Enabling Volunteer Computing for the BESIII Simulation System

IHEP

Wenxiao Kan[1], Wenjing Wu[1], Andrei Tsaregorodtsev[2], Gang Chen[1]

[1] IHEP, 19B Yuquan Road, Beijing, 100049 China

[2]CPPM, 163 Avenue de Luminy, Marseille, 13288 France

CPPM

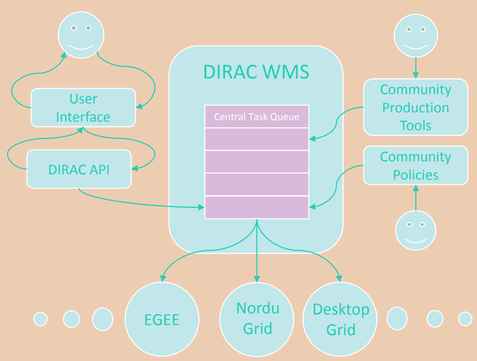
Introduction

Project objective: build a distributed computing system by integrating the idle CPU time of all kinds of heterogeneous desktop computer resources with virtualization technology and deploy high-energy physics data processing on it.

DIRAC

The DIRAC architecture consists of numerous cooperating Distributed Services and Light Agents built within the same DISET framework following the Grid security standards.

- ***Cover all the possible resources***
- ***Loose coupled services***
- ***Light Agents***
- ***A variety of Interface***



BOINC

Berkeley Open Infrastructure for Network Computing

---Developed at UCB Space Science Laboratory by the SETI@home group

Reduce the barriers of entry to public-resource computing:

A project can be run from a single computer running standard open-source software

Share resources among autonomous projects:

Each PC owner can join multiple projects

Results in better resource utilization

Support diverse applications:

Offer various data distribution mechanisms

Support various programming languages

…

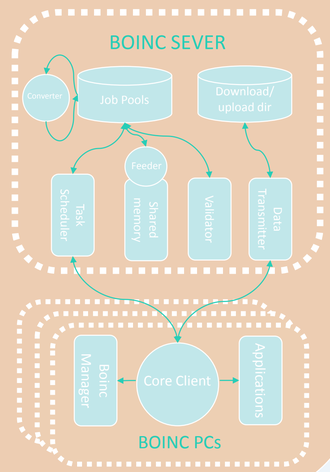
Reward participants:

Mostly by giving them ***credits***

---System must be cheating-resistant

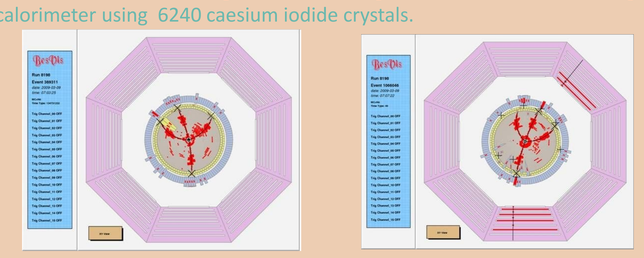
Also by offering nice graphics

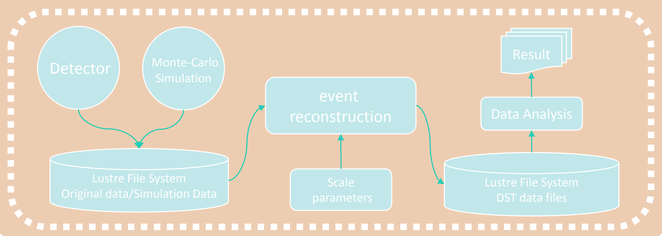
---Great screensavers!



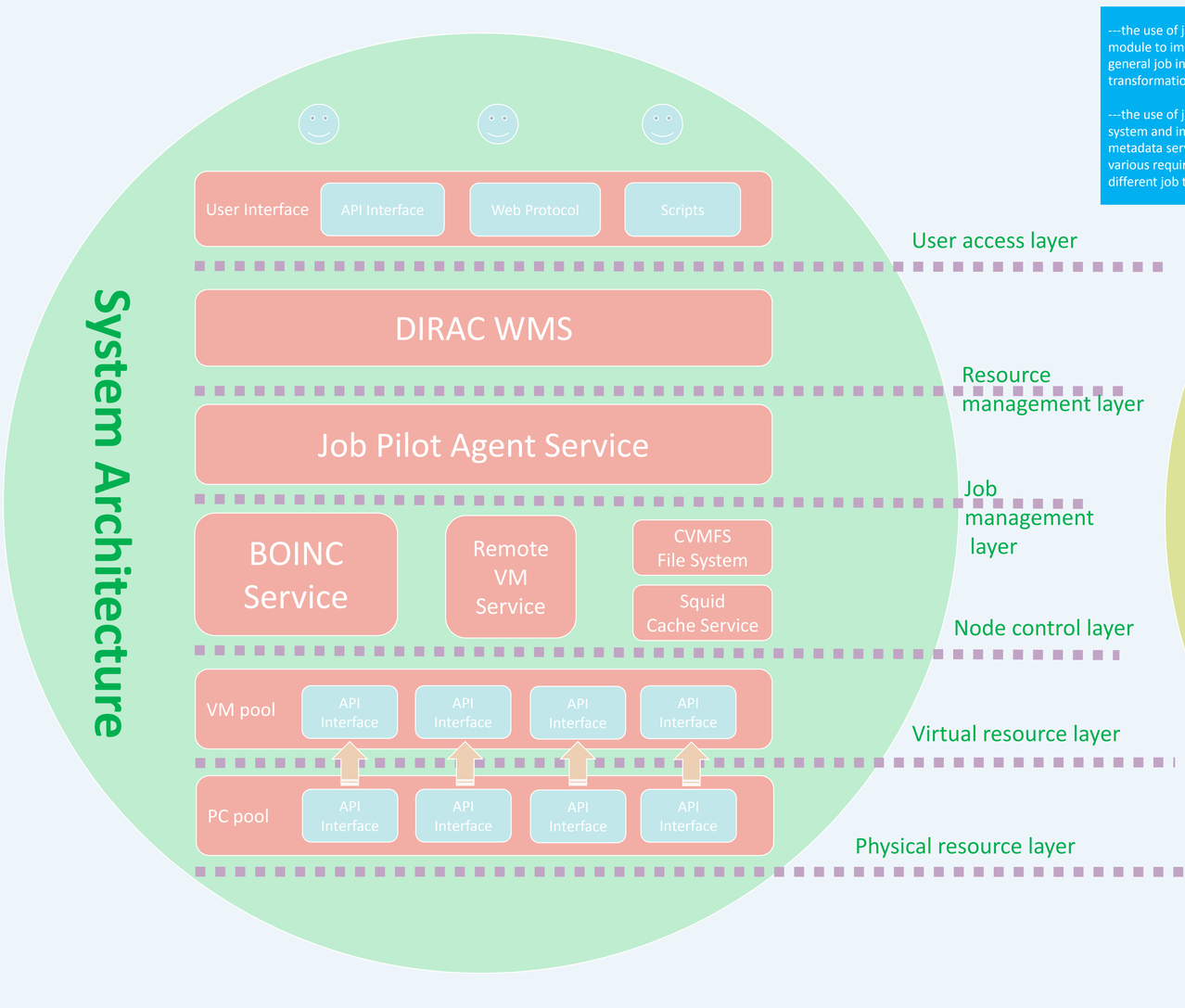
BESIII

BES III uses a large superconducting solenoid to provide a 1-tesla magnetic field, and also features ahelium gas-based tracking chamber and an electromagnetic calorimeter using 6240 caesium iodide crystals.



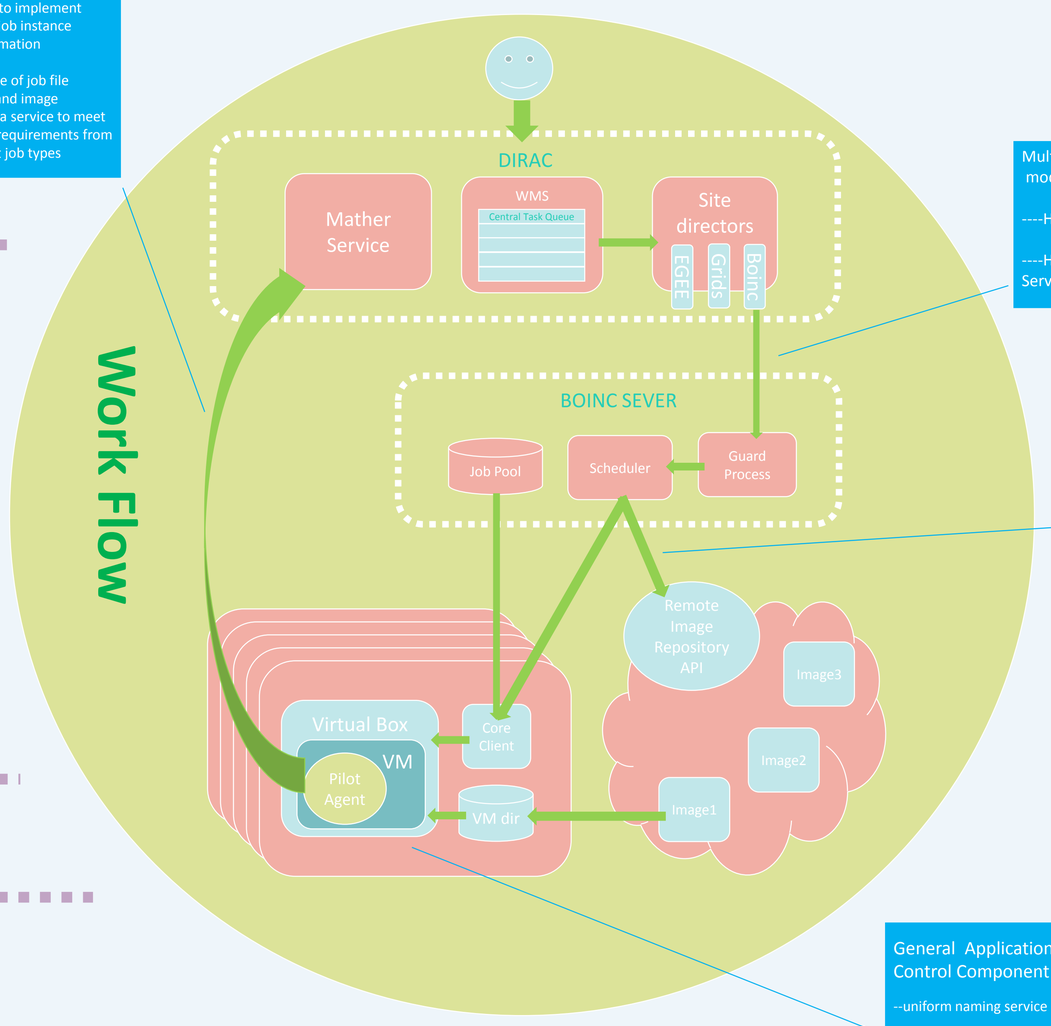


Design and Implementation



---the use of job agency module to implement general job instance transformation

---the use of job file system and image metadata service to meet various requirements from different job types



Multiple communication modes:

----HTTP + SOAP

----HTTPS + Restful Web Service

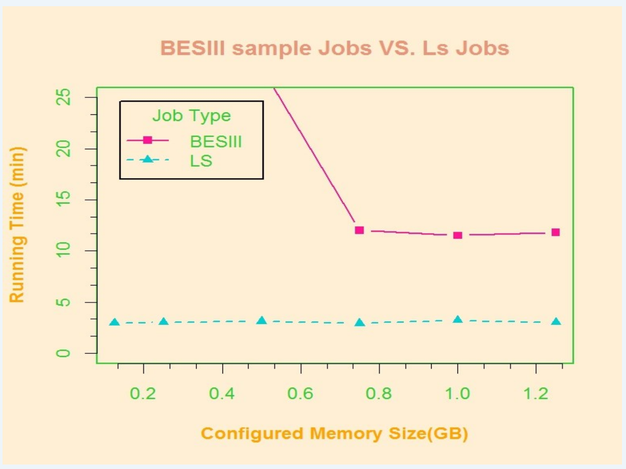
Image API development work mainly base on libcloud API which is provided by Stratus Lab, and for our functionality, we can add parameter for image mark(like the image identity in markentplace) when we submit a job in DIRAC, then the boinc server side accept the job requirement, it will use the marketplaceID to parse the image metadata and get the related information of a image, and then the client can download the image use the information, beside, we also can use the local image.

General Application Control Component :

--uniform naming service

-- shared folder service

Testing and optimization



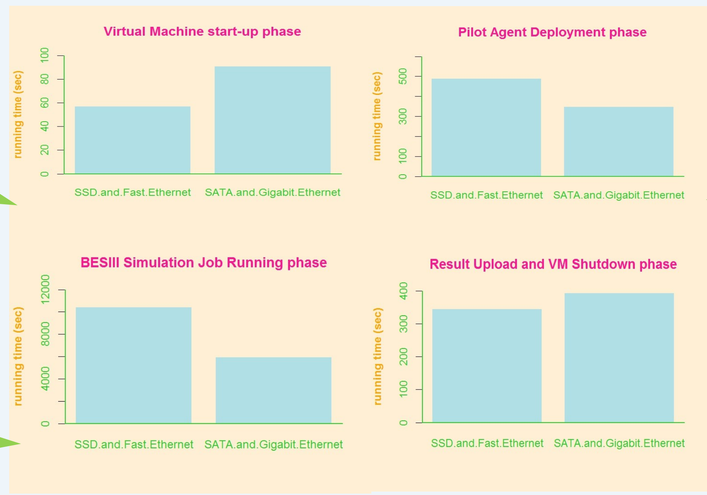
Analysis on different allocated memory size

Two questions:

1. Allocated memory size is an important factor which involved with job running efficiency;
2. Find the best memory size which can't reduce job running efficiency, and also doesn't waste resources;

According to the test result, when the job running efficiency decreases significantly, the critical value is the best memory size we need.

A adaptive memory allocation strategy is implemented by using the test result, when a new application is deploy in the system, a range of jobs with increased memory size will be run locally, then compare the running time, figure out the best configured memory size of this kind of jobs.



Obviously, the hard disk IO performance is the most important influence factor in the virtual machine start-up stage. Why?

In this phase, Boinc client needs copy the vm image to the work slot and start-up…

Obviously, the network bandwidth performance is the key factor in this phase. Why?

In this phase, Pilot Agent will talk to DIRAC Matcher service, and pull the real job, then deploy job running environment from CVFS..

Um.. Why?

BESIII simulation job is running in Virtual Machine now, then it will cost a lot of time to cache the input files, job environment…

Also, the most important, 90% of the whole job running time is on this phase.. We can conclude that the bandwidth is the most important…

It's Ok, because we have concluded that network bandwidth performance is the most influence factor…

Influence factors analysis of a simulation job life cycle on PCs