Always Up, Always Running

Continuous Availability with Progress® OpenEdge™

PROGRESS
SOFTWARE

Contents

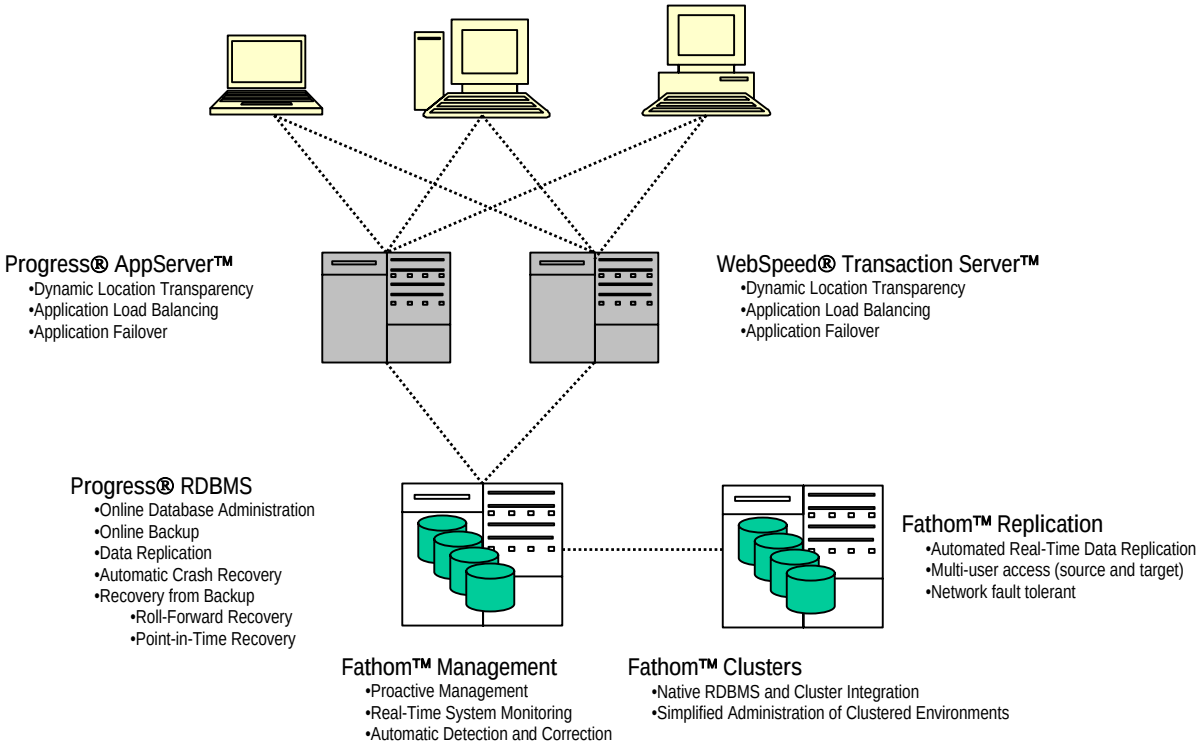
Introduction

Globalization, the need for 24x7 operations brought on by the Internet, increased competition, and increased economic pressures have amplified the need for the continuous availability of business-critical applications. For many businesses even the smallest of interruptions, such as minutes or hours of outage of vital business applications, or interruption in service from a key supplier or outside service partner, can have serious business consequences. Even in environments where continuous availability is not absolutely mandatory, frequent outages or even occasional outages lasting more than an hour or two cannot be tolerated.

As a result, businesses must plan for and mitigate the risk of any type of business process interruption. This means that enterprise applications must be able to avoid or minimize the impact of both planned and unplanned downtime. It also means that the impact from any downtime that does occur is minimized so that applications and data are available as quickly as possible.

Simply acquiring redundant hardware and installing failover software does not guarantee that applications will be available when they are needed. Applications must be built on a resilient software infrastructure that provides greater levels of availability. The Progress® OpenEdge™ business platform provides a software infrastructure that dramatically minimizes downtime and simplifies recovery, ensuring that users have access to the applications and data they need, when they need it.

Figure 1: Continuous Availability with Progress® OpenEdge™



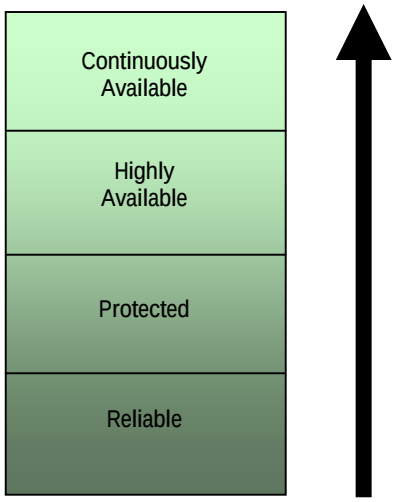
Defining Continuous Availability

In today's environment, interruption in service is becoming more and more unacceptable. Interruption in service, or downtime, can be caused by many types of events, both planned and unplanned. Planned events include hardware maintenance, repairs, and upgrades. Unplanned events include hardware failures and power outages.

There are many strategies to minimize downtime. As shown in Figure 2, the foundation for any availability strategy is to start with a reliable system, one with dependable hardware and software. For a protected environment, the goal is automatic detection and recovery from faults through redundant hardware and software. For example, a system should not shut down because of disk failure. In a highly available environment, the goal is to reduce the amount of time the application is down — fast recovery with minimal impact.

A continuously available environment gives the highest level of availability. The goal of continuous availability is to eliminate interruptions in service so that applications and data are available to users 24x7, regardless of planned or unplanned downtime. It insulates users from the complexities of back-end systems. Recovery is transparent. For example, it should not be necessary to power down the systems for maintenance or administrative activities.

Figure 2: The Availability Continuum



Regardless of where an application falls in the availability continuum, eventually some amount of downtime is bound to occur. Therefore, achieving higher levels of availability also means reducing the complexity of recovering from any outages that do occur so that access to applications is restored more quickly and efficiently.

The Cost of Downtime

Significant cost can be incurred when business-critical applications are down. The cost of downtime will vary from industry to industry, from business to business, and within a business from application to application. Downtime costs include tangible costs that can be measured in terms of hard dollars, such as lost revenue (current revenue, as well as future revenue), lost productivity, lost inventory, late fees and penalties, and legal costs. Downtime costs also include intangible costs that are not easily measured, but they impact the business

just the same. Examples include damaged reputation, impaired financial performance or share value, and decreased customer satisfaction.

Companies that are the most dependent on automated systems, such as energy and telecommunications companies, accrue an average of nearly $3 million in losses for every hour of downtime, based on lost revenue and productivity, according to an October 2000 Meta Group study. Manufacturing companies and financial institutions suffer per-hour revenue losses of $1.5 million to $1.6 million. Health care, media, and hospitality/travel companies lose between $330,000 and $636,000 of revenue per hour.[1] Table 1 shows the average cost of downtime across various industries.

Table 1: The Cost of Downtime

INDUSTRY SECTOR	REVENUE/HOUR	REVENUE/EMPLOYEE-HOUR
Energy	$2,817,846	$569.20
Telecommunications	2,066,245	186.98
Manufacturing	1,610,654	134.24
Financial Institutions	1,495,134	1,079.89
Information Technology	1,344,461	184.03
Insurance	1,202,444	370.92
Retail	1,107,274	244.37
Pharmaceuticals	1,082,252	167.53
Banking	996,802	130.52
Food/Beverage Processing	804,192	153.10
Consumer Products	785,719	127.98
Chemicals	704,101	194.53
Transportation	668,586	107.78
Utilities	643,250	380.94
Health Care	636,030	142.58
Metals/natural Resources	580,588	153.11
Professional Services	532,510	99.59
Electronics	477,366	74.48
Construction and Engineering	389,601	216.18
Media	340,432	119.74
Hospitality and Travel	330,654	38.62
Average	$1,010,536	$205.55

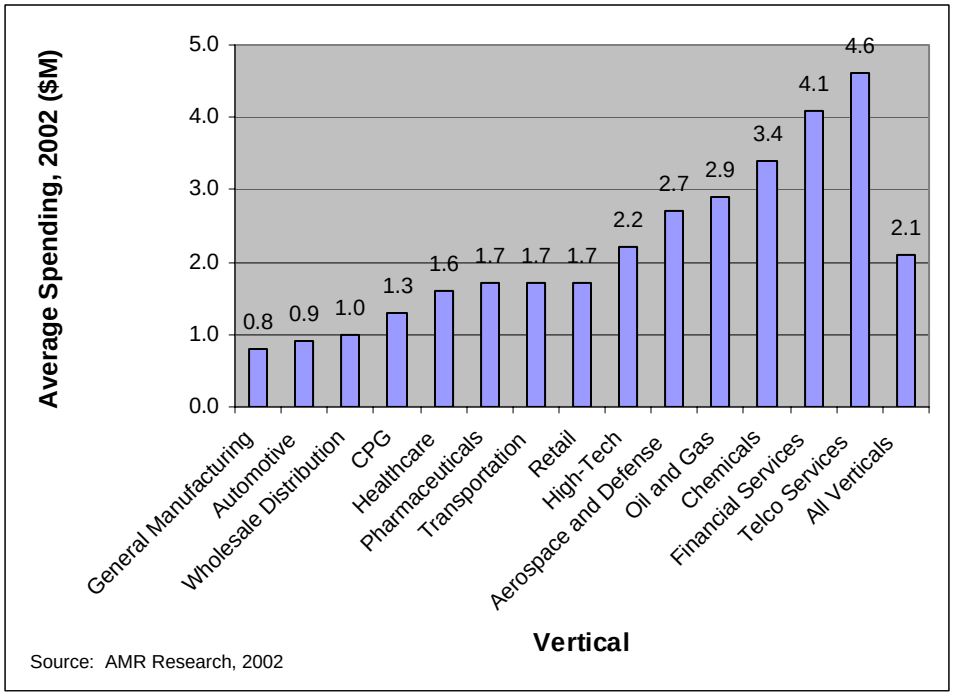
Source: IT Performance Engineering & Measurement Strategies: Quantifying Performance Loss, Meta Group, October 2000.

While downtime cost varies, it is always significant and can often be staggering. Understanding the impact of downtime can not only highlight the necessity of striving for continuous availability, but can also aid in justifying investments to ensure that applications and data are always available.

[1] Jon William Toigo, “Storage Disaster: Will You Recover?”, Network Computing, March 5, 2001

In order to minimize the risk of business disruption, many companies have made business continuity a budget line item and increased their investment in it.[2] As shown in Figure 3, the average spending on business continuity by industry ranges from $0.8 to $4.6 million, averaging at $2.1 million across industries.

Figure 3: Average Spending on Business Continuity



Understanding the costs of business downtime and taking the necessary precautions are vital to business survival. When the real impact of downtime is known, the decision to invest in increasing levels of availability becomes clear.

[2] Colleen Niven and Laura Carrillo, "Business Continuity Planning: How Much Are You Willing To Pay To Reduce Business Disruption?", AMR Research Report, 2002

Achieving Continuous Application and Data Availability

Achieving continuous availability means minimizing both planned and unplanned outages. Furthermore, it means minimizing user impact for any outage that does occur. Progress OpenEdge, including the Progress RDBMS, Progress AppServer™, Progress WebSpeed® Transaction Server™, Fathom™ Replication, Fathom™ Clusters, and Fathom™ Management, delivers capabilities that minimize planned downtime, eradicate unplanned downtime, and simplify recovery. Progress OpenEdge is the key to deploying and managing continuously available applications and data.

Minimize Planned Downtime

Planned downtime is inevitable. Certain tasks, such as software maintenance and upgrades, must be performed from time to time. This is especially true of an application's database. For many applications that are not business-critical, availability during business hours is required, but downtime during non-business hours is acceptable. However, as more business becomes more global, and as more business is conducted over the Internet, an increasing number and type of applications must be up and running 24x7. For these applications, planned downtime must be minimized or even eliminated.

Online Database Administration

Irrespective of an application's chosen database, DBAs are required to occasionally perform essential maintenance and administrative chores. Such activities include maximizing utilization of storage hardware through the optimal distribution of data and corresponding search indexes. This task is ideally achieved at an application's planning stage, but parameters invariably change as databases and their environments mature. This means that administrators may be required to reorganize data tables and indexes periodically through the life of an application.

To provide continuous availability, it is essential that such tasks be performed without service interruption. Progress RDBMS allows most common administrative tasks, including table and index reorganization, to be performed without requiring the database to be taken offline. For applications running on the Progress RDBMS, the primitive method of having users log out of core applications during maintenance and administration is now happily resigned to the history books, and expensive planned downtime is dramatically reduced.

Online Backup

Database backups provide an insurance policy against application failures and site disasters. To be able to recover reliably with minimal loss of current data, frequent database backups must be performed. This is often a scheduled task that requires the database, and the applications that access it, to be offline.

Online backup provides a mechanism to back up the database while it is still in use. Data remains fully available to applications and users. This single feature greatly reduces the amount of planned downtime. Incremental backup reduces planned downtime even further by backing up only the data that has changed since the last full or incremental backup, taking less time and returning the database to full availability for complete transaction processing faster.

Progress RDBMS provides both full and incremental online backup, eliminating the need for scheduled downtime for this recurring task. When an online backup is performed, Progress RDBMS automatically

switches over to the next after-image file so that operations can continue, seamlessly. Progress RDBMS also supports zero-impact backups using database quiet points and split-mirror backup techniques. This allows the Progress RDBMS to be used in the most demanding environments with the latest storage array technology for "non-stop" applications.

Eradicate Unplanned Downtime

While unexpected outages cannot be completely eliminated, masking outages from users creates the perception of uninterrupted availability. There are several approaches to concealing outages from users, so that data and applications are available when needed and business operations can continue.

Data Replication

Replication provides a cost-effective solution for the continuous availability of data. This approach employs two systems, the first being the primary online system and the other being a standby system. The primary and standby systems are identically configured, each with its own database. The primary system's database is automatically replicated, or duplicated, on the standby system's database. If the primary system fails for some reason, the standby system is brought online as soon as the last replication action is completed, and all transactions in process at the time of the failure are restarted on the standby system. All further requests are also routed to this system until the primary system can be brought back up and its database replicated with the standby system's database.[3]

Progress OpenEdge offers several methods of data replication. Progress RDBMS provides log-based site replication to support hot standby systems at either the same or remote sites. The replicated database can be synchronized to reflect the primary database's state within seconds. In the event of a complete site loss (e.g., fire, earthquake), a remote hot standby site can be quickly restarted to provide continuing service.

Fathom Replication extends the replication capabilities of Progress RDBMS by providing transparent, automated, zero-latency database replication from a single host database to one or more target databases. The replicated databases can be used for hot standby failover or database recovery purposes. Fathom Replication fully automates and simplifies the replication process, guarantees data integrity between source and target, and automates recovery and resynchronization. As the replication is transparent, there is no impact to the applications. In addition, with Fathom Replication Plus data can be load balanced, or moved, in real-time to increase performance and efficiency.

Whether replicating data with Progress RDBMS or Fathom Replication, current data is available to applications and users when needed. The result is minimal impact in the event of unplanned downtime or a disaster.

High Availability Clusters

Failover clusters achieve high availability through the use of redundancy. When a failure occurs at a node in the cluster, resources can be automatically redirected and the workload redistributed to an available running node. This minimizes the number of single points of failure, reducing the risk that failure of a single system will result in downtime for an entire application.

---

[3] Carl W. Olofson, "The Truth About Clustering and Other Strategies for Database High Availability and Recoverability", IDC, 2001

Fathom Clusters is a companion to existing cluster solutions to further safeguard mission critical data and ensure business continuity for Progress-based applications. Fathom Clusters allows cluster management software to monitor and manage the Progress RDBMS as a native cluster resource, ensuring high reliability and improved performance. It guarantees a certified failover environment through tight integration with existing cluster managers. With Fathom Clusters, the difficult and costly task of administering in a clustering environment is significantly simplified, saving resources, time, and money.

Dynamic Location Transparency, Load Balancing, and Failover

A critical component for business applications is the application server. These servers implement the core business logic of the application and act as brokers for access to the database. As such, application servers must also have safeguards to assure the continuous availability of business applications and data. To meet this requirement, application servers deploy their own high availability capabilities that are complementary to the balancing and clustering that occurs at the network, hardware, and database levels. OpenEdge provides this level of application and data availability with Progress AppServer and WebSpeed Transaction Server.

The Progress application server architecture for AppServer and WebSpeed Transaction Server includes a NameServer, which directs client requests to available server processes. The NameServer provides applications with three benefits directly related to continuous availability: location transparency, load balancing, and failover. Location transparency means that client applications don't have to know about all of the application servers that are on the middle tier of an application's infrastructure. The client simply makes a connection to a NameServer requesting to execute some business logic on an application server. The NameServer keeps track of where the business logic exists and then re-connects the client with that service. As additional application servers are added to the middle tier or moved to other locations, the NameServer records these events. The client application does not have to be modified. This greatly reduces the amount of maintenance and update on client applications.

Load balancing is the ability to route requests across multiple application servers. For example, in an AppServer configuration, a machine set up to process inventory requests could become overloaded. On a separate machine, another AppServer broker could be created for inventory that is delegated by the NameServer to take 60% of inventory requests while the remaining 40% remain on the first machine. This is all transparent to the client application, which greatly reduces maintenance and administration. The benefit is that the OpenEdge application servers identify and resolve bottlenecks that could lead to denial of service to a client or a long response time.

The NameServer provides two levels of failover. Within a single NameServer, two brokers that provide exactly the same functionality can be created. If the first broker fails for whatever reason (e.g., hardware failure), the NameServer can automatically redirect all subsequent requests to the second broker. Additionally, the NameServer may also be replicated so that in case of a failure connection requests can be routed to another NameServer.

Proactive Management

Another way of avoiding unplanned downtime is to proactively manage critical information assets. Diagnostic tools can help identify and correct database problems before they worsen, as well as maintain optimal levels of performance.

Fathom Management provides comprehensive tools to anticipate, diagnose, and correct availability and performance issues before they negatively impact business processes. Through real-time system monitoring, alerting, and automatic problem detection and correction, Fathom Management provides faster resolution of problems and the ability to avert problems altogether. The result is reduced application and end-user downtime. Fathom Management can also be integrated with other management tools through Simple Network Management Protocol (SNMP) to provide centralized management of decentralized services, simplifying database and system management even further.

Simplify Recovery

No matter what safeguards are in place, some amount of downtime will eventually occur. When this happens, it is vitally important to facilitate a rapid, smooth recovery to minimize any business impact. The recovery process can be complex and labor intensive. Data currency, or freshness of data, is another issue. Therefore, success can be measured in terms of system administration, recovery time, and risk of data loss. Progress OpenEdge offers several solutions that simplify reliable and effective data recovery so that applications and users are back online and fully operational as quickly as possible.

Automatic Crash Recovery

When an application performs an uncontrolled shutdown, or crash, the application must return itself to a consistent state before the application or its users can resume doing work. Progress RDBMS provides efficient and automatic crash recovery when a database is restarted after failure, without user intervention.

Recovery from Backup

The backup operation creates a copy of a database. This copy can be used to restore the database to a known state in the event of media failure or if the database is somehow damaged (for example, from user errors).

The basic recovery model allows the database to be recovered to the point of the last backup (the most recent backup). With this model, the database cannot be restored to the point of failure or to a specific point in time. So, changes since the last backup will be lost.

Progress RDBMS provides an advanced recovery model, which allows the database to be recovered to the point of failure (roll-forward recovery) or to a specific point in time (point-in-time recovery). Basically, online backups are integrated with the RDBMS after-image logging subsystem to allow for recovery of transactions that completed after the last online backup was performed. The initial database is restored from backup, and then after-image log files are applied for either roll-forward or point-in-time recovery. This recovery model provides better protection for data and assures data currency after a failure.

Automated, Zero-Latency Recovery

When instantaneous access and freshness of data are primary business and IT concerns, data must be recovered such that applications and users have access to the most current data as quickly as possible. There can be no delay waiting for the database to be recovered from backup.

Fathom Replication duplicates a source database to a target database, and maintains the target database so that it is always synchronized with the source. Therefore, the target database is an identical copy of the source database at any point in time. In the event of a system failure, Fathom Replication enables clients to connect to the target

database, providing real-time, read-only access to the target database. The target database effectively becomes the new source database. Fully automating the recovery process and guaranteeing data integrity between source and target, Fathom Replication provides the fastest and easiest recovery of current data.

Conclusion

The globalization of the economy and the rise of e-commerce have fueled changes in the way businesses compete with one another and how they sell to their customers. One of the most critical needs created by this environment is the continuous availability of applications and data.

Whatever the cause, the consequences of downtime can be devastating. When a critical business application is supposed to be running, and isn't, it can result in millions of dollars in lost revenue and productivity, as well as far-reaching damage to a businesses' reputation and financial performance.

The Progress OpenEdge business platform provides a variety of cost-effective solutions that help businesses move up the availability continuum. Minimizing both planned and unplanned downtime, and ensuring fast and easy recovery, Progress OpenEdge provides the foundation to attain increased service levels and higher return on investment. OpenEdge delivers the data and application availability you need to help keep your business *in* business.

Take the Next Step

Find out how your business can benefit from Progress OpenEdge. For more information on OpenEdge, contact your local Progress Software representative or visit the Web site at **www.progress.com**.

About Progress Software Corporation

Progress Software Corporation (PSC) (Nasdaq: PRGS) supplies industry-leading technologies for all aspects of the development, deployment, integration and management of business applications. PSC, headquartered in Bedford, MA, operates through the Progress Company, Sonic Software Corporation, and PeerDirect Corporation. PSC can be reached at **www.progress.com**.

Worldwide and North American Headquarters
Progress Software Corporation, 14 Oak Park, Bedford, MA 01730 USA Tel: 781 280 4000 Fax: 781 280 4095

Europe/Middle East/Africa Headquarters
Progress Software Europe B.V. Schorpioenstraat 67 3067 GG Rotterdam, The Netherlands Tel: 31 10 286 5700 Fax: 31 10 286 5777

Latin American Headquarters
Progress Software Corporation, 2255 Glades Road, Suite 300 E, Boca Raton, FL 33431 USA Tel: 561 998 2244 Fax: 561 998 1573

Asia/Pacific Headquarters
Progress Software Pty. Ltd., 1911 Malvern Road, Malvern East, 3145, Australia Tel: 61 39 885 0544 Fax: 61 39 885 9473

PROGRESS SOFTWARE

www.progress.com