Some Questions about the Drivers of Imaging Computational Costs

Author: Duncan Hall, Software and Computing Domain Specialist, SKA Program Development Office

Version and save time and date: Refer to document footer

1 Introduction

[1] These questions are a first pass at establishing and documenting the key factors driving the costs of computation for Phase 1 of the SKA

[2] The required imaging dynamic range (the ratio taken between the flux of a typical L-band source in a field of 100 mJy and 1σ noise level after a 1000-hr integration on a single field) for SKA Phase 1 is 65 dB

[3] Current estimates for the computational costs of imaging for the SKA are in the region of n x 100,000 floating point operations per *uv* point

[4] The number of *uv* points per second is given by the product of (i) the number of unique baselines; (ii) the number of frequency channels; (iii) the number of beams formed; and (iv) the number of correlator dumps required per second to restrict smearing to an acceptable limit, which has been defined as 2% [in the *SKA Design Reference Mission v. 1.0 1.4.3. Correlator Integration Time*]

[5] For SKA Phase 1 [[1]*Garrett et alia*], (i) is ~31,000 for dishes and is 125 for aperture arrays; (ii) for continuum observations has been set at 67,000; (iii) is 1 for dishes and is 480 for aperture arrays; and (iv) can be calculated by reference to receptor diameter, maximum baseline (200 km for Phase 1), frequency of operation, and standard formulae, e.g. [*Thompson, Moran and Swenson 6.4 Effect of Visibility Averaging*]

2 Questions – the answers to which all impact on [3]

A. What is the relationship between the DRM-defined limit of 2% in [4] above and requirement [2]? In other words, what is the provenance of the 2% limit – and is it traceable to dynamic range requirements in the DRM? Is 2% sufficient?

B. [[2]*Varbanescu et alia, 3.1.1. Data and computation*] state that "The size of the [sub region] projection, m × m, is a parameter calculated based on the radio interferometer parameters, and it is anywhere between 15 × 15 and 129 × 129 elements." To achieve requirement [2] is 129 x 129 elements sufficient?

C. [*Varbanescu et alia, 3.1.1. Data and computation*] state that the "convolution function is oversampled, i.e., its resolution is finer than that of the original sampling by a factor of *os* × *os*." Furthermore, they state that a usual oversampling rate is *os* = 8. Is *os* = 8 sufficient to achieve requirement [2]?

D. To achieve requirement [2] how many "major cycles" (in Varbanescu's terminology, "main loops") are required in a worst-case imaging task?

E. Gridding and de-gridding are quoted [e.g. [3]*Bhatnagar and Cornwell*] as consuming the largest proportion of flops per *uv* point for image processing: "For typical sized visibility databases, the gridding costs dominate over the cost of FFTs." What work has been done on the examining and improving the efficiency of algorithms for gridding and de-gridding – or perhaps otherwise dealing with irregularly spaced sampling? An example of the latter is use of radial basis functions described in e.g. [[4]*Carr et alia*].

F. Dealing with non-circularly symmetric side lobes – especially those inherent in an offset Gregorian reflector design – might pose interesting challenges for calibration and imaging. Has any empirical work been reported to achieve requirement [2] in the presence of such effects? For example, how many beam parameters are likely to be required to achieve requirement [2]? What about dealing with pointing errors?

G. Many fields are likely to contain sources that may not be able to be distinguished from imaging artefacts. How does this impact on achieving requirement [2]?

H. What work has been reported on to address "flagging" and removing observed outliers caused by any number of effects such as RFI and instrumental artefacts?

I. Further questions include:
    a. What are the key bottlenecks in data processing / data flow? [Amdahl's law]
    b. What are the required sizes of data caches?
    c. What are the required energy efficiencies of computation and data transport?
    d. Your particular research interest?

---

[1] http://arxiv.org/abs/1008.2871

[2] Scientific Programming 17 (2009) 113–134: "Building high-resolution sky images using the Cell/B.E."

[3] Astronomy and Astrophysics, Volume 426, Number 2, (November 2004) 747–754: "Scale sensitive deconvolution of interferometric images I. Adaptive Scale Pixel (Asp) decomposition"

[4] ACM, Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia (Feb 2003) 119–126: "Smooth surface reconstruction from noisy range data"