

## High-Availability and Fault-Tolerant Solutions Minimize Risk of Unscheduled Downtime

By Craig Resnick

### Keywords

High Availability, Fault-Tolerant, Redundant, Downtime, OEE, Production Management, Collaborative Production Systems

### Overview

Today's manufacturers face immense competitive pressures and increasing regulatory requirements. Many companies have shed their excess manufacturing capacity, making it increasingly critical to improve the effectiveness of their remaining manufacturing assets to satisfy customers, drive new growth, and increase profit margins.

To improve performance, manufacturers have implemented an increasing number of mission-critical applications at the production management level. With fewer production facilities, even slight interruptions to these "collaborative production systems" can have severe and immediate consequences.

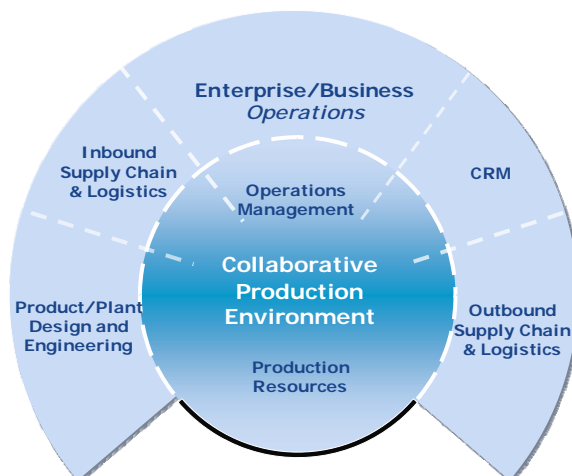
To improve performance, manufacturers have implemented an increasing number of mission-critical applications at the production management level. These include applications for materials management, finite scheduling, quality and performance monitoring, plant historians, and production records, as well as LIMS, asset lifecycle management, facilities management, and even operator HMIs. These comprise what ARC refers to as collaborative production systems.

With fewer production facilities, even slight interruptions to these collaborative production systems can have severe and immediate consequences. The potential impact of any type of adverse event has increased substantially. This drives the need for manufacturers to implement high-availability and fault-tolerant solutions (such as those widely used for their real-time control and safety systems) for collaborative production systems. These solutions minimize the risk of unscheduled downtime and protect manufacturers from potential disasters through effective recovery mechanisms.



## Unplanned Downtime: The Nemesis of End Users

Any downtime is expensive for end users, but unscheduled downtime is the most expensive. It impacts the businesses ability to meet its production schedule and customer commitments. Unplanned downtime, which includes unexpected stoppages resulting from equipment failure, operator error, or nuisance trips, is the nemesis of all end users. Unscheduled downtime is also costly in terms of equipment damage, environmental harm, and worker safety. Dynamic overall equipment effectiveness (OEE), a key performance indicator in many plants, reflect the cost of downtime and helps determine the real-time impact of the performance of any individual process or piece of equipment on the overall efficiency of the plant. Unscheduled downtime is a primary factor that significantly lowers dynamic OEE, which translates to decreased efficiency and profitability for the manufacturer.



**Collaborative Production Systems**

When analyzing the causes of unscheduled downtime, it becomes clear that most incidents occur due to multiple factors, rather than a single factor. Preventable human error is certainly a contributing factor to unscheduled downtime, but hardly the only one. Much unscheduled downtime is also caused by equipment failures, which are also largely preventable. However, this requires a multi-layered, multi-disciplined approach that includes deploying collaborative production systems designed and implemented to deliver very high levels of availability and fault-tolerance. This typically requires effective

data backup mechanisms, redundant processors for critical applications, and industrial-grade software, plus more fault-tolerant server technology to ensure continuous availability of these mission-critical applications.

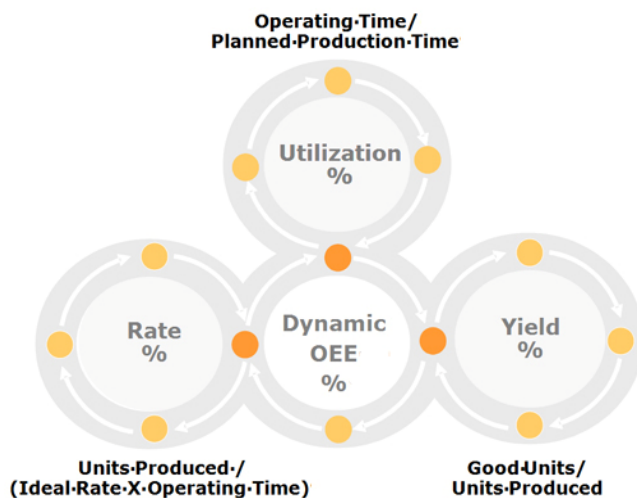
### Separate Solutions for High Availability and Disaster Recovery

Part of a multi-layered, multi-disciplined approach to eliminating unscheduled downtime requires understanding both how the solution will deal with disaster recovery and how it should deal with the high-availability requirements for uninterrupted operation. Separate solutions should be implemented for each. High-availability solutions should be the first line of defense against failures. A geographically separated, fault-tolerant, high-

availability solution provides availability protection for over 90 percent of failures. For the other approximately 10 percent of failures, a disaster recovery solution can be implemented via a wide area network. Long-distance data replication is a legal or regulatory requirement for many companies. However, disaster recovery on its own does not provide a high-availability solution. Many steps must occur for a disaster recovery failover. A disaster recovery solution would treat every failure, even minor ones, as a disaster, requiring the disaster recovery failover process to initiate a recovery.

### Basic Steps End Users Can Take for High-Availability Systems

End users can take some basic steps to provide high-availability solutions for their collaborative production systems. The first step is to maximize their operator's effectiveness in the control room. This is essential to minimize the risks of accidents and maximize production quality, as well as to eliminate unscheduled downtime.



$$\text{Dynamic OEE} = \text{Rate} * \text{Yield} * \text{Utilization}$$

The global process industry loses \$20 billion, or five percent of annual production, due to unscheduled downtime and poor quality. ARC estimates that almost 80 percent of these losses are preventable and 40 percent of those preventable losses are primarily due to operator effectiveness issues in the control room. Maximizing operator effectiveness requires automating as many functions as technology will allow, as well as reducing complexity wherever possible. For example, in many plants, operators still monitor the

processes and collect data manually or semi-automatically using chart recorders. This is both tedious and error prone, and does not provide adequate process insight to instill a sense of ownership among the control room operators.

Another step that end users can take to provide high availability solutions is to create availability objectives, an important first action in formulating a protection strategy. These objectives are typically split up between recovery time objectives and recovery point objectives. Recovery time objectives

involve establishing the time that it would take for an application to be up and running again. Recovery point objectives involve the time it would take to recover the data in case of failure, plus a plan to deal with the potential loss of data. Once both recovery time and recovery point objective baselines are established, the end user can establish and commit to the service level agreements required for the overall company, specific business units, or particular internal groups.

Recovery time objectives involve establishing the time that it would take for an application to be up and running again, while recovery point objectives involve the time that it takes to recover the data in case of failure, plus a plan to deal with the potential loss of data.

For end users to meet their recovery time objectives, they must separate issues by system availability versus recovery. Typically, collaborative production system recovery time should be in seconds or minutes. In this context, recovery involves quickly reestablishing system availability and doing so in an automated, proactive fashion. These systems require protection from storage failures, network failures,

server failures, power failures, and facility outages. However, when widespread outages occur (such as those caused by natural disasters such as hurricanes, tornados, and floods), data retention and recovery is the priority. This process can take hours or days, depending on the damage, as is typically handled in a manual, reactive fashion.

These clear differences between high availability and disaster recovery, based on cause, time differences, automated versus manual, and recovery expectations, are why end users require separate strategies to be effective.

### **Disaster Recovery and Data Protection Add to High Availability**

Another question that end users have to answer is what they want to happen to their application when a disk, network, or server fails, there is a data center power outage or fire, a regional power outage, or a severe natural disaster. A high-availability solution should preempt the need for any sort of long-distance solution. That is why disaster recovery and data protection solutions complement, rather than replace, high availability solutions. Replication products offer failover, but are not designed for high availability. They do not provide a reliable method for detecting failures and, as a result, could induce false failovers. They are asynchronous, so each time a failover is initiated, data will be lost and there is a good possibility of database corruption. While some have an “automated” failover feature, this is not recommended, as their failure detection is weak and failure actions may

not be necessary. This could result in “split brain,” where the application is running on both the primary and backup system at the same time. Replication products do not detect storage or other failures, only network heartbeat failures. They treat all failures, both major and minor, as disasters and require the entire failover process to be invoked, a substantial process. Failback is also a manual process that requires intervention and downtime. For these reasons, disaster recovery complements a true high-availability solution by protecting against the lesser chance of a major disaster.

### Simple Steps End Users Can Take to Reduce Human Error

Human error is a significant source of unscheduled downtime, especially in the area of operator effectiveness in the control room. The good news is that end users can take some very simple steps to reduce human error. Complexity and manual processes increase the chances of human error and downtime. The best way to reduce this is to leverage the power of the collaborative production system to automate as many functions as technology will allow, which will reduce complexity. This will eliminate operator functions such as manual data collection, which often lead to significant human errors. The collaborative production system can simplify and automate processes to the greatest possible extent.

When it comes to protecting the hardware, storage, and networks of their collaborative production systems from unscheduled downtime, one simple



Asset Lifecycle Management Processes

step end users can take is just not to cut corners. End users should select proven, trusted, name brand equipment, as the return on investment will be immediate the second that an unproven piece of equipment fails and requires service or spare parts.

Also, end users must perform recommended maintenance to extend and maximize the lifecycle and financial return on their assets. This is part of what ARC refers to as Asset Lifecycle Management (ALM).

Another key step is to add redundancy to all potential links of peripheral systems that would act as the weak link in the chain to cause failure. This includes power supplies, critical cooling, error-correcting code memory,

etc. Finally, end users should combine device redundancy with redundant array of independent disks storage algorithms to protect data access and data integrity, and add extra disks configured with this protection. This technology allows high levels of storage reliability from PC-class disk-drive components.

Additional steps that end users can take to protect the hardware, storage, and networks of their collaborative production systems from unscheduled downtime include using redundant server connections to eliminate failovers caused by the failure of a single server or network component. In addition, end user's physical network hardware should never share common components.

### **Last Word: High Availability and Fault-Tolerant Solutions Can Help End Users Leverage the Power of Their Systems**

End users should automate the protection of their collaborative production system's servers, storage, and networks, as complexity and manual processes increase the chances of human error and downtime. Point prod-

End users should automate the protection of their collaborative production system's servers, storage, and networks, as complexity and manual processes increase the chances of human error and downtime.

ucts can complicate the sequencing and complexity of preventing or recovering from a failure. That is why it is critical for end users to select solutions that automatically monitor, detect, and reconfigure resources to keep applications running without IT intervention.

Yesterday's redundant collaborative production systems were complex, proprietary, expensive, and based on hardware. Today's redundant systems are flexible, automated, affordable, and based on software. For end users to maximize the value to today's collaborative production systems, they must ensure that redundancy, high availability, and disaster recovery solutions are deployed to fully leverage their power.

*For further information or to provide feedback on this Insight, please contact your account manager or the author at [cresnick@arcweb.com](mailto:cresnick@arcweb.com). ARC Insights are published and copyrighted by ARC Advisory Group. The information is proprietary to ARC and no part of it may be reproduced without prior permission from ARC.*