PROOF at the NAF

Wolf Behrenhoff, Hartmut Stadie
Universität Hamburg

2nd Annual Workshop of the Helmholtz Alliance
"Physics at the Terascale"
November, 27th 2008

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Analyses at the NAF

What is special about the NAF?

Differences between NAF and Grid clusters:

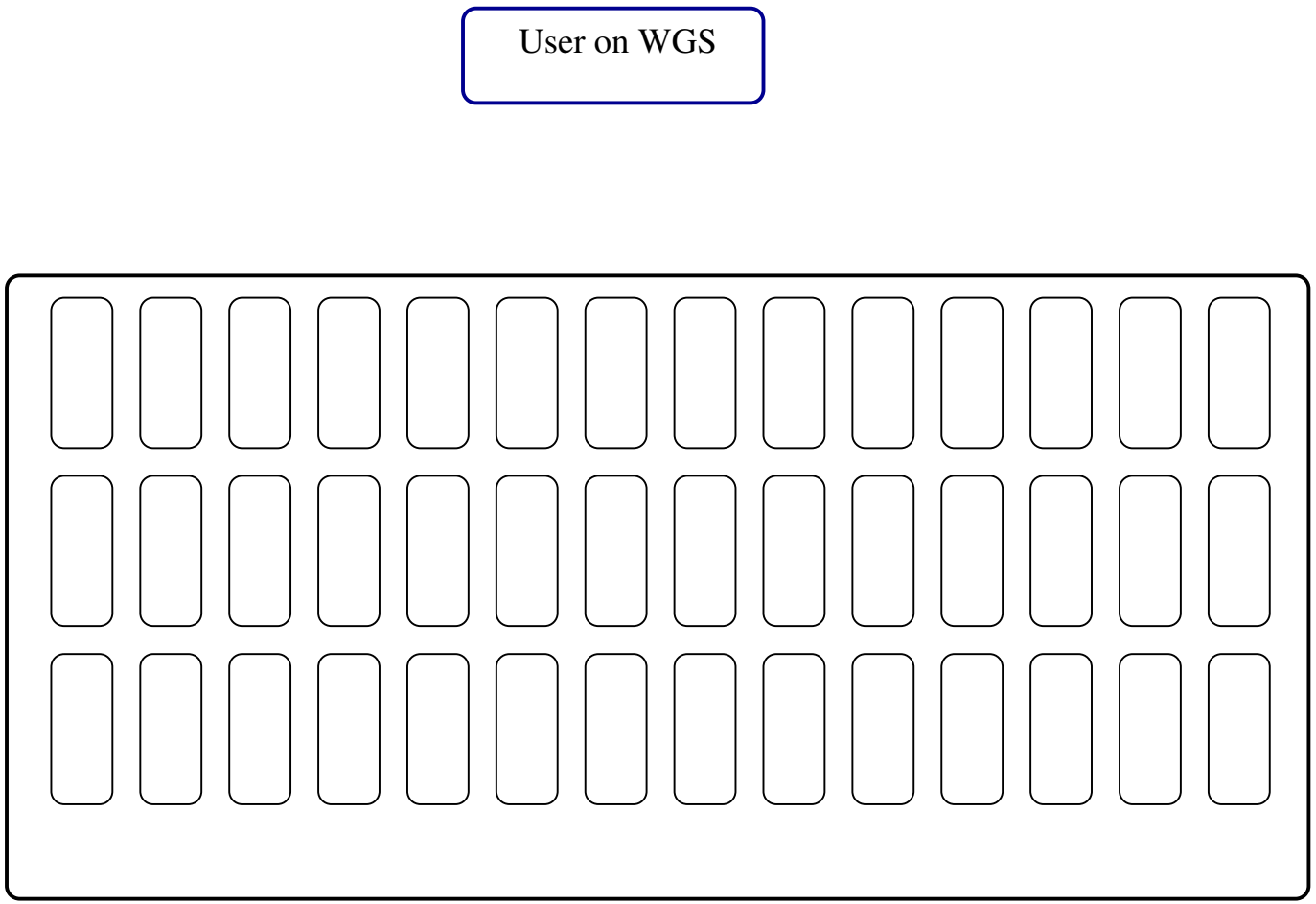
- local batch system (higher reliability)
- worker nodes connected to home directory
- special cluster file system, Lustre

Possible use case:

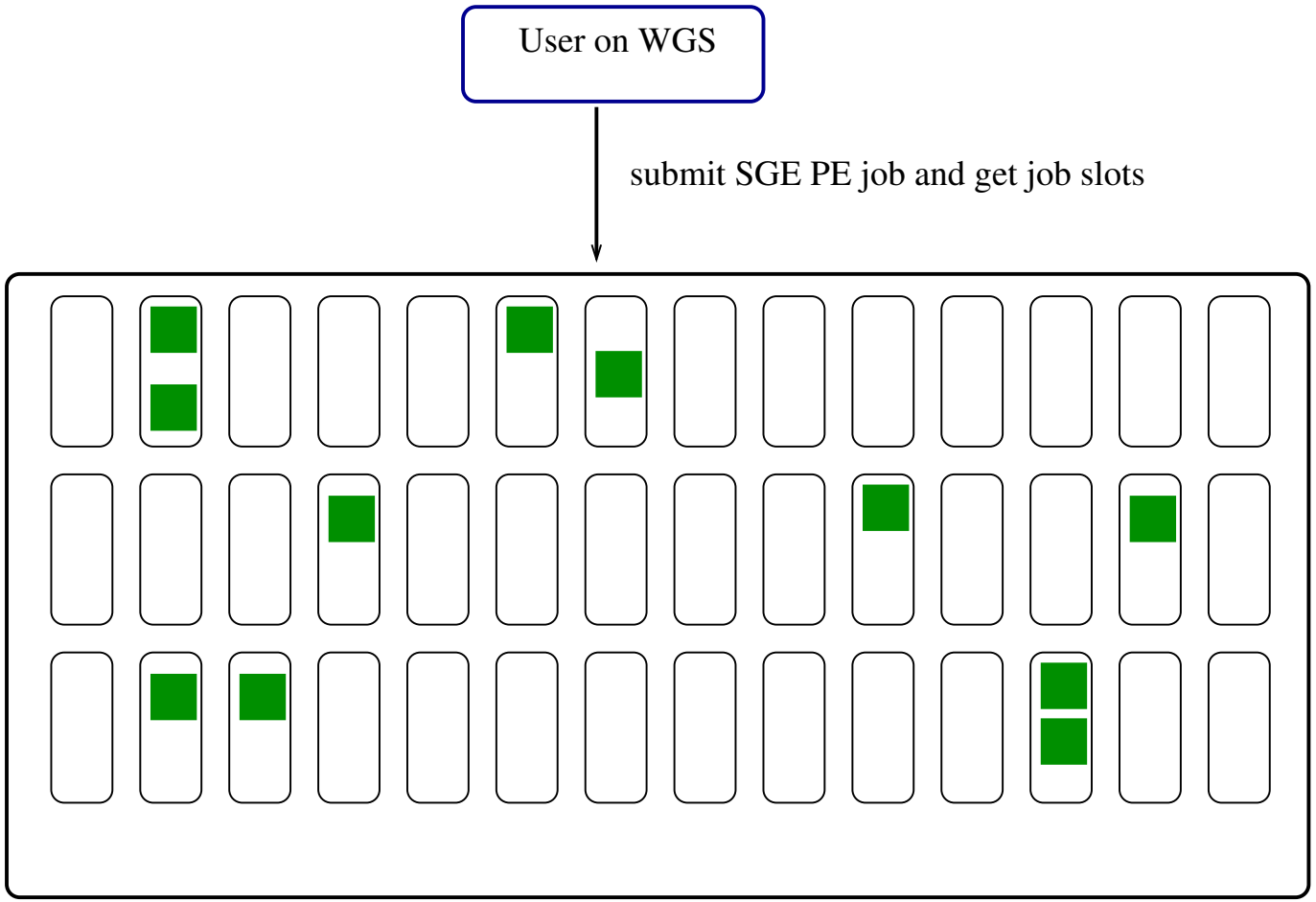
Final data analysis step with:

- many events
- rather large events
- computing intensive algorithms
- need for fast turn-around

PROOF(Parallel ROOT Facility)

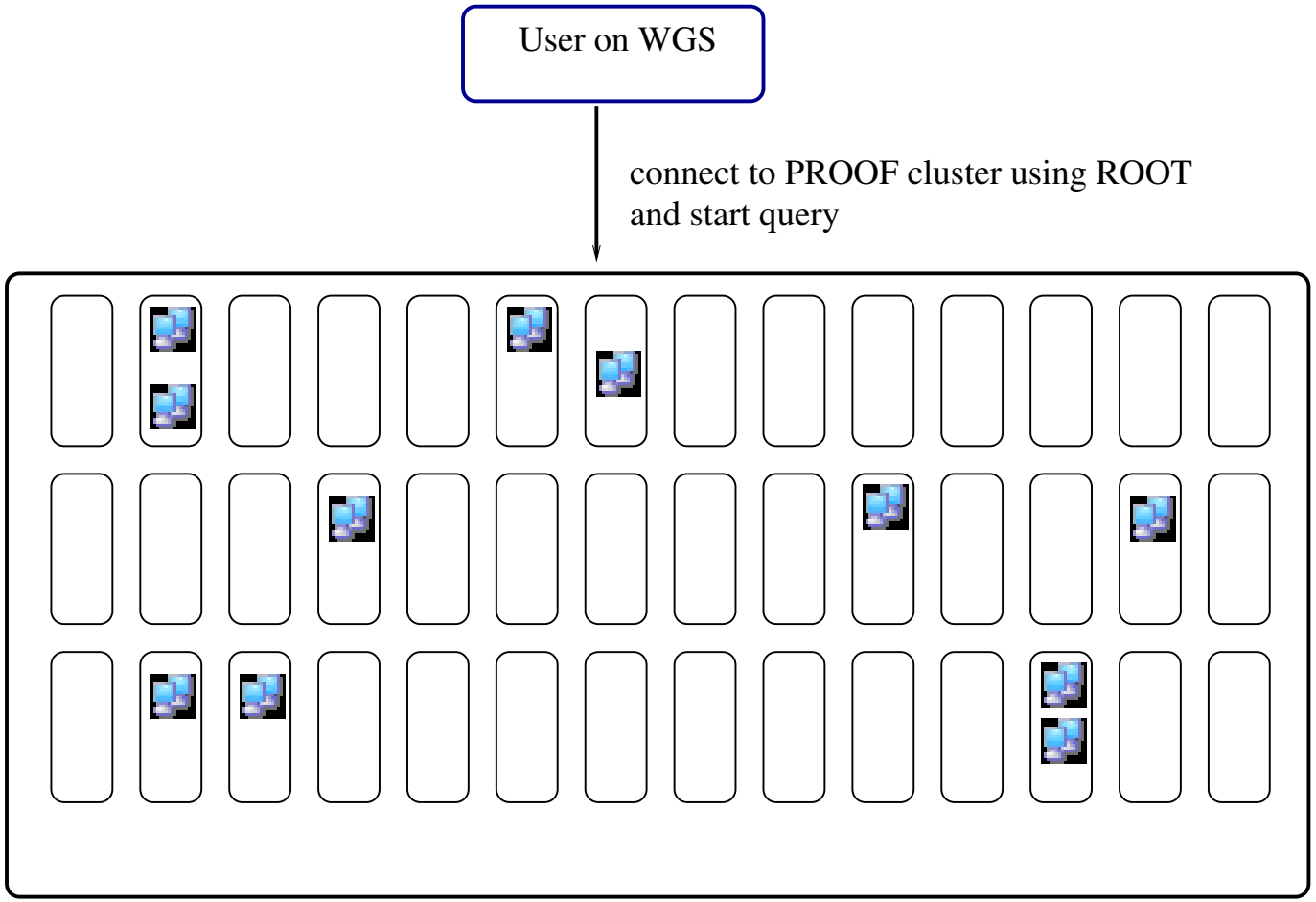


PROOF(Parallel ROOT Facility)



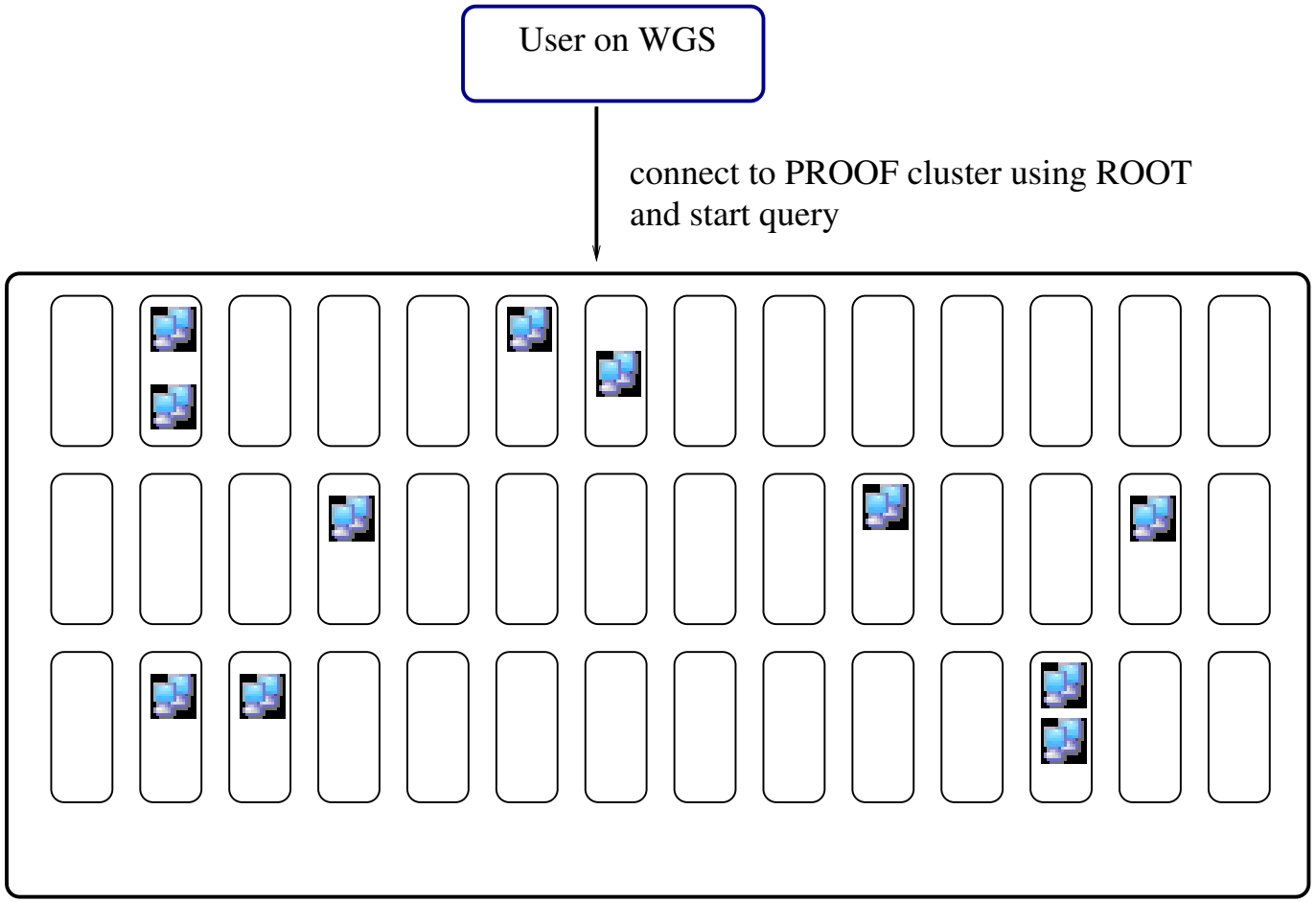
1 start PE job

PROOF(Parallel ROOT Facility)



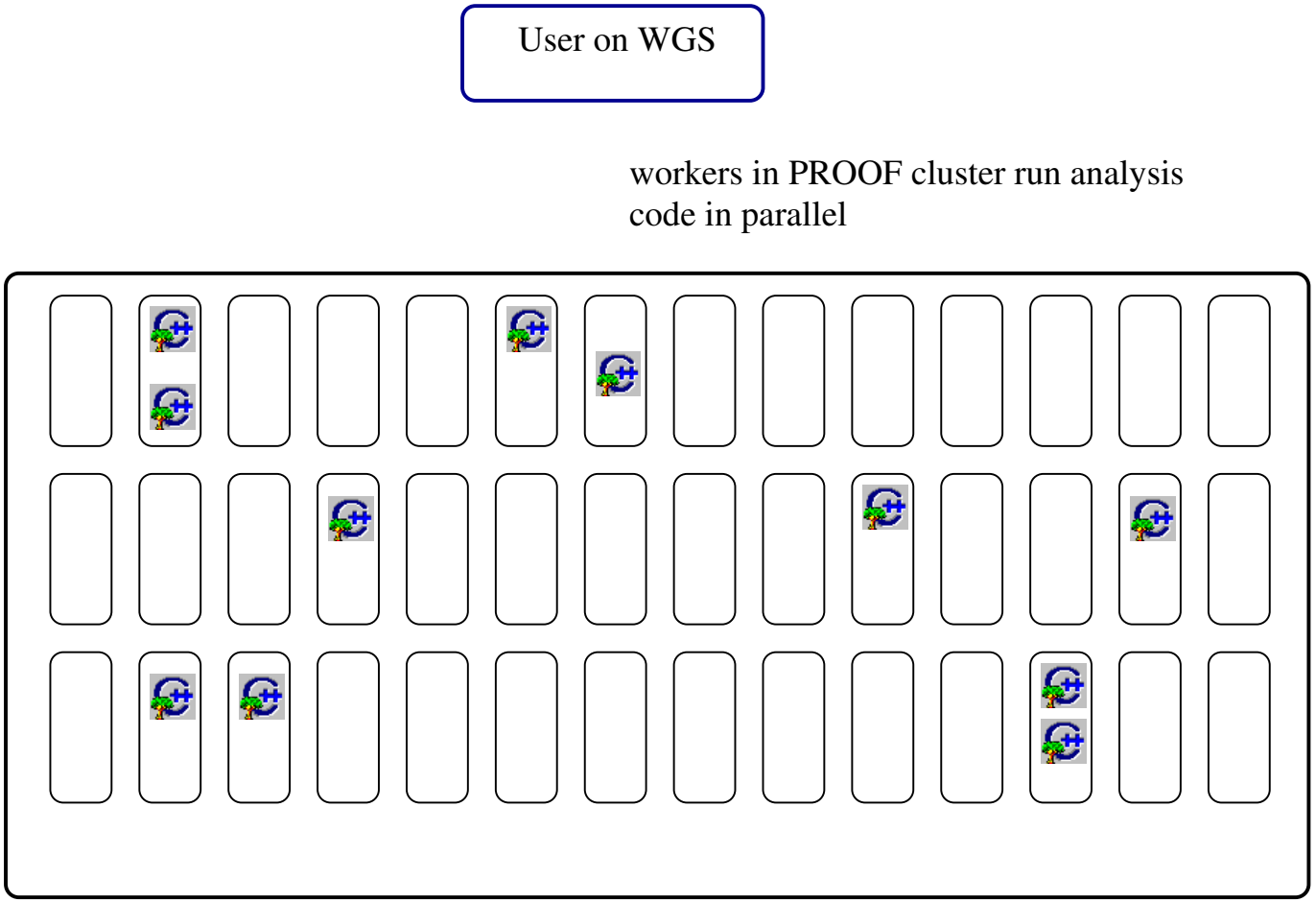
1. start PE job
2. start PROOF cluster

PROOF(Parallel ROOT Facility)



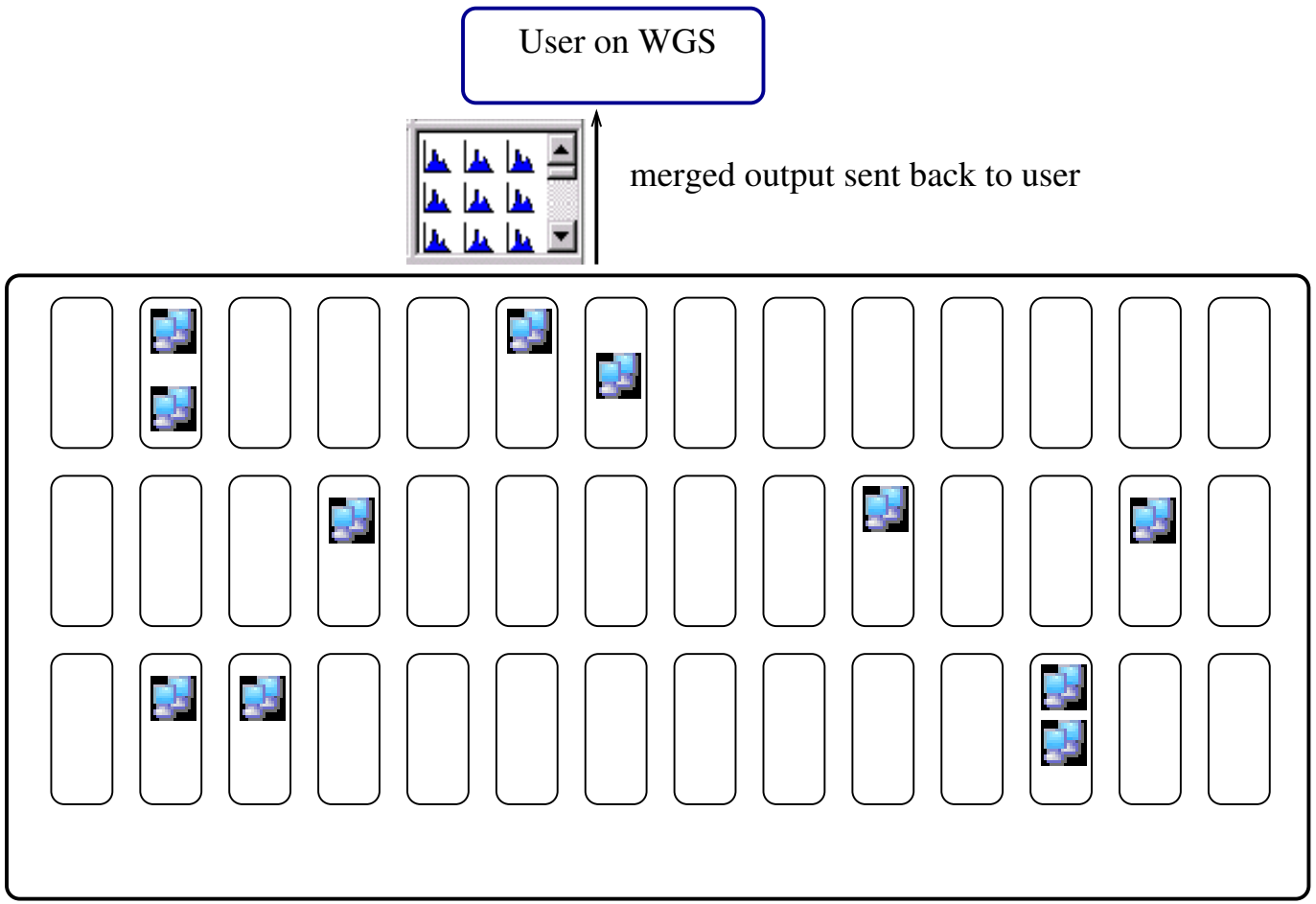
1. start PE job
2. start PROOF cluster
3. connect to cluster within ROOT

PROOF(Parallel ROOT Facility)



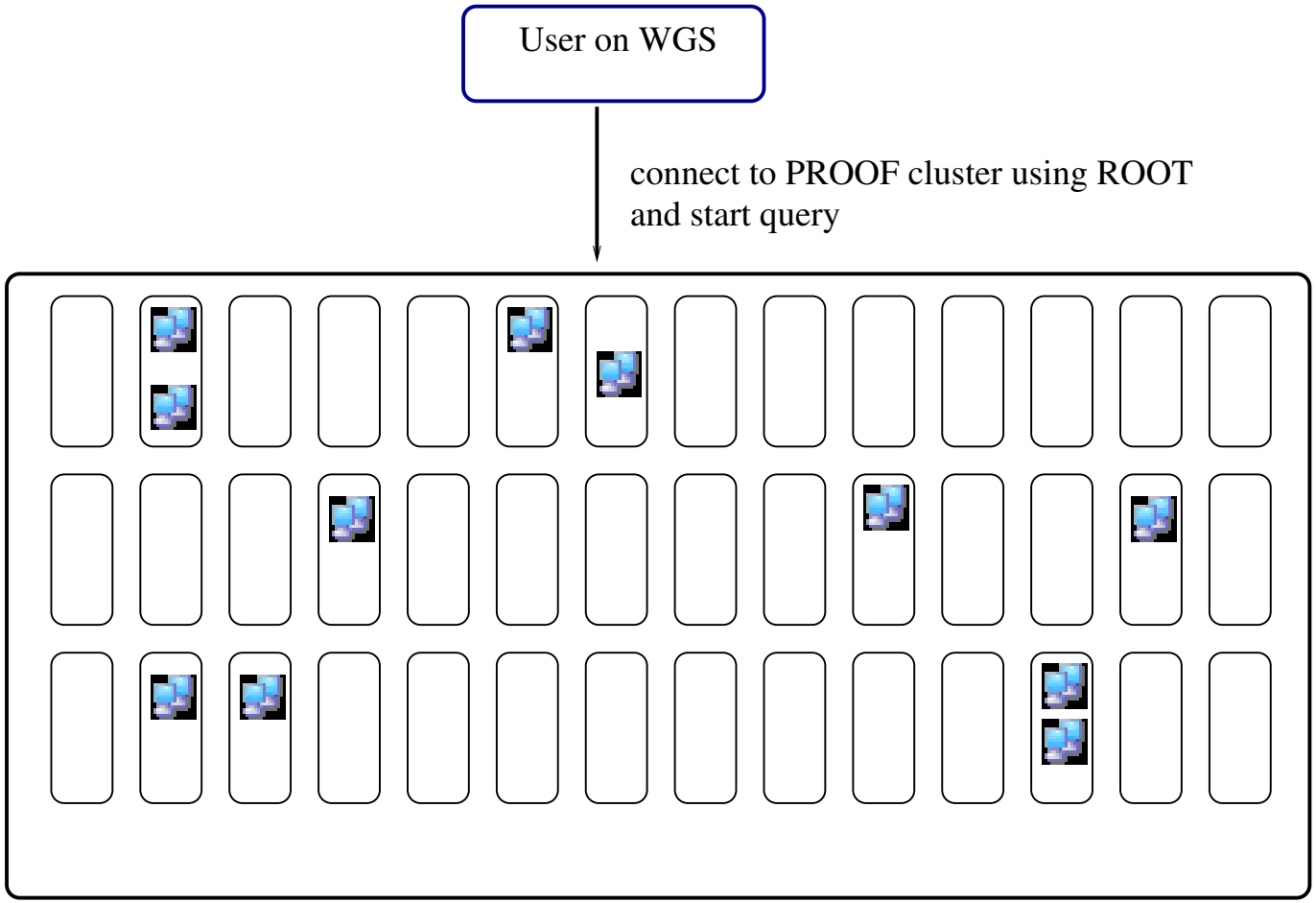
1. start PE job
2. start PROOF cluster
3. connect to cluster within ROOT
4. run your analysis code

PROOF(Parallel ROOT Facility)



1. start PE job
2. start PROOF cluster
3. connect to cluster within ROOT
4. run your analysis code
5. get merged output

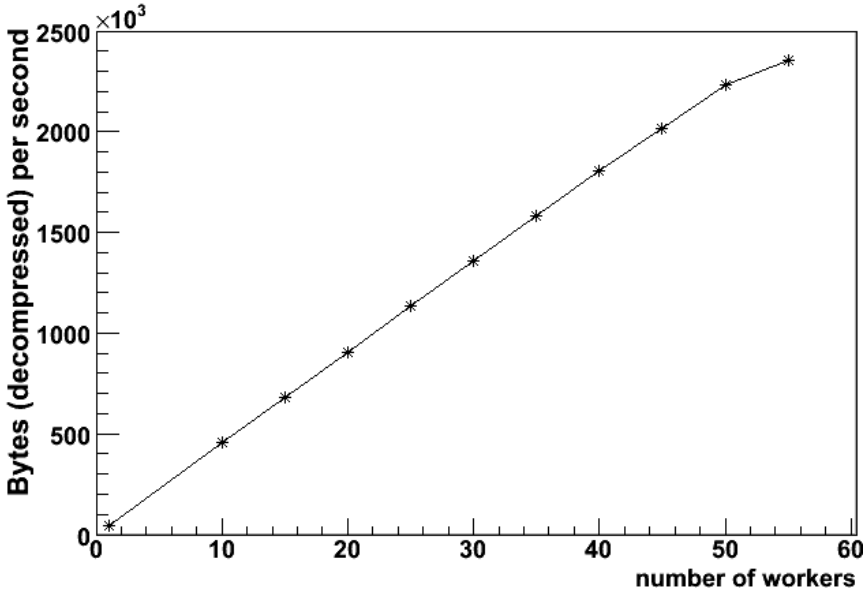
PROOF(Parallel ROOT Facility)



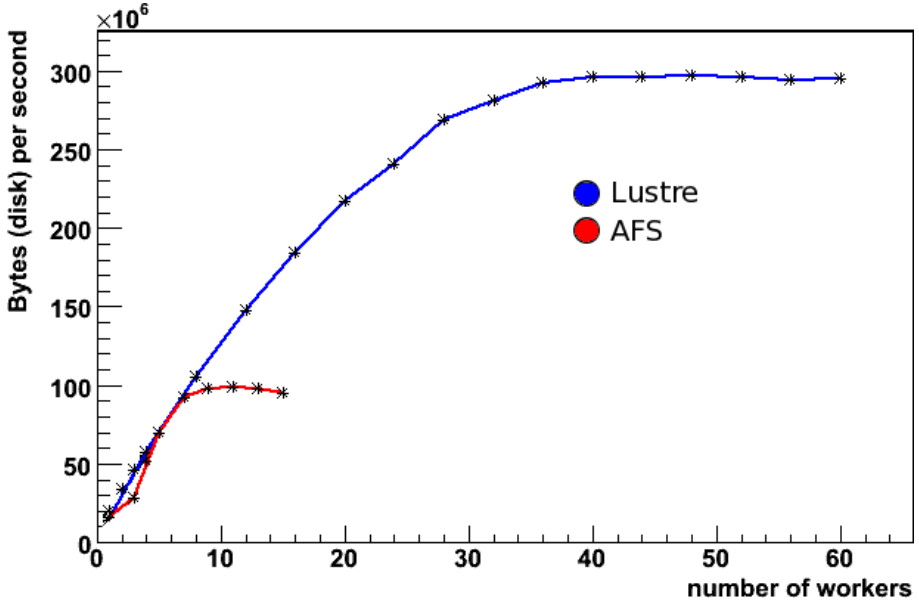
1. start PE job
2. start PROOF cluster
3. connect to cluster within ROOT
4. run your analysis code
5. get merged output
6. study results and repeat 3 to 6

First Performance Plots

CPU-bound task
(reclustering of jets in underlying event analysis)



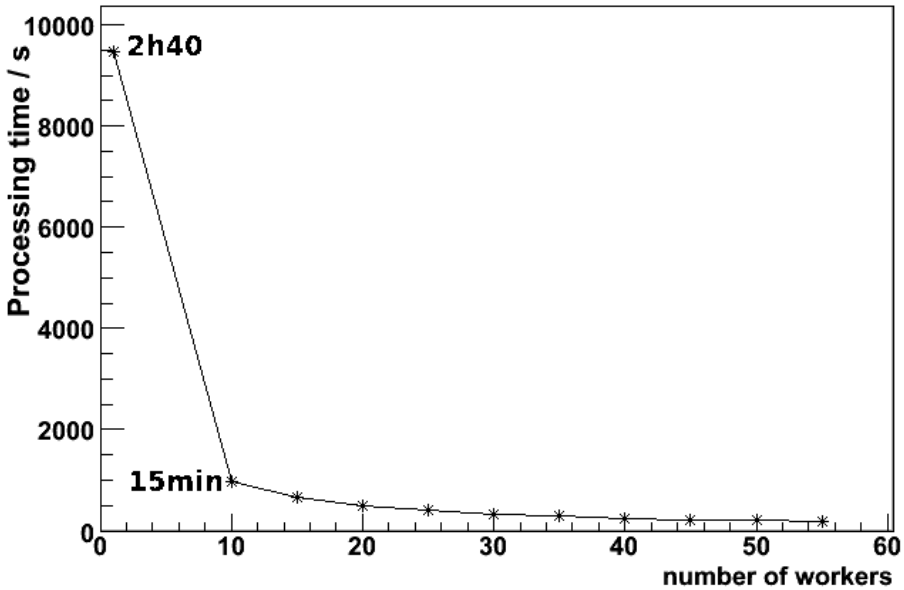
I/O-bound task
(e.g. plot electron p_T in W candidates)



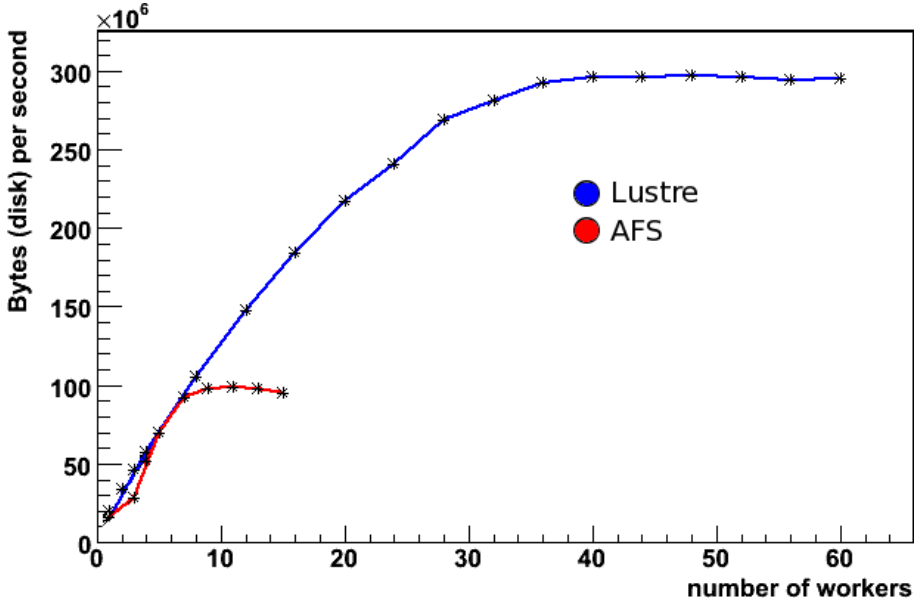
How-to for CMS users: look at CMS NAF page

First Performance Plots

CPU-bound task
(reclustering of jets in underlying event analysis)



I/O-bound task
(e.g. plot electron p_T in W candidates)



analysis time decreased from more than two hours to five minutes!

How-to for CMS users: look at CMS NAF page