Data Driven Investigation of Faults in HVAC Systems

Comparative data mining in energy systems

Balakrishnan (Murali) Narayanswamy, Bharathan Balaji, Rajesh Gupta, Yuvraj Agarwal

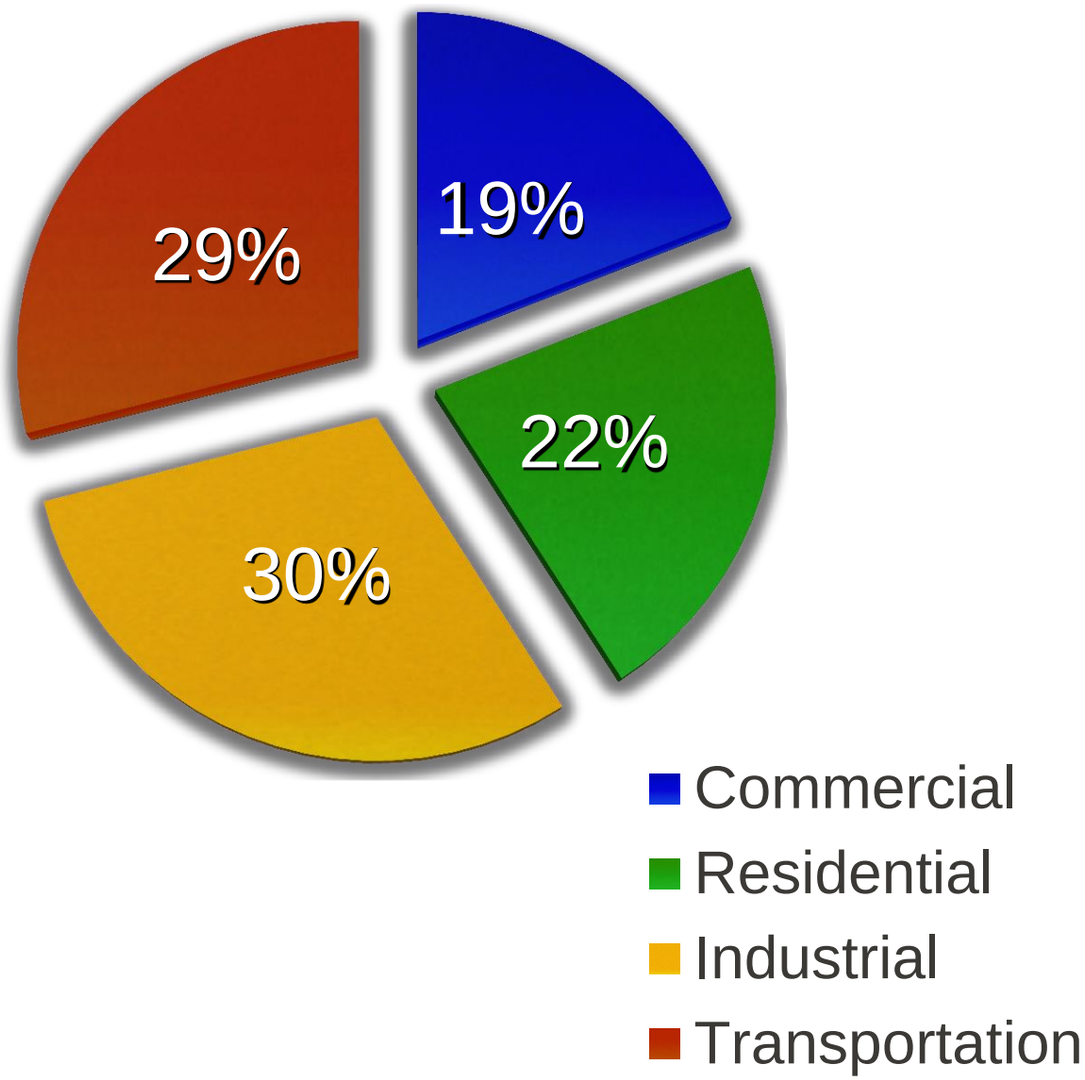
University of California, San Diego *Carnegie Mellon University*

synergy
systems, networking and energy efficiency

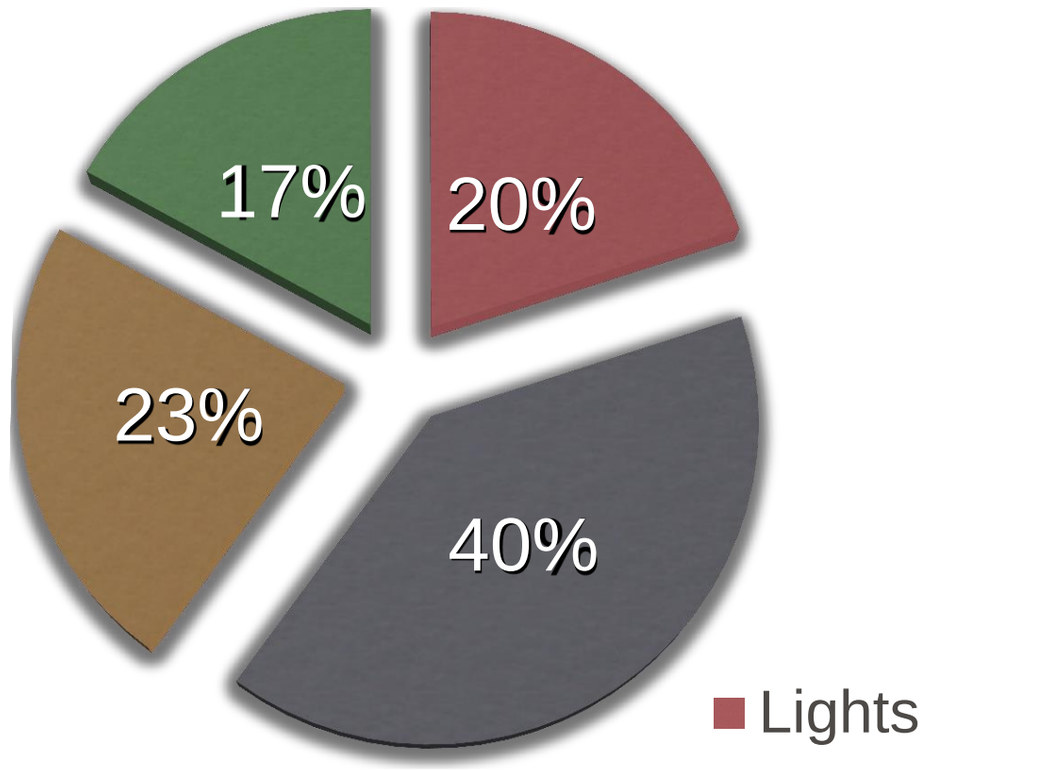
MESL
Microelectronic Embedded Systems Laboratory

Building HVAC is important!

US Energy Consumption

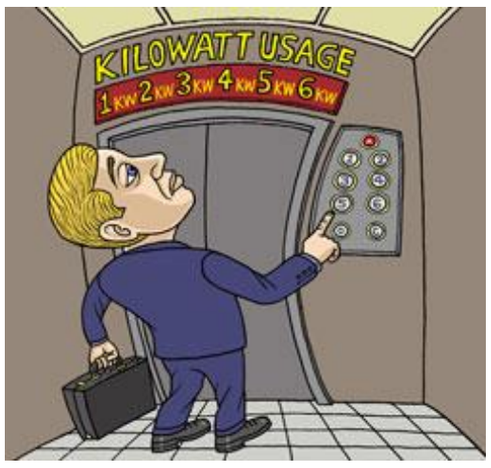


Commercial Buildings Energy Breakdown

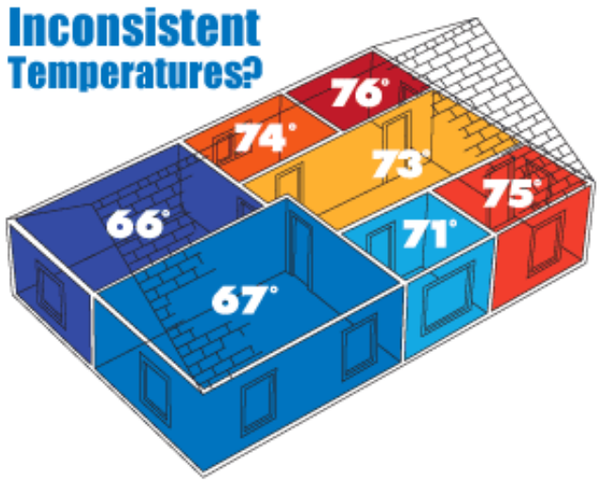


Buildings are already complex

Many thousands of lights

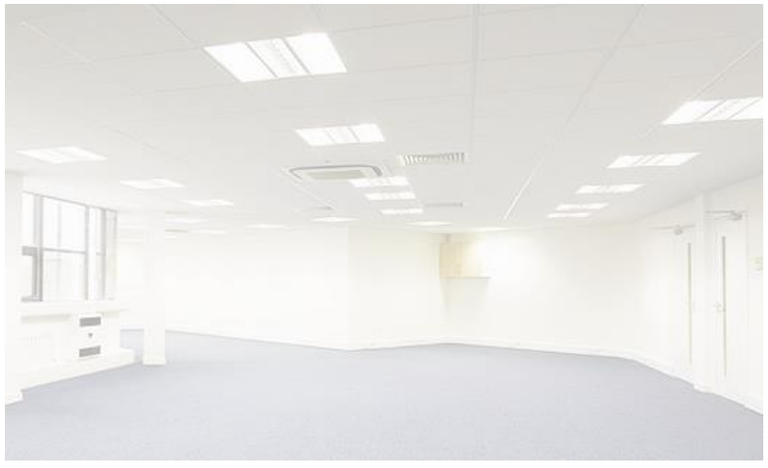


Energy hogging elevators



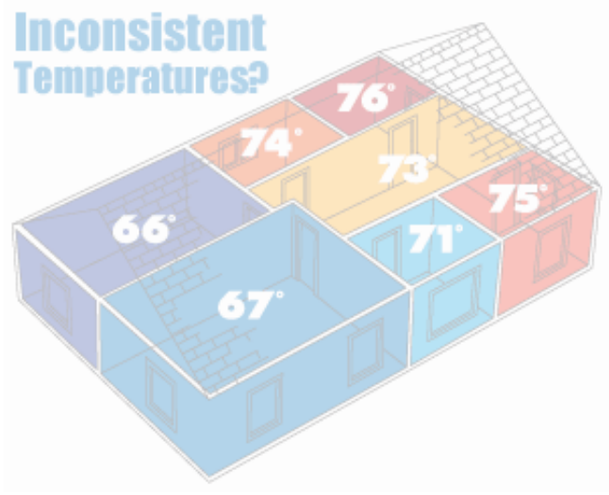
Strong requirements

Buildings are already complex

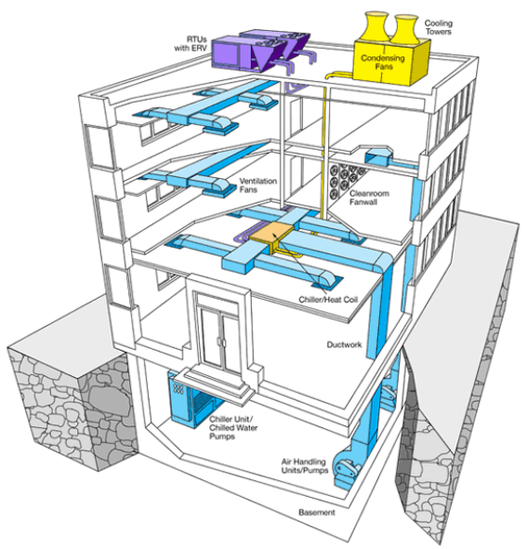


Many thousands of lights

Energy hogging elevators

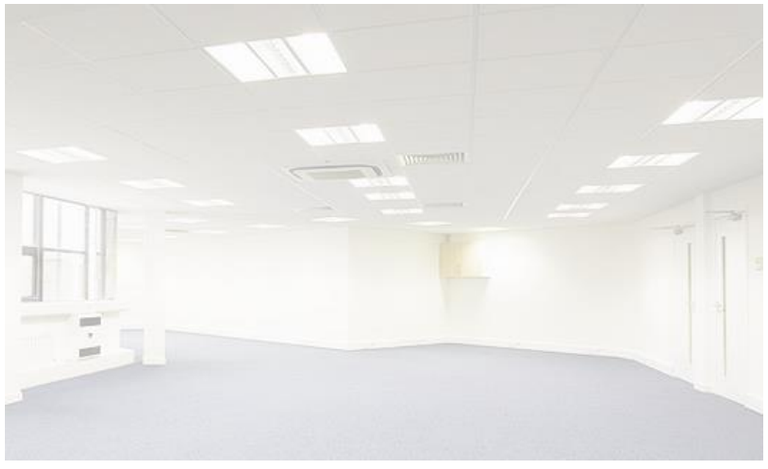


Strong requirements



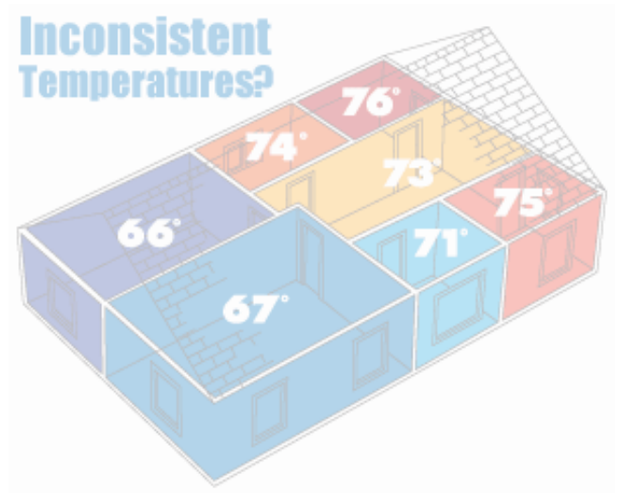
Complex HVAC system with many moving parts

Buildings are already complex

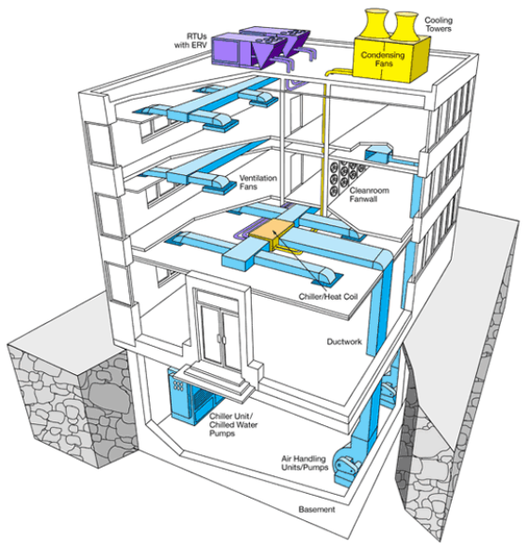


Many thousands of lights

Energy hogging elevators



Strong requirements



Complex HVAC system with many moving parts

many 10s of large fans

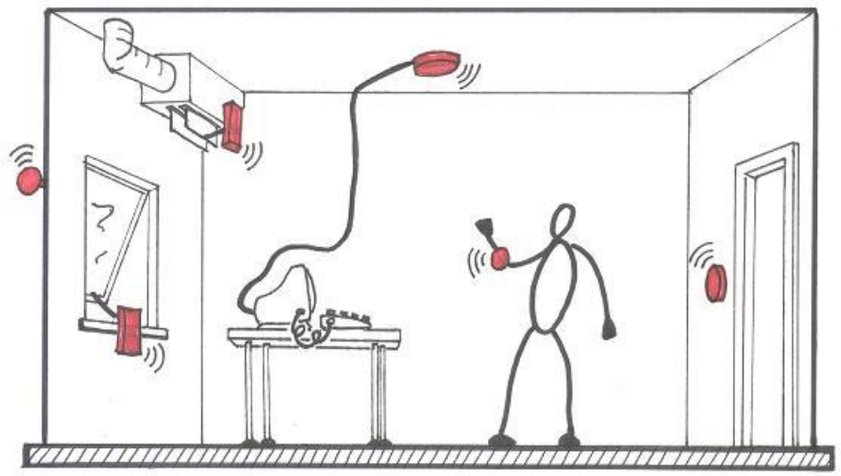
many 100s of dampers

10s of large pumps

Buildings are becoming smarter

Better sensed

More controllable



Smart and personalized

Buildings are becoming more complex

UCSDCSE
Computer Science and Engineering

Built in 2004, 145,000 sq ft, 5 floors

HVAC : VAV with reheat coil, 237 zones

Occupants : Faculty, staff and students

17+ sensors per thermal zone

100s of Air Handler (AHU) sensors

more than 1 year data, every ~5 minutes

new

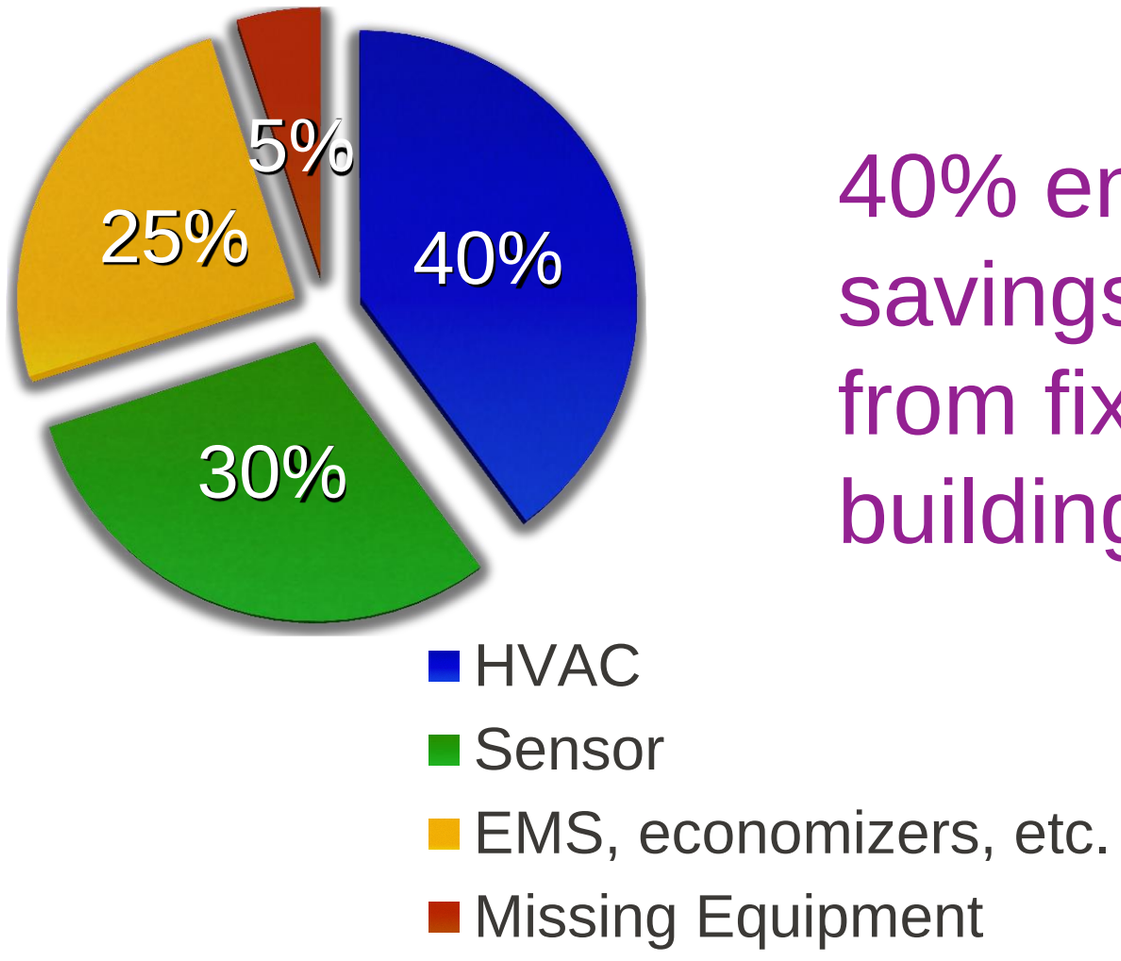
not very big

~5000 sensors
More Complexity!

DATA!!

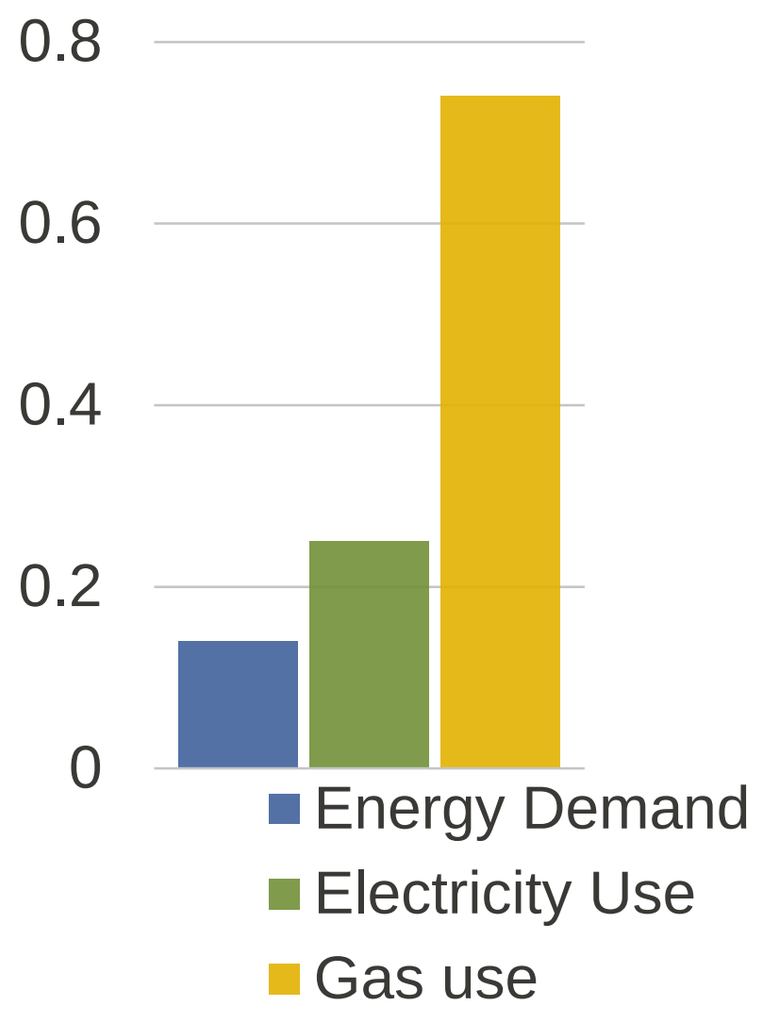
Complexity ↑ Faults ↑

More than 50% of buildings have control faults [1]



40% energy savings potential from fixing building faults [2]

Verified savings [3]

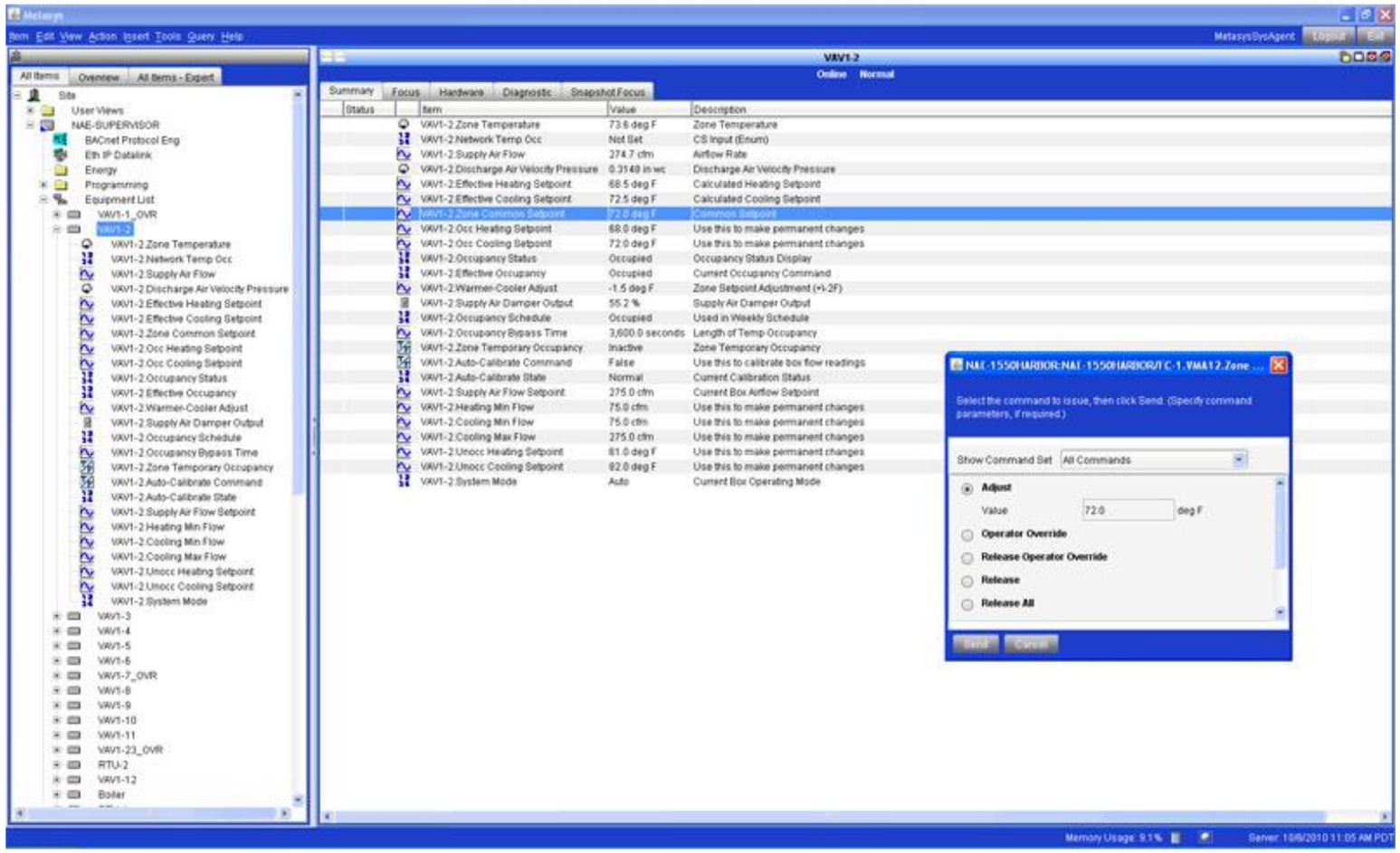


[1] M.A.Piette, et.al. Quantifying energy savings from commissioning: preliminary results from the pacific northwest. In Proc. of National Conference on Building Commissioning, 1994.
[2] J. Schein,et. al.. A rule-based fault detection method for air handling units. Energy and Buildings, 38(12):1485–1492, 2006.
[3] Selch, M. and J. Bradford (2005). Recommissioning Energy Savings Persistence. National conference on building commissioning.

Solution : Data aware energy use!

Fault Detection and Diagnostics (FDD)

Building Management Systems



More than 10,000 alarms per day on the UC San Diego campus.

How do we bridge this gap between theory and practice?

Prior work

Rule based

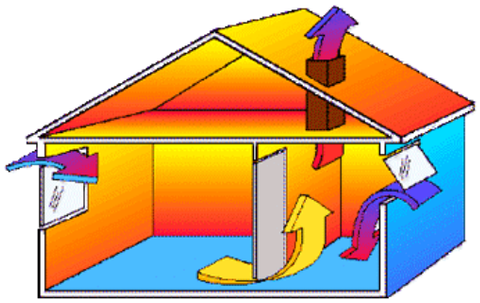
If  but still  then 

Rules : Building and climate specific

Models : Hard to build

Change based : Don't handle long standing faults

Model based

If  then I expect 

but I got  so 

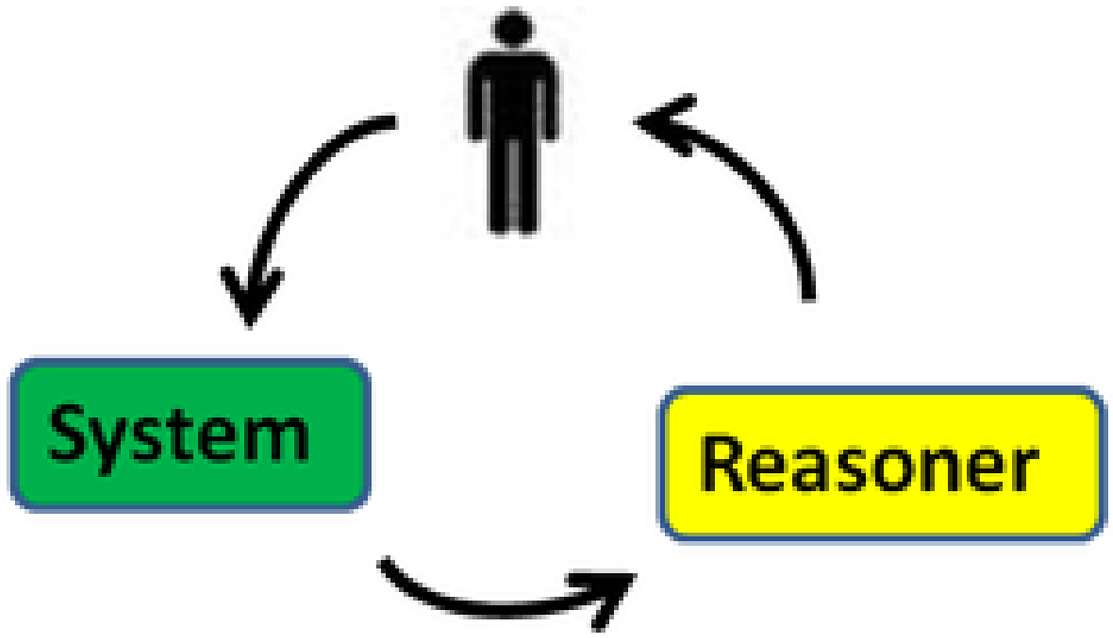
Change based

The last time I did  sensors told me I got 

but this time  so 

What this talk is about

- Qualitative evaluation of different FDD methods for Variable Air Volume (VAV) boxes
- To locate long standing faults we use comparative data mining
  - Comparing black-box zone models across zones can find anomalous zones - MCC
- Intelligent Rules based on data driven fault exemplars can get the best of both worlds (data and model driven)



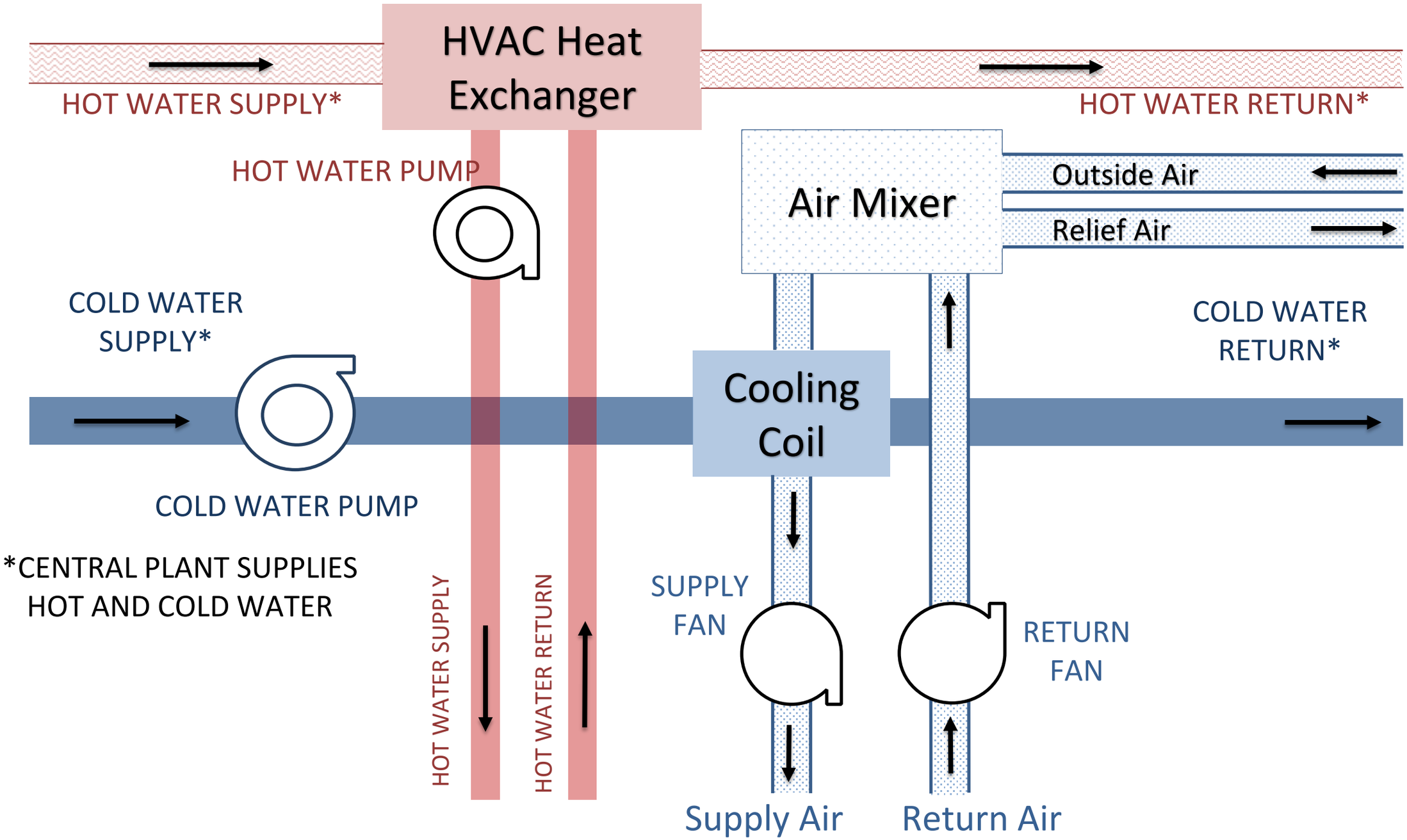
Outline

- Need for building FDD
- Background on building HVAC
- Analysis of what a solution would look like
- Our Approach (Model Cluster Compare - MCC)
- Evaluation
- Research Questions

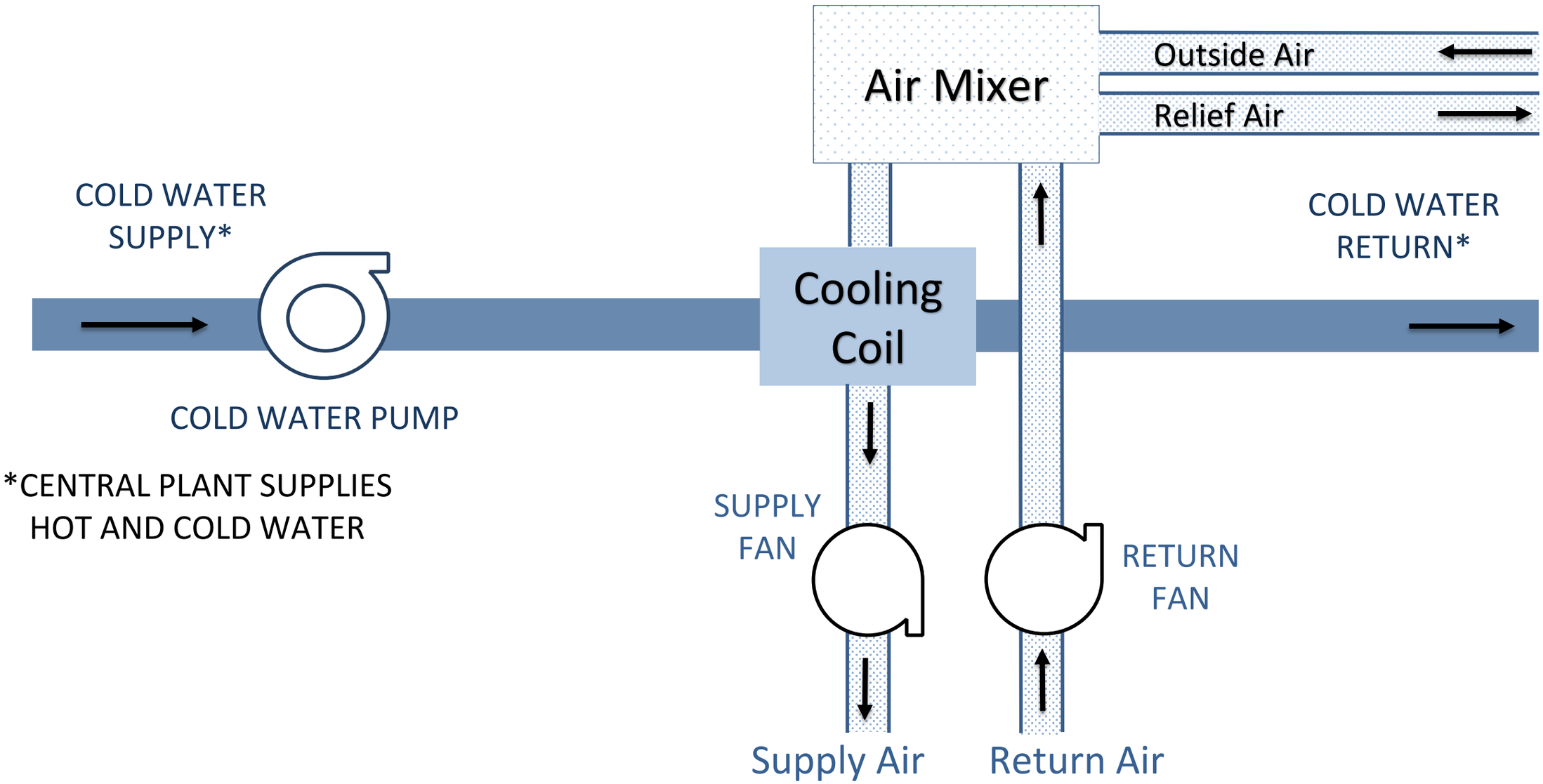
Outline

- Need for building FDD
- Background on building HVAC
- Analysis of what a solution would look like
- Our Approach (Model Cluster Compare - MCC)
- Evaluation
- Research Questions

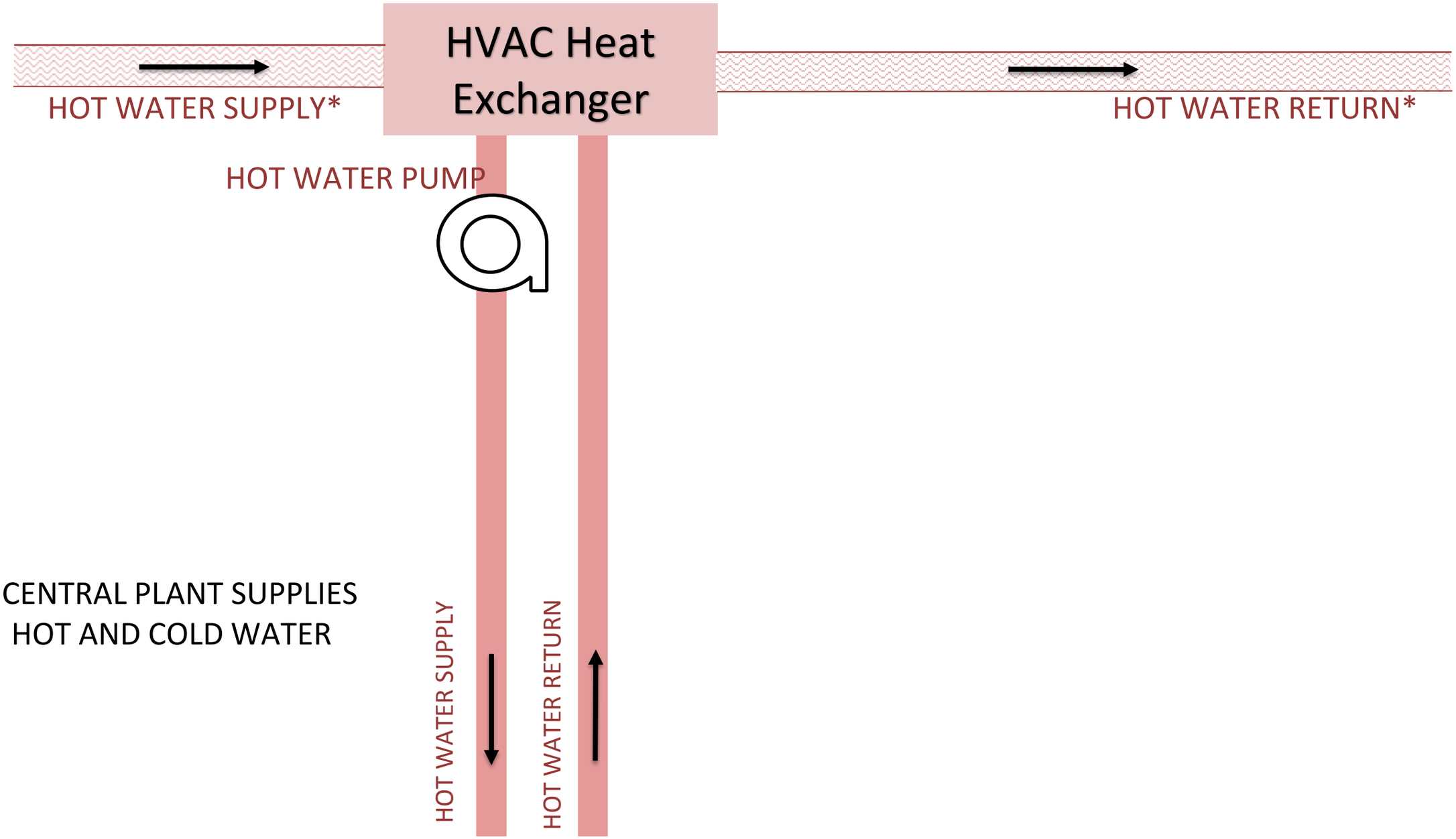
Background : HVAC system



Background : HVAC system

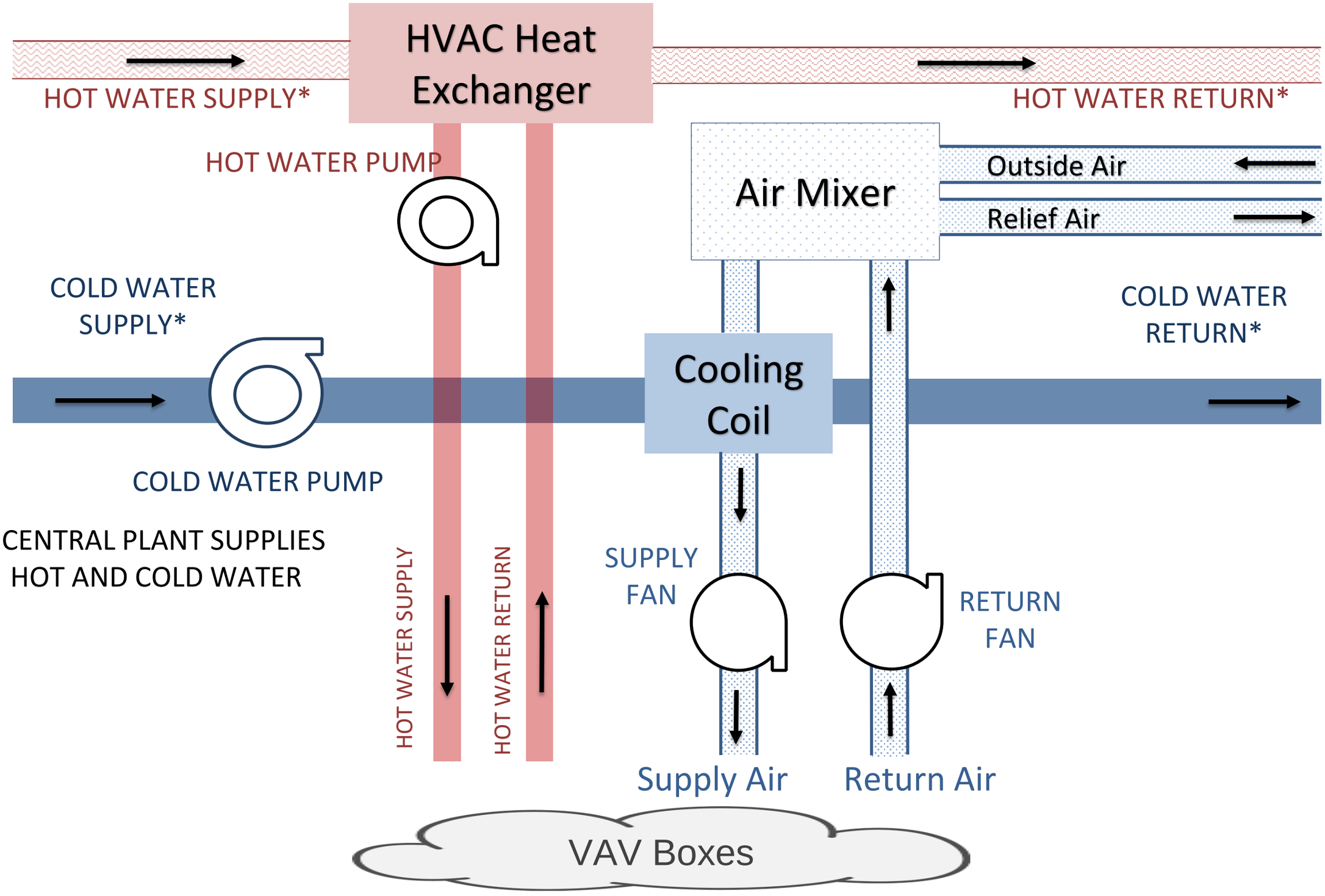


Background : HVAC system

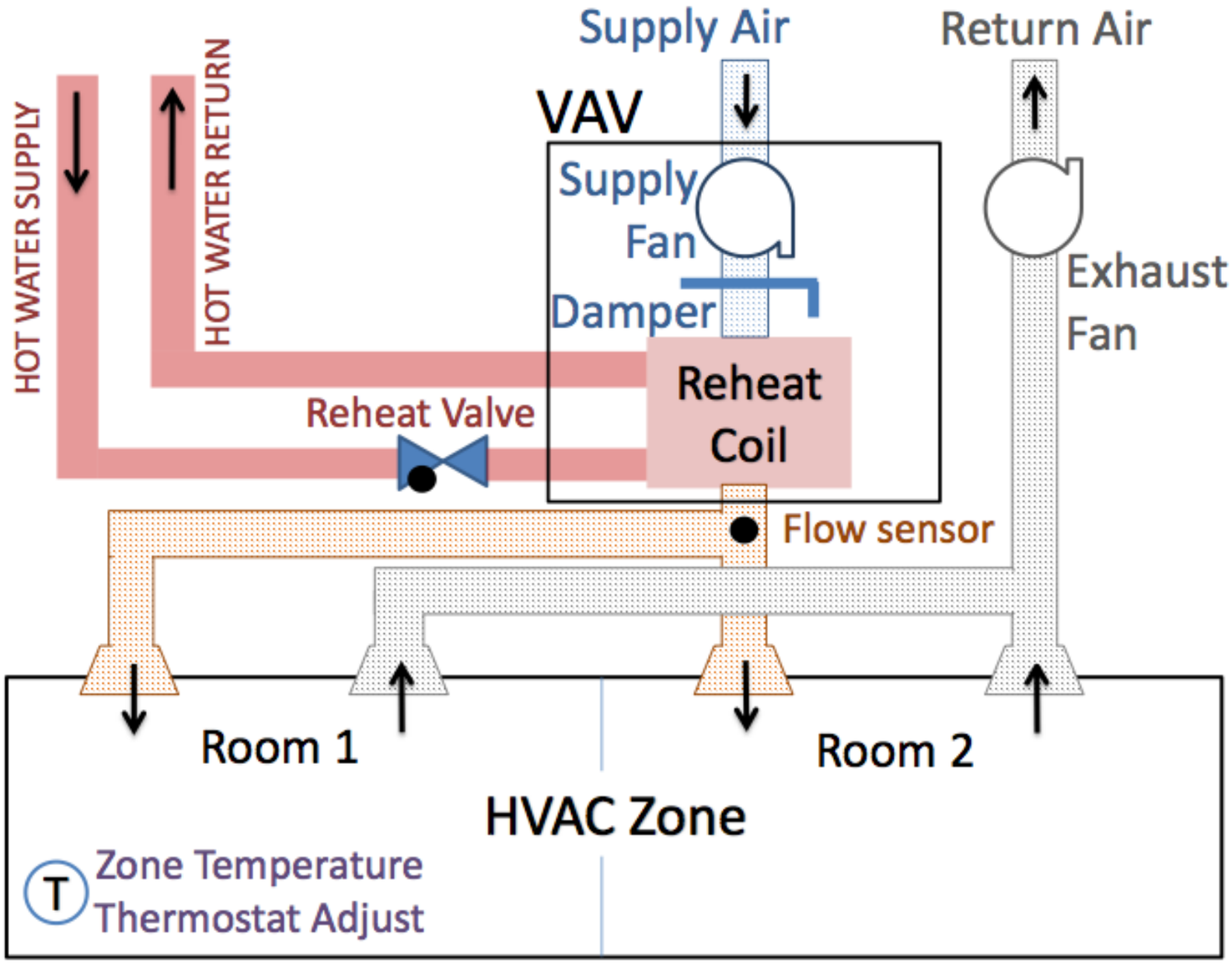


*CENTRAL PLANT SUPPLIES HOT AND COLD WATER

Background : HVAC system

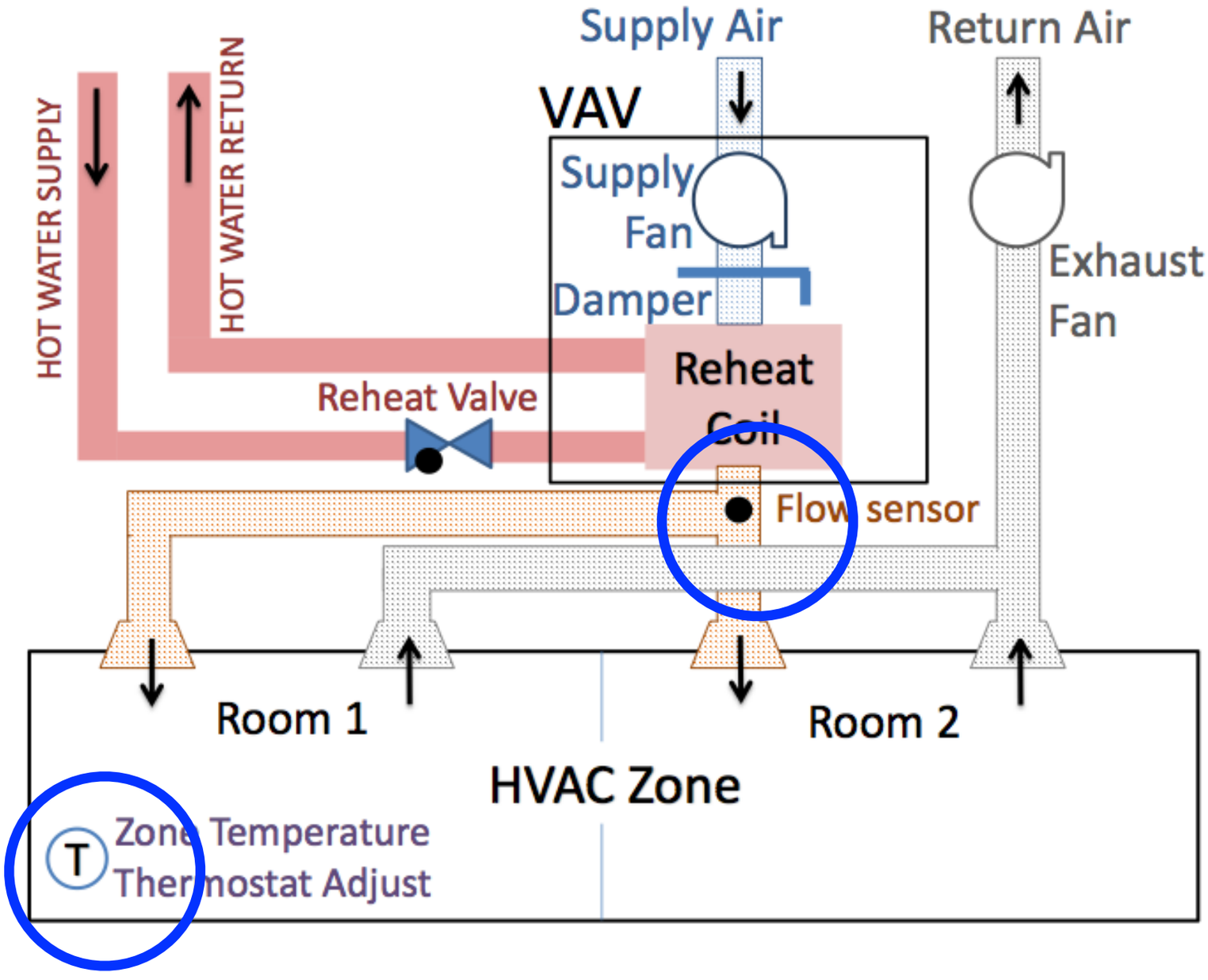


Background : VAV unit



Zone specific flow control devices

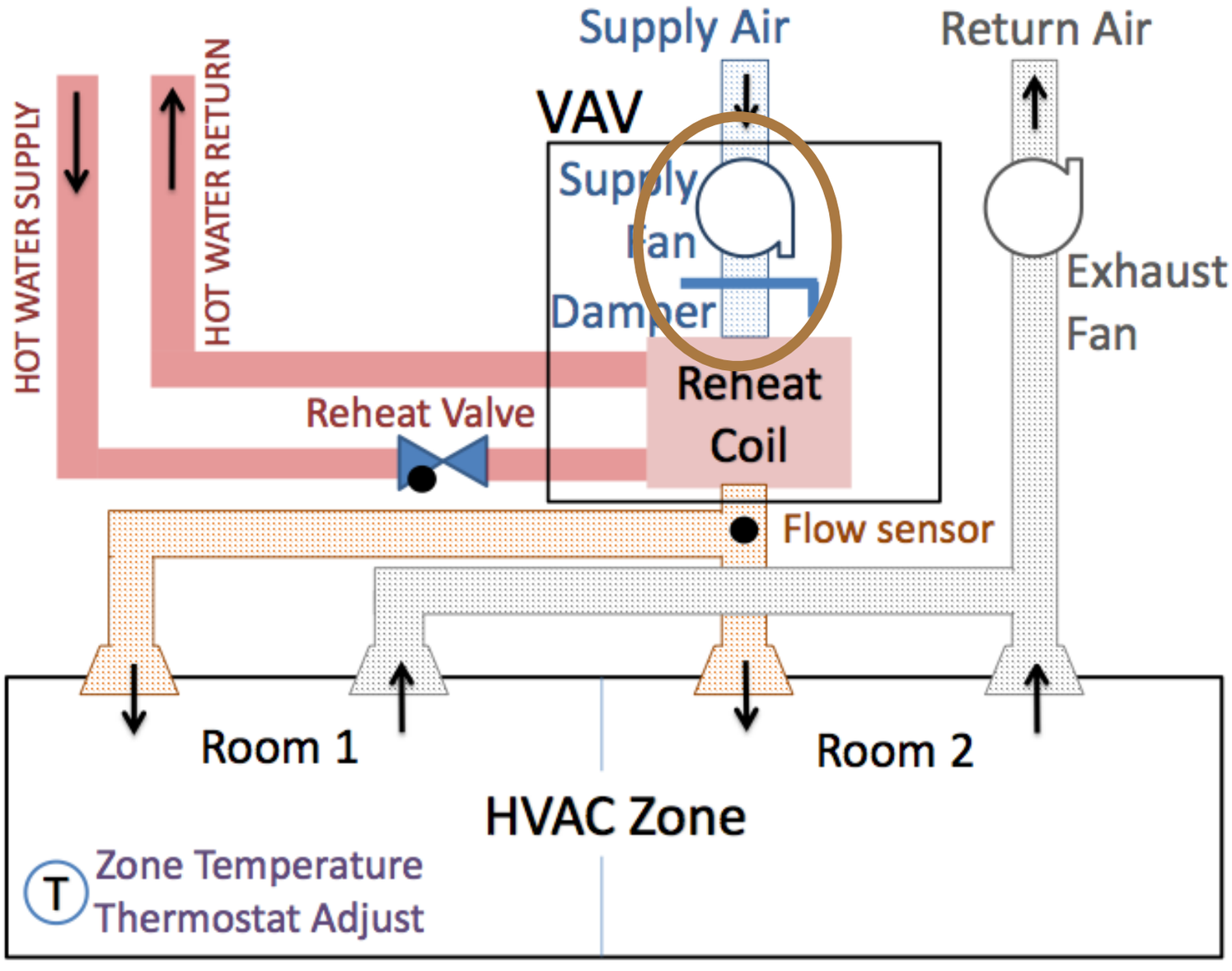
Background : VAV unit



Zone specific flow control devices

Sensors : Temperature, flow

Background : VAV unit

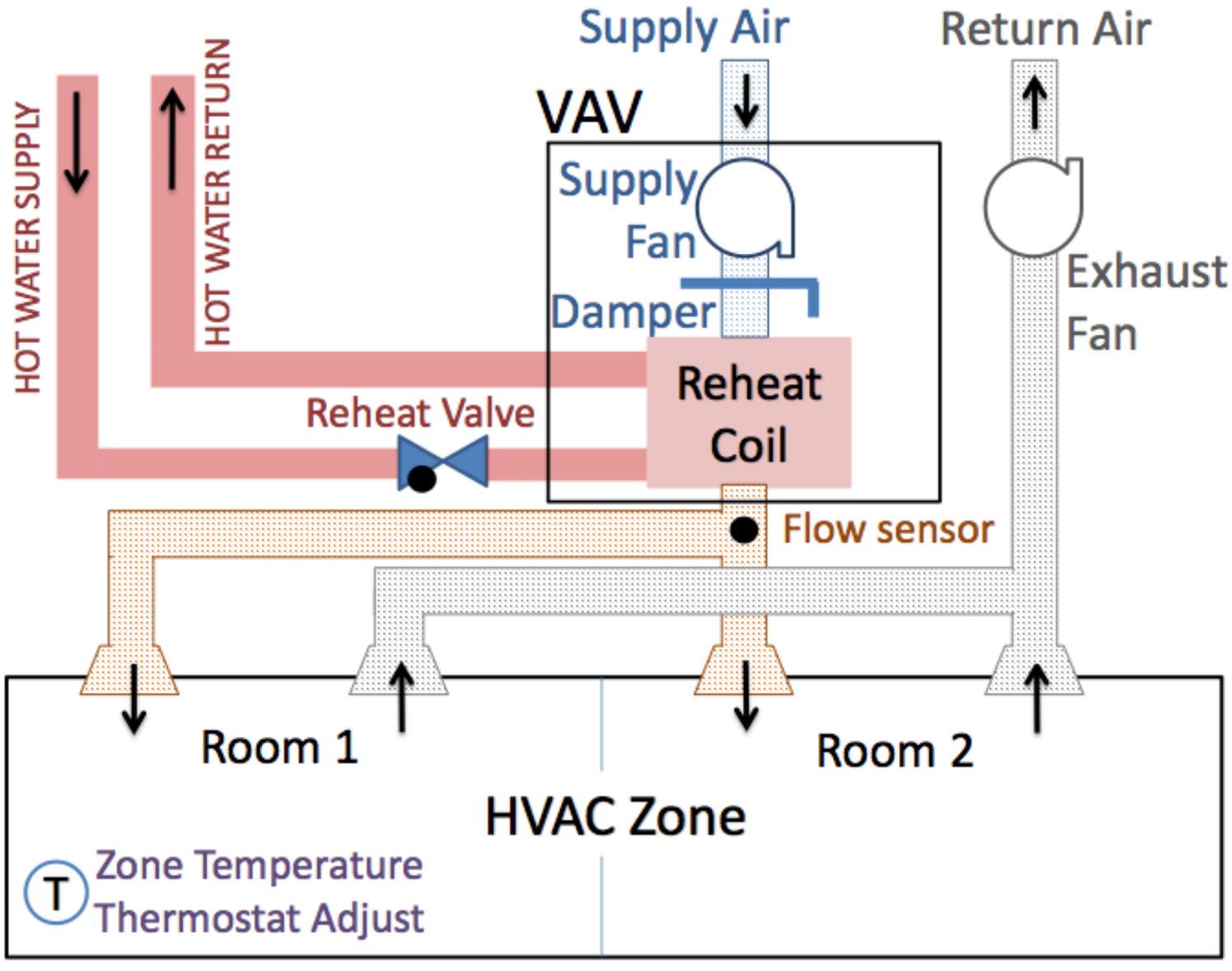


Zone specific flow control devices

Sensors : Temperature, flow

Actuators : Dampers, fans

Background : VAV unit



Zone specific flow control devices

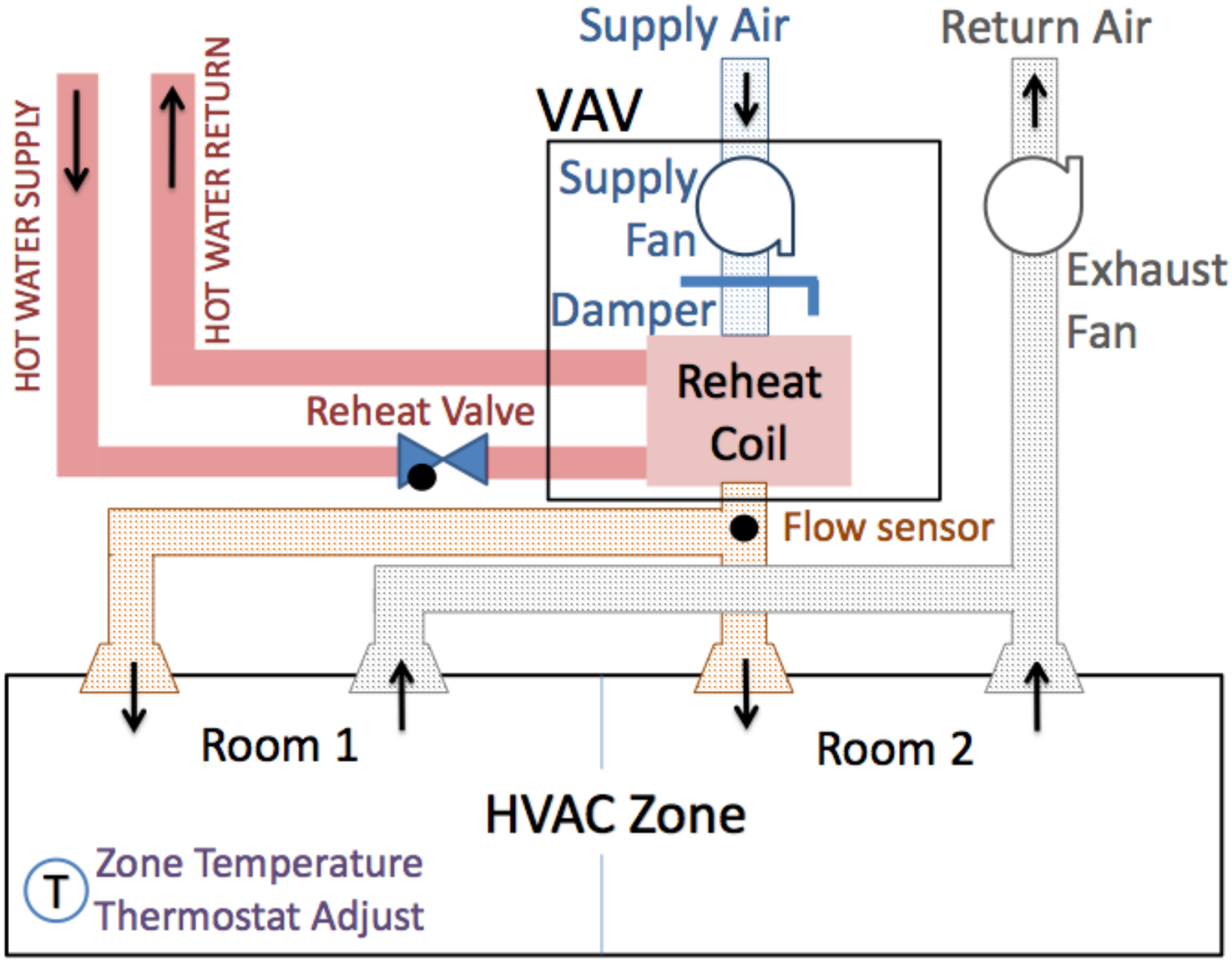
Sensors : Temperature, flow

Actuators : Dampers, fans

Set points : cooling, heating, flow

Control Signals : occupied, heating, cooling

Background : Faults in VAV units



Zone specific flow control devices

Sensor Failure

Sensors : Temperature, flow

Actuator Failure

Actuators : Dampers, fans

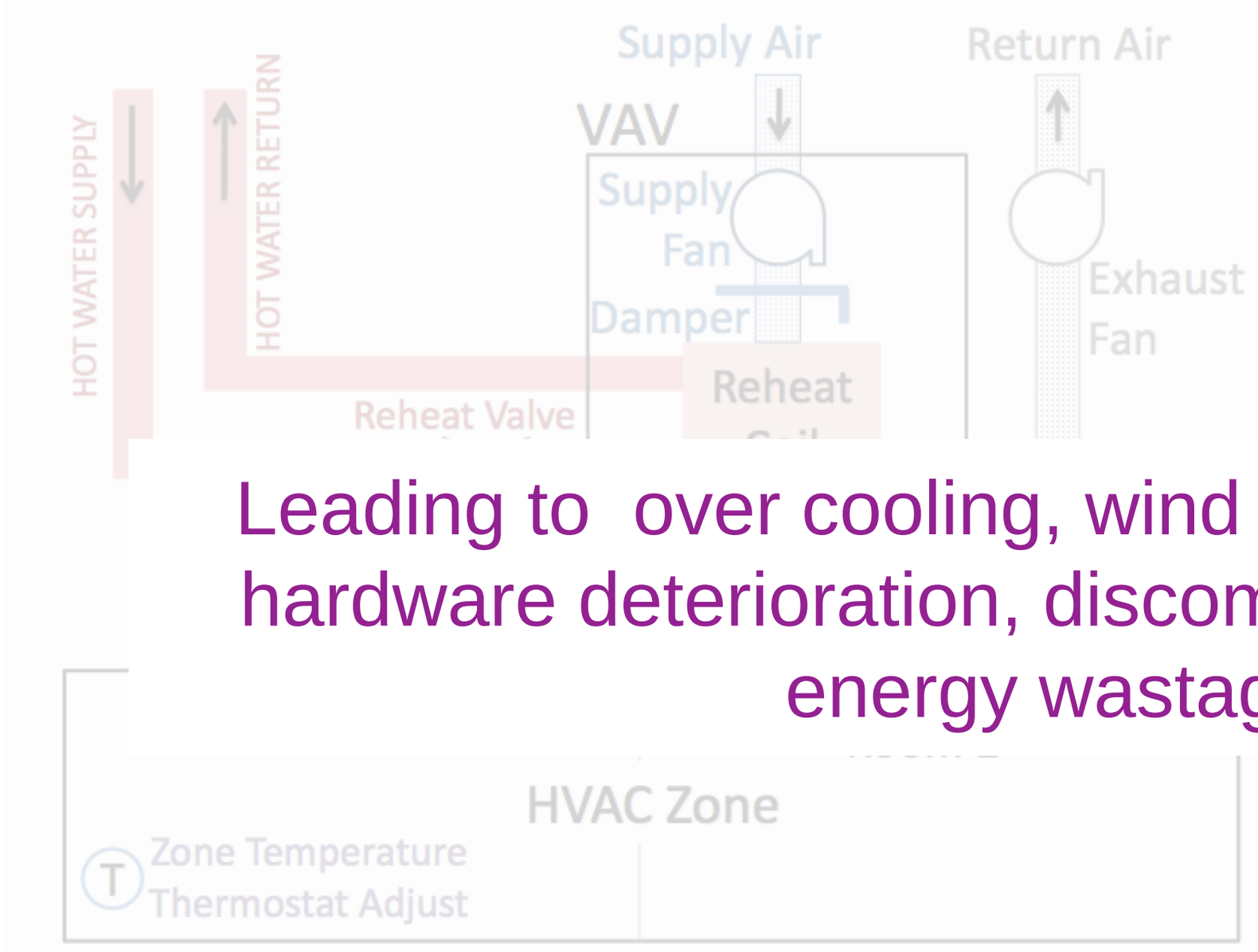
Software/Configuration Errors

Set points : cooling, heating, flow

Communication/Control Problems

Control Signals : occupied, heating, cooling

Background : Faults in VAV units



Zone specific flow control devices

Sensor Failure

Sensors : Temperature, flow

Actuator Failure

Actuators : Dampers, fans

Leading to over cooling, wind chill, undercooling hardware deterioration, discomfort to occupants, energy wastage

Communication/Control Problems

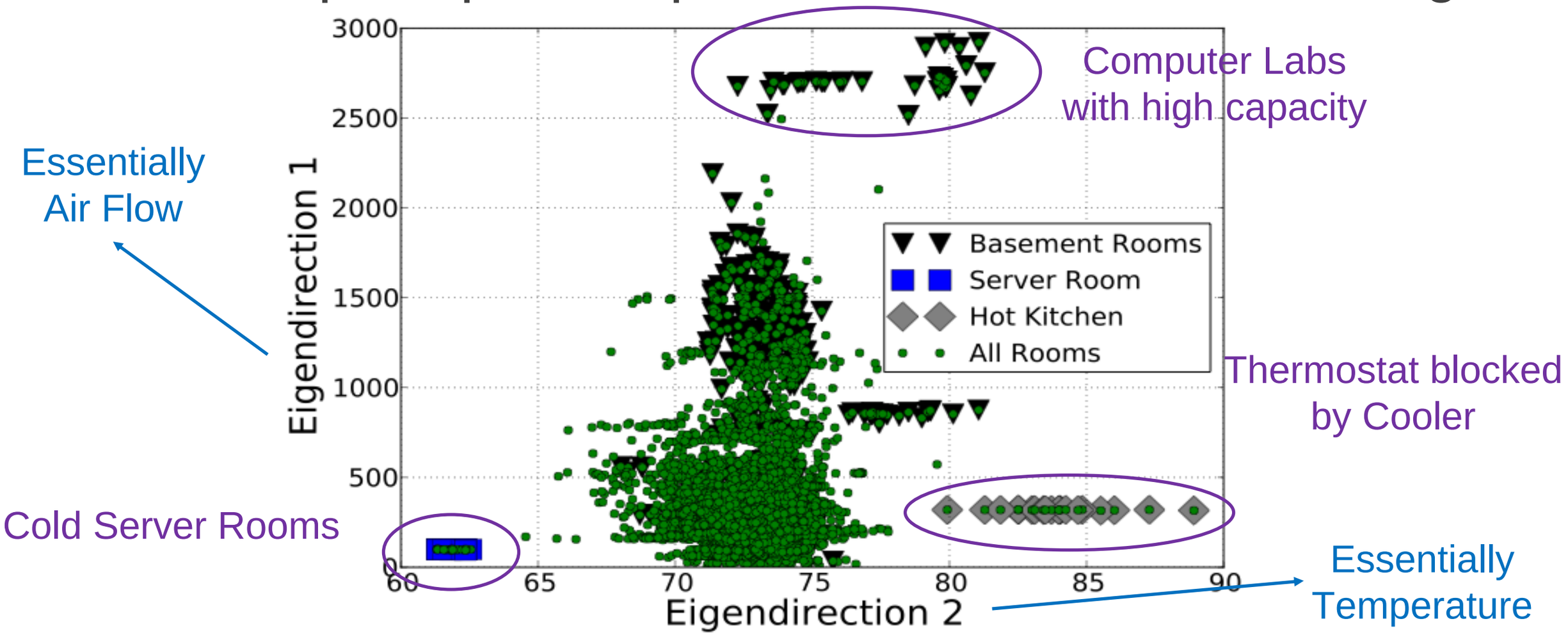
Control Signals : occupied, heating, cooling

Outline

- Need for building FDD
- Background on building HVAC
- Analysis of what a solution would look like
- Our Approach (Model Cluster Compare - MCC)
- Evaluation
- Research Questions

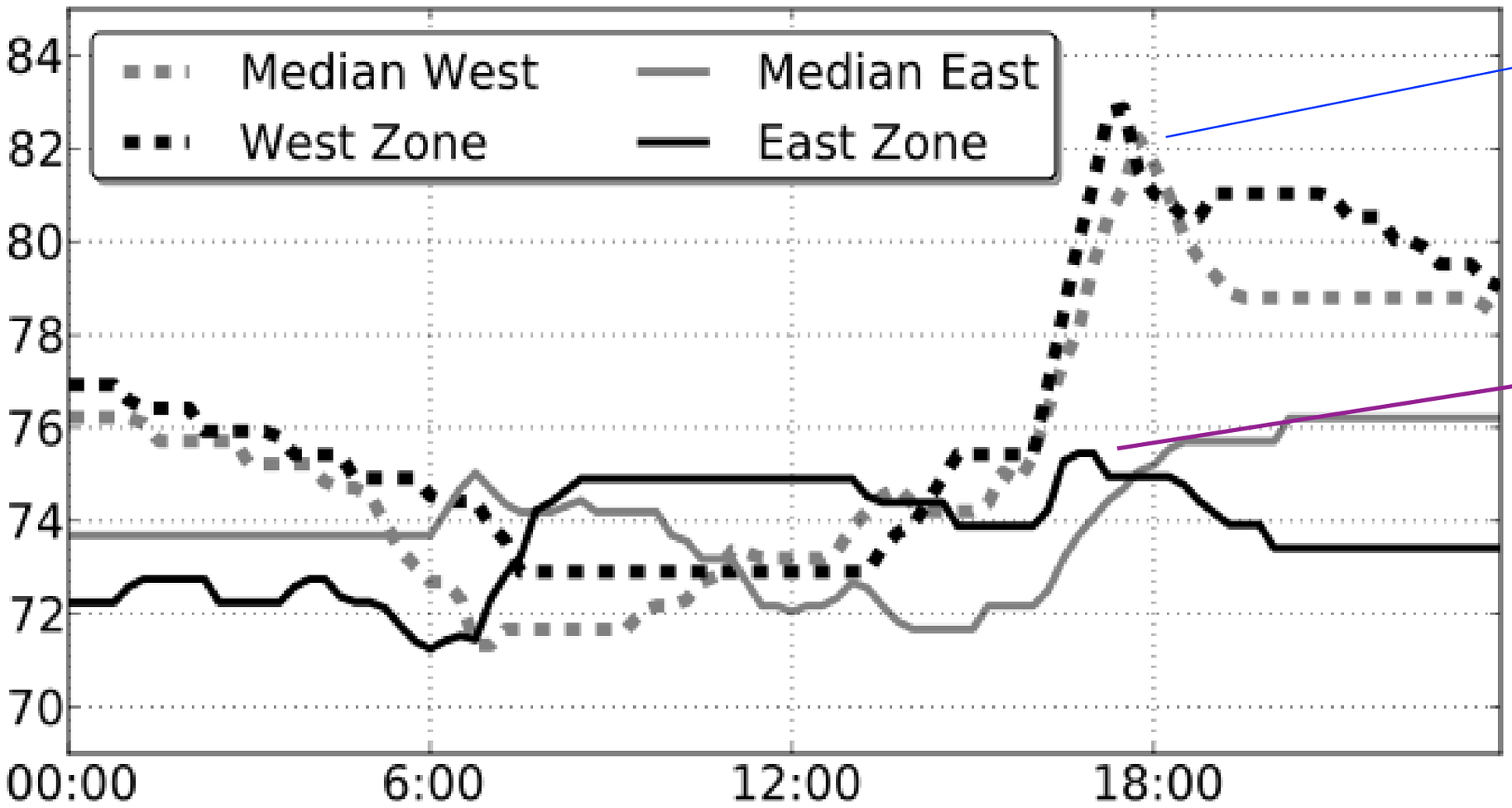
Analysis : Working with sensor readings

Data Driven Approach (PCA)
Find the principle components of the sensor readings



Working directly with sensor readings tends to find extreme measurements

Analysis : Generic rules and models



West facing,
late afternoon peak

East facing,
more temperate

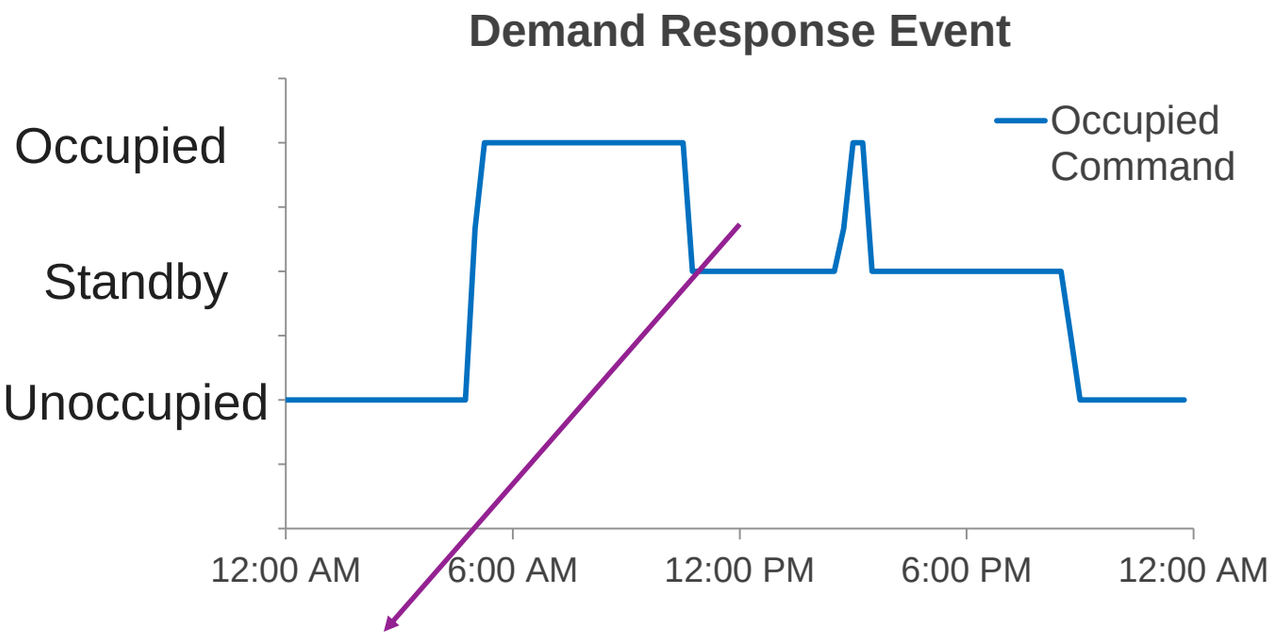
Large differences between zones even on the same day

Analysis : Change based methods

Change detection -Anomalies represent changes in characteristics of process history

Analysis : Change based methods

Change detection -Anomalies represent changes in characteristics of process history

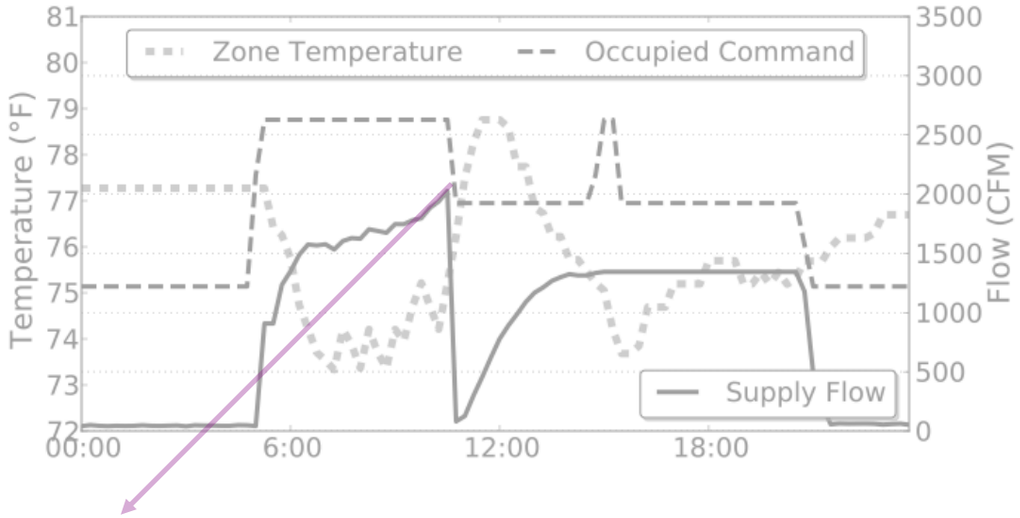


Occupied command changed in response to DR event

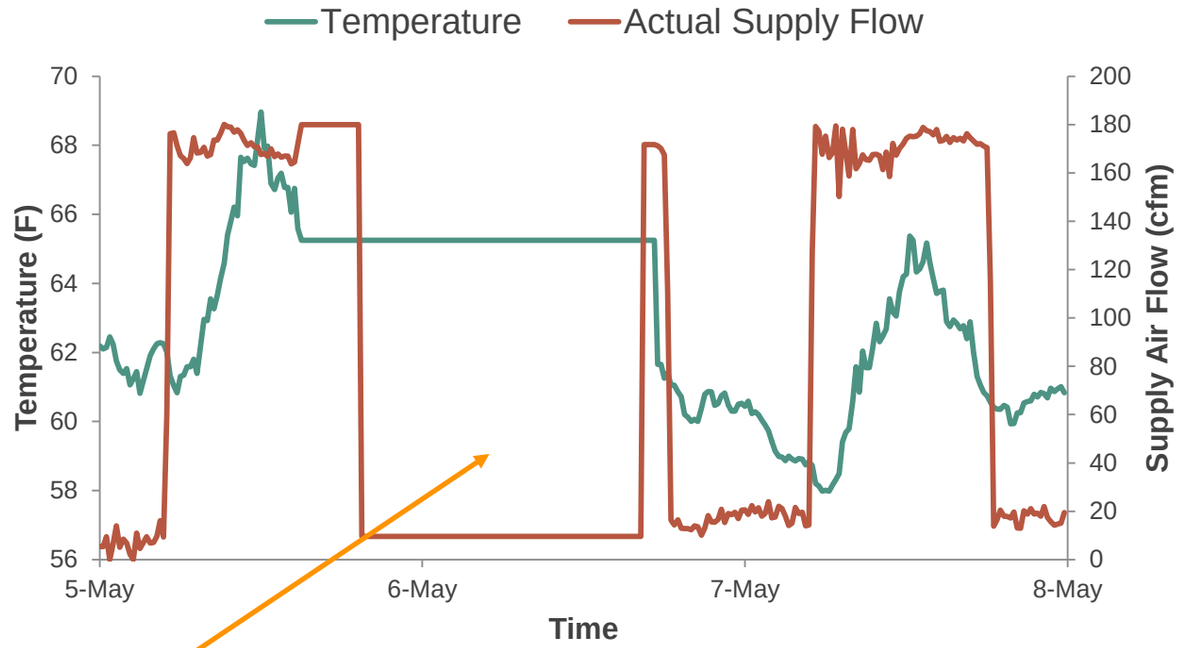
Analysis : Change based methods

Change detection -Anomalies represent changes in characteristics of process history

DR event



Occupied command changed in response to DR event

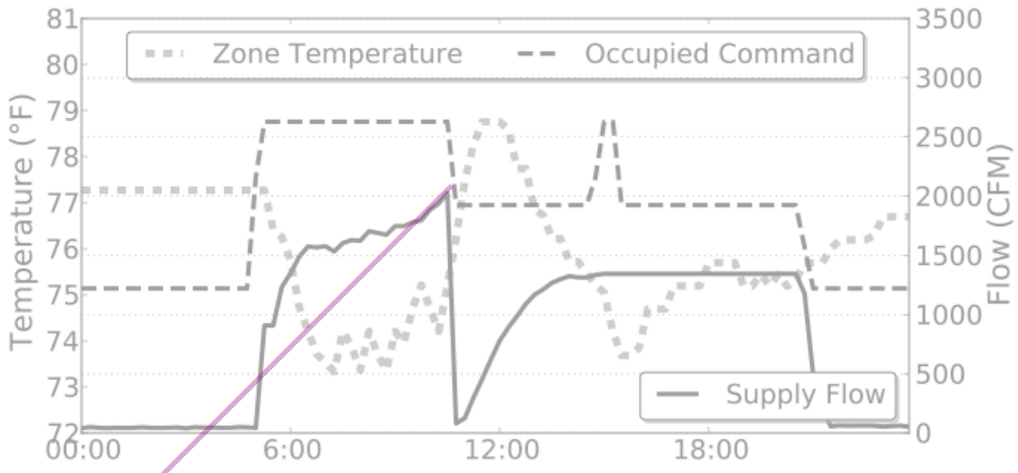


Missing Data

Analysis : Change based methods

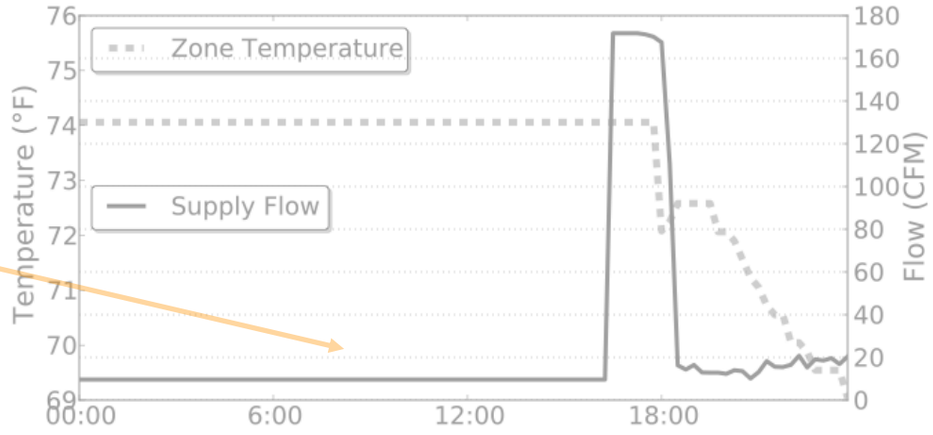
Change detection -Anomalies represent changes in characteristics of process history

DR event



Occupied command changed in response to DR event

Missing Data



Sensitive to confounding parameters, like external events

Analysis : Putting it all together

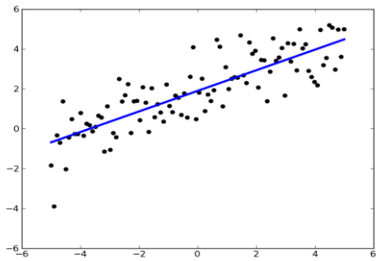
- Working directly with sensor readings tends to find extremes in sensor readings
  - Models that capture on the inter-relationships between sensors and parameters of interest
- Large differences between zones even on the same day
  - Cluster rooms with the same characteristics
- Sensitivity to confounding parameters
  - Compare rooms that have the same confounding parameters

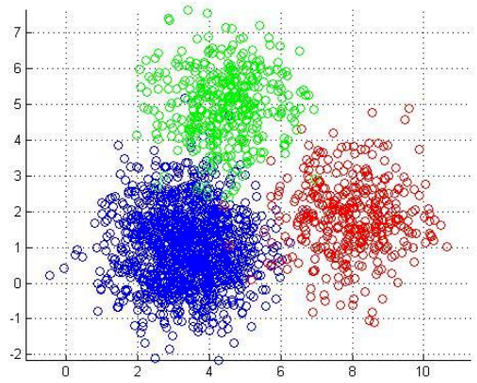
Outline

- Need for building FDD
- Background on building HVAC
- Analysis of what a solution would look like
- Our Approach (Model Cluster Compare - MCC)
- Evaluation
- Research Questions

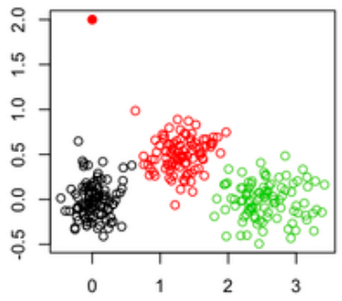
MCC : Model, Cluster, Compare

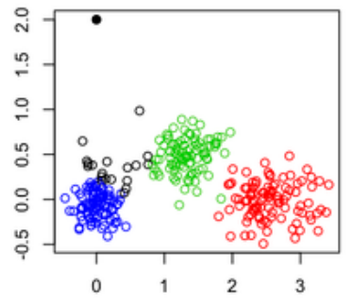
- Build black-box statistical models of the inter-relationship of sensors in a VAV box
  - We use linear models





- Cluster these model parameters
  - We use k-means
- Compare them across time and space to identify anomalies
  - We say far away from any cluster
- Many, many, many alternatives!





MCC : Model

- A zone has multiple sensors and actuators
  - (temperature, air-flow, damper position etc.)
- Sense and actuate the same spatio-temporal space
- Results in analytical redundancy
  - captured by model coefficients : both structure and value

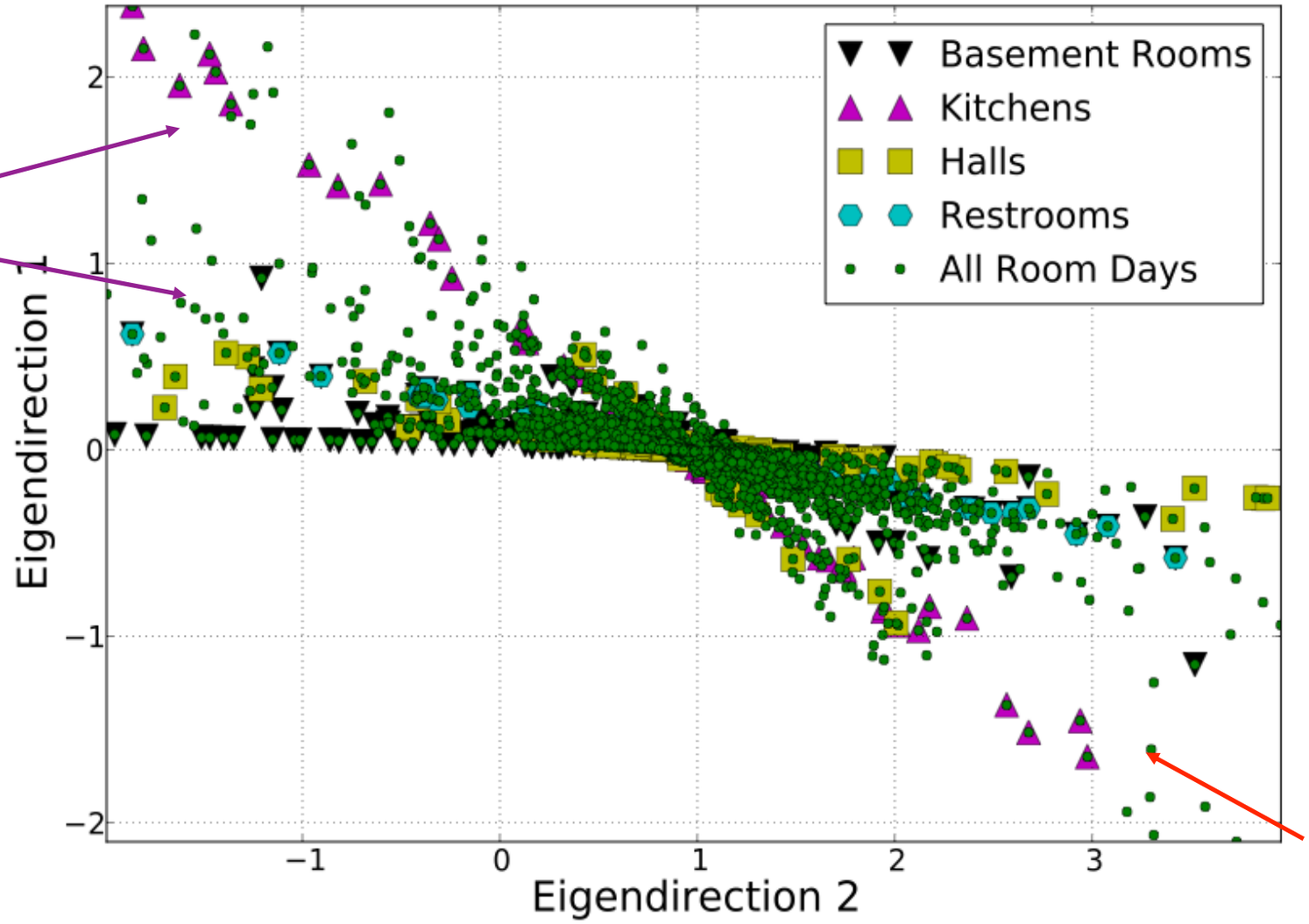
Model

Energy consumption → $E^i_{t+1} = \Theta \vec{x}^i_t + n^i_t$

Sensor + Actuator Readings

MCC : Model

Anomalies ?



Anomalies ?

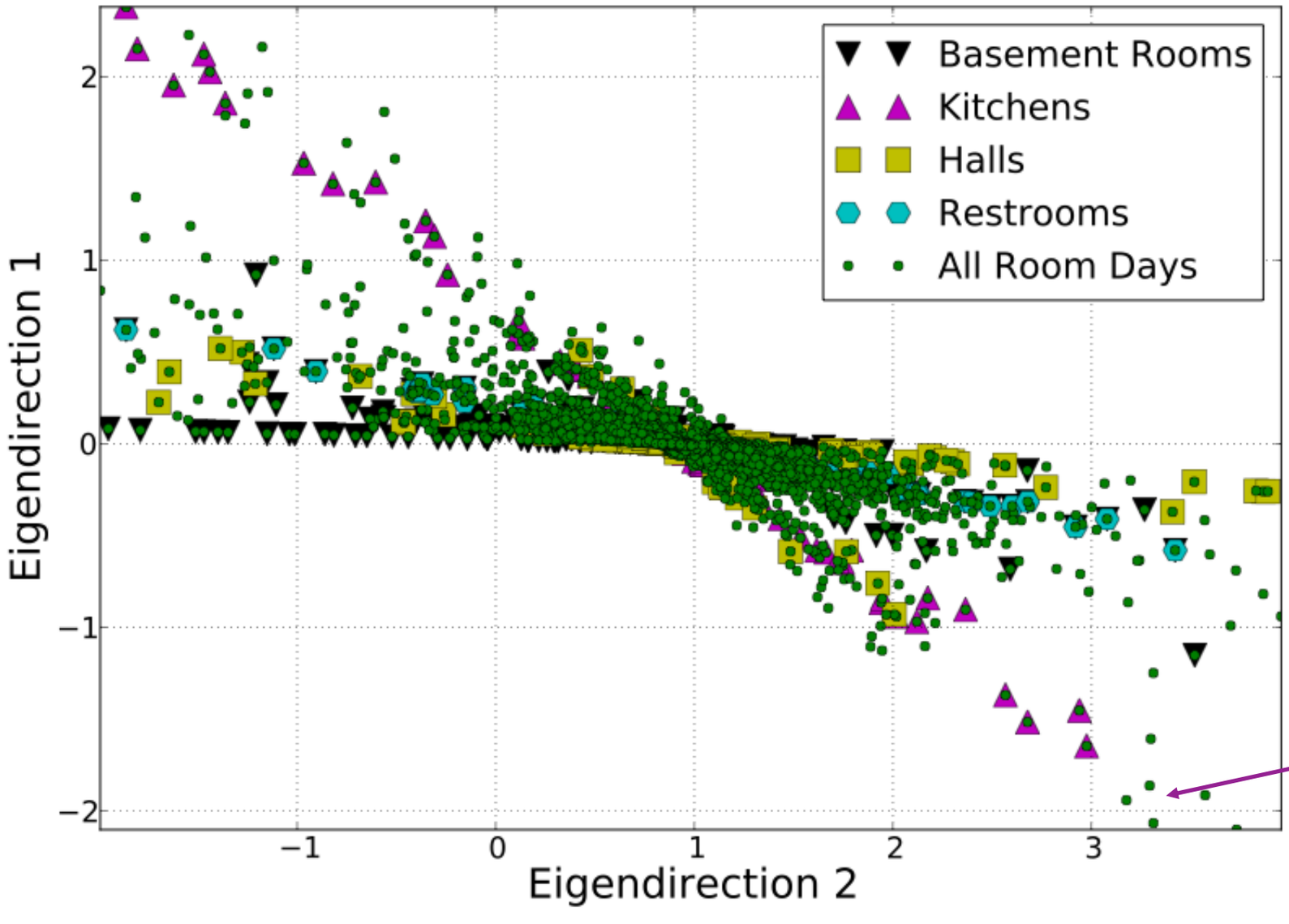
PCA on model coefficients

See a different kind of separation

MCC : Cluster and Compare

Cluster Model Coefficients

k-Means++



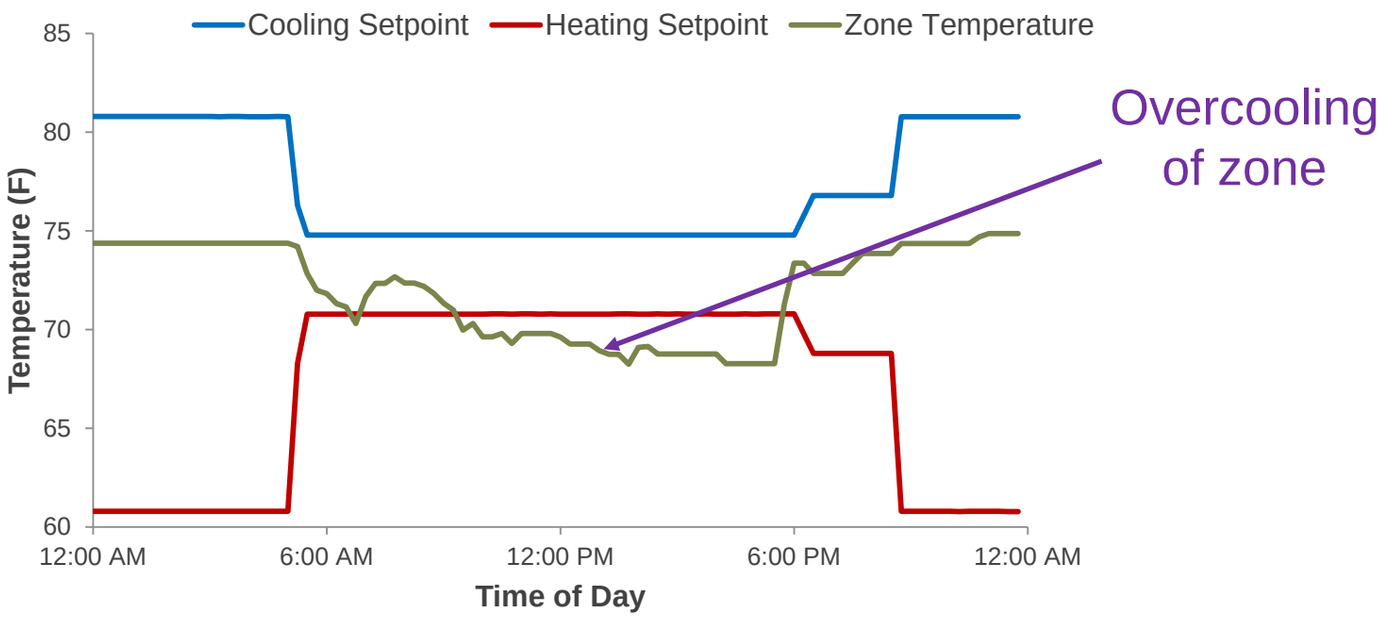
How many clusters ?

Anomalies!

Outline

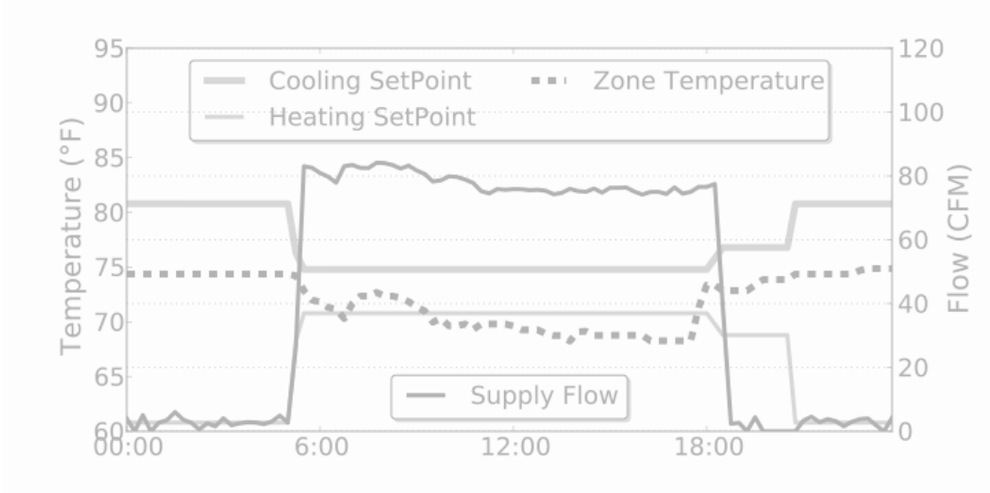
- Need for building FDD
- Background on building HVAC
- Analysis of what a solution would look like
- Our Approach (Model Cluster Compare - MCC)
- Evaluation
- Research Questions

Evaluation : Quantitative

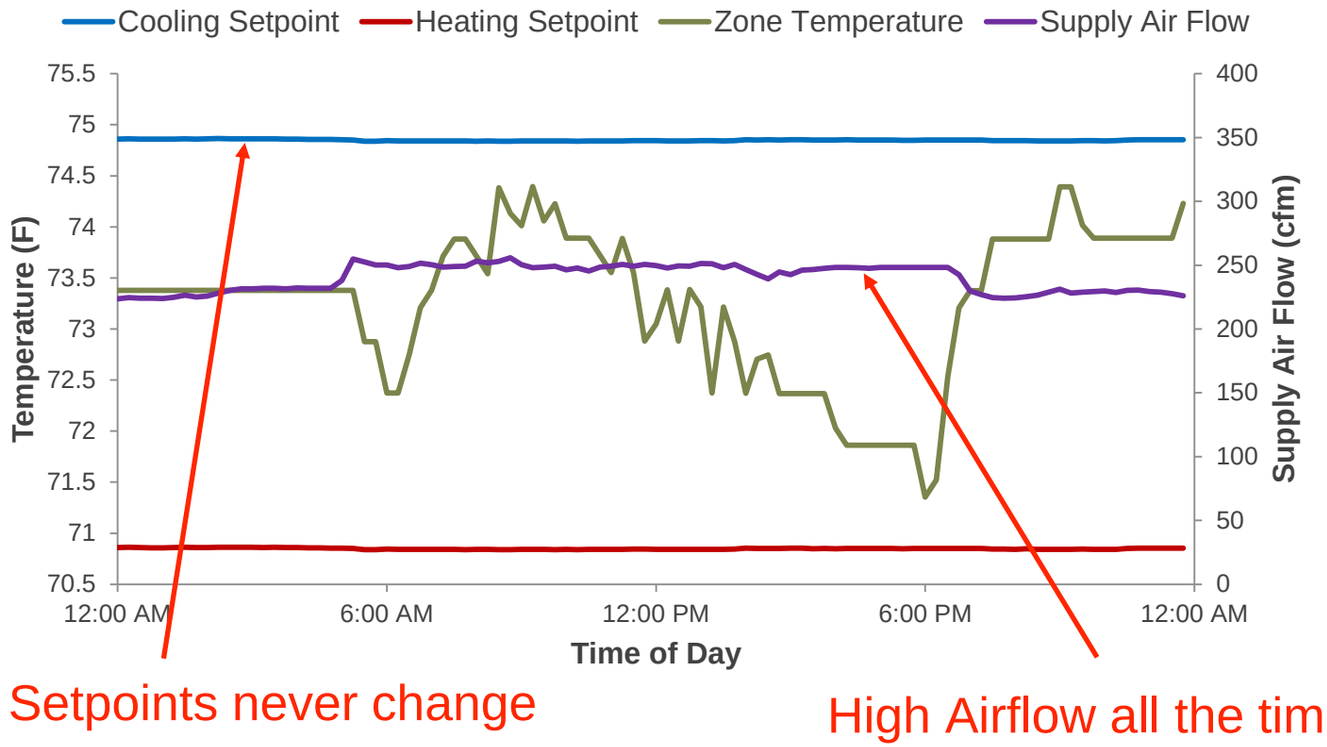


Zone temperature not within limits

Evaluation : Quantitative

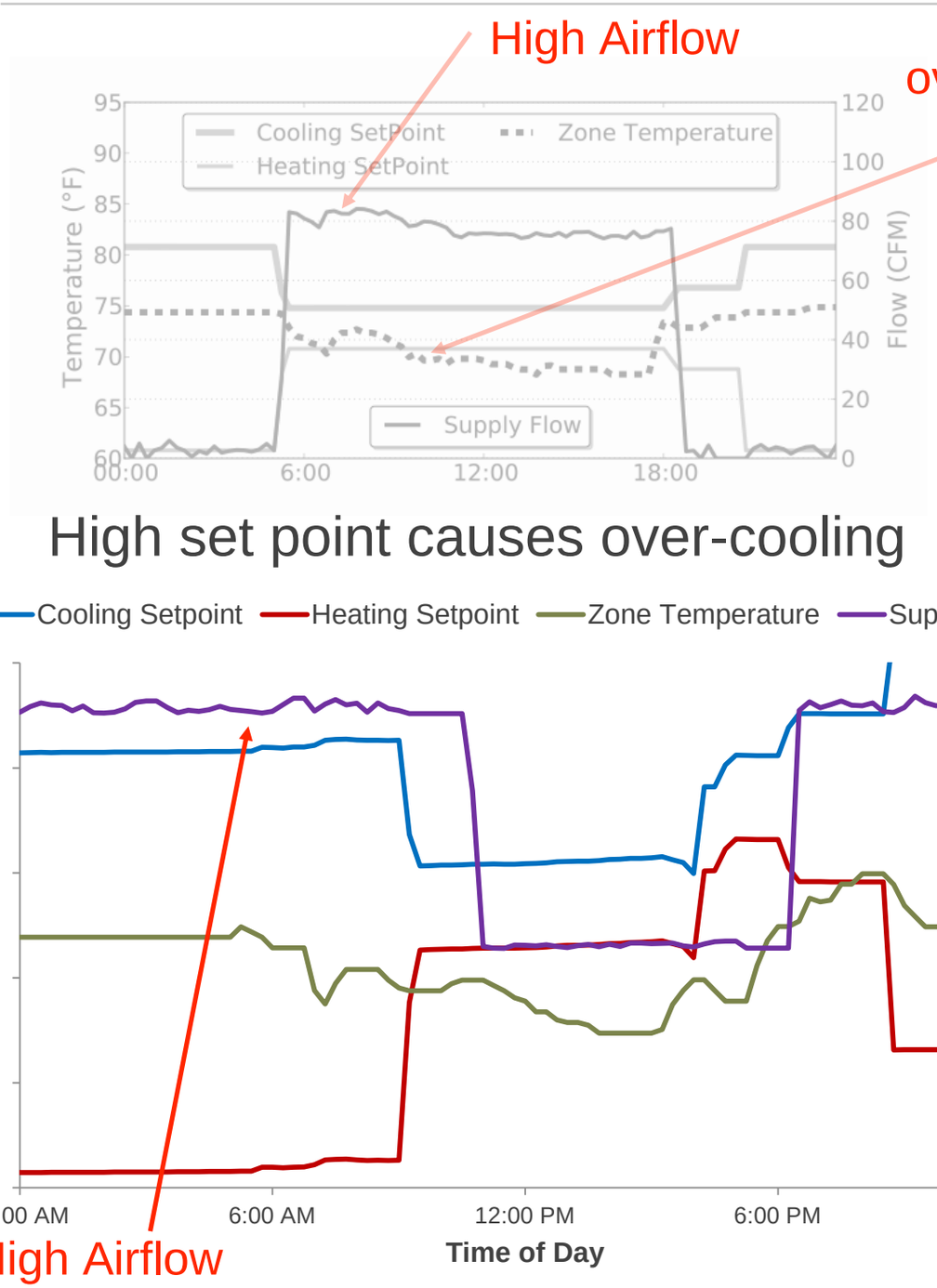


Zone temperature not within limits



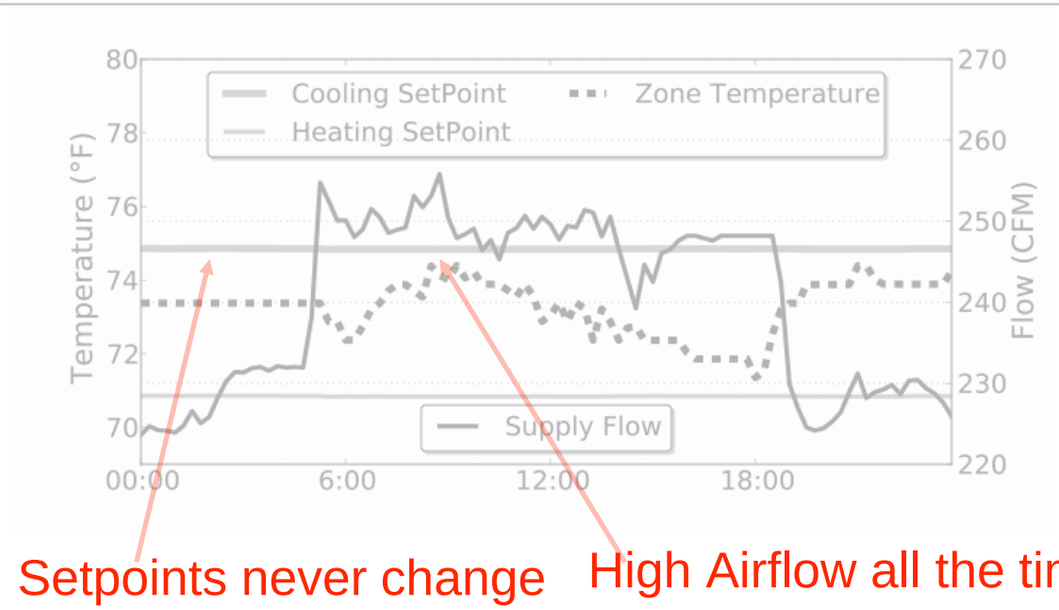
Set points and actuation not in sync

Evaluation : Quantitative



High set point causes over-cooling

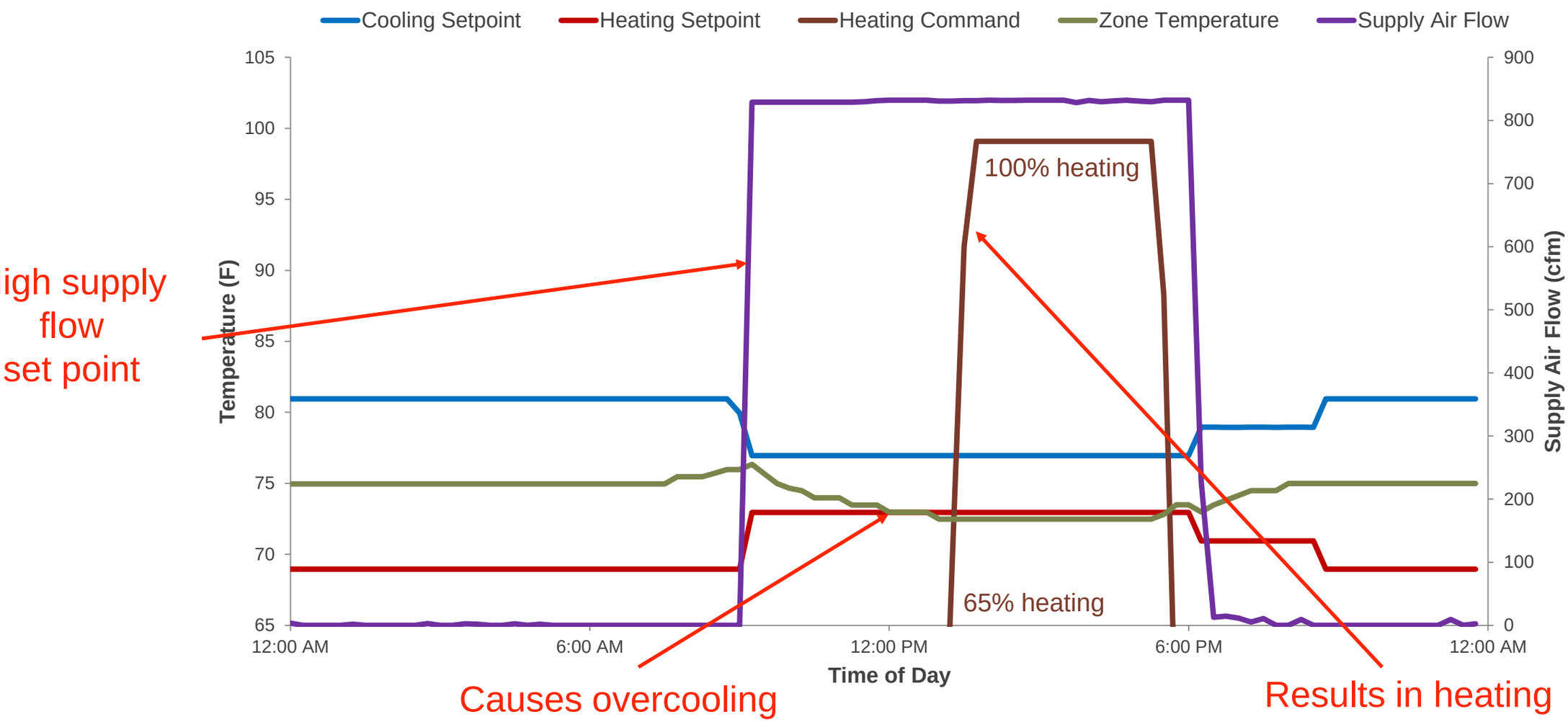
(Unnecessary) high flow when within set points



Set points and actuation not in sync

Even though temperature within set points

Evaluation : Dynamic Anomaly



‘Dynamic’ Anomaly - Not anomalous at any one point

Overcooling - Heating Cycles

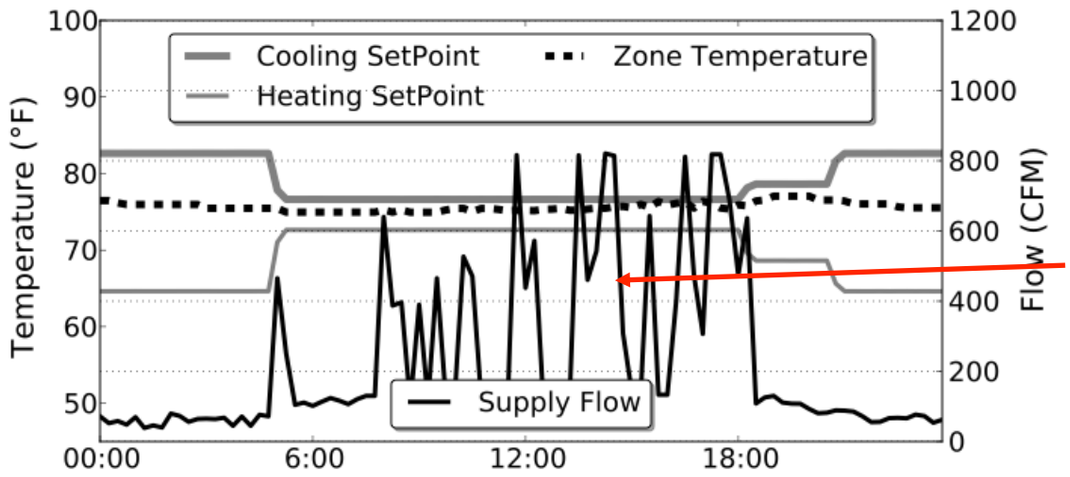
Big Energy Waste! [but hard to fix without occupancy]

Evaluation : A problem with data driven methods

- Problem :
  - A data driven algorithm (PCA, SBS, MCC) does find ‘interesting’ anomalies
    - Different kinds of anomalies - all useful ?
  - However, they don’t find all instances of every anomalies
- Solution :
  - Design rules based on fault exemplars
  - iRules - (don’t sue us) - Intelligent Rules

Evaluation : Qualitative

1. Heating or Cooling is Ineffective
2. Excess Flow
3. Heating-Cooling Anomaly
4. Temperature and Air Flow not within set points



5. Short Cycling Anomaly

Detected by change based methods

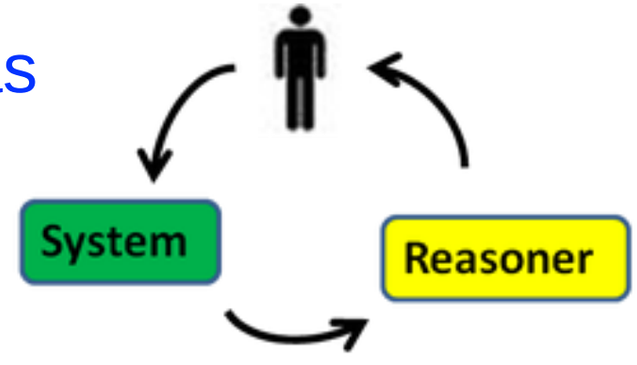
Faults

Method	Data	Ineffective	Excess Flow	Heating-Cooling	Set-Point	Cycling	False Alarms
Manual	Fourth floor, 30 days	7	1	3	2	1	0
Rules	Fourth floor, 30 days	4	1	2	2	0	2
Correlation	Fourth floor, 30 days	1	1	0	0	1	8
PCA	Fourth floor, 30 days	1	0	1	2	1	5
MCC	Fourth floor, 30 days	5	1	3	1	0	2
Rules (top 10 verified)	All Floors, 1 year	0	0	0	10	0	0
Correlation (top 10 verified)	All Floors, 1 year	0	1	0	1	0	8
PCA (top 10 verified)	All Floors, 1 year	1	0	0	1	0	8
MCC (top 10 verified)	All Floors, 1 year	0	2	6	1	0	1
iRules	All Floors, 1 year	18	12	14	27	7	1

Performance

Conclusions

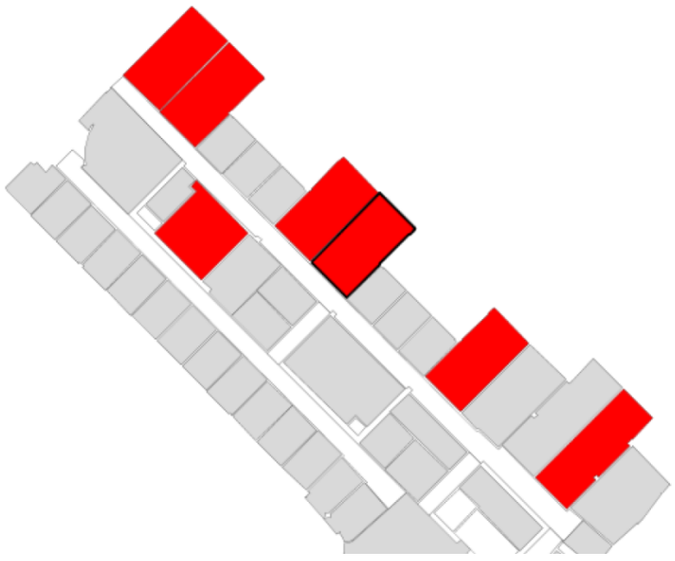
- Buildings are a big source of data
  - Too big ? Data mining, visualization, HCI are critical
- Despite data- Many faults, many long-standing faults, much energy waste
  - Data driven anomaly detection is useful
- Generalization to new buildings, climates is hard
  - Within building comparative data mining is useful
- We find common areas [kitchens, conference rooms] often mis-configured
  - Value of occupant count based control in these areas
- Data mining good, human in the loop better



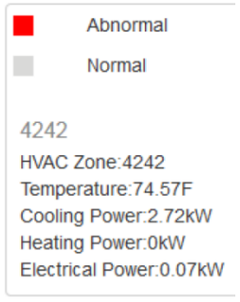
BuildingSherlock: Fault Management

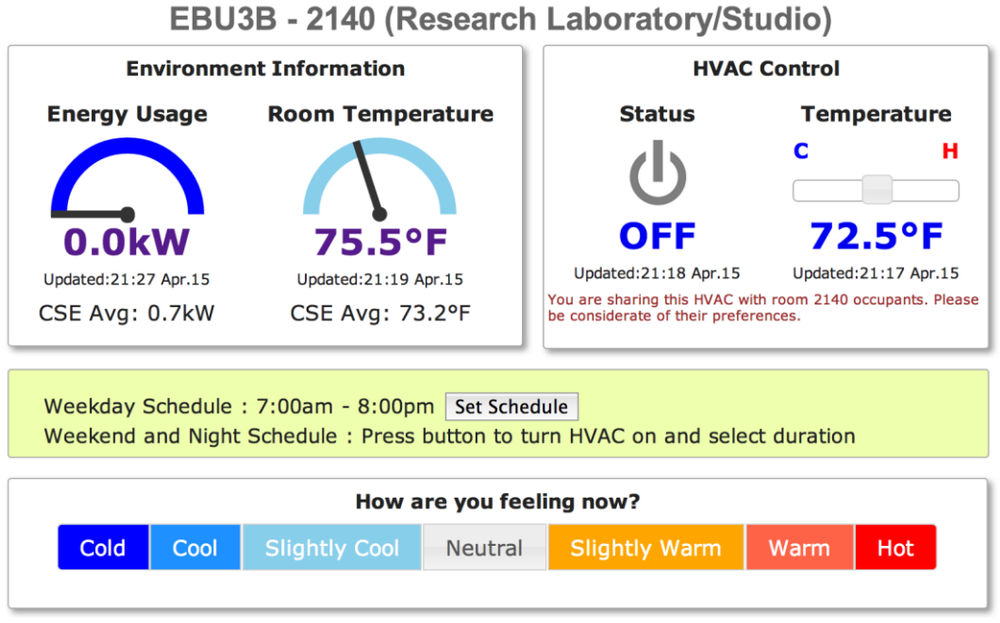
Framework to detect faults and prioritize fixing them

- Make sensor data and metadata available
- Allow complex fault detection algorithms
- Provide contextual information on faults



Fault Dashboard

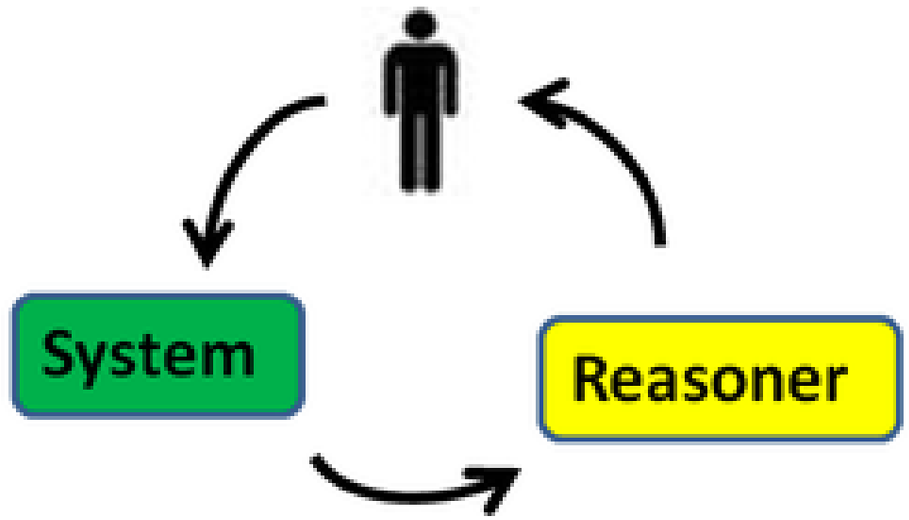




Occupant Dashboard

Future Work

- Specific - Comparative data mining
  - How many clusters ? Other clustering methods ?
  - Fault Diagnosis
  - Better (non-linear) room models
  - Other buildings, climates
- More general - How much human in the loop ?
  - Automatic rule generation
  - What is anomalous ?
  - Automatic fixing of software errors ?
  - Automatic fixing of hardware errors ?



Thank You!

Acknowledgements: Anna Levitt, Rizhen Zhang, Anthony Nwokafor, UC San Diego Facilities Management