Availability Digest
--- achieving 100% uptime
Volume 4
Issue 9
September 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:

Case Studies
Apollo 11 - Continuous Availability, 1960's Style
Never Again
PayPal Fault Takes Merchants Offline
Best Practices
Achieving Fast Failover - Part 2
Recommended Reading
Mission-Critical Network Planning

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

Apollo 11 – Continuous Availability, 1960s Style

Continuous availability may be a hot topic today, but it wasn't in 1969. Yet that was the year we used continuous availability to put a man on the moon - then successfully did it again and again. The Apollo moon missions were all highly automated and represented the extreme in safety-critical systems. How did we achieve continuous availability at the dawn of the transistor age when four decades later, we are still trying so hard to accomplish just that?

I had the good fortune to spend time with Ed Poole, one of the engineers on the Apollo 11 team, to learn about their approaches. Interestingly, it turns out that the techniques used were similar to what we use today – lock-stepped computers and active/active systems. Read our article about Apollo 11 in this issue of the Digest to learn more.

Archeologists have just uncovered a 4,000-year old wall of cut stone in Jerusalem. No one can figure out how stone was cut so perfectly with primitive tools. Maybe a thousand years from now, scientists will be wondering how we put a man on the moon with such primitive technology.

Dr. Bill Highleyman, Managing Editor

Case Studies

Apollo 11 – Continuous Availability, 1960s Style

We talk a lot about continuous availability – systems that virtually never fail – as if this is a brand new concept made possible by today's technologies of fault-tolerant systems, clusters, virtualization, and active/active systems. We talk about task-critical systems, in which a fault will idle workers or production lines, mission-critical systems that can take down a company, and safety-critical systems whose failure can mean loss of life or property.

But forty years ago, continuously available systems put men on the moon and got them back several times without a loss. The entire lunar mission, from launch to translunar coast to landing and then rendezvous and return to earth, were controlled by computers whose failure would mean the loss of the astronauts. Against all odds, all systems worked impeccably for the lunar landings.

How did the technologies of the '60s achieve this near miracle? In this article, we look at the approaches taken back then to ensure continuous availability. We will find that they were not much different than what we do today. The ability to achieve continuous availability in our computer systems has been around for a long time.

--more--

Never Again

PayPal Fault Takes Merchants Offline

PayPal is the world's largest online payment service. By inserting a small PayPal HTML snippet into his web site, a merchant can effortlessly accept customer payments using a variety of credit cards such as Visa, MasterCard, American Express, and Discover as well as PayPal accounts.

Clearly, the availability of PayPal services is of utmost importance to its merchant customers. If a merchant depends upon PayPal for all of its payment services – credit cards and PayPal accounts – a PayPal outage can bring that merchant to its knees.

But on Monday, August 3, 2009, that is what happened. An "internal network hardware issue" crashed the entire PayPal network. No merchant anywhere in the world could conclude a transaction over PayPal. It was hours before PayPal was back in full operation.

The lesson to be learned from these outages is not a provider lesson; it is a user lesson. Databases are not the only things that need to be backed up. Critical services also need to be backed up. Contingency plans must be in place in order for a company to continue operations in the face of the loss of a critical service.

--more--

Best Practices

Achieving Fast Failover in Active/Active Systems – Part 2

Active/active systems provide continuous availability not because they avoid faults but because they can recover from faults so quickly that users don't realize that there has been an outage. This capability requires not only that recovery to a backup component be rapid but that it be reliable.

Recovery in an active/active system is reliable because it is known that the surviving node is operational. After all, it is currently processing transactions successfully.

Rapid recovery in an active/active system requires two capabilities – rapid fault detection and rapid failover. In Part 1 of this series, we discussed three techniques for achieving these goals – transaction redirection, user redirection and network redirection. In this part, we conclude with server redirection and DNS redirection.

With server redirection, nodes monitor themselves. This may lead to a fault indication even though both nodes are operational. As a result, there can be a tug-of-war in which two nodes both try to seize the users connected to the other node. Methods to avoid the tug-of-war syndrome are discussed.

DNS redirection can take several minutes and is not conducive to continuous availability.

--more--

Recommended Reading

Mission-Critical Network Planning

A decade ago, end users of a company's computer systems were located within company walls. However, today's heavy reliance on the Internet has created an environment in which many users are located outside the company setting. When a system goes down, employees may still have little recourse but to file a trouble ticket; but angry customers can take their business elsewhere. Mission-Critical Network Planning, by Matthew Liotine, addresses this new complexity.

A better title for his book might have been Mission-Critical *Continuity* Planning. In his title, Liotine includes in the term "network" any component that may affect the continuity of a mission-critical application – communication networks, servers, storage, data centers, power, and HVAC. Each of these topics deserves a full volume (or more) of its own. Consequently, Liotine cannot delve into great detail on any of them in his 400-page book.

As a result, this book is not intended to be a design guide for high availability. Rather, it provides the background that one new to the field can use to become rapidly conversant in the various technologies involved in high-availability design. It is especially useful to managers who are faced for the first time with converting their legacy systems to high-availability systems capable of recovering from a fault in minutes rather than in hours or days.

--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