SSCL--335

DE91 006352

ATTACHMENT A

PHYSICS AND DETECTOR SIMULATION FACILITY

TYPE 0 WORKSTATION

SPECIFICATIONS

Computer Acquisition Working Group*

Superconducting Super Collider Laboratory†
2550 Beckleymeade Ave.
Dallas, Texas 75237

November 1990

---

*G. Chartrand, L. R. Cormell, R. Hahn, D. Jacobson, H. Johnstad, P. Leibold, M. Marquez, B. Ramsey, L. Roberts, B. Scipioni, G. P. Yost

†Operated by the Universities Research Association, Inc. for the U.S. Department of Energy under Contract No. DE-AC02-89ER40486.

Table of Contents

1. Introduction

1.1 Overview

This document specifies the requirements for the front-end network of workstations of a distributed computing facility. This facility will be needed to perform the physics and detector simulations for the design of Superconducting Super Collider (SSC) detectors, and other computations in support of physics and detector needs. A detailed description of the computer simulation facility is given in the overall system specification document[1]. This document provides revised subsystem specifications for the network of monitor-less Type0 workstations. The requirements specified in this document supersede the requirements given in Ref. 1.

In Section 2. a brief functional description of the facility and its use are provided. The list of detailed specifications (vendor requirements) is given in Section 3. and the qualifying requirements (benchmarks) are described in Section 4.

1.2 Philosophy

The SSCL has adopted a policy that encourages and will foster the use of open environments for computing. The SSCL will require the use of open systems and industry standards where possible for operating systems, languages, utilities, and protocols. To meet this goal the Laboratory has decided to establish a computing environment which emphasizes the use of distributed, networked computing, graphics, and peripherals from multiple vendor sources. The SSCL is a member of the Open Software Foundation (OSF) and is tracking the developments of Unix International (UI). The SSCL supports the efforts of OSF and UI and will require the use of open system technology as it becomes available. In particular, it is intended that either or both of the OSF/1, System VR4 operating systems will be installed as soon as possible. The SSCL will serve as the systems integrator for the facility to ensure that the networked components are interoperably compatible.

2. Functional Requirements

2.1 Functional Model

Functionally the facility requirements can be broken into three major subsystems: 1) a networked front-end for interactive usage, 2) a file server, and 3) a ranch of parallel batch processing compute servers. A diagram of a model which has been developed to meet the functional requirements is shown schematically in Fig. 2.1. Each of the distributed subsystems is networked by one or more high-speed network links to the other subsystems. Details of this model for the facility are provided in Ref. 1. In this section, the requirements for the front-end Type 0 workstations are described.

2.2 Front-End Network

The front-end network shall provide the user the interactive computing that he needs to gain access to the facility, to retrieve and edit files, to compile and run small jobs, and to submit batch jobs. As shown schematically in Fig. 2.1, the front-end shall be comprised of a network of workstations. The front-end will be comprised of three types of workstations: some of these workstations will not have monitors (Type 0), some fraction will be mid-range graphics workstations, and a small number will be high-end high-resolution graphics workstations. Each workstation, ideally one per user, shall provide a host unit for the users logged into the facility. The front-end workstation shall provide the system files for the user. The workstations shall permit usage of the system utilities described in paragraphs 3.2 and 3.3, shall permit usage of the batch system, and shall permit usage of disk and tape storage on the file server.

2.2.1 External Users. Initially the SSC computer simulation facility will be used primarily by outside users who will log on remotely through a wide area network (WAN). Good access is essential if this is to be possible. Remote users will gain access to the facility via the ESNET WAN. The facility shall provide DECnet and TCP/IP interfaces over Ethernet to the WAN.

2.2.2 Local Users. Local users are those Laboratory employees and visitors who use the facility while in residence at the SSCL. The local users' needs will not differ significantly from the external users with just a few exceptions. The local users will gain access to the facility by the local area network (LAN) which is likely to permit them a much larger bandwidth for the transfer of data from the facility to the individual user's workstation or display terminal. Consequently these users will make larger demands on the data transfer and graphics capabilities of the facility. Fully interactive detector visualization will require bandwidths that are currently obtainable only through dedicated processors in a graphics workstation located at the user's desk.

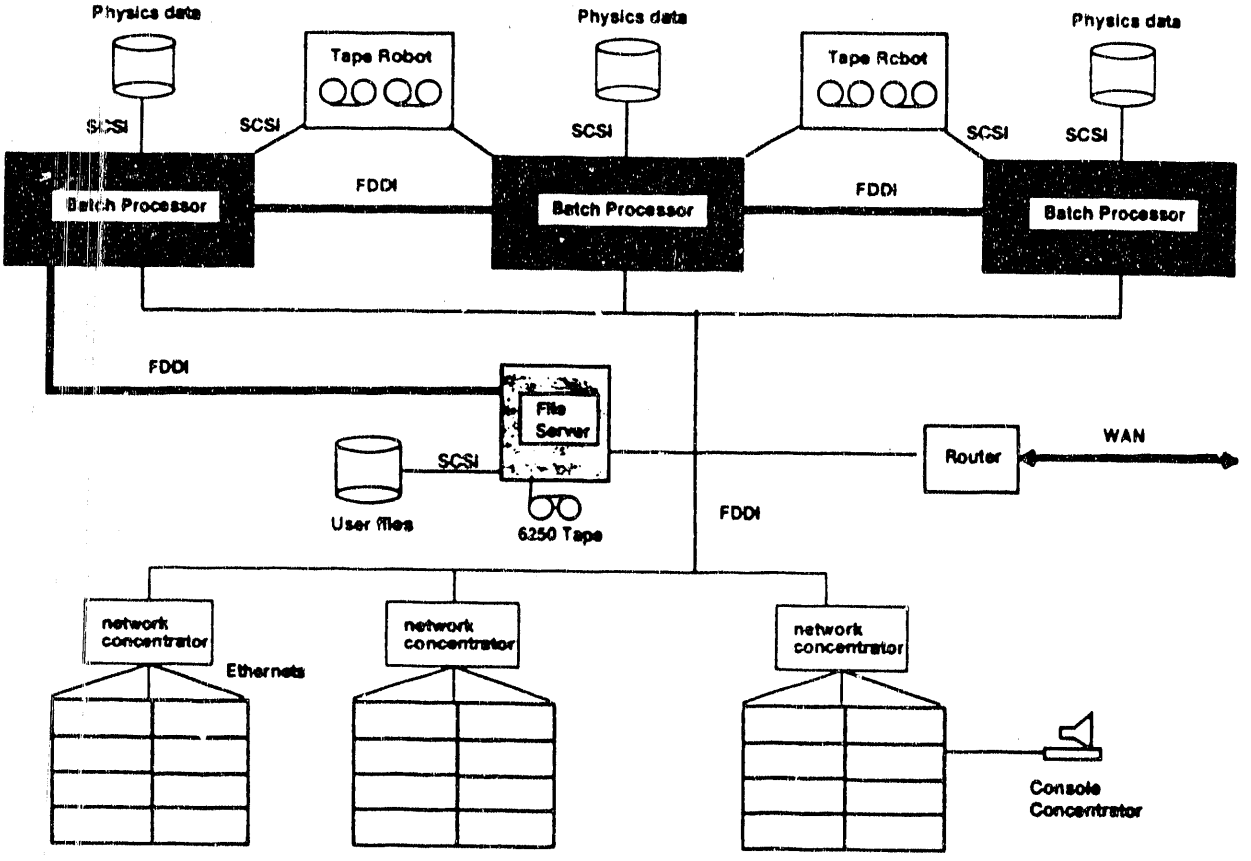


Figure 2.1. Functional diagram (block diagram) of the simulation facility.

2.3 Operational Model

This paragraph provides a description of the operational concept of the facility. Interactive users will log on to the system either remotely (over T-1, etc. and through a router) or locally (via Ethernet). A process resident on a file server will assign an idle "front- end" workstation to the user. Ideally there will be one workstation for each user, thus providing him with a completely dedicated resource. Initially there will be approximately 30 front-end machines subnetted into two groups of fifteen and connected via a bridge. Most of the front-end units will not have monitors and will be placed on rack-mounted shelves. Each subgroup will have a separate interface to the file server. The system is intended to be scalable by adding additional file server/front end/Ethernet groups separated by bridges but accessing the same batch and archival services as the other front end groups.

The eventual goal is that the front end/file server systems be able to access both disk and tape resources containing batch job output independently of the batch processors. Access to disk by dual/multi-ported drives and tape by multi-headed robot-based systems will accomplish this, and is therefore, reflected in the system design. However, for the first phase of development, 8-mm tape carousels will be used by the batch processors for tape storage. This is intended only as a temporary solution to the tertiary storage problem.

The theme of multi-vendor, networked, RISC-based open systems in this environment drives some systems integration requirements as well. In particular, the sharing of binary data by asynchronous processes across multi-vendor platforms argues for data compatibility. To this end all the systems shall support the IEEE floating point standard and interface with the rest of the hardware, network, and software systems with big Endian byte ordering. The precise meaning of this is elucidated in section 3.2.2.

In this section the pertinent hardware, operating system, software, graphics, network, support, and operational requirements are specified for the front-end network, Type 0 workstations. This subsystem forms one portion of the facility as described briefly in Section 2. The specifications for the workstations are stated in the form of minimum requirements. Superior performance at suitable price/performance values is clearly desirable. The minimum required performance of the sub-system CPUs is stated in Section 4 in the form of industry standard benchmarks and physics benchmarks.

3.1 Hardware Requirements

The front-end, Type 0 workstations shall have a console switching device which will allow a single monitor to act as console for any machine within that subsystem. Every machine shall have at least one serial port to receive console messages. The operating system shall not require the console device to be available for output. Every machine shall have a mechanism to enable the OS to be loaded. There shall also be a remote OS installation capability over the network.

3.1.1 Front End Network. The strategy for front end processors centers around partitioned interactive utilization for tasks such as data analysis, test compilations and test runs. The goal is to provide one computer per interactive user. While several types of computers will be purchased for the front-end, only the Type 0 (monitorless) computers are described by this specification.

Type 0 computers shall fit on a 19-inch shelf. The computers shall be self-contained units with the possible exception of the disks which may be external. A monitor, keyboard, and mouse will not be provided, however, these computers shall have the capability to use those facilities.

3.1.1.1 CPU. The minimum CPU performance characteristics of the front-end machines are specified by the benchmark requirements of Section 4.

3.1.1.2 Memory. At least 16 MB of RAM is required in a configuration which is upgradable to 32 MB RAM.

3.1.1.3 Disk subsystem. Secondary storage shall consist of two types, maintained and scratch.. Maintained user files shall be on a server. However, due to the large number of interactive users and the high volume of data being analyzed, network traffic issues dictate a "semi-dataless" configuration, i.e., one with sufficient local disk for OS and swap space plus a large local scratch cache supporting data analysis. Minimum disk capacity shall be at least 600 formatted Mbytes.

3.1.1.4 Network interfaces. Ethernet (IEEE 802.3) network interfaces shall be available for these platforms. The network specifications are provided in Section 3.5.

3.1.1.5 Console ports. The console ports for all front end network machines shall be RS-232. All system console messages and input requests including boot up sequences shall be directed to the port. The front end network system shall function as if a VT100 ASCII terminal using XON/XOFF flow control were connected to the port.

3.1.1.6 Console concentrator. A central workstation using an async multiplexer adapter with at least 32 ports shall be provided. This workstation will function as a console concentrator. The system operator will interact with a particular network front end system by keying into and reading the display of the subwindow corresponding to that front end system. The front end system shall be capable of operating in the absence of a connection to either a VT100 or the console concentrator.

3.2 Operating System Requirements

3.2.1 General Requirements. This section describes the general operating system requirements of the front end network. Additional requirements for each category will be listed in the corresponding subsections.

3.2.1.1 O/S types. Each platform shall execute as its native operating system at least one of the following versions of the UNIX operating system:

Berkeley UNIX - BSD4.3 or a version derived from BSD4.3
AT&T System V - SYS V.3 or a version derived from SYS V.3, or SYS V.4

The name UNIX is used as a generic term for either of these operating systems in the remainder of this document.

3.2.1.2 Required Extensions. Several extensions to UNIX are commonly made by vendors. They are usually based on merging the features of BSD4.3 and SYS V.3.

- Berkeley Sockets. Berkeley style sockets shall be provided as the interface to the communication protocols. The internet domain and the UNIX domain shall be supported. Stream, datagram, and raw (for superuser) sockets shall be supported.
- Asynchronous I/O. Asynchronous I/O will be supported. Processes shall have the ability to initiate read or write operations and not have their execution blocked waiting for completion. A synchronization facility equivalent to BSD4.3 select (2) shall be provided.
- Shared Memory. A shared memory function shall be provided. A user process shall be able to create/attach to/delete a section of shared virtual memory. The system shall support several independent memory contexts simultaneously.

3.2.1.3 Reconfiguration. The UNIX system shall be capable of being reconfigured, that is, the various parameters that affect performance, buffers, memory utilization and devices shall be changeable either on the fly or by a combination of recompiling and/or relinking the kernel.

3.2.1.4 Device Drivers. The UNIX system shall support either BSD 4.3 or SYS V.3 style device drivers. A sample character device driver and block device driver shall be available. A sample networking device driver shall also be provided.

3.2.1.5 Source Code. Source code for the kernel and operating system utilities shall be available. The code will be used to obtain an understanding of system performance/functionality. This may require that an AT&T System V source license be obtained by SSCL.

3.2.1.6 Process Priority. The scheduling priority of a process shall be settable both at the start of execution and while it is executing. Only the superuser shall be able to increase the priority of a process.

3.2.1.7 Internet Domain Name Service. The Internet Domain service is required on all systems. Each system shall provide the resolve(3) functions as well as the named(8) server.

3.2.1.8 Network File System. The Network File System as described by SUN Microsystems Inc. shall be used to interconnect the platforms.

3.2.1.9 Yellow Pages. The yellow pages facility as described by SUN Microsystems Inc. shall be utilized on all platforms to ensure a single user domain.

3.2.1.10 Signals. The vendor shall provide a list of signals. Standard UNIX signals shall be supported.

3.2.1.11 REXEC. It is mandatory that rexec shall be supported.

3.2.1.12 Console Messages. It shall be possible to direct all console messages to a serial port.

3.2.1.13 System Loading Capability The OS software shall support system loading from a locally attached device. It is required that this can be done remotely, over the network.

3.2.2 Data Compatibility

3.2.2.1 IEEE Floating Point Standard. All systems shall support the IEEE floating point standard.

3.2.2.2 Big Endian Compatibility. Systems shall support the big-endian byte ordering. It is required that binary files written by a FORTRAN program on a native big-endian system can be read by a FORTRAN program on the offerer's system. Similarly, the offerer's system will write FORTRAN binary big-endian files that can be read on native big-endian systems. This compatibility between systems shall be transparent to the user, i.e. no file and/or program changes shall be required to perform data conversion.

3.2.3 Application Software Several SSCL FORTRAN applications including ISAJET, PYTHIA, JETSET, PAW, ZEBRA and GEANT must run on the Type 0 workstations. Test runs of these applications shall be run as part of the benchmark and demonstration requirements described in Section 4.

3.2.4 Front End Network

3.2.4.1 File Systems. Either the standard BSD 4.3 or SYS V.3 or SYS V.4 file system shall be used. Enhanced file systems will be allowed if the file system structure is documented and provided to SSCL. Long file names (31 characters) and symbolic links shall be supported.

3.2.4.2 Operating System Size The operating system, system utilities including compilers and linkers, swap space and temporary file directories shall reside within 150 formatted megabytes.

3.2.4.3. User License. A minimum eight-user license per computer is required.

3.3 Software Requirements

3.3.1 Compilers. All compilers shall support variable names of at least 31 characters. The following compilers shall be provided:

- C.
- FORTRAN 77. The FORTRAN compiler shall support variable names including "$" and underscore "_", the character set A–Z, a–z, !, ", %, &, <, >, $, and _. It shall also support Hollerith constants, case folding except in character strings and Hollerith, octal, hex, and binary constants, I/O formats for hex, types such as REAL*4, INTEGER *2, LOGICAL*1, etc., the IMPLICIT NONE capability, and the DO WHILE/END DO capability. It is required that the FORTRAN compiler also support the ability to "include" files.

Modules developed in one language shall be callable from any of the others. Each compiler shall be able to include information necessary for the debugger. Each compiler shall be capable of doing several levels and types of optimization.

3.3.2 Linkers. The linker shall be capable of linking code generated from any of the compilers. The linker shall recognize external names which are significant in all 31 character positions.

3.3.3 Debuggers. The debugger shall work with all languages and support modules consisting of combinations of languages. The debugger shall work in an X Windowing environment with a graphical user interface. That is, the source code shall be displayed in a window with current execution statement and breakpoint information indicated. Commands to the debugger and

status from the debugger shall be displayed in a separate window. The debugger shall be BSD 4.3 DBX. An assembly language debugger shall also be provided.

3.3.4 Network Utilities. The Berkeley UNIX "r" (rlogin, rsh, etc.) commands shall be supported. Network utilities such as netstat(1), ping(1M), and arp(1M) are required. Telnet and FTP shall be supported. On systems with more than one network interface, these utilities shall be supported transparently. The Berkeley protocol services normally maintained in /etc/services shall be supported along with inetd(8).

3.3.5 Backup/Archive. Both tar(1) and cpio(1) shall be provided.

3.3.6 Shell. The C Shell (csh) and Bourne shell (sh) shall be provided. The C shell shall contain Berkeley style job control as well as file name completion.

3.3.7 X Window System. The X Window system servers (for display full systems) and client utilities for version X11R3 or X11R4 shall be supported. The standard window managers (uwm, twm, etc.) shall run on the system. X Window development toolkits must be provided.

3.4 Graphics Requirements

3.4.1 Graphics software. The following graphics requirements apply to Type () computers.

3.4.1.1 Graphics Kernel System (GKS). An ANSI standard compliant implementation of GKS level 2c must run on all workstations and CPU-servers.

3.4.1.2 The X Window System (X). A working implementation of X Window version X11R3 or X11R4 must run on all workstations.

3.4.1.3 User Application Graphics Software. The High Level Graphics system and Zebra (HIGZ) and the PAW application graphics systems must run on all workstations and CPU-servers. The interactive features of these graphics libraries must run on all workstations.

3.5 Network Requirements

Any networking related hardware or software which is to be supplied by a third party will be the responsibility of the offeror. Support and installation for third party products may be supplied by the third party, but the offeror will maintain responsibility for the performance, installation, and maintenance of these products.

3.5.1 Ethernet Hardware Network Requirements. The system(s) shall be able to support and effectively utilize minimum of two Ethernet (IEEE 802.3) interfaces for the TCP/IP protocol suite.

3.5.2 Protocol/Application Requirements. The following network protocols/applications are minimally required to be supplied on all elements of the system and function through Ethernet (IEEE 802.3).

1. TCP/IP
2. NFS/SMTP/FTP/Telnet
3. X Windows
4. DECnet (phase IV end-node)

3.6 Support Requirements

During the initial receipt, installation and testing of the system, on-site technical support shall be provided at the SSCL. This support shall include detailed documentation of all hardware and software components, maintenance on same, and training to familiarize employees with the system. Dependent upon options being exercised, the following items become additional system requirements.

3.6.1 Maintenance. Maintenance for all system hardware items shall be provided. Standard hardware maintenance plans offered by the supplier may be quoted; but as a minimum shall include on-site support provided by the supplier from his support facility throughout the warranty and full life-cycle of the workstations.

The software offeror shall provide maintenance support for all software provided. Software maintenance support shall include all software updates (major and minor) as they become available. These updates shall include executable code delivered on appropriate magnetic media, and supporting documentation. In addition, technical support via telephone on-site, on an as required basis must be provided.

3.6.2 Training. Training must include the availibility of formal classroom training of all commercial-off-the-shelf training courses on a timely and routine schedule for all available courses.

3.6.3 Documentation. All existing documentation applicable to the operation and maintenance of the proposed system shall be provided along with any current and future updates. This is to include, but is not limited to, copies of all system and user-level documentation in hardcopy form with right to copy, plus on-line man (1-8) pages. Advice regarding such documentation shall be provided as part of the technical support option.

3.7 Operational Requirements

This section covers the operational requirements of the hardware described in this document. These requirements are summarized in Figure 3.7.

3.7.1 Footprint

- The maximum footprint for the networked front end (the Type 0 workstation) is based on the premise that the CPUs shall be mounted on shelves in 6 foot cabinets with a minimum of 8 processors per cabinet. The footprint of an individual cabinet shall not exceed 10 square feet and shall not be greater than 4 feet in depth. This footprint does not include surrounding space needed for maintenance procedures.
- The console for each subsystem will have a footprint not exceeding 16 square feet.

3.7.2 Power. AC voltage shall be 120 volts, 1 phase current at 60 Hz. Maximum power utilization for the networked front end (Type 0 machines) will not exceed 750 watts per processor.

The footprint numbers only include space for hardware. The maintenance requirements will be considered by the implementers. The vendors are responsible for putting the appropriate equipment into the designated area. The values stated in the following table are the maximum allowable for a Type 0 workstation.

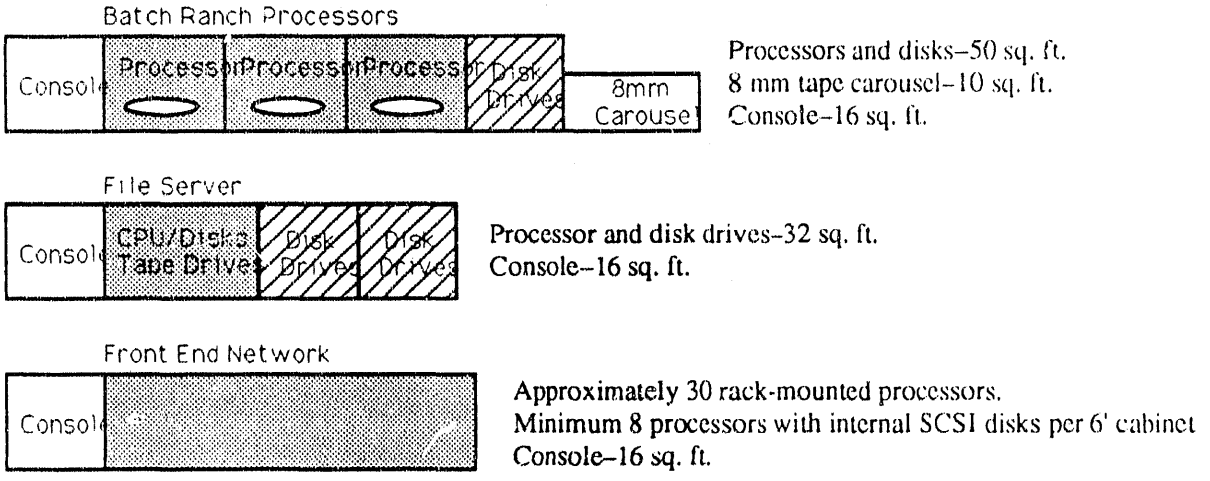
Cabinet Footprint	Power per workstation	A/C per workstation
≤10 sq. ft. ≤4 ft. depth	≤750 W	≤2600BTUs/hr.

3.7.3 Air conditioning requirements. The heat load of the networked front end processors shall not exceed 2600 Btus per hour per computer.

3.7.4 Current. The current for the network front end machines shall not exceed 6 amps at 120 volts.

Example Layout for Computer Room -- Phase I

(Drawing not to scale)



Note: Console space is based on one console per functional unit.

Figure 3.7. Operational requirements for the simulation facility.

4. Qualification Requirements

4.1 Benchmarks

The results of the benchmark tests will form the major criteria in selecting qualifying equipment. The SPECmark results are required; the individual benchmark scores within the SPEC suite will be evaluated separately with a preference being given to the scientific codes. However, the Physics benchmark tests are the most important evaluation criteria.

4.1.1 SPEC

4.1.1.1 SPEC Benchmark Suite. The Systems Performance Evaluation Cooperative (SPEC) benchmark suite will be run on all platform types. The suite consists of the SPEC Benchmark Release 1.0 benchmarks. The benchmarks are described in the SPEC Newsletter available from SPEC. The results from each benchmark will be compared for each system within a given platform type.

4.1.1.2 Front End Network. The minimum acceptable SPECmark for a front end network CPU is 18.

4.1.2 SSCL Physics Suite. This suite consists of six FORTRAN programs based on SSCL math and physics applications. The physics codes consist of three Monte Carlo simulations; namely, an ISAJET test program, a PYTHIA test, and a JETSET test. The math codes consist of Linpack and a Fourier transform program, FOURA. The math codes will be used to compare single- and double-precision speeds. The sixth program, FIO, is a FORTRAN binary data I/O test. All benchmarks must be run with two levels of optimization: 1) minimum optimization (un-optimized), and 2) maximum optimization. The results from both runs must be reported, and the vendor must supply the output from the runs. Accuracy of the results is required. Any necessary vendor modifications to the code to effect this running must be reported to the SSCL.

4.2 Demonstrations

A set of demonstrations shall be performed to verify the compliance of the Type 0 workstations with the requirements os paragraph 3.2.3. These demonstrations will be performed only on those workstations found to be in the competitive range following the initial evaluation of proposals by the SSCL. These demonstrations are not a requirement for qualification, but the successful completion of them is a requirement for final selection. Demonstrations of the following applications is required:

4.2.1 ZEBRA. ZEBRA is a FORTRAN application which provides a common data format and data management capabilities for CERNlib. The ZEBRA test should be run first as the other tests rely on ZEBRA working properly.

4.2.2 PAW. PAW (Physics Analysis Workstation) is a system for the interactive analysis of data. Several PAW examples including PAWEX1 through PAWEX9 shall be run to demonstrate the ability of the workstation to run this application. These examples produce graphical output. A color monitor connected to the Type 0 workstation shall be used to verify this demonstration.

4.2.3 GEANT. A GEANT test program called gexam1 shall be run on the Type 0 workstation to verify compliance with paragraph 3.2.3.

REFERENCES

1. Physics and Detector Simulation Facility Specifications, SSCL-275, Computer Acquisition Working Group, G. Chartrand, et al.,, revised June, 1990.

TRADEMARKS

References to the following list of products has been made throughout this document. It is recognized that the products are proprietary and that the product names are registered trademarks

1. DEC is a trademark of Digital Equipment Corporation.
2. Ethernet is the registered trademark of the Xerox Corporation.
3. OSF and OSF/Motif are trademarks of the Open Software Foundation. OSF/1 is a trademark of the Open Software Foundation.
4. POSIX is a trademark of the Institute of Electrical and Electronic Engineers (IEEE).
5. SPECmark and SPECthruput are registered trademarks of the Systems Performance Evaluation Cooperative.
6. SUN is a trademark of Sun Microsystems.
7. UNIX is a registered trademark of the American Telephone and Telegraph Corporation. UNIX is used as a short form for UNIX Operating System.
8. VAX, VMS, and DECNET are registered trademarks of the Digital Equipment Corporation.
9. Winchester is a trademark of the International Business Machines Corporation.
10. X Window System is a registered trademark of the Massachusetts Institute of Technology.

END

DATE FILMED

02 / 04 / 91