Fermilab

MEMORANDUM OF UNDERSTANDING

Between the

MINOS Experiment and the Computing Division

October 2003

NETWORKING SUPPORT

The LAN at Soudan has been configured and installed by the Fermilab network group. We will require similar help with the Near detector LAN at Fermilab. The equipment and installation costs for the Near detector LAN are listed below. The installation cost is based on the cost for Soudan with $20,000 added. This will be replaced once the real cost is known. The LAN at the near detector will have fiber connected directly to the front-end electronics racks with media converters to avoid any possibility of ground loops. We do not expect to have any special requirements for networking in FCC beyond what is generally available.

Item	Quantity	Cost/Unit	Cost
Cisco 4506	1	3600	3600
Supervisor module	1	9500	9500
Power supply	2	700	1400
128 M Memory	1	600	600
10/100 48 port module	2	3100	6200
10/100 48 port fiber module	1	11500	11500
24 port module for WH12 control room	1	4000	4000
Installation	1	80000	80000
Total			36800

Table I Expected costs of equipment for Near Detector LAN

The strategy for the Soudan Lan support is to incorporate it's operation and management into the Fermilab campus network support effort. The Fermilab policy that defines the campus network as a restricted central service (http://www.fnal.gov/cd/main/cpolicy.pdf) will be extended to include the local network at Soudan. From a practical perspective this will mean:

1. The topology of the Soudan network will be specified by the Fermilab network group and modifications or changes made only in consultation with or at the direction of that group.
2. All active network devices (switches, routers, hubs with manageable components) will be procured and installed under the direction of the Farmilab network group and will be managed remotely by that group.
3. Basic network services (DNS, DHCP) will be provided on servers that will be supplied, installed and remotely managed by the Fermilab network group.

The remote management of the Soudan LAN is predicated on having a simple, easily manageable architecture. This necessitates centralized network connections and a minimum number of active network devices. Unintelligent microhubs may be used to provide extra network connections in situations where all the local network jacks are in use. All active network devices, including unintelligent microhubs, will be of a type, make and model specified or approved by the Fermilab network group.

APPENDIX VII - OFFLINE COMPUTING

INTRODUCTION

This appendix deals with the offline computing needs for the MINOS experiment. The offline needs can be broken into the following areas:

- Data Handling and Storage
- Offline Data Processing
- Monte Carlo Generation
- Offline Analysis
- Software
- Personnel resources

This document will address the resources required in each of the preceding areas. These resources include number of tapes, amount of CPU, disk space and number of people.

OVERVIEW
We would like as much as possible to integrate our computing resources into the existing FNALU/AFS/FARM architecture to reduce the support load. We are very interested in being able to use Terabyte file servers as AFS disk to reduce the cost of making data available over the WAN to collaborators. We expect to use SAM for data handling and if we can convince all groups in the experiment to install it we will be less reliant on AFS as a means of distributing data. We would like to deploy our analysis CPU as batch nodes within FNALU and continue to use the FNALU Linux machines for interactive development. We would expect that our FARM needs would be met by augmenting the existing general farm.

MINOS 5 YEAR RUN PLAN

The experiment has presented a plan to the laboratory management for the number of protons on target per year (pot/year) we would like to receive. Based on this plan the experiment has a strawman run plan for the amount of low, medium and high energy beam that we would like. This is summarized in Table I.

	Protons on target per year x 10^{20}			
Year	Low Energy	Medium Energy	High Energy	Total
2005	1.9	0.4	0.2	2.5
2006	4.0	0	0	4.0
2007	3.5	1.0	0.5	5.0
2008	5.6	0.6	0.3	6.5
2009	7.5	0	0	7.5

Table I MINOS 5 year run plan

This information is used to determine the data rate in the near detector, which is shown in Table II. For the years 2005-2007 we assume that there is one spill every 1.9 seconds. For 2008-2009 we assume that there is one spill every second due to a reduction in the Main Injector cycle time but fewer protons per pulse.

			Events per spill		
Year	POT x 10^{20}	Protons per pulse x 10^{13}	Low	Medium	High
2005	2.5	2.5	25	63	125
2006	4	4	40	100	200
2007	5	5	50	125	250
2008	6.5	3.25	33	81	162
2009	7.5	3.75	38	94	188

Table II Expected number of events per spill for the MINOS run plan

The beam neutrino rate in the far detector has been neglected as it is tiny compared to all other numbers. The rate from the far detector is dominated by 1Hz of cosmic ray μ interactions which will be used both for calibration and for cosmic ray and atmospheric neutrino physics studies. The far detector also has about 12 Hz of noise triggers, which are small events and contribute to the raw data size but are eliminated after reconstruction. There are also pedestal runs, light injection etc which are listed in the “Other” category. The neutrino interaction rate for the near detector from Table II will vary depending on the actual beam being run but for the low energy beam will be from 12-25. About 30% of the events are produced in the calorimeter section and the remaining 70% in the spectrometer section, which is only read out every 5 planes and has 4-way multiplexing. The near detector DAQ system is capable of recording the full 250 Hz of cosmic rays μ seen by the near detector but it is expected that we will record only a fraction of these for full reconstruction, we have assumed 11 Hz, and this is reflected in the numbers in Table IV and V. The far detector assumes 3×10^7 seconds in one year (cosmic rays are always there) and the near detector assumes an effective year of 2×10^7 seconds for beam and 3×10^7 seconds for the cosmic rays. For simplicity we assume that 1Kbyte ≡ 1000 bytes.

Sample	Rate/second (Hz)	Events/year	Raw Even Size (Kbytes)	Data Volume /year (GB)
Cosmic ray μ	1	3×10^7	1.1	33
Noise	12	2.6×10^8	0.2	72
Other				312
Total				417

Table III Event rates and raw data volumes for the Far detector

		Events per year ($\times10^8$)				
Sample	Event Size (Kbytes)	2005	2005	2006	2007	2008
_ (calorimeter section)	0.6	1.17	1.2	2.55	2.64	2.28

_ (spectrometer section)	0.03	2.73	2.8	5.95	6.16	5.32
Cosmic Rays	0.6	3.3	3.3	3.3	3.3	3.3
Total		7.2	7.3	11.8	12.1	10.9

Table IV Event rates for the Near detector

		Raw Data Volume per year (GB)				
Sample	Event Size (Kbytes)	2005	2006	2007	2008	2009
_ (calorimeter section)	0.6	70	72	153	160	137
_ (spectrometer section)	0.03	8	8	18	18	16
Cosmic Rays	0.6	200	200	200	200	200
Total		278	280	371	378	353

Table V Raw Data volumes for the Near detector

The raw data is fairly compressed it expands after event reconstruction when we add the de-multiplexed hits, tracking information and calibration. The current expansion rate is a factor of 44. We need to study this to see if it can be reduced but we have used this number in the planning that follows. Tables VI and VII show the expected reconstructed data volumes for the near and far detectors. We currently write a file of Candidates, an ntuple version and a compressed ntuple. These are included in the calculations. We have used the same event size for cosmic ray events. This is probably an overestimate.

			Reconstructed Data Volume per year (GB)				
Sample	Data type	Event Size (Kbytes)	2005	2006	2007	2008	2009
_ (calorimeter section)	Candidate	29	3400	3500	7400	7600	6600
	Ntuple	5.8	680	700	1500	1500	1300
	Comp. Ntuple	0.8	100	100	210	210	150
_ (spectrometer section	Candidate	1.4	380	390	830	860	740
	Ntuple	0.3	80	80	170	170	150
	Comp. Ntuple	0.04	11	11	25	25	22
Cosmic Rays	Candidate	29	8800	8800	8800	8800	8800
	Ntuple	5.8	40	40	40	40	40
	Comp. Ntuple	0.8	6	6	6	6	6
Total			13500	13600	19000	19200	17800

Table VI Reconstructed data volumes for the Near detector

Sample	Data Type	Event Size (Kbytes)	Reconstructed Data volume per year (GB)

Cosmic Rays	Candidate	44	1300
	Ntuple	9	270
	Comp. Ntuple	1.4	40
Total			1600

Table VII Reconstructed data volumes for the Far detector

DATA HANDLING AND STORAGE

The total raw + reconstructed data volume per year varies from 1.6 TB in 2004 to 21TB in 2008. This data will be stored in the STKen tape robot. The ntuple-level data will be stored permanently on disk, between 1.2 TB (2005) to 2.3 TB (2008) per year. We would plan for DCache disk to hold stage about 20% of the Candidate data at any one time, 3-4 TB per year.

The default mode of operation is to record data on disk locally at the detectors and transfer it over the network to Fermilab/FCC (Feynman Computing Center) into DCache for archiving to tape. We will probably also desire to make several copies of the raw data as the volume is small and tapes are large, meaning a tape failure could result in significant data loss. We are using Enstore to do the tape archiving.

EVENT RECONSTRUCTION

The event reconstruction for both the Near and Far detectors will be done at Fermilab. A summary of the processing needs is given in Table VIII for steady state – keep up with the data and in Table IX for reprocessing. These numbers are based on the performance of the existing MINOS C++ reconstruction code but reduced by a factor of two for expected performance gains. The Far detector numbers are taken from real data, the Near from Monte Carlo. The processing time per event will be given in GHz-seconds per event and the CPU requirements will be given in GHz. We have assumed that in the years 2004-2006 we will do 2 complete reprocessing passes of the data per year that will be completed in 3 months each. In 2007 we assume 1.5 passes and in 2008-2009 we assume 1.2 passes. We assume a farm efficiency of 70%.

		GHz per year					
	GHz-sec/event	2004	2005	2006	2007	2008	2009
_ (calorimeter section)	1.6		24	25	54	57	50
_ (spectrometer section)	0.65		24	24	53	55	48
Cosmic Rays (Near)	1.6		49	49	49	49	49
Cosmic Rays (Far)	8	23	23	23	23	23	23

Total		23	120	121	179	184	170

Table VIII Steady state event reconstruction needs for Near and Far detectors

	GHz per 3 months					
	2004	2005	2006	2007	2008	2009
Number of passes	2	2	2	1.5	1.2	1.2
_ (calorimeter section)		70	73	155	80	28
_ (spectrometer section)		75	70	37	38	33
Cosmic Rays (Near)		206	206	206	103	41
Cosmic Rays (Far)	93	93	93	47	19	19
Total	93	442	442	445	240	121

Table IX Reprocessing needs for Near and Far detectors in a 3 month period

MONTE CARLO GENERATION AND STORAGE

There are two types of Monte Carlo required for MINOS, simulation of neutrino interactions in the detector for oscillation measurements/conventional neutrino physics and simulation of the neutrino beam to understand features of the beam such as beam profiles, flux etc. In both cases the requirements are not precisely known so the numbers here are based on assumptions. We assume here that we will generate the samples at Fermilab but the possibility may also exist to generate them at collaborating institutions and ship them to Fermilab for storage, as is done by CDF and D0.

Physics Monte Carlo

For studies of cosmic ray and atmospheric neutrino events in the Far detector we assume a factor of 10 more Monte Carlo than data, namely 3×10^8 events. For the Near detector we have made a similar assumption but only considered the events produced in the calorimeter section. The execution time for the simulation is dominated by the event reconstruction time. The event size is larger due to storage of the "truth" information" for the event. It may be possible to reduce the event size but we have not studied this yet. The needs per year are summarized in Table X. We have not included any reprocessing requirements for the Monte Carlo YET.

	Year					
	2004	2005	2006	2007	2008	2009
Far detector						
Events ($\times 10^{8}$)	3	3	3	3	3	3
Raw Data (TB) (27 Kbytes/event)	8	8	8	8	8	8
Candidates (TB) (70 Kbytes/event)	21	21	21	21	21	21
Ntuple (TB) (15 Kbytes/event)	4.5	4.5	4.5	4.5	4.5	4.5
Comp. Ntuple (TB) (2 Kbytes/event)	0.6	0.6	0.6	0.6	0.6	0.6
CPU (GHz)	145	145	145	145	145	145
Near detector						
Events ($\times 10^{9}$)		1.17	1.2	2.55	2.64	2.28
Raw Data (TB) (10 Kbytes/event)		12	12	25	26	23
Candidates (TB) (39 Kbytes/event)		47	47	98	101	90
Ntuple (TB) (8 Kbytes/event)		9.6	9.6	20	20	21
Comp. Ntuple (TB) (1 Kbytes/event)		1.2	1.2	2.5	2.6	2.3
CPU (GHz)		174	176	376	391	336

Table X Physics Monte Carlo needs per year

This data will also need to be stored in the STKen tape library. We would like to be able to store the Ntuples on disk, this may require reducing the evnt size.

<u>Beam Monte Carlo</u>

These numbers are 3 years old and need to be updated.
Currently a single run takes about 2.4×10^{5} GHz-seconds to obtain sufficient statistics. The data volume produced per run is about 240 Mbytes. It is expected that a factor of 10 times longer runs will be needed for the physics analysis and that a few hundred of these runs will be required. We currently assume that the runs are generated overt the course of a year. This may be invalid meaning that the GHz per year would increase. The needs are summarized in Table XI.

GHz-sec/run	2.4×10^{6}
Runs/year	200
GHz/year	15
Storage/year	480 GB

Table XI Beam Monte Carlo needs per year

OFFLINE ANALYSIS

We plan to perform physics analysis of the data at Fermilab as well as at other MINOS collaborating institutions. This area of the needs is the most uncertain. We have made some assumptions that we outline here. We assume that there will be of order 30 users doing analysis on a central MINOS facility located at FNAL. Most analyses will proceed from the Ntuples. We assume that most users are using the Ntuples produced on the farm but that a few, of order five might need to produce their own Ntuples from the Candidate files.

Ntuple analysis

We assume 30 users and each makes 6 passes through the Ntuples to create a compressed ntuple. It takes 11000 GHz-seconds to produce a compressed Ntuple from a year of Far detector data. We assume that a user should be able to produce a compressed Ntuple from the data in one day and from the Monte Carlo in one week. This is 25 GHz of CPU for data and 36 GHz of CPU for the Monte Carlo per year. For the Near detector we assume that users take 1 month to create a compressed Ntuple from the data and the Monte Carlo. For the data we include making Ntuple from the Calorimeter section only.

	CPU in GHz				
Year	2005	2006	2007	2008	2009
Data	3	3	4	7	6
Monte Carlo	30	30	40	70	60

Table XII Analysis CPU for the Near detector, Ntuple processing.

Candidate analysis

We assume 5 users each making 2 passes per year through the Candidate data to create Ntuples. It takes 0.4 GHz-seconds/event to create a Ntuple. We assume that in the Far detector this task is accomplished in 1 month for the data and 3 months for the Monte Carlo. This takes 70 GHz of CPU per year for data and 230 GHz for the Monte Carlo. For the Near detector we assume that the data can be processed in 3 months but that only one pass of the Monte Carlo will be done per year.

	CPU in GHz				
Year	2005	2006	2007	2008	2009
Data	62	62	136	135	117
Monte Carlo	300	300	650	677	585

Table XIII Analysis CPU for the Near detector, Ntuple creation from Candidates.

In addition to CPU we also need disk for users to be able to store their created Ntuples. We will arbitrarily assume that each user will be allocated 100 GB of space per year. This will be 3 TB per year.

DATABASE

The database warehouse will be kept in Oracle and will initially reside on a SUN system. We will also need a development machine. Should Linux systems become supported for production

Oracle databases we would migrate to that platform. We plan to have Linux MySQL servers at Soudan, the Near detector and as replicas for the offline analysis and the production farm. They will also be used at collaborating institutions. We will require access to some fraction of the Oracle client licenses, the number will be minimized by the use of the MySQL replicas. The size of the database is expected to be about 500 GB over 5 years. The expected initial cost of the system is about $125,000.

SUMMARY

Year	2004	2005	2006	2007	2008	2009
Tape (TB)	36	120	120	201	206	191
Data Ntuple disk (TB)	0.3	1.3	1.3	2.2	2.2	2.0
MC Ntuple disk (TB)	5	16	16	28	28	23
User disk (TB)	3	3	3	3	3	3
DCache data (TB)	0.3	3.1	3.1	3.7	3.7	3.5
DCache MC (TB)	4.2	14	14	24	24	22
CPU, steady state (GHz)	23	120	121	179	184	170
CPU, reproc (GHz)	93	442	442	445	240	121
CPU, MC (GHz)	145	334	336	536	551	496
CPU, analysis (GHz)	361	691	691	1051	1108	1006
CPU Speed (GHz)	3.5	5.6	8.8	?	?	?
Duals, reco	3	11	8			
Duals, reproc	13	39	25			
Duals, MC	21	30	30			
Total Duals, FARM	37	80	62			
Duals, analysis	32	62	39			

Table IXX MINOS needs for tape, disk and CPU for 6 years of data taking.

Table XX shows the costs associated with the needs in Table IXX. The following assumptions have been made. A dual CPU farm node costs $2200 independent of CPU speed. A 4 TB file server currently costs $14,400. We will assume that the cost remains roughly the same but the capacity increases by a factor of 1.6 each year. We assume that in 2004 the tape cartridge size is 200 GB and costs $75. In 2005 and beyond we assume that the capacity is 400 GB and costs the same.

Year	2004	2005	2006	2007	2008	2009

Tape	13500	22500	22500	38000	38600	35800
Static Disk	30000	45600	29200	29900	18400	14400
DCache disk	14400	38500	24600	24900	14400	14400
Total Duals, FARM	46200	66000	66000			
Duals, analysis	70400	136400	85800			
Total	174500	309000	228100			

Table XX MINOS costs for 6 years of data taking