

# Scalable Node Monitoring

Alex Drotar – University of Colorado

Erin Quinn – Fairleigh Dickinson University

Landon Sutherland – New Mexico State University

Mentors: Mike Mason – HPC3, Jon Bringhurst – HPC3

LA-UR-12-23629

# Project Description

- Build a high performance computer
- Create a tool to monitor node applications in Component Based Tool Framework (CBTF) using code from Lightweight Data Metric Service (LDMS)

LA-UR-12-23629

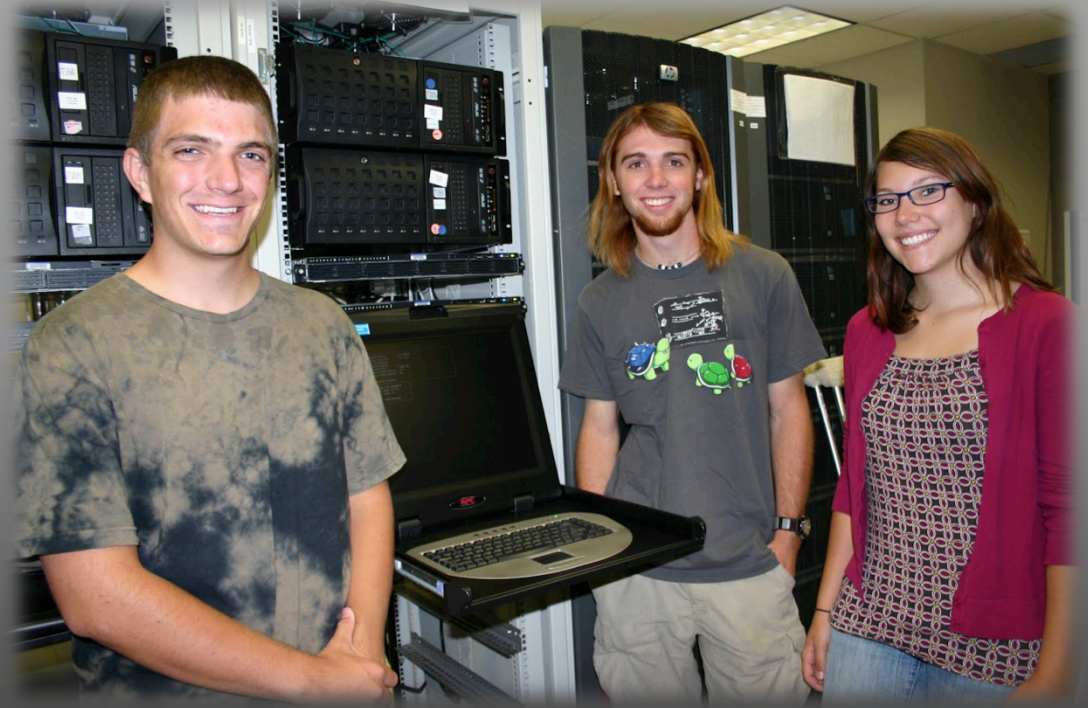
# Importance

- Need a scalable, parallel tool to monitor nodes on clusters
- New LDMS plugins need to be able to be easily added to tool

LA-UR-12-23629

# Our Cluster

- Built from the ground up
  - 8 nodes
  - Running CentOS 6.2
  - 8 Cores each
  - 16Gb RAM each
  - WareWulf-provisioned



LA-UR-12-23629

# CBTF

- CBTF stands for “Component Based Tool Framework”
- Scalable
- Adjusts to different topologies automatically
- Uses MRNet (Multicast/Reduction Network) mechanism for information transport

LA-UR-12-23629

# CBTF

- CBTF is flexible and general enough to be used for any tool that needs to do a task on many nodes
- Components are reusable and “EASILY” added to a new tool

LA-UR-12-23629

# Three Levels of CBTF

- Frontend Node
  - Interacts with user
- Filter Nodes
  - Filters or concatenates information from backend nodes
- Backend Nodes
  - Where the actual work of the tool is done

LA-UR-12-23629

# How CBTF Works

- 3 Main Files
  - Tool File – loads files, MRNet setup, interacts with CBTF network
  - Component (Plugin) File – Components are defined and given instructions
  - XML File – sets up connections between components and directs data streams

LA-UR-12-23629



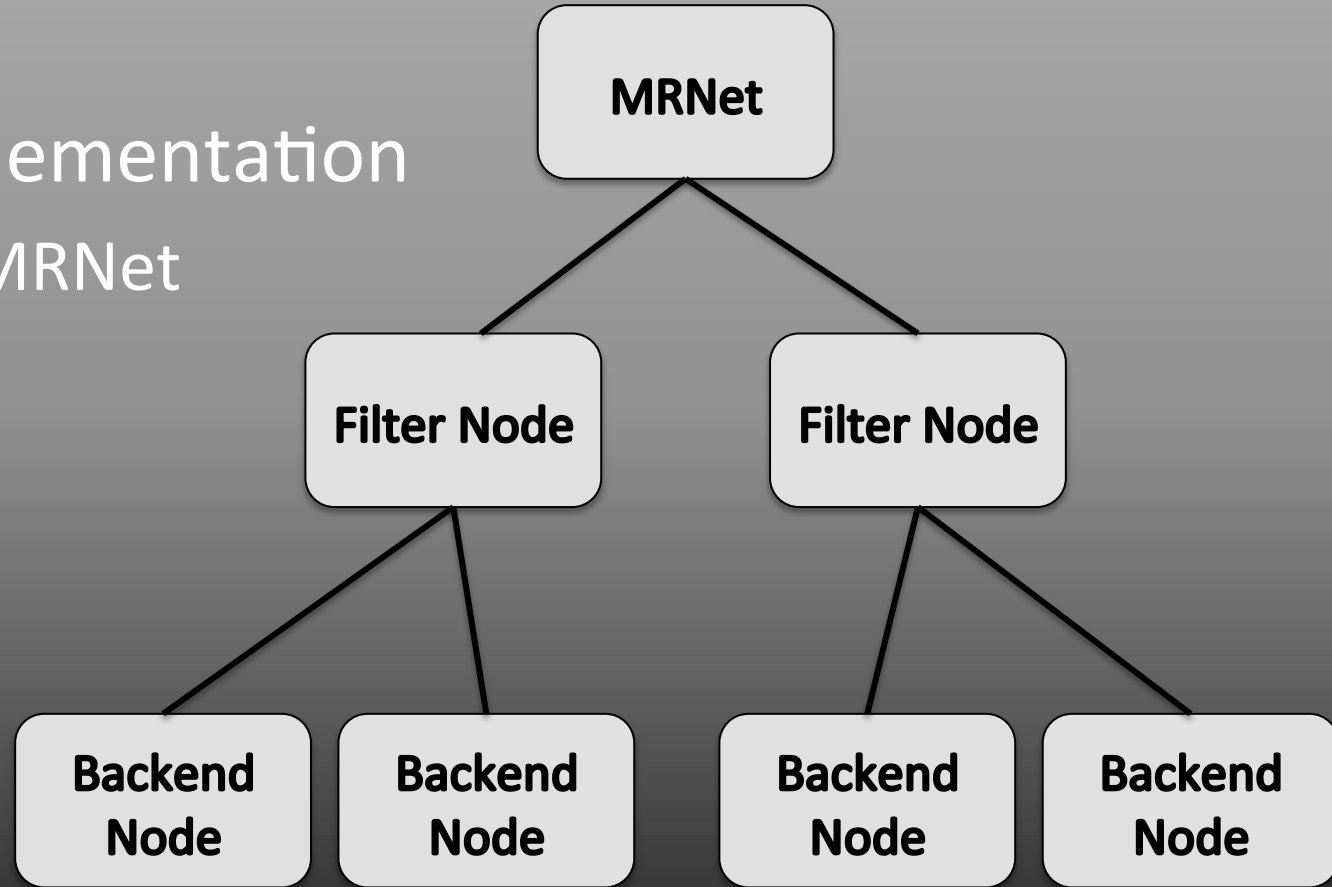
# LDMS

- LDMS stands for “Lightweight Data Metric Service”
- Tool used for monitoring nodes
  - Information from /proc/
    - Vmstat, meminfo ...
- Created an application layer
- Created Dynamic Libraries

LA-UR-12-23629

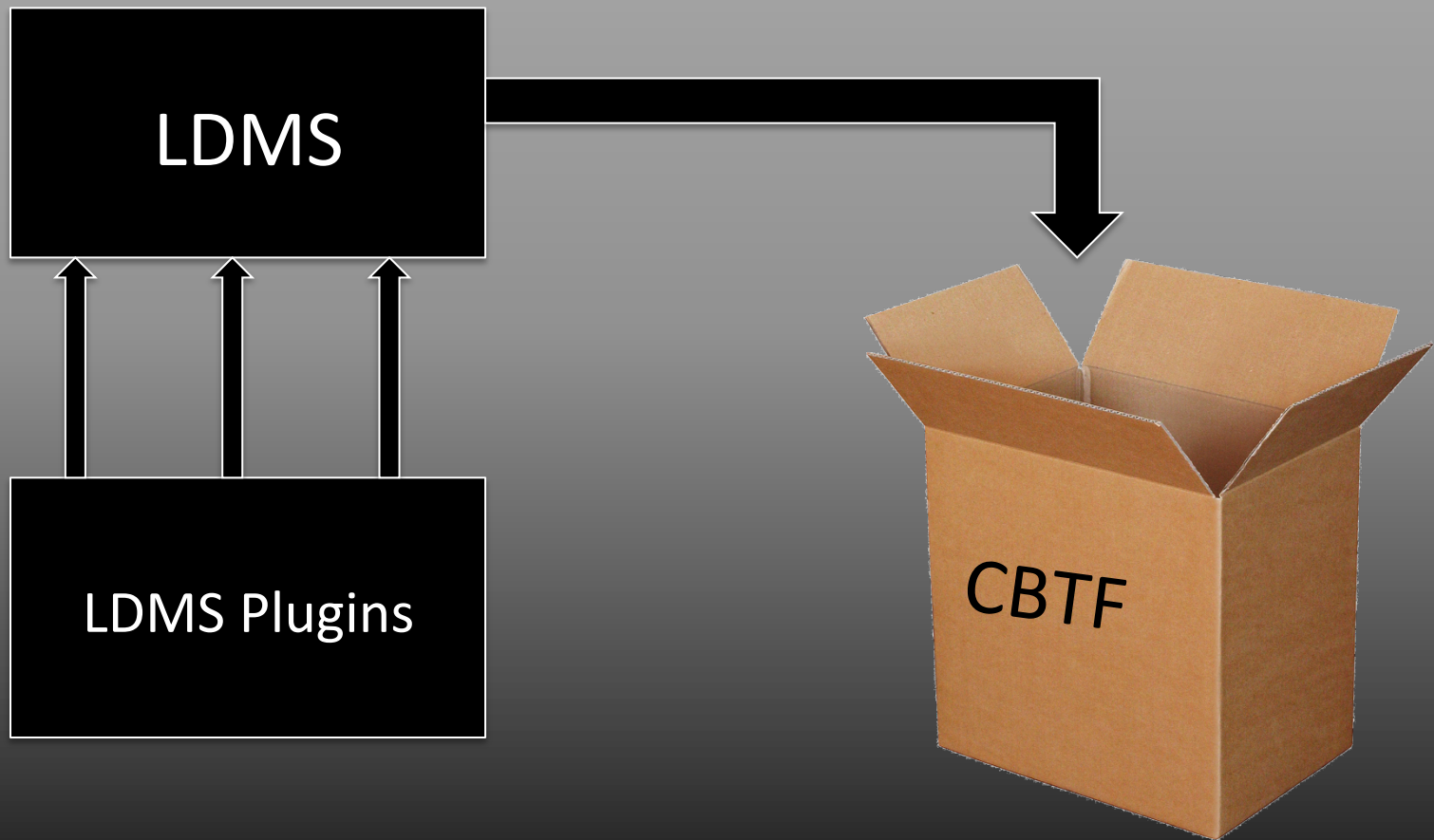
# Why Change LDMS?

- Fat tree implementation
  - CBTF uses MRNet



LA-UR-12-23629

# Implement LDMS Code Using CBTF



LA-UR-12-23629

# Ltool

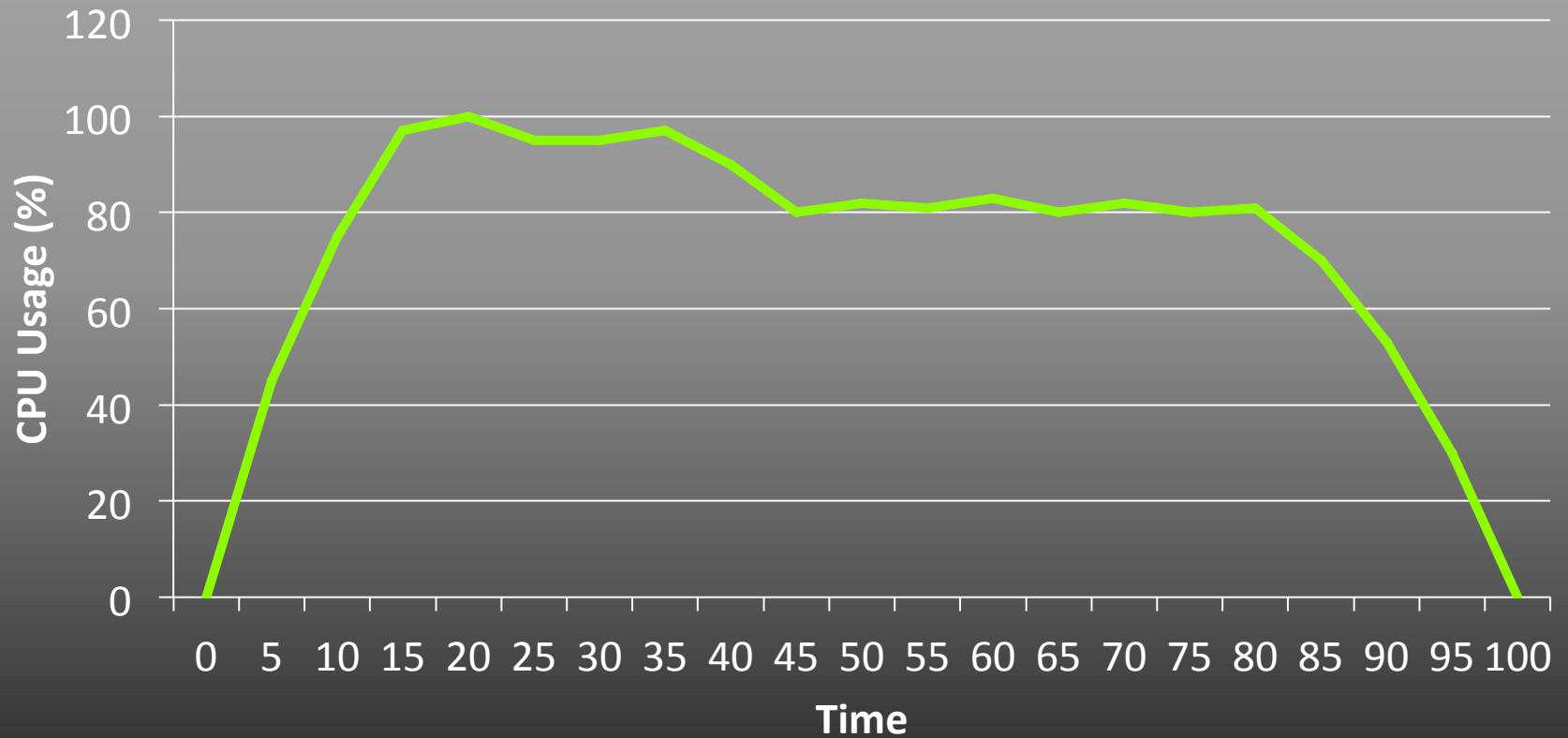
- Ltool is the name of the tool we derived from LDMS
- Dynamically linked
- Includes the following components:
  - Vmstat
  - Meminfo
  - Procinterrupts
  - ...and more

```
[cbtn@black Ltool]$ ./Ltool  
  
Ltool> VMstat  
running 'VMstat' on backends.
```

LA-UR-12-23629

# Expected Results

CPU Trend



LA-UR-12-23629

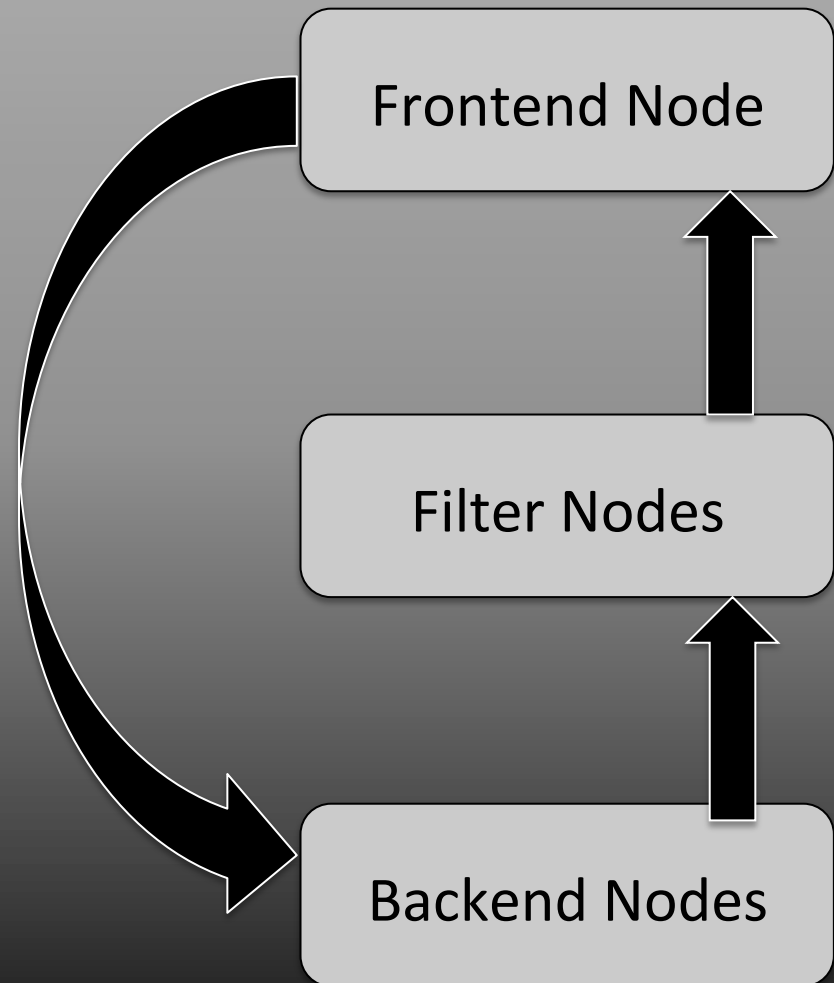
# LDMS to CBTF

- Allows new LDMS plugins to be simply transferred into our tool

LA-UR-12-23629

# How Does It Work?

- Ltool command is run on the frontend node
- Ltool collects information from the backend nodes
- Backend nodes send information to the filter nodes
- Filter nodes concatenate information and send to a database on the front end node



LA-UR-12-23629

# CBTF - Ltool

- Frontend Node
  - Allows user to select plugin to run
- 2 Filter Nodes
  - 1) Collects X amount of messages from each node
  - 2) Adds the information to the database
- Backend Node
  - Runs LDMS application
  - Runs on a specified interval

LA-UR-12-23629

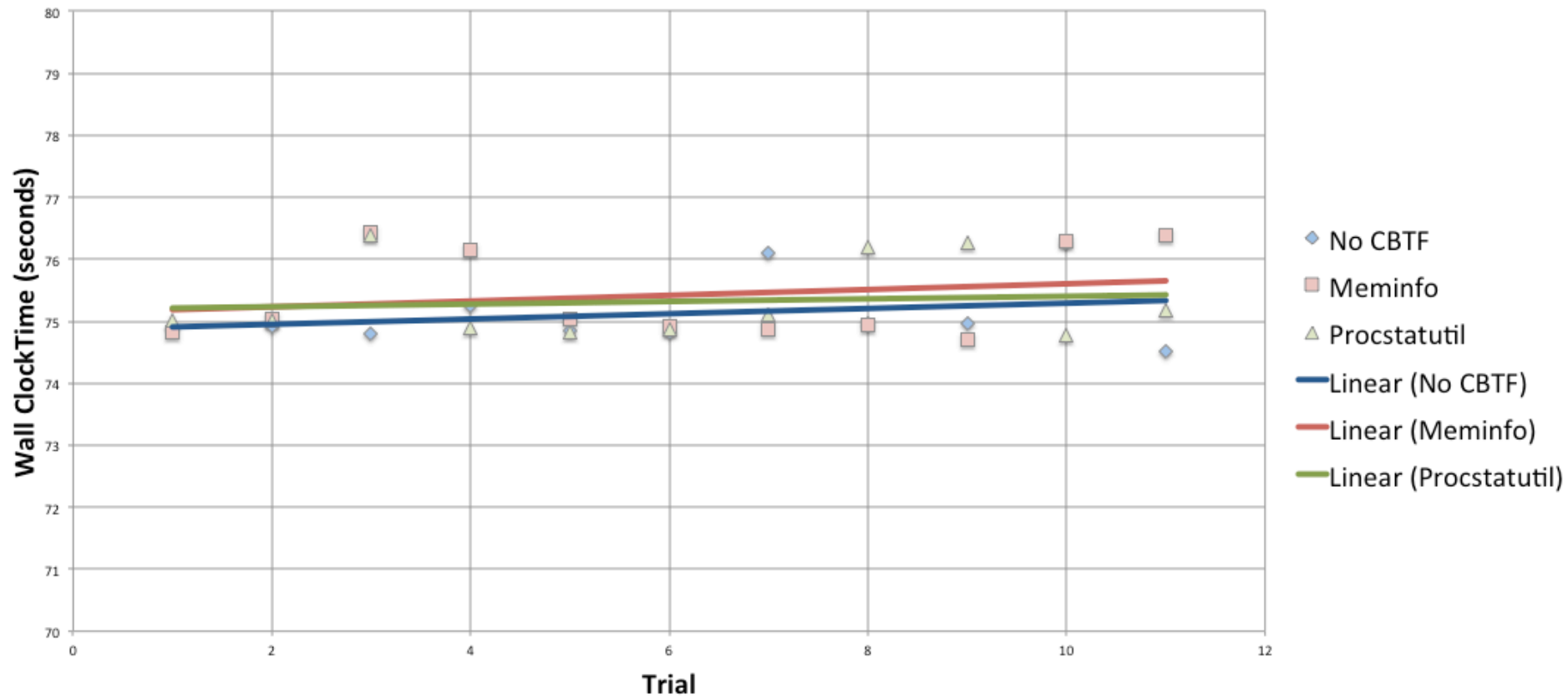


# Database

- MySQL
  - Each plugin has its own table
    - Allows to query by plugin
    - Simplifies future additions
  - Information connected by MySQL dynamic script
    - Allows to be easily moved to new machines / users
      - No predatabase information needed

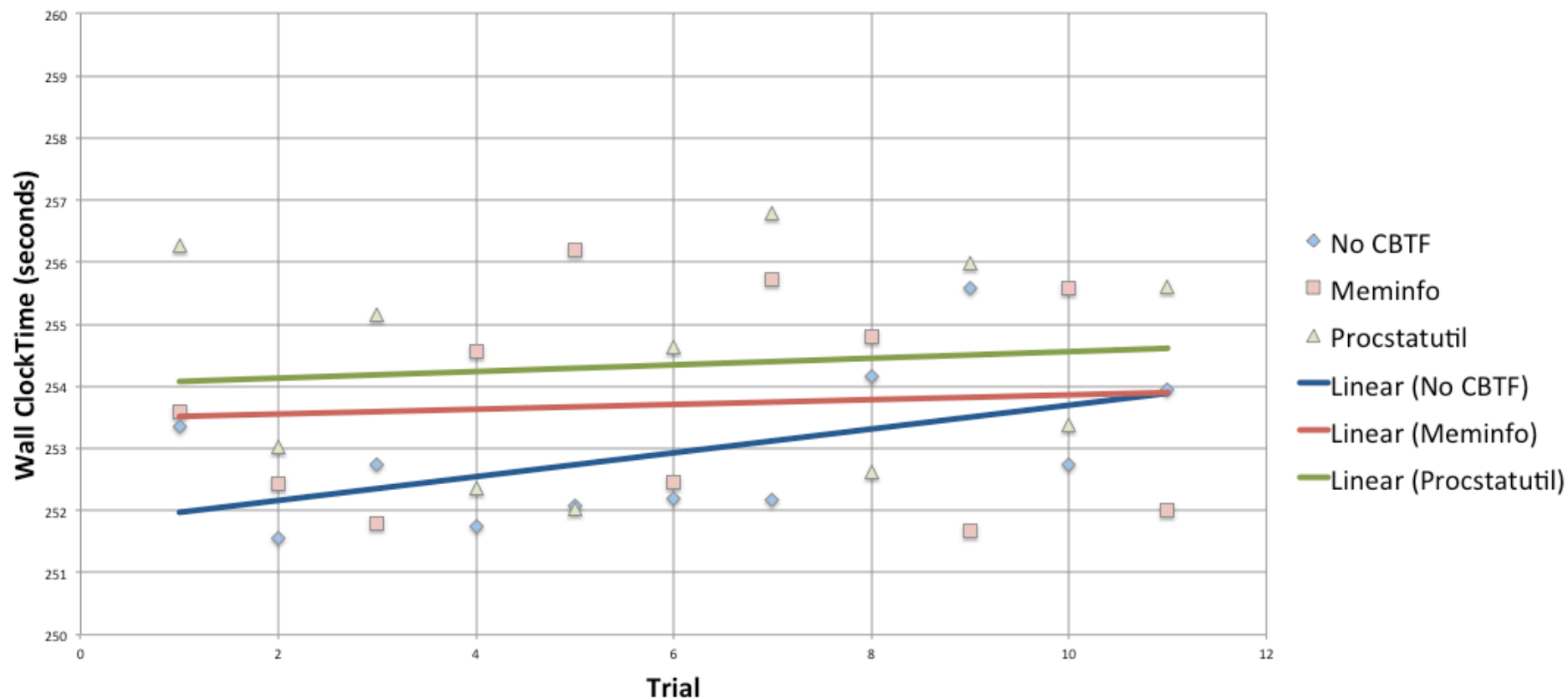
LA-UR-12-23629

## Lttool Overhead (Problem Size = 100)



LA-UR-12-23629

## Ltool Overhead (Problem Size = 150)



LA-UR-12-23629

# Results

- LDMS was successfully implemented into a CBTF tool, and that the overhead involved with running the tool is relatively low.

LA-UR-12-23629

# Conclusions

- Ltool is a useful tool when it comes to monitoring nodes on a cluster because the overhead involved with running the tool is not particularly high and it will automatically scale to any size cluster.

LA-UR-12-23629

# Future Work

- Ltool can be tested to see if it can run consistently for numerous days
- Create triggers for code to run with Ltool to allow multiple components to run at different times
- Multiple components executing at the same time
- Use MRNet to filter more data



LA-UR-12-23629

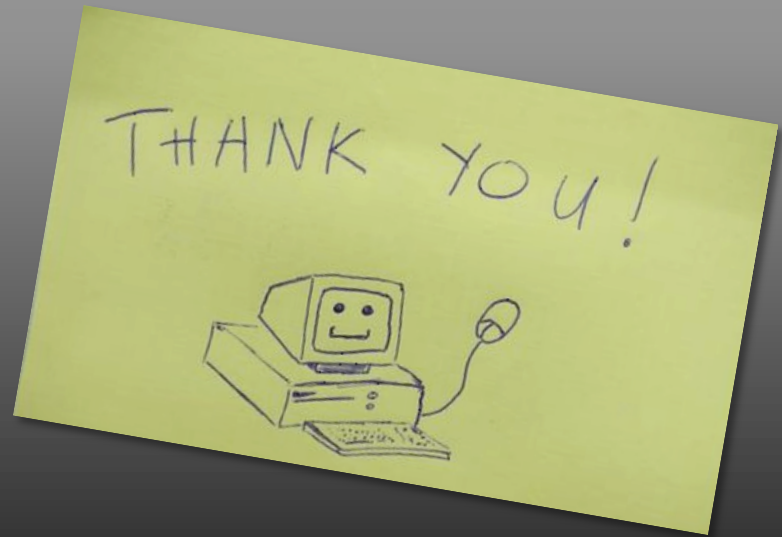
# Any Questions?



LA-UR-12-23629

# Thank You

- Dane Gardner
- Our Mentors: Mike Mason, Jon Bringhurst
- TJ Machado
- Jim Brandt
- Ann Gentile



LA-UR-12-23629