

Volume 3 / Issue 1 2008 - Features

Remote Data Replication

Making a Virtue of Necessity

In the previous issue of Healthcare IT Management, we noticed the growing compulsions faced by hospital IT managers to ensure successful backups of their data – especially in the face of the expected acceleration in data creation as a result of modernisation and e-Health programmes. One strategic approach to the issue of disaster management and business continuity involves low-cost remote data replication. This is in some senses fast becoming a twin, the other face, of disaster management.

The remote data replication landscape, like much else in the IT world, has been marked by significant, mainly incremental technological change – largely to make choices easier and more user-friendly as well as cost-effective. Making all this possible is a fall in costs of storage and network bandwidth, as well as optimisation techniques such as intelligent compression, also known by the more unwieldy term – data ‘deduplication’.

Today, remote replication rarely refers to the synchronous replication of all data – with every input to the database mirrored, as it happens, to the remote site.

Instead, remote data replication is part of a wider approach to data and IT systems management. It is used, for example, in such as consolidating data as well as building datawarehouses. For vendors of remote replication solutions, these direct value-add, user-facing facets are accompanied by taking upfront account of issues like efficiency in the use of network bandwidth, system downtime and flexibility. Yet another trend is the coupling of replication with continuous

data protection (CDP) technologies – since replication is, as its name suggests – little more than providing a mirror of the data somewhere else, and this does not eliminate viruses and Trojans or other threats.

One of the key differences between remote data replication and other forms of disaster management is that the latter are essentially host- or storage-based. This entails dependency on operating systems and architecture, and often therefore on specific vendors – and then not necessarily just one. Remote replication, by contrast, is based on the network.

Storage-Array Systems

There are essentially two approaches: storage array- and fabric-based.

In the past, array-based solutions were inflexible – in other words, the onsite storage (at the hospital) and offsite replication systems had to be compatible. This usually meant they had to be from the same vendor.

However, the situation has been changing, largely due to consolidation in the storage industry; there has been considerable M&A, as well as pressure from users reluctant to change their onsite systems. Many specialist storage vendors now ensure that their remote site works with onsite storage arrays from leading IT systems companies such as IBM, Sun, Oracle or Hitachi.

Fabric-Based Remote Replication

The latest development is network fabric based replication. This is principally by means of enabling software embedded as switches within the SAN (storage area network), which do not have any significant impact on network performance.

Fabric replication also serves to eliminate a ‘spaghetti’ of backup solutions inherited from previous cycles of technology, which like much else in the legacy world is a day-to-day challenge for IT managers.

At present, several specialised vendors are researching means to provide elements of ‘intelligence functionality’ to their switches. These so-called second-generation solutions can scale from 16 to 256 or even 512 ports (in a dual chassis configuration) and provide for larger levels of

storage with greater speed and control.

Large players such as Cisco are adding Quality of Service (QoS) offerings to their SAN switches to dynamically differentiate and prioritise storage traffic based on the specific requirements of the data. For example, QoS capabilities provide priority to latency-sensitive applications such as online transaction processing (OLTP) in favour of throughput-intensive applications such as data warehousing.

Asynchronous and Synchronous Replication

As with much else in technology (and real life), there are trade offs in the manner by which remote data is replicated.

Synchronous replication entails real-time, end-to-end elements. The data has to be transferred from source to destination, and this has to be acknowledged – before, so to speak, the next data shipment is packaged and delivered. The two principal limitations here are network disruptions and distance, and the use of synchronous replication is usually confined to a remote storage site in the basement of a hospital, or a data centre next door.

In the case of asynchronous replication, data is transferred to a local server, which acknowledges receipt. The next step – transfer to the remote storage facility – is done when time and bandwidth permit. Asynchronous replication is also more robust as far as network disruption is concerned since a local copy of the data is maintained on the local server until the network is restored.

Some of the newer solutions also return to only a fixed percentile of the interrupted transfer, to eliminate the need for a full-scale restart of the replication process.

Real World Choices

Many users (including hospitals) have resorted to a mix of synchronous and asynchronous remote replication. They have also assigned priorities for replication to conserve bandwidth, with mission critical information taking centre-stage while other forms of data are queued for remote replication, if required synchronously, during the night shift.

Such real-life choices are further facilitated by using solutions which only replicate changed blocks of information once again, with the option of night-time, synchronous replication of the entire databases.

Vendors are now seeking to allow such priorities and schedules to be set (and reset) via a central console, which also manages all host bus adapters (rather than having to do so separately from individual servers).

New-generation solutions also offer a host of other user-friendly features such as GUI interface and single screen setups, as well as a simple choice for implementing ad-hoc overrides – an all-too-common reality in the life of healthcare IT managers.

Published on : Thu, 3 Jan 2008