Environmental Conditions and Disk Reliability in Free-cooled Datacenters

Ioannis Manousakis, Sriram Sankar, Gregg McKnight,

Thu D. Nguyen, and Ricardo Bianchini

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Microsoft

Problem Statement

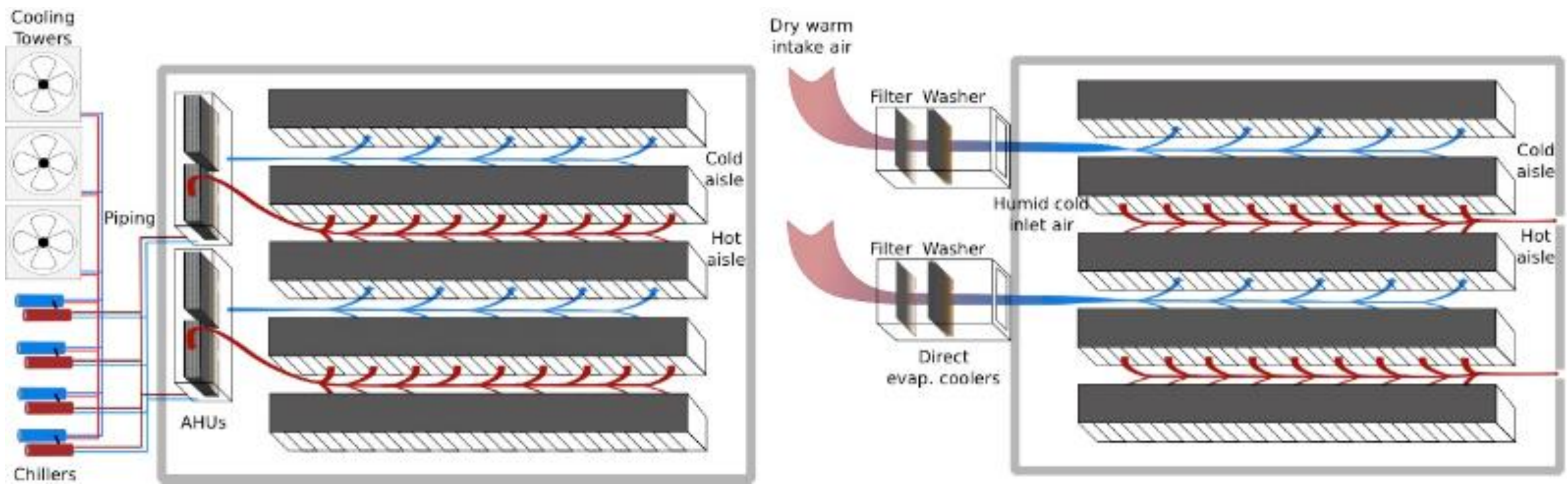
- Datacenters are costly and consume lots of energy

- Evolving cooling technologies in datacenters
  - Chiller-based (traditional)
  - Water-side economized
  - ***<u>Air-side economized (aka free cooling)</u>***

- Unexplored tradeoff: environmentals, reliability, cost

Background

- Cooling technologies

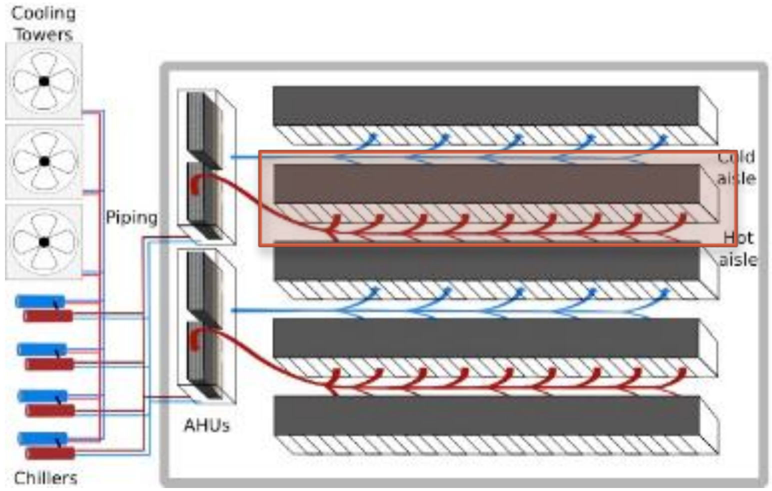


a) Chiller-based Cooling　　b) Free Cooling

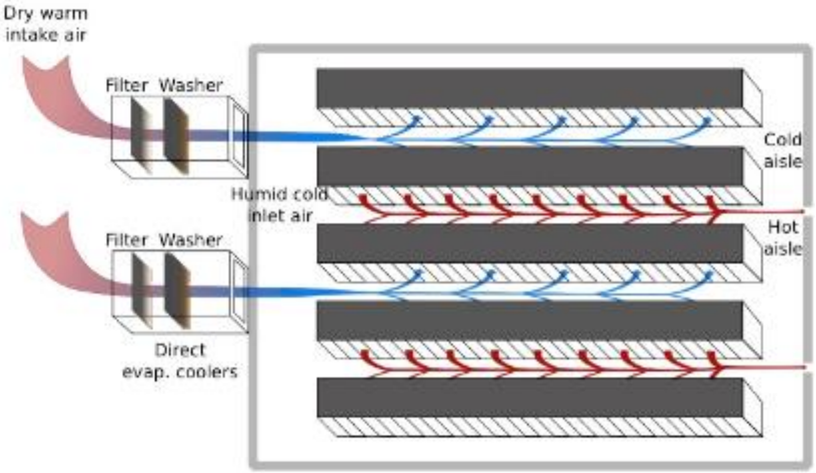
➢ *Chillers and water-side: tightly controlled environmentals*

Background

- Cooling technologies



a) Chiller-based Cooling

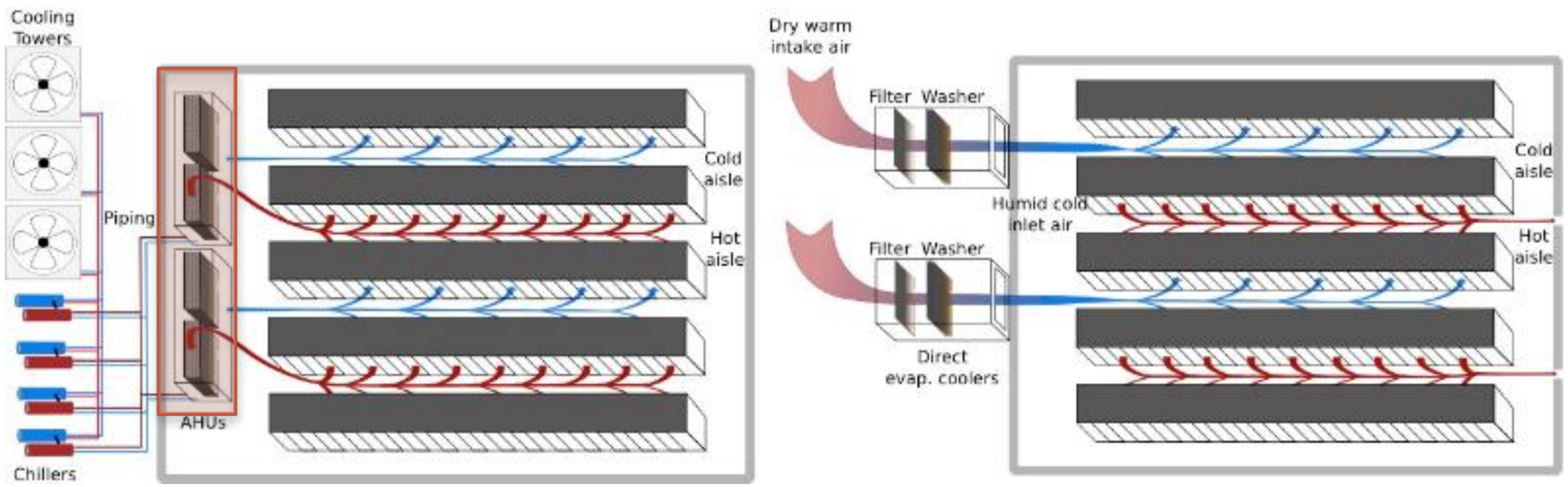


b) Free Cooling

➢ *<u>Chillers and water-side: tightly controlled environmentals</u>*

Background

- Cooling technologies

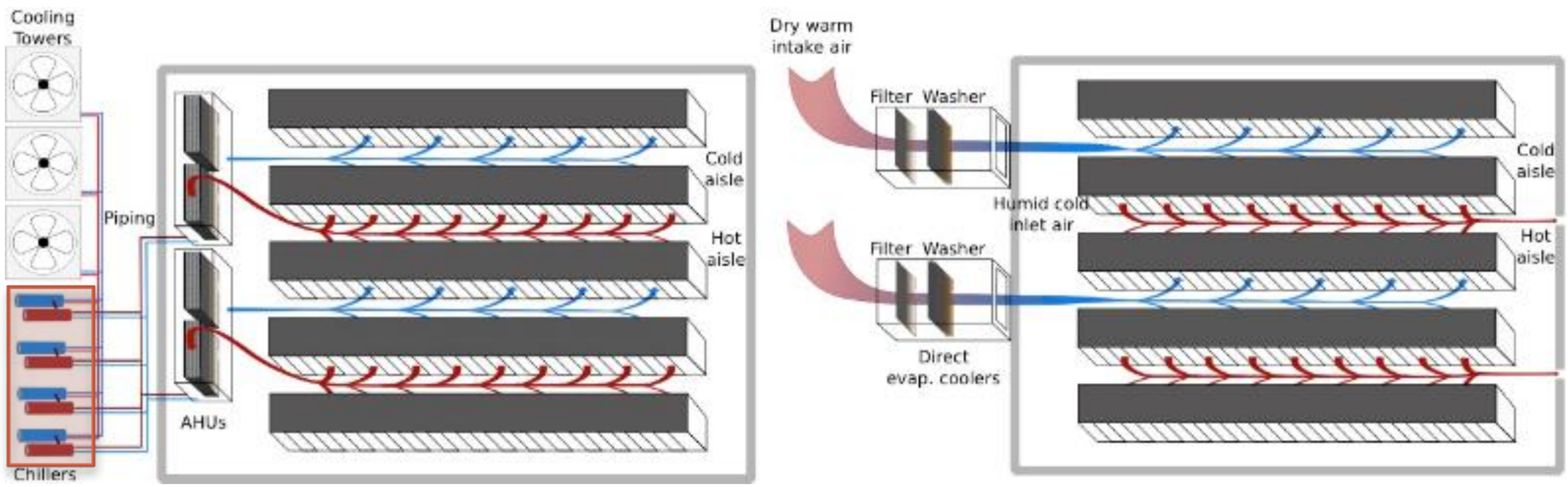


a) Chiller-based Cooling ***b) Free Cooling***

➢ *<u>Chillers and water-side: tightly controlled environmentals</u>*

Background

- Cooling technologies

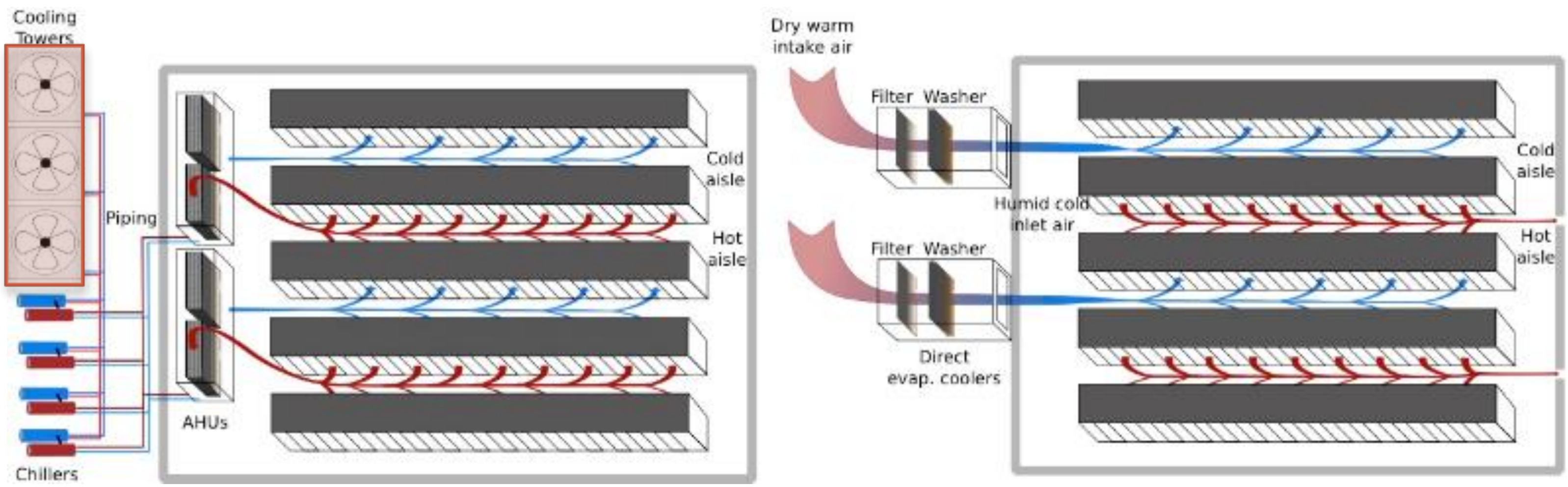


a) Chiller-based Cooling ***b) Free Cooling***

➢ *Chillers and water-side: tightly controlled environmentals*

Background

- Cooling technologies

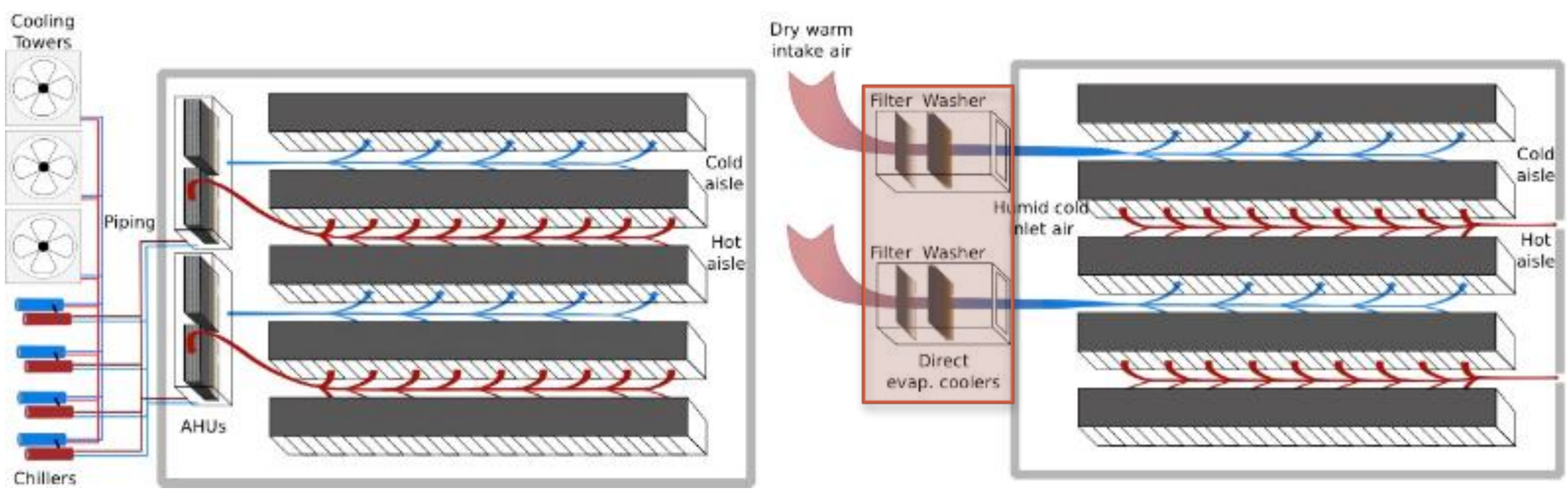


a) Chiller-based Cooling ***b) Free Cooling***

➢ *<u>Chillers and water-side: tightly controlled environmentals</u>*

Background

- Cooling technologies

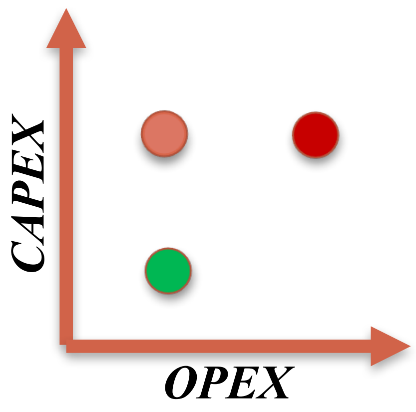


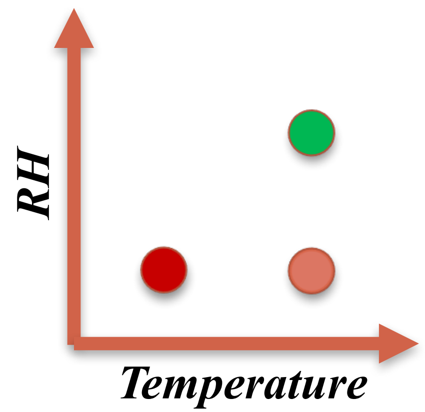
a) Chiller-based Cooling ***b) Free Cooling***

➢ *<u>Free cooling: may expose servers to harsh environmentals</u>*

Technology Characteristics

- Cooling technologies:
  - Chiller-based
  - Water-side economized
  - Free cooling





Prior Work

- Hard disk failure studies in datacenters
  - Pinheiro[FAST'07], El-Sayed[SIGMETRICS'12], Sankar[ToS'13]

- *Focused on temperature and temperature variation*
  - Chiller-based (traditional) datacenters

- What's new with our work
  - Three types of cooling
  - Wider (more aggressive) environmental envelopes
  - Primary focus on (relative) humidity

Contributions and Roadmap

1. ***<u>Impact of environmentals on disk failure rates</u>***
2. Root causes
3. Cooling vs reliability vs cost tradeoffs
4. Modeling of failure rates
5. Design considerations

Methodology

- Collect large traces from hard disks
  - Nine datacenters (2-4 years), 1M HDDs
  - All types of Microsoft datacenters

Tag	Technology	Population
CD1	Chiller	117K
CD2	Water-side	146K
CD3	Free-Cooled	24K
HD1	Chiller	16K
HD2	Water-side	100K
HH1	Free-Cooled	168K
HH2	Free-Cooled	213K
HH3	Free-Cooled	124K
HH4	Free-Cooled	161K
Total		**1.07M**

CD1

cold *dry*

HH1

hot *humid*

Methodology

- Collect extensive hard disk operation traces
  - Logged and archived by Microsoft Autopilot

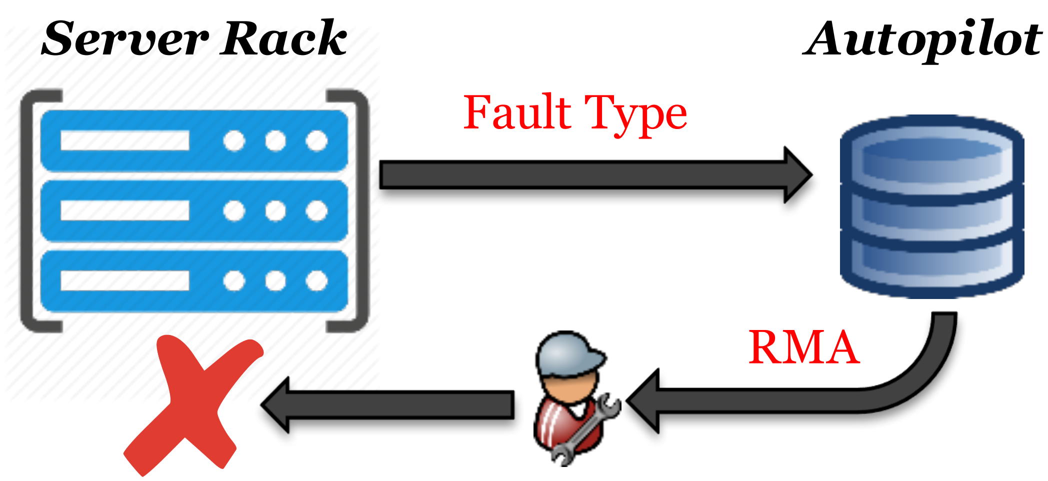
1. I/O communication faults (dead controller / TX-RX error)
2. Behavioral SMART faults (read-write, sectors, seek, etc.)
3. Age-related SMART faults (max hours, on-off cycles, etc.)

Server Rack

Autopilot

Methodology

- Collect extensive hard disk operation traces
  - Logged and archived by Microsoft Autopilot
  1. I/O communication faults (dead controller / TX-RX error)
  2. Behavioral SMART faults (read-write, sectors, seek, etc.)
  3. Age-related SMART faults (max hours, on-off cycles, etc.)



Results

- Annual Failure Rate summary
  1. Dry datacenters exhibit low AFRs (1.5 – 2.3%)
  2. Humid datacenters exhibit higher AFRs (3.1 – 5.4%)

DC Tag	Technology	AFR	Increase wrt 1.5%
CD1	Chiller	**1.5%**	0%
CD2	Water-side	**2.1%**	40%
CD3	Free-Cooled	**1.8%**	20%
HD1	Chiller	**2.0%**	33%
HD2	Water-side	**2.3%**	53%
HH1	Free-Cooled	**3.1%**	107%
HH2	Free-Cooled	**5.1%**	240%
HH3	Free-Cooled	**5.1%**	240%
HH4	Free-Cooled	**5.4%**	260%

CD1

cold *dry*

HH1

hot *humid*

Annual Failure Rate (AFR) Results

1. Dry datacenters show low AFRs (1.5 – 2.3%)
2. Humid datacenters show higher AFRs (3.1 – 5.4%)

DC Tag	Technology	AFR	Increase wrt 1.5%
CD1	Chiller	**1.5%**	0%
CD2	Water-side	**2.1%**	40%
CD3	Free-Cooled	**1.8%**	20%
HD1	Chiller	**2.0%**	33%
HD2	Water-side	**2.3%**	53%
HH1	Free-Cooled	**3.1%**	107%
HH2	Free-Cooled	**5.1%**	240%
HH3	Free-Cooled	**5.1%**	240%
HH4	Free-Cooled	**5.4%**	260%

CD1

cold *dry*

HH1

hot *humid*

Annual Failure Rate (AFR) Results

1. Dry datacenters show low AFRs (1.5 – 2.3%)
2. Humid datacenters show higher AFRs (3.1 – 5.4%)

DC Tag	Technology	AFR	Increase wrt 1.5%
CD1	Chiller	**1.5%**	0%
CD2	Water-side	**2.1%**	40%
CD3	Free-Cooled	**1.8%**	20%
HD1	Chiller	**2.0%**	33%
HD2	Water-side	**2.3%**	53%
HH1	Free-Cooled	**3.1%**	107%
HH2	Free-Cooled	**5.1%**	240%
HH3	Free-Cooled	**5.1%**	240%
HH4	Free-Cooled	**5.4%**	260%

CD1

cold *dry*

HH1

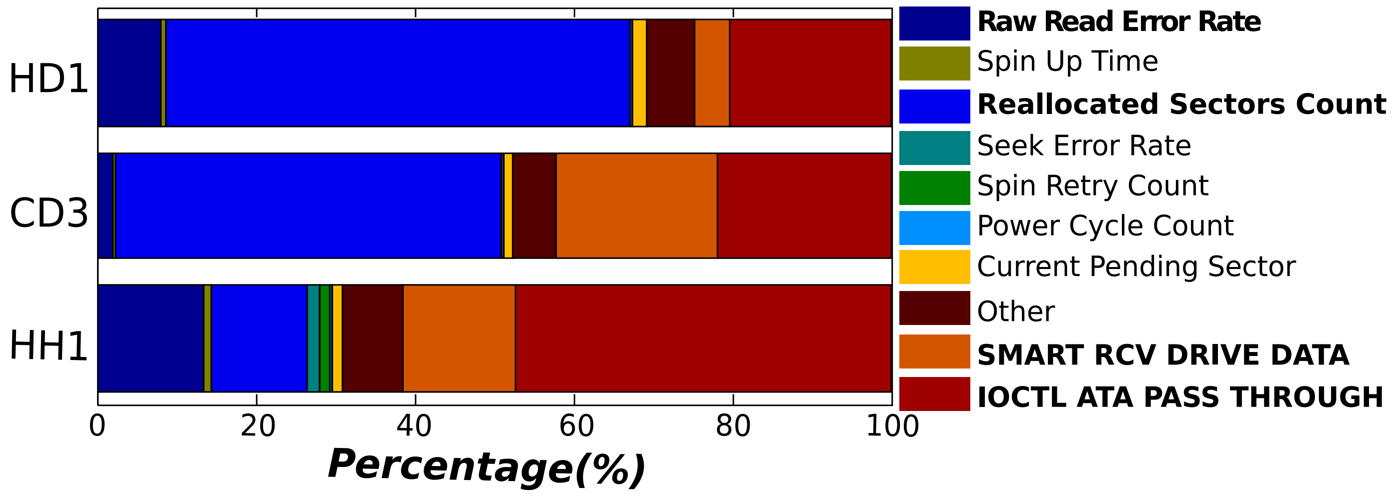
hot *humid*

Contributions and Roadmap

1. Impact of environmentals on disk failure rates
2. **Root causes**
3. Cooling vs reliability vs cost tradeoffs
4. Modeling of failure rates
5. Design considerations

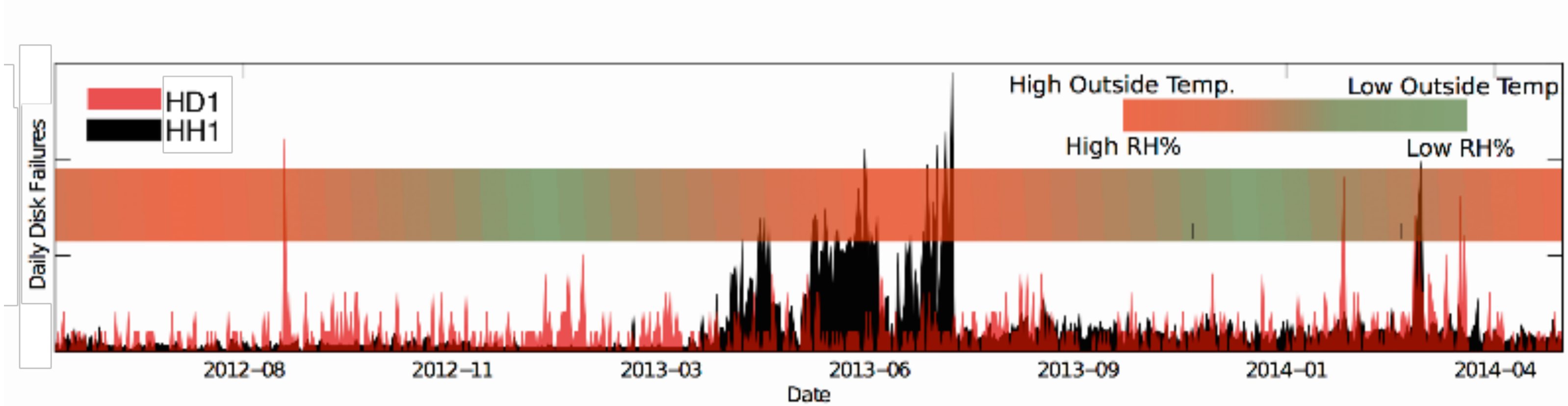
Root Causes: Error Breakdown

- Dry DCs → Bad sector count (mechanical): ~50-60%
- Humid DCs → Controller (connectivity): ~60%



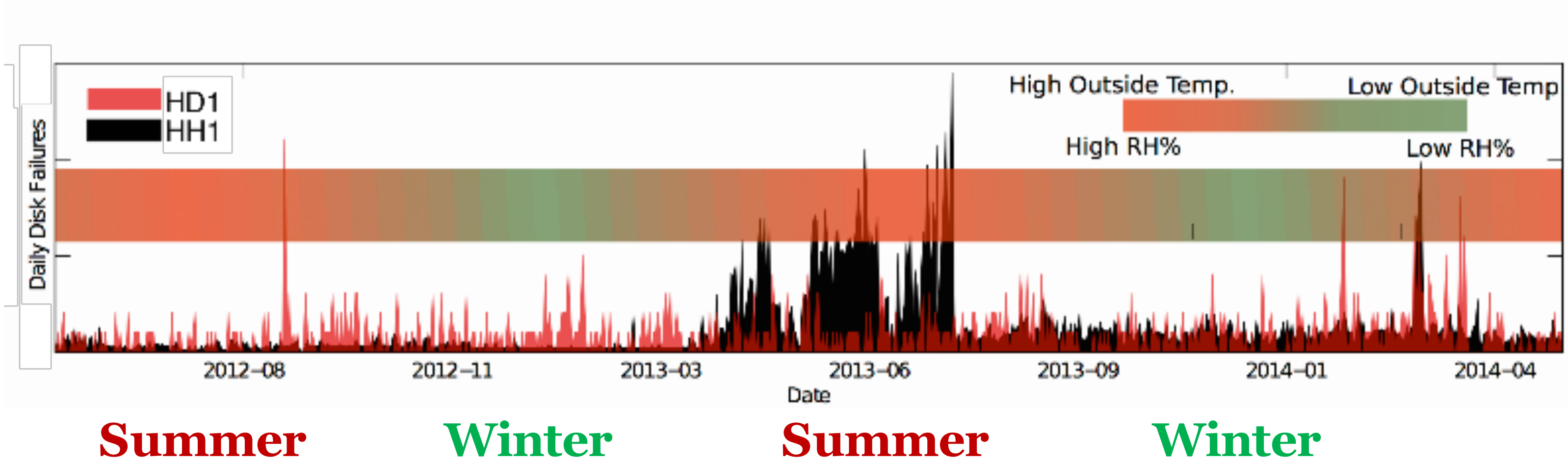
Root Causes: Temporal Clustering

- Significant temporal clustering on **HH1**
- No temporal clustering on **HD1**



Root Causes: Temporal Clustering

- Significant temporal clustering on **HH1**
- No temporal clustering on **HD1**



➤ **<u>Data suggests a lifetime failure process</u>**

Root Causes: Correlating to Environmentals

- Failure rate regressions for **<u>HH1</u>**
  1. Discover trends – variables that change together
  2. Split into 4 groups **P1 – P4** (total population = 170K)

coefficient a

Popul.	%	Temp.	RH	CoV Temp.	CoV RH
P1	30.1	$-6.4*10^{-3}$	$5.1*10^{-2}$	$-1.7*10^{0}$	$-9.0*10^{-0}$
P2	25.6	$-1.6*10^{-2}$	$5.3*10^{-2}$	$-1.0*10^{-1}$	$-1.6*10^{-1}$
P3	23.3	$6.3*10^{-3}$	$9.9*10^{-2}$	$-8.4*10^{0}$	$3.5*10^{0}$
P4	19.6	$3.3*10^{-2}$	$11.5*10^{-2}$	$-3.9*10^{0}$	$-1.3*10^{0}$

$$y = a^*x + b$$

coefficient b

Popul.	%	Temp.	RH	CoV Temp.	CoV RH
P1	30.1	$5.1*10^{-5}$	$1.2*10^{-4}$	$-7.9*10^{-3}$	$-6.5*10^{-3}$
P2	25.6	$-1.9*10^{-5}$	$1.0*10^{-4}$	$-9.0*10^{-3}$	$-3.7*10^{-3}$
P3	23.3	$1.4*10^{-3}$	$2.1*10^{-4}$	$-4.9*10^{-2}$	$-4.4*10^{-2}$
P4	19.6	$1.7*10^{-3}$	$4.4*10^{-4}$	$-1.3*10^{-1}$	$-8.0*10^{-2}$

$$y = a^*e^{(b^*x)}$$

Root Causes: Correlating to Environmentals

- Failure rate regressions for **<u>HH1</u>**
  1. Discover trends – variables that change together
  2. Split into 4 groups **P1 – P4** (total population = 170K)

coefficient a

Popul.	%	Temp.	<u>RH</u>	CoV Temp.	CoV RH
P1	30.1				
P2	25.6				
P3	23.3				
P4	19.6				

$$y = a^*x + b$$

coefficient b

Popul.	%	Temp.	RH	CoV Temp.	CoV RH
P1	30.1	$5.1*10^{-5}$	$1.2*10^{-4}$	$-7.9*10^{-3}$	$-6.5*10^{-3}$
P2	25.6	$-1.9*10^{-5}$	$1.0*10^{-4}$	$-9.0*10^{-3}$	$-3.7*10^{-3}$
P3	23.3	$1.4*10^{-3}$	$2.1*10^{-4}$	$-4.9*10^{-2}$	$-4.4*10^{-2}$
P4	19.6	$1.7*10^{-3}$	$4.4*10^{-4}$	$-1.3*10^{-1}$	$-8.0*10^{-2}$

$$y = a^*e^{(b^*x)}$$

Root Causes: Correlating to Environmentals

- Failure rate regressions for **<u>HH1</u>**
  1. Discover trends – variables that change together
  2. Split into 4 groups **P1 – P4** (total population = 170K)

coefficient a

Popul.	%	Temp.	<u>RH</u>	CoV Temp.	CoV RH
P1	30.1				
P2	25.6				
P3	23.3				
P4	19.6				

$$y = a^{*}x + b$$

coefficient b

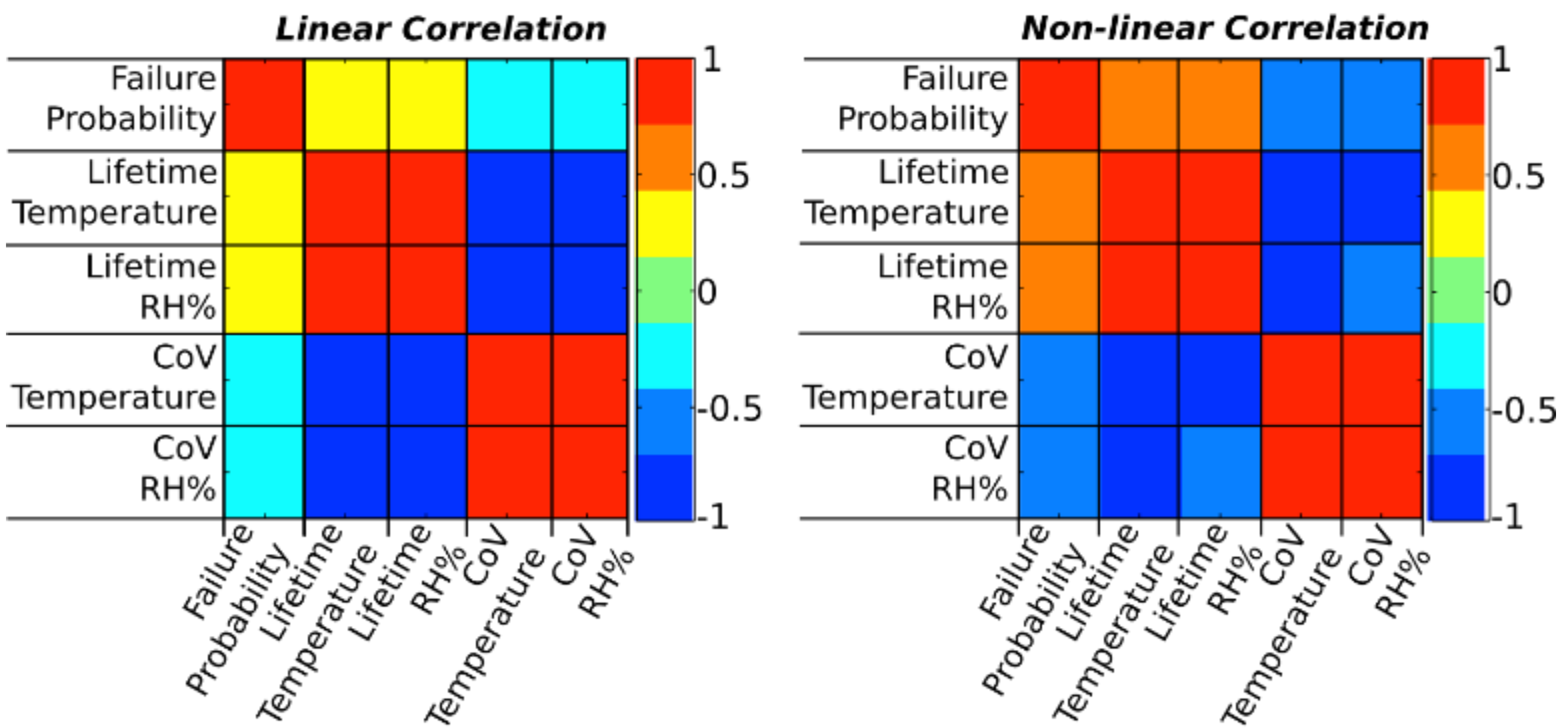
Popul.	%	Temp.	<u>RH</u>	CoV Temp.	CoV RH
P1	30.1				
P2	25.6				
P3	23.3				
P4	19.6				

$$y = a^{*}e^{(b^{*}x)}$$

➢ **<u>RH% seems to have the strongest impact</u>**

Root Causes: Correlating to Environmentals

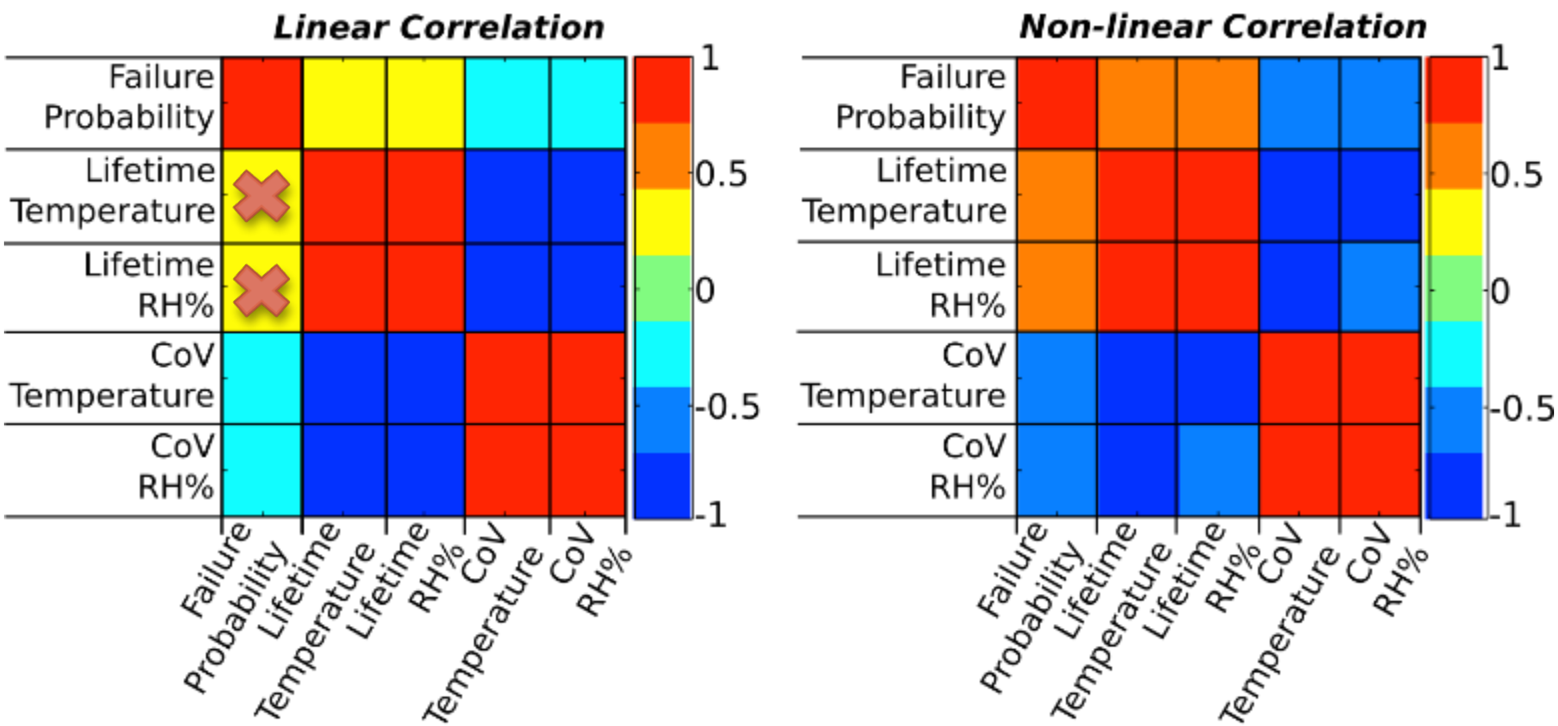
- Failure rate correlations for <u>**HH1**</u>
  1. Discover trends – variables that change together
  2. Correlation matrix



a) Pearson's correlation ***b) Spearman's correlation***

Root Causes: Correlating to Environmentals

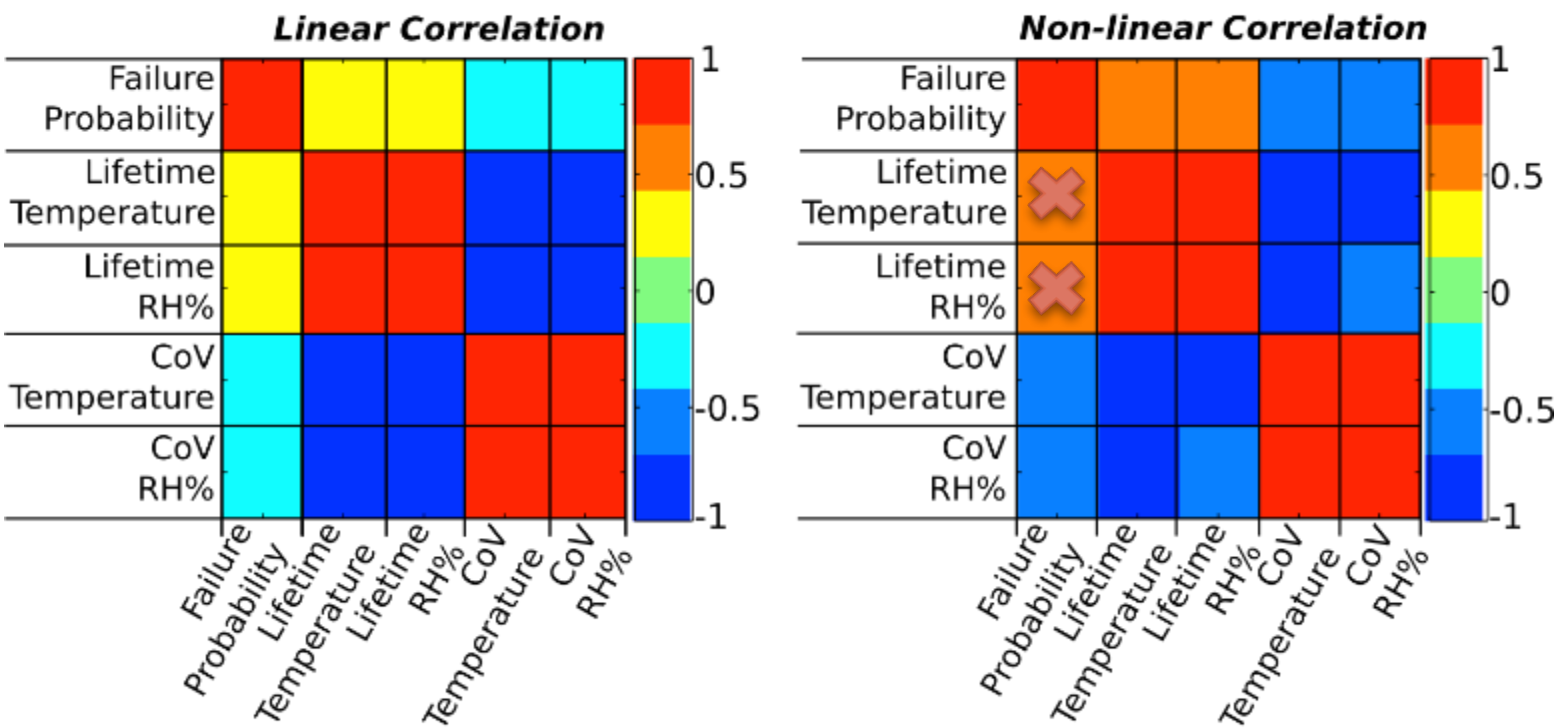
- Failure rate correlations for **<u>HH1</u>**
  1. Discover trends – variables that change together
  2. Correlation matrix



a) Pearson's correlation *b) Spearman's correlation*

Root Causes: Correlating to Environmentals

- Failure rate correlations for **<u>HH1</u>**
  1. Discover trends – variables that change together
  2. Correlation matrix



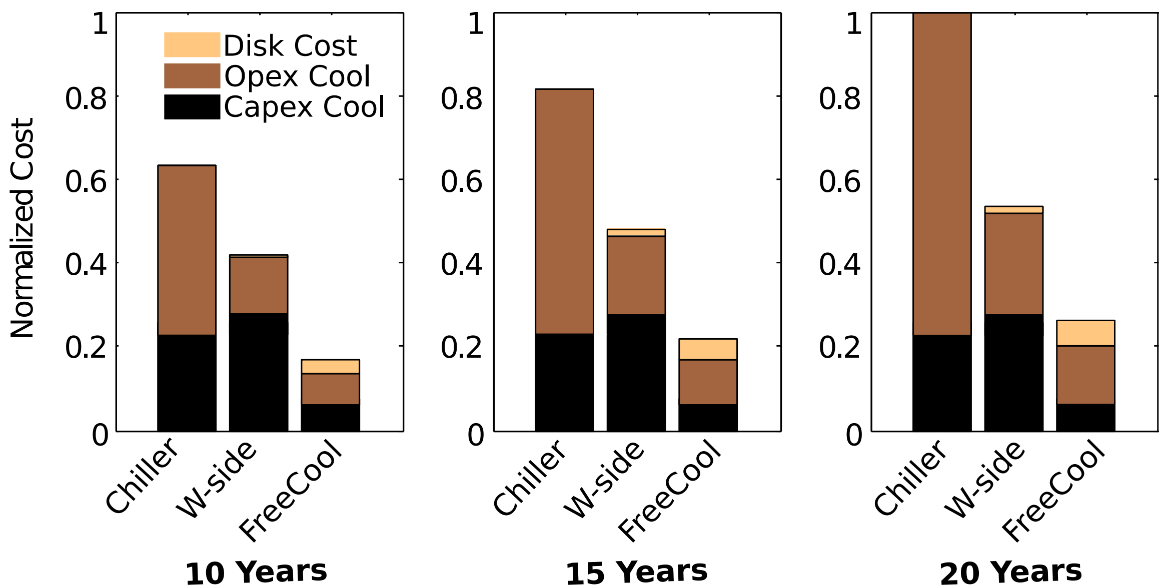
a) Pearson's correlation *b) Spearman's correlation*

Contributions and Roadmap

1. Impact of environmentals on disk failure rates
2. Root causes
3. **Cooling vs reliability vs cost tradeoffs**
4. Modeling of failure rates
5. Design considerations

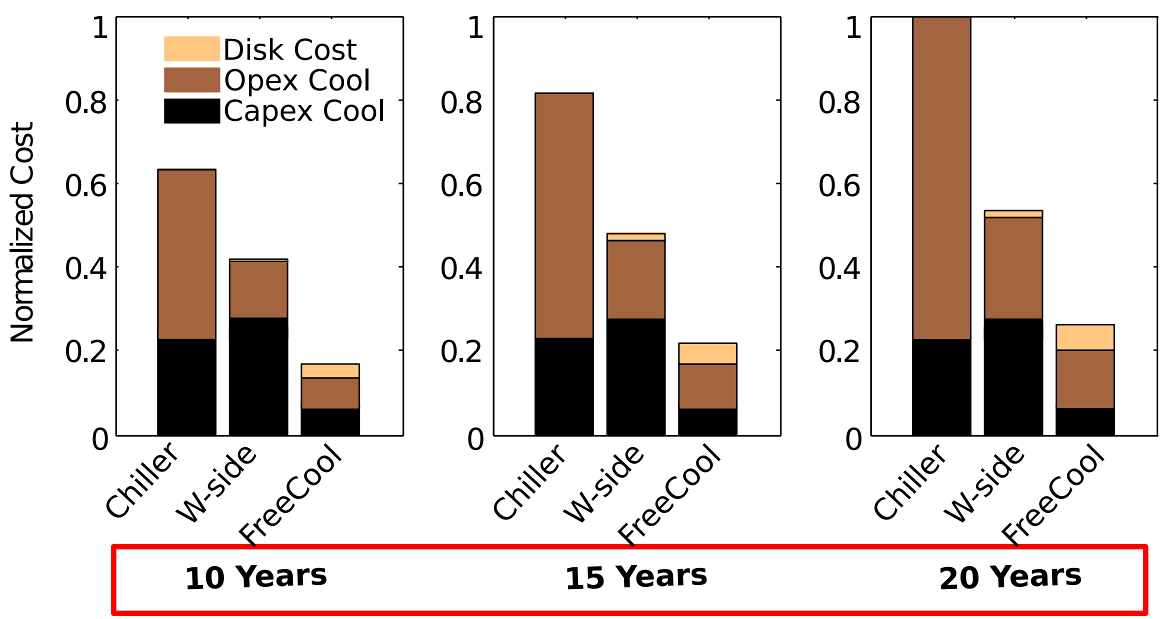
Cooling-Related Cost Tradeoffs

- Cooling technologies vs costs
  - Free cooling results in higher HDD costs
  - Operator will pay the extra HDD costs (or warranties)



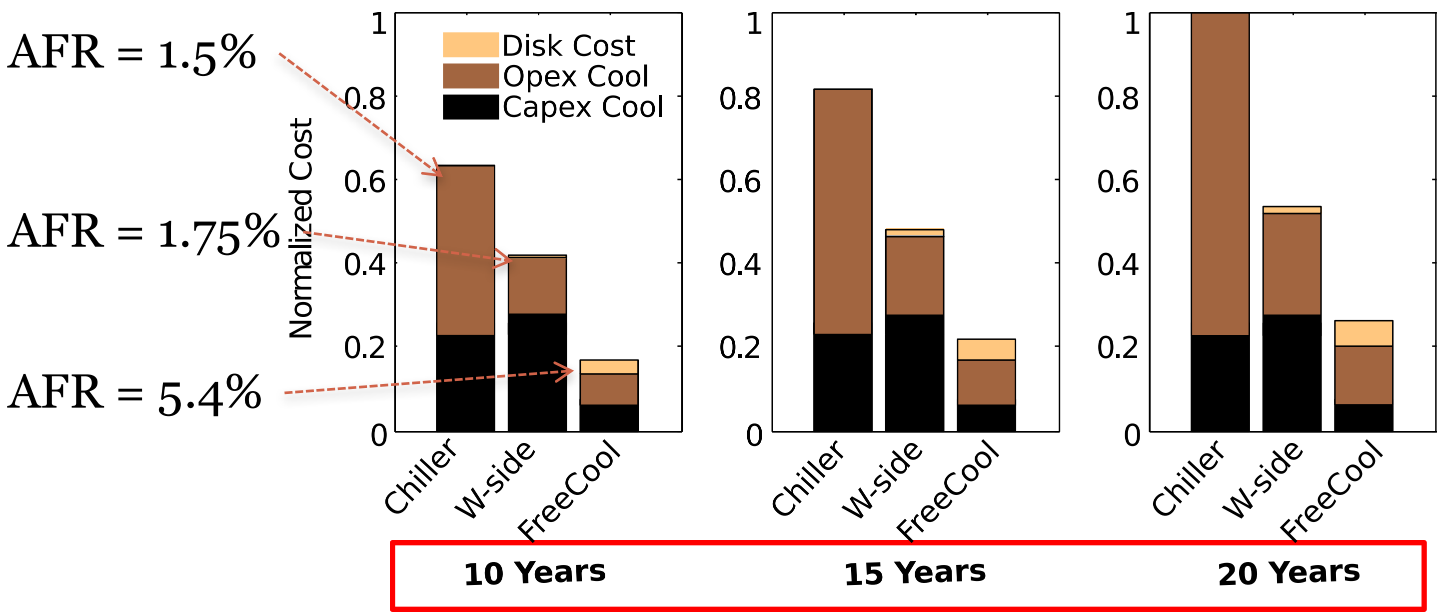
Cooling-Related Cost Tradeoffs

- Cooling technologies vs costs
  - Free cooling results in higher HDD costs
  - Operator will pay the extra HDD costs (or warranties)



Cooling-Related Cost Tradeoffs

- Cooling technologies vs costs
  - Free cooling results in higher HDD costs
  - Operator will pay the extra HDD costs (or warranties)



Summary of Observations

1. Failures correlate with environmentals
   - RH appears to be the dominant effect
2. Impact different parts of the HDD
   - Temperature → mechanical & controller
   - RH → controller

Summary of Observations

1. Failures correlate with environmentals
   - RH appears to be the dominant effect
2. Impact different parts of the HDD
   - Temperature → mechanical & controller
   - RH → controller
3. Failures do not occur instantly
   - Match a lifetime model
   - Lifetime is "consumed" depending on environmentals
4. Free cooling still cheaper, despite the higher AFRs

Contributions and Roadmap

1. Impact of environmentals on disk failure rates
2. Root causes
3. Cooling vs reliability vs cost tradeoffs
4. **<u>Modeling of failure rates</u>**
5. Design considerations

Model Construction

- Estimate AFRs
  - Various server and datacenter designs/conditions/locations

Model Construction

- Estimate AFRs
  - Various server and datacenter designs/conditions/locations

1. Modeling HDD mechanical degradation

$$AF_T = e^{\frac{E_a}{k} \cdot \left(\frac{1}{\overline{\overline{T_b}}} - \frac{1}{\overline{\overline{T_e}}} \right)}$$

2. Modeling corrosion
   - Extension of Arrhenius equation
   - Accounts for combined temperature and RH effects

$$CR(\overline{T}, \overline{RH}) = \text{const} \cdot e^{\left(\frac{-E_a}{k \cdot \overline{T}} \right)} \cdot e^{(b \cdot \overline{RH}) + \left(\frac{c \cdot \overline{RH}}{k \cdot \overline{T}} \right)}$$

Model Construction

- Lifetime Acceleration Factor (AF)
  - Compared to a baseline (AFR=1.5% @25C and 50% RH)
  - AF_1 : Temperature - AF_2: RH and Temperature

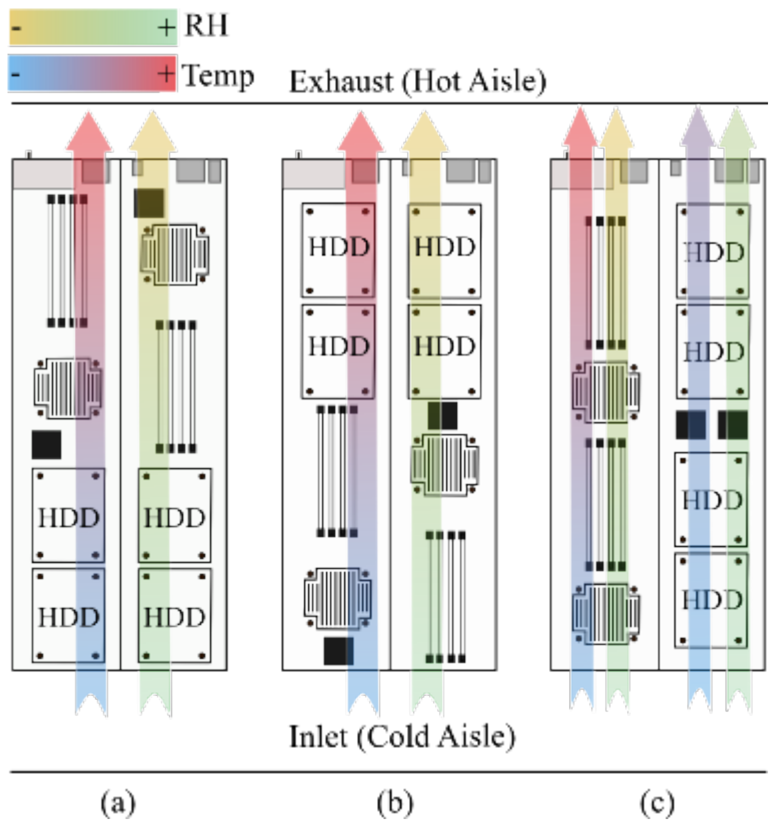
- Validation
  - Collect hourly environmentals in other datacenters
  - Use the model constructed in **P1** to predict failure rates
  - **Validated with P2, P3, P4, and CD3, HD1 (within 5%)**

Contributions and Roadmap

1. Impact of environmentals on disk failure rates
2. Root causes
3. Cooling vs reliability vs cost tradeoffs
4. Modeling of failure rates
5. **Design considerations**

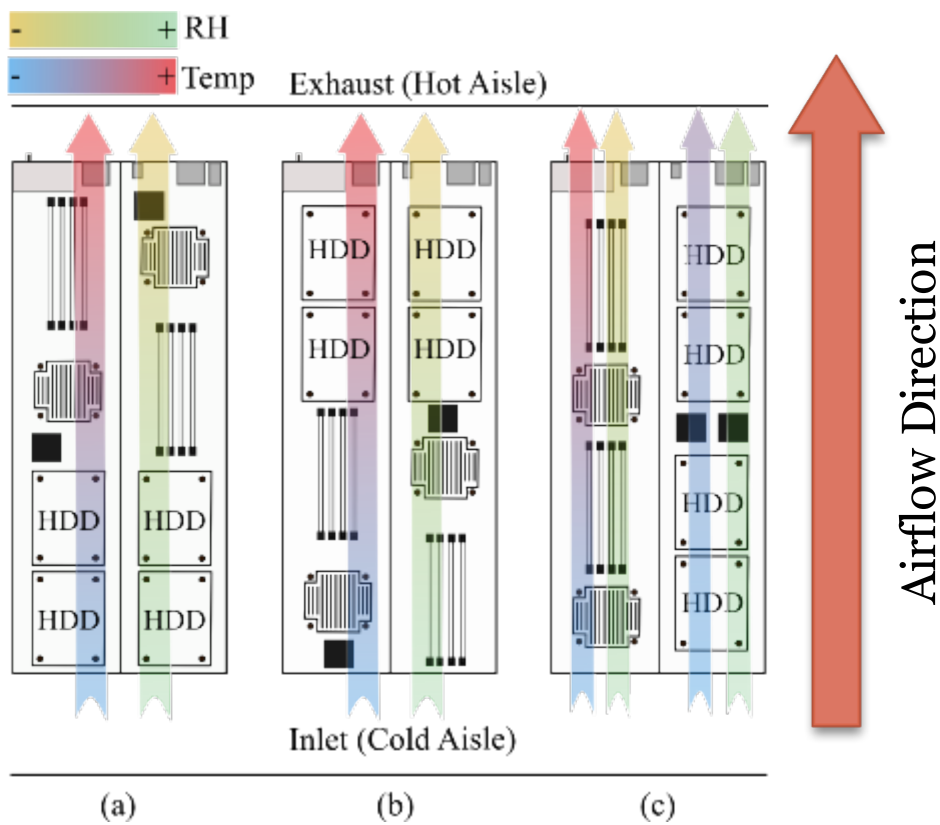
Design Considerations

- Disk placement affects HDD failure rates



Design Considerations

- Disk placement affects HDD failure rates

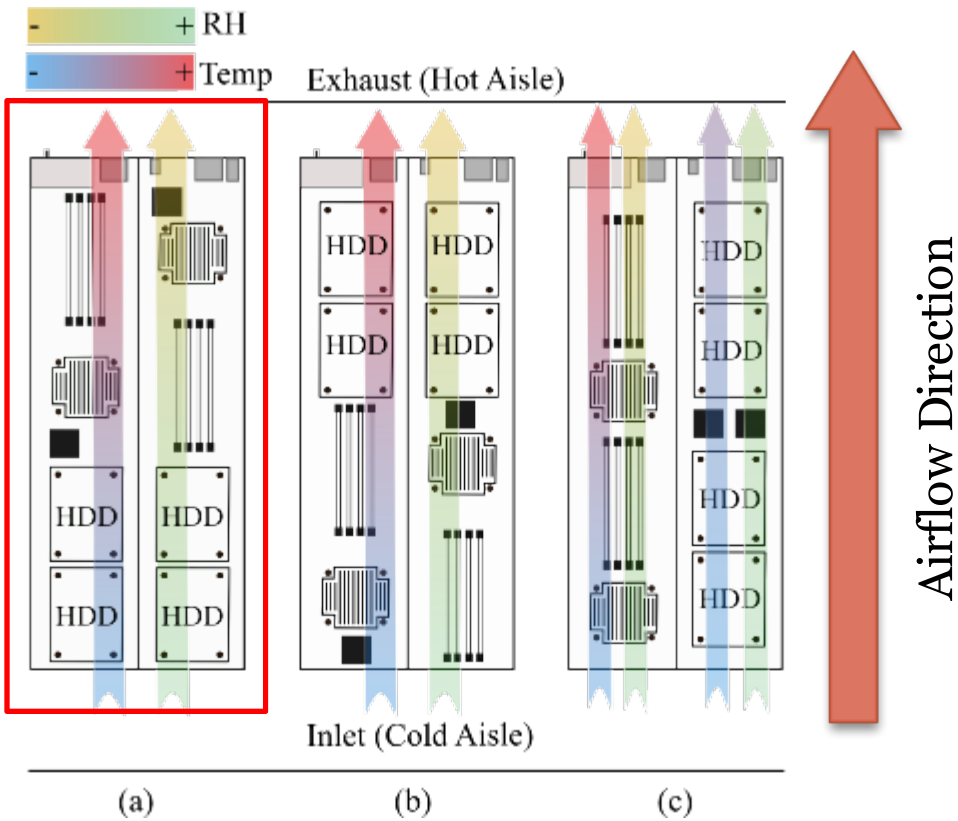


Design Considerations

- Disk placement affects HDD failure rates

Disk at the front

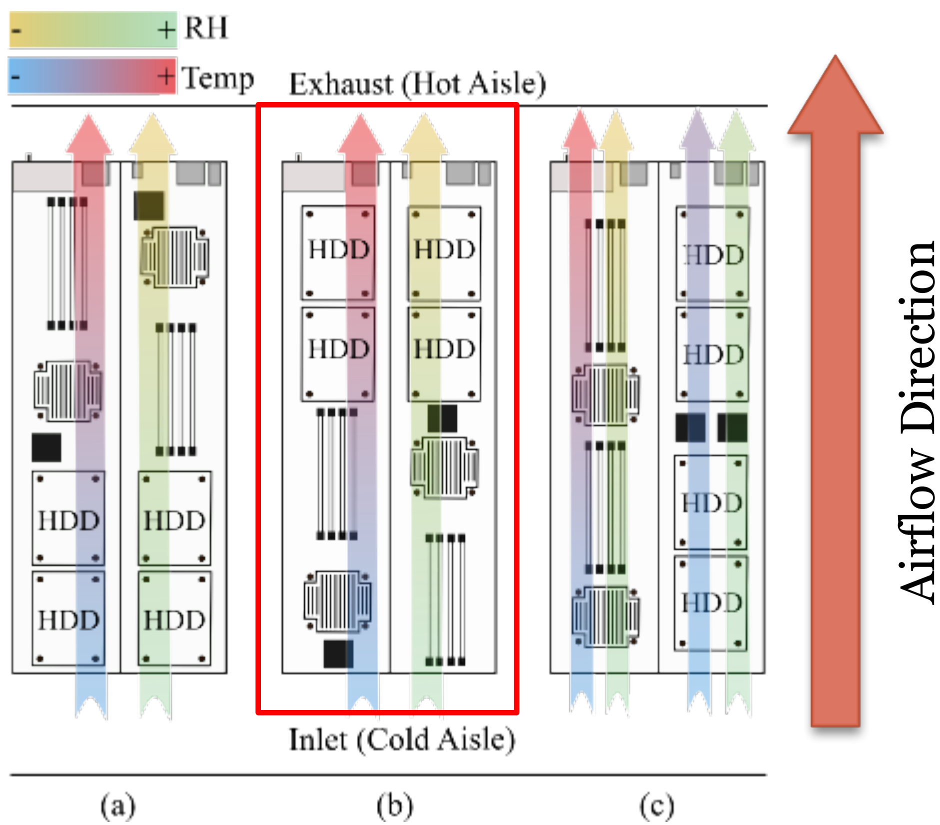
- **Low Temp**
- **High RH%**



Design Considerations

- Disk placement affects HDD failure rates

Disk at the back
- **High Temp**
- **Low RH%**

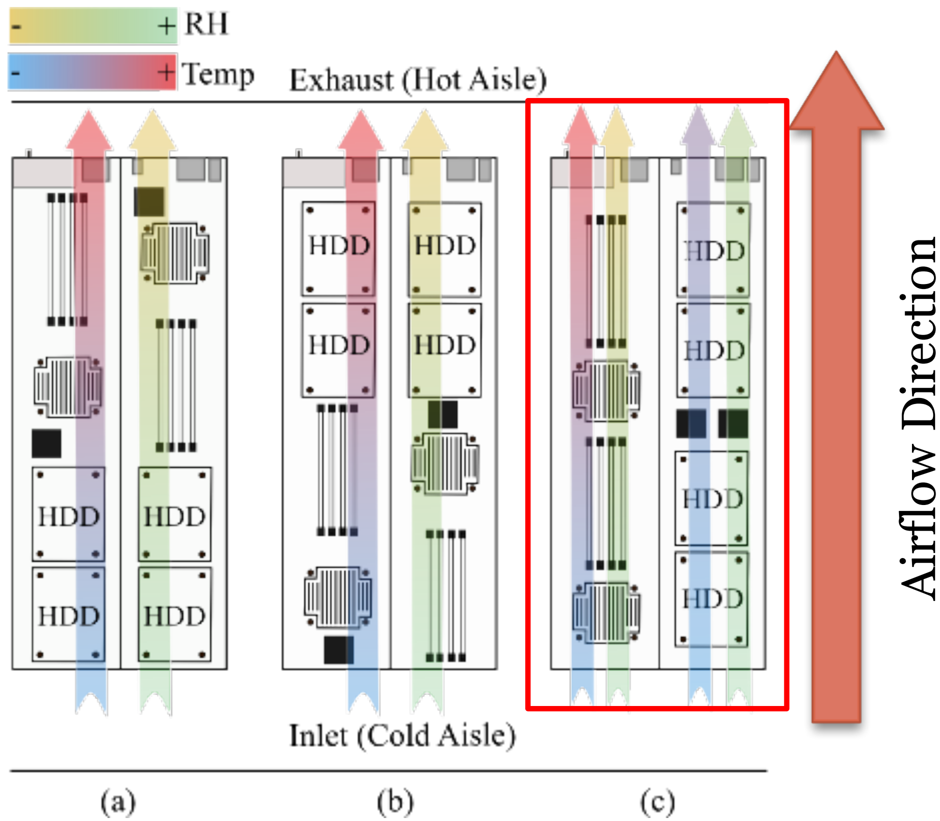


Design Considerations

- Disk placement affects HDD failure rates

Disk at the side

- **Var. Temp**
- **Var. RH%**



Conclusions

- Explored HDD reliability vs environmentals
  - 9 datacenters with 3 cooling technologies, 1M disks
  - AFRs impacted by enviromentals, especially high RH
  - Tradeoff favors free cooling: costs down, despite higher AFRs
- Developed an accurate model from real failure data
  - Combines corrosion and temperature
- Learned lessons
  - Server layout has a significant impact on HDD AFRs
  - More lessons in the paper

Thank you

Questions?

RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

Microsoft