

High availability technologies for DB2

By James Gill
Principal Consultant & DB2 Expert, Triton Consulting
March 2009

Availability for a brave new world

With the advent of Software as a Service (SaaS) more businesses are relying on the ability to access their business data through web based applications. In addition to the rise of SaaS and Cloud Computing, our businesses are increasingly operating on a global scale. When once you could schedule your maintenance updates for Sunday night, this now affects users across the other side of the globe.

When downtime is unplanned however, these issues multiply ten-fold. These outages are a lot more visible to users and the public at large with potential ramifications to revenue, brand image and customer satisfaction.

In this paper we will look at the various solutions to the application availability issue for DB2 databases and how they meet the demands of our ever changing global operations.

Availability solutions for DB2 databases

There are a number of high availability & disaster recovery solutions which have been in the marketplace for some time.

Active-passive clustering is a good general purpose high availability solution within a local environment. It typically provides a warm standby solution – i.e. an outage in the primary server is detected by the backup, which then takes over. The main stumbling block with this method is that it cannot work over a long distance and so is really only suitable for a single location solution.

With an active-passive clustering solution an organisation typically has an active or primary server and a passive or standby server. The TCO of this solution can be relatively high with expensive hardware resources sitting idle on the standby server. In addition to the warm standby server some organisations set up an additional standby within a separate DR site.

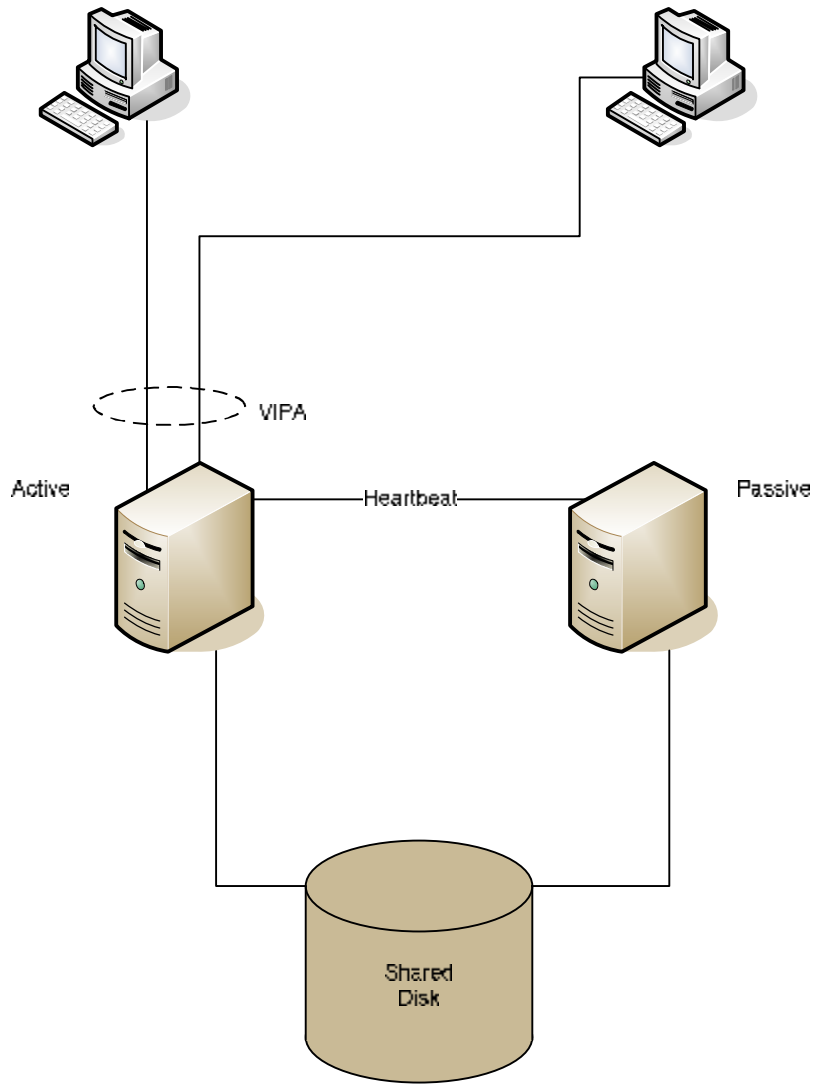


Figure 1 – Shared Disk HA Solution

A heartbeat between servers detects when the primary server goes down and moves services across to the failover server. There is generally an outage experienced where the primary server has failed, and the standby server detects this change in state.

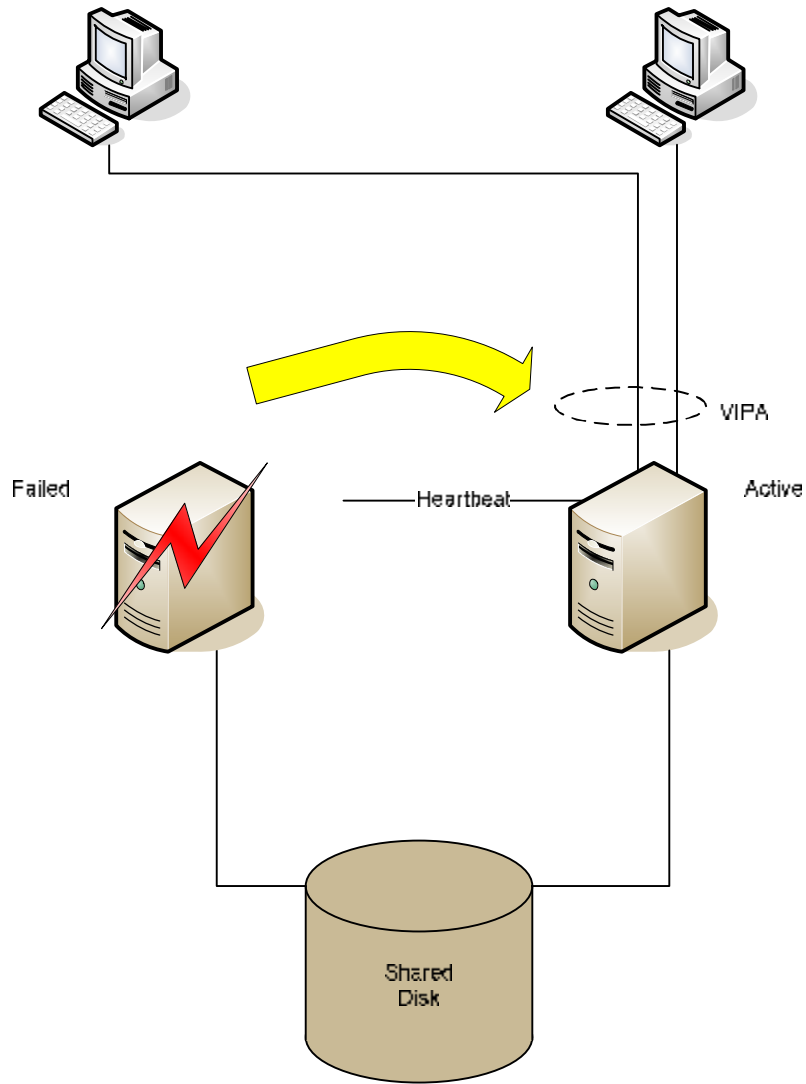


Figure 2 – Shared Disk HA Failover

However, this is a solution used by many organisations across Europe and the US, especially within the banking sector.

Examples of active-passive implementations are the AIX HACMP, and DB2 UDB for Linux Unix and Windows HADR.

HADR or High Availability Disaster Recovery for DB2 from IBM works in a similar way with a primary server and a standby server. The difference here is that the primary server processes transactions and ships logs to the standby server. The standby then stores these and applies the log buffers from the primary. Whilst this results in two copies of the database, this isolates the customer from disk subsystem failures. On failover the standby becomes the new primary. HADR is a good system and one that has been deployed across many customer sites. It does, however, still rely on the Active-Passive database set-up meaning that expensive resources are left idle.

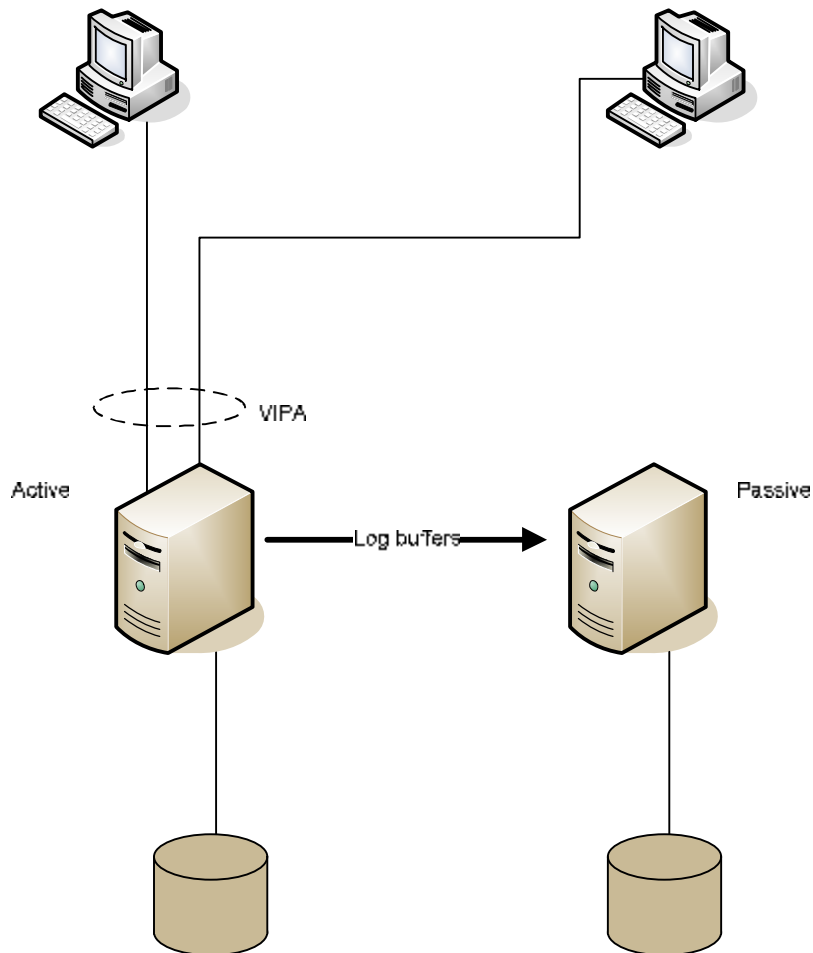


Figure 3 – HADR Configuration

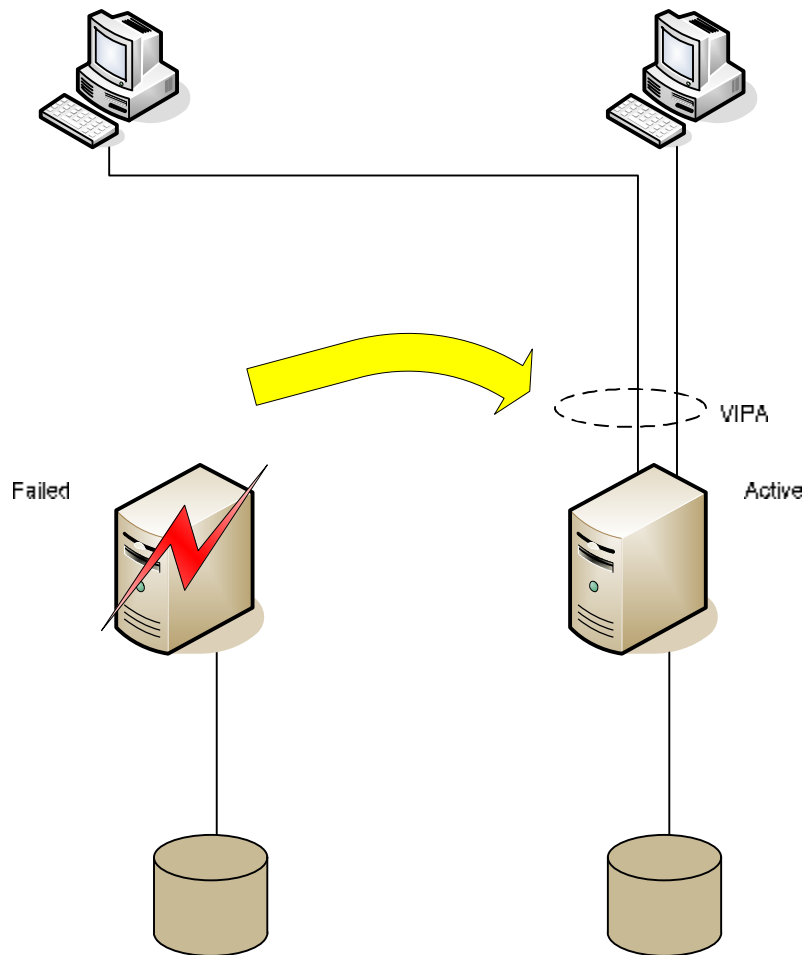


Figure 4 – HADR Failover

HACMP runs at the operating system level, with a heartbeat signal ensuring that the services are still available. The heartbeat can be implemented over the network, or through a serial connection or even shared disk. If the passive server does not receive regular heartbeats from the active server, it will take over services.

Services are provided to networked requesters over a virtual IP address (VIPA), and it is this which is moved over in the event of take-over processing.

Note that HACMP solutions usually utilise a shared SAN solution, so that the database is as up to date as possible. When the heartbeat is lost, the active server must assume that it has lost connectivity and start closing its services, to ensure that they can be successfully restarted on the passive server.

Similarly, the passive server must wait for a pre-arranged period to ensure that the active server has completed shutdown processing.

The total delay, then, from loss of service on the primary, and restoration of service on the secondary can be several minutes.

Note also that takeover does not occur on the first lost heartbeat, but typically the third. This is to ensure that network or server workloads do not cause “false” takeovers.

To mitigate against network load issues, the heartbeat is usually delivered over separate physical adapters through a private physical LAN.

HADR is a similar technology to HACMP, but is implemented in the database server, rather than the operating system. The reliance on shared SAN is dropped, with the active database shipping log buffers to the passive copy to apply. These buffers are then applied on the passive copy, ensuring that it is kept nearly in sync with the primary copy.

Note that HADR relies on automation to affect the switch over from the primary to the slave.

Peer to Peer Clustering, or 2-Way Replication allows two or more active database servers to provide read / write access to application data. Data updates are delivered over the replication solution to the other members of the replication cluster in an asynchronous manner – i.e. transaction performance is not impacted, but a finite time exists between the updates appearing on the source and target servers.

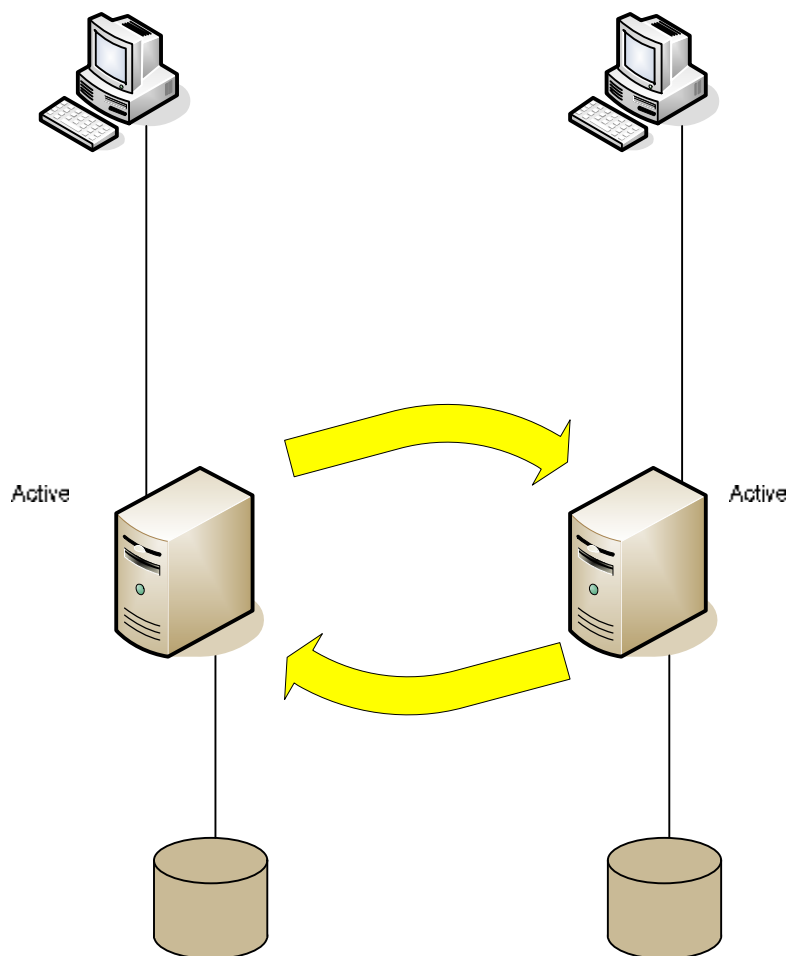


Figure 5 – 2-Way Replication

As there is no shared locking strategy, the weakness of this solution is that the same data can be updated on two replication cluster members at the same time leading to data collisions. An example of this may be that a room booking system is updated by two people – the CEO and the cleaner. Both book a room for the same time, the cleaner from the London office, and the CEO from the Edinburgh office. The CEO's booking commits on the Edinburgh server and is replicated to the London office as the cleaner's booking commits from the London office and is replicated to Edinburgh. Which booking ends up being applied will depend on how conflicts are resolved by the

replication tooling. Typically, it is the last update that wins, and whilst this could lead to some red faces in our example, the issues are more marked with, e.g. a financial services system.

To overcome this problem, customers will often logically partition their data, so that updates are applied on a regional basis, removing the risk of a collision. Whilst providing a solution to the immediate problem, management of this solution can be awkward with different business units having different service requirements, and changes in regional responsibilities can be difficult to implement.

Examples of replication tools that would support this sort of solution are DPROF and Informatica.

DB2 for z/OS Data Sharing is an all active, shared memory clustering solution based on the zSeries Parallel Sysplex technology. The parallel sysplex coupling facilities are used to cache locking information and buffered data, making these available to all of the members of the cluster.

This is the pinnacle of high availability solutions for DB2, additionally supporting seamless capacity upgrades as well as a 99.999% up time with a mean time to failure of 60 years.

Mainframe technology has been focused for some time now on high availability and zero outage solutions, and the combination of parallel sysplex, DB2 data sharing and DASD mirroring technologies has combined to provide a robust solution platform.

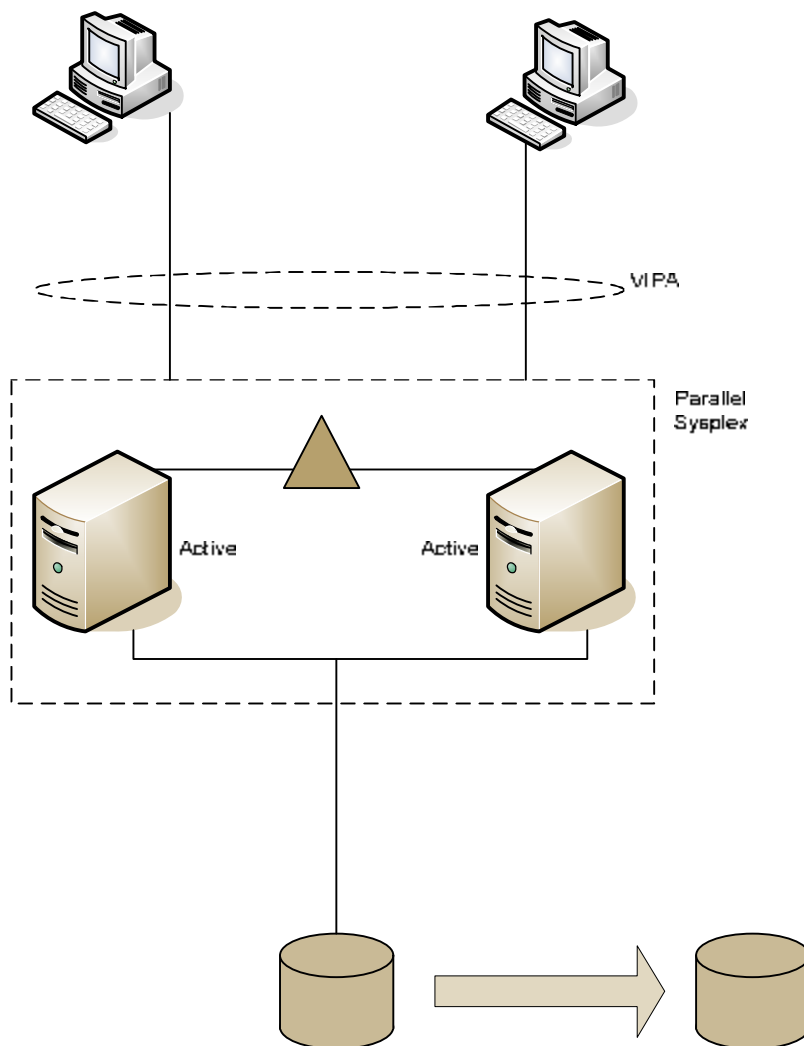


Figure 6 – DB2 for z/OS Data Sharing

A New Solution

Database virtualisation utilises an integration server in front of multiple full copies of the database to provide a single view to the application server or client. Updates are written to all of the copies, whilst reads are handed off to the server with the best resource availability and anticipated response times.

Consistency is managed by the integration layer, which can apply logged updates to database servers that are temporarily withdrawn from service – whether this is planned or unplanned.

Virtualisation allows the customer to implement a lower cost infrastructure to support their workloads, utilising multiple low cost servers to provide the same read capacity as would have traditionally been supplied by a single large server.

“Database virtualization provides you with the ability to access your data, all of the time, no matter where the data is physically stored. Database virtualization provides fast access, even in the event of a database server failure, and also provides the ability to add capacity on demand as your workload grows. If you have a system that needs 100% uptime, you need to look at database virtualization.”
Dwaine Snow – Senior DB2 Evangelist, IBM.

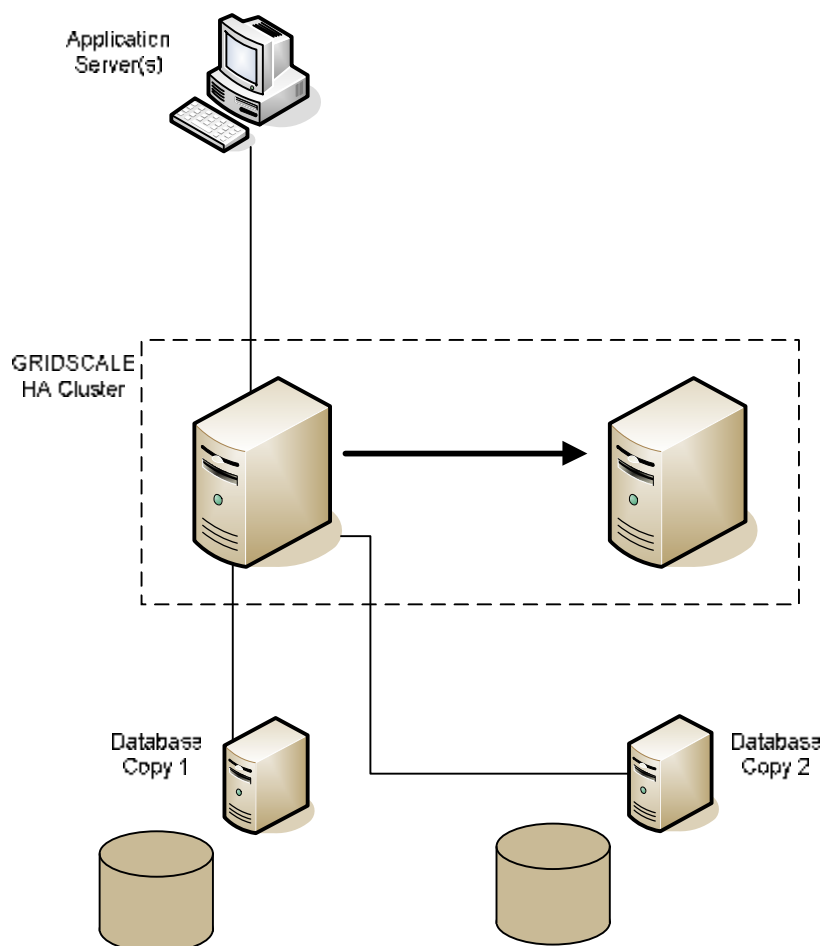


Figure 7 – xkoto GRIDSCALE Database Virtualisation

Because all servers have to process updates, the profile of the application is important when considering this method.

A further benefit of this strategy is that additional capacity can be introduced by increasing the number of database servers.

"My thoughts are that this is how databases will be implemented as time passes. The days of the tight cluster and single database instance are coming to an end." Robin Bloor – Founder, Bloor Research.

Availability into the future

Looking forward it is certain that our need for availability will only grow. Downtime and outages will become less and less acceptable to users. In this time of mergers and acquisitions, corporations across the world are needing to join up their IT systems and work with users in disparate locations. All of this points to a growing need for availability solutions which can span geographies and keep applications available to users across the globe 24/7.

xkoto Inc and Triton Consulting, leading DB2 and Information Management specialists, have joined forces to jointly deliver continuous availability and disaster recovery solutions for critical applications for the UK market. This partnership combines xkoto's revolutionary database virtualization technology with Triton's industry renowned experience working with DB2 and other IBM Information Management products for a wide range of UK companies.

For more information visit <http://www.triton.co.uk/continuousavailability/>