameters of the algorithms needed. The cost of computing overheads in practical implementations of these algorithms is included via coefficients in the equations. In order to make SofC estimates, these coefficients are measured by running existing implementations in the CASA package on simulated Measurement Sets (MS) for each selected KSG.

Major Cycle Measurements and Results

Given the large size of data sets for ngVLA KSGs, current CASA software, available NRAO testing hardware, and the time available to complete these measurements, we performed trials using smaller size data sets that are nevertheless representative of the KSGs. The measurements obtained from these trials are then used to obtain scaling laws. The three primary classes of metrics we measured are floating point operations (FLOPs), I/O usage and memory usage. We use two primary tools for gathering test metrics of trial imaging runs: native CASA instrumentation, and the Score-P infrastructure (https://www.vi-hps.org/projects/score-p). The results are summarized in Figures 2 and 3.

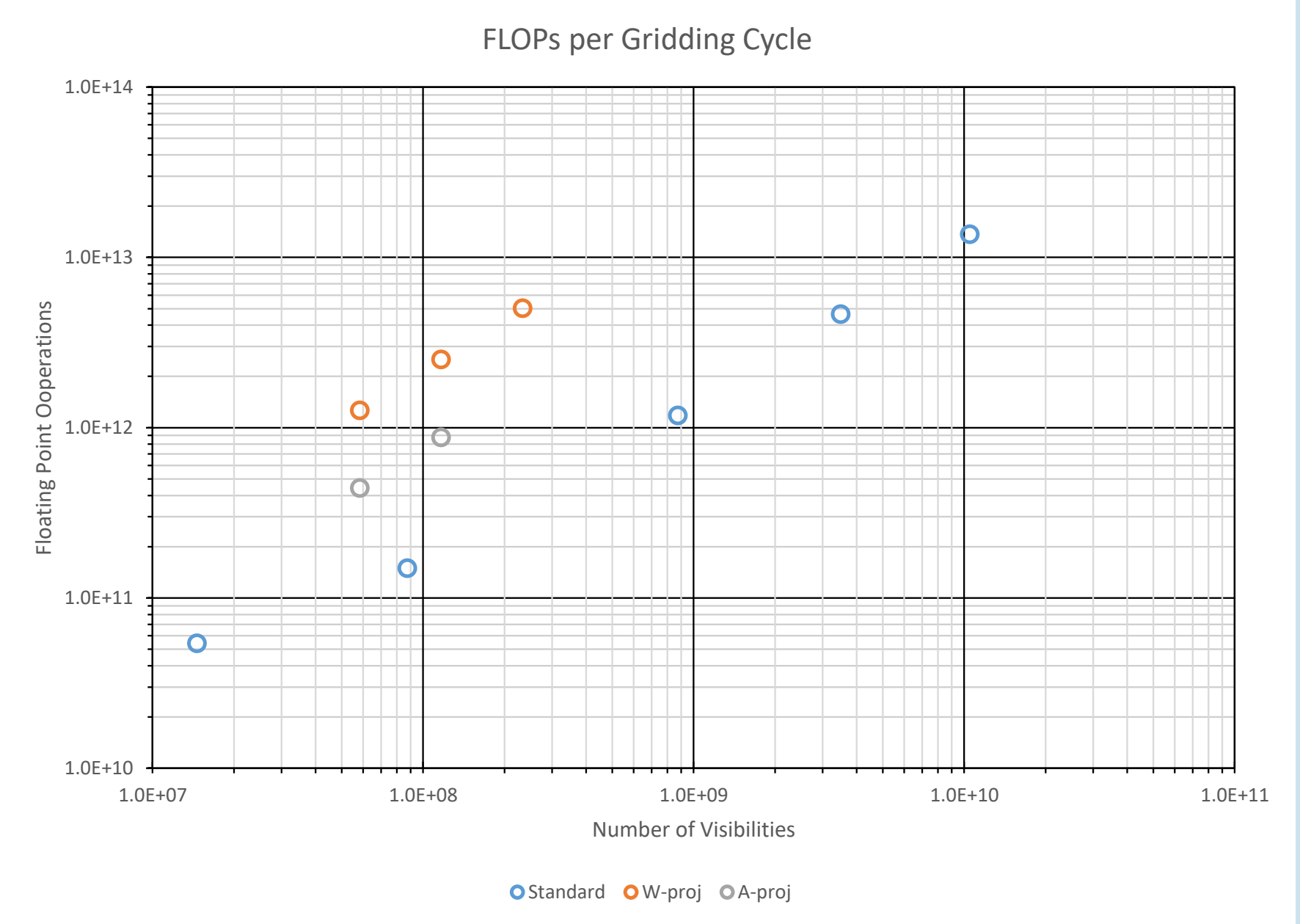

Figure 2. Number of floating point operations from executing one gridding cycle over datasets of different sizes, for different algorithms (Standard, W-projection, and A-projection). The slopes show good accordance with the expected theoretical results. For standard gridding the expected slope should be one complex multiply and one addition times the size of the convolution kernel times 2 polarizations: (7x7)x24 = 1176. The measured slope in the plot is 1280.8, the difference explained by a few additional unaccounted operations in the implementation (e.g., FP comparisons). Similar results can be found for W-proj and A-proj.

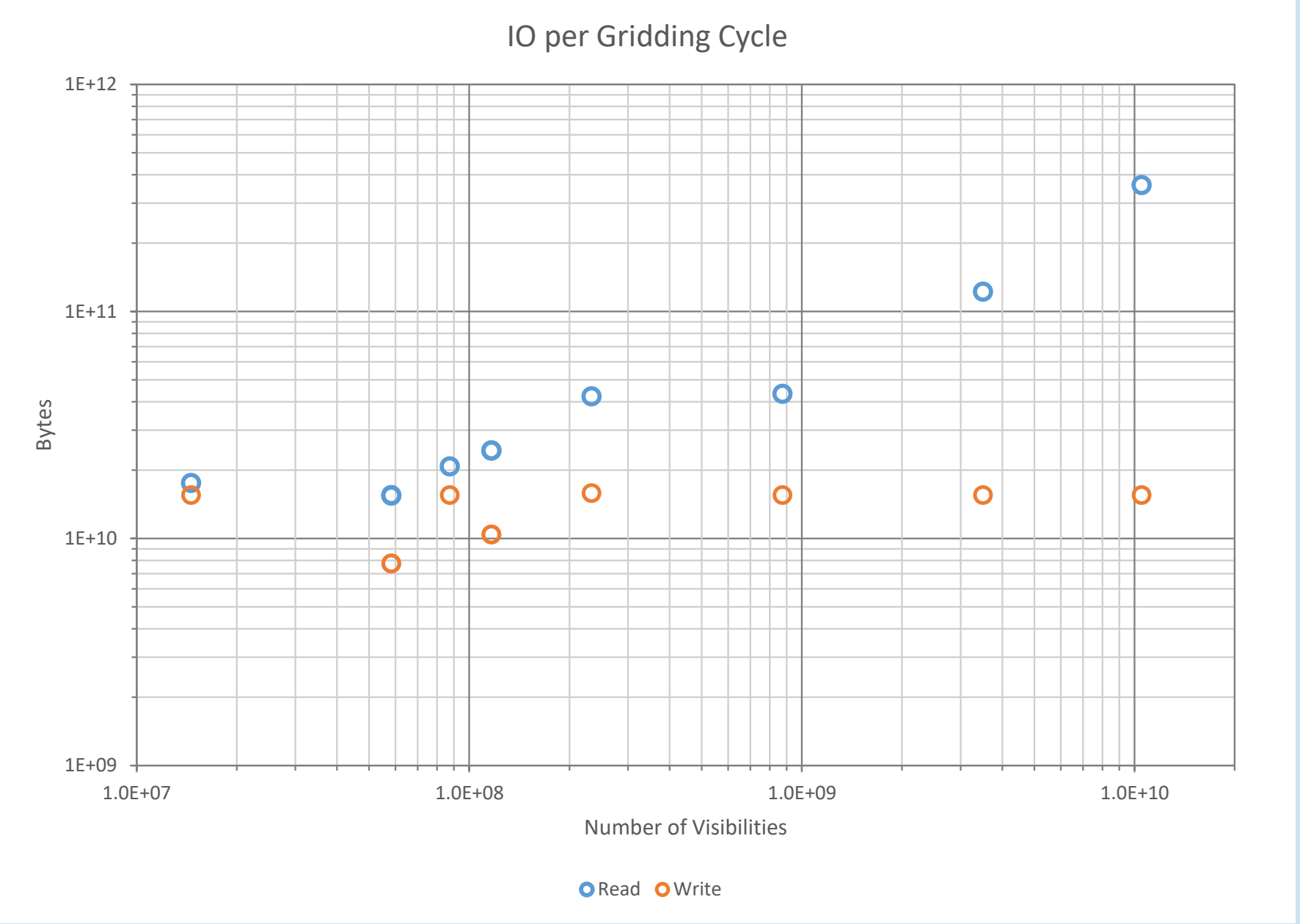

Figure 3. Input/output results from executing one gridding cycle over datasets of different sizes. These results have good accordance with theory as well. Writes are relatively constant (corresponding mostly to writing images, independent of the raw visibilities), while the slope of the reads is 32.16, a good match for 2 complex numbers for the visibilities (16 bytes) and 2 floats for the weights (8 bytes). This gives arithmetic intensities of 40 FLOPS/byte for standard, 233 FLOPS/byte for A-proj, and 670 FLOPS/byte for W-proj.

Integrating these figures with the science requirements shown in Table 1, and using a single-core efficiency of 10% (measured) and a parallelization efficiency of 90% (assumed for now), **it is estimated that the ngVLA will require a system capable of handling 61.4 PFLOPs/s, in order to process a data rate of 6898.09 Gvis/hour**. The raw data rate is 7.6 GB/s and the storage rate 20.16 PB/Month.

The National Radio Astronomy Observatory is a facility of the National Science Foundation
operated under cooperative agreement by Associated Universities, Inc.