

en.community.dell.com

eHarmony Loves SharePlex for Data Replication

by *DELL-Bill Br*

5-6 minutes

Online Relationship Site Trusts Dell Software to Match 20 Million Profiles

In all the years I've been working around databases, I've never heard anybody say that data replication or information availability put them in the mood for romance. That's because not many things in IT are romantic, of course, but thanks to our [recent customer success with global online relationship site eHarmony](#), at least they're moving in the right direction.



eHarmony has implemented [SharePlex replication and data integration software](#) to keep its secondary database current within seconds of the production customer database. I'll explain why that's important and why eHarmony has entrusted replication of the data on its 20-million-plus users to SharePlex.

Like most organizations that deal in big data, eHarmony has an entire team constantly running database-intensive analysis and reporting tasks. The company conducts ongoing research on its data to personalize services for customers, inform its business decisions and keep its matching model current. If researchers did their work on the main production database, they would slow down the user experience for online customers, so instead eHarmony replicates to a secondary database that is inaccessible to users.

The biggest problem in replicating a database is managing it with reasonable costs. There's a lot of work (see below) involved in finding, sending and verifying the changes in, say, a 20-terabyte production database that users are constantly updating, and that work takes time. The greater the lag, the more stale the data in the secondary instance becomes, so the replication software needs to work fast.

Remember: replication isn't about backup, which you can get away with now and then throughout the day. Replication is continuous and runs in near-real time.

eHarmony was using Oracle Streams, a mechanism for sharing and propagating data among databases, for their secondary database, but they were unhappy with the lag time they saw. Worse yet, Oracle announced that Oracle Streams would not be supported beyond Oracle Database 12c. eHarmony saw the writing on the

wall and looked for an alternative. They evaluated Oracle GoldenGate, a high-performance application for real-time data integration, but the software license was expensive and the consulting fees even more so.

Ultimately, eHarmony chose SharePlex, which satisfies their replication requirements at about one third the cost – both up-front and ongoing – of Oracle GoldenGate.

Why is SharePlex such a good match for eHarmony?

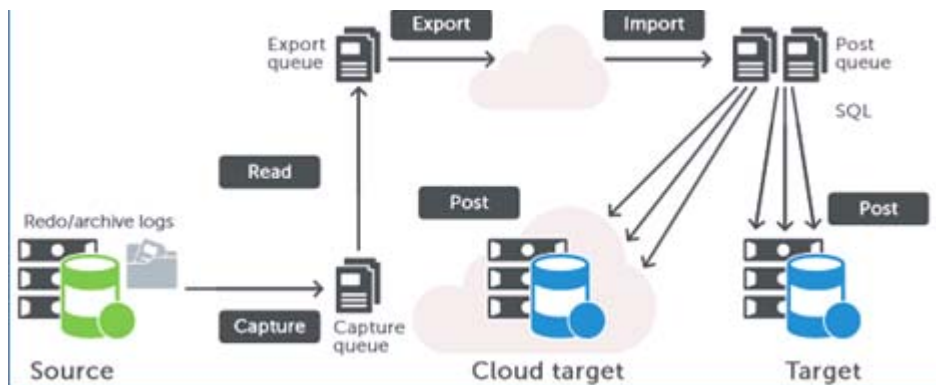
SharePlex replicates only the changes made to the source data, which is easier said than done. I'll give you a look under the hood.

For reliability, SharePlex uses a series of processes to identify, send and verify changes between databases:

- Capture – Reads the redo logs or archive logs on the production database for recent changes, then writes the data to the capture queue.
- Read – Reads data from the capture queue, adds routing information and sends it to the next queue.
- Export – Reads data from the export queue, then sends it across the network to the secondary database.
- Import – Receives data at the secondary database and builds a post queue.
- Post – Reads the post queue, constructs SQL statements to apply or “post” the replicated changes to the secondary database.

This diagram depicts these processes, with replication from a production (source) database to secondary databases (target) in and/or out of the cloud:





For speed, SharePlex uses an asynchronous stream protocol with TCP/IP connections for Export and Import that are efficient for large data transfers and tight on communication bandwidth. Instead of working on a commit or refresh schedule like other replication products, SharePlex replicates changes as they occur, which avoids spikes in network performance.

This design results in minimal lag time between production and secondary databases, which is a huge priority for eHarmony.

Once they had implemented SharePlex and seen how well it met their business needs, eHarmony evaluated and purchased [Toad for Oracle](#), an environment its database developers use to analyze subscriber and operational data. Best of all, they worked with Dell Financial Services to spread out their investment in SharePlex and Toad over one year.

So, are data replication, information availability and everything else about IT romantic? Well, maybe not quite. But eHarmony still thinks it's pretty cool that they can keep data on more than 20 million users synchronized across databases for one third the cost of Oracle Streams.

- Read the entire case study, "[Online Matches Made in Heaven.](#)"
- Want to look under the hood some more? See "[How SharePlex Works](#)" in the SharePlex for Oracle Administrator Guide.