



Volume 3  
Issue 12

# Availability Digest

--- achieving 100% uptime

December 2008

The digest of current topics on Continuous Processing Architectures. More than Business Continuity Planning.

BCP tells you how to **recover** from the effects of downtime.

CPA tells you how to **avoid** the effects of downtime.

## In this issue:

### [Never Again](#)

[Innocuous Fault Leads to Weeks of Recovery](#)

### [Best Practices](#)

[Google's Extreme-Green Data Centers](#)

### [Availability Topics](#)

[Tussling with the Word "Redundant"](#)

### [The Geek Corner](#)

[Configuring to Meet a Performance SLA](#)

Complete articles may be found at

<http://www.availabilitydigest.com/articles.htm>.

## Our Community Connect Europe 2008 Active/Active Seminar Now Available

Last month, Mannheim, Germany, hosted Community Connect Europe 2008, the joint conference and trade show sponsored by Connect, the Independent HP Business Technology Community users' group.

As has been the case for the past several years, continuously available systems using active/active technology were highlighted. Lloyds TSB Bank and US Bank presented descriptions of their active/active systems. Lloyds Bank proudly emphasized that it has not had a failure in its active/active POS/ATM system in the system's fifteen years of operation.

The OpenVMS folks reviewed upcoming enhancements to OpenVMS split-site active/active clusters. Data-replication vendors Goldengate and Gravic Shadowbase, both with booths on the floor, described the use of their products to implement active/active systems. Both HP and Gravic announced plans to support synchronous data replication.

Our multi-hour seminar, "Active/Active Systems: Theory and Practice," was well attended, as usual. If you would like a copy of it, just let us know at [editor@availabilitydigest.com](mailto:editor@availabilitydigest.com).

Dr. Bill Highleyman, Managing Editor

---

## Never Again

### Innocuous Fault Leads to Weeks of Recovery

The bank's ultimate horror started with a single disk failure on one node of a three-node, geographically-distributed system. Through a sequence of unimaginable events, this presumably innocuous fault spread through all three processing nodes run by the bank, taking them all down. The international bank suddenly found that its POS and ATM services had come to a halt.

It would take weeks to recover, and full recovery was impossible. Significant amounts of data were lost forever, though some of it was recoverable from other incompatible systems. Manual reconciliation of disputes carried on for months.

[--more--](#)

---

## Best Practices

### Google's Extreme-Green Data Centers

Google has recently filed for a patent describing an all-green data center. No man-made energy is used for powering the data center nor for cooling it. Rather, power is generated from the wave motion of the ocean; and cooling is accomplished using seawater. Essentially, Google's proposed green data centers are barges that house cargo container-sized server farms

A major focus of data-center design today is the minimization of its energy footprint. Google's floating green data centers may be a major step in this direction.

[--more--](#)

---

## Availability Topics

### Tussling with the Word "Redundant"

A common language between peoples can nevertheless be confusing. Case in point: the word "redundant." The Availability Digest has always used "redundant" to mean "backup." However, it has been pointed out to us that to the British, it may mean "unnecessary." That certainly changes the way we would look at a "redundant" system.

[--more--](#)

---

# The Geek Corner

## Configuring to Meet a Performance SLA – Part 1

If the performance of a data-processing system has degraded to the point that it is no longer useful to its users, the system is, for all practical purposes, down. It is therefore common to specify a performance requirement in the Service Level Agreement (SLA) for the system. The performance requirement is often expressed as a probability that the system's transaction-response time will be less than a given interval. For instance, "98% of all transactions must complete within 500 milliseconds."

The normal queuing relationship with which we are all probably familiar,  $\text{response time} = \text{service time}/(1-\text{load})$ , gives only the average response time to be expected. This simple relationship tells us nothing about the distribution of response times, which is what we need to know in order to determine what percentage of response times will be in excess of some value specified by an SLA.

In this series of three articles, we explore the solution to this question. In Part 1, we review the simple queuing equation for a single server. Part 2 extends this analysis to multiple servers. In Part 3, we address the response-time distribution, which leads to the SLA solution.

[--more--](#)

---

## Would You Like to Sign Up for the Free Digest by Fax?

Simply print out the following form, fill it in, and fax it to:

Availability Digest  
+1 908 459 5543

**Name:** \_\_\_\_\_  
**Email Address:** \_\_\_\_\_  
**Company:** \_\_\_\_\_  
**Title:** \_\_\_\_\_  
**Telephone No.** \_\_\_\_\_  
**Address:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

The Availability Digest may be distributed freely. Please pass it on to an associate.

To be a reporter, visit [www.availabilitydigest.com/reporter](http://www.availabilitydigest.com/reporter).

Managing Editor - Dr. Bill Highleyman [editor@availabilitydigest.com](mailto:editor@availabilitydigest.com).

© 2008 Sombers Associates, Inc., and W. H. Highleyman

