11.9 HARNESSING THE SPARE COMPUTING POWER OF DESKTOP PERSONAL COMPUTERS FOR IMPROVED SATELLITE DATA PROCESSING AND TECHNOLOGY TRANSITION

Ingrid C. Guch[*1], A. S. Jones[2], R. R. Ferraro[3], M. J. Kane[1], C. W. Karlburg[1], and S. Q. Kidder[2]
[1]NOAA/NESDIS Office of Satellite Data Processing and Distribution, Suitland, MD
[2]Colorado State University / Cooperative Institute for Research in the Atmosphere (CIRA), Fort Collins, CO
[3]NOAA/NESDIS Office of Research and Applications, Camp Springs, MD

1. INTRODUCTION

A system has been created to generate mapped satellite imagery for NESDIS operational customers that can be run on any collection of NOAA Office Personal Computers (PCs) running Windows 2000. In addition, the system can be modified experimentally by more sophisticated users who have software compilers on their office PCs. The system is composed of a single application that fits on a standard CD-ROM and can generate a variety of imagery data derived from microwave, visible and IR data from NOAA and DMSP Polar-orbiting satellites. It has been installed on dedicated PCs for operations as well as the office PCs of maintenance personnel, and research collaborators. The maintenance personnel PCs are used to process parallel development data streams during non-working hours. The benefit of this system is that underutilized NOAA Office PCs can process satellite data streams in support of research, development and operations. If successfully and securely implemented on larger scales in the future this could dramatically increase the computing power NOAA currently has available to process satellite data, particularly considering current and planned connections of NOAA to the Next Generation Internet.

The work described here is based on the Data Processing Error and Analysis System (DPEAS), a parallel data processing system for cross-sensor satellite data merger and analysis (Jones and Vonder Haar, 2002). DPEAS was created with support from the Department of Defense and NOAA. It takes advantage of NASA HDF-EOS format and technology. This paper summarizes the setup and various benefits of the system at NOAA/NESDIS related to accessing the unused computing cycles of NOAA office PCs for processing satellite data in near real-time.

2. SETUP

Ten NOAA Office PCs were used for this initial study. Two of them were enhanced with extra memory (512mb extra) and processing speed (doubled) as well as extra software (Windows 2000 Server, Batch Job Server, FTP Control). These PCs were considered the operational component and the server for the worker-node PCs. There are 5 worker-node PCs. Each is a regular NOAA Office PC with a batch job server and FTP control software added. These process satellite data during nights and weekends (when they would normally be turned off). At these times, they can be part of the operational processing loop or they can process using experimental algorithms or techniques. There are 3 developer PCs. These are regular NOAA Office PCs with programming software installed (FORTRAN, C/C++). Developers on these PCs can run and change the DPEAS executable manually but are not part of the operational processing loop.

The most significant extra cost for NOAA/NESDIS on this project was for the batch job server (~$500 each for the 2 operational machines and 5 worker nodes). To truly save money, this work must be done on a larger scale to take advantage of site license prices.

3. NEW SATELLITE DATA PROCESSING CAPABILITIES

Three of the largest communities using satellite products from NESDIS are climatologists, weather forecasters and weather modelers. Frequently these communities have conflicting needs for products. Climatologists may want products from an older algorithm in order to minimize variables in trending charts that began many years before. Weather forecasters may want products from the latest algorithm in order to have the most accurate information possible to predict the next storm. Weather modelers may want no products at all, just the algorithm to put into their model. NESDIS satisfies the first two groups by running product streams in parallel for a set amount of time (on the order of weeks to months) before switching to the latest algorithm. It usually has not been possible to run the parallel system for more than 1-2 months due to computing limitations, although users desire a full year to assess the impact of the algorithm change in each of the four seasons. NESDIS must work with the third customer group (modelers) very closely to ensure the algorithm is used properly in the model and the output product is as expected. Quality control of the product can be very difficult as it is sometimes unclear what problems are

* *Corresponding author address:* Ingrid C. Guch, NOAA/NESDIS/OSDPD SP13 RM 2322 FB4; 5200 Auth Rd Suitland MD 20746-4304; e-mail: Ingrid.Guch@noaa.gov

related to the NESDIS algorithm and what problems are related to the model attributes.

With DPEAS, the computing limitations to running parallel algorithms are avoided by accessing the unused computing cycles of NOAA Office PCs during evenings, weekends and other off-hours. Modelers are presented with the NESDIS operational product software on a CD-ROM along with the algorithm so they can easily compare the product output from the model to the product output from the operational software.

4. IMPROVED TECHNOLOGY TRANSITION

Traditionally at NESDIS, research for new satellite products begins at the Office of Research and Applications (ORA) and the associated Cooperative Institutes (e.g., CIRA). They frequently have access to current operational software that runs at the Office of Satellite Data Processing and Distribution (OSDPD) and modify it to meet their research needs and associated computer platforms. The results of the research are presented to potential users (e.g., National Weather Service) at scientific conferences, workshops and via imagery on the Internet. Requests for NESDIS to produce the resultant product operationally (meaning that the product is supported 24x7) are forwarded to the NESDIS Satellite Products and Services Review Board (SPSRB). If approved, the SPSRB and associated oversight panels work to satisfy the request. Software must be rewritten for the operational environment so that personnel at OSDPD can effectively run and maintain it. The amount of rewriting necessary differs depending on how closely the researchers were working with OSDPD at the time of development. Collaboration beginning at the earliest concept phase is encouraged but not always possible.

With the DPEAS system, ORA and CIRA researchers can obtain the operational system running at OSDPD on a CD-ROM and immediately run it on their NOAA office PC. The researchers can modify and run their own version on their PC. They can give their version on a CD-ROM back to OSDPD, who in turn can immediately (if desired) run the new software operationally. The re-engineering step can be eliminated because all parties are using the same platforms and the software is extremely portable. The DPEAS system is a single executable. While the re-engineering step can theoretically be eliminated (figure 1), NESDIS has found that minor improvements to the software are necessary (improved comments and streamlining) to ensure maintenance personnel can quickly alter the software when the capabilities of satellites and/or sensors change.

5. SECURITY ISSUES

A secure system is critical to successful data processing at NOAA/NESDIS. DPEAS is a peer-to-peer (P2P) distributed application with centralized management. Parallel processing is temporally dynamic with worker nodes entering and exiting the processing

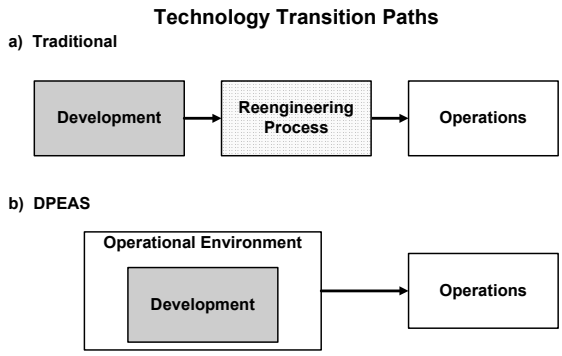

Figure 1. DPEAS modifies the technology transition from a) a traditional approach to b) one that merges the operational and development environments. This simplifies the technology transition path by eliminating the reengineering process of the traditional technology transition approach.

cluster at will. The primary coordinating node of the cluster can also migrate from node to node automatically. There is significant interaction between worker nodes (any number of NOAA Office PCs) and the processing clusters (2 dedicated Windows 2000 Server PCs). The processing clusters must also have access to the input satellite data stream. Each interaction must be secure. Currently this is performed at the domain-level via the OS file system security mechanisms. DPEAS' behavior conforms to the P2P Working Group definition (http://www.peer-to-peerwg.org). Additional groups are also active in facilitating new P2P security standards (Jones and Vonder Haar, 2002). For example, new approaches based-on XML and SOAP using web services would enable the system to more easily scale across the web to fully exploit the new NGI capabilities. During this exploratory stage, the NOAA Office PCs running DPEAS do not link up to all of the NOAA Office PCs on the Local Area Network. This allows for a more controlled and thorough analysis of the full security implications of the P2P approach within the NOAA Office environment.

6. CONCLUSIONS

DPEAS can be used to tap into NOAA Office PCs during non-working hours to provide a large amount of CPU time for satellite data processing. It is an example of a multi-agency (DoD/NOAA/NASA) application that is currently enhancing the ability of NOAA/NESDIS to meet customer needs from various user groups (i.e. NWS, Navy, Army) in a more cost efficient manner.

REFERENCES

Jones, A. S., and T. H. Vonder Haar, 2002: A dynamic parallel data-computing environment for cross-sensor satellite data merger and scientific analysis. *J. Atmos. and Oceanic Technol.,*.**19,** 1307–1317.

ACKNOWLEDGMENTS

This work was supported by a grant from the NOAA High Performance Computing and Communications (HPCC) Program. The work at CSU/CIRA was supported by contract NA17RJ1228 under CIRA's cooperative agreement with NOAA.