

Weather Service Harnesses Whirlwind of Data With New Cluster

Improved productivity, accurate, cost-effective, real-time online updates compared to traditional UNIX Solutions.



Products:

- » Altus 1000E Opteron™ server
- » Altus 3200 Opteron™ server
- » Scyld Beowulf® Series
29 Linux clustering software

Key Benefits:

- » Increased Productivity
- » Easy to Implement, Manage
- » Fast, Powerful Computing

Advice from NSW:

1. When evaluating Linux clusters, look for a company that provides full support for Linux and Beowulf cluster experience.
2. Make sure the company certifies the equipment to run under Linux with few special drivers needed.
3. Make sure you do a cost/benefit analysis of what the total cost of ownership is of a cluster to understand any hidden administration costs or certain types of end user or level of system administration support.

“It was exactly what we needed in terms of power. What we did not expect were the productivity gains resulting from the management features of the Scyld Beowulf software.”

– Paul Kirkwood, chief of the NWS Southern Region Dissemination Enhancement Team

Summary:

Since one-seventh of the U.S. economy (about \$1 trillion a year) is weather sensitive, the National Weather Service (NWS) really wants to stay on top of its data. But it's not easy converting huge quantities of data from every nook and cranny of the planet into real-time calculations and then current weather updates and future forecasts on a publicly accessible website that averages 15 million hits per day for the NWS Southern Region alone! And who wants to give a fishing fleet the wrong information when they're sailing into a storm?

So the Southern Region team chose Penguin Computing and Scyld Software for a powerful computing system capable of dynamically processing vast amounts of data. As a result, the team can track and generate hourly updates of accurate information for a region covering the area between Albuquerque, N.M. and San Juan, Puerto Rico with less cost and effort than traditional UNIX systems. That makes for a pretty good forecast going forward.

The Challenge:

The NWS collects huge quantities of data from Doppler weather radars, data buoys for marine observations, surface observing systems, instruments for monitoring space weather and air quality as well as satellites operated by its parent organization, the National Oceanic and Atmospheric Administration (NOAA).

This enormous stream of information is fed into sophisticated computer models to generate accurate up-to-the-minute weather updates and warnings for the Southern Region. Paul Kirkwood and his Southern Region Dissemination Enhancement Team were responsible for getting the information out to the public through heavily used applications,

including the web-based Graphical Point Forecast.

The team needed a powerful cluster for this kind of job. However, the existing UNIX servers had reached the end of their lifecycles and the team needed replacements with serious back-end processing capabilities. Without that feature, NWS could not maintain real time translation of data from the Channel Definition Format (CDF) into HTML for the website. And, even though the cluster had to run very complex calculations, there was not a great deal of additional system administrator time – is there ever? – to manage new cluster software.

The Dissemination Enhancement Team also wanted the new system to be powerful enough to handle enormous data sets dynamically at a lower cost, making the search even more complicated. But because the NWS goal is “producing and delivering information you can trust when you need it,” the team explored every option for the right combination of hardware, software, and support to meet its needs.

The The Solution:

During initial investigations, the team soon realized that a Linux cluster would deliver the power of traditional UNIX systems at a fraction of the price. It would also be comparatively easy to implement since the Dissemination Enhancement Team had already standardized its computing system on Linux. The team also focused its search on AMD Opteron™-based systems, because the team members knew these processors would be able to handle the existing 32-bit and future 64-bit custom applications NWS was running within the required time-frames.

Weather Service Harnesses Whirlwind of Data With New Cluster

Improved productivity, accurate, cost-effective, real-time online updates compared to traditional UNIX Solutions.

The Solution (cont.):

With the criteria in place, the Penguin Altus server-based cluster with Scyld Beowulf® would meet the team's cost and performance needs compared to other systems. Kirkwood was drawn to Penguin because of its Linux and Beowulf expertise and because it was one of a few turn-key Linux vendors supporting the AMD Opteron processor. The new Penguin/Scyld cluster was exactly what the team needed in terms of power, plus it was backed by Penguin Computing's strong customer support and unique expertise in implementing the Scyld Beowulf software cluster. What the team did not expect were the productivity gains resulting from the management features of the Scyld Beowulf software itself.

The commercial-grade Scyld software used in the NWS cluster decreased the amount of time it takes to manage and update the system. Advanced features of the Scyld Beowulf software, such as the single system view and boot, were especially valuable. In addition, the ability to load Linux once on the master node as opposed to on every single node saves the team time and ensures that the

cluster is running consistent versions of Linux, preventing version skews.

Using the cluster on Linux Kernel Version 2.4, the team can now track and generate hourly updates of accurate information for the public in an easy-to-use format for a region covering the area between Albuquerque, N.M. and San Juan, Puerto Rico. As a result of this powerful new computing system, the team can focus on higher level issues like developing and implementing new applications that make it easier for people to use the Web site.

The Installation

As configured for the NWS Southern Region Dissemination Enhancement Team, the Penguin Computing Altus servers cluster is composed of 7 compute nodes and primary and secondary master nodes with dual AMD Opteron processors

and 4GB of memory in each node. The Scyld Beowulf Series 29 Linux Clustering Software chosen by NWS is a fully integrated software platform for extreme ease of deployment and manageability of a Linux cluster.



About NWS

The National Weather Service (NWS) provides weather, hydrologic, and climate forecasts and warnings for the United States, its territories, adjacent waters and ocean areas, for the protection of life and property and the enhancement of the national economy. NWS data and products form a national information database and infrastructure which can be used by other governmental agencies, the private sector, the public, and the global community.

National Weather Service

» <http://www.nws.noaa.gov>

National Weather Service Southern Region

» <http://www.srh.noaa.gov>

About Penguin Computing and Scyld Software

Penguin Computing is the leading innovator of highly scalable, powerful Linux cluster, server and workstation platforms based on open standards hardware and software. The company has focused 100% on Linux since its inception and as a result, its solutions are robust and reliable, with a price/performance that's hard to beat. Scyld Software, a subsidiary of Penguin Computing, provides the industry leading Linux clustering software, Scyld Beowulf™, whose superior ease of use dramatically simplifies the deployment and management of Linux clusters. The company's extensive customer base includes Fortune 1000 companies, government agencies and educational institutions. Founded in 1998, Penguin Computing is headquartered in San Francisco, California. For more information, please visit:

Penguin Computing

» <http://www.penguincomputing.com>

Scyld Software

» <http://www.scyld.com>