



Volume 4
Issue 3

Availability Digest

--- achieving 100% uptime

March 2009

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.
www.availabilitydigest.com

In this issue:

[Never Again](#)

[Has Gmail Become Gfail?](#)

[Availability Topics](#)

[Eavesdropping on the Internet](#)

[Recommended Reading](#)

[Business Continuity Today](#)

[The Geek Corner](#)

[Meeting a Performance SLA - Part 4](#)

Complete articles may be found at
<http://www.availabilitydigest.com/articles.htm>

Sign up for your free subscription at
<http://www.availabilitydigest.com/signups.htm>

Availability is More Than MTBF and MTR

Availability is more than failure intervals and repair times. Anything that compromises the usefulness of data-processing services to end users is an availability concern. In this issue, we track seemingly unrelated factors – Internet vulnerabilities and performance – that seriously compromise system availability.

The Border Gateway Protocol (BGP) is the primary routing protocol for long-haul Internet traffic. A vulnerability (not a bug) in BGP allows a nefarious web site to hijack your traffic and dump it into a black hole. Even worse, it can eavesdrop on your traffic even as your messages are sent to their rightful destinations.

On another topic, a system is less useful if its performance is poor. Therefore, SLAs often contain a performance specification in the form of “99% of all transactions shall complete in less than 100 msec.” How do you know if your system will meet this SLA specification? Our five-part series on configuring for a performance SLA addresses this question.

Internet vulnerabilities and performance issues are two challenges that compromise availability. Others are obvious, some are less overt. The Availability Digest is always seeking those unexpected factors that rob a system’s usefulness. Know of any? Please share them with us at editor@availabilitydigest.com.

Dr. Bill Highleyman, Managing Editor

Never Again

Has Gmail Become Gfail?

Gmail is a highly popular free email service offered by Google. It is the third most popular email service (after Yahoo! and Hotmail) with about 100 million subscribers. Gmail launched in April, 2004, as an invitation-only beta release and became available to the general public in February, 2007.

Gmail's business model is advertising instead of user subscription fees. Gmail scans incoming email and posts textual advertising messages related to the email's content. Gmail is a component of Google Apps, which is an ad-supported free suite of useful collaboration, messaging, and office productivity services.

Unfortunately, Gmail has had a dismal availability record. It has been down six times, often for a day or more, in the last eight months, displaying an availability of less than 99%.

The bright news is that unlike many other SaaS vendors, Google works hard to communicate the status of problems to its customers. It has recently made available the Apps Status Dashboard to provide frequently updated online information concerning outages and other problems.

[--more--](#)

Availability Topics

Eavesdropping on the Internet

At the 2008 DEFCON hackers' conference, security researchers Anton Kapela and Alex Pilosov demonstrated a fairly simple technique to divert Internet traffic to an eavesdropping site on its way to the intended recipient. Though this vulnerability had been earlier predicted, no one had ever been able to demonstrate it. Kapela and Pilosov showed that they could eavesdrop on DEFCON traffic with their own server to the cheers of the hacker crowd.

Eavesdropping can be used by corporations for competitive purposes and by governments for surveillance purposes. Messages can even be modified in transit by the eavesdropper.

There is no easy way to detect that your traffic is being monitored, nor is there any easy way for the Internet to prevent this sort of attack. Your best bet is encryption so that your traffic has no value to an eavesdropper.

The problem is in a vulnerability of BGP, the Border Gateway Protocol. BGP is the routing protocol used to distribute global routing information throughout the Internet. It is used by all major ISPs as well as by many smaller providers and other organizations.

In this article, we review those elements of BGP that lead to the possibility of eavesdropping.

[--more--](#)

Recommended Reading

Business Continuity Today

Business Continuity Today is a freely-available living eBook covering a broad range of topics relating to business continuity. We say “living” because the book is a perpetual work-in-progress. Published by Vision Solutions, a new chapter is added every few months. Six chapters have been published since February, 2007; and the seventh is due out soon. The book is downloadable from www.businesscontinuitytoday.com.

Each of the chapters is an overview of some pertinent field. Some of the chapters are written by Vision Solutions staff, and others are written by guest experts in their fields. As overviews, they are short on specifics but long on the broad coverage of their subjects. Though Vision Solutions is an IBM Premier Business Partner focusing on business continuity solutions for IBM systems, the book is wonderfully nonvendor specific.

[--more--](#)

The Geek Corner

Configuring to Meet a Performance SLA – Part 4: Single Server with General Service Time

Many applications carry with them a performance Service Level Agreement (SLA) that specifies the response times they must achieve. 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, “When handling 50 transactions per second, 98% of all transactions must complete within 500 milliseconds.”

In Part 1 of this series, we derived the basic average response-time expression for a single-server system. In Part 2, we extended that result to a multiserver system in which multiple servers work off a common work queue. In Part 3, we showed how to configure a server with exponential service time to meet an SLA specification.

If service time is not exponentially distributed, the solution is more complex. Here in Part 4, we show how to size a system to meet a performance SLA for servers with general service times.

Though the derivation of SLA sizing is involved, and the calculations are complex, SLA sizing is reduced to simple graphics supported by an easy-to-use spreadsheet.

[--more--](#)

Sign up for your free subscription at
<http://www.availabilitydigest.com/signups.htm>

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.
Managing Editor - Dr. Bill Highleyman editor@availabilitydigest.com.
© 2009 Sombers Associates, Inc., and W. H. Highleyman