# VarSys

## Managing variability for high performance computers, cloud computing systems, and computer security.

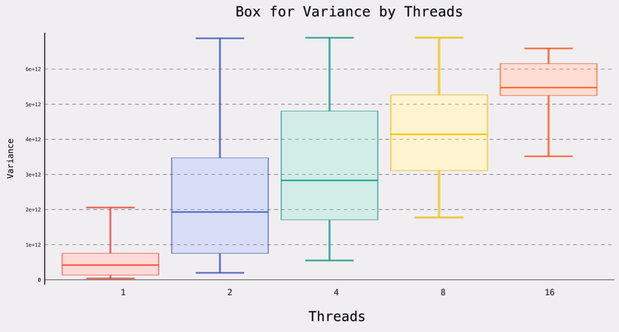
## Motivation

What is variance?

$$\sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n-1}$$

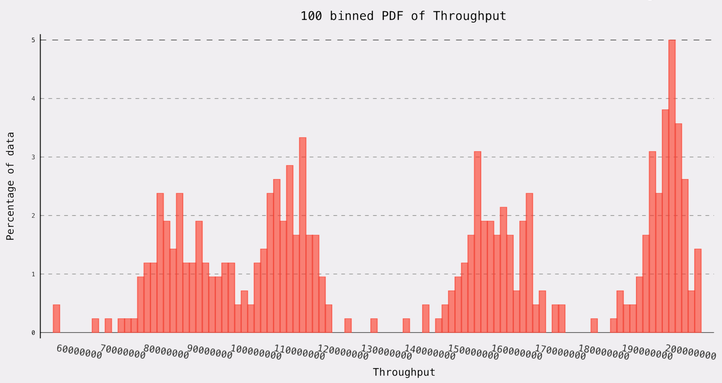
Why is variability important?

Computational performance varies!




What about distributions?

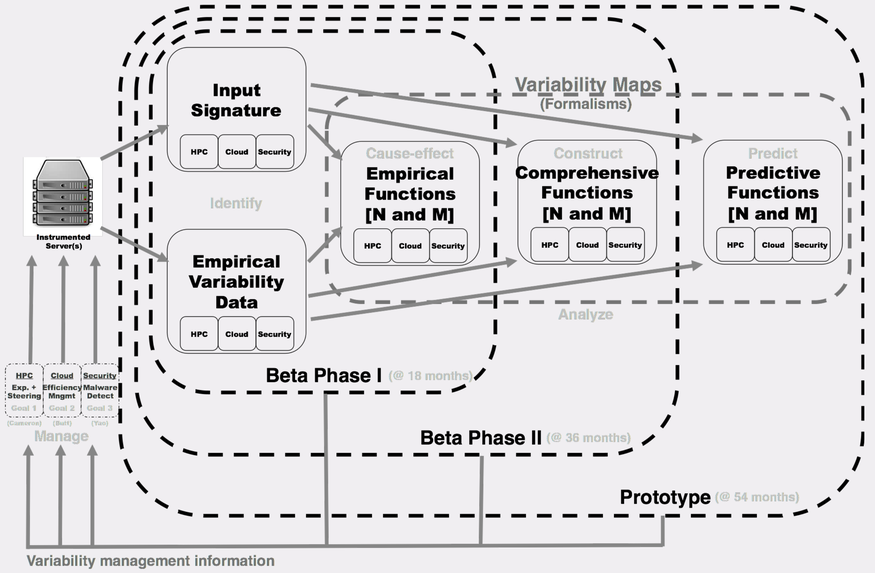
A distribution (rather than a statistic) tells us a lot more about the behavior of a system.




## Modeling the System

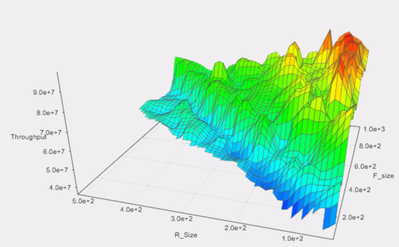
| System Parameter | Values |
|---|---|
| File Size | 64, 256, 1024 |
| Record Size | 32, 128, 512 |
| Thread Count | 1, 2, 4, 8, 16, 32, 64, 128, 256 |
| Frequency | $\{12, 14, 15, 16, 18, 19, 20, 21, 23, 24, 25, 27, 28, 29, 30, 30.01\} \times 10^5$ |
| **Response Values** | |
| Throughput Mean | $[2.6 \times 10^5, 5.9 \times 10^8]$ |
| Throughput Variance | $[5.9 \times 10^{10}, 4.7 \times 10^{16}]$ |

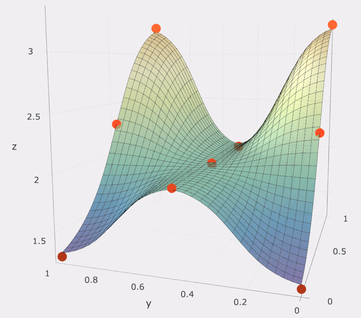
First, we describe the system in terms of controllable *parameters* (numbers and strings). Then, we can build statistical and mathematical models of the performance of the system.

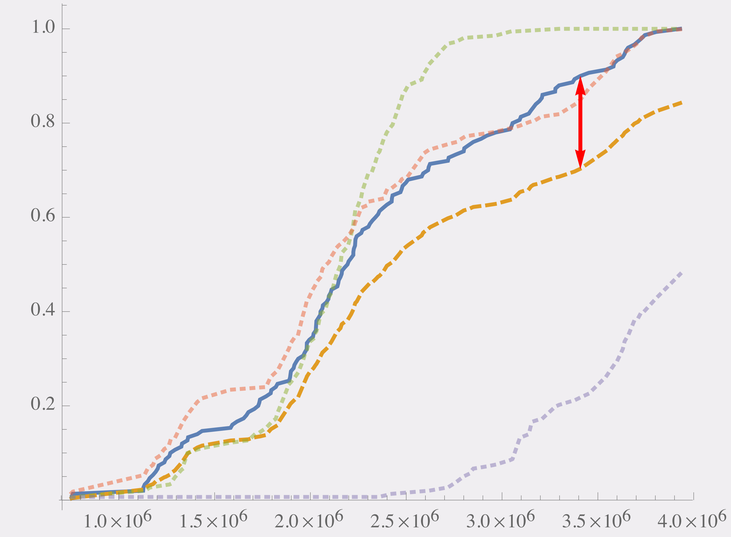



## Interpolation and Regression

Using mathematical and statistical techniques (some from machine learning) we can predict the performance of a system and the *distribution* of performance of a system.




## High Performance Computing

HPC systems consume a lot of energy and time, both are functions of how the system was built and configured. Models can be used to optimize a configuration.

## Cloud Computing

Small savings in compute time and performance magnify greatly when 1000's of machines are involved. Service Level Agreements (SLAs) can be tightened.

## Computer Security

A strong understanding of variability can improve defenses against malicious users by demonstrating new vulnerabilities, and helping prevent side channel attacks.