

“Virtualisation and Parallelisation for using Opportunistic Idle CPU Resources”

MD. Mamunur Rashid
(PhD Student ,LHCb)

Optimistic..... Sure you need more hands



Task > Available Resources

Not Practical

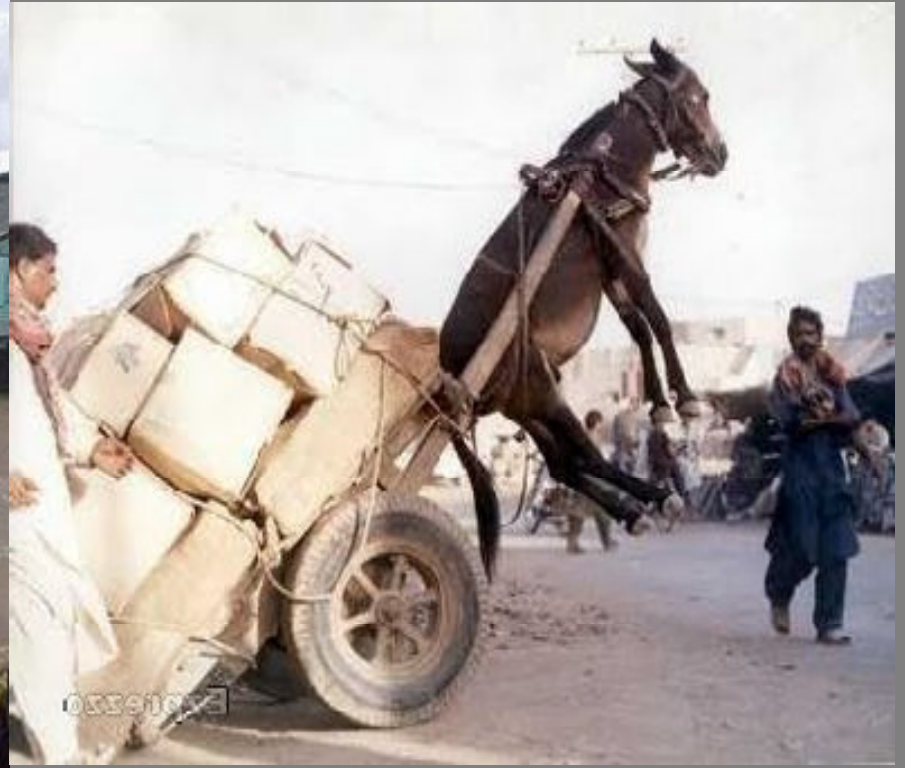


Too practical



What can be consequence!

Consequence can be



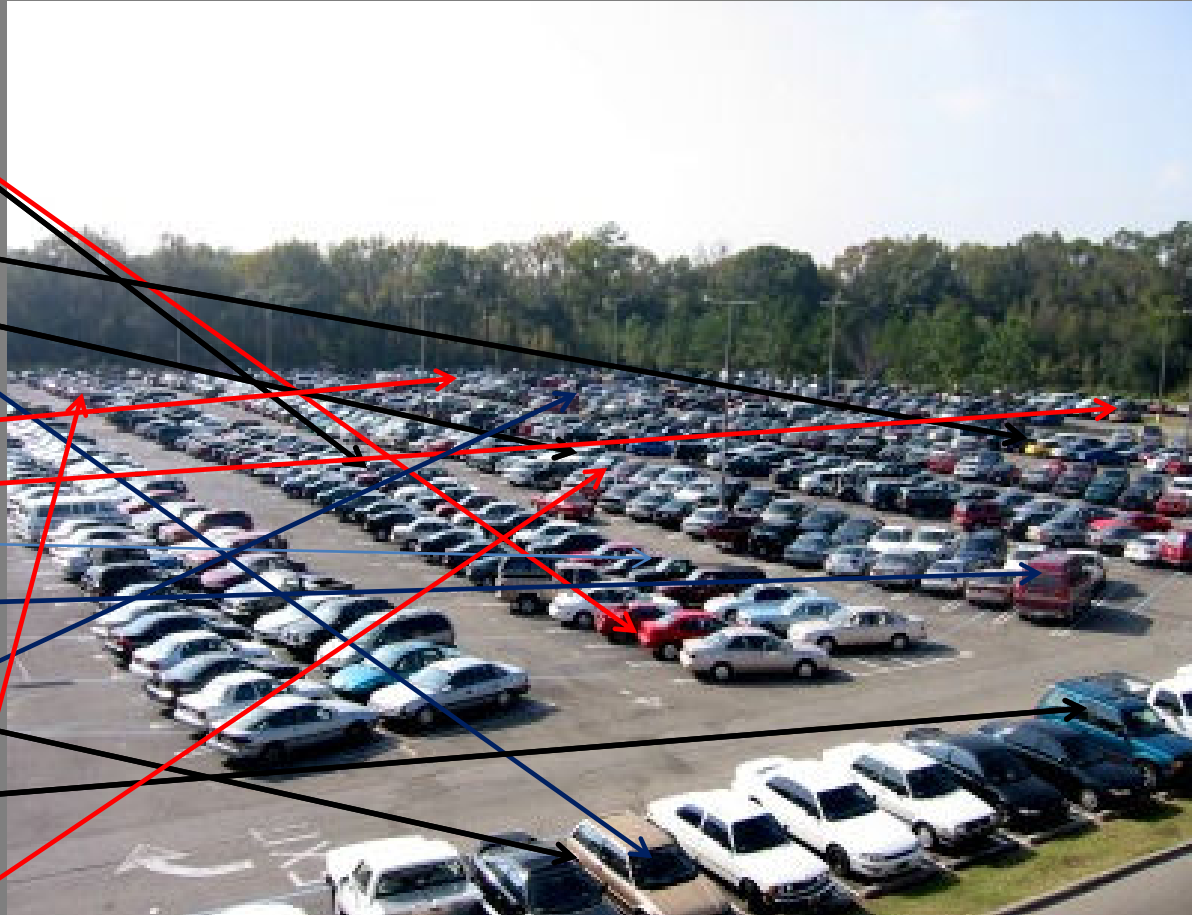
We are missing something from all these pictures

*So what we missing ,is a harmony, a synchronization of using
available resources*

We have these

Parking Catalogue

- Parked on Friday
- Parked on 2/2/2008
- Parked on Monday
- Parked 2/4/2009
- Parked on Friday
- Parked on.....
- Parked o.....
- Parked
- Paarke.....
- Park.....
- Par.....
- P.....
-
- Parkedn



Proposal: Please don't live your car here
let some one use it you are not using.

Why this approach is failed:

Approach is not Practical :

Car is very costly

Potential Risk of Damage

People have a habit of thinking it as very personal belongings

*So we had a prototype to solve a problem and it fails
lets See if we can implement it any where else as its a
promising approach*

Problem

“We need to find more computational resources for accomplish our physics experiment”

1. On condition we don't want to spend all our money to buy all the CPU.
2. What ever we buy today may be not enough for tomorrow.
3. So we need some thing for free or preferably cheap and unlimited

we can't use these machines , they are occupied



Lets try some where else

What about these ones ..



No occupied + Available + Idle

What about these ones.....



Or more interestingly these ones.....



Occupied but not used May be you gave up. but your machine is able to get some thing to do as they are not as lazy as we are. Lets called them Idle Machine

Who are already using this idle resources

*SETI@home and some other projects using BOINC
And they are :*

Active: 329,830 volunteers
574,820 computers.

24-hour average: 2,623.34 TeraFLOPS.

Our contribution/Approach in this research field :

“Virtualisation and Parallelisation for using opportunistic idle CPU resources”

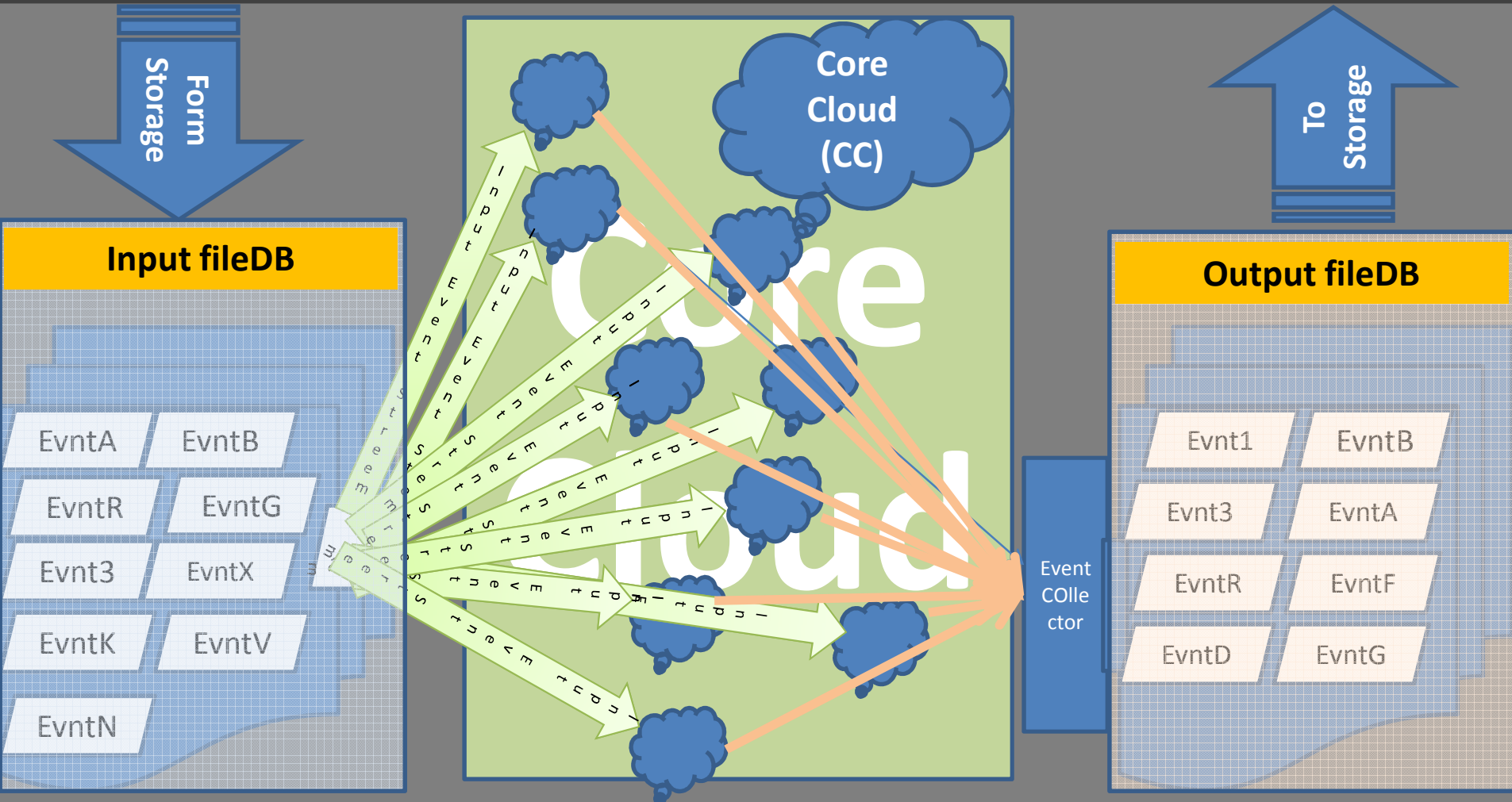
What we have for Parallelization :

We try to keep all the available cores busy to ensure full parallelization in instruction stream.

Implementation level will execute event/core:

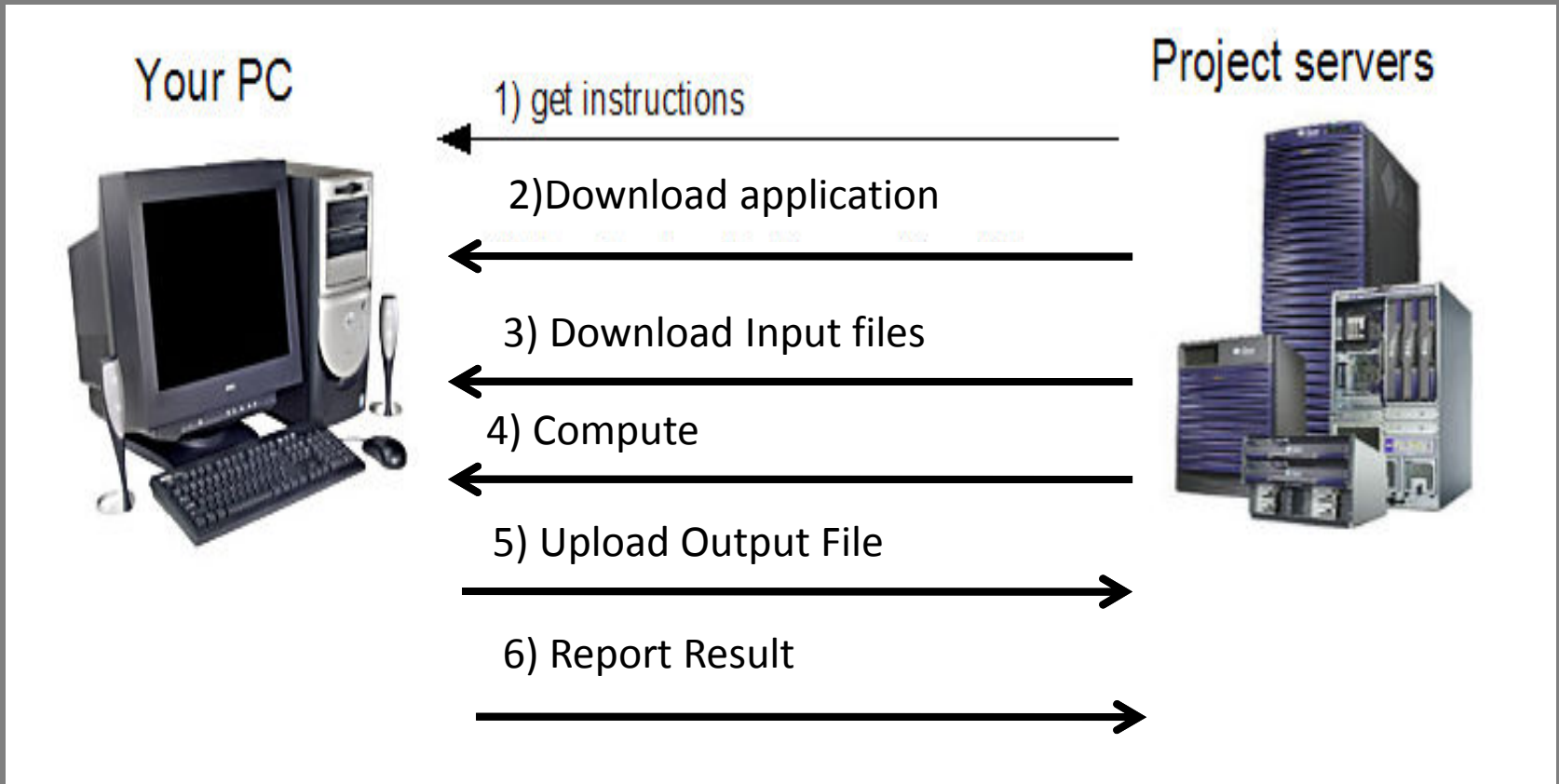
- Events size of 30-35 k
- Recorded in 2 GB files
- The first stage reconstruction
- Takes 1-2s per event

Basic Flow



Basic structure of BOINC

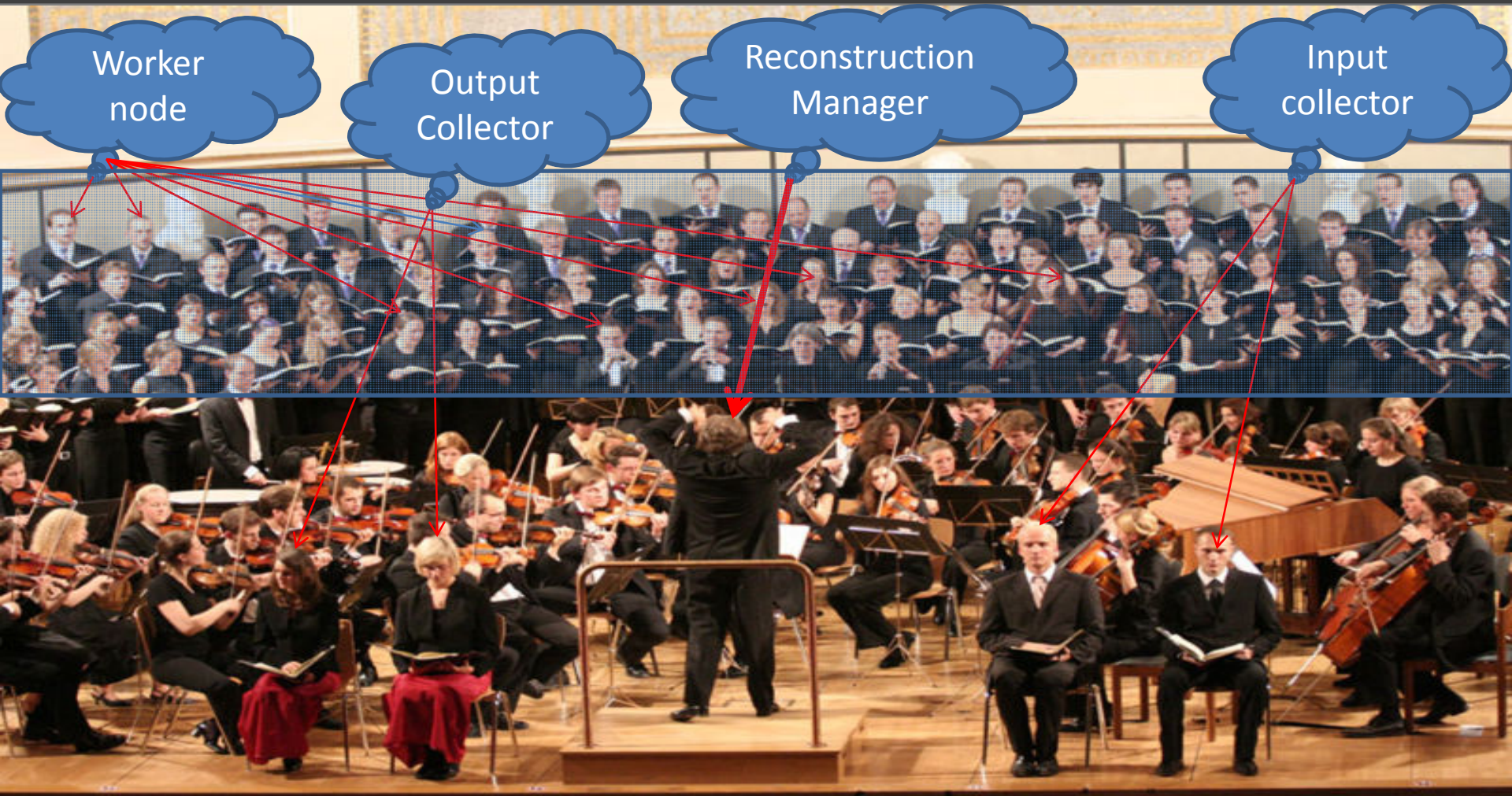
Interaction between client and server



Technical/Scientific benefit of CernVM:

1. It will ensure the amount of exact acquired resources in the remote machine. (Soon Virtual machine is installed we know the system configuration(ram/hard drive size) and how much computation it can perform.)
2. It will facilitate access to use heterogeneous resources (MAC/LINUX/Windows..etc)
3. It will give a freedom to the volunteer user from installing every update of any available software by encapsulating operating system environment all in one (CernVM).
4. The virtual machine will come up with only strictly required software to minimise the overload of package size.(We can say it ThinVM)

Where we want to be!



Where all the available resource is being used and it would not effect us much if any of them is not performing as it will be a large collaborated resource sharing

Where we can exercise this method

1. In Our online farm

2. Machine with multiple cores

3. Over Amazon/Google cloud @£/\$/€(with CernVM)

*4. Over the internet for opportunistic idle resource hunting
using BOINC and CernVM*

Conclusion :

So to run this project we have :

DIRAC

Single machine multiple core access mechanism

CernVM

BOINC (As an option)

We can start with simulation

So all we have to do integrate all these components together and monitor its performance and always try to find a better means to improve its performance .

Is it all free!

Many Thanks for Attending this Session :

Question /Comments?

My Comment : "lets Science Use it, when you don't"