AMS Note AMSnote-2002_01_02
February 4, 2002

Computing Facilities for the AMS-02 ISS Mission

V. Choutko, A. Klimentov and M. Pohl

Abstract

The AMS-02 detector is scheduled to take data on board of the ISS from the year 2004 onward (UF-4 flight), for at least three consecutive years. In this note, requirements for the Payload Operations and Command Center (POCC) for this mission as well as the AMS-02 Science Data Center (ASOC) are presented. Requirements concerning detector commanding, data monitoring, event processing, data storage and physics analysis are discussed and solutions are proposed. A preliminary cost estimate is evaluated and an action plan is drawn up.

Introduction

The AMS-02 detector is scheduled to begin data taking in year 2004 (UF-4 flight) and continue to operate through at least 2007. Facing the increased complexity of the new detector design, the long running time and the way how the data will be transmitted from the detector to the production facilities, a modernisation and extension of the existing data center hardware and the purchase of new computers is required. To estimate the scale of the upgrade, we will assume that AMS-02 software will be developed following the same principles as used for AMS-01 one [1].

The on board computing equipment is concentrated in the providing experiment health and status information to the ISS crew and AMS ground shifters, buffering AMS science and slow control data. The computing equipment on ground can conceptually be divided into the following functional units:

- Ground Support Computers (GSC): receive transmitted control and science data from POIC[1)]; buffer science, housekeeping and NASA ancillary data for transmission to POCC and ASOC (more about ground data handling can be found in [2]).
- Payload Operations and Command Center (POCC): center where AMS operations take place, including commanding, storage and analysis of house keeping data and partial science data analysis for rapid quality control and feed back.
- AMS Science Data Center (ASOC): receives and stores all AMS science and house keeping data, as well as ancillary data from NASA, ensures full science data reconstruction, calibration and alignment; keeps data available for physics analysis; archive all data.

The GSC equipment will be located at Marshall Space Flight Center in all phases of the mission. The rest of the equipment must be relocated during the different phases of the mission. To give an example, we expect that the POCC and ASOC will first be set-up at the AMS-02 detector integration site, then move to KSC together with the detector itself for the pre-flight tests, then to JSC or MSFC to setup POCC for the detector activation period , and be moved to their final location only after successful installation on ISS and running-in of the experiment. It is also envisaged to distribute the functions of the ASOC between two sites, one at MIT and one at CERN. Experience in operating the center will allow to correctly balance the loads between the two sites. Note that POCC and ASOC may be located at the same or at different sites [2].

AMS-02 Raw Data Volumes and Event Rates

Table 1 shows the AMS-02 related data transmitted between ISS and AMS centers.[2)]

The average cosmic ray fluxes for the ISS orbit for the years 2004 and following, as well as the estimation of the AMS acceptance is discussed in [3]. It yields an average event rate 200 to 250 Hz after the Level 3 trigger. Full reconstruction will be possible for 80 to 85% of the total events, giving a final event rate of 180 ± 20 Hz.

[1)] Payload Operations and Integration Center at MSFC, Huntsville, Alabama

[2)] It is important to mention that AMS data also will be recorded to the HDD of onboard computer and later shipped to the ground [4]

Stream	Bandwidth	Data Category	Volume (GB/day)	Volume (TB/year)
High Rate	3-4 Mbit/s	Scientific Calibration	30.1-41.0 0.03	11-15 0.01
Slow Rate	16 kbit/s	House Keeping	0.2	0.06
NASA Anciliary Data			0.03	0.01

Table 1: AMS-02 Data Volumes

Ground Support Computers at Marshall Space Flight Center

AMS computers placed at MSFC will receive the data from Payload Operation and Integration Center (POIC), store them until retransmitted to the AMS Scientific Data Center (ASOC) and if needed to the remote AMS POCC. The GSC computers will have to run unattended 24h/day, 7 days/week, thus system reliability will be the vital issue. The system should be able to store about 2 weeks data, i.e. 600 Gbyte. The configuration will include :

- dual-CPU 1.5^{+}GHz;
- 512 Mbyte RAM;
- 650 Gbyte of disk space (600 Gbyte data disks, 50 Gbyte system, software and scratch area);
- 2×100 Mbit/s (or 1Gbit/s) network controllers;
- UPS unit;
- Linux operating system.

Two identical systems (computer and disks) will be installed and run simultaneously and independently. Thus, in case of computer failure the data transmission will be continued without interruption. The choice of the vendor will be done in consultation with MSFC computing staff and based on the maintenance contracts MSFC has. Thus, most likely Compaq or Sun systems will be used in agreement with MSFC requirements.

Monitoring and Commanding Computers at POCC

AMS computers at the POCC have to handle all activities related to the operation, control and monitoring of AMS detector:

- send commands to AMS detector;
- receive data from POIC (and from GSC at MSFC);
- monitor data and run quality control programs;
- process about 10% of data in “real-time” mode to provide fast information to the shift taker.

We intend to use TReK[3)] workstations (WS) [5] for commanding and slow control monitoring. TReK software satisfies the NASA computing security requirements and been designed for the remote communication between payloads control room and POIC. Extensive tests of data transmission between MIT and CERN using TReK have already been made [2]. It is important to stress that commanding and detector control will be done only from one single place, the POCC. However, detector monitoring programs can be run by experts from different geographical locations.

The proposed TReK work station configuration includes:

- dual-CPU 1^{+}GHz Pentium;
- 1 Gbyte RAM;
- 100 Gbyte of disk space;
- 3×100 Mbit/s (or 1Gbit/s) network controllers;
- Windows 2000 operating system

We will need five workstations for commanding and monitoring: primary and backup for the commanding and three WS to monitor the AMS status and slow control parameters, graphical event display and subdetectors status displays. The monitoring workstations will run under Linux, while TRek requires a Windows operating system.

Five more computers (see "processing node" and "disk-server node" in the next section) will be used to run online data processing and quality assurance programs. This will allow to process approximately 10% of AMS data with a delay less of than 15 minutes after data arrive to the POCC. A total of 0.5 TBytes of disk space will be available for the online computing. These online computing facilities will be fully integrated with the offline data production and analysis computing systems at the ASOC, using the same operating system, software tools and packages. The same program will be used for data processing in both centers.

The AMS Science Operations Center

The AMS Science Data Center (ASOC) will provide all facilities required to reconstruct AMS-02 science data, calibrate and align the detector and store results of reconstruction as well as Monte-Carlo simulations. In this section, we discuss the requirements associated with each of these functions and estimate the required resources.

Data Transfer and Networking

It is important to understand if public Internet links are sufficient to serve AMS-02 needs for data transfer from MSFC to remote AMS centers or if dedicated links are required. To collect data on this issues, test have been made to simulate the AMS-02 data transfer between USA and Europe [2]. From these tests it was concluded that in its current configuration the Internet provides sufficient bandwidth to transmit data from MSFC to a remote POCC at a rate approaching 4 Mbit/s, and that retransmission is needed for less than 0.1% of the data. From current estimates the required throughput for AMS-02 data would use less than 5% of the bandwidth of the existing CERN/USA line, and less than 3% of Very High Performance Backbone Service (vBNS) to which MIT and most of the US labs are

[3)] Telescience Resource Kit

connected. This fraction of usage will be acceptable even when the LHC experiments begin operation. We thus conclude that dedicated networking is not required for AMS-02 operation.

AMS-02 Event Production

The event data production is described in more details in [6]. We just briefly repeat the key ideas. The AMS-02 event reconstruction facility should fulfil the following tasks:

- ensure the full reconstruction of about 10% of the input events in the real time mode with an average time delay of no more than half an hour;
- after an initial period of detector calibration and alignment[4)], ensure the full reconstruction of 100% of the events in a quasi real-time mode, i.e. with a typical time delay time of less than one day.

As is shown in reference [6], the total power of the AMS-02 processing farm must be equivalent to 400x450MHz Pentium II processors, including 20% contingency and reprocessing. A homogeneous farm of multi-processor Pentium (AMD) based systems running Linux is proposed as a cost effective way to provide this processing power. Depending on the processor clock speed, the farm will contain 45 to 50 independent nodes. There will be two types of nodes, “pure” processors and servers, with the following typical configurations:

- “processing node”:
  - dual-CPU 1.5^{+}GHz Pentium/AMD;
  - 1 Gbyte/CPU RAM;
  - IDE RAID controller, 6x120GB+ IDE disks;
  - 2×1 Gbit/s network controllers;
  - 1×100 Mbit/s (or 1Gbit/s) network controllers;
  - Linux operating system
- “server node” :
  - dual-CPU 1.5^{+}GHz Pentium/AMD;
  - 1 Gbyte RAM;
  - 2 Tbyte of disk space, combination of IDE disks and SCSI RAID arrays;
  - 2×1 Gbit/s network controllers;
  - 1×100 Mbit/s (or 1Gbit/s) network controllers;
  - Linux operating system.

AMS-02 Data Base

Tests with an ORACLE database to store AMS-02 catalogues, event tags and production status are done at CERN and INFN Milano. Two computers (primary and backup) will run as database servers. These will be Sun or Pentium PCs with a configuration similar to the disk-server PC used for the AMS Production Farm. Access to the database will be given to all members of the collaboration via a WWW interface. The approaches to such an interface are currently under development.

[4)] This will probably requires one to three months from the beginning of the data taking

AMS-02 Data Storage[5)]

Access to the raw data must be provided for the following purposes:

- detector verification studies; the access will be granted to the Laboratories participated in building detector and readout electronics;
- calibration (access to selected data, for example data from tracker, TRD or ECAL);
- alignment (access to selected data);
- event visualisation (access to a very small fraction of data);
- data processing by the general reconstruction program; it will produce the basic event parameters, event summary data (ESD) and do events classification (access to all the data, once)
- data reprocessing (access to all data once and to selected data more frequently).

Event tag information must allow event classification according to the needs of physics analysis. Each event must be unambiguously tagged in terms of reconstruction program version, calibration and alignment data version. Tag information for all events during the whole period of data taking must be kept on direct access disks.

Based on the experience of the first flight the calibration information and raw data taken during the last 9 months should be also on direct access disks at any given time. In addition, 30% of all ESD are the data of interest and are supposed to be kept on disks for the same period of time. All taken and reconstructed data must be archived and we will need a proper archiving system. The DLT solution used for the first flight will not be able to archive Tbytes of data. In summary, direct access must be provided to 20 Tbyte of data, 60Tbyte of data must be archived and accessible by demand.

The data volume sizes given in Table 2 are derived from Table 1 and by multiplying the ESD size by 200Hz (upper limit on the event rate as discussed in [6]).

Data Type	Total size per year (Tbyte)	Data On Direct Access disks (Tbyte)	Data to be archived (Tbyte)
Raw Data	11-15	8.3 - 11.3	11-15
Event Summary Data	44	8.3	44
Event Tag	0.6	0.6	0.6
Total	55.6 - 59.6	17.2 - 20.2	55.6 - 59.6

Table 2: AMS-02 data volumes

Proposed Solutions

Archiving system: A DVD RAM of 9.4 Gbyte is available since 2001 with an access time of 150 ms, but tape robots with 200 GB (LTO or SDLT) cartridges still look more attractive in comparison. The market situation will be reassessed in the beginning of 2003 to prepare a decision.

Direct access disks: 40 Gbit/inch^2 are available now, so 200 Gbyte disks appear on the market this year, 1.5 year yealier it was predicted. The disks organisation is discussed in the details in [6].

[5)]The definitions used in this section are presented in [6]

An attachment of 3 TByte RAID-Array disks per dual-CPU Linux "disk-server" in combination with IDE disks on the processing nodes will probably be a realistic solution.

An alternative is to use HPSS[6)] or CASTOR[7)]. The HPSS and CASTOR allow to have fast access to the Pbytes of disk space and organise the data archiving without additional investments into a robotic system. However, but we are not currently discussing the usage of HPSS as a main data storage for AMS, because of the uncertainty on the location of the ASOC and the related compatibility issues.

AMS-02 Physics Analysis Facility

The main mission of the AMS-02 Physics Analysis will be to allow fast batch and interactive access to the bulk of the AMS-02 data. Allowing for a 100% increase in the number of actively working physicists and an 100% increase of the average event access time leads to a system equivalent to 20×1 GHz Pentium III processors and 2 Tbytes of disk space.

A combination of SMP system and RH Linux dual-CPU PCs (with the configuration similar to "AMS processinbg node", but with the reduced number of IDE disks) are proposed as a possible solution. Depending on the processor clock, the system will contain 8 to 16 processors. Both Sun or Compaq Alpha platforms will probably qualify for this purpose.

AMS-02 Computing Facilities Cost Estimation

The costs given in this chapter are preliminary and do not include the cost for desktop equipment and the corresponding software for the collaboration members. We do not anticipate any AMS personnel to operate the data center. We rather envisage that this will be done via regular shifts of physicists, with the help of two persons, a Data Processing Coordinator and a Computing Coordinator. One FTE software engineer will be able to run all AMS computing facilities.

Table 4 presents the situation at the time of the AMS-01 mission and the projected characteristics in 2003 for memory and processors. The projection is based on our own experience and predictions taken from [7, 8]. The estimation of CPU power is given in SPECint95[8)].

AMS-02 Computing Facilities

The cost estimate of the computing facilities is derived by multiplying the requirements by the expected unit costs of 50 US$/SPECint95 of processing power. It is proposed to use inexpensive IDE disks for the computers of the POCC and the bulk storage at the ASOC. More expensive SCSI disks with the possibility to build RAID-Arrays are proposed for the GSE at MSFC, for ESD storage and the physics analysis computers. We summarize the AMS-02 computing facilities and their cost in Table 4.

[6)] High Performance Storage System

[7)] http://wwwinfo.cern.ch/pdp/castor/Welcome.html

[8)] Processor benchmark, used for most of HEP codes. 1 SPECint95 = 40 SPECint92 = 10 CERN units = 40 MIPS. A Compaq quad-CPU Alpha Server AS4100 is rated 130 SPECint95 at 600 MHz, a Pentium II with Intel SE440BX2 Motherboard is rated 24 SPECint95 at 600 MHz.

Components	1998	2001	2003
Memory	0.35 um feature size 64Mbits/s chip 22 US$/Mbyte	$\rightarrow$	0.13 um feature size 4000 Mbits/chip 0.40 US$/Mbyte
High-end processors	10 SPECint95 18 kUS$/system	$\rightarrow$	100 SPECint95 5 kUS$/system
Magnetic Disk	100 US$/Gbyte 6 Mbyte/s - disk	55 US$/Gbyte Tbyte RAID array up to 50Mbyte/s -array	8 US$/Gbyte multiTbyte RAID array up to 100Mbyte/s - array
Intel PC	dual-CPU Intel PII rated at 400 MHz, 384 Mbyte RAM 7.5 kUS$	dual-CPU Intel PIII rated at 1 GHz, 1 Gbyte RAM 7 kUS$	dual-CPU Pentium rated at 2GHz, 2Gbyte RAM 6 kUS$

Table 3: AMS-01 situation and projected characteristics for disks, processors and computers

Milestones and Spending Profile

It is desirable to buy the bulk of computing equipment as late as possible due to the improving price-performance ratio. However, from the management point of view it will be better to build the infrastracture gradually. We propose a spending profile and milestones as given in Table 5 (in case that UF-4 launch date is shifted to the right, the purchasing and installation of disk pools in ASOC and MSFC will be adjusted accordingly). After the initial capital investment, annual costs arise due to maintenance, yearly archiving of 15 Tbyte of data, and the corresponding increase of data volume. Processors have a limited lifetime and we assume that they will be replaced by a biannual upgrade of 20%. This can be estimated as an additional 20kUS$/year for the year 2004 and 30kUS$/year after 2004.

References

[1] V. Choutko, A. Klimentov "OO Software For The AMS Experiment. (3 Years Of the Development)", AMS note 1999_10_03;

[2] P. Fisher, A. Klimentov "Ground Data Handling for AMS02 ISS mission", AMS note 2001_05_01;

[3] V. Choutko "Particle Rates for Year 2003", AMS note 2000_02_03;

[4] M. Capell "AMS Electronics", presentation to R.Nygren, CERN , Oct. 2000.

[5] TReK. Telescience Resource Kit. http://trek.msfc.nasa.gov;

[6] V. Choutko, A. Klimentov "Data Processing and computing for the AMS-02 ISS mission.", AMS note 2000_02_04;

Function	Computer	Qty	Disks (TBytes)	Cost (kUS$)
GSC @ MSFC	Compaq or Sun dual-CPU 1+GHz	2	1.4 RAID-array	55
POCCx2	Pentium or AMD dual-CPU 1.5+GHz	20	1.0 RAID-array IDE disks	150
ProductionFarm	Pentiums or AMD dual-CPU 1.5+GHz	50	10.0 RAID-array IDE disks	350
Database Servers	Pentium or Sun dual-CPU 1+GHz	2	0.2	50
EventStorage and Archiving			25.0 Raid-Array	200
Analysis	multi-CPU SMP computer, 4Gbyte RAM, 300 SPECint95	1	RAID Array x000, 2TB	55
Total		75	40	860

Table 4: Cost estimate of the proposed AMS-02 computing facilities

[7] PASTA - The LHC Technology Tracking Team for Processors, Memory, Architectures, Storage and Tapes, Status Report, August 1996.

[8] L. Robertson, 3rd LHC Computing Workshop, Marceille, August 1999.

Year	kUS$	Comments
2000	18.4	AMS02 reconstruction and simulation code development,
		farm studies, purchase 1 dual-CPU rated at 900MHz
		risk analysis for quad-CPU Pentiums
	7.7	install TReK workstations at CERN and MIT to evaluate SW,
		data transmission tests between CERN and MIT using TReK
	15.4	computers maintenance
subtotal	41.5	
mid 2001	11.5	production farm prototyping with various type of processors (
		AMD, Itanium) and HW architecture (singleCPU, dual-CPU machines)
		to choose the most cost effective products
		Setup the farm testbed at CERN
	17.5	computers maintenance
end 2001	35.0	10% prototype of processing farm, code devopment for
		the distributing data processing, database prototype.
		Process AMS01 data with new SW framework
	24	2TB disks to store MC02 files
	14.5	2 tape drives with capacity 200GB (or more)
	7.5	running cost (racks, UPS, cables)
subtotal	110.0	
beg 2002	15.0	Installing computer in POIC at MSFC, to do realistic
		tests between MSFC and MIT, MSFC and CERN
end 2002	105.0	40% prototype of processing farm, use the farm for MC production
	20.0	setup database servers
	14.5	computers maintenance
	7.5	running cost
subtotal	162.0	
beg 2003	70.0	60% prototype of processing farm
		choice of archiving system
	130.0	purchase 12 Tbyte disks to equip processing farm disk pool
		purchase the archiving system
		make the decision of SMP vs distributed computing and
		make choice of the analysis computer
mid 2003	52.0	purchase and setup AMS POCC computers and disks
	21.0	purchase ans setup final configuration in MSFC
	170.0	setup processing farm and database servers in final configuration
end 2003	83.0	purchase and setup analysis computers and RAID system
	13.0	computers maintenance
	15.0	running cost
subtotal	554.0	
beg 2004	70.0	purchase disks to setup 13 Tbyte disk pool
	20.5	purchase disks to setup disk pool for GSC @ MSFC
	13.0	computers maintenance
end 2004	20.0	Upgrade of computers purchased in Years 2001, 2002.
	15.0	running cost
subtotal	138.5	
Grand Total	1006.0	for years 2000 - 2004

Table 5: Spending Profile for the Computing Infrastructure