Proposal

+++

BCFpc: Batch Computing Facility based on PCs

****

Corrections:

August 20,1997

November 3, 1997

November 28, 1997

February 12, 1998

Andrei E. Chevel

shevel@pnpi.spb.ru

Petersburg Nuclear Physics Institute

188350, Gatchina, Leningrad district,

Russia

This version preparation date: February 12, 1998

Abstract

This proposal is devoted to realization and practical use a General Purpose Batch Computing Facility based on a PC cluster (farm) for High Energy Physics research.

For up to date information please see "http://hepd.pnpi.spb.ru/pcfarm/".

Problem: Needs for high performance computing for a low cost per unit of power.

Introduction

Several considerable facts might be taken into account:

- decreasing the financial resources in HEP;
- increasing the requirements for computing;
- existing of a cheap operating platform like Linux;
- cheap commodity hardware (PC);
- existing batch systems CODINE, NQS, LSF, and others which are relatively easy available in many leading HEP laboratories;
- almost all important applications in HEP might be running on Linux;
- new languages: Java;

We might consider General Purpose Computing Facility on a base of PCs. There are a range of such successfully realized projects [please see our proposal www pages "http://hepd.pnpi.spb.ru/pcfarm"].

It is discussed PC farm consisting of N nodes (PCs) for batch processing with relatively weak communications between nodes. N might be considered in range from several PCs to probably 100 or more. There is already known computing system based on 9K (!) processors PentiumPro200 (Intel).

Reliability of an existing hardware is enough to support PC farm consisting of 100 nodes by one person (for hardware, operating system and general application software).

Scalability: we can increase a number of PCs to desired value in step by step mode by small quanta.

The value about **400 Pentium Pro 200 MHz or 200 Pentium II 300 MHz** can be discussed as really achievable maximum of computing power (in case when four or two processor motherboards will be used accordingly).

A number of SCSI ports and disk drives may be various and will reflect concrete needs.

Technical consideration

We can consider large installation like a constellation of simple farms (see Fig. 2).

However before realization of the large farm we should create pilot relatively small experimental Batch Computing Farm (see Fig. 1).

There are several technical questions which may be mentioned now:

- how to maintain and load one OS version in every computer;
- how to use one (or two) keyboard and monitor for a hundred of computers;
- how to display the status of the whole large farm in compact and consistent manner.

It is possible to continue this list however those technical details will be solved later because good part of them are already realized in a range of previous projects (FNAL: computing farm, DESY: Hermes PC Farm, etc.). We also have to take into account that new equipment, software, prices are appeared every month.

Real Experimental Batch Farm

To test scalability and other parameters discussed proposal it required to deploy small Batch PC Farm consisting of at least 3 PCs.

Main aim: to get real experience in deploying the batch system (CODINE, LSF, etc.) on small PC Farm and to collect user feedback, measurement the real farm parameters: comparing the execution time for selected jobs. It is planned such a farm will be used for simulation and microDST analysis.

In any case this pilot project will show for which type of calculations this farm will fit more.

It has been deploied experimental installation:

- two PCs:
  - Pentium II 266 MHz;
  - 512K cache;
  - 128MB main memory;
  - 5.2 GB EIDE disk drive;
  - video card (on mother board);
  - no monitor;
  - Fast Ethernet card;
- Hub Fast Ethernet: 3COM Super-Stack II 100 TX (12 port RJ-45);
- APC Smart-UPS 2000;
- Linux RedHat distribution;
- batch systems (CODINE, LSF);
- CERNlib (PAW and ROOT are included);
- central PC:
  - Pentium II 266 MHz;
  - 512K cache;
  - 256MB main memory;
  - disk drive 5.2 GB;
  - two Fast Ethernet card;
  - CD ROM;

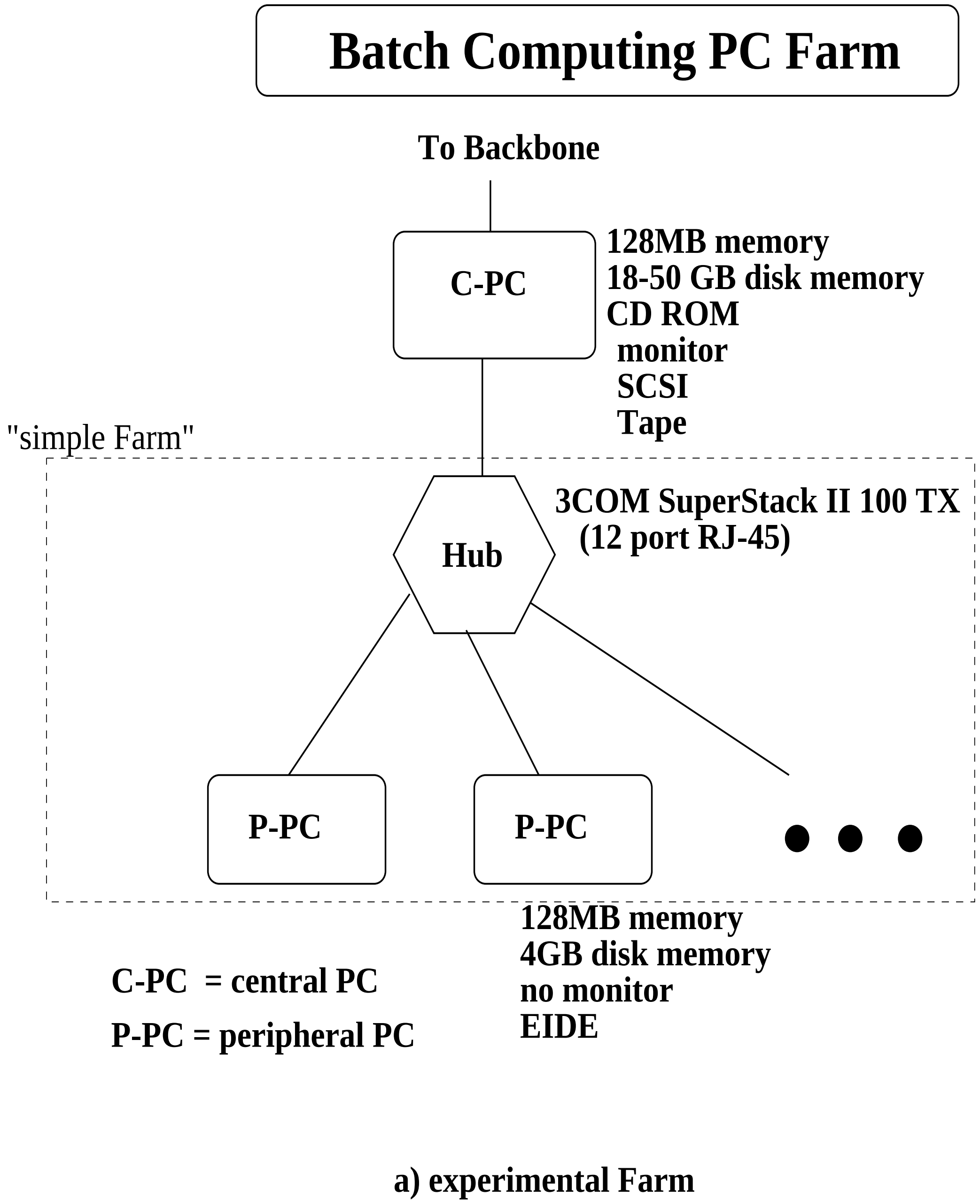
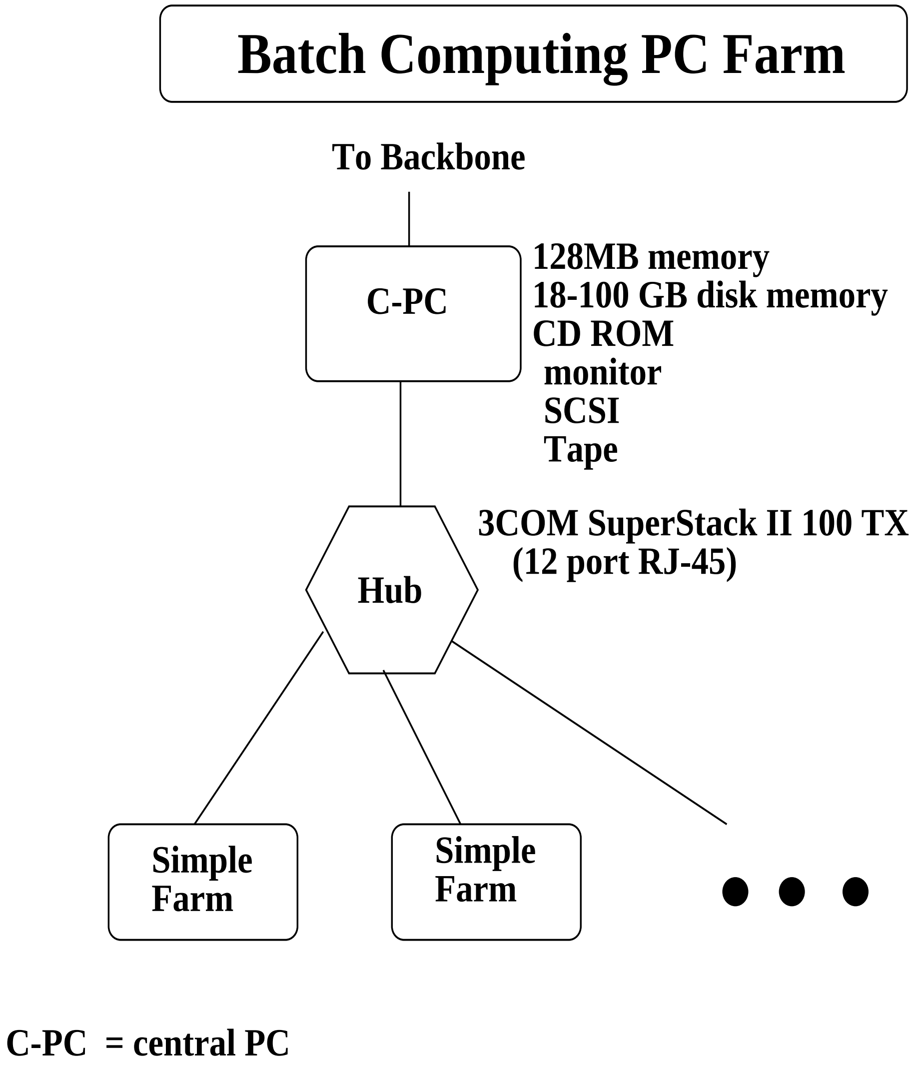


Fig. 1. Experimental Batch Computing PC Farm.



b) Batch Computing Farm

Fig. 2. A Scheme of the Batch Computing PC Farm.

- monitor 15";

Things which should be done yet (12th Febryary 1998):

- hardware installation:
  - disk drive subsystem (30-100 GB);
  - tape drive DLT 7000;
  - special cabinet for the PC farm.
- software installation and tuning:
  - batch software (agreement on license);
- measurement of execution time for the selected jobs and other parameters;
- organizing the user access (over the Internet as well).

Schedule:

Week	Job
0	Project start
4	Hardware delivery at PNPI
6	Linux + CERNlib + other programs, libraries, utilities
8	Start User access
10	Report paper about first results

Remark. The proposal was started 10th January 1998.

Spending:

What	Who
Manpower	PNPI
space to deploy	PNPI
equipment	PNPI + ?
batch system	GENIAS ?

How to participate in the project

Introduction

In order to increase the value of our proposal we try to use collaborating efforts in project realization.

It is clear that Batch Computing Facility might be used over the Internet as well. That means any laboratory can buy one or more PCs to add to current BCFpc installation and to obtain a range of user accounts to use whole BCFpc power in batch mode.

It is clear that BCFpc shall have standard program library configuration for HEP (i mean CERNlib and related software).

An official report on BCFpc usage will be available on regular base (once a month or week).

Why to collaborate

If somebody plans to collaborate in our proposal his laboratory will save a reasonable man power to maintain the computing facility (hardware, application software, and administrative efforts as well).

How to be registered on the Batch Computing Farm

If you feel you need to use Batch Computing Farm for your computing you are required to be registered. To do that please fill the registration form and send it by fax (+7-81271-360-40). The pattern of the registration form can be found out in ”http://dbserv.pnpi.spb.ru/ volkova/”.

How to pay for a PC (PCs) or other components

To avoid transportation and probably custom problems we propose to use special procedure. In contact with our partners we find out a PC and after send the offer (invoice) to our partner. When the offer (invoice) will be paid we are able to receive the PC and add it to our Batch Computing Farm (BCFpc).

More or less standard configuration for PC we discuss is:

- Pentium II (fastest);
- 256 MB main memory;
- 100 Mbit Ethernet;
- EIDE disk drive (5-10 GB);
- NO MONITOR - NO KEYBOARD - NO MOUSE .

In total less $3K (December 1997).

It might be discussed any other (more advanced) PC configuration and prices.