Designing a Control System for High Availability

Art Pietrzyk, TUV FSExp, Rockwell Automation
Brian Root, Redundancy Marketing Manager, Process Initiative, Rockwell Automation
Paul Gruhn, P.E., CFSE, Training Manager, ICS Triplex

Introduction

When hearing the term “high availability,” many engineers think of redundancy as the only method for achieving higher availability. However, redundancy increases the number of components, which increases the number of potential component failures. Therefore, redundancy, if not applied properly, can actually *decrease* system availability. So, should redundancy remain top-of-mind or should alternate methods be considered?

This paper will discuss redundant and non-redundant methods for achieving high availability of control systems, as well as improvements in control technology and recommended control system designs. The paper will also highlight features within the Rockwell Automation Integrated Architecture™ platform and ICS Triplex product lines that can help achieve higher availability.

Please note that this paper focuses on high availability for *control* systems. This paper does not address safety system concepts or design philosophies.

What is High Availability?

At the most basic level, availability can be defined as the probability that a system is operating successfully when needed. Availability is often expressed as a percent. Expressed mathematically, availability is one minus the unavailability.

Availability (A) is calculated using the formula A = MTBF / (MTBF + MDT), where MTBF is Mean Time Between Failure and MDT is Mean Down Time. MDT is often assumed to be the same as MTTR, the Mean Time to Repair. MTTF, Mean Time To Failure, is often considered interchangeable with MTBF, although there are subtle differences. Another common term in the field of reliability engineering is failure rate (λ) which is expressed as 1/MTBF.

The term high availability has been used to encompass all things related to productivity, including reliability and maintainability. So let’s take a closer look at these terms as well.

Reliability

Reliability can be defined as the likelihood that a device will perform its intended function during a specific period of time (often called the mission time). It is a measure of system success over a time interval.

To help make sure that products meet customer expectations, reliability can be designed using techniques such as Component Derating and Design through Six Sigma. Diagnostics are necessary in order to detect faults and alert personnel when faulty hardware needs to be replaced. This helps achieve high availability. The level of diagnostics for the Allen-Bradley® ControlLogix® programmable automation controller (PAC) from Rockwell Automation exceeds 90 percent, meaning most failures can be detected and appropriate actions taken. ICS Triplex Trusted™ and AADvance systems have diagnostic coverage in the range of 99 percent.

But even the most robust and reliable system may not be the most available. To be available, a system must also be easy to troubleshoot, modify and repair during the mission time, which may exceed a decade or more.

The Impact of Maintainability on Availability

Maintainability is the ability of a system to be changed or repaired.

Factors that affect maintainability include:

- System and component-level diagnostics for detecting and isolating failures
- Annunciation of faults
- Tools for troubleshooting
- Trained personnel
- Accessibility
- Time to replace or repair
- Ability to add components or make changes

Maintainability of a system significantly impacts the end user's perception of availability. For example, today's automobiles have diagnostic features that can improve availability (e.g. tires that heal and/or monitor pressure, electronic ignitions, diagnostic message displays that indicate a malfunctioning system and the ability to call for roadside assistance).

Some of the features within Rockwell Automation products that help improve maintainability include things such as system-level diagnostics, wire-off detection, auto-tuning I/O and deterministic communications. To assist with troubleshooting, onboard LED indicators, network monitoring tools, graphical programming languages and HMI displays help quickly identify and remedy problems.

Key to keeping a system maintained is to make sure there are qualified and trained personnel. Less obvious ─ but just as important ─ are physical characteristics that affect maintainability. Modules or components should be capable of being removed, replaced or added to the system without interrupting the mission. Replacements should not need rewiring or reprogramming.

Features like online edits, partial downloads, adding I/O online, and removing and inserting modules under power help make the ControlLogix PAC, and the ICS Triplex Trusted and AADvance systems more available. The ability to add tags to the HMI online also helps improve the availability of the HMI and information layers.

Control technology features that have improved maintainability, include:

- Adding or removing modules under power
- Adding I/O online
- Online edits and partial downloads
- Soft switching of processor's producer/consumer communication
- Internal diagnostics to detect failures
- Diagnostics of field circuit problems: open circuit, short circuit, etc.
- Configurable fault response: hold last state or turn off
- HART and other fieldbus technology with sensor and actuator diagnostics
- Self-learning or inherent machine diagnostics
- Adding sensors, I/O and tags online without interruption

In today's world of increasingly complex semiconductor and software-based devices, it can be more challenging to predict failures. It is not a question of *if* a failure is possible, but how often it can be detected and whether the mission can be completed.

HART and other fieldbus devices communicate with more intelligent sensors, instruments and actuators, and provide their own level of device diagnostics. This diagnostic functionality, along with additional process data that these devices provide, are married to software that benefits users with upfront alarms, calibration and model information for easier replacement and inventory management, making sure the parts are replaced quickly and correctly.

Other innovations, such as state-based control and self-learning diagnostic routines, have raised the ability of the controller to detect, annunciate and describe problems within the machinery. For many users, the ability to maintain and revise the system without shutting down offers an acceptable level of availability, especially if the change or repair can be made in minutes.

Achieving High Availability Through Redundancy and Fault Tolerance

Customers or critical applications that cannot tolerate impact to the mission may find redundancy or fault tolerance necessary.

Redundant components needed for high availability include:

- Uninterrupted power supply (UPS)
- Redundant power supplies
- Redundant components
    - Chassis
    - Processors
    - I/O modules
    - Sensors and actuators
    - PCs/HMI
    - Networks
    - Media
    - Servers
    - Databases

Some form of redundancy or fault tolerance is generally used if a control system shutdown or loss of visibility causes a major loss of revenue, loss of equipment, injury to people or a disruption to public services. ***Redundancy*** in these situations means the duplication or triplication of equipment that is needed to operate without disruption, if and when the primary equipment fails during the mission. ***Fault tolerance*** is the ability of the system to tolerate faults and continue operating properly. There are subtle differences between redundancy and fault tolerance.

In terms of electrical equipment, the most important place to start for guaranteeing reliable operation is by providing continuous power. Most power comes from the electrical power grid to a plant, where at one point in the electrical delivery system a single transformer supplies an area within a plant. A lightning strike on the power grid will certainly have an impact. Bad power is sometimes filtered by a plant's electrical infrastructure, as bad power can cause unexpected behavior to running microprocessor-based equipment. Thus, the control system can only be assumed as reliable as the power provided to it. The good news is that there are many suppliers of quality uninterruptible power supplies (UPSs) that can help provide a constant power of acceptable quality.

The key is to attach the output power of the UPS to the primary controller, which filters surges and minimizes recovery of the system when power is re-established by the plant. Rockwell Automation systems are available with redundant power supplies and will work with DC and AC voltage options to accommodate plant, battery and UPS supplied power. This is critical when powering the primary power supply with a UPS and isolating the second redundant power supply with standard plant line power. Redundant power supplies can be installed on both the redundant controller pairs and in a remote ControlLogix I/O chassis.

In addition to the power considerations, redundant components may be required to provide or maintain control in the event of failure. The ControlLogix system is an example of a platform that provides this assurance. This system provides redundant power supplies, controllers and network modules (both ControlNet and EtherNet/IP versions) that reside in a separate chassis. This separation and duplication provide for a more complete and maintainable controller redundancy.

AADvance systems can be fault tolerant and utilize dual or triple redundancy of processors and I/O modules. Trusted systems utilize triple modular redundancy (TMR) for the highest level of fault tolerance.

All Rockwell Automation systems, employing redundancy or fault tolerance, only require a single configuration. The only additional setup needed with the ControlLogix system, other than the hardware, is to check the redundant controller parameter box in the Rockwell Software® RSLogix™ 5000 programming software. All tag data and I/O can be updated between the primary and secondary controller at a user-defined rate.

In the event of a failure of one of the modules in the primary's chassis, the system switches control and HMI function with a bumpless transfer on the control to the secondary chassis. When the faulty module is replaced in the disqualified secondary chassis, the system re-synchronizes automatically without any extra operator action. Redundant controllers can be migrated to the next release without shutting down control, which answers one of the most significant concerns for an end user.

HMI availability is mission-critical due to the fact that most customers require some form of HMI to provide visibility into the process or equipment 100 percent of the time. Typically, the network of choice for the information layer to an HMI, data historian or MES system is Ethernet, but if redundant media is specified, ControlNet is the best choice.

Redundant Field Devices

When it comes to determining if redundant I/O modules are required for achieving higher availability, sensors and end devices must be addressed as well. Most system designers know that the reliability and diagnostics from sensors and actuators is a magnitude less than that of the logic solver. They often implement redundancy on the input side by monitoring the same process variable with two separate sensors that are wired to two separate I/O modules. A comparison is made in the logic to determine if there is a mismatch. Maintenance can perform a check to isolate the anomaly or automatic testing schemes can be implemented to isolate a problem to the sensor, wire or module.

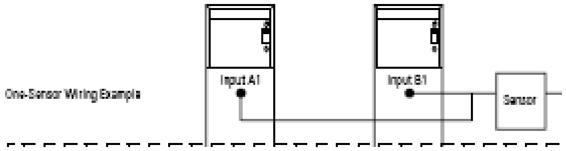
Redundancy With the ControlLogix System

If redundant I/O modules are desired to achieve availability and maintainability, then several methods utilizing terminations are possible. Each method and associated terminations offer various degrees of diagnostics, ease of use and costs. The levels discussed in this paper are based on the following:

1) Allen-Bradley Bulletin 1492™ terminal blocks and termination boards

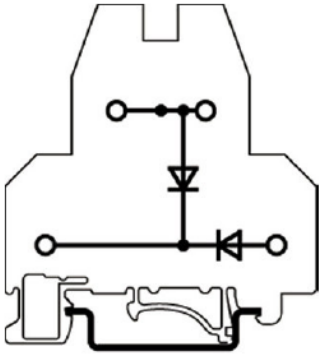
In Figure 1, simple termination blocks are used to simplify the wiring from a single sensor to two input modules. Two separate modules should be used so that if one fails, the other can still provide input data while the failed module is replaced. For both discrete and analog inputs, no diode isolation is required. The analog signal from a sensor must be converted to a voltage so it can be read by two input modules in parallel.

Figure 1.



Terminations with blocking diodes can be used for discrete and analog outputs. The diodes isolate the final output drive devices so that a failure to OFF or GND does not bring the other output low. Note: The polarity of the diodes will be based on whether the modules are sinking or sourcing. Most outputs are sourcing so the P side of the diode should normally go to the output (see Figure 2).

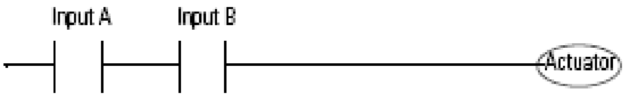
Figure 2.



There are some obvious limitations to this simple form of redundancy. First of all, outputs which fail ON are not isolated and will remain on unless an alternative method for removing power is designed into the system. The failure modes covered are those outputs which turn off and the second output can continue to provide power.

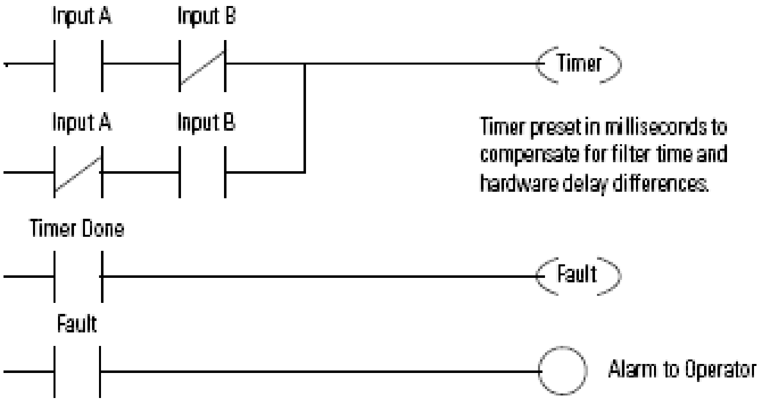
There are no diagnostics to detect if an output has failed unless a provision, such as monitoring outputs with inputs, is included. Inputs can be compared for agreement in ladder logic (see Figure 3). If miscompare is detected, the user must troubleshoot the input states and voltages at the termination panel and troubleshoot to isolate the failure.

Figure 3. Application logic can compare input values or states for concurrence.



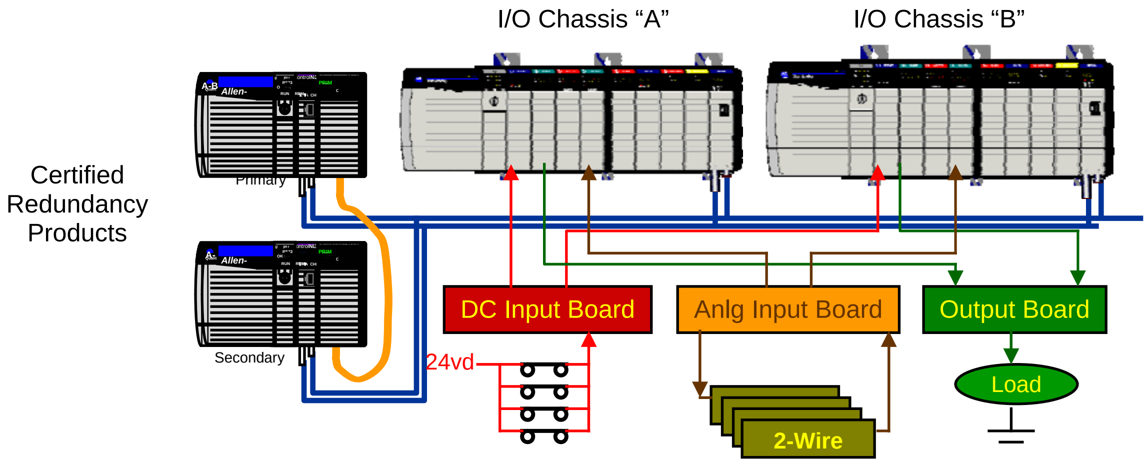
The user program must also contain rungs to annunciate a fault in the event of a sustained miscompare between two points (see Figure 4).

Figure 4.


Input A
Input B
Timer
Input A
Input B
Timer preset in milliseconds to compensate for filter time and hardware delay differences.
Timer Done
Fault
Fault
Alarm to Operator

The next example of a form of I/O redundancy is through the use of complex terminations, which have on-board circuitry and are able to provide more diagnostics and other functionality. Some of the features available include: over-voltage protection, fuses, and LEDs for troubleshooting. Figure 5 shows a restricted form of redundancy for achieving fault tolerance of I/O for a SIL-rated ControlLogix system.

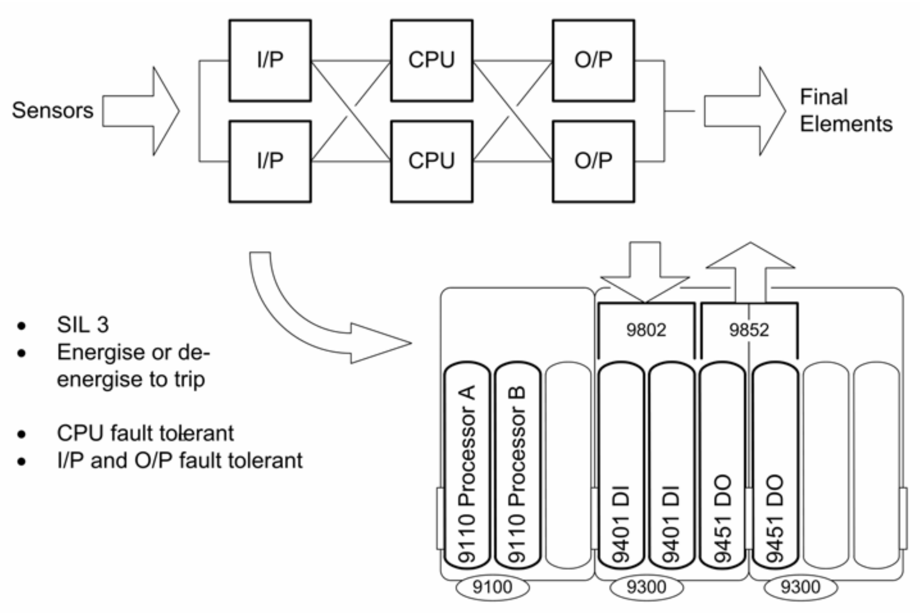
Figure 5. Fault Tolerant ControlLogix System



Fault Tolerance With AADvance

The AADvance system, supplied by ICS Triplex, a Rockwell Automation company, can have individual portions configured simplex, dual or triplicated. Different levels of fault tolerance can be provided depending upon the users requirements. A 1oo2D arrangement (1 out of 2 with diagnostics) is fault tolerant and can survive single module failures. In a dual configuration, if a single module were to fail, the configuration degrades to simplex. If the last module were to fail, the system would shut down. This is referred to as a 2-1-0 degradation mode.

Figure 6. Redundant (1oo2D) AADvance System



AADvance can also have portions configured in a triple modular redundant (TMR) architecture for greater fault tolerance. In a triplicated configuration, when a single module fails, the configuration degrades to dual. If another module were to fail before the first failure is repaired, the configuration degrades to single. If the last module were to fail, the system would shut down. This is referred to as a 3-2-1-0 degradation mode.

Fault Tolerance With Trusted System

The ICS Triplex Trusted system utilizes a triple modular redundant (TMR) architecture. Triplication eliminates the possibility of any single component failure causing a spurious or false trip. This achieves the highest level of availability. High levels of internal diagnostics, as well as errors detected through discrepancies, allow the system to continue running in the presence of faults and annunciate faults for operator action. Modules can be easily replaced online without affecting the process.

A difference in the triplication between AADvance and Trusted systems is that the Trusted majority votes data in hardware. Therefore, it is not possible to run on a single “slice”. This is referred to as a 3-2-0 degradation mode. If spare modules are installed, the degradation mode becomes 3-3-2-0.

The Trusted TMR system provides:

- The highest level of internal diagnostics
- Tolerance to multiple failures
- No time repair restrictions
- Reduced operating system size and complexity

Achieving High Network and HMI Availability Through Design

Many times, customers can achieve an acceptable level of availability through design, which includes the controller, HMI and information system. The designer must be willing to accept the fact that anything can fail and design the facility around this notion.

The equipment or plant can be designed to continue running if a machine were to fail. This has often been referred to modular distributed design, and involves the following areas:

- Distributed control architectures
- Distributed control design with independent line, zones, etc.
- Distributed HMI
- Distributed databases

In continuous and batch processing operations, following the S88 model helps achieve availability by allowing recipes and procedures to be ported to various equipment, lines and plants. Human interface into a process or operation is not only crucial, but in many cases it is an absolute requirement that has been achieved over the years using hardwired indicator lights and manual controls.

Today, these are being replaced with the more cost effective CRT-based HMI. Each technology has its positive and negatives. The negative aspects of electronic, and especially PC-based HMI, is that it is a single and complex device. Microsoft-based displays have the additional challenges of being open and subject to security, virus and other related issues.

Since most process customers cannot see the machinery and product, which is spread out over large areas and in closed pipes and tanks, visibility is crucial. Most customers use one of several different methods to achieve visibility and most, if not all, include some form of redundancy. One typical method is to design using diversity. This might mean diverse HMI devices, such as an Allen-Bradley PanelView™ terminal with FactoryTalk® View HMI software or a CRT-based HMI, are used in conjunction with local indicator lights. Another method is the use of client server configurations, with multiple clients linked sometimes with redundant servers.

If data is critical, than provisions should secure it from data loss such as, if a single server would go down. One of the most economical methods is to keep some of the more recent data stored in the controllers. If a network is determined to be a weak link, then redundant or fault tolerant communications should be utilized. This could include Ethernet rings, with or without redundant media. Networks can be configured with redundant paths using switches and/or routers.

Security is also important, as security breeches will impact availability.

Conclusion

This paper explored the redundant and non-redundant methods for achieving high availability, as well as improvements in control technology and recommended control system designs. Although redundancy is the traditional method for achieving high availability, the means to achieving high availability require more than just thinking about redundant components. A system with no redundant components can still be very available.