



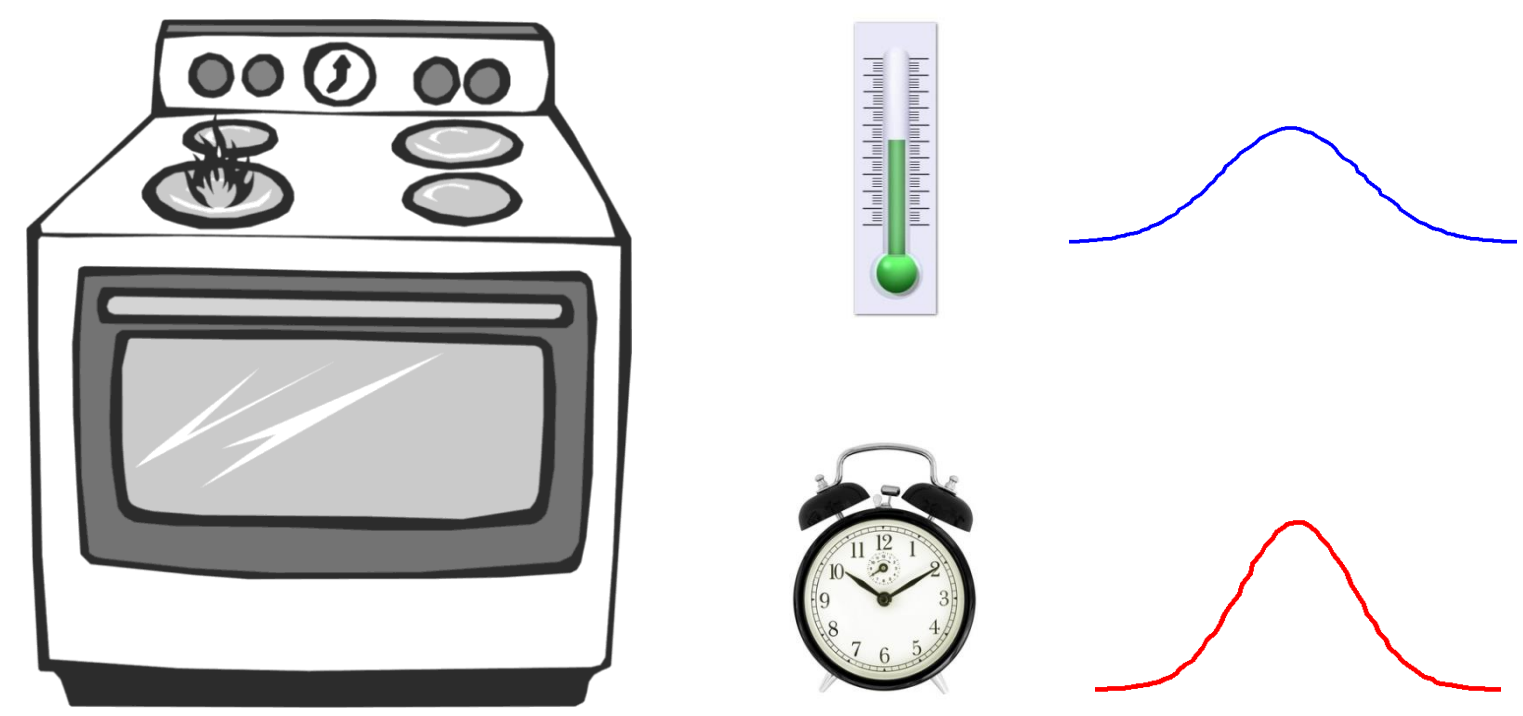
System as a Simulation

Jun He[§], Duy Nguyen[#], Andrea C. Arpaci-Dusseau[§], Remzi H. Arpaci-Dusseau[§]

[§]Department of Computer Sciences [#]Department of Statistics
University of Wisconsin - Madison



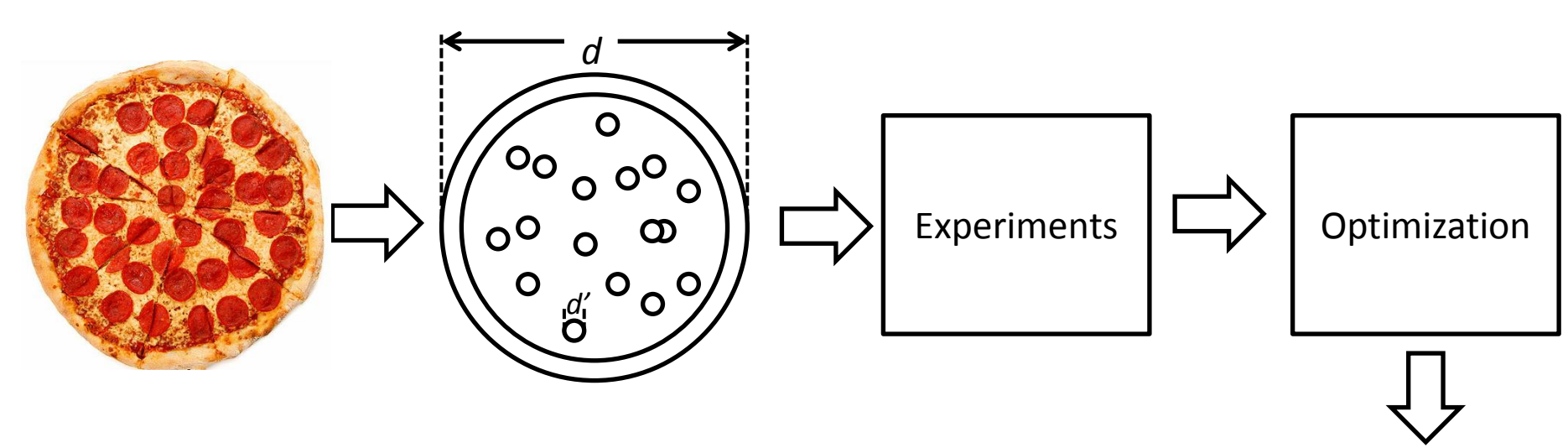
What's a Good Pizza?



How to Produce Good Pizza?



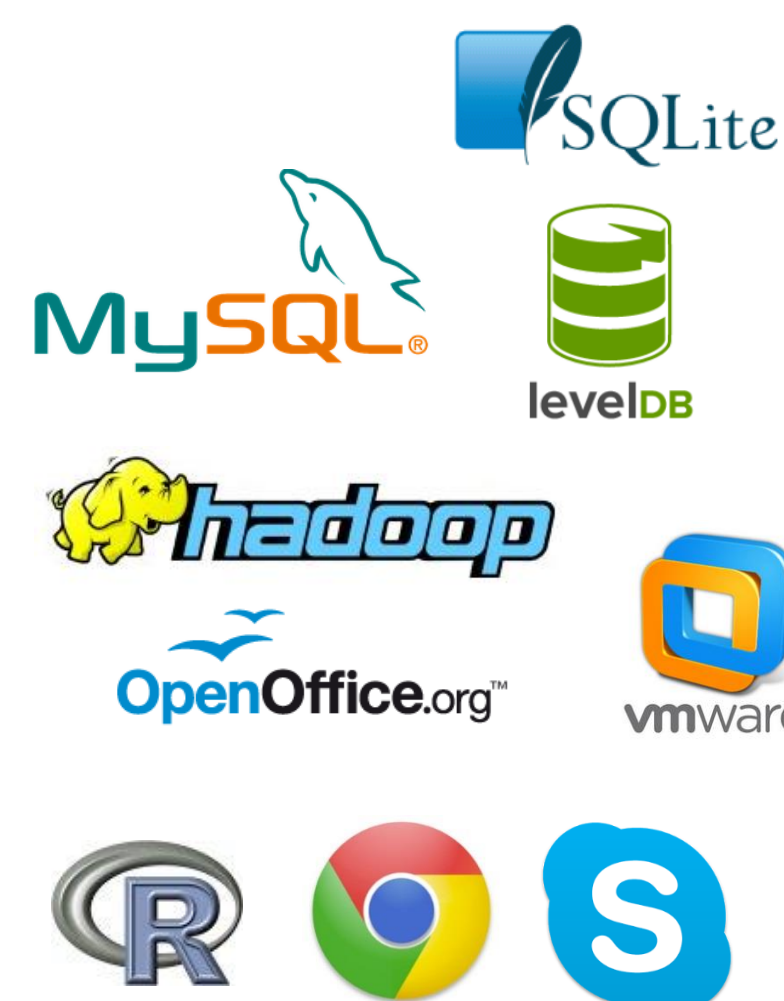
We can build a pizza simulator.



$taste = f(\text{crust}, \text{cheeze}, \text{pepperoni}, \text{temperature}, \text{time})$

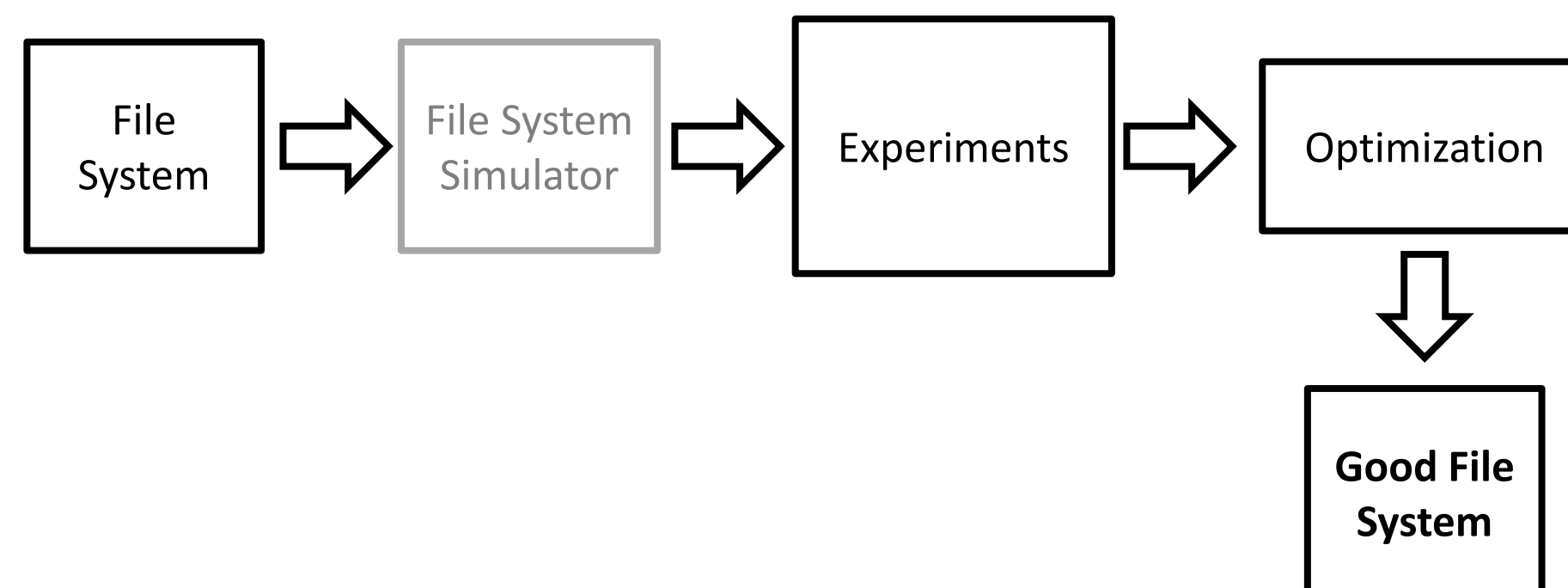
Better examples:
Nuclear bomb, prosthesis, medicine, ...

What's a Good File System?



- file size
- fsync
- sequentiality
- # of writes
- disk size
- directory
- # of cores
- overwrites
- ...

We can build a file system simulator?

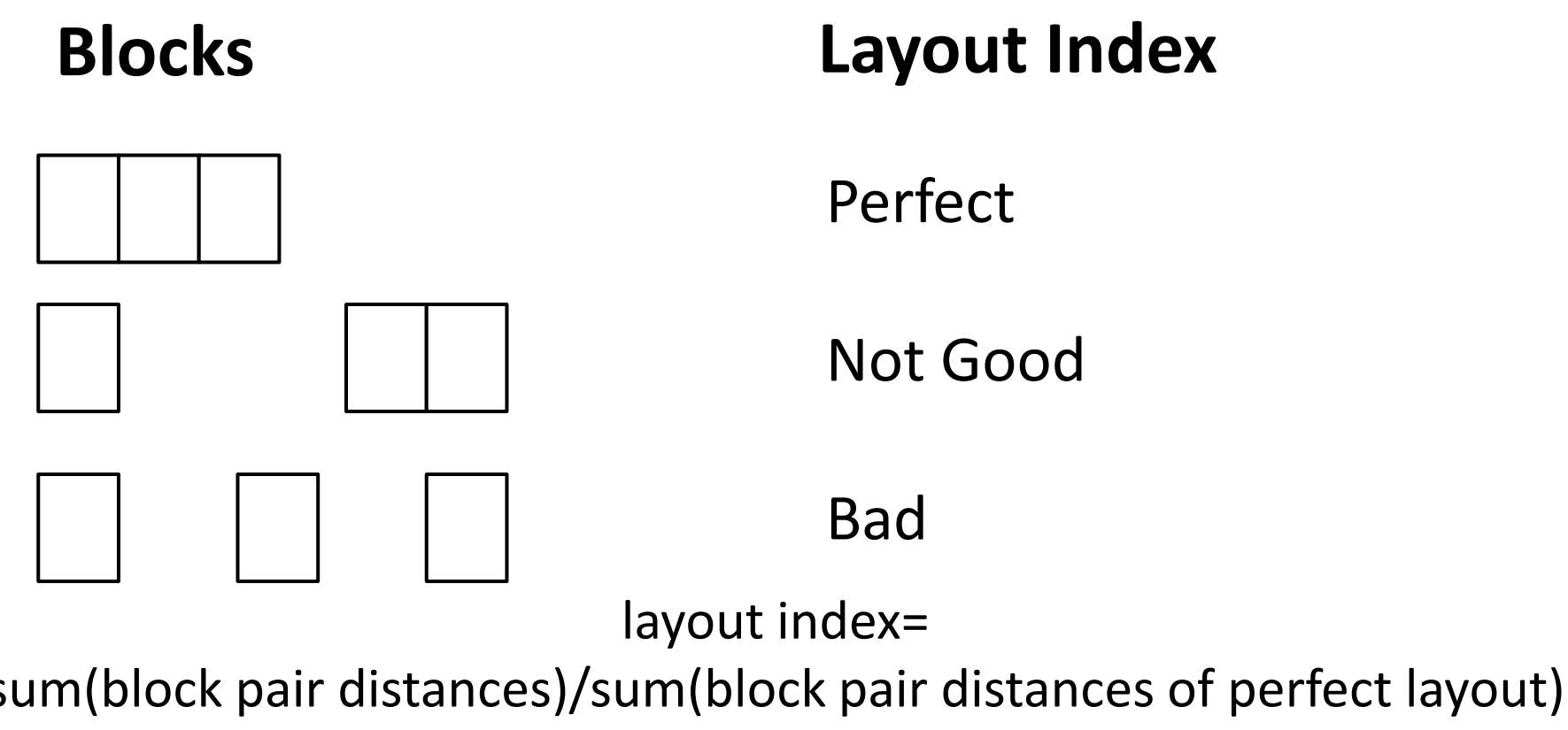


We want to help:

- File system designers
 - diagnose design weakness
 - find best policy/configuration
- File system users
 - match their workload with file system
 - use file system efficiently

We focus on block allocation in this study.

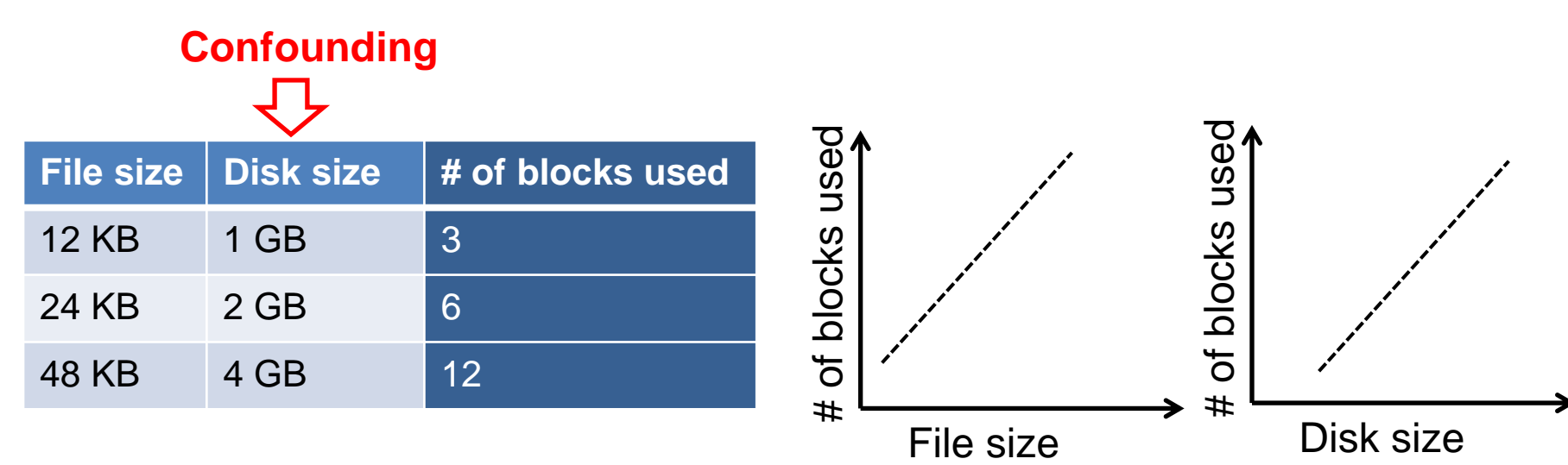
Layout index is a metric that is able to differentiate the following cases



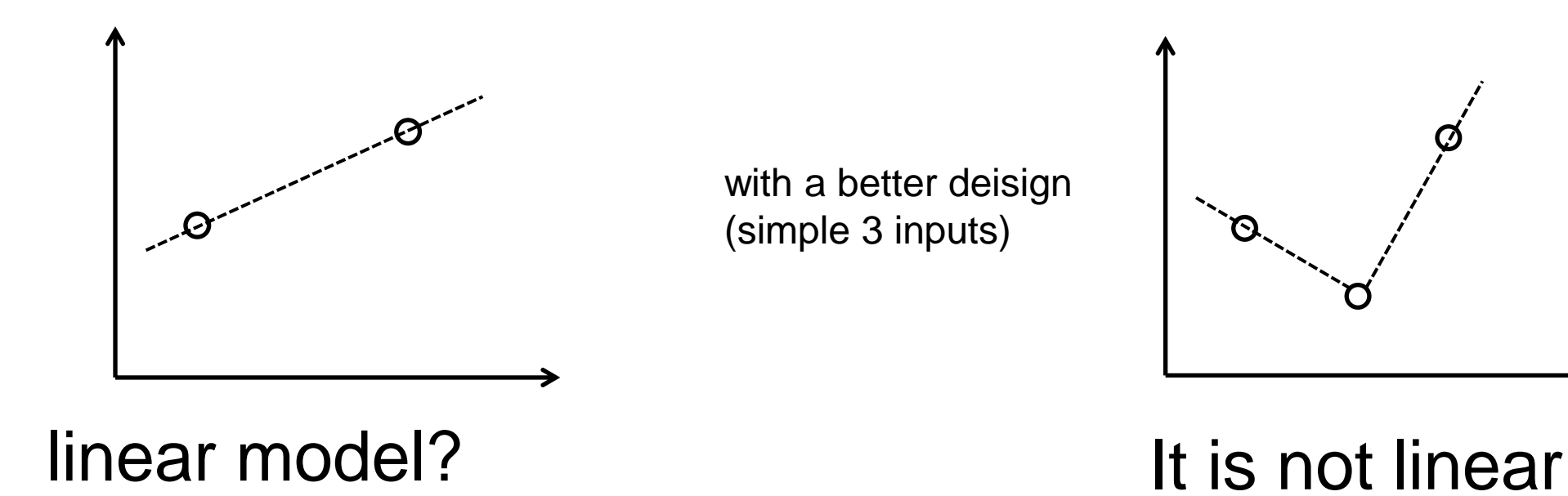
We want to find out how much the following factors affect layout index

- directory span
- disk size
- disk used
- file size
- fullness
- sparse/overwrite
- number of files
- number of chunks
- write order
- fsync style
- sync style

Experimental design helps avoid confounding factors



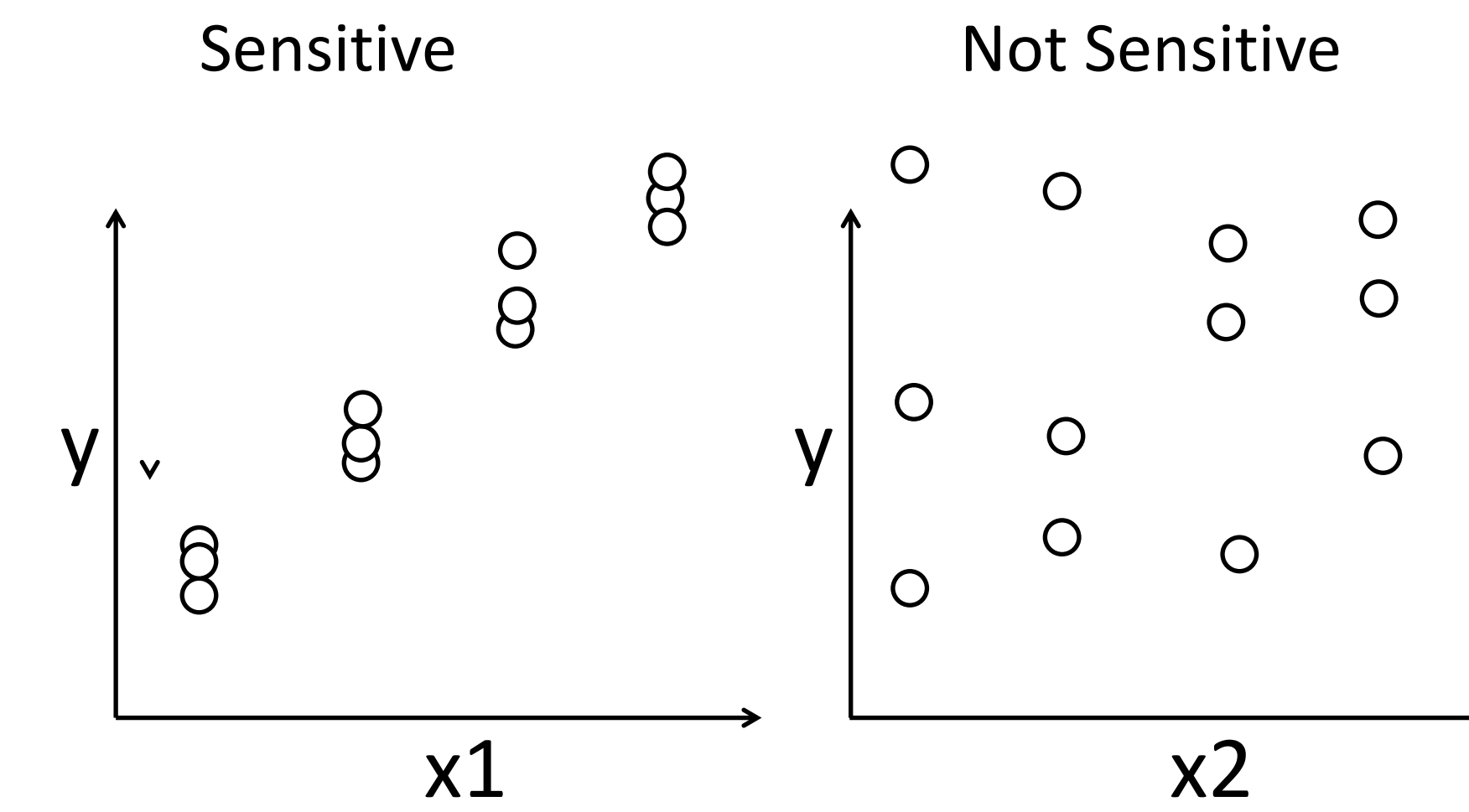
Experimental design helps detect model bias



We use a nested experimental design

runs	number of chunks	fsync sync order	shared factors
n	1	$OA(n, 2^1 2^0 (1!)^1)$	Latin Hypercube Design
n	2	$OA(n, 2^2 2^1 (2!)^1)$	
n	3	$OA(n, 2^3 2^2 (1!)^1)$	
n	4	$OA(n, 2^1 2^0 (1!)^1)$	

We conduct sensitivity analysis

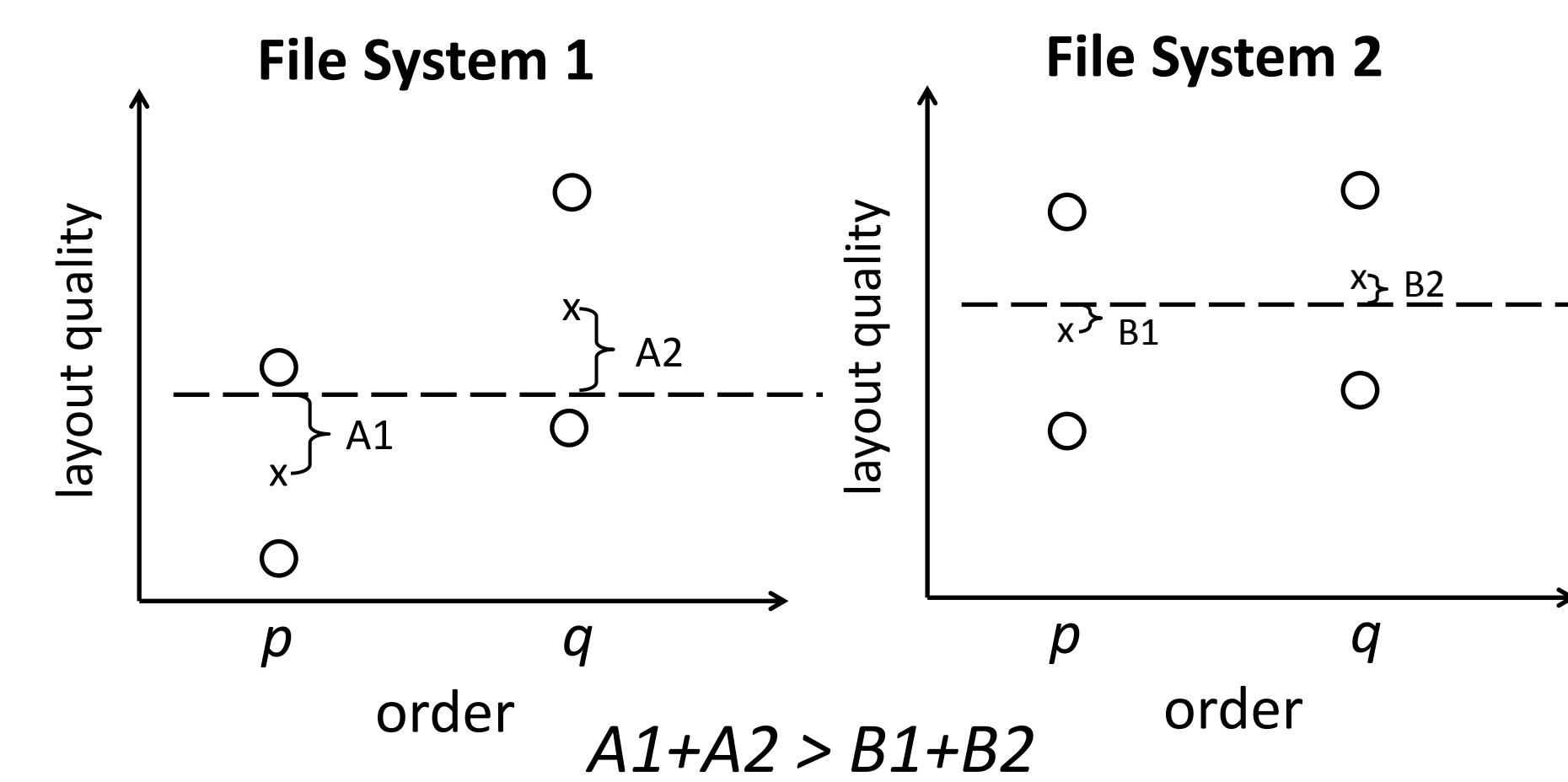


Our analysis can reveal effects of interactions

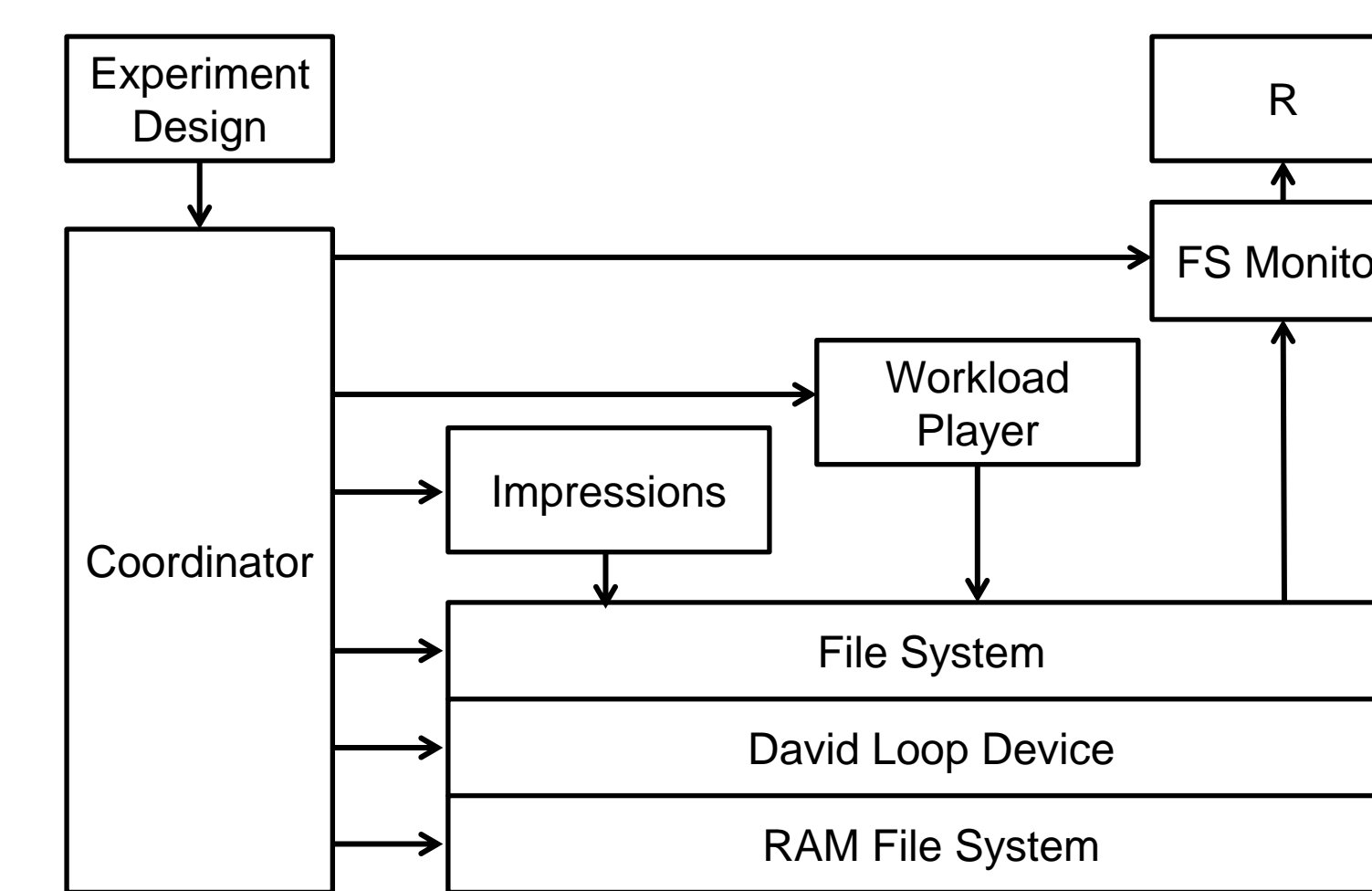
Without interaction	Order=Sequential	Order=Nonsequential
file.size=12KB	Good	Bad
file.size=96KB	Good	Bad

With interaction	Order=Sequential	Order=Nonsequential
file.size=12KB	Good	Bad
file.size=96KB	Bad	Good

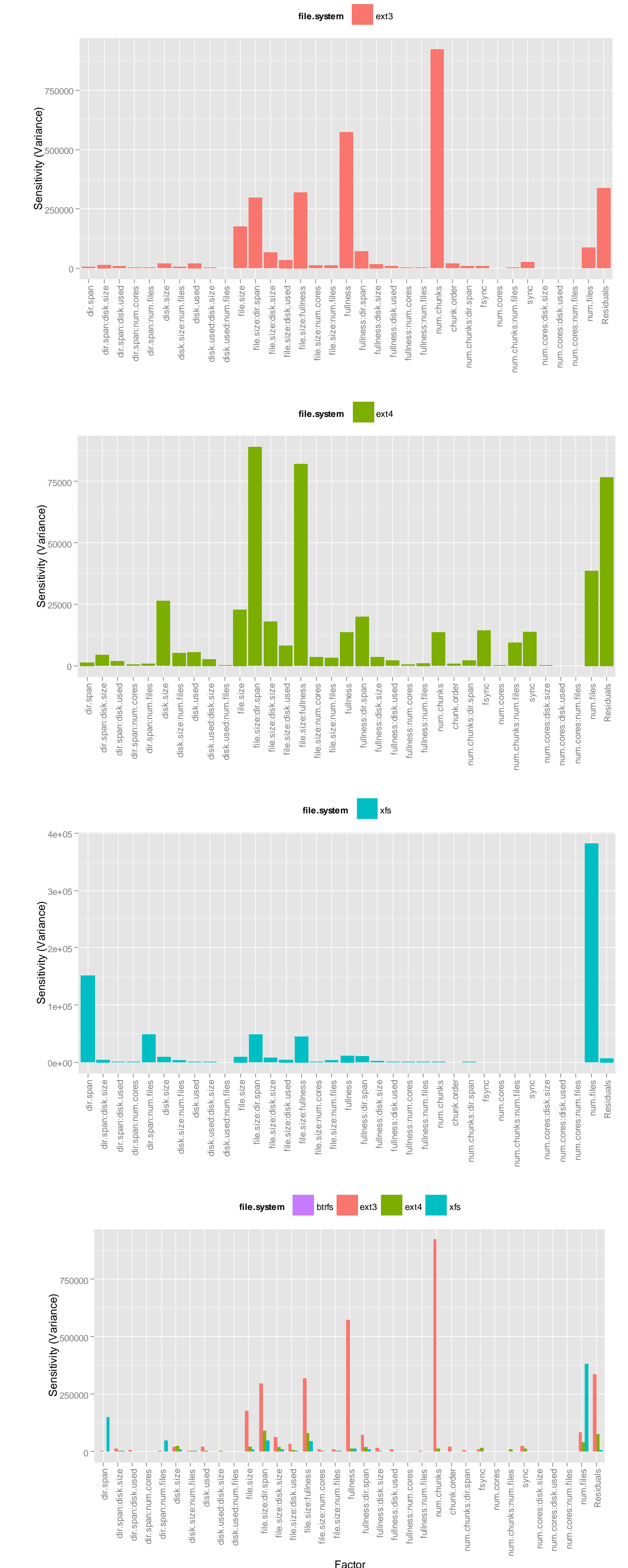
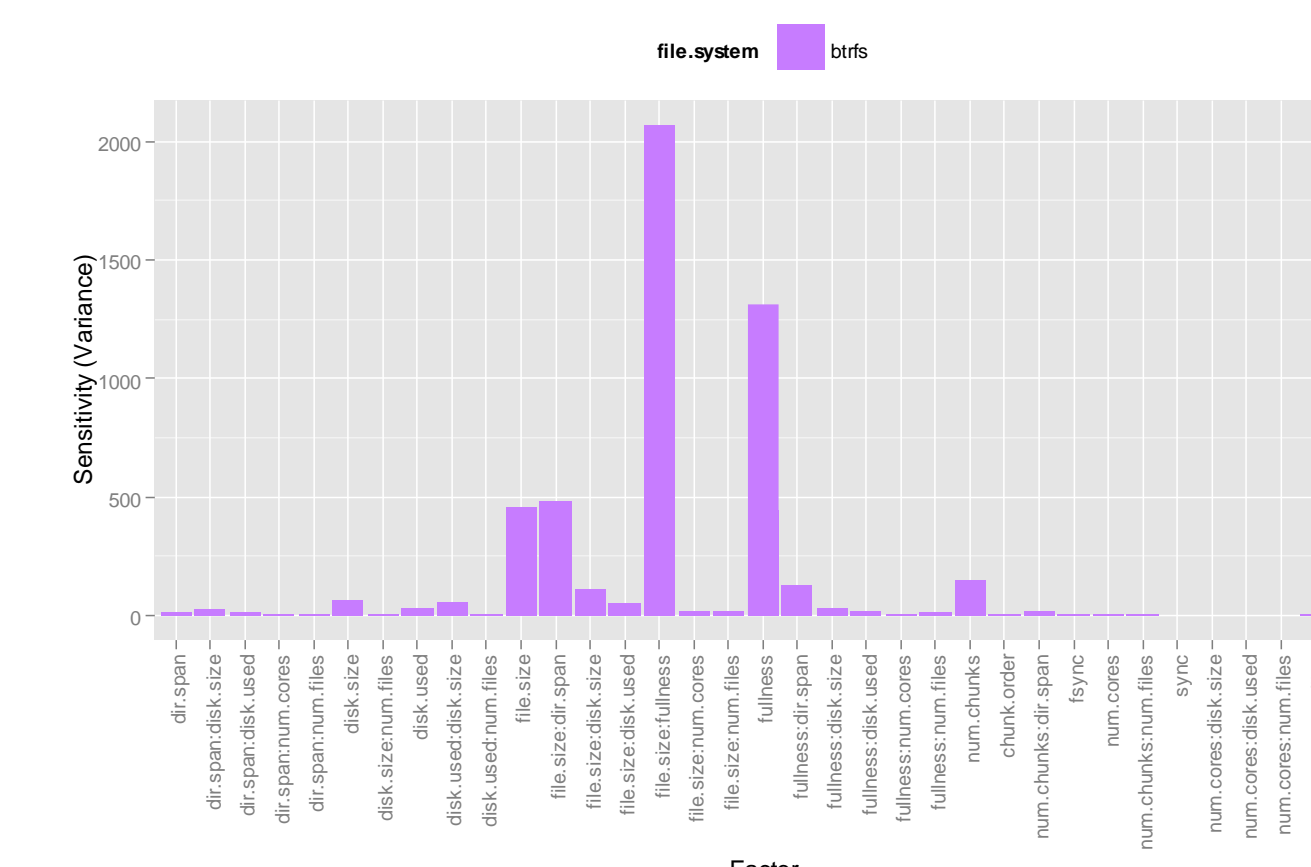
ANOVA-base Sensitivity



We built a realistic and fast test framework



Results



Conclusions

ext3 is the most sensitive. Btrfs is the least sensitive. Our "System as a Simulation" approach helps diagnose problems and guide system/application design.

Acknowledgements

We would like to thank PRObE project for providing the testbeds. We would like to thank Peter Qian from Department of Statistics for his help.