Recommended Design Criteria for Data Center Hardware for ~~Back Office~~Train ~~Control Systems Hardware~~Systems ~~Back Office Systems~~
~~New~~Revised ~~20~~19~~20~~23 (6 Pages)

A. Purpose

The purpose of this Manual Part is to provide general recommendations to allow business decision makers to determine how best to meet the needs for back office systems* hardware and supporting infrastructure based on system scalability, availability and reliability, and service life requirements.

B. Scalability

1. The scalability requirements of the back office systems hardware relate to the overall objectives of the railroad. If the objective is to maintain the current rail system with little or no chance of expansion, then server sizing should consider the maximum system load and allow spare capacity for application updates. If the business plan is more dynamic, with planned expansion through addition of territory or tenants, or additional functionality, a more open ended server model should be considered that permits growth in system capacity.

   Items that should be considered for scalability are:
   - Loading on the electrical system and backup power
   - Physical space
   - Data storage
   - Networking
   - Processing power
   - Cooling

C. System Availability and Reliability

1. System Availability is the ability of a product to be in a state to perform the required function under given conditions over a given time interval assuming the required external resources are provided.

2. System Reliability is the ability of an item to perform a required function under given conditions for a given time interval.

*The term "Back Office Systems" as used herein refers to any server application used for the purpose of Communication, Signaling or Network Operations to support the aggregation and dissemination of data.

Formatted: List Paragraph, Left, Indent: Left: 0", Border: Bottom: (Single solid line, Auto, 1.5 pt Line width)

Formatted: Font: Arial, 12 pt

Formatted: Indent: Left: 0"

3. System Availability and Reliability are the aggregate of the individual sub-systems' availability and reliability. It is based on the business/service requirements of the railroad, which may also include external requirements to fulfill interoperability.

4. An analysis should be performed to determine the overall System Availability and Reliability requirements, its interfaces and the facility infrastructure. This assessment will determine the criteria the hardware must meet. This analysis evaluates three key characteristics: the operational requirement; the availability requirement; and the impact of system downtime.

D. Operational Requirements Determination

1. The first step in identifying the operational requirements is to quantify the windows of opportunity that are available to perform system testing and maintenance, based on the railroad's operations. Table 2451-1 provides a suggested relationship between scheduled shutdown hours and the railroad's back office system operations.

Table 2451-1: Time Available for Scheduled Shutdown

ANNUAL SCHEDULED SHUTDOWN HOURS	BACK OFFICE SYSTEM OPERATIONS
>400	System is operational less than 24 hours a day and less than 7 days a week. Scheduled maintenance downtime is available during working hours and off hours.
100-400	System is operational less than 24 hours a day ~~and less than~~, up to 7 days a week. Scheduled maintenance downtime is available during working hours and off hours.
50-99	System is operational up to 24 hours a day, up to 7 days a week, and up to 50 weeks per year. Scheduled maintenance downtime is available during working hours and off hours.
0-49	System is operational 24 hours a day, 7 days a week for 50 weeks or more. No scheduled maintenance downtime is available during working hours.
0	System is operational 24 hours a day and 7 days a week. No scheduled maintenance downtime is available.

2. The second step is to determine the operational availability requirements, or the total uptime, that the system must support without disruption. Availability is sometimes expressed as a percentage, or "nines of availability". Table 2451-2 gives the tolerance categories for unscheduled downtime.

Formatted: List Paragraph, Left, Indent: Left: 0"

Formatted: Indent: Left: 0.5"

Table 2451-2: Tolerance Categories for Unscheduled Downtime

UPTIME PERCENTAGE
<99.0
99 to 99.9
99.9 to 99.99
99.99 to 99.999
99.999 to 99.9999

The third step is to determine the Operational Level, which is a function of scheduled shutdown hours relative to the percentage uptime. Table 2451-3 provides an Operational Level scale of 0 to 4, depending on these two criteria.

Table 2451-3: Identify Operational Level

SCHEDULED SHUTDOWN HOURS *(FROM TABLE 2451-1)*	UPTIME % *(From Table 2451-2)*				
	<99.0%	99 To 99.9	99.9 To 99.99	99.99 To 99.999	99.999 To 99.9999
>400	0	1	2	3	3
100-400	1	2	3	3	4
50-99	2	3	3	4	4
0-49	3	3	4	4	4
0	3	4	4	4	4

3. The fourth step is to identify the impact or consequences of system downtime. Not all equipment downtime will result in the same impact to mission critical functions and identifying the impact aids in determining the mitigation of downtime risk. Table 2451-4 identifies five impact classifications.

Table 2451-4: Downtime Impact Classifications

CLASSIFICATION	IMPACT OF DOWNTIME
System	Affecting the entire system, or resulting in a significant disruption or delay to operations.
Multiregional	Affecting a major portion of the system territory, but not in its entirety, or resulting in major disruption or delay to operations.
Regional	Affecting a portion of the system territory, but not in its entirety, or resulting in a moderate disruption or delay to operations.
Local	Affecting only a single site, or resulting in minor disruption or delay to operations.
Minor	Affecting only a single function or operation, resulting in a minor disruption or delay to a non-critical aspect of operations.

The final step is to combine the three key characteristics to arrive at a usable expression of system availability which will be used as a guide to determine the features required to support system operation. Table 2451-5 provides a matrix mapping the operational level against the impact of downtime.

Table 2451-5: Determining the System Availability Category

IMPACT OF DOWNTIME *(From Table 2451-4)*	OPERATIONAL LEVEL *(From Table 2451-2)*				
	0	1	2	3	4
System	SA-1	SA-2	SA-3	SA-4	SA-4
Multiregional	SA-1	SA-2	SA-3	SA-3	SA-4
Regional	SA-1	SA-2	SA-2	SA-3	SA-3
Local	SA-0	SA-1	SA-2	SA-3	SA-3
Minor	SA-0	SA-0	SA-1	SA-2	SA-2

E. Server Hardware Design Considerations

1. The choice of the server hardware is influenced by the availability of maintenance windows. The fewer maintenance windows are available, the more system redundancy is required to permit shutdown of servers when performing upgrades and periodic maintenance.

2. For railroads which cannot afford scheduled maintenance windows consideration should be given for use of blade servers and chassis. These provide hot swap capabilities, where additional server processor units, power supplies, storage units and cooling fans can be installed while the system remains fully operational.

3. Table 2451-6 provides a list of server design considerations for the range of system availability criteria.

Table 2451-6: Server Design Considerations

SERVER DESIGN CONSIDERATION	SYSTEM AVAILABILITY CRITERIA *(From Table 2451-5)*				
	SA-0	SA-1	SA-2	SA-3	SA-4
Server Redundancy	Cold Standby	Warm Standby	Warm Standby	Hot Standby	Hot Standby
Redundant Power Supplies	Optional	Optional	Recommended	Recommended	Recommended
Physical Network Interfaces	1	1	2	2	2
Data Storage Integrity	Raid-1	Raid-2	Raid-4	Raid-5	Raid-5
Rack Environmental Monitoring	Optional	Optional	Recommended	Recommended	Recommended
Hot Swappable Modules	Optional	Optional	Optional	Recommended	Recommended

F. Supporting Infrastructure Considerations

1. ANSI/TIA-942 is a Telecommunications Industry Association standard for Data Centers. This standard covers all aspects of the physical data center (back office) including site location, architecture, security, safety, fire suppression, electrical, mechanical and telecommunications. It is recommended as a guideline for developing the desired level of system availability for the locations that house back office systems.

2. The System Availability of a back office is limited by supporting infrastructure for the system hardware.

3. ANSI/TIA-942 describes four Rating/Tier levels in which data centers can be classified. Below is the high level description of each Rating/Tier level. Detailed specifications are given in the ANSI/TIA-942 standard.

Rated-1/Tier-1: Basic Site Infrastructure

A data center which has single capacity components and a single, non-redundant distribution path serving the computer equipment. It has limited protection against physical events.

Rated-2/Tier-2: Redundant Capacity Component Site Infrastructure
A data center which has redundant capacity components and a single, non-redundant distribution path serving the computer equipment. It has improved protection against physical events.

Rated-3/Tier-3: Concurrently Maintainable Site Infrastructure
A data center which has redundant capacity components and multiple independent distribution paths serving the computer equipment. Typically, only one distribution path serves the computer equipment at any time. The site is concurrently maintainable which means that each and every capacity component including elements which are part of the distribution path, can be removed/replaced/serviced on a planned basis without disrupting the services to the End-User. It has protection against most physical events.

Rated-4/Tier-4: Fault Tolerant Site Infrastructure
A data center which has redundant capacity components and multiple independent distribution paths serving the computer equipment which all are active. The data center allows concurrent maintainability and one fault anywhere in the installation without causing downtime. It has protection against almost all physical events.