Data Replication Buying Guide

YOUR QUESTIONS ANSWERED

How to Choose a Data Replication Solution

IT professionals are increasingly turning to heterogenous data replication to modernize data while avoiding the costs and risks of migration projects. Data replication has many benefits, including making data available to the latest and greatest applications while leaving mission-critical data flows in place, moving queries off legacy systems where they are most costly, and reducing stress on IT resources (read: you).

So, how do you choose the right replication solution for your needs?

Stelo put this guide together to help you do exactly that. This document will help you develop a short list quickly, ask probing questions to test suitability, and to ultimately decide if a data replication solution will meet your needs both now and in the future. Before delving into the specifics, here are few truisms about sourcing software as they relate to data replication:

One

Ask The Right Questions

Just read on. We have done the hard work for you in this document.

Two

Know Your Buying Criteria

When it comes to data replication, most IT pros value reliability above all with total cost of ownership, performance, and deployment speed not too far behind. It is important to determine and weight your buying criteria before you source a solution.

Three

Speak With Current Customers

Current customers are a great source of information even if you know vendors will connect you with only their happiest clients.

Four

Try To Break The Software During Your Trial

When doing a trial, put the software through every potentiality for your use-case. These stress tests can help you ensure that there are no surprises post-purchase and help you better understand the ins and outs of the software. If you break it, don't buy it!

This document will help you efficiently separate the contenders from the pretenders and find a robust and reliable data replication solution that meets your needs. As you go through this process, we hope that **Stelo's** solution will make your short list as we believe we offer the best tool on the market. In the meantime, happy replication hunting…

Product Features + Performance

Do you support my replication pairing?

This is the most fundamental question and a quick way to develop your short list. Almost all data replication solutions support Oracle, IBM, and Microsoft, but less popular RDBMS such as Informix or modern data stores such as Snowflake or Delta Lakes may or may not be supported by a specific vendor.

Data Mirroring or Streaming replication?

Transactional databases rely on data mirroring to populate data warehouses or to off-load reporting, but this is not always appropriate for no-SQL environments where it is crucial to stream large amounts of data quickly and reliably for analytics, and AI. Instead of building a data ecosystem with a patchwork of replication tools, it is vital to find a unified data replication solution for both your existing relational databases and future data streaming requirements.

Where does the processing happen?

Replication requires substantial processing power and should be offloaded to an intermediate server. This minimizes load on your source and destination systems and has the added benefit of maintaining a unified copy of your data in case of a database failure. Replication solutions that do not utilize an intermediate server should be avoided as they are inherently less stable and will cost you far more time, money, and problems over the long term.

How long does it take to deploy?

This is a key area where you can separate the wheat from the chaff. The best replication solutions deploy in under 24 hours. The worst require a week or more of configuration. Deployment time is also a good way to assess reliability. If the initial configuration is complex and time-consuming, you can expect stability and performance problems to arise in the future. Benchmark copies of data should not require more than 24 hours to create unless you are replicating an unusually large amount of data. If initial customizations take a lot of time upfront, you can expect similar lags when customizing in the future as well. High performance software does not require much tweaking and you should be worried about the reliability and accuracy of solutions that do.

What is your latency/how real-time is real-time?

Change data capture strategies often demand a tradeoff between latency tolerance and change volume requirements, so finding a solution that provides the ideal balance between fast, high-volume performance and lossless data capture is essential. The best real-time solutions maintain latency of under one second. Anything more than that and currency problems are likely to arise.

For No-SQL environments where cost can be a significant factor, your data replication tool should allow dynamic tuning to either reduce latency or minimize compute resource utilization (aka cost) as your data requirements and budgets shift.

Can I customize what I replicate?

Make sure to ask if you can choose which rows, columns, or tables are replicated. Choosing what you want to replicate reduces load and ensures that data is shared appropriately. The very best solutions let you not only choose what data to replicate but let you transform it as well. This ETL functionality can solve all sorts of data integrity problems and resolve format differences easily. Avoid solutions that simply replicate everything in bulk as this creates security challenges, increases processing requirements, and leads to additional data management costs

How does your replication solution react to schema changes?

Some replication solutions are responsive to schema changes, others are not. You want software that detects changes in real-time and replicates them automatically. When other DDL alter processing occurs, you want a notification alert to merge or reject each schema change. If your solution does not react to schema changes, you can expect a lot of wasted time due to avoidable rebuilds.

What happens if connectivity is lost?

Some replication solutions require a full rebuild of the pairing, which is the kind of nightmare that keeps IT professionals up at night. The best solutions treat the unified copy on the intermediate server as the master and automatically resume replication when the source and/or destination databases come back online. These solutions can become a key component of your disaster recovery strategy as well.

Support

How good is your support?

No vendor will say that their support is anything less than stellar so it can be tricky to assess the quality of support before purchase. To start, ask for key performance indicators such as average support ticket resolution time and average years of experience on the support team. If the vast majority of support tickets (>90%) are not resolved in under a business day, take a moment to imagine a multiple-day interruption to your mission-critical data flows and move that vendor to the reject pile.

How proactive is your support?

You want an expert that will take ownership of your challenge or obstacle, not simply give you information so that you can hopefully handle the problem on your own. You want a support team that is easily reachable via email, phone, or web and utilizes a modern support portal such as Zendesk or Freshdesk so that you can track progress and review your support history as needed. The best support teams will utilize remote desktop sharing so that they can walk you through or even execute the fixes that you need. This kind of responsive, expert, and efficient support can turn a potential catastrophe into a minor bump in the road. Too much depends on your data to expect anything less.

Is your support end-end?

Heterogenous data replication is inherently multi-platform and that runs counter to the vendor lock-in strategy of Oracle, IBM, and Microsoft. They want your organization to work 100% within their ecosystem and, if you do not, you are likely to end up half supported when it comes to data replication. For example, IBM Infosphere supports replication from IBM Db2 to MS SQL Server but their support team will offer only limited SQL Server expertise at best, leaving you to troubleshoot the SQL Server side of your replication pairing on your own. Working with data from multiple source formats can be challenging so make sure that your replication vendor offers the expertise to support you end-to-end.

Do you offer 24x7 Support?

Around the clock support is essential for mission-critical deployments. Make sure your data replication partner is always available to keep your operations running smoothly no matter the hour.

Cost of Ownership

What is your pricing model?

Some data replication tools use the core count of the source and destinations to determine pricing, but this can be costly when you are only replicating a small dataset. Others lock replication pairs, so your options are limited to a singular use-case. The primary reason why software vendors price by the source system or constrain replication pairs is that it can increase prices exponentially with minimal additional costs for the vendor. In other words, good for them, bad for you. Depending on your data architecture, software priced according to load could save you tens or even hundreds of thousands of dollars versus source systems pricing.

What if my needs change over time?

Make sure that the replication tool meets your current requirements and allows for future use as new database platforms emerge on the market. Data ecosystems inherently shift over time, and you want a solution that can move your data anywhere it is needed without breaking the bank or reinventing the wheel. The right replication partner encourages you to make improvements by integrating technologies that allow you to use your data better, not lock you in to stale solutions.

Does your software require a dedicated engineer?

Some data replication solutions require that you pay for a dedicated engineer, which increases cost substantially and suggests that the software may be buggy. Good replication software does not require a babysitter.

What is your total cost of ownership over five years?

This question will help identify any other hidden costs and to accurately project the cost of a solution. You should mention any foreseeable changes in your needs to accurately project how costs may change over time.

Conclusion

If you ask all these questions, you should find a solution that meets your buying criteria. Again, be sure to speak with current customers and to stress-test the software during your trial. We hope that this buying guide has been helpful.

Our team believes that our product, **Stelo Data Replicator** is the best data replication software on the market, and we can't wait to help you get started. If you would like more information about Stelo, please contact us at **stelodata.com** or call **+1 415.669.9619**.

www.stelodata.com
+1 415.669.9619