PERFORMANCE ANALYSIS OF DISTRIBUTED ORTHO GENERATION USING ERDAS ORTHO ACCELERATOR (LOA)

Yasemin Kuzu-Sinram

ERDAS, Inc. 5051 Peachtree Corners Circle, Norcross, Georgia 30092-2500 - yasemin.sinram@erdas.com

Commission IV, WG IV/3

KEY WORDS: Distributed, Automation, Processing, Digital, Orthorectification, Performance

ABSTRACT:

The demand for digital ortho imagery is rapidly increasing with the growth of Internet and GIS applications. Orthophotos are a valuable source of information for government agencies, engineers, planners, land managers and individuals. Orthophotos are required to be updated periodically and the delivery times are decreasing rapidly. The size and the number of the images increase with the demand for more detailed and accurate imagery. Since ortho rectification projects are becoming very data heavy, ortho production is a bottleneck for photogrammetric shops due to huge processing times which is often too high for single processor architectures.In this paper we are examining the benefits of the ERDAS Ortho Accelerator by analyzing the performance of the distributed processing. In this experiment, we have chosen large orthorectification jobs to show the time savings of distributed processing. Distributed orthorectification processing performance metrics are analyzed with respect to the number of processing nodes applied. The results are also compared to stand-alone processing performance. This paper shows distributed processing is necessary with large photogrammetric projects as processing time can decrease significantly with the number of nodes used.

1. INTRODUCTION

Adoption rates of distributed processing in the photogrammetry industry have not been universal. Photogrammetric processing is largely workstation-based. With the advances in computer and network hardware, current inexpensive computer systems can handle larger amounts of data and computer prices are getting relatively cheaper. In order to generate orthos in reduced delivery times demanded by the market, photogrammetric workflows can use automated procedures such as distributed computing. Distributed computing can be defined as the ability to divide large processing jobs into smaller tasks and running them on multiple processors or multiple machines. Distributed processing harnesses the processing power of several CPU nodes to increase throughput. It results in greater throughput and frees up operators to perform other tasks. Also, purchasing nodes instead of softcopy workstations will result in cost savings.

ERDAS Ortho Accelerator (LOA) is developed under a strategic development agreement between Leica Geosystems Geospatial Imaging (now ERDAS) and GeoCue Corporation (formerly NIIRS10). LOA was developed to speed up the rate with which digital orthophotos are produced. LOA integrates ERDAS photogrammetric processing components into the GeoCue geospatial process management framework to create a very efficient, enterprise-enabled orthophoto production environment. LOA is an extension (CuePac) to GeoCue. A CuePac is a collection of GeoCue menus, checklist and auxiliary programs that implement a "canned" set of workflows.

GeoCue is not workstation centric. The command dispatch system allows distributing ortho jobs to remote nodes. With orthorectification and mosaicking capabilities, ERDAS Ortho Accelerator takes advantage of the distributed and scheduled workflow processing capabilities as well as process management tools provided by GeoCue.

LOA provides below solutions to several market problems which exist in orthophoto production processes:

- Orthorectification jobs are distributed to individual machines (ortho nodes).
- Dispatched tasks can be run on a machine other than the one on which it was launched
- Orthorectification and mosaicing jobs can be scheduled to run at user defined times
- Multi user access to the same project from any workstation in the network.
- Accurate and real-time monitoring of scheduled tasks.
- Each manipulation of the data is incrementally saved via transaction processing against the database.
- The system is protected from unauthorized access to data
- User-defined metadata can be associated with the imagery

LOA consists of 3 modules which are: Project Importer, Ortho CuePac and Mosaic CuePac. LOA is built on LPS and ERDAS mosaicing components.

Project Importer enables introducing a photogrammetric project into the GeoCue managed workflow. LOA supports LPS block files, BAE SOCET SET® projects, INPHO MATCH-AT projects and Intergraph Z/I Imaging® and ImageStation® Automatic Triangulation (ISAT) projects. When a photogrammetric project is imported, graphical objects are created in GeoCue with geographically correct footprints and locations in the Map View. Also, metadata is attached to each entity that includes fields such as interior and exterior orientation, image size, camera information and so fort.

Ortho CuePac is an enterprise ortho production system that is

based on LPS technology. ERDAS ortho tools are launched automatically within LOA. Ortho CuePac uses GeoCue to define production steps. Various orthorectification parameters can be defined as a part of the workflow. Distributed processing is also an important aspect of Ortho CuePac. Users can run large orthorectification jobs on different machines in order to reduce overall processing time. Orthorectification jobs can be submitted from different machines. LOA sorts them in a queue, allows users to choose which machine and when to run jobs. Ortho jobs can be scheduled at a later time. Data locking mechanism prevents data collisions in a multi-user environment. On client machines users can see the real-time status updated of the submitted ortho jobs. Users are informed when the jobs are completed.

Mosaic CuePac is a GeoCue environment that integrates ERDAS MosaicPro into the GeoCue geospatial process management framework. LOA launches ERDAS MosaicPro as an external process to generate seam lines, perform color adjustment and decide on parameters as feathering and smoothing options. ERDAS MosaicPro improves the seam editing portion of mosaicing as well as offering additional radiometric adjustment capabilities. Figure 1 shows orthophotos generated on LOA.
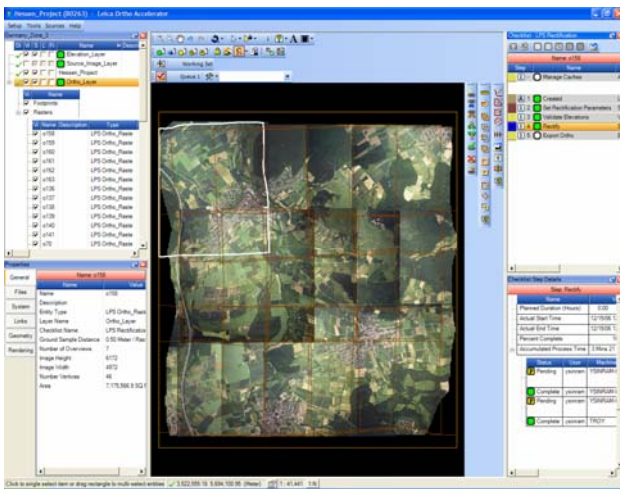

Figure 1: An LOA project showing ortho images of Hessen, Germany. DOP, (c) HVBG 2007

2. SYSTEM SETUP

2.1 System Requirements

In this experiment, the LOA system setup consists of a Server machine, a client machine and 5 ortho rectification nodes

The GeoCue Server is the heart of the system providing the central control and management functions. This is the most important machine, which matches jobs with selected machines. GeoCue uses a database for storage of project metadata (data about data), project data and the configuration information that defines a workflow environment. High density data such as image and elevation data are stored in flat files in distributed data storage within the network connected file system. GeoCue maintains references to these externally stored data.
GeoCue Server can run with MSDE or with full SQL Server installation. MSDE limitation is 5 nodes. If more nodes will be used full SQL server should be installed. In this setup, the server is running with MSDE.
In this experiment, a Pentium 4, 2.40 GHz with 1.00 GB of RAM running Windows XP is used as the Server machine. During the performance analysis no jobs has been run on GeoCue server.
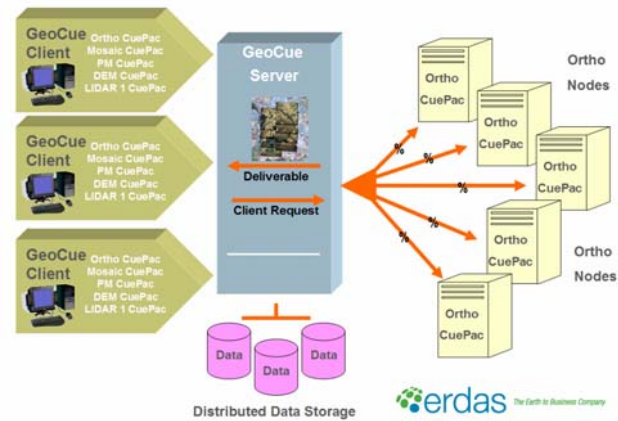

Figure 2. System Components

The GeoCue Client is the machine that submits ortho or mosaic jobs to the nodes. The project status may also be tracked on these machines. Client machines are the production workstations. There should be only one Client Machine per human operator dedicated to LOA. Depending on the production needs several Client machines can be used simultaneously. The GeoCue Client is a window into the project while the data processing is performed on the ortho nodes. In this study one workstation is used as the Client machine. The ortho jobs have been submitted using client machine.
In this experiment the GeoCue Server and the Client are the same machine.

Ortho Nodes are the workers that perform the computation. Ortho nodes can be any system on the network. They can be installed on a production machine not in use or a dedicated node in a server room. However, ortho nodes are recommended to be dedicated to ortho processing only whereas on client machines operators can run alternative processes.
Ortho nodes do not have a user interface, but they report feedback and success or failure information back to the client machine. The number of the ortho nodes is highly scaleable depending on the production needs.
In this experiment, system setup consists of 5 ortho nodes. Each node has the same specification: Pentium 4, 3.20 GHz with 2.00 GB of RAM running Windows XP. No other jobs were run during the performance experiment on ortho nodes.

Networking: Network is the connection between the GeoCue Server, client worksations, ortho nodes and the data storage locations. The installation is done on a network domain. The machines are connected in a local area network. Interconnection between the nodes, client machine and the file server was done using gigabit Ethernet. Gigabit bandwidth refers the data rate supported by the network connection. The performance will be greater when larger bandwidth is used. However, bandwidth is not the only factor that affects the speed. Network traffic might cause latency in the transmission. Also addition of the too many

nodes might reach the point of diminishing returns.

This experiment was not done in a dedicated network, therefore some network traffic and bottlenecks due to I/O has been examined and reported in the result summary.

2.2 LOA Setup

The work required to install LOA is minimal and IT expertise is not required. The user needs to select the Server machine that will host GeoCue Server application, the workstations that will host GeoCue Client application and ortho nodes that will run the jobs. The following steps explain how to setup LOA.

- Install GeoCue Server software on the server machine.
- Install MSDE if Server machine is not running with full SQL Server.
- Run database manager and install GeoCue database
- Select Common Folder and share it
- Decide data storage locations (GeoCue WareHouse) and share them
- Import Licenses
- Install LPS and GeoCue Coordinate systems
- Install GeoCue Client software on the production workstations and ortho nodes and configure it.
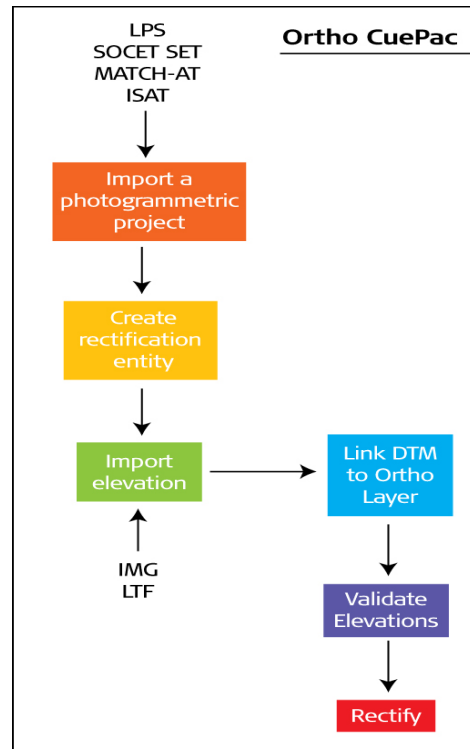- Install LOA on Client workstations and ortho nodes
- Install LPS support modules

3.   ORTHORECTIFICATION WORKFLOW IN LOA

LOA workflow starts with setting up a project. Then a photogrammetric project (e.g. LPS block file) is imported into GeoCue. The photogrammetric project should be a triangulated block. The next step is to generate rectification entity. In this step ortho layer will be generated and rectification parameters such as output pixel size, resampling method, projection information will be defined. An elevation file then be imported and linked to ortho layer. Multiple elevation files can be imported. The linking step will establish the elevation files to be used for rectifying each image. Validating elevations will update the footprint of the ortho entities by intersecting ortho entities with the linked elevation files. This step will invoke the dispatch dialog. Validate elevations step can be on several nodes.

Once the validate elevation step has been completed, the next step is orthorectification. On dispatch dialog, the user can select the machines to run orthorectifaction jobs. The status of the jobs can be tracked on dispatch monitor. The dispatch monitor displays jobs that are pending, in progress and completed. Once the rectification is finished on any of the nodes the raster images will be displayed on the map view. A typical distributed run is shown in Figure 6.

Figure 3. Ortho CuePac workflow chart

4.   PRACTICAL RESULTS

For this experiment, two triangulated blocks with **240** and **858** frame images have been selected to show the timesavings as a result of using the distributed processing in LOA. For further referencing, these blocks will be referred to as Block 240 and Block 858 respectively.

Source Images
Both blocks contain TIFF format 8 bit frame images and they are **962 MB** each. File caching was enabled during these tests. The source images were stored in the file server and then copied to each node's caching folder to perform orthorectification. Once orthorectification is complete, the ortho images were then transferred to the GeoCue WareHouse on the file server.

DTM used in Orthorectification
The DTM was physically loaded onto each workstation for each orthorectification process. The DTM is copied from the server to the local hard drive of the node each time an image is orthorectified. The **10m** cell size DTM is **27 MB** in size. This means that in addition to the source images, 27 MB of DTM is also copied from the server to each node's caching folder.
Table 1 summarizes the characteristic of the project.

	Block 240	Block 858
Capture GSD (m)	0.11	0.11
Output GSD (m)	0.15	0.15
Forward Overlap (%)	60	60
Source Data Size (GB)	226	806
Project Area (km^2)	91	735

Table 1: DataSet Summary

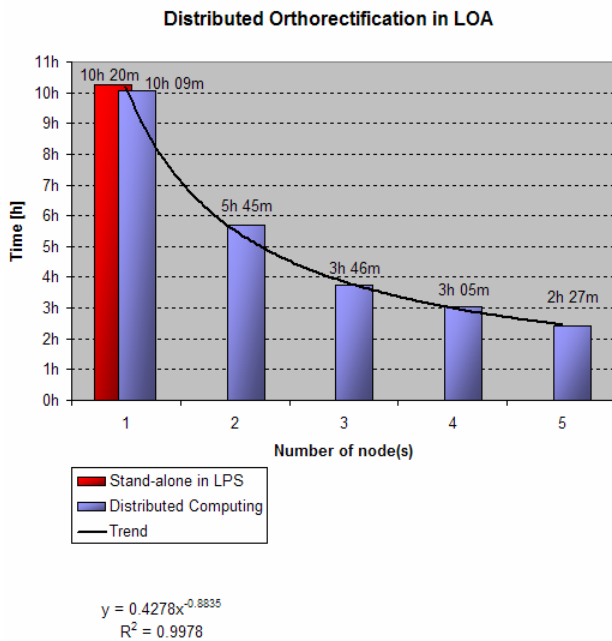Figure 4 and Figure 5 depicts the time spent for distributed orthorectification using several numbers of nodes.

$$y = 0.4278x^{-0.8835}$$
$$R^2 = 0.9978$$

Figure 4. Orthorectification time in hours and minutes for 240 images

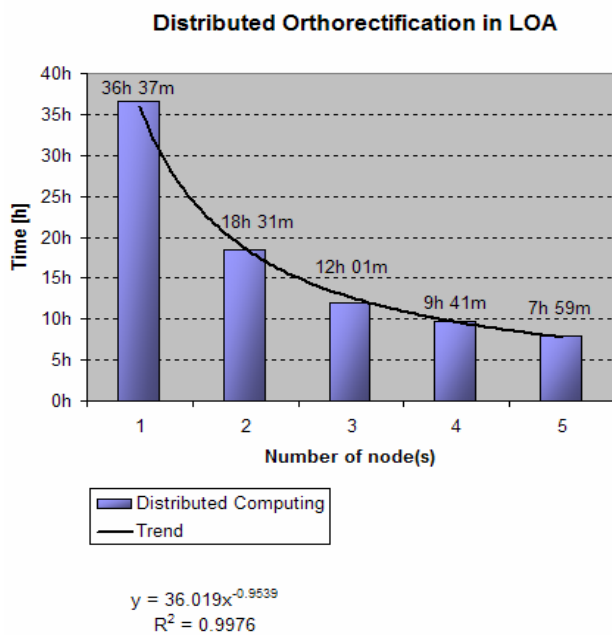
$$y = 36.019x^{-0.9539}$$
$$R^2 = 0.9976$$

Figure 5: Orthorectification time in hours and minutes for 858 images.

For Block 240, orthorectification in stand-alone mode using LPS was performed on one of the ortho nodes and took 10 hours 20 minutes. The source data is located on the file server and for stand-alone mode the ortho images created were written to the local hard drive. Orthorectification in the distributed computing environment using 1 node took 10 hours 09 minutes. Using 5 nodes accelerated the orthorectification time by factor 4.2 to 2 hours 27 minutes.

For Block 858, orthorectification using **1 node** took **1 day 12 hours 37 minutes**. Using **5 nodes** decreased the processing time by a factor of **4.6** to **7 hours 59 minutes.**
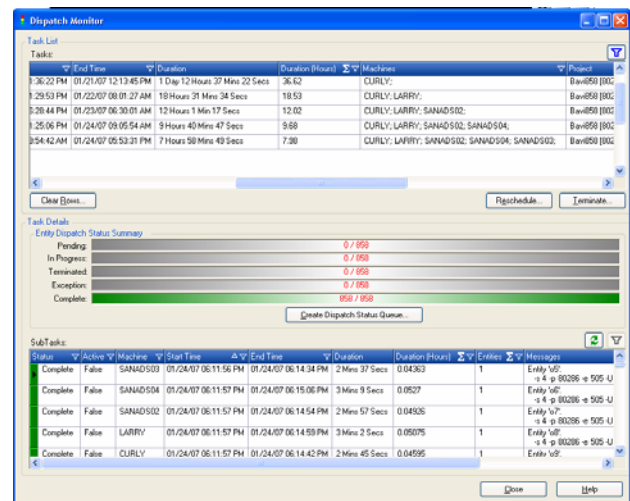

Figure 6: Dispatch Monitor

5. RESULTS SUMMARY

This experiment showed that distributed processing can be very beneficial when working with large orthorectification projects. As could be seen in **Figure 4** and **Figure 5**, processing time decreases significantly with the number of nodes used. It is clear that distributed processing will be more efficient with larger ortho projects.

Orthorectification is accelerated by a factor of 4.2 to 4.6 times using a 5 node cluster environment.

When file caching is turned off for the 5 node run, it slowed down the process by only 3%. File caching might be necessary in order to prevent file writing errors.

The system setup used for this experiment is sufficient. However it is not the ideal setup because the experiment was not done in a dedicated network. The ideal setup would be all ortho nodes; server and client machines and the data storage are in a dedicated network. This will prevent the network traffic that has been examined during this experiment. Also, the results would be faster and more reliable. However bottle-necks can always happen even in a dedicated network.

As a future studies, this experiment will be repeated using a dedicated network.

REFERENCES

Graham, L. 2006. Enterprise Geospatial Production Revisited, Photogrammetric engineering and remote sensing, vol. 72, n°5, pp. 486-492

ERDAS Ortho Accelerator,
http://www.erdas.com/products/Product_Rdr.aspx?CURRENTI D=114

GeoCue Corporation, http://www.geocue.com/