

PROBLEM DIAGNOSIS & VISUALIZATION

Jiaqi Tan, Soila Pertet, Mike Kasick, Xinghao Pan, Keith Bare, Eugene
Marinelli, Rajeev Gandhi

Priya Narasimhan

Carnegie Mellon University



Automated Problem Diagnosis

- Diagnosing problems
 - Creates major headaches for administrators
 - Worsens as scale and system complexity grows
- Goal: automate it and get proactive
 - Failure detection and prediction
 - Problem determination (or “fingerpointing”)
 - Problem visualization
- How: Instrumentation plus statistical analysis



Exploration of Fingerprinting

- Current explorations

- *Hadoop*

- [HotCloud 09, HotMetrics 09, WASL 08, SysML 08, NOMS 10, ISSRE 09, ICDCS 10, CCGrid 10]

- *PVFS & Lustre*

- High-performance file system (Argonne National Labs) [HotDep 09, FAST 10]

- *RUBiS*

- Three-tier web service (Rice University) [NOMS 10]

- Studied

- Various types of problems
 - Various kinds of instrumentation
 - Various kinds of data-analysis techniques



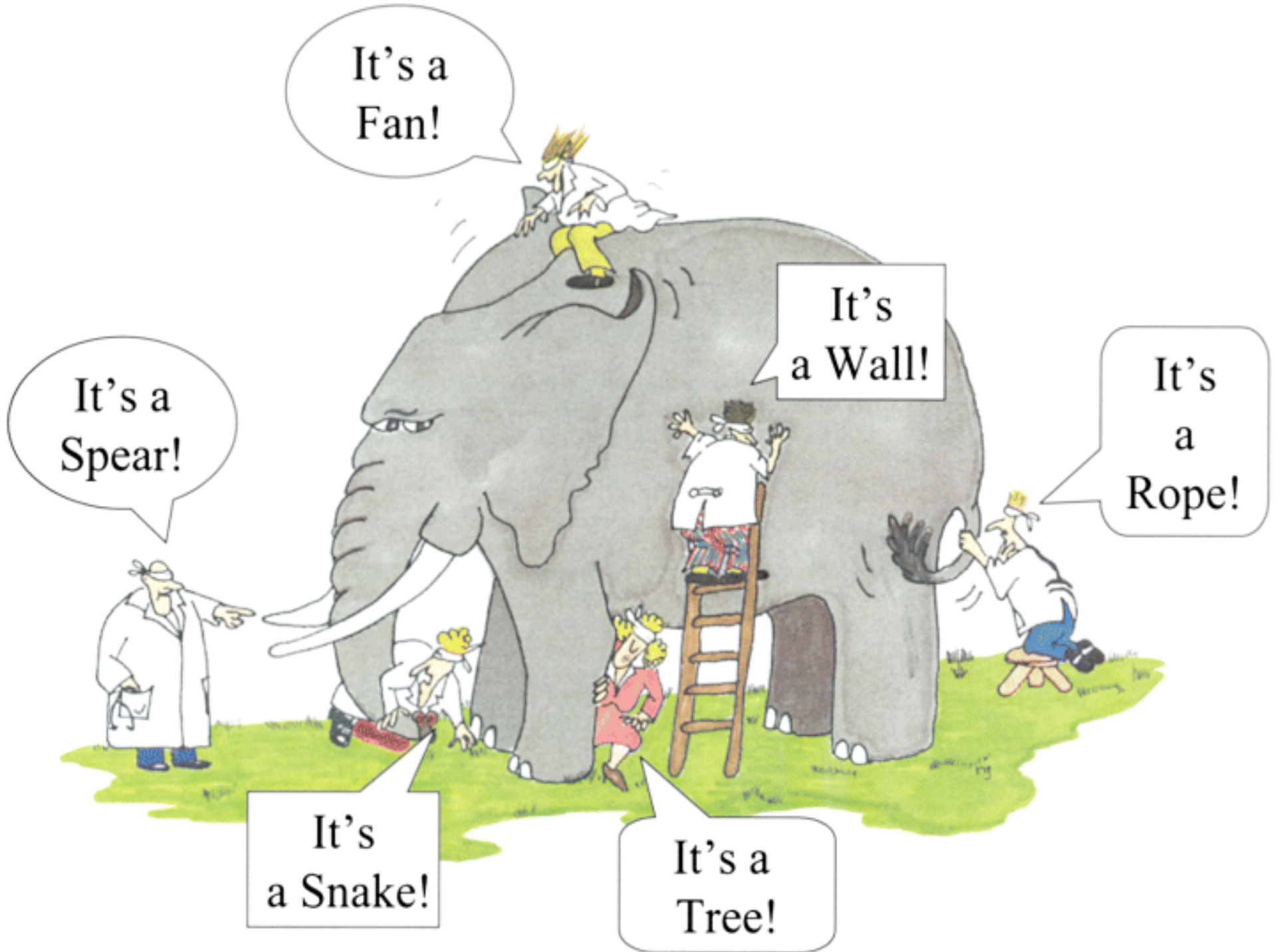
Performance Problems (Hadoop)

	Fault	Description
Resource contention	CPU hog	External process uses 70% of CPU
	Packet-loss	5% or 50% of incoming packets dropped
	Disk hog	20GB file repeatedly written to
	Disk full	Disk full
Application bugs Source: Hadoop JIRA	HADOOP-1036	Maps hang due to unhandled exception
	HADOOP-1152	Reduces fail while copying map output
	HADOOP-2080	Reduces fail due to incorrect checksum
	HADOOP-2051	Jobs hang due to unhandled exception
	HADOOP-1255	Infinite loop at Nameode

Studied Hadoop Issue Tracker (JIRA) from Jan-Dec 2007

Performance Problems (PVFS)

- Problems motivated by PVFS developers' experiences
 - From Argonne's Blue Gene/P PVFS cluster
- "Limping-but-alive" server problems
 - No errors reported, can't identify faulty node with logs
 - Single faulty server impacts overall system performance
- Storage-related problems:
 - Accidental launch of rogue processes, decreases throughput
 - Buggy RAID controller issues patrol reads when not at idle
- Network-related problems:
 - Faulty-switch ports corrupt packets, fail CRC checks
 - Overloaded switches drop packets but pass diagnostic tests



It's a Fan!

It's a Wall!

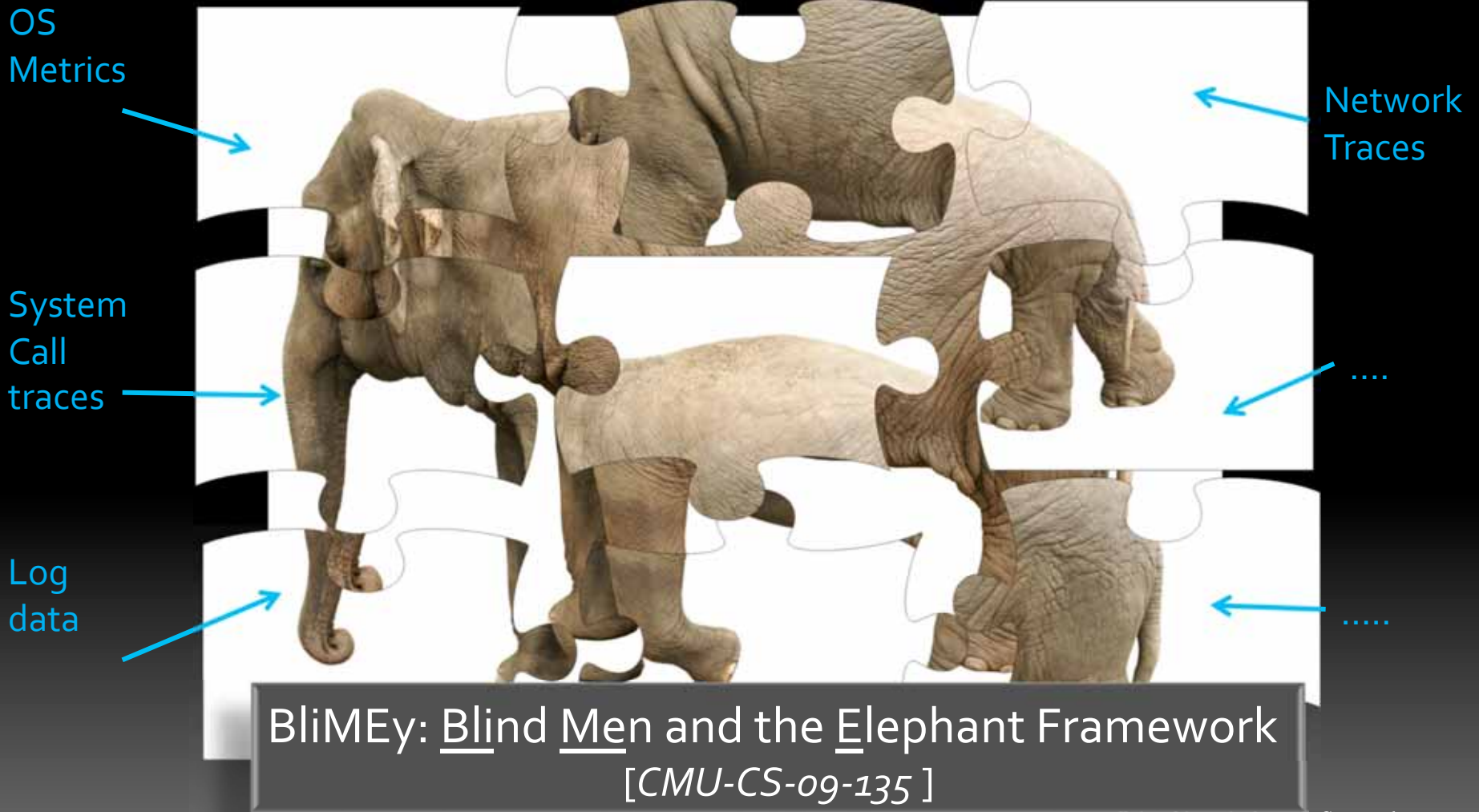
It's a Rope!

It's a Snake!

It's a Snake!

It's a Tree!

Putting the Elephant Together



Visualization

- To uncover Hadoop's execution in an insightful way for admins
- To reveal outcome of diagnosis on sight
- To allow developers/admins to get a handle as the system scales
- Value to programmers
 - Allows them to spot issues that might assist them in restructuring their code
 - Allows them to spot faulty nodes

Other Problems

- Understanding the limits of black-box fingerprinting
 - What failures are outside the reach of a black-box approach?
 - What are the limits of “peer” comparison?
 - What other kinds of black-box instrumentation exist?
- Scalability
 - Scaling to run across large systems and understanding “growing pains”
- Visualization
 - Helping system administrators visualize problem diagnosis
- Trade-offs
 - More instrumentation and more frequent data can improve accuracy of diagnosis, but at what performance cost?
- Virtualized environments
 - Do these environments help/hurt problem diagnosis?

Summary

- Automated problem diagnosis, both online and offline
- Current focus: Hadoop, PVFS, Lustre
 - Real-world bug databases, problems in the wild
- Long-term
 - Scalability, scalability, scalability,
 - Expand fault study
 - Improve visualization, working with users
- Additional details
 - [USENIX WASL 2008](#) (white-box log analysis)
 - [USENIX HotCloud 2009, ICDCS 2010](#) (visualization)
 - [HotMetrics 2009 & ISSRE 2009](#) (black-box metric analysis)
 - [NOMS 2010](#) (black-box vs. white-box analyses)
 - [NOMS 2010](#) (hardware-counter instrumentation)
 - [CCGrid 2010](#) (M45 data analysis for performance prediction)
 - [HotDep 2009](#) (system-call analysis for PVFS)
 - [USENIX FAST 2010](#) (black-box analysis for PVFS & Lustre)

FOR MORE INFORMATION:

[HTTP://WWW.ECE.CMU.EDU/~FINGERPOINTING](http://www.ece.cmu.edu/~fingerpointing)

priya@cs.cmu.edu