The KLOE Computing Environment

The KLOE Collaboration
Presented by Matthew Moulson

***Abstract*— The KLOE experiment at DAΦNE is largely devoted to the study of discrete symmetries in the $K_S K_L$ system. During 2001–2002, KLOE collected 1.4 × 10^9 ϕ decays, for a total data volume of nearly 200 TB. We describe the KLOE computing environment, outlining our solutions to issues of calibration, reconstruction, data reduction, and data access for analysis. We discuss the performance of the system, with particular emphasis on the data-processing and transfer rates, and also describe the generation of large Monte Carlo samples. The KLOE computing environment has been demonstrated to be flexible and scalable in the presence of rapidly evolving data-taking conditions.**

***Index Terms*— Computer facilities, data acquisition, data handling, data management, data processing, elementary particles, simulation.**

I. INTRODUCTION

KLOE is a large, general-purpose detector permanently installed at DAΦNE, the Frascati ϕ factory, an e^+e^- machine with $W \approx m_\phi \approx$ 1.02 GeV. KLOE is optimized for the study of discrete symmetries in the neutral-kaon system, but the physics program of the experiment includes studies of charged-kaon decays, radiative decays of the ϕ meson, η decays, and the measurement of the cross section for $e^+e^- \to$ hadrons below 1 GeV. The DAΦNE design luminosity is 5×10^{32} cm^{-2} s^{-1}, which corresponds to a ϕ-production rate of 1.7 kHz. Given the breadth of the KLOE physics program, all ϕ decays are interesting, and as much of this rate as possible must be acquired and reconstructed.

The KLOE detector consists essentially of a large drift chamber (DC) surrounded by an electromagnetic calorimeter (EmC). The drift chamber [1] is strung with 12 582 stereo sense wires. We measure the signal arrival time for each sense wire, and the integrated current for each group of 12 wires. The lead/scintillating-fiber calorimeter [2] consists of a barrel and two endcaps. The modules are read out on both sides by a total of 4880 photomultiplier tubes; we record the time and amplitude of each signal. In total, the detector has about 23 000 front-end electronics channels. The readout system is discussed in detail in [3].

To complete its physics program, KLOE needs to collect a data set of at least a few fb^{-1} of integrated luminosity. KLOE took its first events in 1999, but significant DAΦNE luminosities were first achieved in 2001. During 2002 data taking, the maximum DAΦNE luminosity was 7.5×10^{31} cm^{-2} s^{-1}. Although this is lower than the design value, the performance of the machine during 2002 was much improved with respect to previous years, and the KLOE experiment was able to collect as much as 4.2 pb^{-1} per day. The combined KLOE 2001–2002 data set amounts to about 450 pb^{-1}, or 1.4×10^9 ϕ decays. This data set is characterized by a signal-to-background ratio that varies significantly as a function of time.

A series of recent upgrades to the machine, including an overhaul of the interaction region inside the KLOE detector, is expected to increase the peak luminosity to 2×10^{32} cm^{-2} s^{-1}—significantly closer to the design value. Data taking with KLOE is scheduled to restart by early 2004. We expect to collect a data set of about 2 fb^{-1} during the upcoming running period.

II. DATA FLOW

A good starting point for the discussion of KLOE computing is the data-flow diagram presented in Figure 1.

Raw data is read out from the detector by the data-acquisition (DAQ) system and formatted and written to raw-data files of size 1 GB by the online system. A typical run might consist of about 20 such files. Calibration processes run on the online system concurrently with the acquisition and sample events from DAQ memory buffers. These processes perform non-CPU-intensive tasks such as adjustment of the global time and energy scales for the EmC and monitoring of the time-to-distance relations for the DC. The results are validated automatically, typically within an hour of run closure. A complete and time-consuming calibration of the time-to-distance relations for the DC is started from closed files only if the calibrations have changed significantly. This typically happens once every few weeks, essentially when the atmospheric pressure changes by more than 1%. In this case, the results are validated manually.

Once the calibrations for a run have been validated, the reconstruction program is automatically started. The raw-data

Manuscript received 14 November, 2003.

The KLOE collaboration: A. Aloisio, F. Ambrosino, A. Antonelli, M. Antonelli, C. Bacci, G. Bencivenni, S. Bertolucci, C. Bini, C. Bloise, V. Bocci, F. Bossi, P. Branchini, S. A. Bulychjov, R. Caloi, P. Campana, G. Capon, T. Capussela, G. Carboni, G. Cataldi, F. Ceradini, F. Cervelli, F. Cevenini, G. Chiefari, P. Ciambrone, S. Conetti, E. De Lucia, P. De Simone, G. De Zorzi, S. Dell'Agnello, A. Denig, A. Di Domenico, C. Di Donato, S. Di Falco, B. Di Micco, A. Doria, M. Dreucci, O. Erriquez, A. Farilla, G. Felici, A. Ferrari, M. L. Ferrer, G. Finocchiaro, C. Forti, A. Franceschi, P. Franzini, C. Gatti, P. Gauzzi, S. Giovannella, E. Gorini, E. Graziani, M. Incagli, W. Kluge, V. Kulikov, F. Lacava, G. Lanfranchi, J. Lee-Franzini, D. Leone, F. Lu, M. Martemianov, M. Matsyuk, W. Mei, L. Merola, R. Messi, S. Miscetti, M. Moulson, S. Müller, F. Murtas, M. Napolitano, A. Nedosekin, F. Nguyen, M. Palutan, E. Pasqualucci, L. Passalacqua, A. Passeri, V. Patera, F. Perfetto, E. Petrolo, L. Pontecorvo, M. Primavera, F. Ruggieri, P. Santangelo, E. Santovetti, G. Saracino, R. D. Schamberger, B. Sciascia, A. Sciubba, F. Scuri, I. Sfiligoi, A. Sibidanov, T. Spadaro, E. Spiriti, M. Testa, L. Tortora, P. Valente, B. Valeriani, G. Venanzoni, S. Veneziano, A. Ventura, S. Ventura, R. Versaci, I. Villella, G. Xu

M. Moulson is with the Laboratori Nazionali di Frascati dell'INFN, 00044 Frascati RM, Italy (e-mail: moulson@lnf.infn.it).

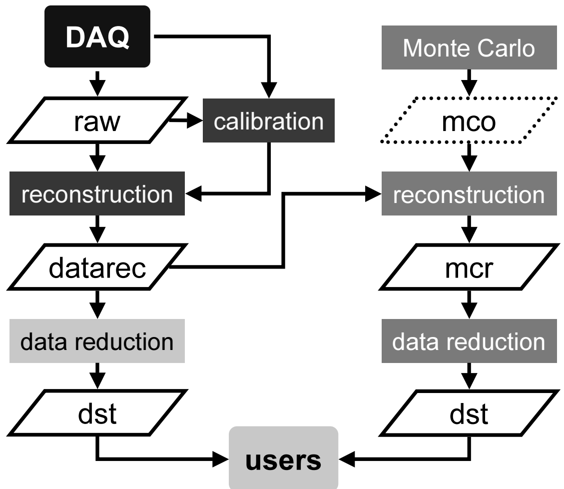

Fig. 1. KLOE data flow

files for the run are processed in parallel on the offline production farm, one per available CPU. As discussed in Section VII, the number of CPUs available is sufficient to allow the reconstruction to keep pace with the acquisition of new data. A single raw-data file typically takes about two hours to process on the offline farm; this determines the turnaround time for the processing of a single run. As output, the reconstruction program writes reconstructed-data files known as *datarec* files. Reconstructed events are significantly larger (12–15 kB) than raw events (∼2.5 kB). The reconstruction program includes filters to reject machine-background and cosmic-ray events; as a result, the total volumes of raw and reconstructed data are comparable, notwithstanding the growth in event size during reconstruction. Reconstructed events are classified into streams and written to separate files as discussed in Section V-A, so that there is one datarec file per stream for each input raw-data file. These datarec files are on the order of a few hundred MB in size, with significant variation from stream to stream.

Finally, the datarec files are used to produce data-summary files (DST files, or DSTs) that contain an essential summary of physics-related quantities for each event, and in which the event size is much smaller (∼3 kB). Datarec files undergo some consolidation at the DST production stage: there is one DST file per stream for each run. The DSTs are the starting point for most user physics analysis.

Monte Carlo (MC) production follows a similar scheme. The simulation executable generates events and writes them to Monte Carlo output (MCO) files, which are then reconstructed in a separate process. During the reconstruction, background extracted from reconstructed events in the data set is overlaid with the events as generated. The information in the MCO files is completely subsumed by that in the reconstructed Monte Carlo files (MCR files), so the MCO files are deleted as the MCR files become available. Due to the need to retain information describing the events as generated, events in MCO and MCR files occupy about 20 and 30 kB, respectively, i.e., they are significantly larger than their counterparts in the data

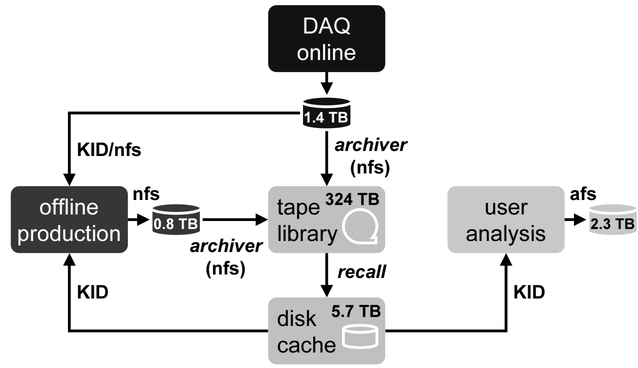

Fig. 2. Schematic view of KLOE data handling

set. MCR files are also subject to data reduction; MC DSTs are produced in a manner identical to that for datarec files (see Section V-B).

KLOE data are written in YBOS format [4]. In datarec, MCR, and DST files, YBOS-formatted events are compressed with ZLIB routines [5], which reduces the data volume by 40%.

III. Data Handling

A diagram of the data-handling scheme is presented in Figure 2. The raw-data files are written to 1.4 TB of disk locally mounted on the online servers. These files are then asynchronously archived to the tape library over an NFS mount. The archiving processes are tailored to minimize the number of tape mounts while guaranteeing enough space on the disk pool. In most cases, reconstruction is performed while the raw-data files are still resident on disk. For input to the reconstruction processes from the online disk, events are either read across an NFS mount or served by the data-handling system using a custom TCP/IP protocol, which is provided by the *KLOE Integrated Dataflow* package (KID) [6]. Reconstruction output is written via NFS to an offline disk pool, from which it is asynchronously archived to tape. DSTs for each run are produced from the corresponding datarec files, often immediately after the run has been completely reconstructed. In this case, the reconstructed events may be read back in across the NFS mount for the data reduction processes. When files already archived and deleted from the online or offline disk pools must be processed on the offline farm, the data-handling system restores the relevant input from tape to a 5.7 TB recall disk cache, from where they are served to the offline processes using the KID protocol. The same access model applies to user jobs running on the analysis farm. In principle, users may need to analyze raw-data, datarec, or DST files. If the files requested are resident on the online or offline disk pools, they are copied to the recall disk cache by the data-handling system to be served to the user processes; otherwise, they are restored to the recall disk cache from tape. Generally, however, only DST files are of interest to the users. A large fraction of the recall disk cache is therefore allocated to DST files to ensure their prompt availability.

The division of the datarec and DST output into separate file streams described in Section V-A increases the total number of files in the data set, but decreases the volume of data that must be read in the context of any particular analysis. The use of the recall disk cache provides some protection from the excessive tape-positioning operations that might otherwise result from this division. Other features of the data-handling system provide additional optimizations. The archiving procedures ensure that individual cartridges in the tape library contain only files corresponding to a single stream. User and production processes normally request many files at once. In most cases, the files corresponding to a given run and stream are stored on a single cartridge; KID implements an automatic prefetch feature based on this fact. Finally, where necessary, production processes request files by order on tape.

A central database based on IBM's DB2 [7] is used to keep track of the locations of the several million files comprising the data set. Each file is logged in the database when it is created. The database entry also contains the reconstruction status of the file, allowing files that require processing to be easily identified.

The backbone of the data-handling system is the KID package, which consists of two pieces: a centralized data-handling daemon, which coordinates the distributed file-moving services; and a client library, with an easy-to-use URL-based interface that allows access to files independent of their locations. KID URLs may incorporate SQL queries used to interrogate the bookkeeping database. Examples of such URLs include the following:

- All raw-data files in the stated run range that have not yet been reconstructed:
  dbraw:run_nr between 23000 and 24000 and analyzed is not null
- All reconstructed files in the $K_S K_L$ stream for a given run:
  dbdatarec:run_nr = 23015 and stream_code = 'ksl'

The database and data-handling server is an IBM F50 running AIX with four 166-MHz PowerPC CPUs and 2 GB of RAM. Two IBM H80s running AIX, each with six 500-MHz RS64-III CPUs and 2 GB of RAM are used as file servers for the offline disk pool, recall cache, and tape library. With the two file servers working in concert, aggregate I/O rates of over 100 MB/s have been obtained. Each file server is connected to 1.8 TB of Fibre Channel (FC) and 1.4 TB of SSA disks (configured in striping mode), as well as to the experiment's IBM 3494 tape library, which has 12 Magstar 3590 tape drives, dual active accessors, and space for about 5400 60-GB cartridges. The tape library is maintained using IBM's Tivoli Storage Manager [8].

Network connections are routed through a Cisco Catalyst 6000 switch. The file servers are connected to the switch via Gigabit Ethernet, as are the servers for the AFS cell, which provides disk space to users and working groups (see Section VI). Connections to all other nodes are via Fast Ethernet.

IV. Online Computing

The DAQ and online systems were designed to sustain a bandwidth of 50 MB/s from detector to tape. The online farm consists of seven IBM H50 servers running AIX, each with four 332-MHz PowerPC 604e CPUs and 270 GB of local disk. In 2002, three of these servers were used for tasks strictly related to the acquisition of data, while the others were used for calibration and monitoring. If needed, all seven servers can be used for DAQ-related tasks.

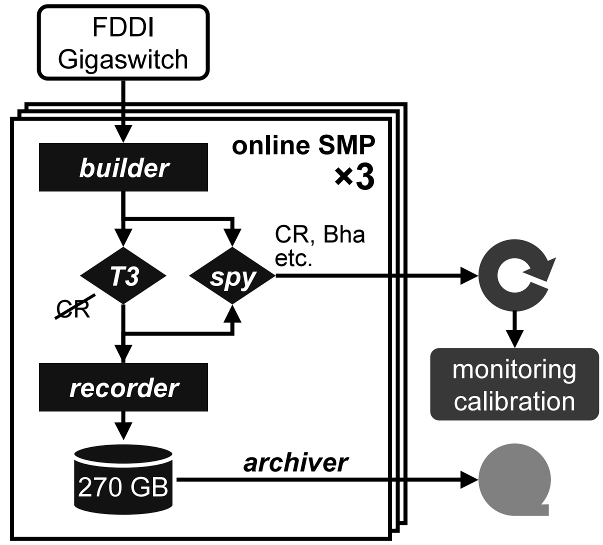

Fig. 3. Processes running on a single online server.

The processes running on each of the three servers dedicated to online tasks are illustrated in Figure 3. Sub-events from the acquisition chains are routed through a Digital FDDI Gigaswitch to the online servers in such a way as to distribute the load evenly among the nodes. On the servers, the sub-events are assembled into complete events and formatted by an event-builder process. Events flagged as cosmic rays (CRs) by the hardware level-2 trigger are then subject to review by the level-3 trigger implemented in software. For some types of events (e.g., $e^+e^- \to \mu^+\mu^-(\gamma)$ and $e^+e^- \to \pi^+\pi^-(\gamma)$), enforcing the CR veto at level 2 incurs notable inefficiencies. At level 3, a fast reconstruction is performed, which allows for better discrimination of CR events. Events passing the level-3 trigger are recorded and asynchronously archived to tape. A number of spy daemons review buffered events before and after the level-3 trigger and select certain event types (Bhabha, CR, etc.) for calibration and monitoring purposes. These events are then written to a circular buffer, from which they are served by KID to the appropriate processes.

Each online server can handle a rate of at least 2.4 kHz. For comparison, in 2002, with an average luminosity of 4×10^{31} cm^{-2} s^{-1}, the level-2 rate (without enforcement of the CR veto) was about 3.5 kHz. This rate was divided among three servers, with up to seven available. The installed online CPU power is adequate to meet the needs of the experiment for the foreseeable future.

V. Offline Computing

The machines for the offline production and analysis farms are a mix of IBM B80 servers running AIX, each with four

375-MHz Power3 CPUs; and Sun E450 servers running Solaris, each with four 400-MHz UltraSPARC II CPUs. In all, 23 B80s and 10 E450s are available, and provide a total processing power equal to about 110 of the processors installed in the B80s. Currently, 16 to 19 B80s and 8 E450s (76 to 90 B80 CPU equivalents) are used for the offline production farm. The partitioning is flexible enough to allow machines to be transferred between production and analysis tasks as the need arises. The offline production and analysis software framework is provided by ANALYSIS_CONTROL [9], with various customizations for KLOE, such as the introduction of the KID interface.

A. Reconstruction and DST Production

The tasks performed during event reconstruction are summarized in the upper part of Figure 4. Raw events are first reconstructed in the EmC. The EmC reconstruction is fast (4 ms per event[1]), and yields cluster coordinates, energies, and times. This information is sufficient to allow rejection of a significant portion of the machine background, as well as of essentially all cosmic-ray events residually present. The effectiveness of the filter depends somewhat on machine conditions; in 2001–2002 running, about 40% of the input rate passed the filter. Reconstruction in the DC includes pattern recognition, track fitting, and vertex finding, and is the most CPU-intensive reconstruction task. On average, DC reconstruction takes about 40 ms per event for events passing the filter, where this number is a sample-weighted average of the reconstruction times for Bhabha events ($\sim$30 ms), ϕ-decay events ($\sim$120 ms), and a small fraction of unrejected background events (15–40 ms). After reconstruction in the DC, events are classified into streams on the basis of topological information and written to separate datarec files. At present, five streams are defined for K^+K^-, K_SK_L, $\rho\pi$, radiative ϕ, and Bhabha events. Events may be written into more than one stream in principle; in practice, the overlap between streams is at most at the level of a few percent. The files for each stream (with the exception of that for Bhabha events) are then used to make DSTs.

DST production proceeds in three stages as illustrated in the lower diagram of Figure 4. First, a more refined event selection may be applied, in order to reduce the stream volume. Second, any topology-specific algorithms necessary to complete reconstruction of the stream are run. Finally, a set of bank-selection and compression routines are applied. These routines discard banks containing information on individual EmC and DC hits and consolidate information concerning reconstructed topologies such as EmC clusters or DC tracks. In the DST production for any given stream, the first two stages may or may not run; the last stage runs always. As a performance example, consider the K_SK_L stream. The completion of all three stages takes about 12 ms per event and results in a reduction of the data volume by a factor of more than six. When tracks are first reconstructed in the DC, the pion mass is used for the evaluation of multiple-scattering and energy-loss corrections. Because of the low velocities of charged kaons at KLOE ($\beta \approx 0.2$), K^+K^- events must be completely retracked once identified, with the correct mass hypotheses. As a result, DST production for the K^+K^- stream takes 115 ms per event. In addition, since the K^+K^- DSTs contain the results of an additional track-reconstruction pass, the DST volume for this stream is slightly larger than it is for the other streams.

[1]Throughout this paper, all CPU times are referred to a single 375-MHz Power3 processor installed in an IBM B80 server

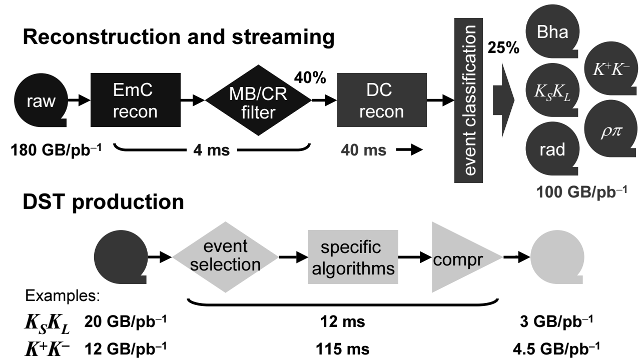

Fig. 4. Flow charts for offline production. Top: Reconstruction and streaming. Bottom: DST production. All numbers are indicative and refer to 2001–2002 averages.

B. Monte Carlo Production

The other significant task that runs on the offline production farm is MC production. In particular for the study of physics backgrounds to some of the rarer decay channels, we require MC event samples that are of a statistical significance comparable to that of the data set itself. MC production campaigns currently running include $\phi \to$ all final states at a luminosity scale of 1:5 (2.6×10^8 events) and $\phi \to K_SK_L \to$ all final states at a scale of 1:1 (4.3×10^8 events). An interesting aspect of KLOE MC production is that runs in the data set are individually simulated. Specifically, an MC sample is generated for each run in the data set, with the number of events proportional to the integrated luminosity of the run under simulation, and such parameters as machine energy, momentum of the collision center-of-mass, beam-spot position, map of dead detector elements, and trigger thresholds set to correspond to the run conditions. The set of MC events for any group of runs thus automatically reflects the correct average over these time-variable conditions. The simulation itself is based on GEANT 3.21 [10], with a full suite of dedicated ϕ and secondary decay generators and full digitization of the detector response. Accidental activity in the EmC and DC is sampled from $e^+e^- \to \gamma\gamma$ events in the data set and inserted into the MC events run-by-run, with care taken to preserve the temporal profile of the background levels within each run. During reconstruction of MC events, the filtering and event-classification algorithms register their decisions in the data stream. The filters are not enforced, however, and the reconstructed events are not streamed into separate files. Instead, streaming is performed at the DST-production stage. In addition to the events classified on the basis of reconstructed quantities, each MC DST stream contains all events with MC-generated topologies relevant to the physics of the stream.

VI. Analysis Environment

In comparison to other experiments of similar scale, KLOE has relatively modest requirements in terms of CPU resources for analysis. This is in large part because of the relative simplicity of KLOE events; the small DST event size, the use of simple, tabular data structures, and the high level to which DST data are processed (with stream-specific reconstruction algorithms already applied) all contribute as well. Users use the offline analysis farm to produce histograms and Ntuples for their analyses. The final stages of physics analysis, based on this output, are usually performed on individual desktop Linux PC systems, which further lightens the load on the offline farm.

The offline analysis farm consists of four to seven IBM B80 and two Sun E450 servers (corresponding to 20–32 B80-equivalent CPUs). Three of the B80s are typically reserved for batch jobs, with queues managed by IBM's LoadLeveler [11].

Analysis jobs usually use DSTs as input. For the 2001–2002 data, the set of DSTs occupies 4.1 TB; MC DSTs occupy an additional 3.1 TB. Copies of the bulk of the DSTs are normally available on the 5.7 TB recall disk cache for prompt access (Section III). Output is written to user and working group areas on the KLOE AFS cell. The AFS cell is served by two IBM H70 servers, each with four 340-MHz RS64-III CPUs, 850 GB of SSA disks, and 250 GB of FC disks (2.3 TB total). Users can access the AFS cell from their desktop PCs.

As an example, consider the analysis of the 2001–2002 $K_S K_L$ data set, which consists of 7×10^8 events, or 1.4 TB of DSTs. The typical time needed for a user to analyze the entire data set on the farm is six days elapsed with six jobs running in parallel (the default maximum). The output size ranges from 10 to 100 GB, which can be accessed *in situ* on the AFS cell or copied off to a user's desktop PC.

VII. CPU Power and Mass Storage Requirements

During the years 2001 and 2002, the average DAΦNE luminosities were about 2×10^{31} and 4×10^{31} cm^{-2} s^{-1}, respectively. In 2001, the input rate to the offline system was determined by the level-2 trigger rate, and was 2.0 kHz on average. In 2002, the implementation of the level-3 trigger (which obviated the need for collection of a sample of prescaled CR-veto events), together with decreased machine-background levels, reduced the offline rate to 1.6 kHz. The DAΦNE upgrade carried out in the first half of 2003 is predicted to further increase the peak luminosity, to 2×10^{32} cm^{-2} s^{-1}, with favorable background conditions. In the following, our predictions for the 2004 run assume an average luminosity of 1×10^{32} cm^{-2} s^{-1} and background levels identical to those of 2002. The predicted level-3 rate is then 2.1 kHz. Current plans call for approximately a year of running, which would allow for the collection of 2 fb^{-1} in 2004.

TABLE I
CPUs required for offline tasks

	2001	2002	2004 (est.)
Reconstruction	42	30	71
DST production	2	5	13
Monte Carlo	11	23	63
TOTAL	55	58	147

A. Processing Requirements

For the 2001 and 2002 data, about 40% of the input rate required reconstruction in the DC, and about 20 ms were needed to reconstruct an event on average. For 2004, as a result of the predicted increase in luminosity, the fraction of events reconstructed in the DC is expected to rise to 60%, which leads to an event reconstruction time of 33 ms. We will also have to produce DSTs for a fraction of the input rate corresponding to $\sim$330 Hz, at 38 ms per event summed over all streams. Finally, assuming that MC production plans remain the same, we will need to generate, reconstruct, and produce DSTs for MC events at an equivalent rate of $\sim$170 Hz, with $\sim$200 ms per event required for generation, and $\sim$375 ms per event required in total. These considerations lead to Table I, in which we list the number of CPUs needed for the various offline tasks to keep up with the data acquisition. We can usually allocate at most 76 CPUs to offline production. While this has been more than sufficient to handle the load so far, for 2004, we will have to roughly double the amount of offline CPU power available for production.

CPU resources needed for analysis are not included in Table I. The CPU demand for analysis should scale roughly linearly with the integrated luminosity. Thus it might be estimated that for the analysis of 2 fb^{-1} of 2004 data, the number of CPUs available for analysis jobs would need to be increased from the current 20–32 to approximately 100. Analysis jobs, however, are usually I/O limited rather than CPU limited; a more pressing concern is the expansion of the recall disk cache for prompt access to the new DSTs.

B. Mass Storage Requirements

In Table II, we list the total volume of KLOE data by file type (DST volumes for data and MC are summed in the table). For the 2001–2002 data, background events contribute significantly to the raw-data volume, and raw-data files occupy the most space. Because of the increased luminosity and favorable background levels predicted for 2004, the volume of the reconstructed data is expected to surpass that of the raw data, and MC events are expected to occupy a considerable volume. The maximum capacity of the current tape library is 324 TB. For 2004, we will need at least an additional 300 TB of tape storage capacity. In addition, the need to maintain a significant fraction of 32 TB of DSTs on disk calls for the recall disk cache to be significantly enlarged.

TABLE II
SIZE OF KLOE DATA SETS, IN TB

	2001–2002	2004 (est.)
Raw-data files	102	100
Datarec files	51	192
MCR files	25	86
DST files	7	32
TOTAL	185	410

VIII. PLANNED UPGRADES

Current plans call for the addition of ten new IBM p630 servers to the offline farm, each with four 1.45-GHz Power4+ processors. This solution would increase the total offline processing power from 110 to 225 B80 CPU equivalents; it would supply the doubling of processing power needed to handle offline production tasks for the 2004 data while increasing to about 75 the number of CPUs available on the offline farm for analysis. We note that the KLOE offline software is portable, and that we were not bound to IBM platforms when planning for the upgrade. After bids were solicited from various suppliers, however, this solution proved to be cost-competitive with alternate solutions based on Linux systems with x86-compatible CPUs, and offered additional convenience in terms of ease of integration with our current offline system. Together with the CPU upgrade, we plan to add about 20 TB of FC disks to the recall cache and AFS cell.

Finally, we will extend our long-term storage capacity. The solution under consideration is the installation of a second IBM 3494 tape library. The new library would use six Magstar 3592 drives, which can write 300 GB to a cartridge in native format, and which can read and write at 40 MB/s [12]. The initial installation would include 1000 cartridges with space for 3600, which would allow the total capacity of the new library to be expanded from an 300 TB to 1.08 PB as needed. The new library will have an FC interface. Since the library must be installed in a separate building 200 m from the KLOE computer center, a storage area network will be deployed to connect the library to the file servers.

IX. CONCLUSIONS

The KLOE computing environment is flexible and scalable. In normal running, the system smoothly serves data to hundreds of contemporary processes related to the storage, reconstruction, and analysis of a data set consisting of nearly 200 TB. Our experience concerning the scalability of the environment leads us to believe that with the extensions described here, the system should easily be able to handle data collection rates of greater than 10 pb^{-1}/day when data taking resumes after the DAΦNE luminosity upgrade.

REFERENCES

[1] M. Adinolfi, F. Ambrosino, A. Andryakov, A. Antonelli, M. Antonelli, C. Bacci, *et al.*, "The tracking detector of the KLOE experiment," *Nucl. Instrum. Meth.*, vol. A488, pp. 51–73, 2002.

[2] M. Adinolfi, F. Ambrosino, A. Antonelli, M. Antonelli, F. Anulli, G. Barbiellini, *et al.*, "The KLOE electromagnetic calorimeter," *Nucl. Instrum. Meth.*, vol. A482, pp. 364–386, 2002.

[3] A. Aloisio, F. Ambrosino, M. Antonelli, C. Bini, V. Bocci, F. Bossi, *et al.*, "Data acquisition and monitoring for the KLOE detector," *Nucl. Instrum. Meth.*, vol. A516, pp. 288–314, 2004.

[4] D. Quarrie and B. Troemel, *YBOS Programmer's Reference Manual*, Fermi National Accelerator Laboratory, 1992, CDF Note 156, version 4.00.

[5] J. Gailly and M. Adler, *ZLIB 1.1.4 Manual*, 2002. [Online]. Available: http://www.gzip.org/zlib/manual.html

[6] M. Adinolfi, A. Aloisio, F. Ambrosino, A. Andryakov, A. Antonelli, M. Antonelli, *et al.*, "KID—KLOE Integrated Dataflow," in *Proceedings of CHEP 2001*, Beijing, Sept. 3–7, 2001, pp. 228–231.

[7] DB2 Product Family. IBM. [Online]. Available: http://www.ibm.com/software/data/db2

[8] IBM Tivoli Storage Manager. IBM. [Online]. Available: http://www.ibm.com/software/tivoli/products/storage-mgr

[9] M. Shapiro and D. Quarrie, *A Beginner's Guide to ANALYSIS_CONTROL and BUILD_JOB*, Fermi National Accelerator Laboratory, 1988, CDF note 384, version 1.01.

[10] *GEANT– Detector Description and Simulation Tool*, European Organization for Nuclear Research, 1995, CERN Program Library Long Writeup W5013, version 3.21.

[11] LoadLeveler V2.2. IBM. [Online]. Available: http://www.ibm.com/servers/eserver/ecatalog/us/software

[12] "IBM TotalStorage Enterprise Tape Drive 3592," G225-6982-00, IBM, 2003. [Online]. Available: http://www.storage.ibm.com/tape/drives/3592