Experiments in Integrated Data Center Power and Cooling Analysis

Shaun Brandt, Hema Kumar, Michael Novak, Mandar Patil, Yu Yang

Portland State University

About the Project

- The need for large scale computing continues to rise, and companies are turning to data centers to meet those needs

- Large clusters of servers are expensive to run, and expensive to cool

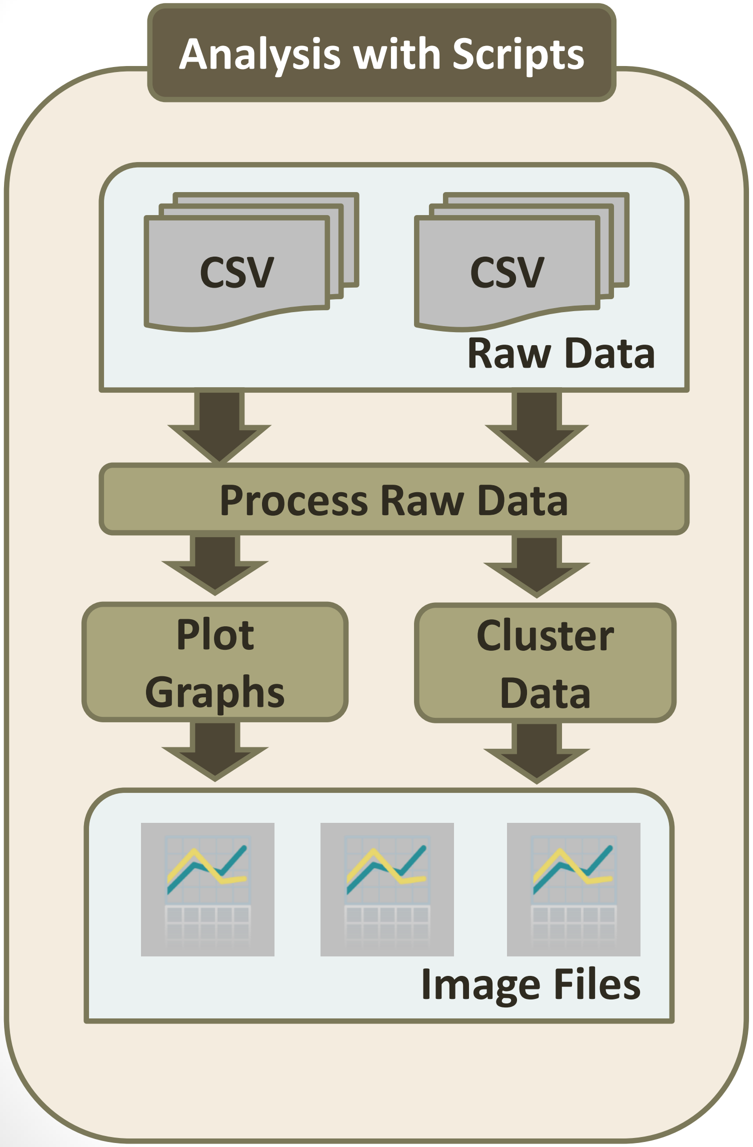
- Cooling devices (especially chillers) consume lots of power and are subject to physical wear

- Data center operators need access to as much information as possible!

About the Data

- Data for this project was collected from a number of sensors at the PNNL Energy Smart Data Center
- 8 server racks (1 network, 2 air cooled, 5 liquid cooled)
- Two application loads (high density / low density)
- Three different application test runs

Project Design



Analysis with Scripts

CSV

CSV

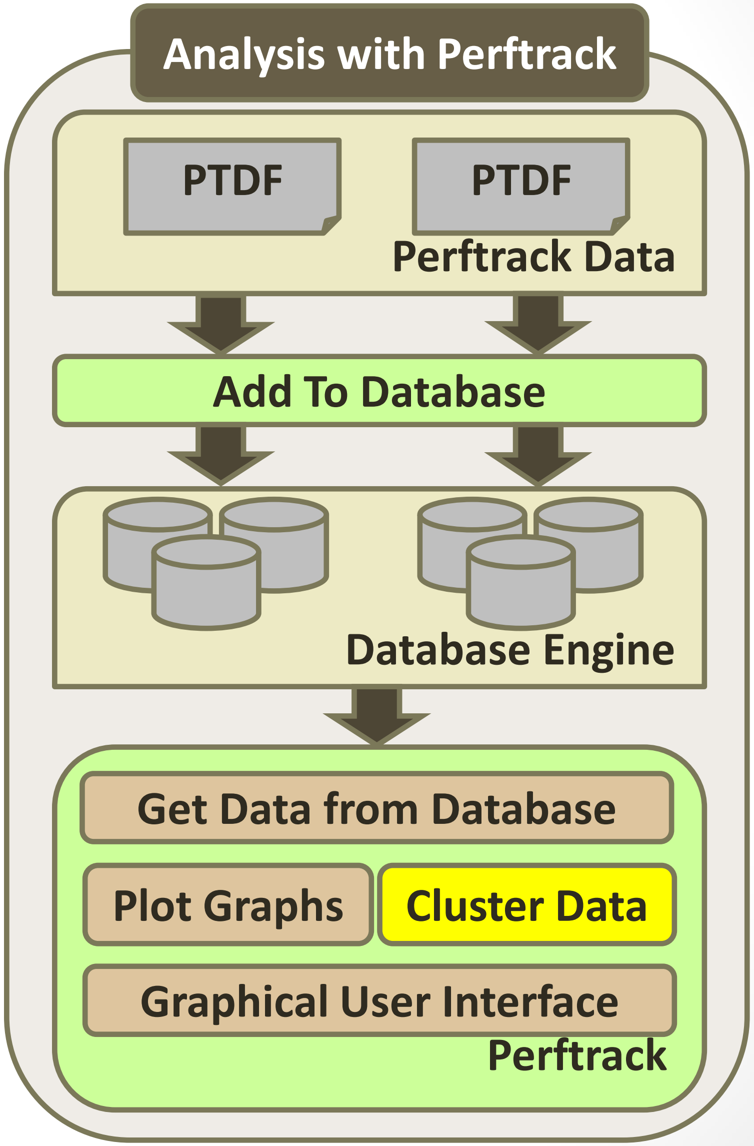
Raw Data

Process Raw Data

Plot Graphs

Cluster Data

Image Files



Analysis with Perftrack

PTDF

PTDF

Perftrack Data

Add To Database

Database Engine

Get Data from Database

Plot Graphs

Cluster Data

Graphical User Interface

Perftrack

Approach 1: Analysis with Scripts

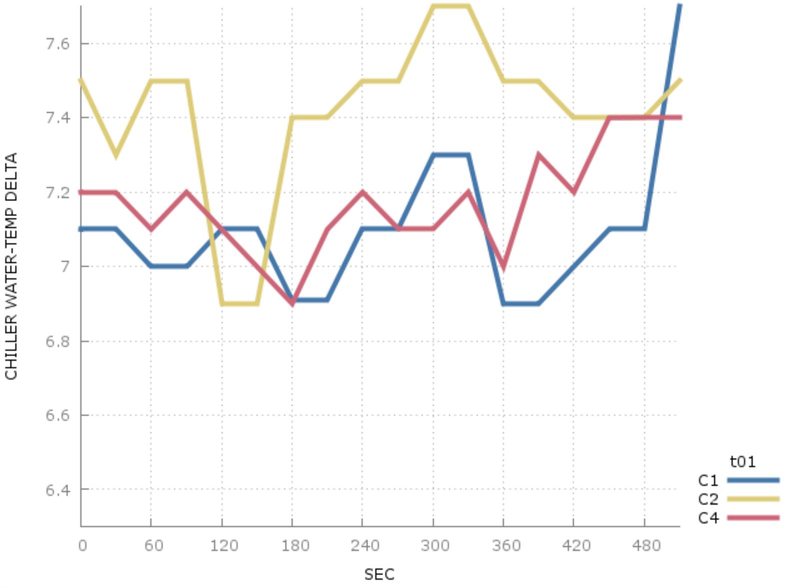
- A set of custom tools was written to work with sensor data provided in CSV format
- Mostly Haskell with calls to Gnuplot (earlier prototypes also used C, Perl and shell)
- Generates a complete report including tables and graphs
- Calculates and graphs metrics such as temperature deltas, COP (coefficient of performance) and power usage

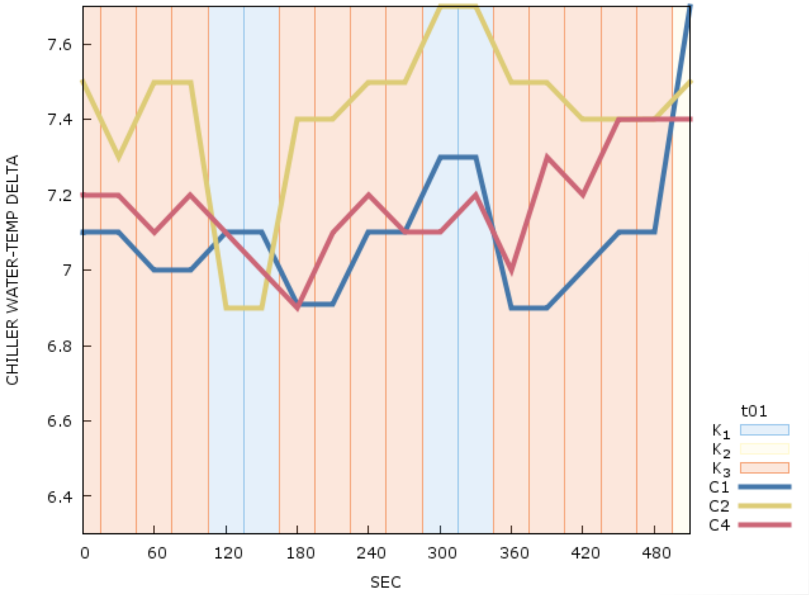
Analysis with Scripts (cont.)

- Cluster analysis based on output of metric calculations
- Our work uses ELKI, an open source data mining tool with many customizable options (such as choice of clustering algorithms)
- We tried clustering on various metrics, including chiller water temperature deltas, coefficient of performance (COP) and cooling load
- We developed a method to overlay cluster information over previously generated graphs

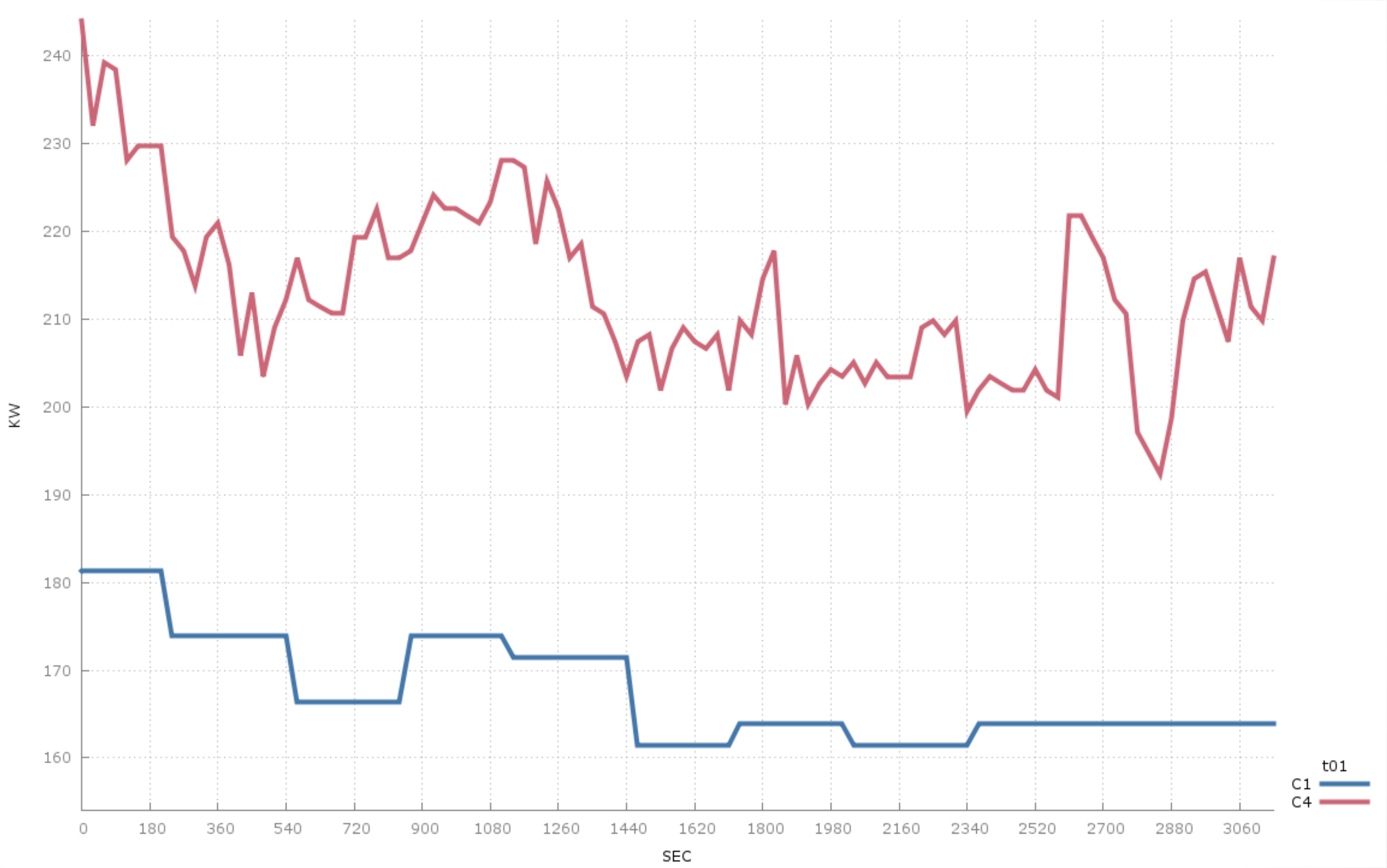
Time-correlated Cluster Charts

- On the left – low density chiller temperature delta chart
- On the right – the associated time-correlated chart

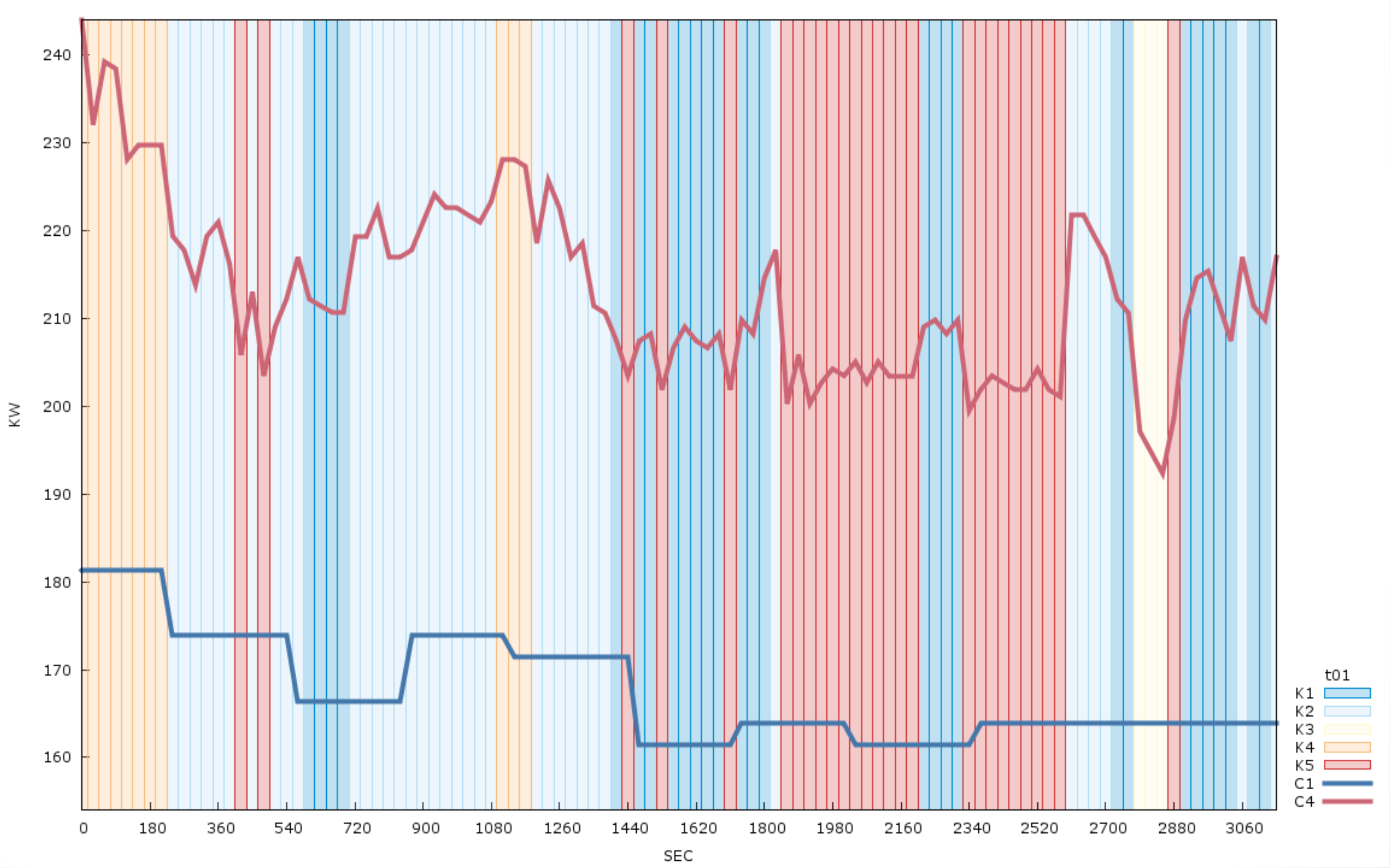




Chiller Power (KW) Chart, HD Data



Associated Time-correlated Cluster Chart



Perftrack: Data Analysis Tool

- Data store and interface for managing performance data
- Compares the data collected in different locations and formats in single performance analysis session
- Includes interfaces to data store and scripts for automatically collecting data describing each experiment
- Includes a graphical user interface
- Written in C++, QT and Python

Approach 2: Perftrack Modifications

- ELKI: Open source, Java-based clustering package
- New menu item ‘Cluster Data’ is added to Perftrack
- Perftrack is modified to generate the ELKI-compatible raw data from user filtered experiment data using GUI
- K-Means clustering algorithm is executed on ELKI-compatible raw data using ELKI package
- Perftrack displays the clusters

Future Work

- Need more (and more accurate) data!

- Perftrack:
  - Full integration of script work
  - On-the-fly data generation for ELKI
  - Tightly-coupled integration of ELKI
  - Graphical user interface enhancement

- Modification of scripts to support direct connection to FRED in place of CSV file access

- Experiment with motif-based cluster analysis (this would require more data)