The application of high performance computing in rock art documentation and research

Introduction
This article discusses the application of high performance computing (HPC) to the remote processing of high quality image based models of rock carvings in the field, using Agisoft PhotoScan / Metashape Professional software. Most work to date has been carried out on the Danish e-Infrastructure Collaboration National High Performance Computer Abacus 2.0, hosted as part of the eScience Centre at The University of Southern Denmark (SDU). More recently, the author has been testing the cloud processing facility available within Metashape Professional, that utilizes Amazon Web Services (AWS) EC2 HPC cluster. This work forms part of the fieldwork component of the author’s PhD project at Aarhus University, Denmark, investigating regionality in Southern Tradition rock art through a programme of extensive documentation of rock carvings, principally on the island of Bornholm, Denmark, and also parts of Central and Southern Norway and Western Sweden. The successful deployment of this technology as part of a programme of surface based documentation breaks new ground in the fields of rock art research, archaeology, and high performance computing in equal measure. Therefore, the workflow discussed in the following pages has wide ranging implications for a number of disciplines.

Using a wireless router connected to the mobile phone network, files and projects can be sent to an HPC from a laptop in the field, for remote processing in an interactive interface. This takes place using a virtual network computing (VNC) connection. VNC allows screen sharing to be set up on a virtual machine operating a graphical user interface (GUI) on the HPC, allowing the process to be controlled, monitored and inspected as desired, or, run automatically as a batch job using a pre-programmed script. Results can be inspected on the virtual machine and finished products exported back to site for study of the rock surface.

This new way of working with image based modelling compliments and augments existing methods, adding another powerful tool to the rock art researcher’s arsenal. Moreover, it retains the focus of the study of rock art as an on site dialogue between the surface itself, the documentation and interpretation produced by scholars. This is achieved by reaping the benefits of drastically reduced processing times for image based models, when applying fast hardware and massively parallel computing to the application of the embarrassingly parallel problem of image based modelling (Stott et al. 2018: 207-209).

Processing on HPC via the cloud will become more accessible and more widely applied in future. Therefore, this article supplies definitions of essential key terms, concepts and processes having relevance
for our understanding of the operation of the technology for those who may not be so familiar with them.

**Why and what we document: the study of the surface**

Observations made in the field, including documentation and registration, are the building blocks of all research. It is inevitable that rock art will, sooner or later, decay to the point where nothing remains (Swantesson 2005: 58; 104). Therefore, it is imperative to strive for the highest standard of documentation possible, even though the daily reality is that perfection can never be achieved and that compromises have to be made in response to pressures of resources and time.

Rock art documentation should be based upon the raw evidence itself and subject to as little manipulation and interpretation as possible (Meijer and Dodd 2018: 291). Within rock art, this means that whole surfaces should be documented, not just individual figures (ibid). The creation of surface based documentation promotes scientific transparency by allowing the veracity of the claims made to be tested by colleagues. However, it is also a valuable tool in the field itself, as a source of information assisting in interpretation. The more sources of information available at our fingertips, the more informed our conclusions become.

Almost ever since people first attempted to record rock art, we have always been aware of the need to record and display rock art in three dimensions (Coles and Gräslund 2000: 11-89). Only recent inventions make it possible to record such levels of detail.

**Key concepts**

**What is a High Performance Computer?**

The Oxford English Dictionary defines supercomputer as “a computer, or series of computers with exceptionally high processing power or speed”. Technically speaking, no set definition exists. Confusingly, this word is sometimes be used to describe workstations, which are not the subject of this research. In terms of hardware, HPC might be differentiated from a “conventional” computer in a number of ways.

The configuration that most readers will be familiar with is a series of components assembled as one entity. An HPC consists of a collection of multiple elements, connected to one another in varying ways, often disambiguated from one another, the resources of which are managed by a resource management system.

The resource management system means that it is not necessary for jobs to be executed on one set group of components, but by different groups of components in the cluster, working in unison. A specific group of resources forming part of the greater whole is known as a node, and the whole system is referred to as a cluster. Batteries of central processing units (CPUs), local solid-state drives (SSD), flash-based random-access memory (RAM) and graphics processing units, known as GPUs (if included), comprise the nodes. These machines can be used together simultaneously for a particular task, significantly cutting the time taken. The acceleration of the task is achieved by parallelization, discussed in the next subsection.

The nodes of an HPC are connected together using switches, which are linked by a high-speed internal network, to form a cluster. Both the switching and the network can be configured and managed (by job management via the switches) in several ways according to the intended use of the HPC. The cluster is further connected to external server-based storage systems and external networks. HPCs also have complex cooling and power systems to minimize loss of performance due to heat.

**Parallelism in computing**

Both HPC and the creation of image based models employ parallel computing extensively in their workflows. Parallel-
ism in computing can be likened to the splitting of a river into multiple channels to increase flow through an area, such as within flood alleviation schemes. Sluices, weirs and reservoirs can be added to the system to further control the flow, or store water, either side of the channels. So, when the same analogy is applied to a computer, tasks can be broken up into many small jobs and executed simultaneously by multiple processors, thus speeding up the process markedly. The machine’s controller acts like a sluice, regulating the flow of tasks to the processors as well as to the system’s memory, both prior to and following computation.

In this setup, the time taken to perform the operations is largely dependent on the speed at which calculations can be processed by individual processors. Adding multiple processors increases the number of potential channels, consequently decreasing the time taken to execute a task. In HPC, these principles are extended on a very large scale. As the size of the cluster increases, the infrastructure required to operate the HPC increases accordingly.

**Image based modelling with Agisoft PhotoScan & Metashape**

Image based modelling is the construction of 3D models from 2D images. The software used by the author, Agisoft PhotoScan and Metashape Professional, uses the multiple view variant of the structure from motion implementation of photogrammetry (Ullman 1979).

The reoccurrence of points in a set of images, whose common points have previously been matched, permits the position of a 3D point to be triangulated by projecting the position of points in two, but ideally three, or more views (Cipolla and Robertson 2009: 1-4; Figure 13.5). Triangulation is used to derive the location of a 3D point, but in order to do so, it is necessary to choose a solution based on the least amount of squared error between observed and predicted positions of the pixels (ibid: 13-14; Figure 13.5).

The next stage, the most critical one in terms of the detail of a subject such as a rock surface, is dense surface reconstruction. Agisoft do not give many details of how this is performed, but it involves the computation of depth maps for each image (Semyonov 2011; Agisoft 2018b). Corresponding points within one or more images can be placed in 3D space. In this way, a dense point cloud can be created. The depth maps can also be used to generate a triangular mesh, which is achieved using nodes connected by edges and infilled with faces. Within the latest versions produced by Agisoft, it is possible to create meshes of the same density as that of the pixels in the original 2D images (Semyonov 2019).

Alternatively, still providing a level of detail equivalent to the dense cloud, digital elevation models can be generated. These can be viewed in 2D within Agisoft’s software itself, as a hill shade or slope visualisation, or, more recently, in 3D, using the plugin Qgis2threejs (Available at: https://github.com/minorua/Qgis2threejs, Accessed 2/1/19) for the open source GIS software QGIS 3, or, within QGIS’s commercial competitor, ArcScene by ESRI. Lastly, but nonetheless useful, all photos can be exported as one orthomosaic, to give an overview of the whole scene.

**Problem: the processing log jam**

As the reader may imagine, a great number of calculations need to be made, many billions of times, in order to construct a 3D model from digital photographs. As the number of photographs and the geometric complexity of the subject increases, so do the calculations and the range of possible values for any given point in the model. This computational difficulty arises as a consequence of the need to estimate location of each point in relation to not only its neighbours, but also all the others in the model. Objects with a high degree of geometric complexity are often described as possessing strong geometry, for example statues, buildings, or, in the case of rock art, large stone blocks with
rock carvings on multiple faces. The largest flatter (or planar) surfaces, also present computational problems, although the threshold at which this becomes a limitation is much higher. In order for a digital reconstruction to be made of any subject, all software used today always needs to calculate and save the position of every pixel in the model in relation to all the others.

All these factors combine to make processing of 3D models limited by computational power and storage, at all levels of the memory hierarchy (which is comprised of the caches, RAM and longer term storage systems).

Remote processing: a brief history
Image based modelling has hitherto been considered a very portable solution for the documentation of archaeological remains, but one whose results are redundant by the time they are available for use. Long processing times and a lack of sufficient computational power available for local processing have been seen as major shortcomings of image based modelling, both as a method of documentation and as an investigative method in the field (Bertelson et al. 2015: 7; Meijer and Dodd 2018: 294). When local processing is conducted away from site, full advantage cannot be taken of the information captured until one has lost the possibility to directly influence the course of investigations and the basis of interpretations. The model should not be a passive product but an actively used tool.

More generally, significant reductions in processing times have been made possible since the advent of processors specially designed for the processing of graphics: GPUs (graphics processing units). The architecture of GPUs is more suited to the processing of multiple repetitive actions than the CPUs (central processing units) (Krewell 2009). HPC systems expand GPU processing on a massive scale, using special designed units. At present, these consist of Nvidia’s Tesla series.

Andresen and colleagues (Stott et al. 2018) recognized the possibilities of cluster processing of image based models on HPC, but saw that the major drawback preventing the technique realizing its full potential was transferring the images to the HPC. Therefore, they decided to test connecting to an HPC cluster using a 4G wireless router and conducting processing remotely from the field.

In the case of the excavation of the archaeological soil layers and artefacts investigated by Stott et al., the ground sampling distance (GSD) did not require sub-centimetre precision, and therefore images could be downsized. However, as Stott et al. pointed out, this is unsuitable for faint inscriptions, i.e. rock art (ibid: 208). Thus, informative though the experiments are, processing times experienced in the author’s project are significantly longer than experienced by Andresen and colleagues, due to the images not having been resized. In rock art, the much smaller variations in depth on the sub-millimetre scale, and subtlety of textures, mean that a similar approach is impossible if one wishes to acquire a high quality result. Resizing would doubtless lead to information about the surface not being identified, as the most subtle variations would be smoothed out.

When comparing the results of the tests conducted by Stott et al. on two small datasets, containing 29 and 183 images respectively, with the results of the author’s experiences of models comprising over 1000 images, a number of additional observations can be made concerning performance. The modal number of cameras within the author’s work is approximately 160, although it is not unusual for between 200 and 400 images to be necessary to create a model of many of the medium sized panels, or free-standing stones. Large surfaces range between 600 and 1500 cameras (photographs). The greater the number of cameras, the greater the gains that HPC can provide - Amdahl’s law (Amdahl 2007 reprinted from 1967) com-
ing into effect at a higher threshold. There were already indications regarding the operation of Amdahl's Law in the HPC environment in the original tests, although it was thought by Stott et al. that the gains made would plateau and stay somewhat constant. This was not an illogical conclusion based on observation of the plateaus Stott et al. experienced after adding more than a certain number of nodes to the job.

From observations made by Emiliano Molinaro (pers. comm., 2018), research developer at the SDU eScience Centre, it appears that only the alignment stage is fully parallelized across all available GPUs and / or CPUs. During dense cloud generation, the creation of the first stage of the process (the depth maps) only uses one GPU for processing (although the author has observed that the other GPU does seem to be used for writing and reading of certain data). Also, the author has found, during further testing, that GPUs are not used during mesh generation and texture mapping. All CPU’s are used to generate the mesh, whilst only 1 CPU core is utilized to decimate the mesh and generate texture. This explains why Stott et al. found that there is little discernable benefit in performance when operating on multiple nodes for these two stages.

### Description of equipment and workflow

#### Field setup

Figures 1 and 2 illustrate the workflow and setup in the field. Images captured on a camera are downloaded to a laptop, ideally flash based for the sake of robustness. The laptop is connected to a wireless router supporting speeds up to 4G, which is usually connected to an antenna, essential for increasing reception and transmission speeds during the upload of photographs and download of exported products from the model. Transfer of files and control of processing on the HPC takes place over a Secure Shell (SSH) con-
nection (public key pair and password). The photos are uploaded to the GPFS (IBM General Parallel File System, now known as IBM Spectrum Scale) of the HPC cluster. There are many ways of doing this, however the author, for reasons of familiarity and ease of setup, uses SFTP within the open source program Cyberduck (https://cyberduck.io).

An interactive session is initiated from the command line on a processing node. Compute nodes containing GPU nodes have been used for the majority of the work. Initially, the alignment is run and then pulled back down from the HPC for verification. Key points and tie points are saved by default to assist the addition of photos, should it be found necessary. On the laptop, coverage is checked and markers are added to the scale rulers included in the scene. These are given false co-ordinates, thus permitting scaling of the model. The remaining processing stages are then run automatically as a batch job, saved at each stage. This includes retention of the depth maps, which eases memory pressure and decreases the time taken for a rerun should the process fail for some reason. This course of action is only possible due to storage space not being a problem on Abacus 2.0, and follows the advice given by email to the author by Diana Ovod (2018), Technical Analyst, Agisoft LLC.

A recent refinement to the workflow has been to run dense cloud generation, mesh, and texture generation for larger models on compute nodes without GPUs. Instead, compute nodes with increased RAM capacity and local SSD storage have been utilized. This has been found to be a good a solution for the larger and more complex models that exceed the standard configuration of GPU nodes currently used. Once again, thanks are extended to Emiliano Molinaro for suggesting this improvement.
Following the completion of processing, the mesh of the model and any other products needed for study on the panel are exported back to the user in the field. The laptop can be taken around site and placed on the rock (with a little help from a laptop stand from IKEA incorporating rubber feet) and the model used in an ongoing dialogue to actively assist in the study of the surface, and vice versa. Visualisations can be applied using open source applications to the model in 3D, in Meshlab (such as directional lighting or radiance scaling), or in 2D, as a raster in QGIS.

Description of HPC clusters used

**Danish e-Infrastructure Collaboration National HPC: Abacus 2.0**

As part of the Danish e-Infrastructure Collaboration’s (DeIC) activities to encourage the development of Denmark as an eScience nation (www.deic.dk), a number of supercomputing facilities have been established. The largest of these is Abacus 2.0 which is the most powerful HPC cluster currently available for research purposes in Scandinavia, hosted at the DeIC National HPC Centre. The Centre’s partners comprise: University of Southern Denmark (SDU), Aarhus University (AU) and Aalborg University (AAU). The facility is located at the SDU eScience centre in Odense.

Abacus 2.0 is comprised of a number of elements: compute nodes, system and network storage. Figure 3 illustrates the arrangement and nature of connections between the compute nodes. For the purposes of advanced graphic processing, GPU nodes have been used as the main workhorses during the project, with the greater on board memory of fat nodes being used for generating the meshes of the largest and most complex models. The full specification of Abacus 2.0 can be found at on the eScience Centre, SDU webpage (https://escience.sdu.dk/index.php/hpc/ Accessed 21/12/2018). Figure 4 summarizes the key stats for GPU and fat nodes on Abacus, as well as the key details of the hardware that has been used in other experiments for comparison, on Amazon’s EC2 cluster. The system runs on the CentOS Linux 7 operating system (Stott et al. 2018), that can be easily exploited by the CUDA API. The allocation of resources is controlled using SLURM workload manager (https://slurm.schedmd.com Accessed 2/1/2019).

**Amazon Web Services (AWS) EC2**

Details are somewhat scarce concerning the setup of AWS’s EC2 instances. The information known concerning the specification of the hardware is listed in Figure 4. Not all of AWS’s data centres contain the hardware to offer all services. In the case of GPU nodes, it is believed that processing of models in Europe currently takes place either in the Frankfurt or Ireland region of AWS.
Evaluation of the setup and workflow

Workflow
The technological possibilities offered by use of the mobile telephony network to transfer files, coupled with VNC connections, allow files to be uploaded to HPC clusters, processed remotely in an interactive session and the exported products sent back to the field. In the case of small and medium sized rocks, the model is available for in the field study after 3 to 4
Figure 6: Orthophoto of part of Bardal 1, Steinkjer, Trøndelag. Processed on the DeIC HPC, Abacus 2.0. J. Dodd in collaboration with NTNU Vitenskapsmuseet. Detail of Figure 6 shown below.
hours. When documenting larger surfaces, the alignment can be checked in the field to ensure that data acquisition has been successful.

Running several concurrent jobs maximises the benefits offered by the technology, optimises use of resources and exploits optimal lighting conditions to the maximum. The latter factor is particularly important, as strong lighting has been found to lead to an 89% increase in the number of points in the sparse point cloud, which can be further increased if the photos are taken with lighting from a low angle (between 25 - 40 degrees for a flat surface). Further efficiencies can be made by pre-planning and selecting the rocks to be documented based on consideration of when the angle for the most optimal lighting might be experienced, the expected size of the model (based on the number of figures and / or the total surface area) together with the proximity of panels relative to each other.

In this optimal scenario, conducted under sole working conditions, as the first model is processing, the second one is prepared and photographed. The photographs for the second are sent to the server, by which time the first model has completed alignment. The coverage is checked and more images added if necessary. Then, these can be aligned using incremental alignment performed locally. Following this stage, the rotation and scaling of the model can be set. This involves placing, checking and adjusting the projection as necessary of the location of ground control points placed on scale bars included in the scene. This is followed by an optimization of the alignment to take account of non-linear distortion (Agisoft 2018a: 58). Next, incremental image alignment is run, if necessary for a second time. Rerunning the alignment has the useful effect of recalculating camera locations in light of the a priori information fed into the model during scaling. These two steps can potentially increase the precision of the model. Finally, the revised files are then

Figure 7: Slope visualisation in Metashape of the upper surface of Store Almegård 1, South West Bornholm. Some of the long, elongated depressions may be foot-soles. Processed on the DeIC HPC, Abacus 2.0. J. Dodd, in collaboration with Bornholms Museum.
uploaded to the server and the remaining processes run as a batch job, using a saved xml script.

The processes are repeated and more nodes used until the processes running on the first rock is finished. To balance the use of bandwidth and time, the mesh (with vertex colour without texture), or the DEM, is exported back to site. Meshes are exported in Stanford PLY, due to its compactness, and GEOTIFF for the DEM, due to the ease of import into the author’s GIS platform: QGIS.

Let us consider the operation of this methodology within an example drawn from practice. During the documentation of the concentrated locality of Hammersholm, North Bornholm, with clear skies and full sunlight during early June, 8 medium sized models (ranging from 150-300 cameras) were taken during a 7 hour period between 1400 and 2100. Each model was processed using 1 GPU node, using up to 5 nodes running concurrently. As a result of combining such effective working with the power of Abacus, all processing could be completed by midnight, with the meshes of 6 models processed and exported back to the field for on-site use.

**Software optimization**

*Incremental image alignment*

Providing that the checkbox to store key and tie points is ticked before commencing alignment, versions 1.4 and onward of PhotoScan / Metashape permit additional images to be added to the model and processed separately should the results of alignment indicate dissatisfactory coverage, or if automatic alignment fails. Manual placement of target points can be used to feed a priori information into the alignment process, which can force alignment to succeed.

*Saving of depth maps*

For the generation of large models in high quality, memory pressure should be considered. As outlined in the section describing setup and workflow, depth maps should be saved if sufficient storage is available.

**Democratization**

Previously, access to cluster and network processing has been restricted. The launch, in October 2018, of cluster processing on AWS directly from the Metashape Pro GUI has drastically changed this situation, giving many more the possibility of running on fast hardware than has previously been the case. The emergence of this capability has had the effect of shifting the emphasis of limiting factors away from the hardware, towards the connections with the cluster.
**Challenges**
The sub millimetre detail required for the study of rock art, and the large size and complex geometry of the materials it is found upon, will always mean that faster hardware will deliver greater returns. The first challenge is to get the data to the cluster, the second to handle the data as fast as possible, and then be able to view the results in a useful manner.

The number of connections and capacity of the internet between the user and the cluster influences data transfer times. Bandwidth is just as much a limiting factor as the calculation and storage capacity of the hardware used during processing. Resource management by the operators of both the mobile telephony network and the internet is highly desirable, if not essential.

**Conclusion**
This paper has outlined the application of cluster processing on HPC within the documentation of Scandinavian rock art. The use of a WLAN router with a laptop to send photographs and receive processed models in return, enables the benefits of processing image based models on HPC Linux clusters to be fully harnessed. The possibility to have the model available for study in the field, on the rock itself, is a major step forward, as the model has become a powerful tool, actively assisting our abilities to study the surface and augmenting existing methods at our disposal. The use of as many sources of surface based documentation as possible, in combination with one another, is accepted professional practice within rock art studies, as no single technique exists that can capture all the details of the surface and the surrounding environment (Meijer and Dodd 2018).

The methodology outlined in this paper is the first study of its kind to employ HPC out in the field as part of an extensive, systematic programme of high quality 3D documentation. The findings have significant implications for a range of scientific disciplines far beyond rock art research. In future, we can expect the use of HPC to grow within rock art documentation and for processing and workflow to undergo further optimization and extensive development.

*James Dodd*
*PhD Scholar, Department of Archaeology and Heritage Studies, Aarhus University.*
*james.dodd@cas.au.dk*

**Note**

1 The software changed name in October 2018. In this article, to avoid confusion, PhotoScan refers to results processed on version 1.4.3 and lower. Metashape refers to results processed on version 1.5 and onward. Models processed on Abacus 2.0 use version 1.2.4 of PhotoScan. The possibility to test newer versions of Agisoft’s software on the Abacus setup has not been made available to the author.

**References**


Painted images of Tepsej I rock art site and some questions on their chronology

Key words: rock art; Minusa basin; Tepsej, painted rock art images; the earliest petroglyphic stratum of Minusa basin; chronology of rock art.