

# COMMUNICATING WITH SAM

## Enabling High Availability with Network Strategies



**Sam Johnston**

### Question:

Recent world events have put business resiliency planning back on our management agenda as a high priority. In the past, our main focus as a company was contingency planning around a system failure, but it is clear that in addition to system availability we also need plans that address interruptions to our actual work space if we are to provide business continuity.

As a component of our overall business resiliency plan, I have been burdened with the task of creating a high availability environment for our mission critical business applications that reside on the iSeries, both in terms of system availability, and secondly how staff could continue to access these systems even if our offices are not functional. Recently our company has consolidated our interactive web site onto the iSeries system that also supports our Internet Web Application presence, so any business interruption that impacts our data centre will also impact our crucial self service business and outages will be apparent to our customers. The challenge we are faced with is how to physically and logically configure our system and network to support high availability. We are looking at two options: high availability within a single data centre as well as considering the more costly and complex solution of high availability through multiple datacenters. Is it possible to create this environment and provide seamless availability to our customers?

### Answer:

The 21<sup>st</sup> century so far has not been kind to business managers that need to plan for unexpected events. Technology has been critical in creating a globally integrated economy, generating efficiencies and economies of scale that we benefit from daily. However, with these benefits come new costs of operating a business. Markets, enabled by technology, move at a lightning pace, meaning outages, even slight ones, have a significant impact not just on short-term profits but often in terms of survival itself. While CEOs enjoy the rewards of creating a single product for markets the world over, the notion that the world is a smaller place and events routinely ripple throughout the world means that the risk side is an exponential increase in the probability of an event impacting the continuity of your business.

The good news is that many of our businesses in Canada have transformed to knowledge based businesses, where workers can contribute in a virtual manner. While there is no denying that a manufacturing plant would be catastrophically impacted by a plant being inaccessible by employees, there are

many workers that can at least on a temporary basis perform at or near capacity if they simply have basic tools available on a remote basis.

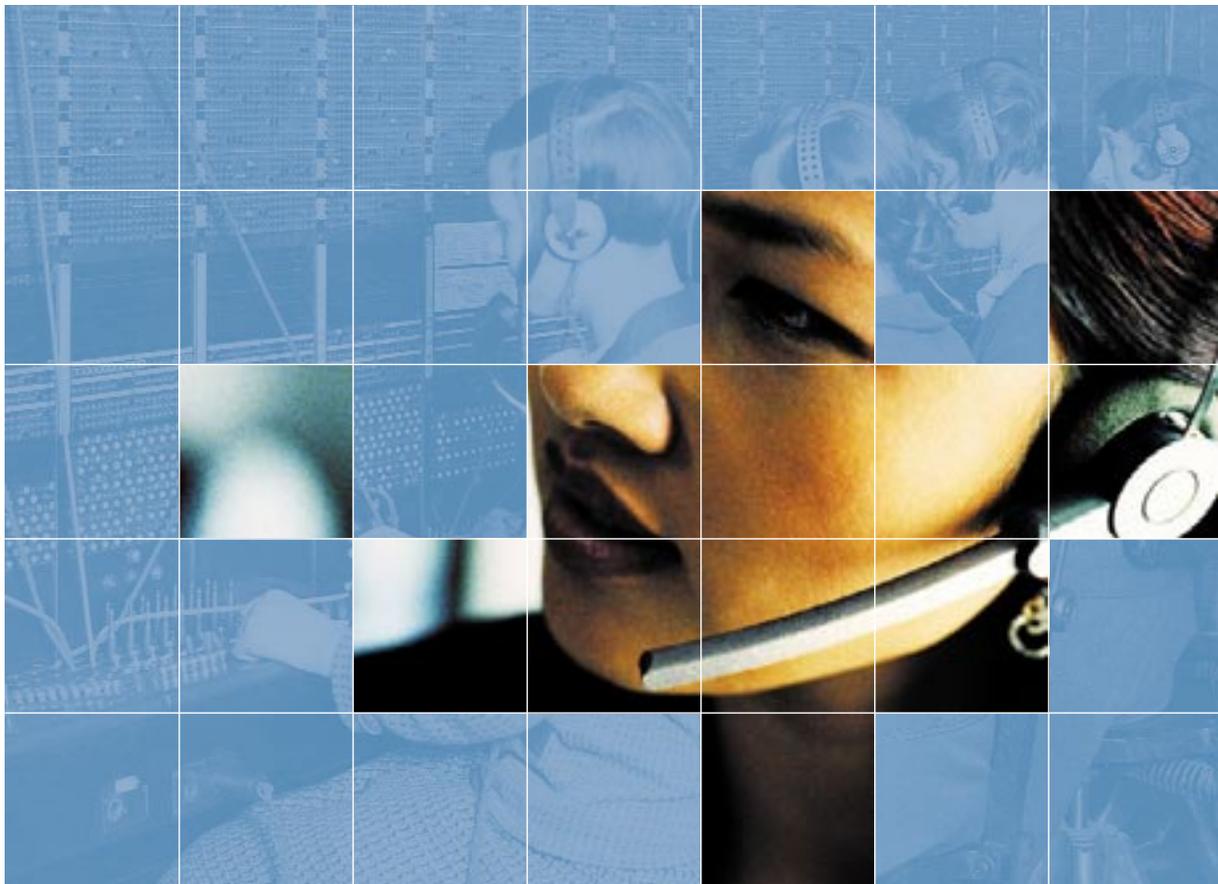
Often overlooked in the discussion on business continuity is the network. Communications and networking are crucial to both enabling the underlying high availability technologies, and can in themselves play a significant role in delivering business resiliency solutions.

The traditional view, for a call centre as an example, was to find a disaster recovery centre that could provide seats in the event of a failure, but this may not be affordable for all companies. Recent events have re-enforced the need for strategies that can be critical to worker continuity in the event of a crisis. While it would be impractical from a cost perspective to build full redundancy in terms of the work place, there are a couple of things you can do to ensure that knowledge workers are protected from work place quarantines as an example.

First, even if you have not taken the decision to adopt a full telecommuter approach, it is essential that all knowl-

edge workers have tools that permit them to contribute regardless of where they are working from. Secondly, your telecommuter strategy needs to address all the business tools – computer, files and most importantly the phone. This is where convergence can play a role, whereby IP call processing and call centre solutions can make it seamless to customers in the event of a disaster by remotely enabling users on the production phone system. With a VPN connection that provides connectivity to both data applications and phone services, call centre agents could be sitting at home handling calls with customers completely unaware of the fact that your office is closed. While this can be enabled in traditional voice solutions, they are more easily executed in an IP world. While it would be impractical to cover all potential situations, the reality is that simple use of communication technology could provide significant protection to your business against events that would prohibit workers from functioning in their normal work space.

A business continuity strategy for workers and their physical workplace is important, and certainly gaining a visibility previously unseen, but we should



# Moving Forward

MAKE A FRESH START WITH YOUR COMMUNICATION.

Let INTESYS and Cisco Systems' AVVID architecture help you create a new world system for your new work space to take advantage of Internet enabled telephony technologies.

- Converge your Internet, data, voice and video traffic onto one network
- Implement a Cisco Call Manager IP telephone system
- With Cisco Unity, provide a single, unified messaging platform
- Add wireless in-building IP Telephones
- Collaborate globally with the Cisco Conference Connection server
- Bring it all together with the Cisco's Customer Response Solutions

For a free initial consultation to discuss how you can transform your communication needs, call INTESYS, the convergence experts 1.800.736.6531, or e-mail [sales@intesys-ncl.com](mailto:sales@intesys-ncl.com).



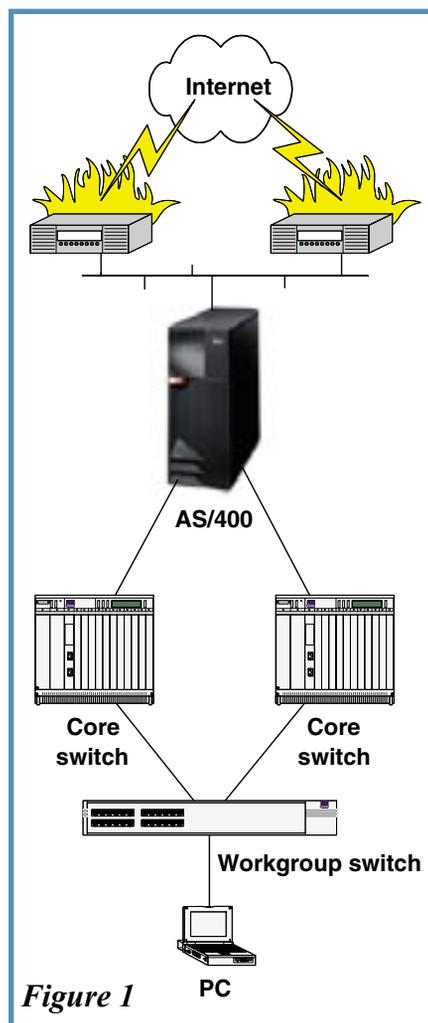
not lose sight of the fact that system and network outages are more likely to occur. Globalization, server consolidation and Internet based applications are driving the need for 24/7 application and data availability. Despite of great improvements in hardware and software reliability and concurrent maintenance abilities, systems are not immune to physical site disasters or network service provider failures. Many businesses today have complex agreements with customers and suppliers that demand service levels such as “just in time” deliveries and electronic payments. The penalties for missing the targets associated with these types of agreement can be extreme. You may lose a customer forever. The cost of an outage isn’t just the cost you incur during the outage - the cost is the future value of the business relationship. You may be able to recover your systems from a disaster but you may never recover the business. To prevent this, many companies have recognized the need for a systems high availability solution, and potentially the next step of a redundant data centre.

While it is important to have a strategy that enables workers to perform even if they cannot go to their office, it is also critical to have a system high availability solution as a key component of an IT disaster plan and the IT plan as a part of an overall business continuity strategy. Systems, Networks, Employees, Customer communications, Supplier shipments and payments all must continue to function in the event of a disaster.

Building a redundant data centre is a business decision. Assuming that the cost of an outage is high, and the probability of an event, albeit low, must be protected against, then you may have the ROI to justify a completely redundant system along with the associated network infrastructure. Let’s assume this is the case for your business, and now you need to evaluate the impact of the decision in terms of whether to locate the high availability system in a secondary site.

Developing a system high availability solution for your iSeries is unique versus

a Windows 2000 clustered environment, but made easy by off the shelf High Availability applications that mirror and synchronize your systems. These products and a second system in another location will allow your business to continue operations on a 24/7 basis. The solution will manage the mirroring of vital application and data objects from the production system to the backup system efficiently and effectively.



It is essential that the connectivity enable simultaneous updates to ensure that no transaction is lost in a disaster, which can be challenging in terms of cost if the sites are located remote to one another. One key factor in minimizing latency between systems is the network that connects the two machines. It must be robust enough to handle the volume of objects being replicated. A typical 500 user environment could generate 100GB

of replicated objects per day. Wouldn’t it be nice if this workload were spread evenly over a 24 hour period, but this can’t always be the case so make sure your network provisioning accounts for peaks. Part of the planning process should access the journaling requirements by time of day based on the transaction volume of the production system. With this information network engineers can calculate the network requirements to ensure both systems can be in sync with each other at all times.

While the basic network set-up described above will ensure that your systems are adequately protected from failures, it does little to protect you from the perhaps the most frequent source of interruptions – the production network.

With respect to Network High Availability design to support system high availability there are a number of considerations that must be weighed. Regardless of whether you can justify a redundant data centre you must consider factors such as internal network connection support versus external network connection support (for example Internet commerce). There is little benefit of a highly available server if the server cannot be reached by key users, internal or external.

In considering any high system availability solution, you must also keep in mind the following key network design considerations:

- Ensure that your data connections to your network are highly available
- Ensure that your data path between systems is robust and sized in accordance with your data replication requirements.
- Ensure that your data network equipment is also highly available, building redundancy at single points of failure
- Ensure your clients have a level of high availability access to the server farm.

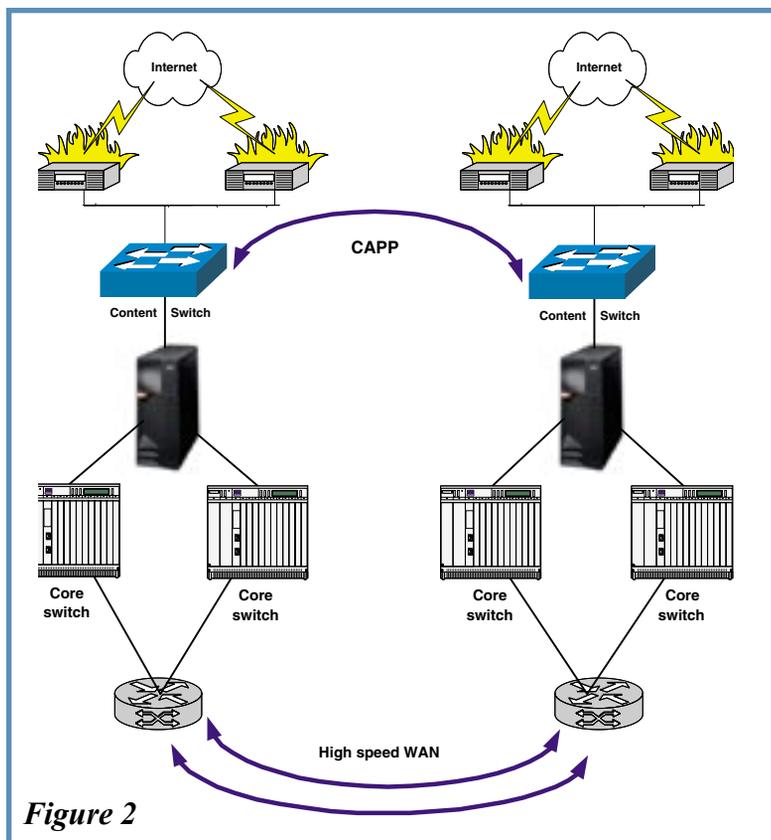
From a networking perspective, clearly the single data centre option with dual servers and dual network infrastructure is the easiest and cheapest to provision.

However, in light of recent world events you will need to weigh the nature of your business and assess the impact of losing an entire site versus the incremental investment to create a second data centre. If you do not select a redundant data centre model, you may still want to purchase disaster recovery services to offset the risk. For a single data centre solution, you should provision a highly available network leveraging Gigabit Ethernet as a minimum for adapters connecting into a robust network chassis solution that features a redundant chassis with redundant switching engines and power supplies. In doing so you can have your iSeries system connect to the switch using multiple NIC interfaces. Failover within the NICs can be configured in multiple ways and many software providers will provide primary and secondary options. In order to support a highly available Internet connection, a redundant firewall model should be used with or without a content switching engine to fail over the automatically the inside connection point to the new server (Figure 1).

A high availability solution with redundant data centers provides for more challenges. The primary challenges that need to be overcome are bandwidth between sites to support effective data replication and the ability to service Internet Commerce from two different sites seamlessly to the end users. In order to answer the first question of bandwidth, working with your high availability software partners, you should be able to determine, based upon your system, the required bandwidth for replication. To connect the two sites, the preference would be to use utility based fibre between locations if the sites are within

the same city, or a high speed ATM or MPLS network if they are geographically disparate. These WAN connections should be provisioned with reliability to support your plan, ideally via redundant paths and carriers. The same should be true of your production WAN to link remote sites to the two centers to ensure that they can failover seamless in the event that the secondary data center becomes primary.

The challenge of provisioning Internet access for web commerce to support multiple sites is a bit more challenging, but can be accomplished seamlessly. The base assumption is that you would provision two separate Internet connections with separate IP address space for each site. Then you can use tools such as Global Load Balancing on as an example a Cisco CSS Content Switch or Global DNS Load balancing on other equipment. Cisco's implementation of Global Load balancing works by enabling peering CSS Switches to communicate status of the web sites using CAPP. Combine this with the built in DNS server and requests can be sent in a load balanced fashion or in your case, redirected related with availability of the server and service connection. The control of this mechanism may be dependent on how your high availability is implemented from the system standpoint. The CSS switch is smart enough to know when the end system is available to offer services based upon detailed keep-alives that could be deployed. Remember with this solution to also provision reliable Internet connections at both sites to support the connection needs of each solution. Figure 2 highlights a theoretical connection for the high availability solution.



Ultimately your high availability model will only be as good as your network. If your network strategy does not mirror the system availability model, you may find that at your time of need, while the system is purring away, users cannot reach it to perform mission critical tasks. High availability is about predicting outcomes and protecting against those negative events. Remember, the network is the most likely spot to cause you grief due to the number of inter-dependencies and factors outside of your control. While the CFO will see high availability as insurance, and a cost of doing business, a little creativity in how you provide high availability by using solutions such as remote agents may even have a positive impact on your daily operations by creating a flexible and more cost effective work model.



**Sam Johnston** is a partner and Chief Technology Officer of Intesys Network Communications Ltd., providing value-added networking and e-commerce solutions to the iSeries community. He can be reached at (416) 438-0002 or via e-mail at [sjohnston@intesys-ncl.com](mailto:sjohnston@intesys-ncl.com). Any TUG member wishing to submit a question to Sam can forward their typewritten material to the TUG office, or to Intesys. The deadline for our next issue is Friday June 13th, 2003.