DELIVERING ALWAYS-ON INDUSTRIAL AUTOMATION

For an **Always-On** World

www.stratus.com

Delivering Always-On Industrial Automation

Developed in conjunction with AITS Australia

Abstract

Unplanned downtime is an issue that many organizations in the Industrial Automation (IA) sector struggle with. As SCADA and other IA systems have evolved from being analog-based to digital-based, and gathering simple process data to handling more complex data at higher speed, the impact of unplanned downtime and any resulting data loss is much more significant.

To overcome the challenge of downtime, IA systems such as SCADA have become more reliant on software applications that run on industry standard servers. Moreover, knowing that the failure of a server is a major issue and the risk of downtime is unacceptable, SCADA and IA vendors have developed "redundancy" features to increase the reliability of their systems. This does not resolve the problem of unplanned downtime and what's more, it significantly increases the complexity of the systems.

These SCADA and IA systems gather data from field devices and sensors that include millisecond timestamps and rely on the accuracy of this data. The need for reliability and availability is paramount. If an Application, HMI Alarm, Historian, or a MES Server goes down, a multitude of process data points, an entire sequence of alarm and event history, or hundreds of production transactions will be lost. Production will be stopped, critical reports will not be generated and root cause analysis cannot be carried out. Even the smallest amount of unplanned downtime can escalate to costly, dangerous, or even potentially life-threatening consequences.

Prior to Stratus' continuous availability (always-on) offerings, IA Managers might have mitigated the risk of unplanned downtime by implementing store and forward processes, manual cold-backup/recovery or cluster failover systems. These are all approaches that involve recovering from failure rather than preventing it in the first place.

Relying on a recovery plan in an emergency is not good enough. It's crucial to ensure that servers are up and running all the time with a robust always-on solution. It is important in SCADA and IA environments that unplanned downtime is prevented from happening in the first place.

In this paper you'll learn about availability, the exact risks of unplanned downtime, the less obvious costs that can be incurred, and the different approaches to availability.

What Exactly is "Availability"?

The term "availability" in the context of application servers is defined as the percentage of time the applications are operational and accessible to users. There are three main approaches to availability:

1. **Backups and restores**: This is achieved by having basic backup, data-replication, and failover procedures in place, which can be deployed with conventional servers. This method delivers approximately 99 percent availability, which equates to an average of 87.6 hours of downtime per year, or more than 90 minutes of unplanned downtime per standard working week.
2. **High availability (HA)**: With a HA solution, applications are further protected from downtime by utilizing duplicate resources. HA solutions include cluster systems (which focus on quick recovery after a failure) and software solutions (which aim to prevent downtime from happening). High availability translates into 99.95 percent—99.99 percent uptime, or 4.38 hours to 52 minutes of downtime per year.
3. **Continuous (or fault-tolerant) availability (CA)**: A CA solution's goal is to eliminate downtime altogether through software or a specialised server system, by preventing it in the first place. CA solutions are the ultimate in availability (99.999+ percent) and result in just 1 to 5.5 minutes of downtime per year.

A solution that delivers 99 percent availability may sound acceptable at first but it is inadequate in many SCADA, process automation & control environments as they mandate continuous operation.

Why Downtime Cannot Be Risked In SCADA, Process Automation & Control Environments

It may seem that it's not crucial to have SCADA, Historian and MES servers up and running all the time, but what's the likelihood of a critical alarm occurring while the SCADA server is down?

What if the production manager needs to generate a production report and a critical piece of information is missing, or he needs to re-reschedule a production plan to fit in an urgent customer request?

What if any of those incidents happen and the system is NOT available? Is it good enough? No, it is not. It is likely that the risks associated with the server being down are probably much greater than we realize. Additionally, what happens when:

- A process HiHi alarm, a fire or gas leak alarm or any process-related HSE hazard, cannot be displayed by the alarm annunciator screen?
- Real-time visibility and access to all the devices and sensors inside the plant are not available?
- The pressure inside a batch tank reaches a critical level and damages the tank and peripheral devices?
- An excessive vibration in a steam turbine in a gas plant, a power plant, or a refinery goes unnoticed for a long time?
- Oil leaks from a pipeline laid in the ocean?

The potential safety, financial and reputation costs are considerable. Missing such important events will often lead to reducing plant availability by limiting production uptime and increasing the plant operation and maintenance cost.

Furthermore, industries that are governed by strict regulation and compliance such as pharmaceutical or food and beverage, could be subjected to high penalties or even loss of license to operate, if some important data is missing—data that's needed to create a product genealogy report, for example.

Moreover, unattended events might lead to health and safety or environmental catastrophes.

Which Availability Option?

Now that the importance of Availability of SCADA, process automation and control applications is understood, let's explore the different approaches to availability and what each one delivers.

High Availability Clusters

High availability (HA) clusters aim to recover from downtime as quickly as possible, as opposed to preventing it from occurring in the first place. HA clusters are a custom-built system composed of two or more servers that run with the same configuration. They are connected with cluster software and shared storage to keep the application data available to both/all servers.

Servers in a HA cluster communicate with each other by continually checking for a heartbeat that confirms if other servers in the cluster are up and running. If a server fails, another server in the cluster (designated as the failover server) will automatically take over; ideally with minimal disruption to users. While HA clusters improve availability, their effectiveness is highly dependent on the skills of specialised IT personnel. Clusters can be complex and time consuming to deploy, and they may require programming, testing, and continuous administrative overhead. Because there are two or more servers, a software license may be needed for each of them which contributes to a high total cost of ownership (TCO). It is also important to note that the risk of downtime is not eliminated with HA clusters. In the event of a server failure, users who are currently connected to that server will lose their connections which can result in loss of in-flight data.

> The potential safety, financial and reputation costs are considerable. Missing such important events will often lead to reducing plant availability by limiting production uptime and increasing the plant operation and maintenance cost.

High Availability Software

High availability (HA) software is designed to prevent downtime, data loss, and business interruption from occurring. HA software is equipped with predictive features that automatically identify, report, and handle faults before they become problems and cause downtime. Two important features of high availability software are that it works with standard x86 servers and doesn't require the skills of highly advanced IT staff to install or maintain. HA software is designed to configure and manage its own operation, making the setup of application environments easier and more economical than clusters. There is a key difference between HA clusters and HA availability software—the software continuously monitors for issues to prevent downtime from occurring, whereas cluster solutions are designed to recover after a failure has already occurred. The most effective HA solutions provide more than 99.99% availability, which translates to less than one hour of unscheduled downtime per year.

Continuous Availability (Always-On) Software

Continuous availability (CA) solutions not only prevent downtime from occurring in the first place, but also provide the cost-saving benefits of an always-on solution based on standard x86 servers. Each application lives on two virtual machines. If one machine fails, the applications continue to run on the other machine with no interruptions or data loss. If a component fails, it is replaced by the healthy component from the second system. Some software can also deliver continuous availability across geographically separated sites as 1 logical system. It prevents data loss, is simple to configure and manage, requires no special IT skills, and delivers upwards of 99.999 percent availability—all on standard x86 servers.

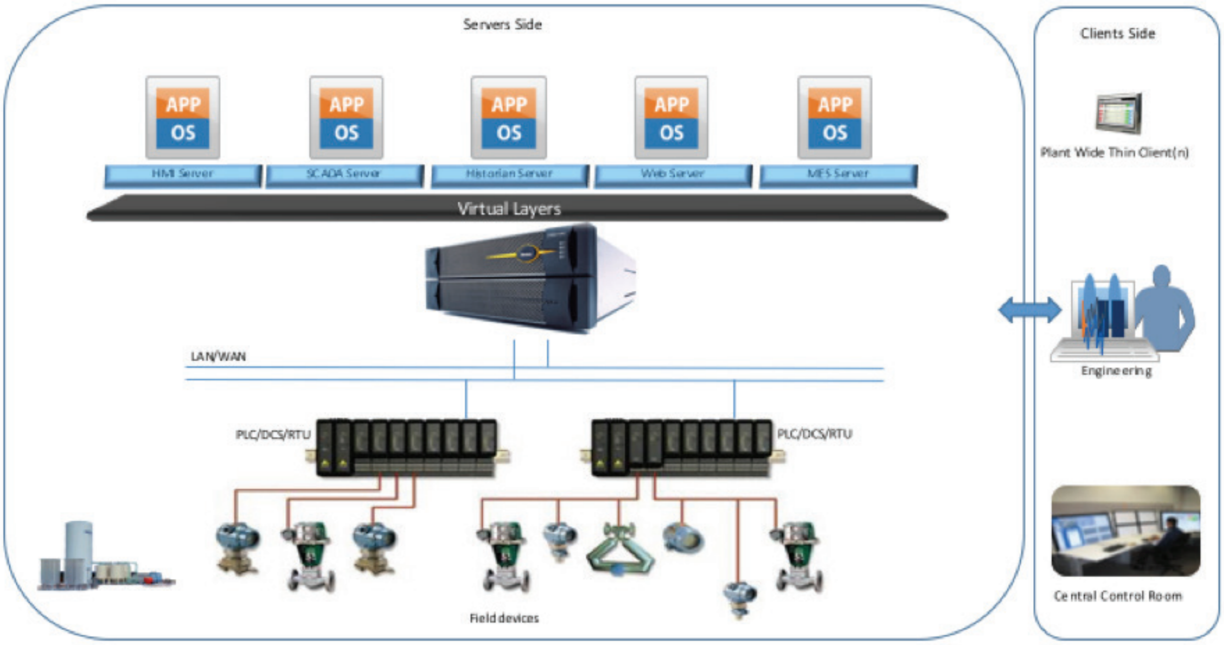
Continuous Availability (Always-On) Servers

A continuous availability (CA) or always-on server system is truly turnkey in that its hardware, software, and services are all integrated for easy deployment and management. This type of solution relies on specialised servers that are purposely built to prevent failures from ever occurring.

These servers are managed just like standard servers, so they don't require specialised IT personnel to manage them. The sophisticated technology simply runs under the operating system and is transparent to the application. No special system administration skills are required.

These solutions include redundancy of components and proactive error detection system software. Automatic fault detection and correction is engineered into the design so that most errors are resolved without the operator even knowing they existed. Additionally, CA servers can run in a virtualized environment, and you can expect greater than 99.999 percent availability if you choose this option.

Architecture for an Industrial Automation Plant



In the End, "Always-On" is the Only Option

There are many applications which organizations rely upon that can tolerate hours of downtime. However, when people's safety, reputation, operational cost, or entire businesses are at stake, trust in your infrastructure's availability is truly needed. Solutions that are designed to prevent downtime from happening in the first place—such as Stratus solutions —are ideal for those business critical applications that simply can't tolerate any downtime. Don't' cut corners and keep your fingers crossed when deciding on the infrastructure to run your critical SCADA or process automation and control applications. In this paper we have concentrated on the technical and availability benefits of deploying an always-on solution but there are also commercial benefits that can be gained when comparing TCO and potential license costs of trying to deploy a HA solution.

For example, an Historian deployment in a HA configuration may require double the expense in the Historian license costs. So if your operation mandates availability consider an always-on solution.

> There are many applications which organizations rely upon that can tolerate hours of downtime. However, when people's safety, reputation, operational cost, or entire businesses are at stake, trust in your infrastructure's availability is truly needed. Solutions that are designed to prevent downtime from happening in the first place—such as Stratus solutions—are ideal for those business critical applications that simply can't tolerate any downtime.

Stratus Always-On Solutions

Stratus offers both platform and software solutions that meet the definition of continuous availability.

Stratus' ftServer® prevents downtime and data loss before it occurs, without a performance hit to your systems or applications. Leveraging the latest Intel® processor technology and supporting Windows®, Linux® and virtualization technology (VMware®, Hyper-V™ and KVM), ftServer's always-on capabilities are achieved with advanced lockstep technology and real-time management and monitoring. This operationally simple solution is quickly and easily installed and is serviced by system generated replacement parts ordering. Famous for its longevity—giving you the miles you need out of your system, ftServer runs and keeps on running, eliminating the need for frequent technical refreshes.

Stratus' everRun® provides continuous availability for business critical applications in an easy to use solution that monitors the entire infrastructure stack from metal to application. The SplitSite capability provides downtime prevention against localized disasters and ensures application availability across geographically separate sites still presented to the applications as 1 logical system. everRun ensures business continuity and maintains data integrity without IT staff physically located where the applications are. It provides cost-effective continuous availability efficiently with two off-the-shelf x86 servers and your customer's choice of everRun managed storage in one easy to manage product.

For more information on the Stratus approach to continuous availability visit **www.stratus.com**.

Stratus™

www.stratus.com

Specifications and descriptions are summary in nature and subject to change without notice.