White Paper

embarcadero®

Ten Questions to Consider When Choosing an Embedded Database

Embarcadero Technologies

Americas Headquarters
100 California Street, 12th Floor
San Francisco, California 94111

EMEA Headquarters
York House
18 York Road
Maidenhead, Berkshire
SL6 1SF, United Kingdom

Asia-Pacific Headquarters
L7. 313 La Trobe Street
Melbourne VIC 3000
Australia

The Truth About Embedded Databases

While large enterprises require complex database management systems that can support thousands of users, manage terabytes of data, and need a highly trained support staff, the needs for embedded databases are quite different. Embedded databases are typically tightly integrated with the application and are usually hidden from the application's end users. The embedded database not only needs to offer high performance and small size, but also needs to be highly reliable and self-tuning. Furthermore, embedded databases may be deployed in "unmanaged" systems or devices, making automatic crash recovery and data security features especially important. Finally, an embedded database must be relatively low-cost to meet the price/performance requirements of the end application.

This paper examines some of the features that are crucial to the success of developing with and deploying an embedded database. It will highlight key features of InterBase® and contrast InterBase with other database alternatives for embedded applications.

10 Key Questions To Consider When Choosing An Embedded Database

- Is the database high performance? Will it scale to meet my application demands?
- Does the database protect data from loss in the event of a system crash or failure?
- Can data be secure in deployed environments? Is it secure when communicating with other devices?
- What are the system resource requirements for the database?
- Can the database be silently installed and run?
- Does the database offer a range of connectivity options?
- Does the database have Unicode support for international deployments?
- How easy is it to identify and diagnose performance problems?
- What are the licensing costs and options?
- What level of support is available from the database vendor?

InterBase: The Ideal Embeddable Database

InterBase combines simple installation and automatic crash recovery, with a self-turning engine that makes it ideally suited for embedding with a deployed application. It is ideal for OEMs, VARs and ISVs who require a database that is very compact in memory and disk size without compromising sophisticated features to support embedded database applications.

The following section examines how InterBase stacks up against the key questions every developer should ask regarding their embedded database technologies.

Performance and Scalability

InterBase is one of the fastest native databases available, and its performance scales well on higher-end system with its built-in SMP (symmetric multiprocessing) capabilities. InterBase scales to multiple CPUs, including multi-core CPUs, with up to a total of 32 CPU/cores per machine (support for 8 CPU/cores is included with the standard license). While other databases may have SMP support, that support may require additional licensing. InterBase also includes a number of performance and optimization features to help ensure maximum database performance no matter the range of system configurations it may be deployed on.

Crash Recovery

In the event of a system crash or failure InterBase recovers automatically when the system is restarted. There's no additional manual restart process or system maintenance required. When InterBase is restarted, all active transactions are automatically rolled back and the database is returned to a logically consistent state without any additional action on your part.

Data Security

InterBase provides security and encryption features for enhanced data protection. Database and Column-level encryption allow encryption of the entire database and/or individual columns in a database. InterBase also supports encryption for backup files to ensure ultimate security.

InterBase supports both strong encryption (AES) as well as weak (DES) encryption standards (128/192/256-bit for AES and 56-bit for DES). AES is the latest strong encryption standard that is being adopted worldwide. AES is fast in both software and hardware, is relatively easy to implement, and requires little memory.

Over-the-wire (OTW) Encryption encrypts the data as well as any other communications (commands, etc.) between servers and clients. This is important if there are concerns about unauthorized parties or systems seeing this communication.

System Resource Requirements

Memory and disk space requirements for embedded databases are important because they impact the end-user's minimal system configuration, hardware resources, and ultimately the purchase and adoption of an embedded application. A database with a small footprint not only minimizes system requirements, but makes it easier and less time consuming to install – which is especially important if the application is delivered via the web.

The InterBase Desktop Edition uses less than 4MB of RAM at runtime. And a compact installation of the Desktop Edition requires only 19MB of disk space (compare that to some alternatives that use over 1GB). The ultra-compact InterBase ToGo Edition is a

completely embeddable, full-featured, zero administration database that enables OEMs, VARs and ISVs to bring their applications and solutions to market faster. The ToGo Edition allows deployment of applications that utilize the InterBase engine DLLs. No separate database server process is running; the database engine is executing in the application process space. InterBase ToGo provides the benefits of InterBase in an extremely small footprint of less than 3MB.

Silent Install & Run

Whether you're deploying an application, or providing an appliance or device with an embedded database, you don't want the end-user to even be aware that a database is being installed or run. InterBase provides silent installation and licensing so end-users of the embedded application will never be aware of, or have to care about, the database component.

Minimize Development Time

A major component of database TCO is the cost of developing, maintaining and enhancing your application. If your database does not provide the tools and features that make development easy, your development team can waste a lot of time coding around the limitations of the database. InterBase provides a rich feature set that makes developing any application fast and easy. Here are just a few:

Database Connectivity Options

Developers using RAD Studio, which includes Delphi®, C++Builder®, and RadPHP® and a multitude of other tools can leverage a wide range of standard database connectivity options through ODBC, JDBC, ADO.NET, and IBX.

Identify Performance Issues Prior to Deployment

The worst time to find out about performance issues is after the application has been deployed. InterBase includes a sophisticated performance monitor called IBConsole that offers both graphical and SQL interfaces. Not only can you see exactly what is happening in the InterBase database at any time, you can also log selected parameters over time for later review. This makes diagnosing performance problems a snap.

Should you encounter reports of performance issues after the application has been deployed, IBConsole can also be used to remotely connect to the database and provide information about its status and activities including:

- A summary of database threads, procedures, transactions and memory use
- View memory use by graph, pool or heap
- List all the tables and views in the database
- Display a list of all recent transactions
- View a list of all recent procedures and triggers
- Display the InterBase log file and statistics list

VENDOR TECHNICAL SUPPORT

Reliable technical support is a basic requirement when working with database vendors. In today's highly competitive business environment, reliable, multi-channel (phone, email, and web), support is imperative.

Open source database solutions are not backed by organized technical support. At best, these projects provide online forums where developers can submit questions and bug reports. The developer has no assurance that a response will be provided, and no contract is in place that regulates the response time of the resolution. The organization must use its own development resources to solve issues relating to installation, configuration, performance, and reliability.

From the time you begin your application development and throughout the life of deployment of your application (which may include mission critical applications) the support from your database vendor should be available whenever you need it. Embarcadero offers a full suite of support options from 'per incident' support to 'around the world around the clock' support agreements.

INTERNATIONAL DEPLOYMENTS: LANGUAGE SUPPORT

InterBase supports Unicode, the universal character set, plus other international character sets including support for International character sets via Unicode (UCS-2) and UTF-8. Also included is support for additional character sets including ISO8859_2 (Latin2), ISO8859_15 (Latin9), and KOI8-R. Unicode and International character support allows InterBase to be deployed to multinational / multi-language markets.

DATABASE MAINTENANCE

A database that requires routine maintenance is simply not an option for a variety of embedded database deployment scenarios. Database maintenance also adds to costs. You may be able to automate the maintenance tasks but automating routine maintenance adds to development time and cost. If the database requires maintenance that cannot be automated, you will need staff to perform the maintenance tasks and the cost will be ongoing.

InterBase requires no maintenance. InterBase has features that are designed specifically for use in applications where no DBA or other support is available. Because InterBase is self-tuning and self-maintaining there is nothing for a DBA (or developer or user) to do during normal operation.

MULTI-PLATFORM SUPPORT

If you're developing applications for multiple platforms, you want to minimize development costs by using a database that's supported on multiple platforms. InterBase is available on Windows®, Mac OS®/X, Linux and Solaris™ to give you platform flexibility now and in the future.

LOW COST LICENSING & EASY DEPLOYMENT

Licensing costs for InterBase are among the lowest of all commercial databases. OEMs, VARs, and ISVs can easily embed and deploy InterBase Desktop Edition, InterBase ToGo Edition or InterBase Server Edition with their applications and end-users will never have to deal with the hassle of a separate database installation. Utilizing InterBase's "silent install" bypasses the normal registration process for seamless installation with the application.

The InterBase ToGo Edition is perfect for small devices, public kiosks, as well as VAR and OEM applications. The InterBase SMP ToGo Edition is a truly embedded database, provided as a DLL that runs in the application space. The ToGo Edition features a small footprint (less than 3MB) that gives Windows developers the ultimate flexibility to include an embedded, highly reliable SQL compliant, transactional database.

A LOOK AT ALTERNATIVES

There are other databases you can consider embedding into your applications but few were designed specifically for the unique requirements of embedded applications. InterBase alternatives can be divided into commercial products, such as Oracle®, Microsoft® SQL Server, and MySQL, and free, open source databases such as PostgreSQL and Firebird.

Note that while MySQL is an open-source database, in most cases it can only be distributed free of charge as part of your application if your entire application is distributed under the GPL or an OSI approved open source license. Since very few commercial applications are distributed under an open source license, MySQL is included with the other commercial databases.

Commercial databases are developed and supported by a single organization that can give you quick in-depth support in an emergency because the support staff has direct and immediate access to the R&D engineers who develop the product. If you encounter a bug or other problem with a commercial database you can often get a patch or work-around without waiting for the next scheduled product release.

Free, open-source databases appear to have a cost advantage because there are no licensing costs for development or deployment. However, the apparent cost advantages of open source databases are typically offset by higher hidden costs and significant disadvantages in terms of support.

Most open-source databases are developed and maintained by groups of volunteers. There is no single vendor organization to provide support; a commercial vendor which must answer to and serve the needs of its paying customers. Instead, you must rely on support from an open-source community and/or a third-party support organization that does not necessarily have direct access to the development team which you find with commercial products.

Because few commercial or open source databases were designed specifically to support embedded application requirements, they all have significant weaknesses when compared to InterBase. The following sections examine some of their limitations.

ORACLE

Oracle is one of the most powerful, complex and expensive databases available today. If you need to support thousands of users and terabytes of data Oracle is an excellent choice. But most embedded and smaller enterprise applications and do not need to scale to that level. Oracle has hundreds of tuning and configuration parameters and is so complex Oracle certifies DBAs at four different skill levels. This level of complexity and support is simply not feasible in most embedded applications.

Oracle Express (XE), a lighter-weight version of an Oracle database, can be deployed at no cost. But is has the following limitations:

- One database per machine
- One instance per machine
- Uses only one CPU
- Uses a maximum of one gigabyte of memory
- Database size is limited to four gigabytes of user data

In reality, Oracle XE cannot be used in any but the smallest of application scenarios due to the first two restrictions. Consider what happens when you try to add a second application to a machine that's already running another application using Oracle XE. Since Oracle XE only allows one database per machine and one Oracle XE instance per machine, you cannot install a second application that uses Oracle XE.

Your only choices are to install another application server or upgrade to the full Oracle product. Suddenly Oracle XE has become a very expensive choice. In addition, Oracle XE has a disk footprint of over 1.1 gigabytes. Compare that to InterBase ToGo Edition which requires less than 3MB and InterBase Desktop and Server editions which only require about 4MB.

MICROSOFT SQL SERVER

Microsoft SQL Server is another database with the features and complexity necessary to handle very large numbers of users and very large databases. SQL Server uses a

complex locking model that can cause poor concurrency if developers do not understand how to avoid its traps.

Developing applications with Microsoft SQL Server typically requires substantially more code compared to InterBase. One reason is Microsoft SQL Server's trigger implementation. Microsoft SQL Server does not have before triggers. If you need to calculate or modify values being inserted or updated, you must create an instead of trigger and write the code to perform the insert or update in addition to your data validation or calculation code. However, even this option may not be available. "Instead of, update, and delete triggers are not allowed on tables that are targets of cascaded referential integrity constraints". In addition, only one instead of trigger is allowed for each action so you cannot modularize your code. Microsoft SQL Server triggers fire once per statement, not once for each affected row. So in addition to the code you would have to write with InterBase, you must declare a cursor and write code to iterate over the affected rows. Microsoft SQL Server 2008 Express Edition can be deployed at no cost but it has limitations.

- Uses only one CPU
- Database size is limited to 4 gigabytes
- Buffer pool memory is limited to 1 gigabyte

Many of today's applications must store more than four gigabytes of data. If you need to store BLOB data such as text documents, scanned images, photographs, sound or video, you can fill a four gigabyte database very quickly. If there is any possibility that any of your applications may exceed these limits you will have to monitor your installations carefully so you can upgrade to the full version of Microsoft SQL Server before you run out of space in the database or performance degrades. Since SQL Server 2005 Express Edition is just a limited version of Microsoft SQL Server, developers require the same level of knowledge and skill to develop applications using the Express Edition that they must have to work with the full product.

MYSQL

MySQL is not a typical database management system but rather a front end for a family of nine different storage engines. All of the storage engines have different table formats, concurrency control mechanisms, and other features. This means that you cannot simply decide to use MySQL. You must evaluate the storage engines against your application's requirements and select the one whose characteristics best meet your application's needs. For example, the MyISAM engine was designed for fast processing of SELECT statements but uses table locks for insert, update and delete statements. This can cause poor performance in a multi-user environment where data is being changed.

MySQL has many other limitations that make it a poor choice for embedding into applications.

- MySQL has very limited performance monitoring.
- MySQL does not support multi-instance installations. That makes migration from one version to the next more difficult and prevents isolation of one application from other applications. It also prevents using multiple instances for load balancing.
- MySQL does not have an integrated mechanism to notify clients of database events.
- MySQL does not have incremental backup. Instead, you must start a new binary log then copy the old log files to the backup location.
- MySQL does not support user-defined data types.
- MySQL includes over 100 system variables plus numerous configuration parameters that control its operation; it is not self-tuning.
- MySQL has a query cache, key cache, table cache, InnoDB buffer and InnoDB log buffer. You must decide how to allocate memory among these five caches to get the best performance.

FIREBIRD

Firebird is an open-source database based on an early version of InterBase that has limited scalability because it does not have multiprocessor support. To take advantage of multiple CPUs you must use the "classic" version, which spawns a separate process for each user. However, the classic architecture does not scale well. Having a separate Firebird server process and database cache for each user consumes resources rapidly as the number of users grows.

A more serious limitation is that Firebird provides limited data recovery in case of media failure or any other event that damages the database. You can restore from your last backup but all changes since the last backup will be lost. The only way to provide recovery with little or no data is to purchase a third-party replication tool and implement replication to another server. This will protect your data but it also increases development time and deployment cost and degrades performance.

Firebird has no performance monitoring tools. Diagnosing performance problems in deployed systems is very difficult. Tasks that can be done easily with temporary tables require additional coding when using Firebird.

Like other open-source databases, the primary source of support for Firebird is a community of volunteers and some third-party organizations. Unfortunately, third-party support providers do not have the access to or integration with the engineers developing the product; this limits both the level of support you can get and the speed with which you can get it.

POSTGRESQL

PostgreSQL is another open-source database that uses versioning architecture. However, PostgreSQL is neither self-tuning nor low maintenance. PostgreSQL has over 170 configuration options that must be set to control performance. In addition, it does not

automatically perform garbage collection as data is accessed. Instead, you must perform a vacuum operation frequently to remove old row versions from the database.

PostgreSQL logs information to a log file. However, the log output tends to be voluminous (especially at higher debug levels) and you won't want to save it indefinitely. You need to "rotate" the log files so that new log files are started and old ones removed after a reasonable period of time.

Like Firebird and other open-source databases, support for PostgreSQL is only available from third-party providers. Consider carefully whether you can get the level of support you need before you commit to any open-source database.

CONCLUSION

The embedded database you choose can have a crucial impact on development time, cost, performance, and overall success of your application. The right choice can ensure that your application avoids downtime and performance issues while satisfying the needs for reliability and scalability needed to keep your customers and users satisfied. The wrong choice can lead to longer development times, application failure, disgruntled users, and damage to your reputation as a software vendor or solution provider.

For over 20 years, InterBase has been proven in thousands of embedded, desktop, and business-critical server applications. With its high-performance, self-tuning architecture, small footprint, and easy installation, InterBase provides you with a resilient, lightweight, low-cost database platform that is ideal for embedding into applications. Key features include:

- High-performance and scalability, with built-in support for Symmetric Multi-Processing (SMP) for multi-core CPUs
- Simple installation
- Small footprint
- Silent installation
- Automatic crash recovery
- Multiple connectivity options
- Near zero maintenance
- Low deployment costs
- Strong encryption (AES), journaling, and other data protection and security features

For more information or to download a free developer edition or trial of InterBase, visit http://www.embarcadero.com/interbase.

Embarcadero Technologies, Inc. is the leading provider of software tools that empower application developers and data management professionals to design, build, and run applications and databases more efficiently in heterogeneous IT environments. Over 90 of the Fortune 100 and an active community of more than three million users worldwide rely on Embarcadero's award-winning products to optimize costs, streamline compliance, and accelerate development and innovation. Founded in 1993, Embarcadero is headquartered in San Francisco with offices located around the world. Embarcadero is online at www.embarcadero.com. 301012