CNS /Update Newsletter Feature

Open Up! Capacity on Demand

CNS Document ID: u021109a
Last Updated: 11/21/02

UF Information Technology

UFIT

2046 NE Waldo Rd, Suite 2100
Gainesville Florida 32609-8942
(352) 392.2061
<editor@cns.ufl.edu>
<table>
<thead>
<tr>
<th>Table of Contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Up! Capacity on Demand</td>
<td>.................................................................3</td>
</tr>
</tbody>
</table>
Open Up! Capacity on Demand

There is a new catch-phrase in the computer industry: "capacity on demand." It allows businesses and installations like NERDC the flexibility to respond to unexpected growth or temporary usage spikes.

How does it work? An installation can approach this function in several ways. It can contract ahead of time for a projected upgrade. Then, when the upgrade is needed, it can be shipped right away, allowing for an immediate response to increased usage.

An installation can purchase a multiprocessor machine which contains more processors than are currently needed. That way, the "spare" processors can be turned on only when needed. The customer does not pay for the extra processors until they begin using them. Or, they can acquire additional on-site servers that lie in reserve until needed.

The latter arrangement is the one used for the NERSP computer complex. The Open Systems group maintains a small bank of processors that are used for testing, development work, and for additional capacity when usage spikes occur. This establishes a standby reserve of units that are always available to bolster capacity needs. If a needed unit is engaged in testing or pre-deployment for another project, it is immediately recovered and configured to respond to current capacity requirements.

Each semester, the UF e-mail load gets a little bigger. NERDC must estimate the size of this growth and adjust the system to accommodate customers' needs. However, there are other situations and events that can cause a sudden increase in traffic over a server that we cannot always predict. For example, if there are exceptionally active viruses sending messages, that can increase the amount of work that the server has to do by increasing the total number of messages. Or, as happened at the beginning of this semester, a related system can experience a problem that impacts on the performance of other servers.

That week, the newly installed virus scanning system developed a bottleneck that resulted in a backlog of messages to be sent via the smtp.ufl.edu (outgoing mail) server. As the queue of backlogged messages grew, the staff knew there was a problem. Within a very short time, additional servers were deployed to deal with this unexpected activity surge.

One component that allowed the on-reserve servers to be used in this manner is the Network Dispatcher component of IBM's Websphere Edge Server. This was described in the February 2002 Open Up! column in /Update. This software accepts incoming network traffic and then redistributes it to cluster of application servers, allowing a group of servers to work in concert to support a single user application.

This gave OSG staff the opportunity to diagnose the problem and make corrections while mail was still being processed. Additionally, the staff maintains a map of each server and careful installation records for each service on each server. These preparations facilitated the quick deployment of extra servers containing the correct image/configuration to alleviate the temporary strain on the smtp server.

Your Comments are Welcome

We welcome your comments and suggestions on this and all UFIT documentation. Please send your comments to:
UF Information Technology

UFIT

2046 NE Waldo Rd, Suite 2100
Gainesville Florida  32609-8942
(352) 392.2061
<editor@cns.ufl.edu>