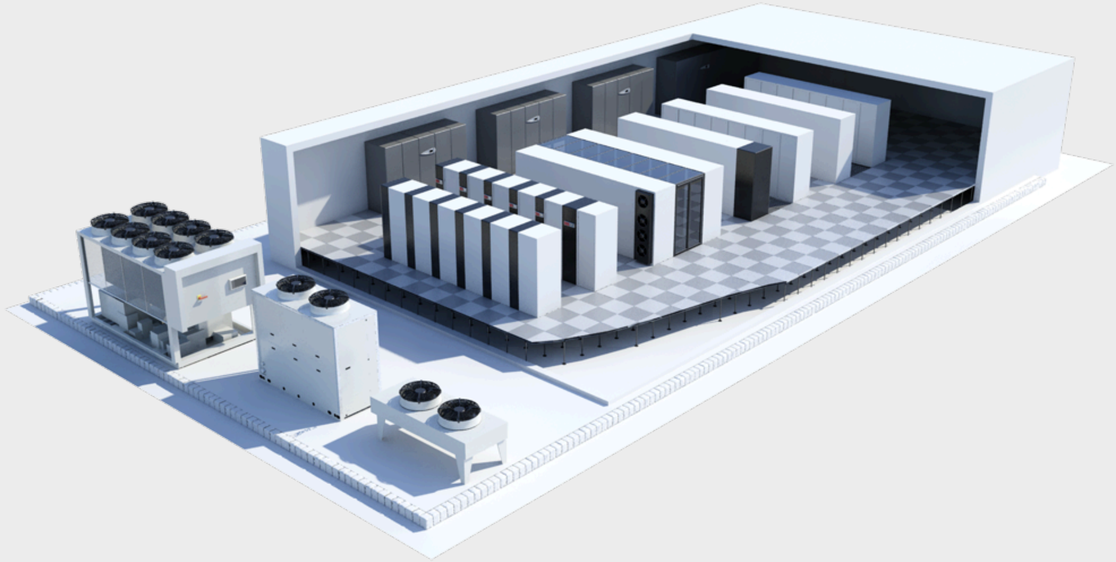
Reduced Cooling Redundancy: A New Security Vulnerability in a Hot Data Center

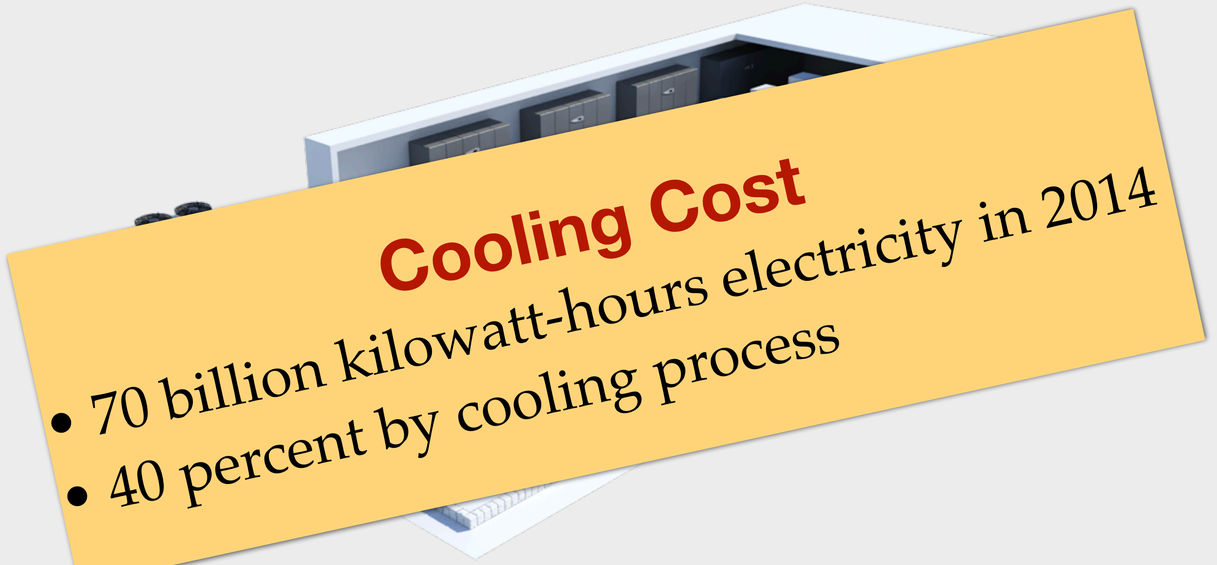
Xing Gao[1,2], Zhang Xu[2], Haining Wang[1], Li Li[3], Xiaorui Wang[3]

1. University of Delaware
2. College of William and Mary
3. Ohio State University

Data Center Cooling Systems



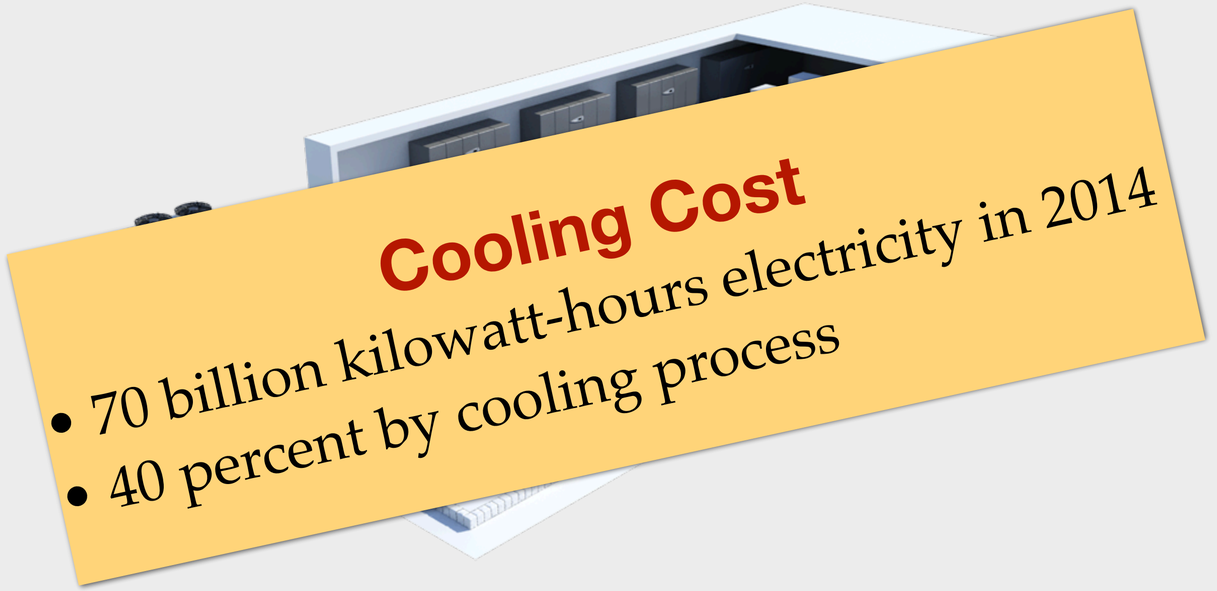
Data Center Cooling Systems



Cooling Cost

- 70 billion kilowatt-hours electricity in 2014
- 40 percent by cooling process

Data Center Cooling Systems

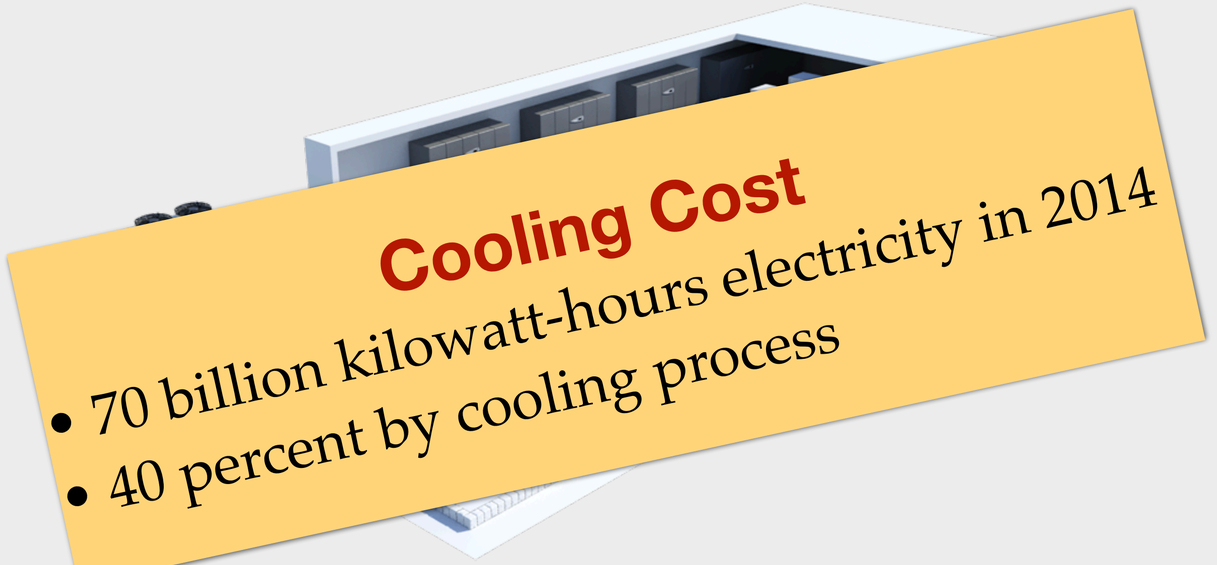


Cooling Cost

- 70 billion kilowatt-hours electricity in 2014
- 40 percent by cooling process

CRAC Cooling

Data Center Cooling Systems



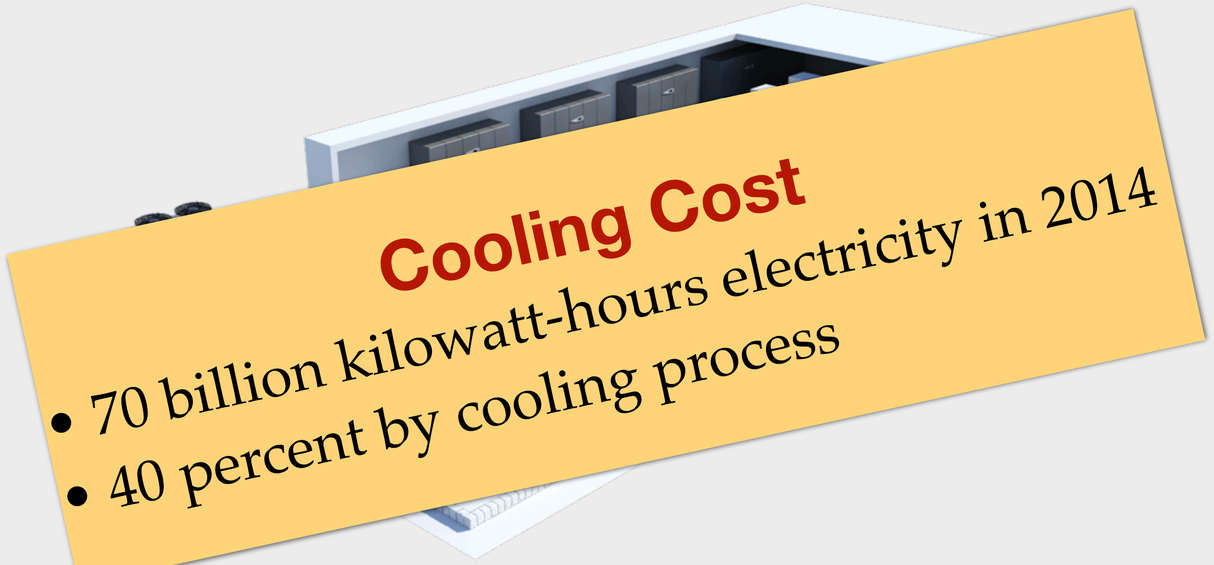
Cooling Cost

- 70 billion kilowatt-hours electricity in 2014
- 40 percent by cooling process

CRAC Cooling

Liquid Cooling

Data Center Cooling Systems



Cooling Cost

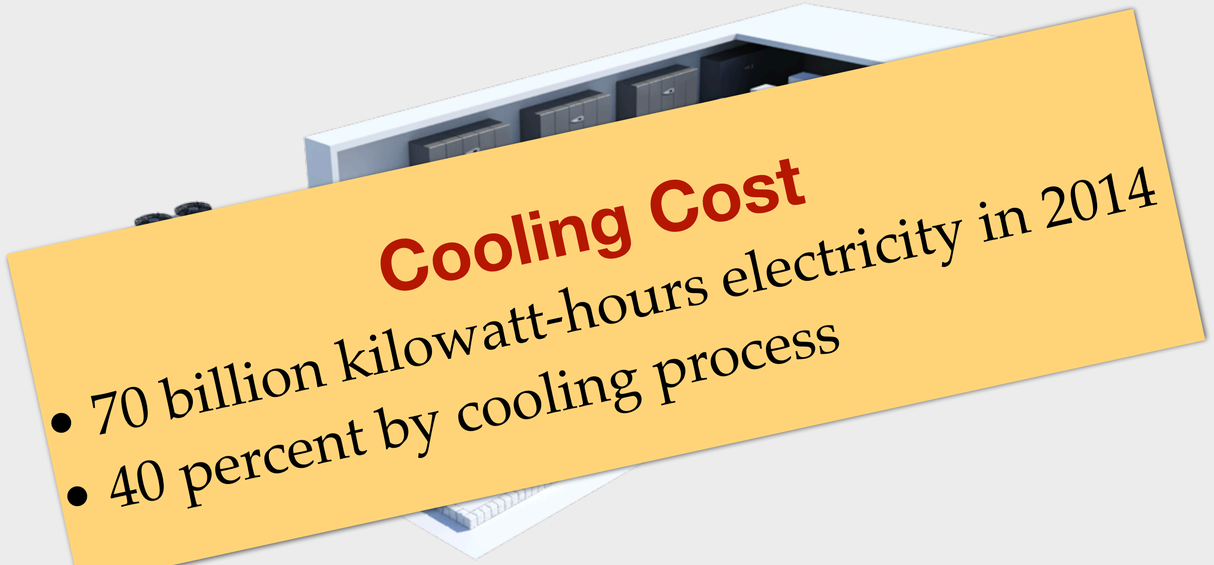
- 70 billion kilowatt-hours electricity in 2014
- 40 percent by cooling process

CRAC Cooling

Liquid Cooling

Free Air Cooling

Data Center Cooling Systems



Cooling Cost

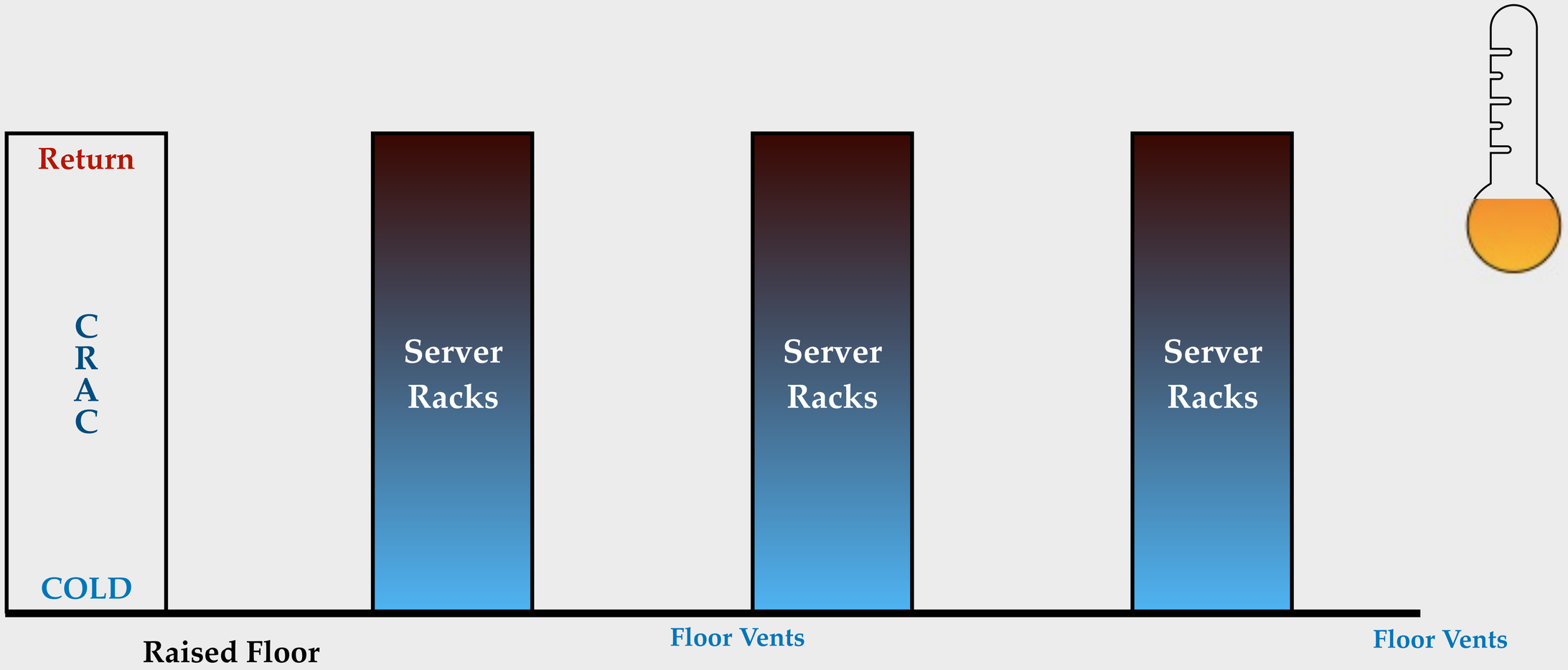
- 70 billion kilowatt-hours electricity in 2014
- 40 percent by cooling process

CRAC Cooling

Liquid Cooling

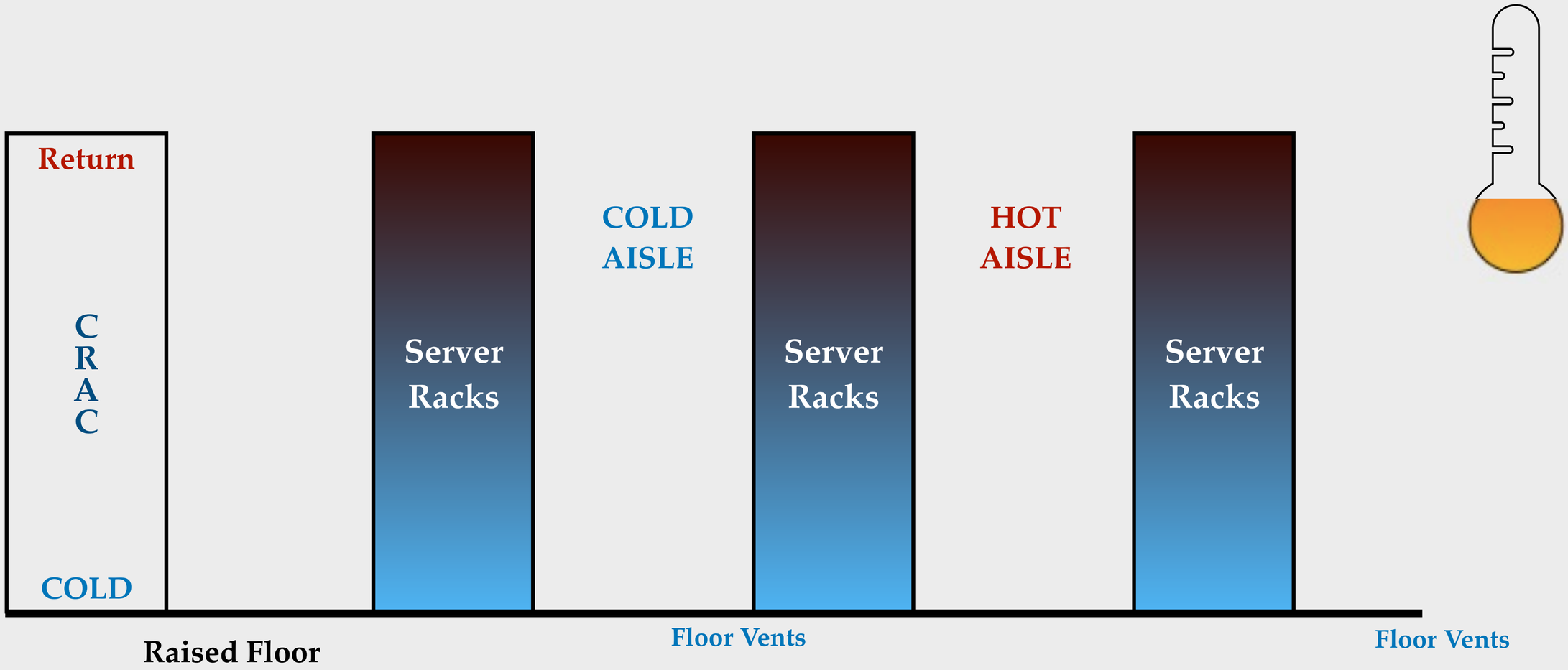
Free Air Cooling

CRAC Cooling System



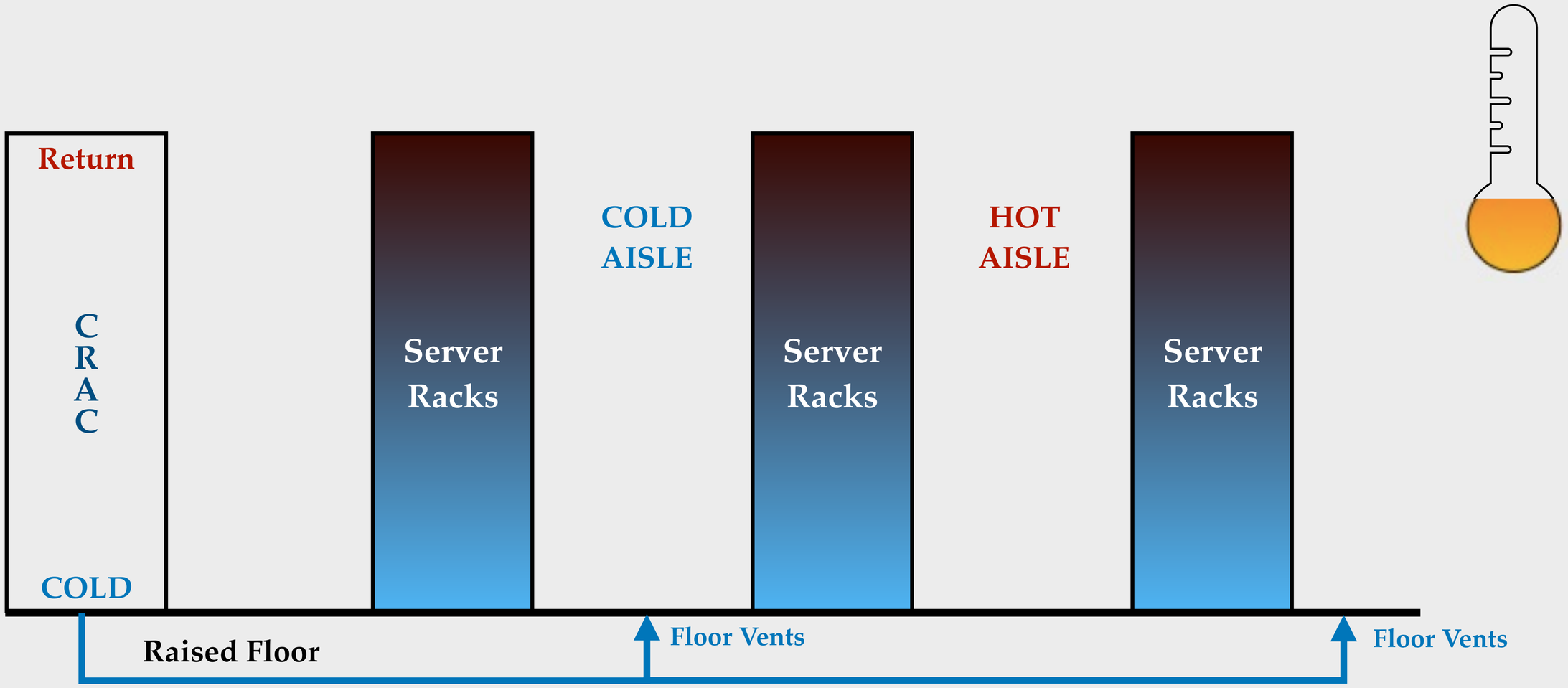
Credit: Data center cooling methods from Submer

CRAC Cooling System



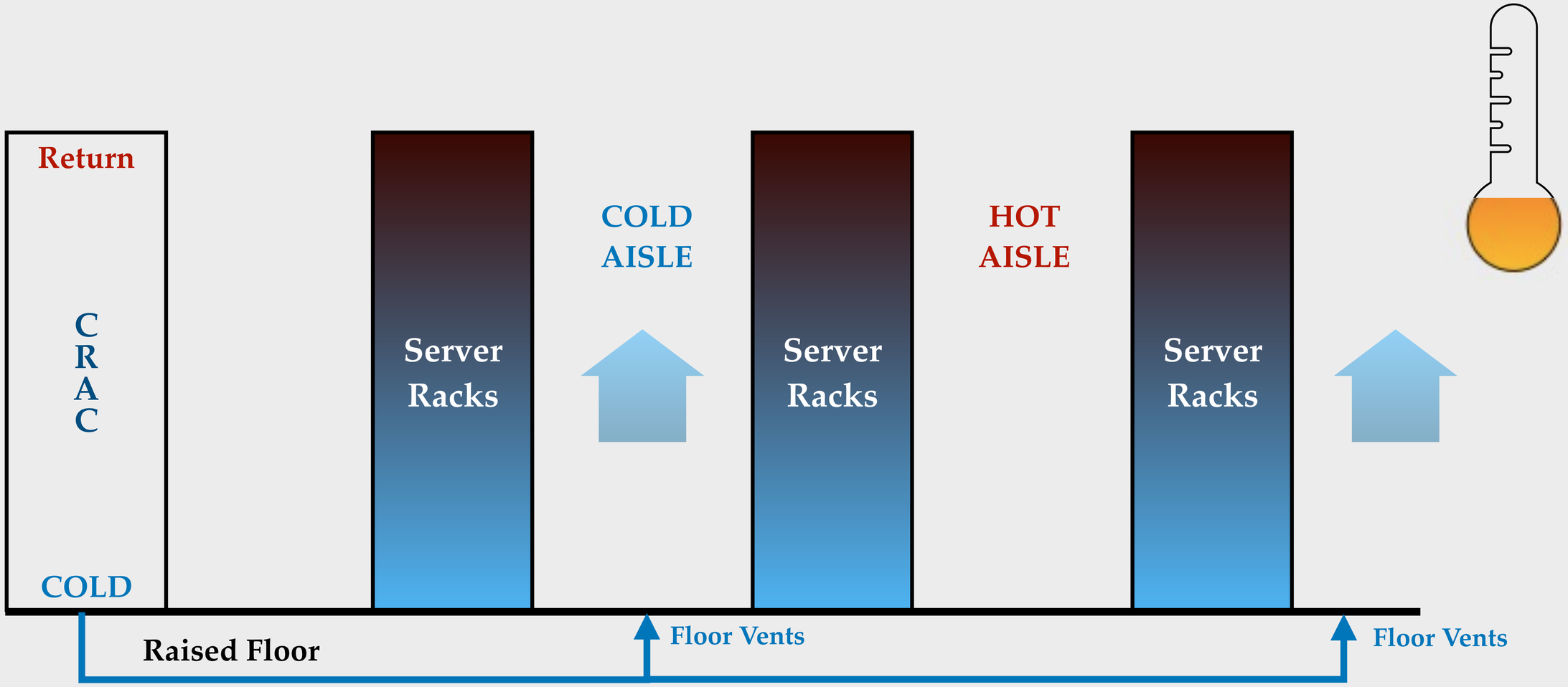
Credit: Data center cooling methods from Submer

CRAC Cooling System



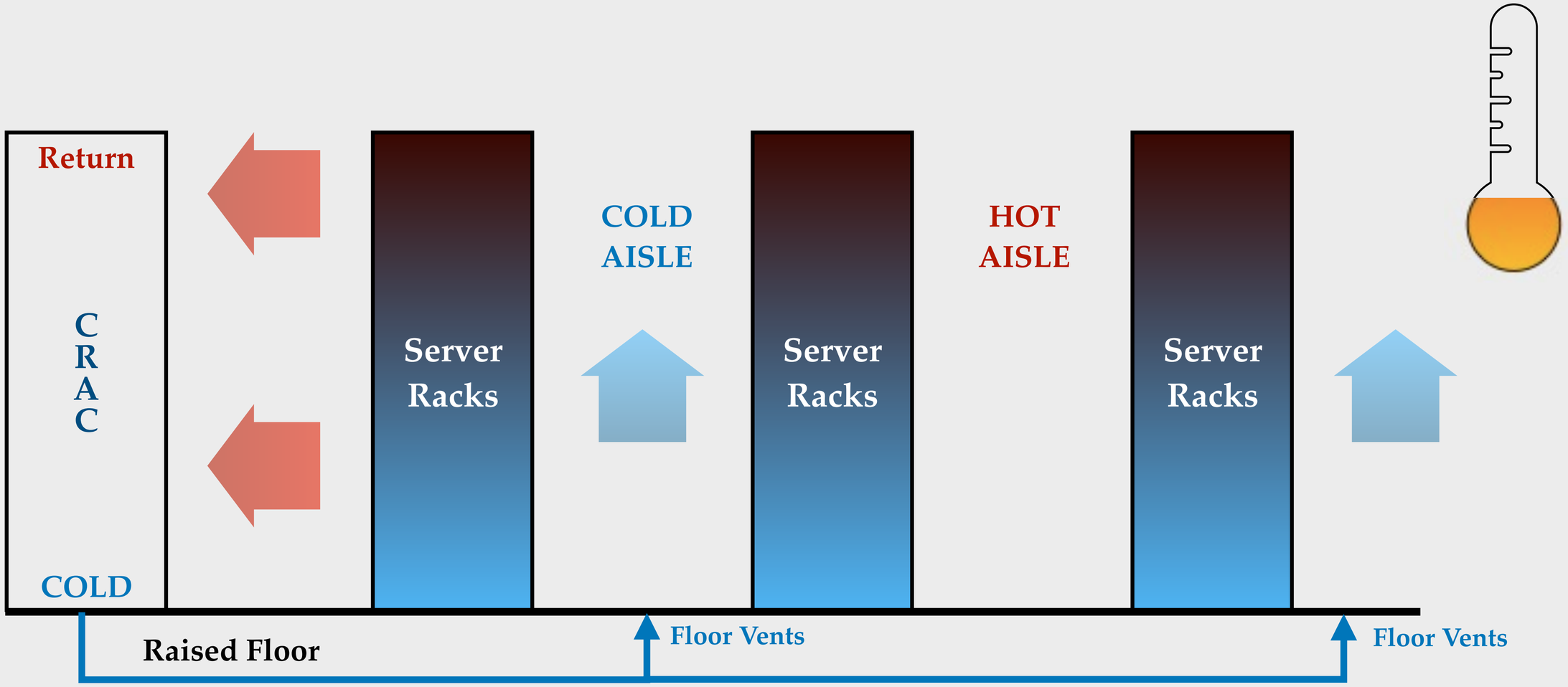
Credit: Data center cooling methods from Submer

CRAC Cooling System



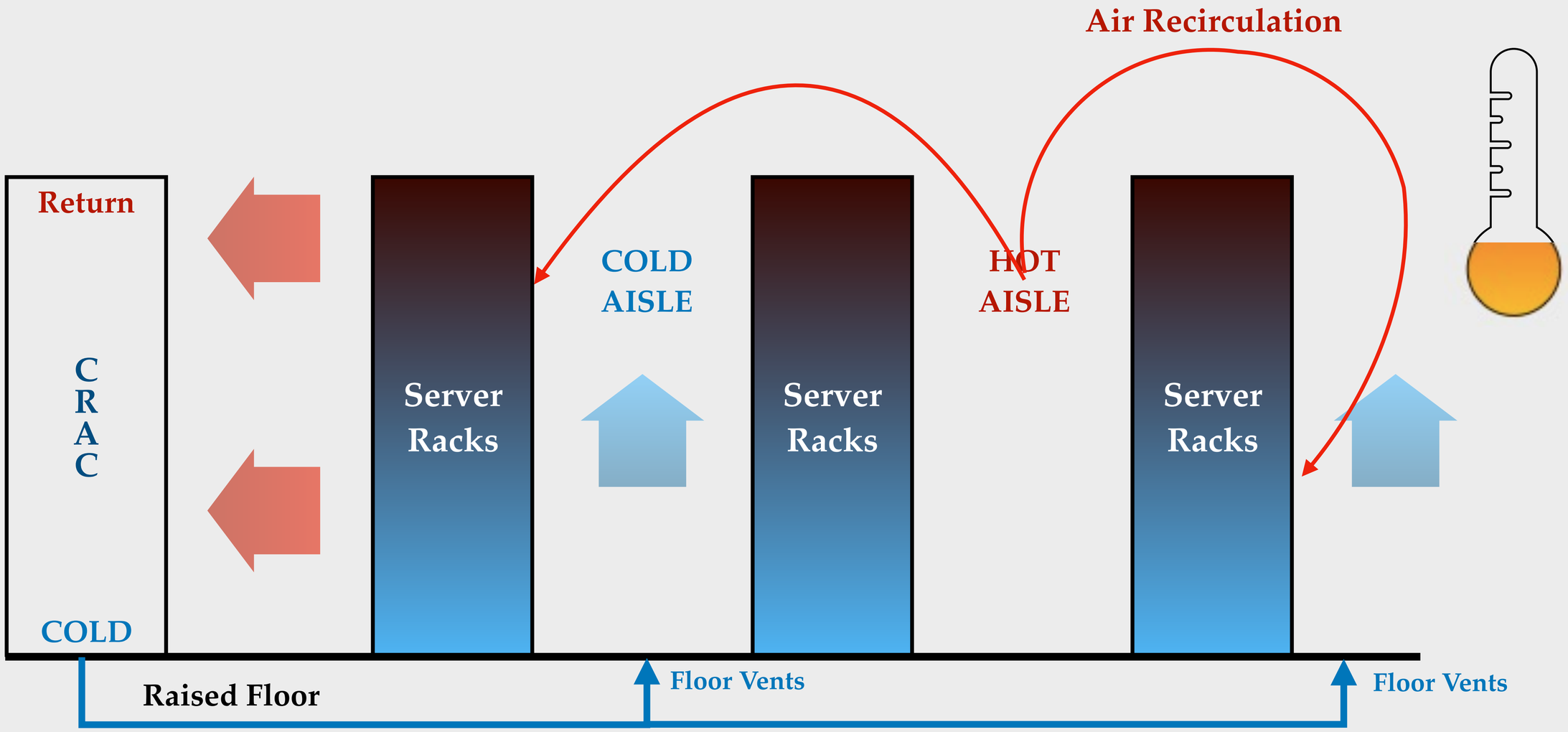
Credit: Data center cooling methods from Submer

CRAC Cooling System



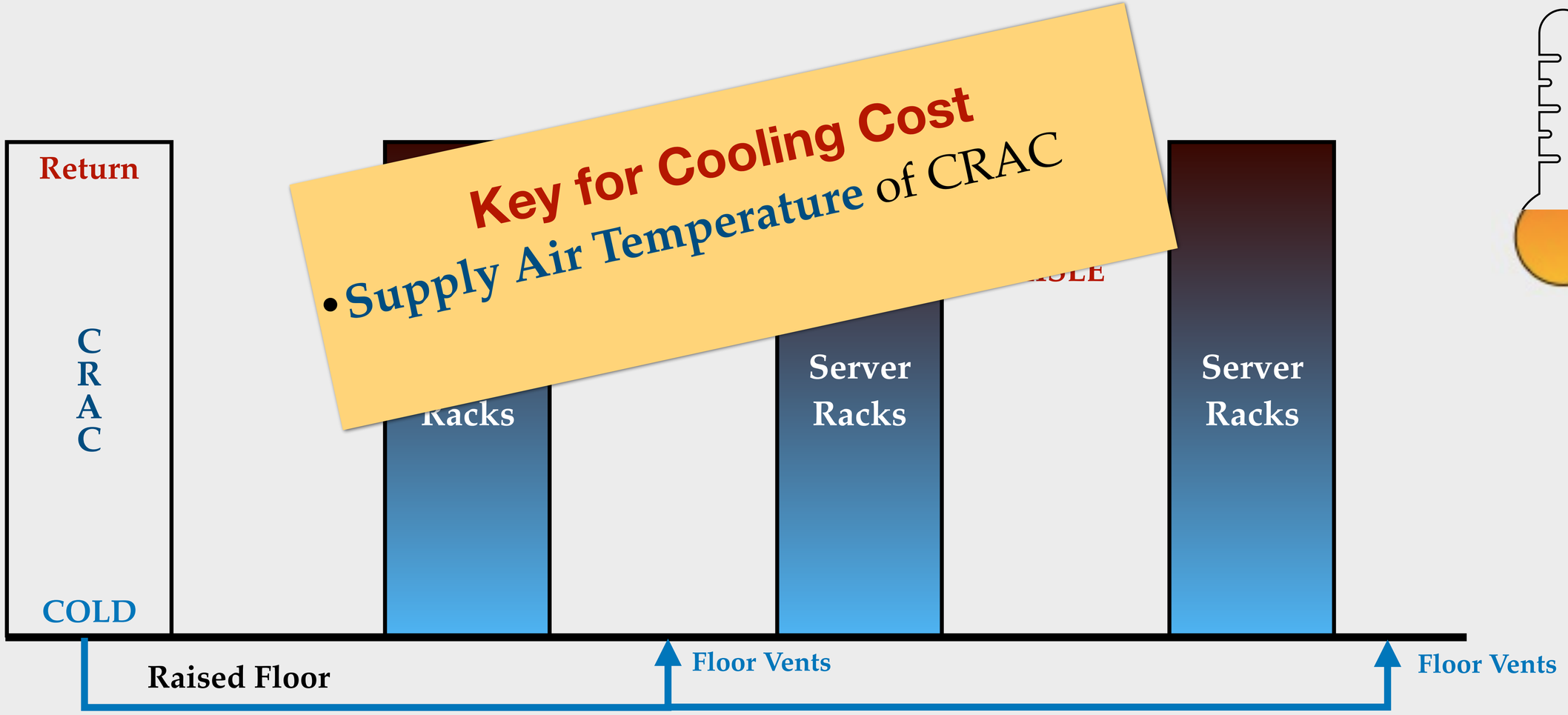
Credit: Data center cooling methods from Submer

CRAC Cooling System



Credit: Data center cooling methods from Submer

Cooling Cost

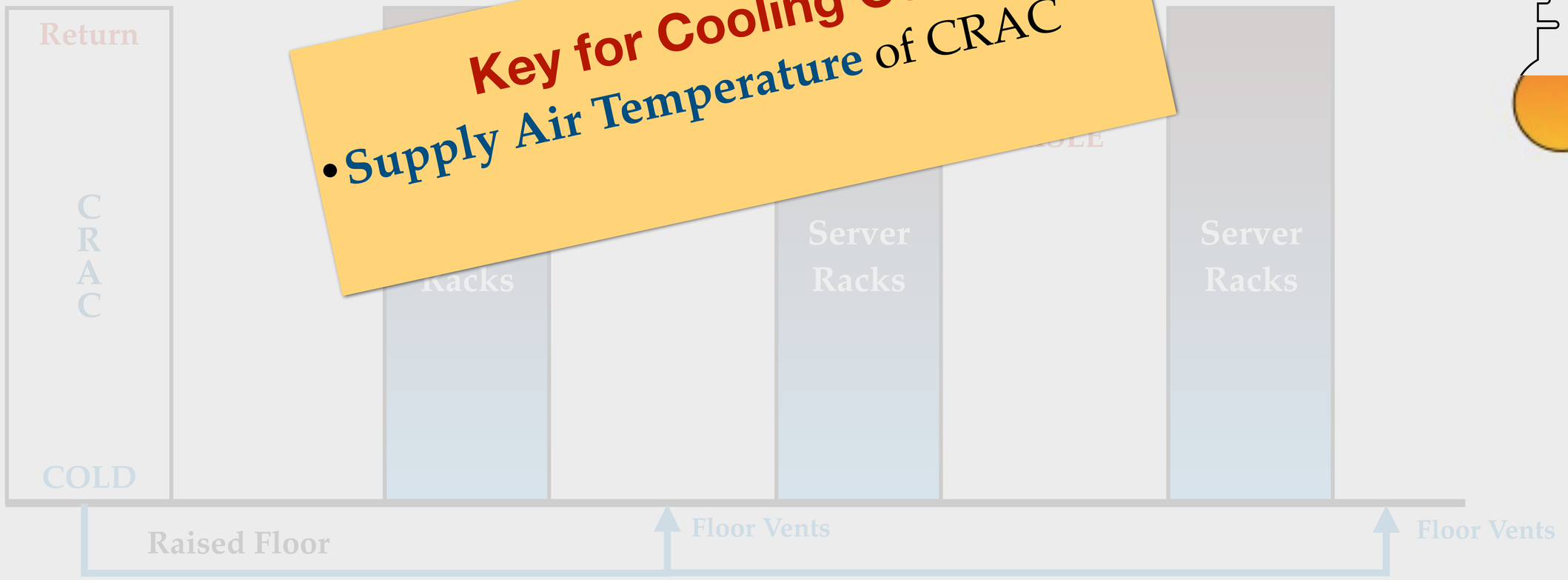


Key for Cooling Cost

- **Supply Air Temperature** of CRAC

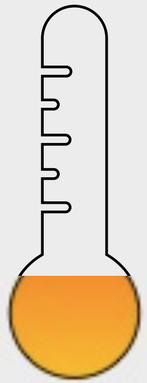
Credit: Data center cooling methods from Submer

Cooling Cost



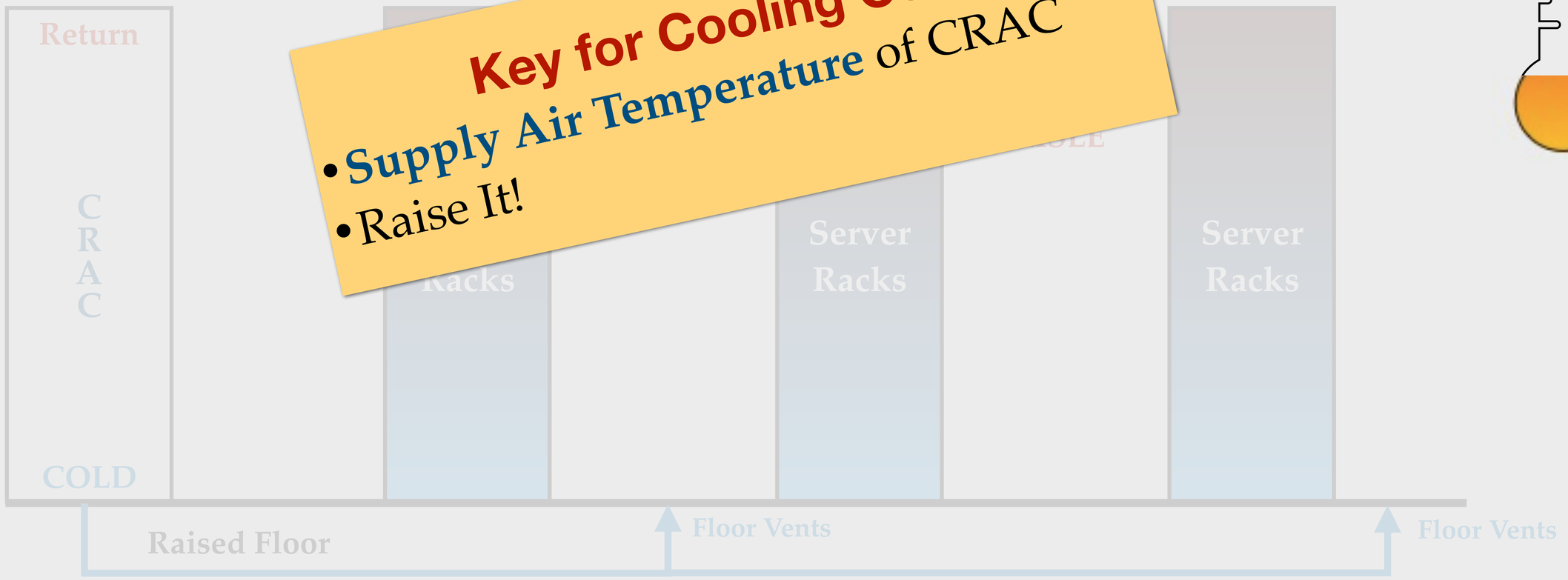
Key for Cooling Cost

- **Supply Air Temperature** of CRAC



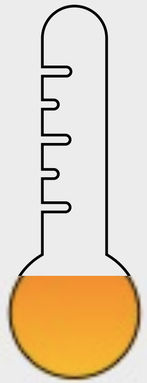
Credit: Data center cooling methods from Submer

Cooling Cost



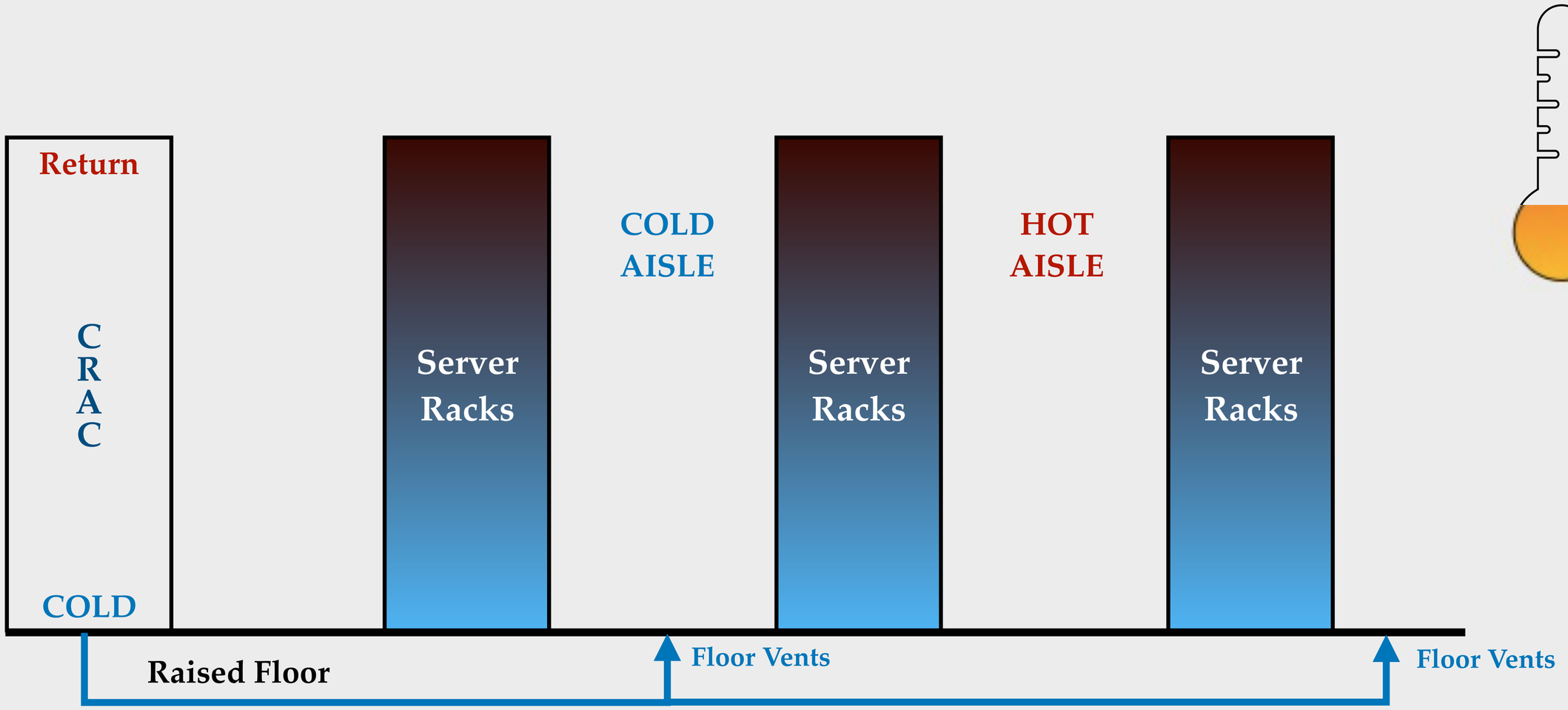
Key for Cooling Cost

- **Supply Air Temperature** of CRAC
- Raise It!



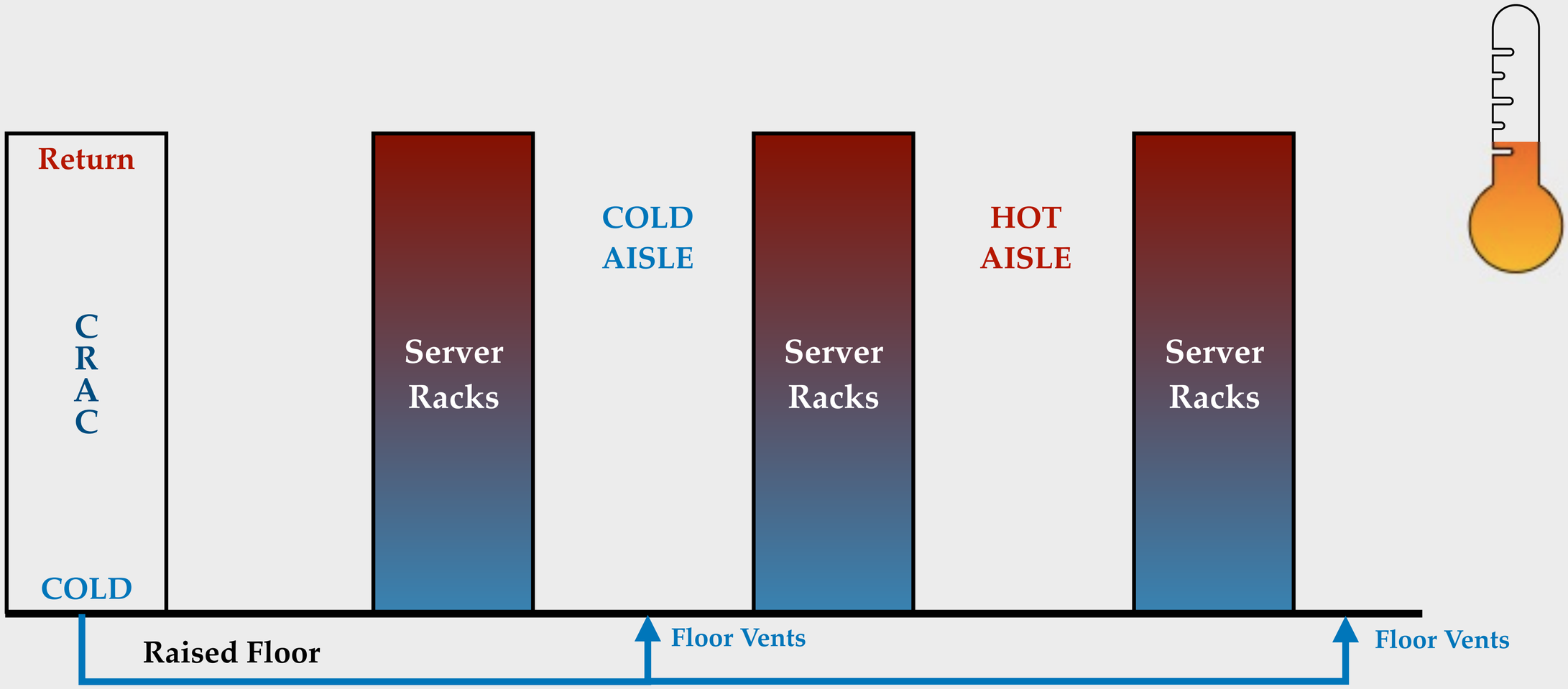
Credit: Data center cooling methods from Submer

Cooling Cost



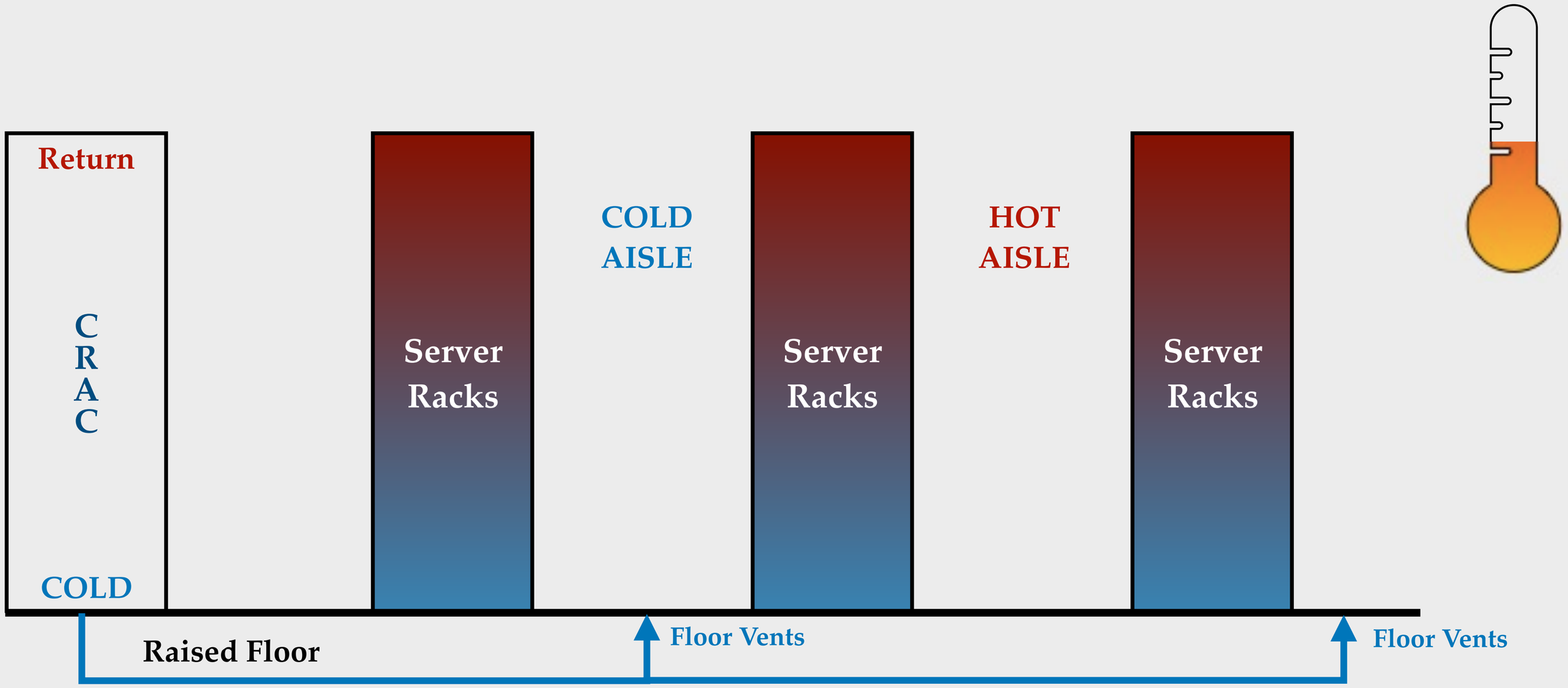
Credit: Data center cooling methods from Submer

Cooling Cost



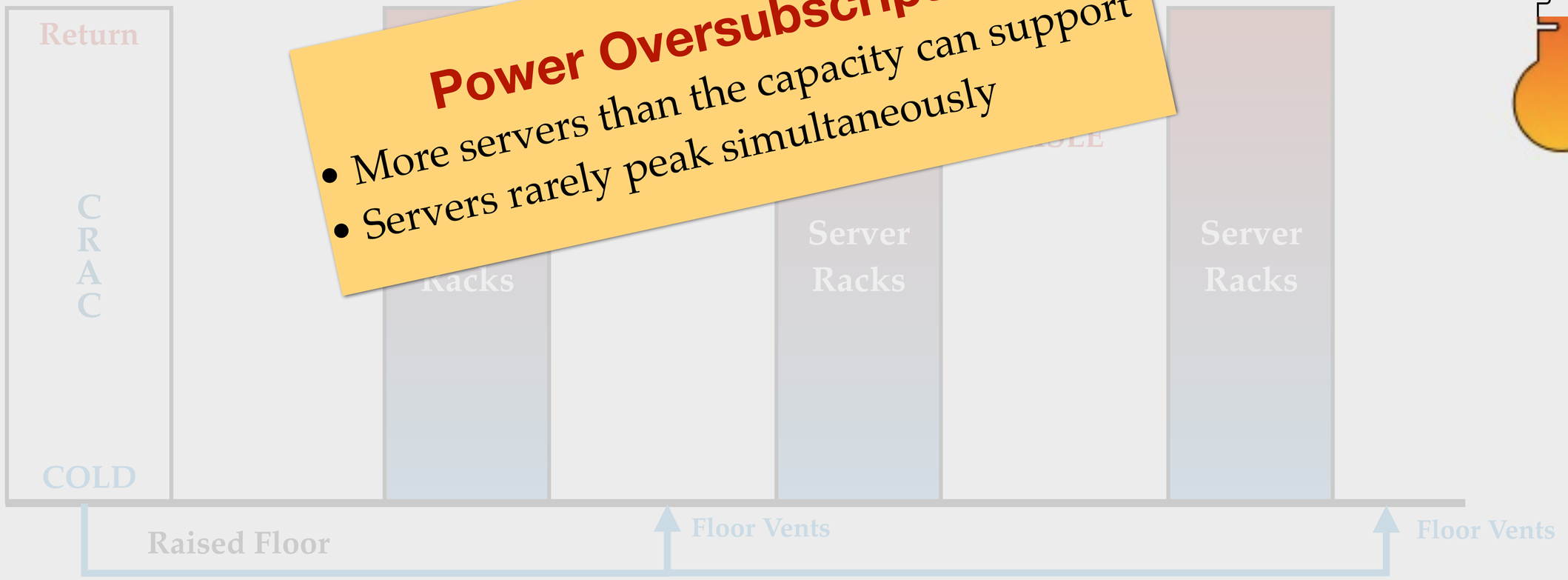
Credit: Data center cooling methods from Submer

Power Oversubscription



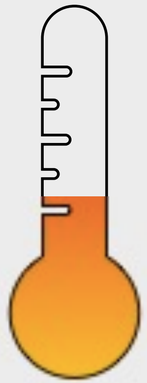
Credit: Data center cooling methods from Submer

Power Oversubscription



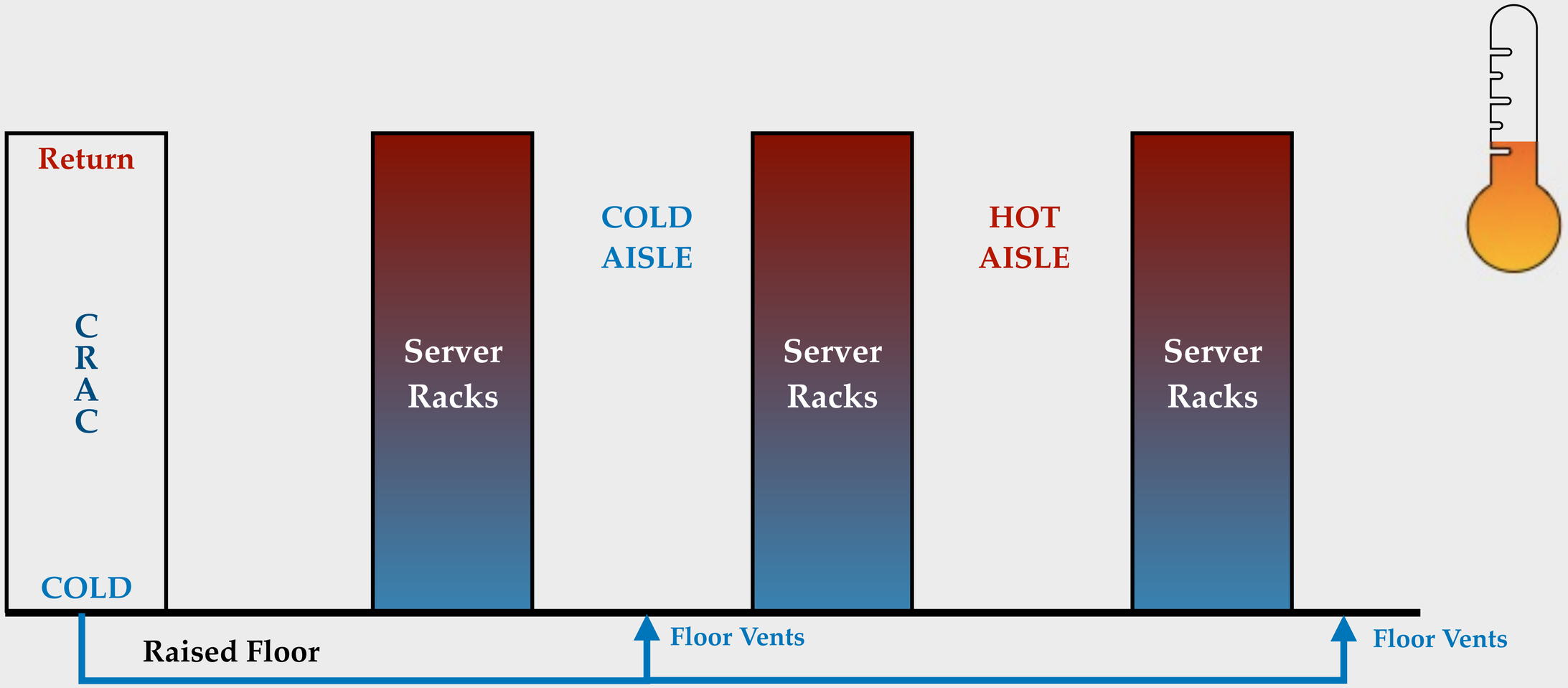
Power Oversubscription

- More servers than the capacity can support
- Servers rarely peak simultaneously



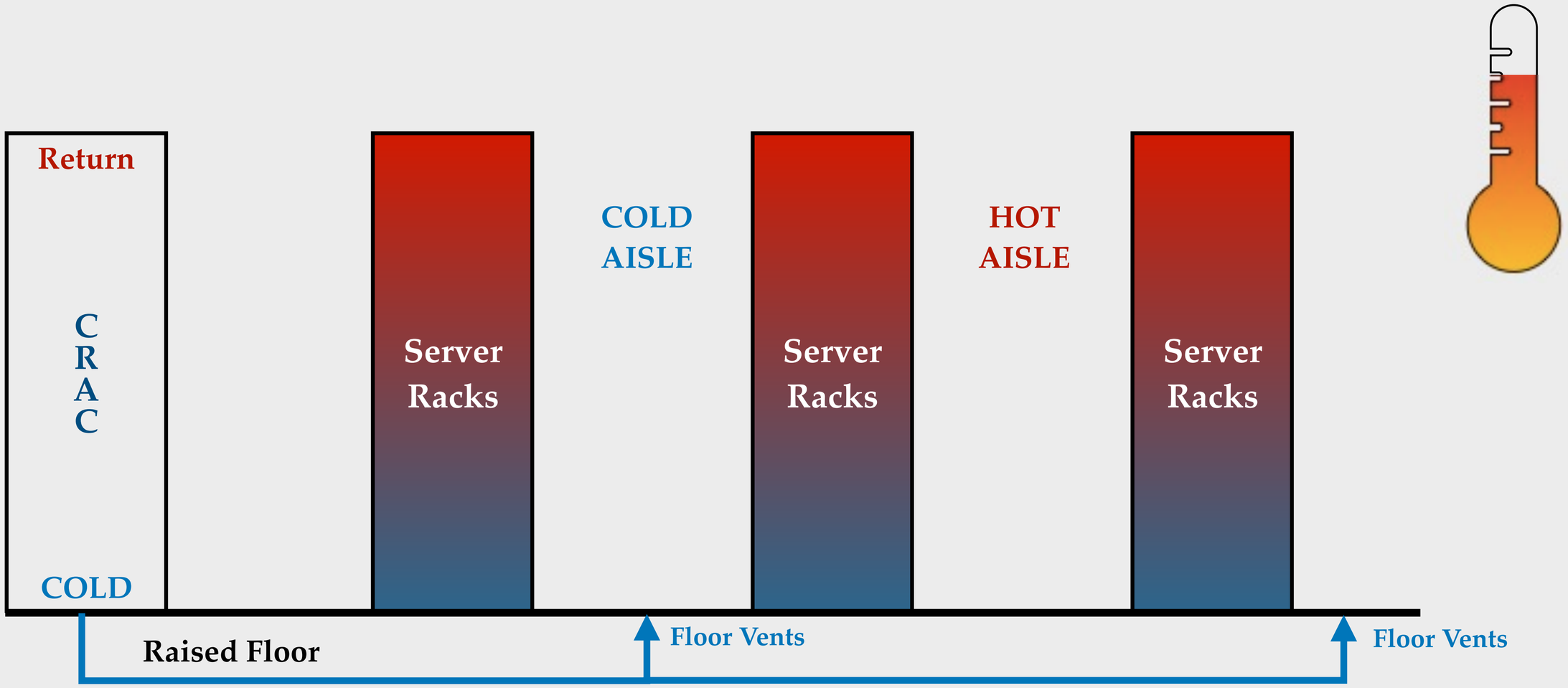
Credit: Data center cooling methods from Submer

Power Oversubscription



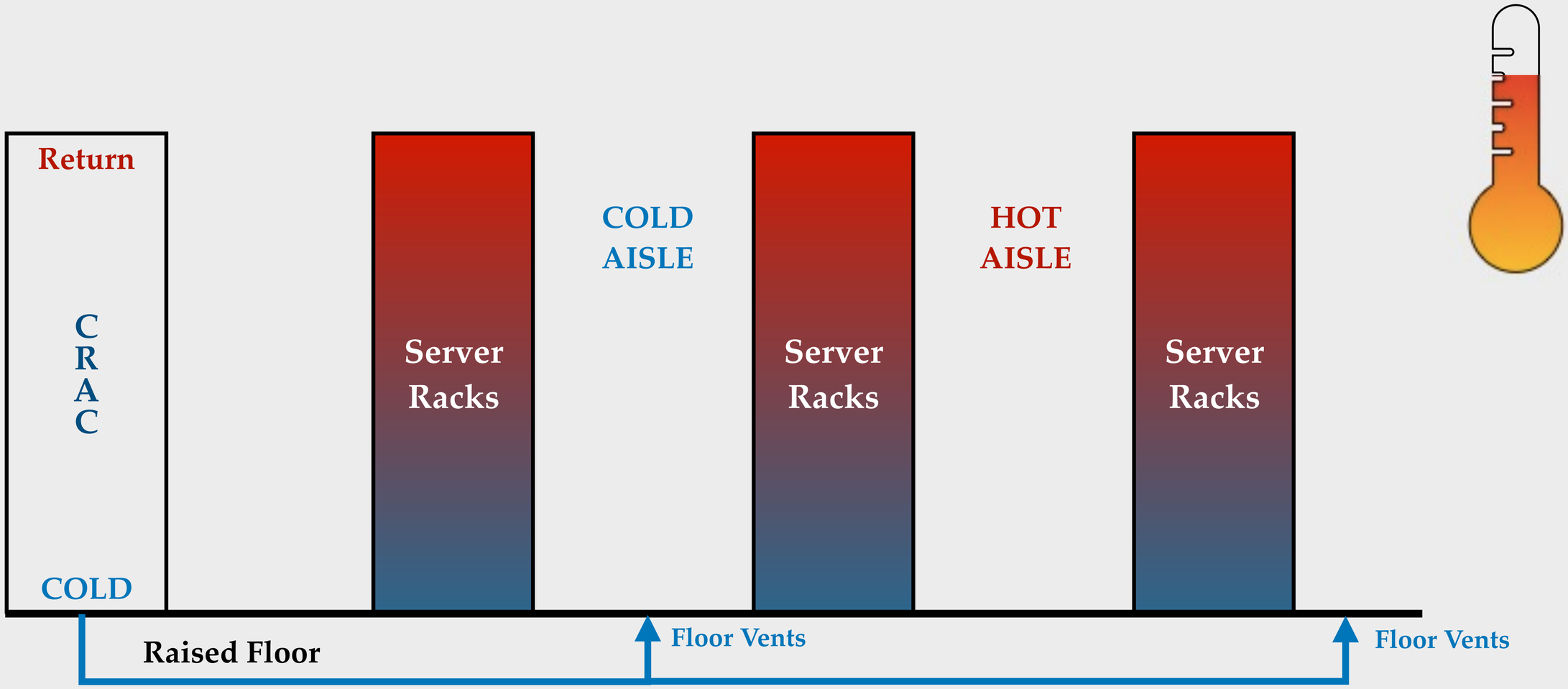
Credit: Data center cooling methods from Submer

Power Oversubscription



Credit: Data center cooling methods from Submer

Reduced Cooling Redundancy

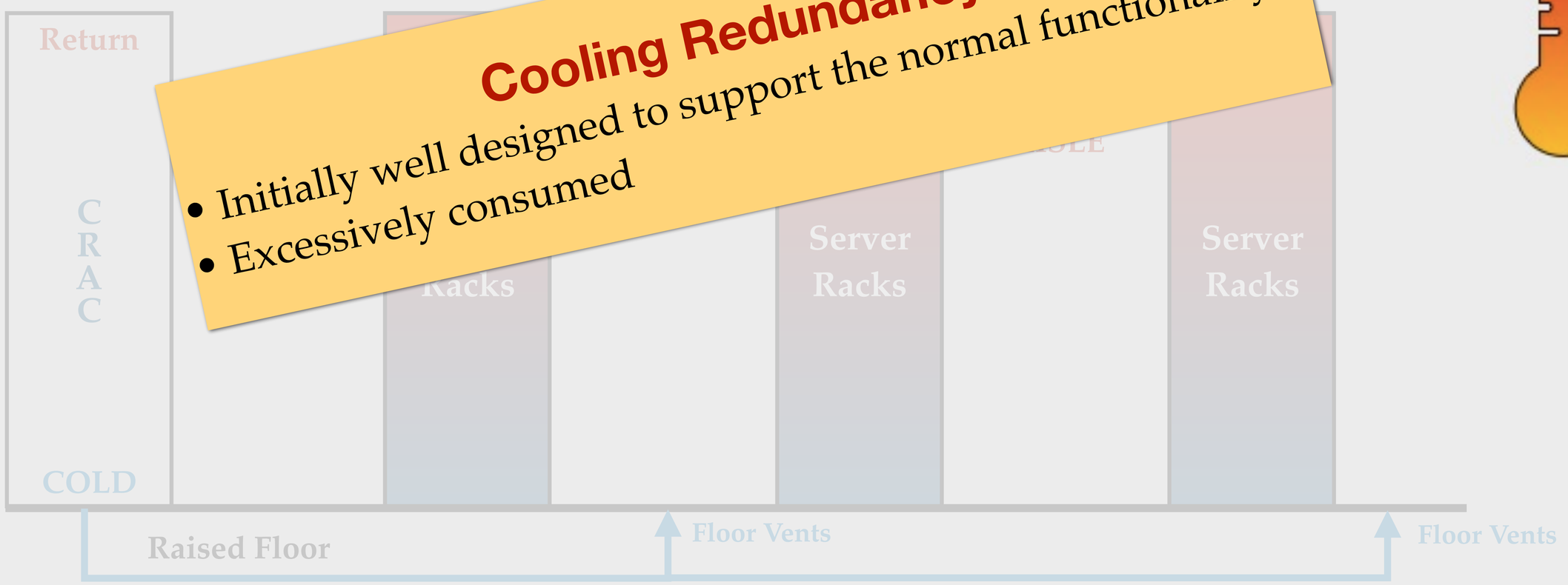


Credit: Data center cooling methods from Submer

Reduced Cooling Redundancy

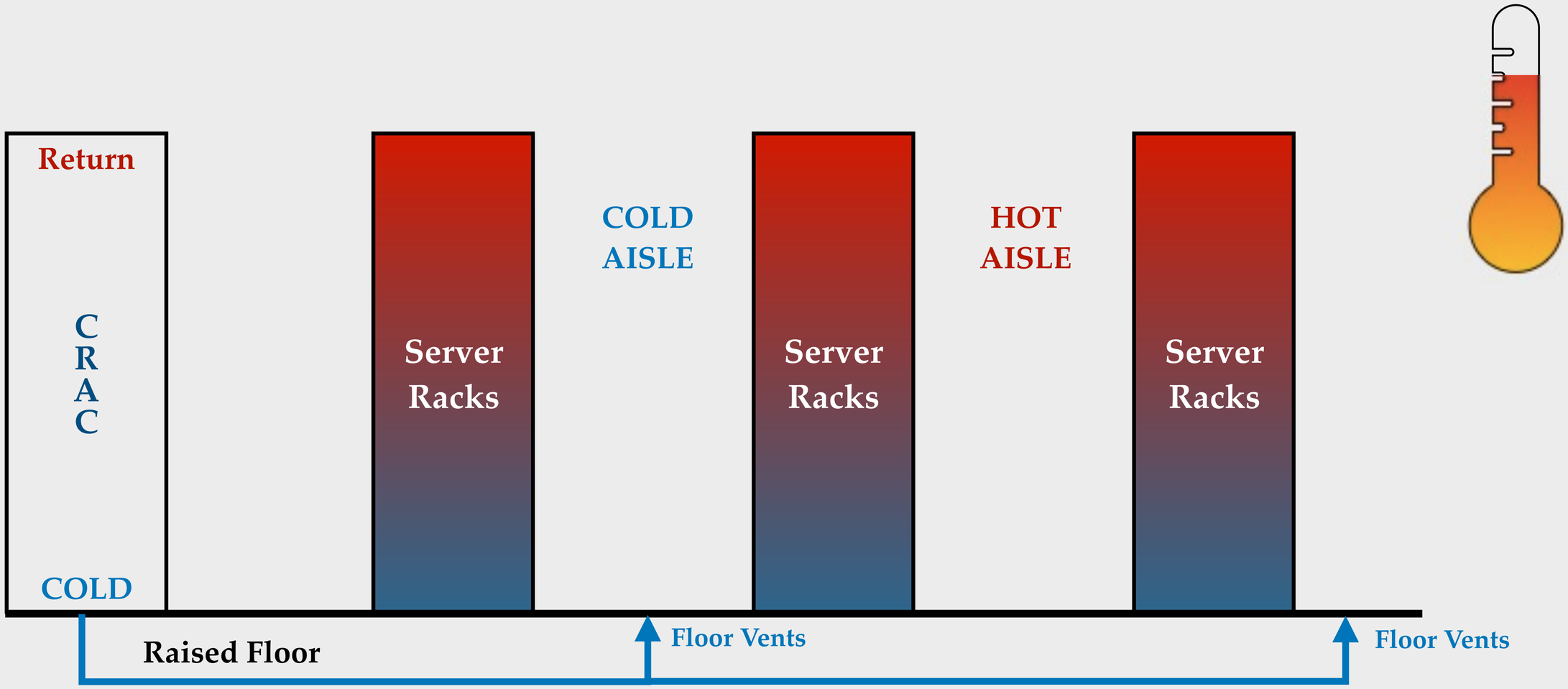
Cooling Redundancy

- Initially well designed to support the normal functionality
- Excessively consumed



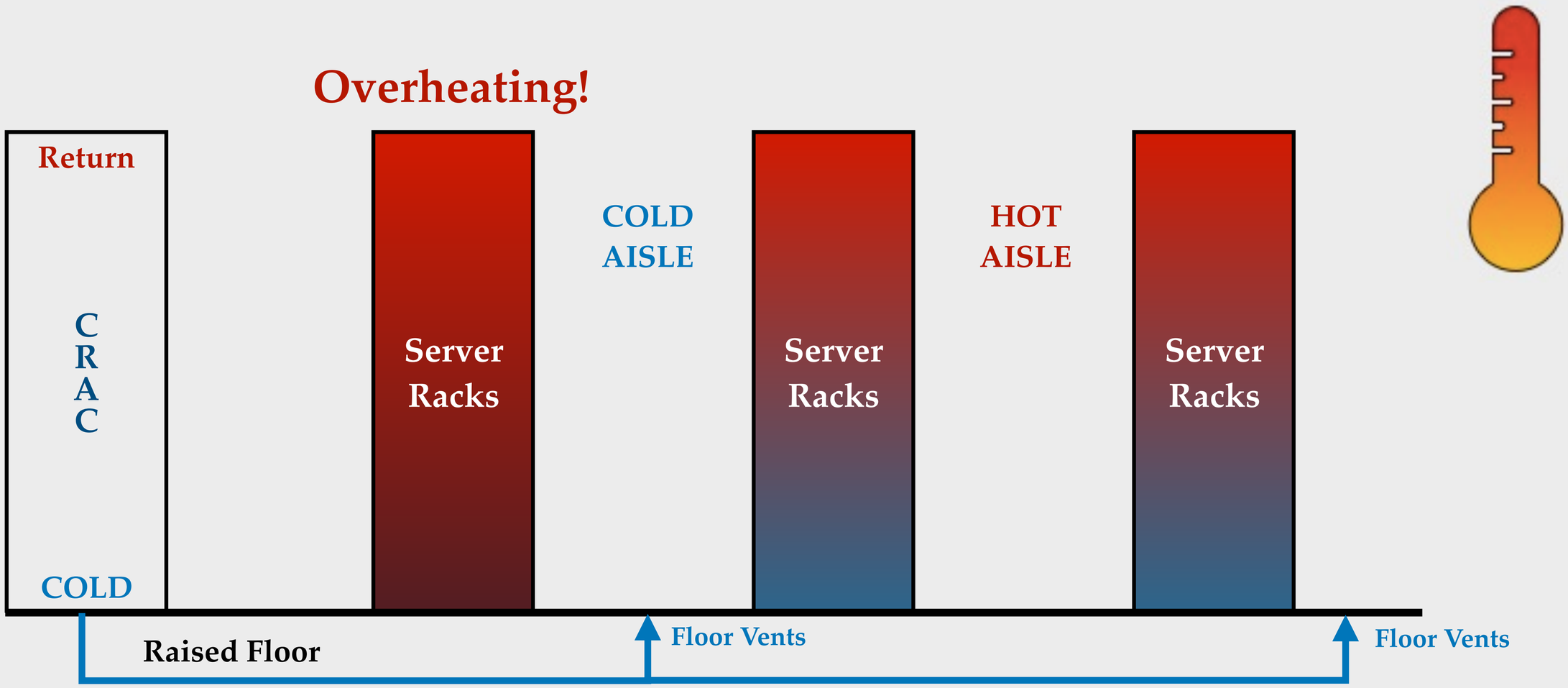
Credit: Data center cooling methods from Submer

Reduced Cooling Redundancy



Credit: Data center cooling methods from Submer

Reduced Cooling Redundancy



Credit: Data center cooling methods from Submer

Overheating – Problems?

Overheating – Problems?

- ***Affect reliability***
  - E.g., hardware failure
  - Electromigration, Time dependent dielectric breakdown, Thermal cycling, Disk failure

$$\lambda_{EM} = A_0(J - J_{crit})^{-n} e^{(-E_a/kT)}$$

Overheating – Problems?

- ***Affect reliability***
  - E.g., hardware failure
  - Electromigration, Time dependent dielectric breakdown, Thermal cycling, Disk failure

$$\lambda_{EM} = A_0 (J - J_{crit})^{-n} e^{(-E_a/kT)}$$

- ***Local hotspots***
  - Air recirculation
  - Gaps between insulation material and the server racks

Overheating – Problems?

- ***Affect reliability***
  - E.g., hardware failure
  - Electromigration, Time dependent dielectric breakdown, Thermal cycling, Disk failure

$$\lambda_{EM} = A_0 (J - J_{crit})^{-n} e^{(-E_a/kT)}$$

- ***Local hotspots***
  - Air recirculation
  - Gaps between insulation material and the server racks

- ***Raise cooling cost***

Overheating – Problems?

- ***Affect reliability***
  - E.g., hardware failure
  - Electromigration, Time dependent dielectric breakdown, Thermal cycling, Disk failure

$$\lambda_{EM} = A_0 (J - J_{crit})^{-n} e^{(-E_a/kT)}$$

- ***Local hotspots***
  - Air recirculation
  - Gaps between insulation material and the server racks

- ***Raise cooling cost***

- ***Cooling failure***
  - Server shutdown!

Overheating – Problems?

- ***Affect reliability***
  - E.g., hardware failure
  - Electromigration, Time dependent dielectric breakdown, Thermal cycling, Disk failure

$$\lambda_{EM} = A_0 (J - J_{crit})^{-n} e^{(-E_a/kT)}$$

- ***Local hotspots***
  - Air recirculation
  - Gaps between insulation material and the server racks

- ***Raise cooling cost***

- ***Cooling failure***
  - Server shutdown!

Overheating – Problems?

- ***Affect reliability***
  - E.g., hardware failure
  - Electromigration, Time dependent dielectric breakdown, Thermal cycling, Disk failure

$$\lambda_{EM} = A_0 (J - J_{crit})^{-n} e^{(-E_a/kT)}$$

- ***Local hotspots***
  - Air recirculation
  - Gaps between insulation material and the server racks

- ***Raise cooling cost***

- ***Cooling failure***
  - Server shutdown!

Overheating – Problems?

- ***Affect reliability***
  - E.g., hardware failure
  - Electromigration, Time dependent dielectric breakdo
    Thermal cycling, Disk failure
- ***Local hotspots***
  - Air recirculation
  - Gaps between insulation
- ***Raise cooling cost***
- ***Cooling failure***
  - Server shutdown!

Potential Threats:

- Intentionally cause overheating?

$$J_{crit})^{-n} e^{(-E_a/kT)}$$

Overheating brings down Microsoft data center

14 March 2013 | By Penny Jones

MANAGE > UPTIME

Wikipedia's Data Center Overheats

Wikipedia was offline for several hours yesterday after a cooling problem in Wikimedia's European data center led to a heat condition that caused a server shutdown. A DNS problem during failover caused additional downtime.

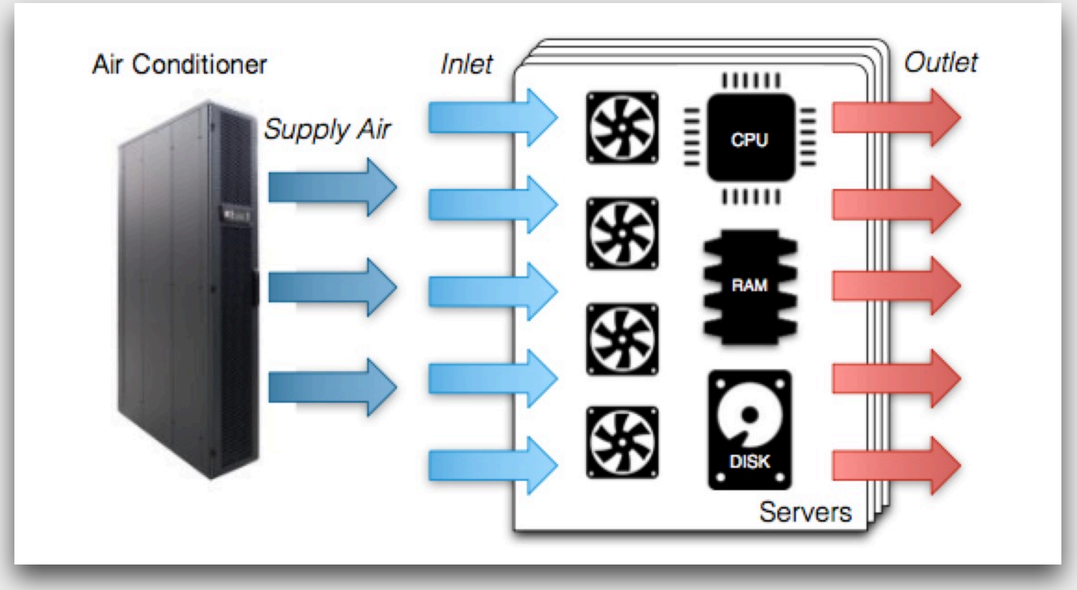
Rich Miller | Mar 25, 2010

DESIGN > POWER AND COOLING

Heatwave, Cooling Failure Bring iiNet Data Center Down in Perth

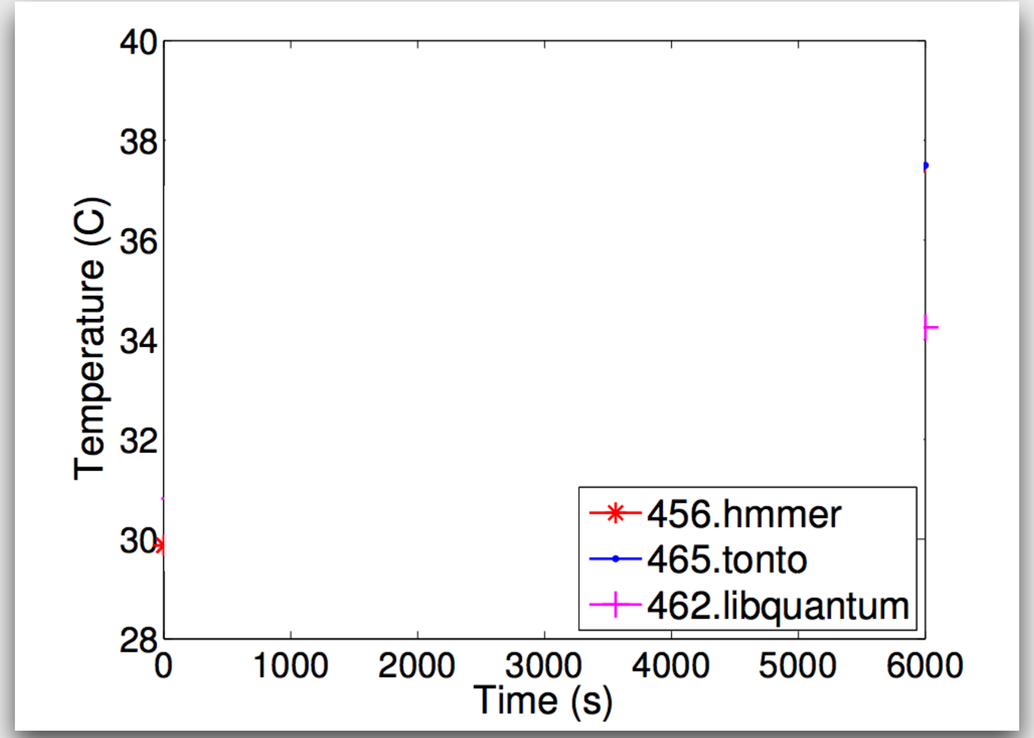
Testbed Measurement Setup

- Sealed Environment
- Supermicro server
- “Go!Temp” temperature probe



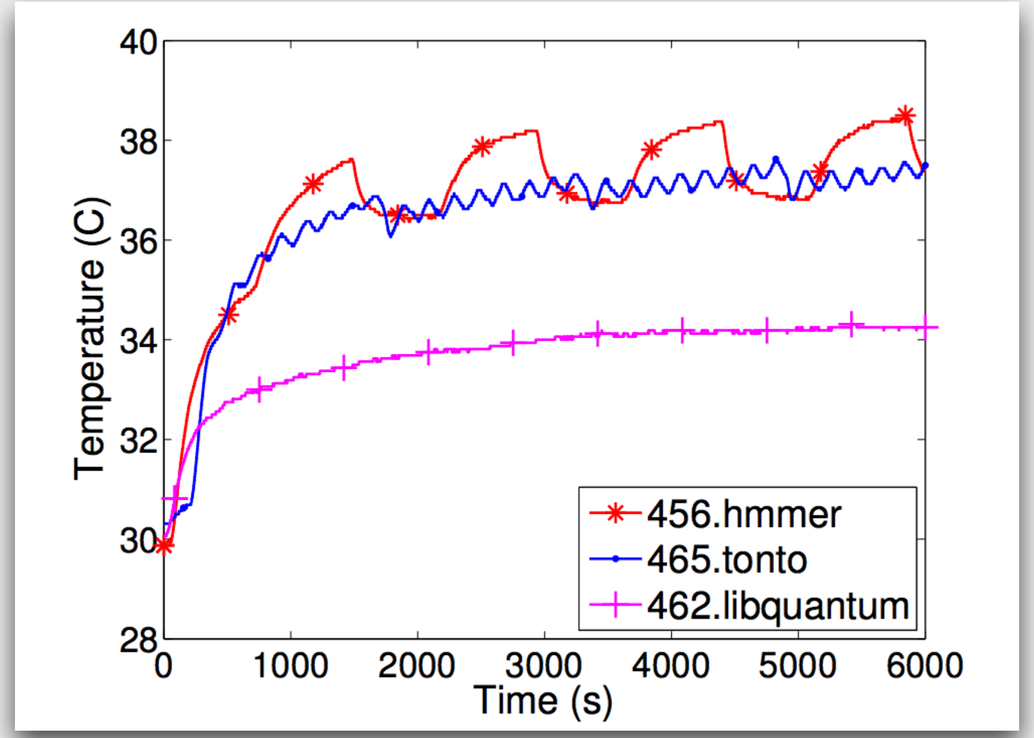
Thermal Characteristics

Similar system resource consumption



Thermal Characteristics

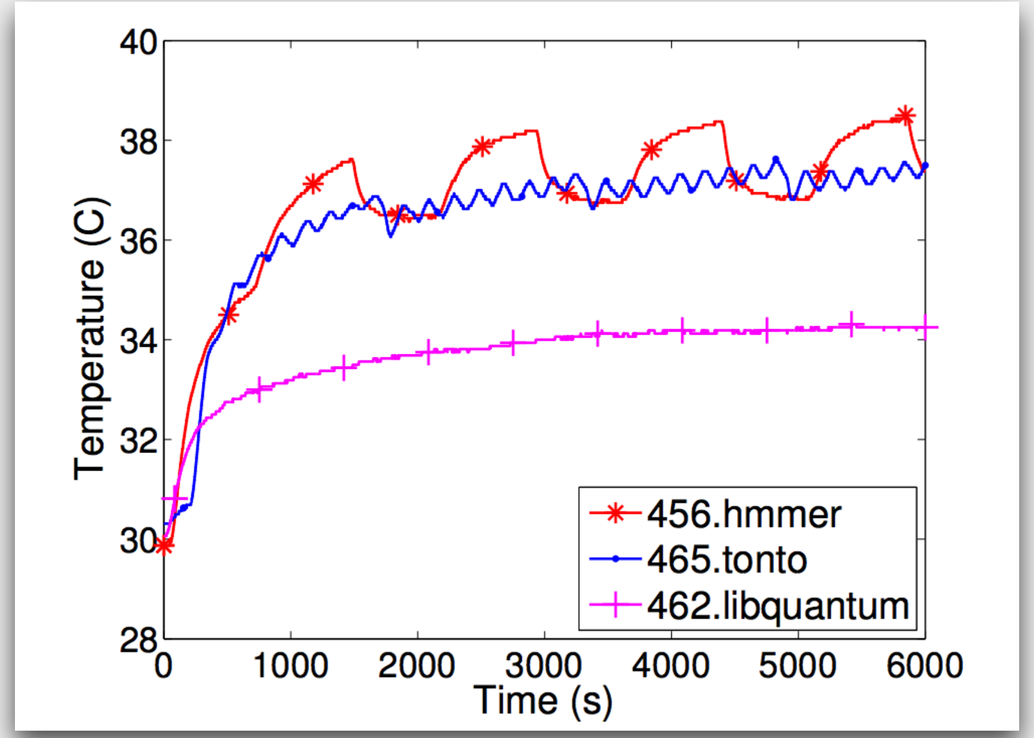
Similar system resource consumption



Thermal Characteristics

Similar system resource consumption

- **leads to different outlet temperature.**

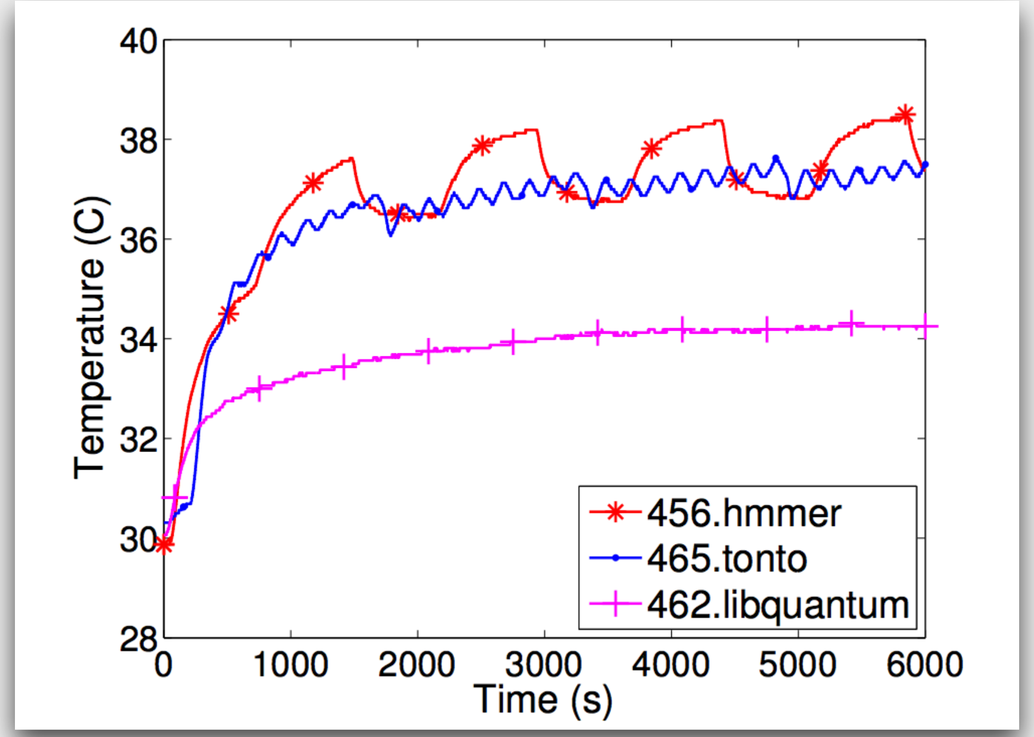


Thermal Characteristics

Similar system resource consumption

- **leads to different outlet temperature.**

Reason: underlying pipeline flows are different
Further cause CPU halt and leave function units idle



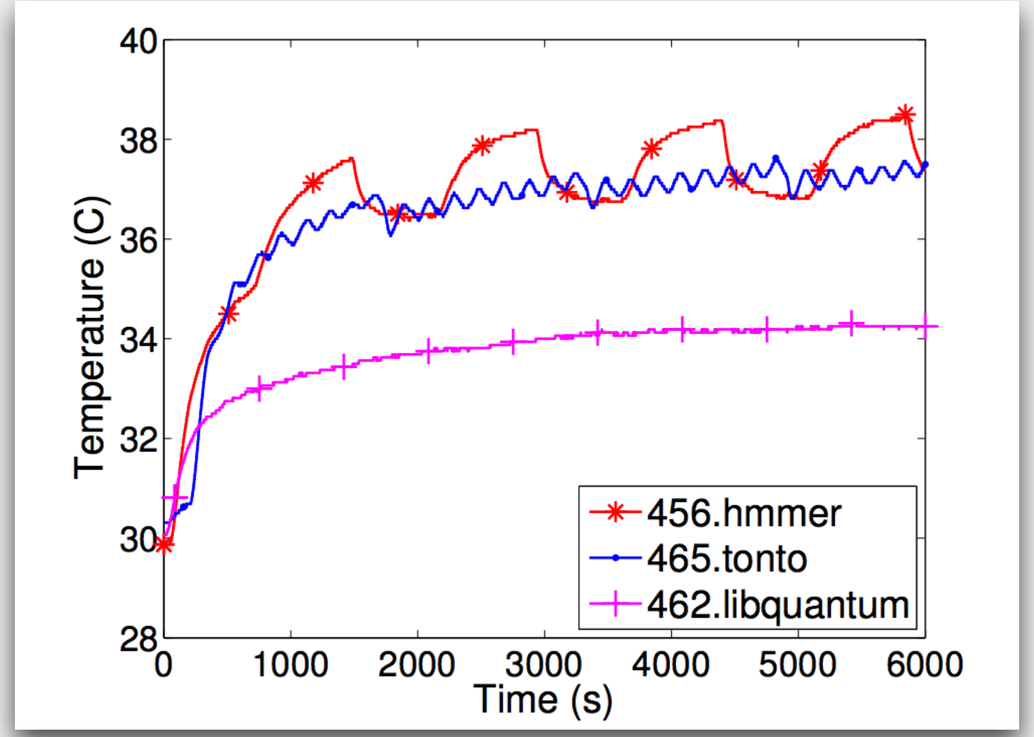
Thermal Characteristics

Similar system resource consumption

- **leads to different outlet temperature.**

Reason: underlying pipeline flows are different
Further cause CPU halt and leave function units idle

Thermal-intensive workloads can generate much more heat without consuming more system resources.



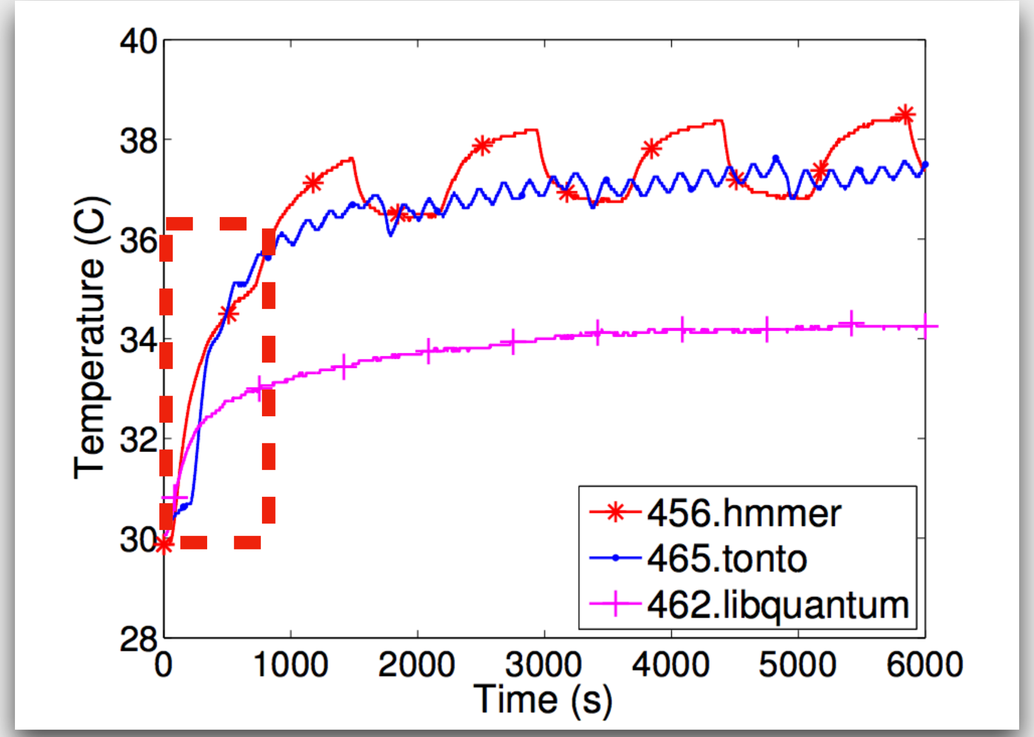
Thermal Characteristics

Similar system resource consumption

- **leads to different outlet temperature.**

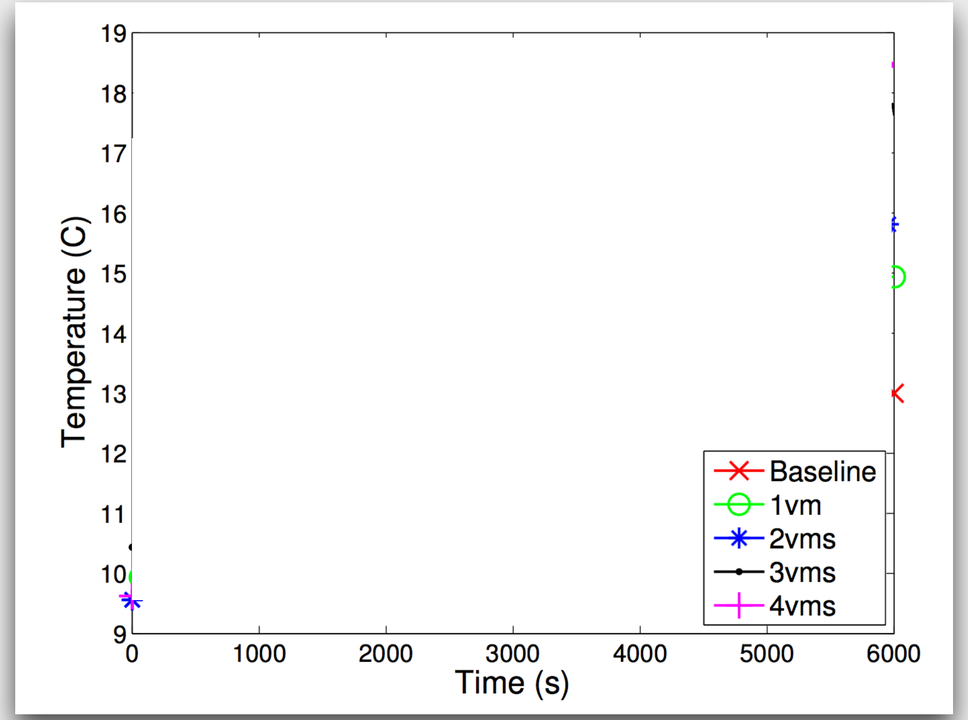
Reason: underlying pipeline flows are different
Further cause CPU halt and leave function units idle

Thermal-intensive workloads can generate much more heat without consuming more system resources.



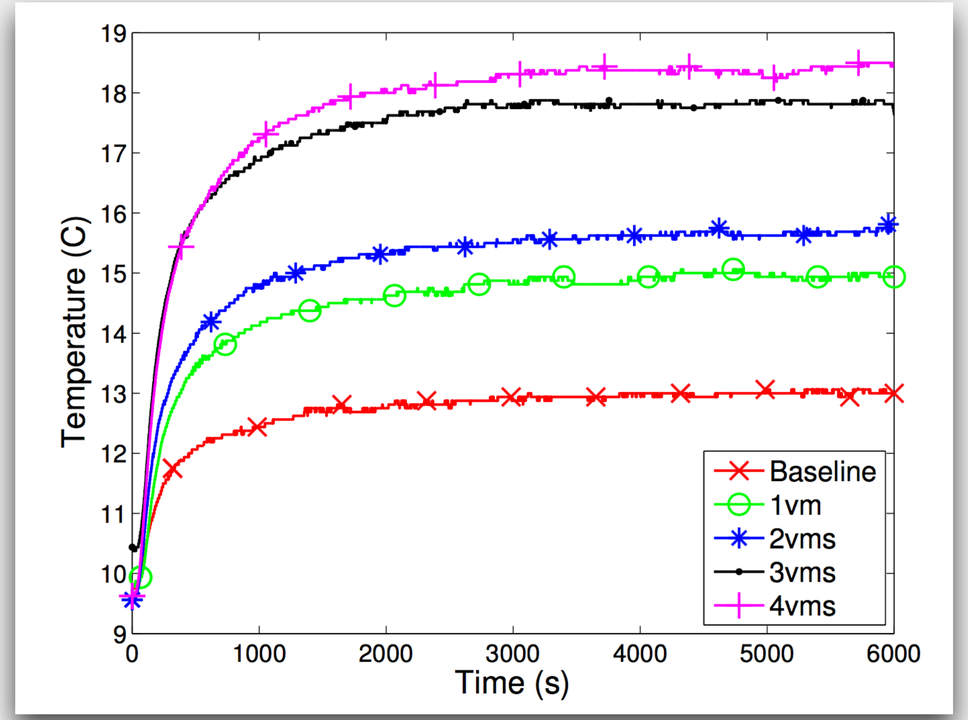
Thermal Attacks on Virtual Machine

- Xen Hypervisor / VM with 4vCPU
- Baseline: 25% utilization



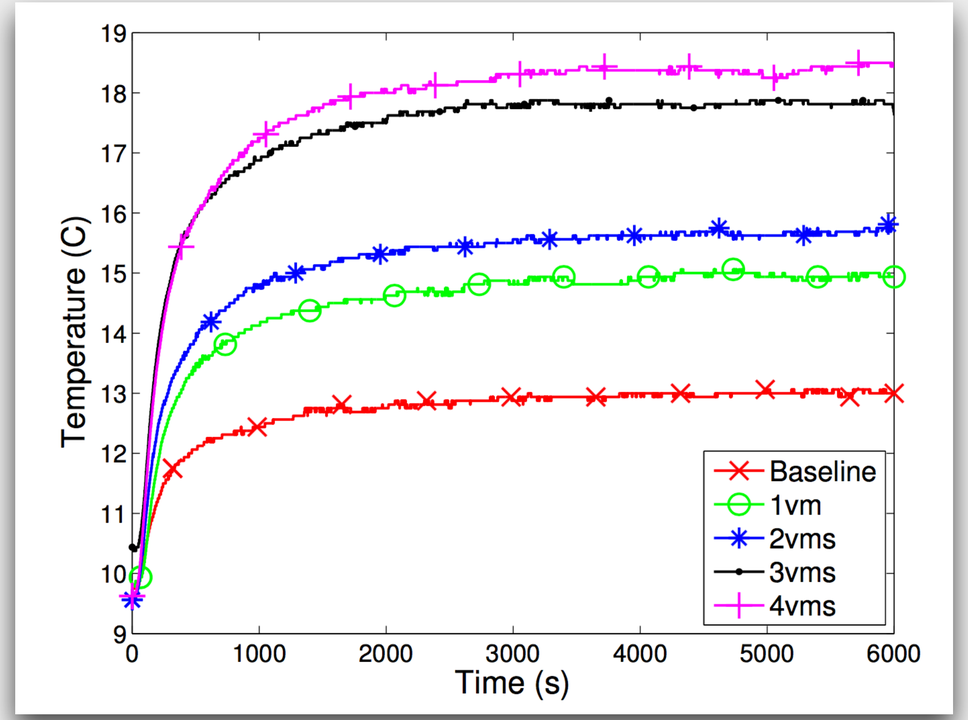
Thermal Attacks on Virtual Machine

- Xen Hypervisor / VM with 4vCPU
- Baseline: 25% utilization



Thermal Attacks on Virtual Machine

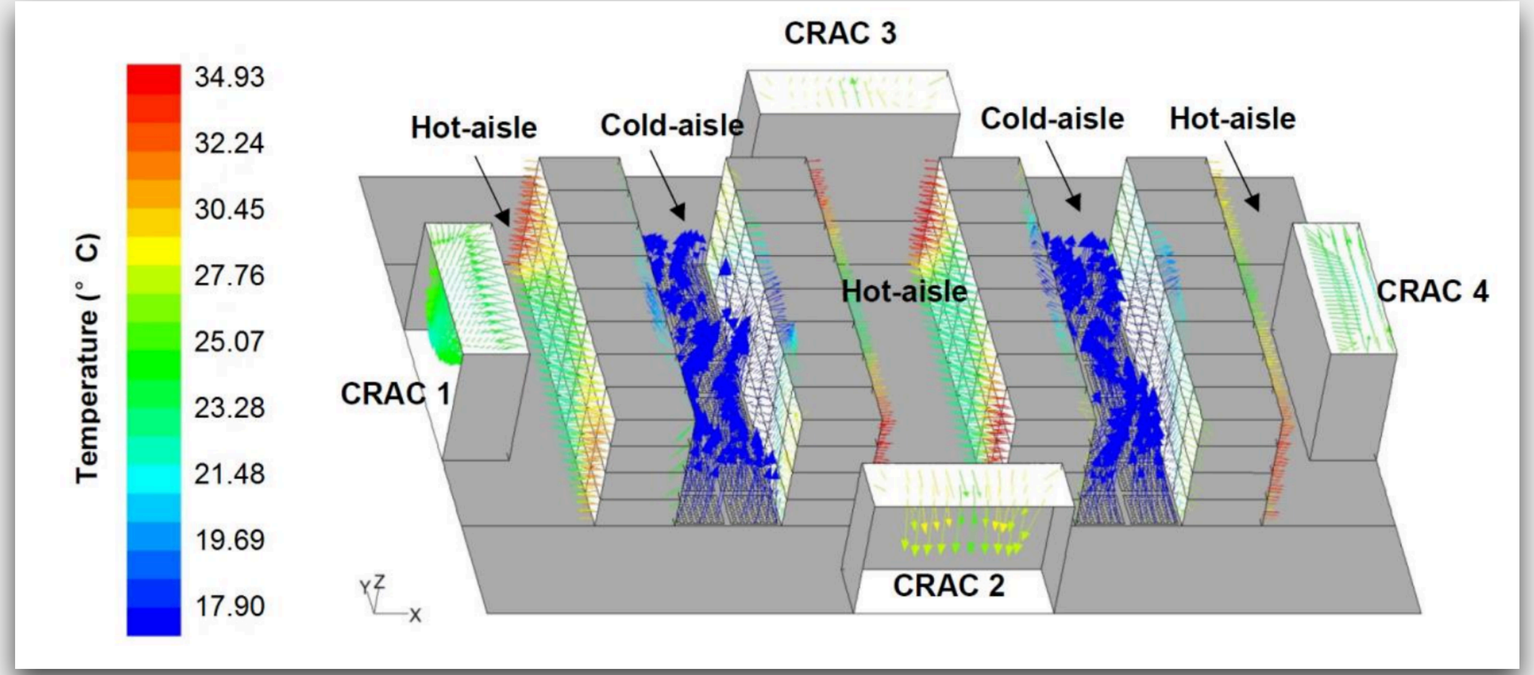
- Xen Hypervisor / VM with 4vCPU
- Baseline: 25% utilization
- Co-resident VMs can produce more heat



Data Center Level Simulation

Data center level: trace-driven computational fluid dynamics analysis.

- CFD package, Fluent, to simulate the thermal environment.
  - Air recirculation.
  - Air density.
  - Air flow rate.

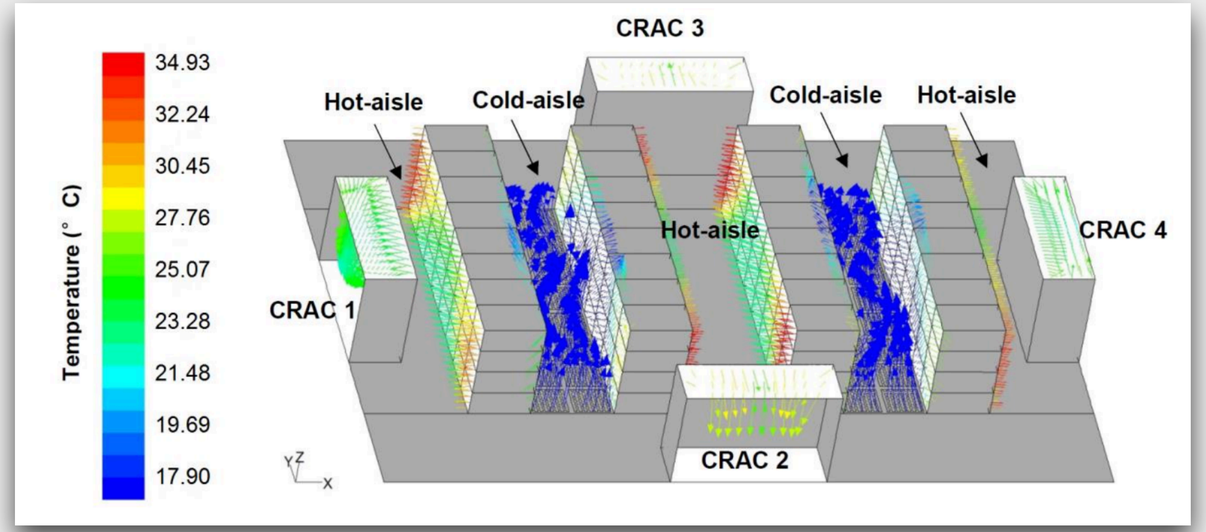


Data Center Level Simulation

Data center level: trace-driven computational fluid dynamics analysis.

- The layout:
  - 4 rows of servers;
  - 8 racks in each row;
  - 40 servers per rack —> 1280 servers.

Power-oversubscribed



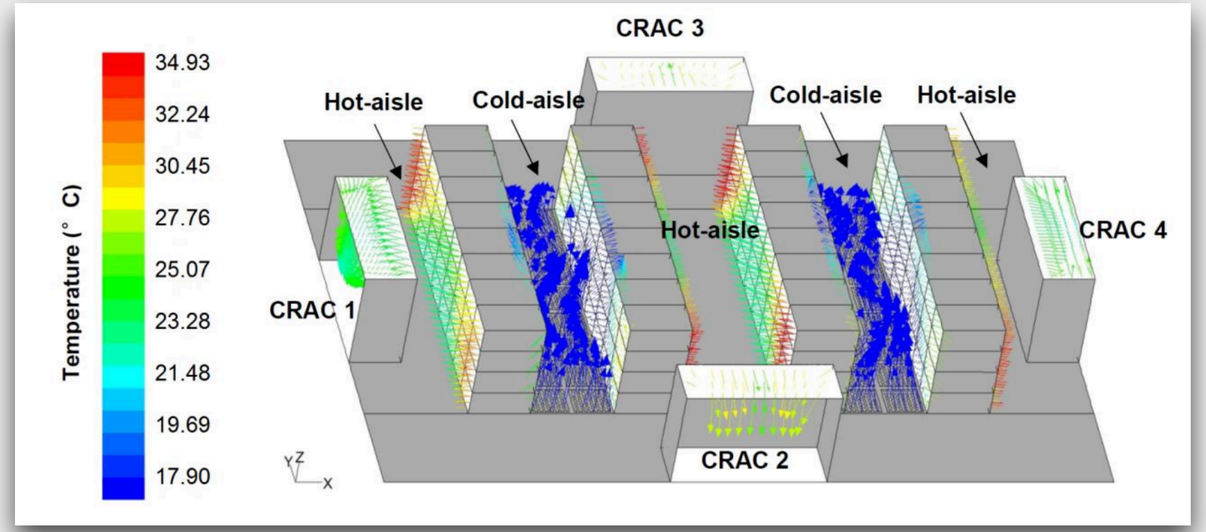
Data Center Level Simulation

Data center level: trace-driven computational fluid dynamics analysis.

- T_{outlet} is affected by outlet temperature, power consumption of all servers, and T_{sup}.
- T_{inlet} depends on T_{sup} and T_{outlet}
- One-week trace from real data centers.

$$K_i T_{out}^i = \sum_{j=1}^{N} h_{ji} K_j T_{out}^j + (K_i - \sum_{j=1}^{N} h_{ji} K_j) T_{sup} + P_i,$$

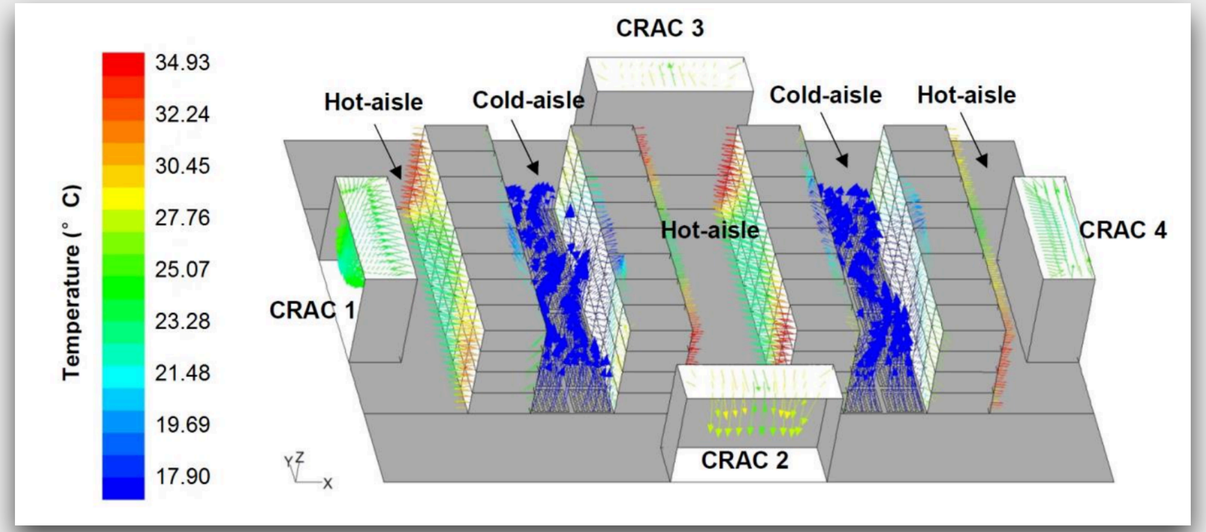
$$T_{in}^i = \sum_{j=1}^{N} h_{ji} * (T_{out}^j - T_{sup}) + T_{sup},$$



Data Center Level Simulation

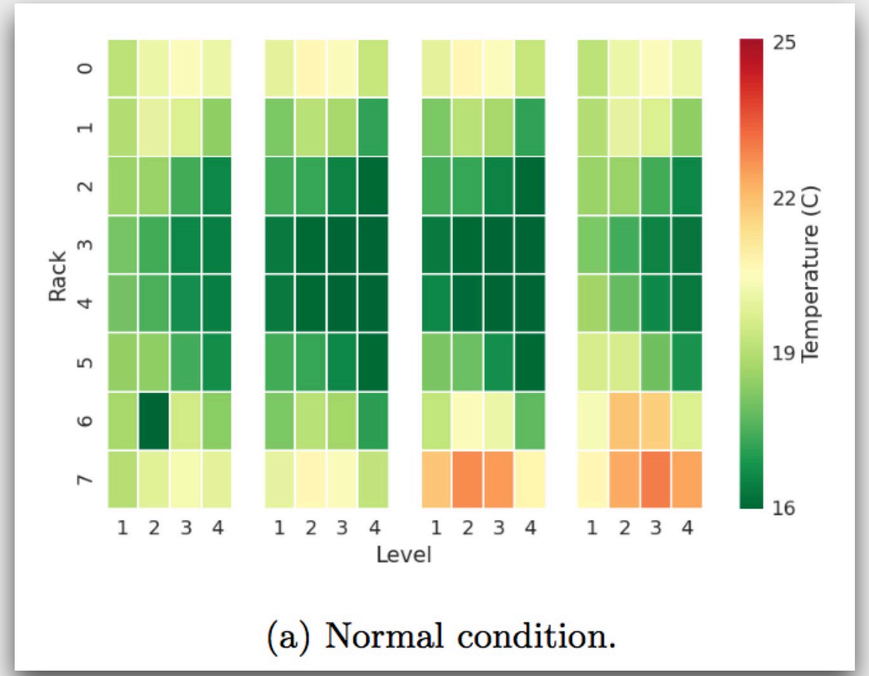
Data center level: trace-driven computational fluid dynamics analysis.

- $T_{sup} = 16°C$
- $T_{redline} = 25°C$ [1]
- Goal: T_{inlet} of all servers < $T_{redline}$



[1] Data center cooling best practices

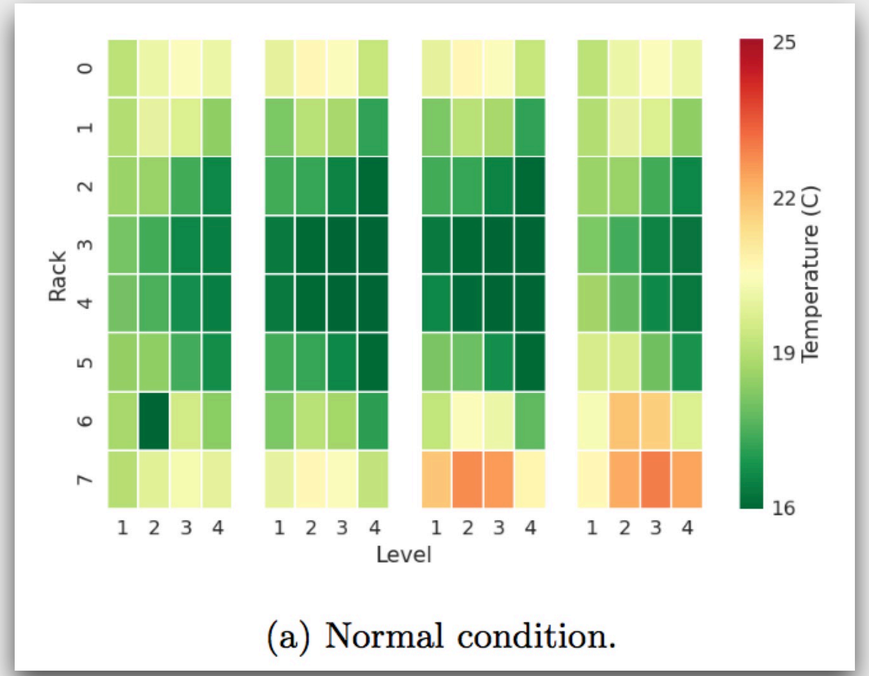
Global View



(a) Normal condition.

Global View

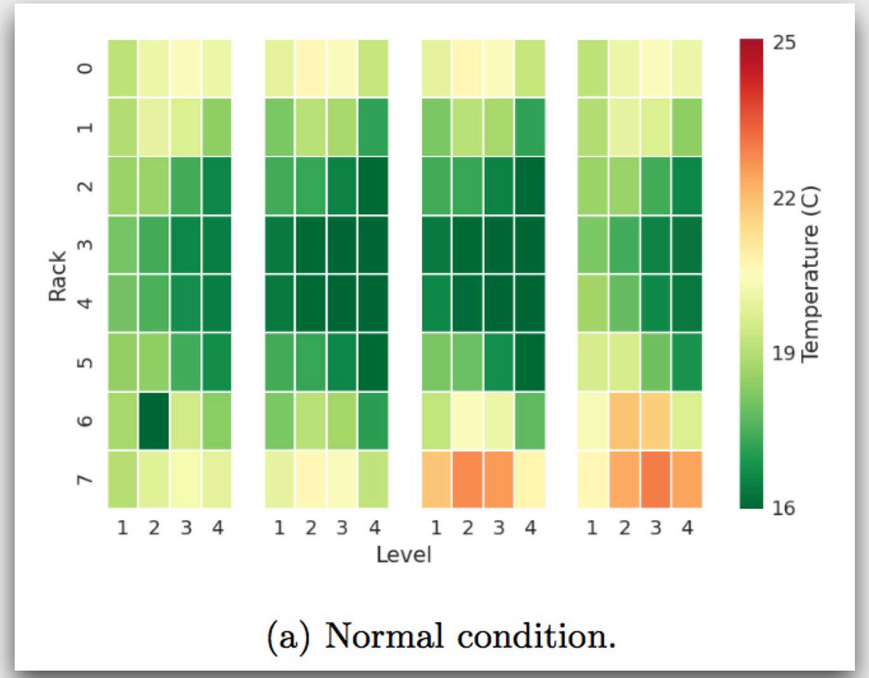
- Switch to thermal-intensive workloads



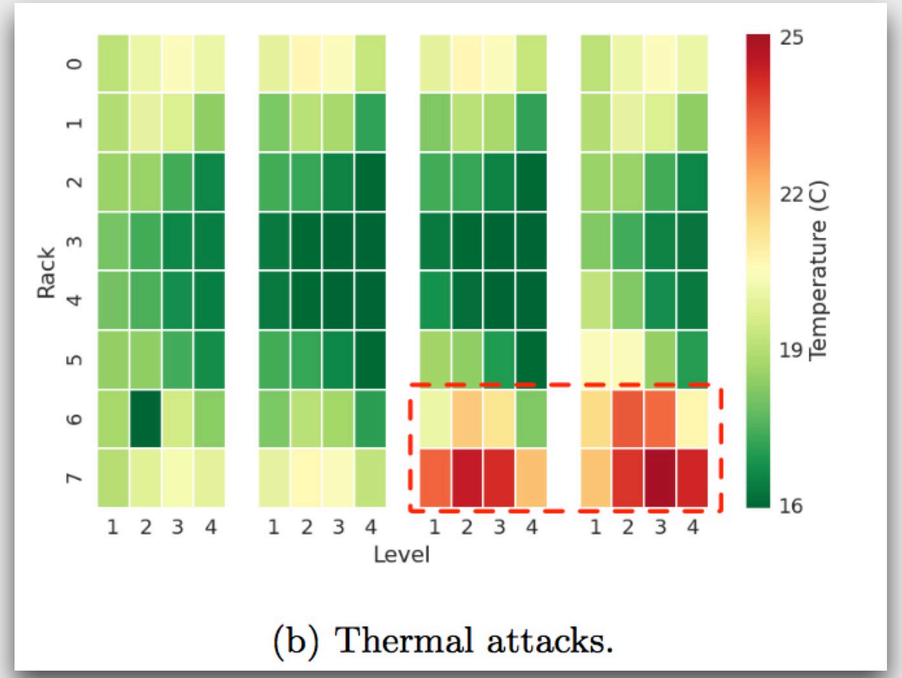
(a) Normal condition.

Global View

- Switch to thermal-intensive workloads



(a) Normal condition.

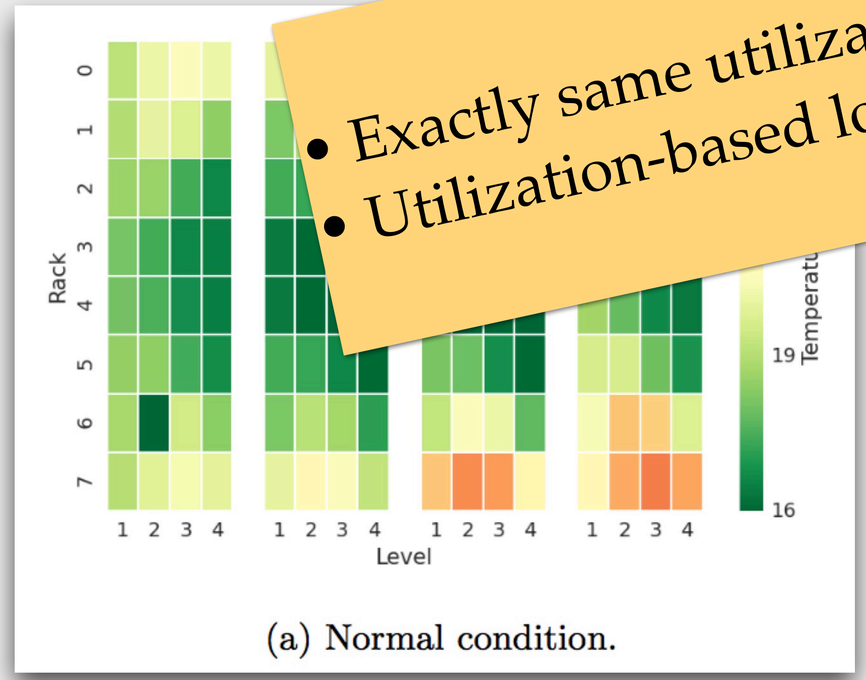


(b) Thermal attacks.

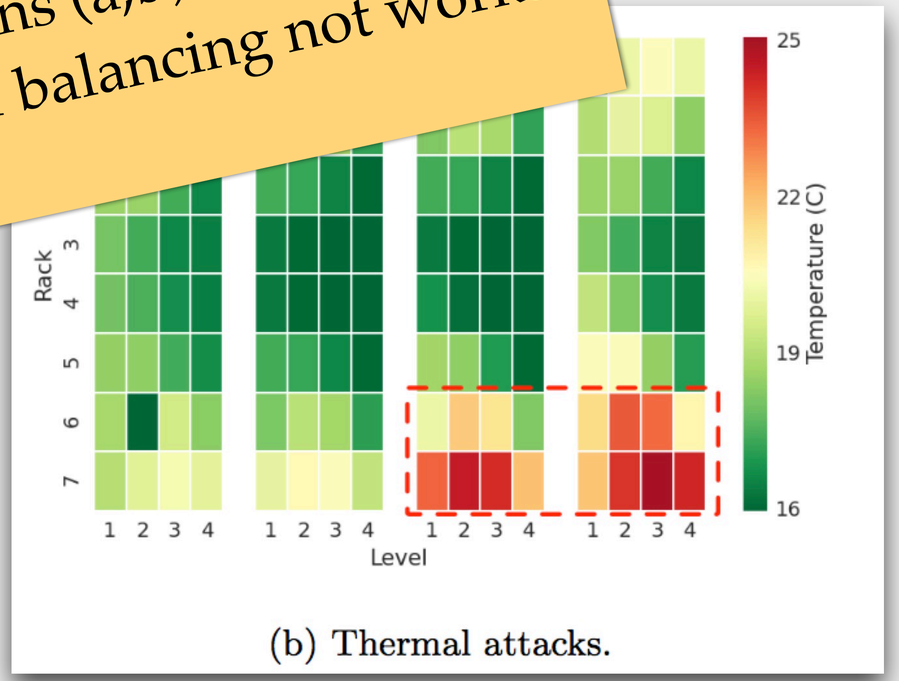
Global View

- Switch to thermal-intensive workl

- Exactly same utilizations (a,b)
- Utilization-based load balancing not works!



(a) Normal condition.



(b) Thermal attacks.

Attacking Efforts

Attacking Efforts

- ***Three types of attackers***
  - *Random:* randomly
  - *Rack-level:* achieve rack-level co-residence
  - *Hotspot*: attackers can roughly infer the servers located in hotspots

Attacking Efforts

- ***Three types of attackers***
  - *Random:* randomly
  - *Rack-level:* achieve rack-level co-residence
  - *Hotspot*: attackers can roughly infer the servers located in hotspots

- ***Three background workloads scenarios***
  - High: 60%
  - Medium: 40%
  - Low: 20%

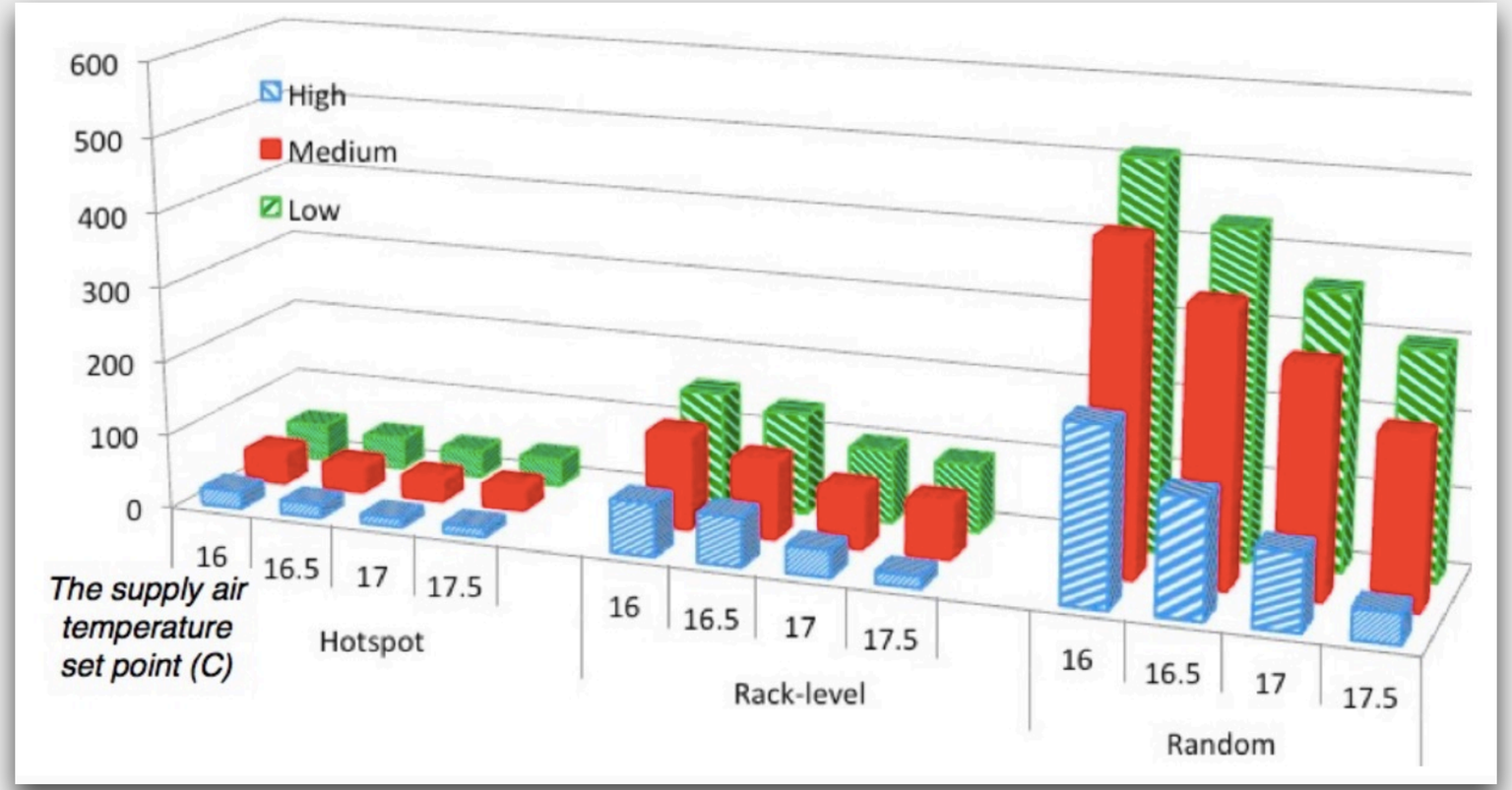
Attacking Efforts

- ***Three types of attackers***
  - *Random:* randomly
  - *Rack-level:* achieve rack-level co-residence
  - *Hotspot*: attackers can roughly infer the servers located in hotspots

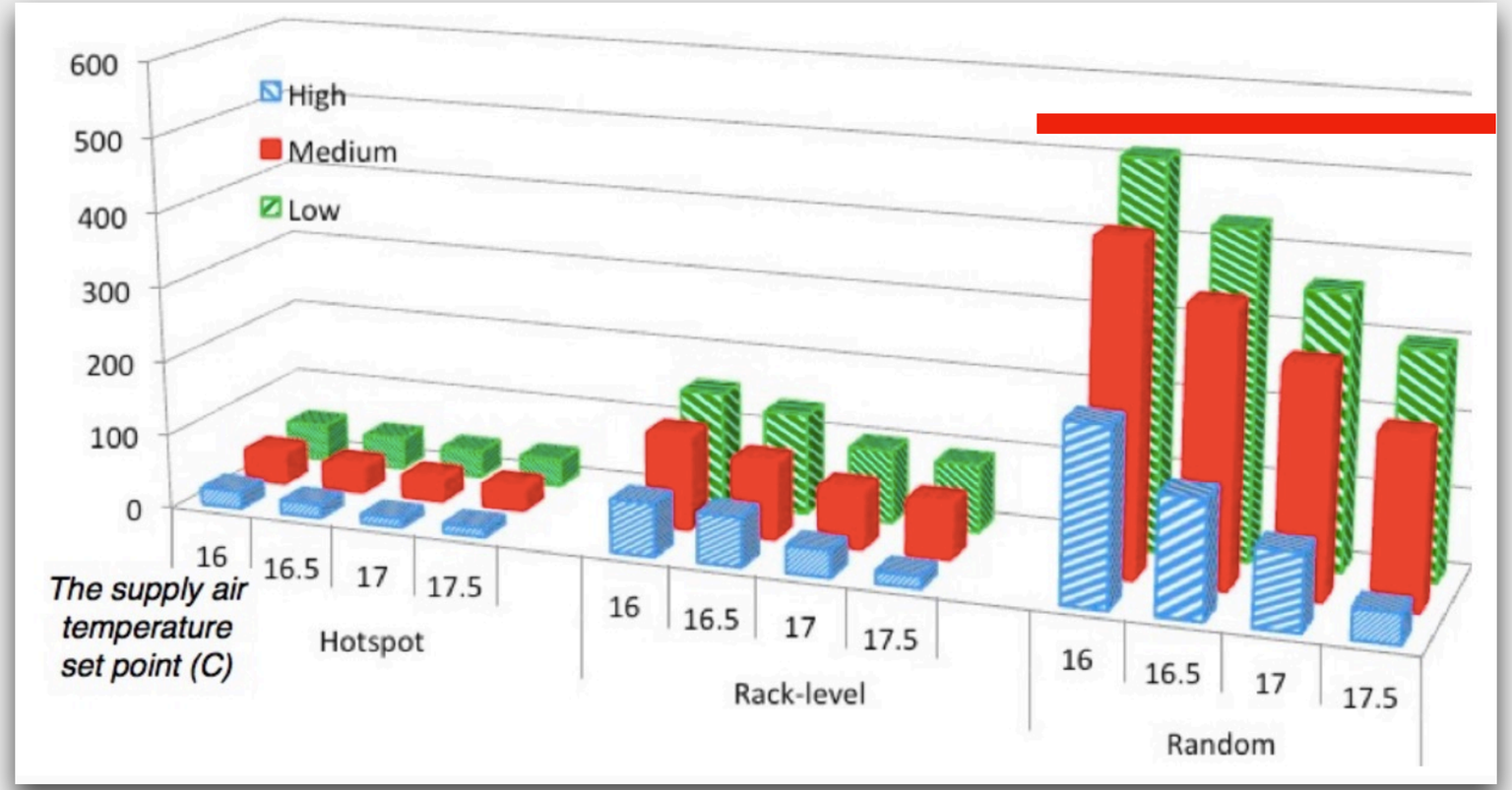
- ***Three background workloads scenarios***
  - High: 60%
  - Medium: 40%
  - Low: 20%

- ***Four different Tsup***
  - 16°C, 16.5°C, 17°C, 17.5°C

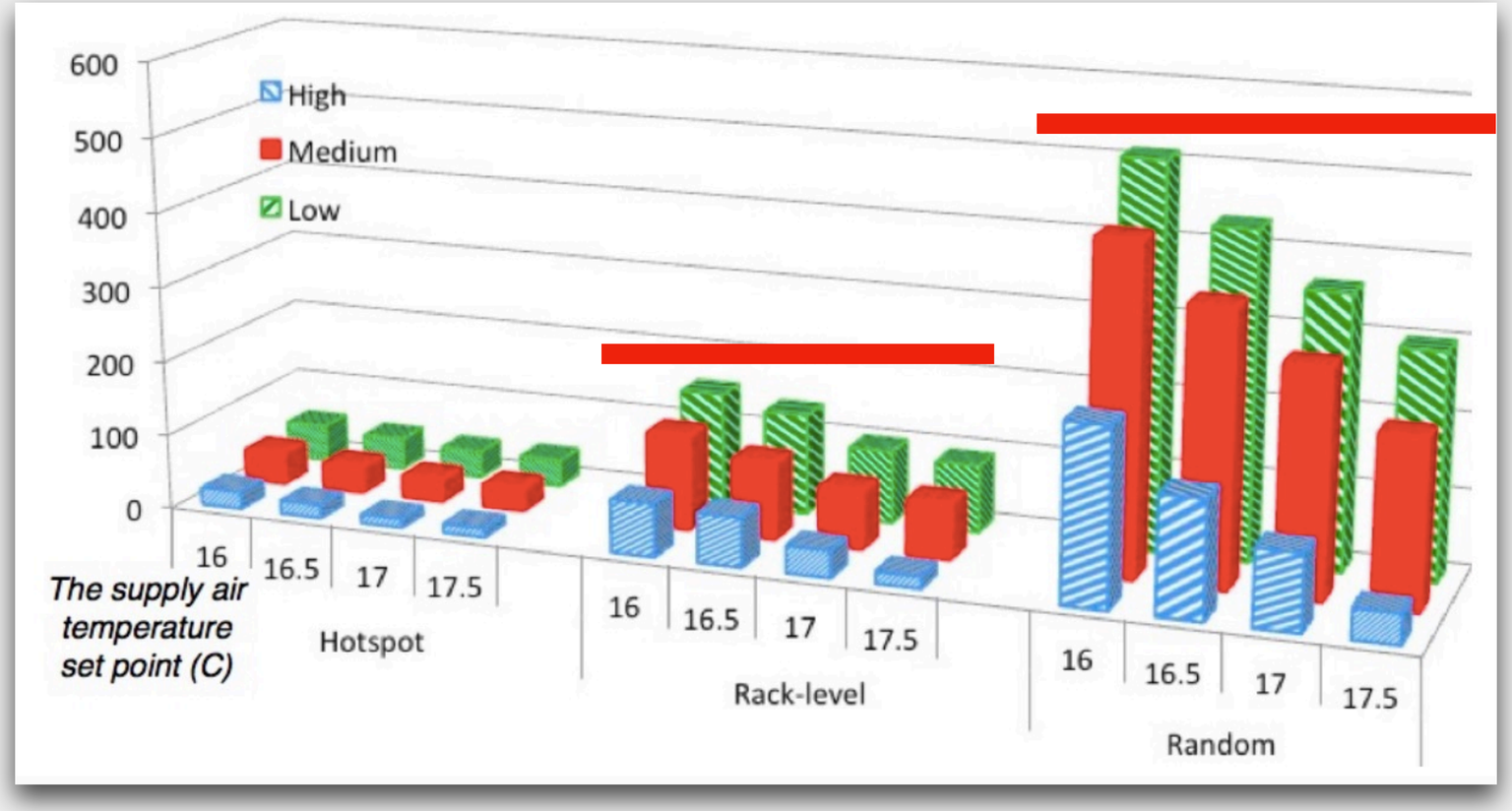
Attacking Efforts



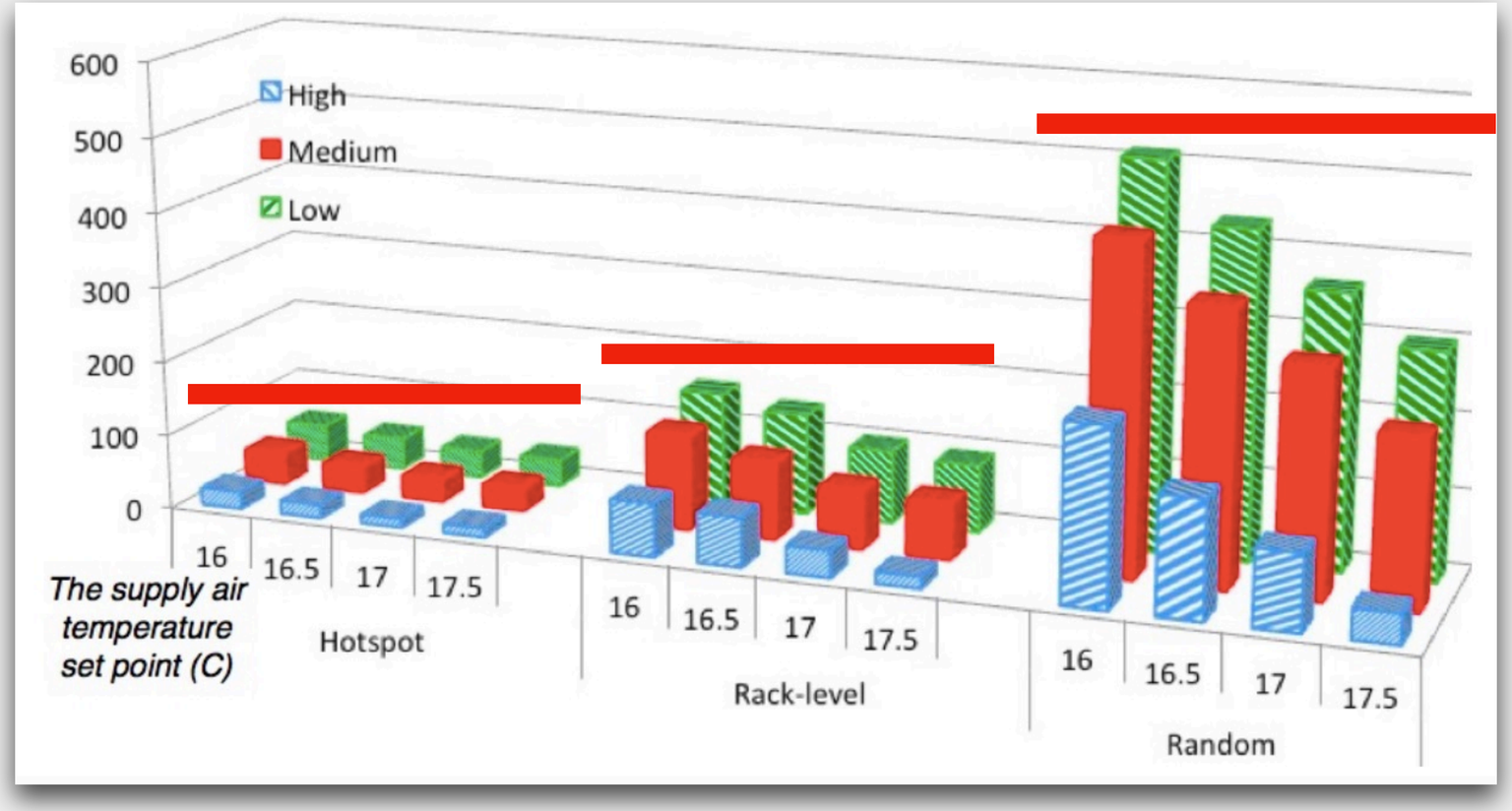
Attacking Efforts



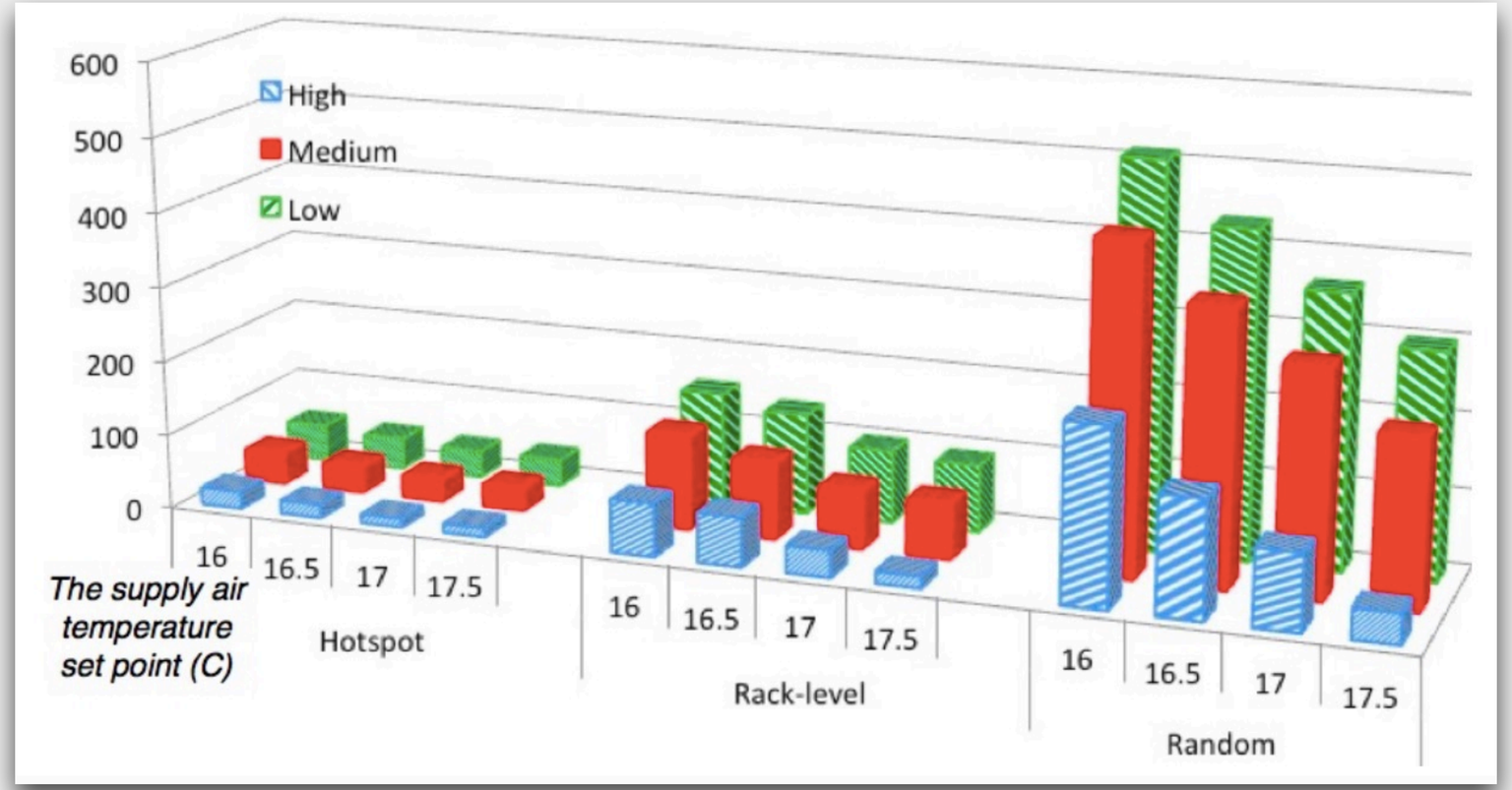
Attacking Efforts



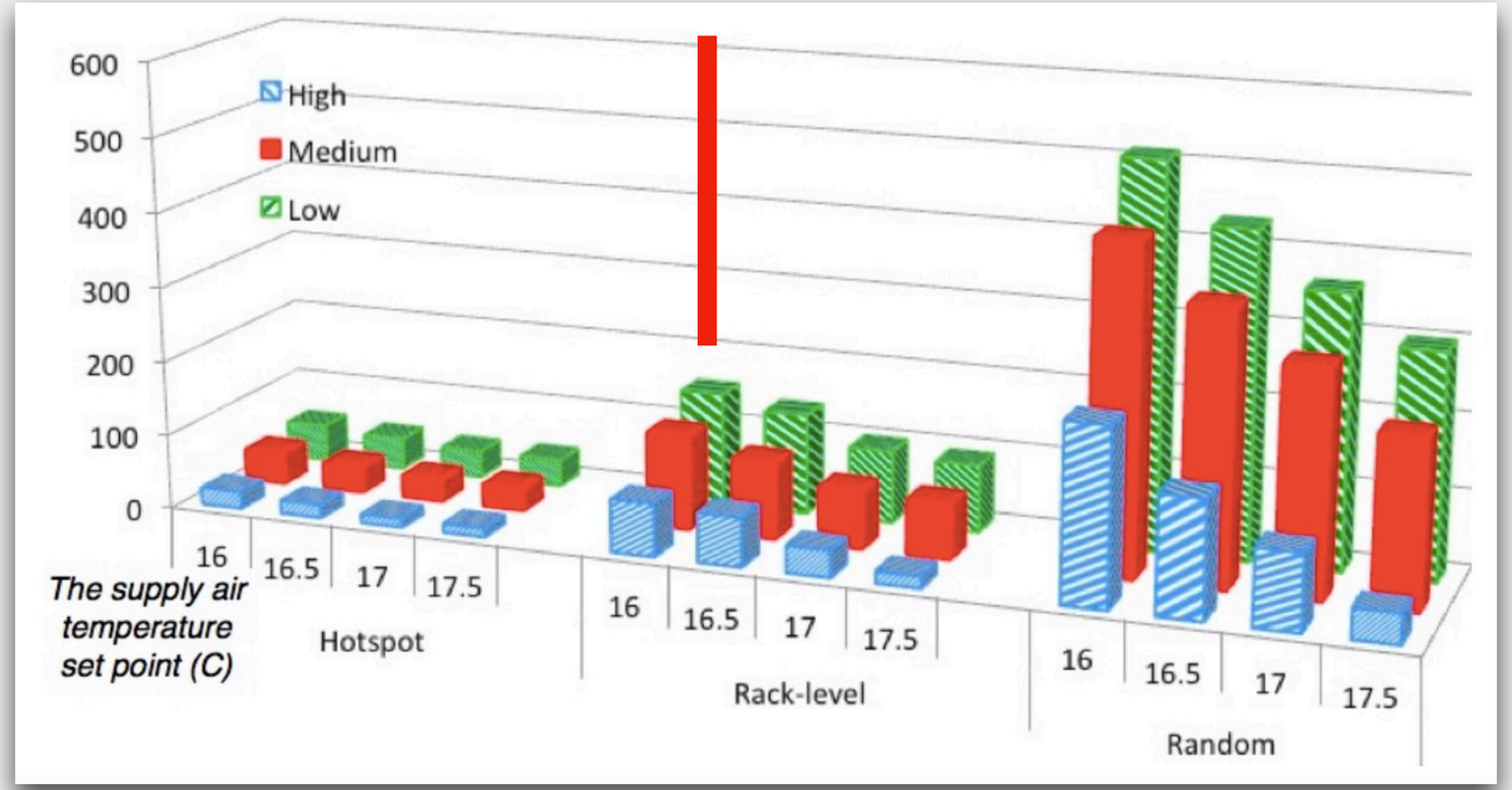
Attacking Efforts



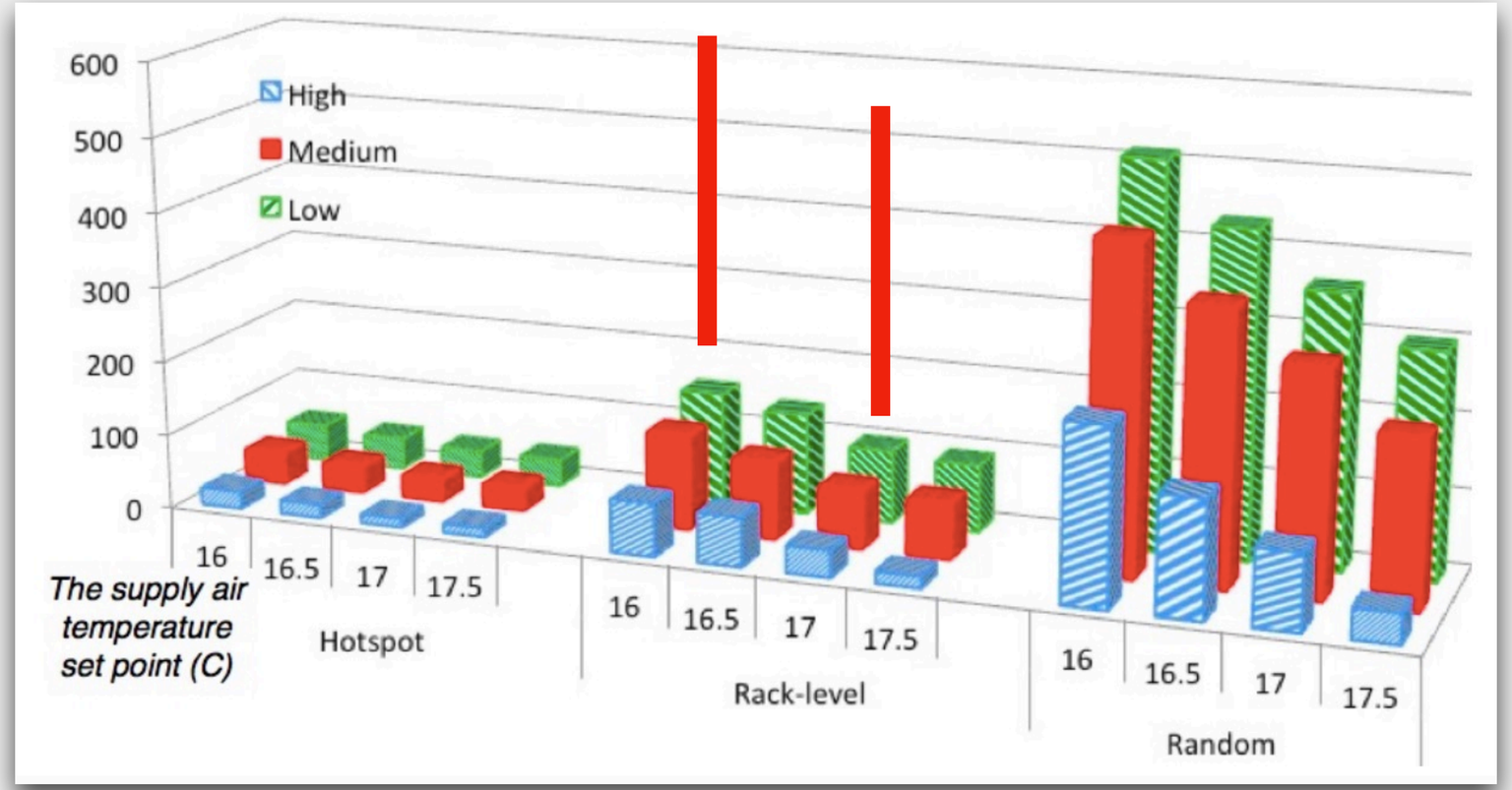
Attacking Efforts



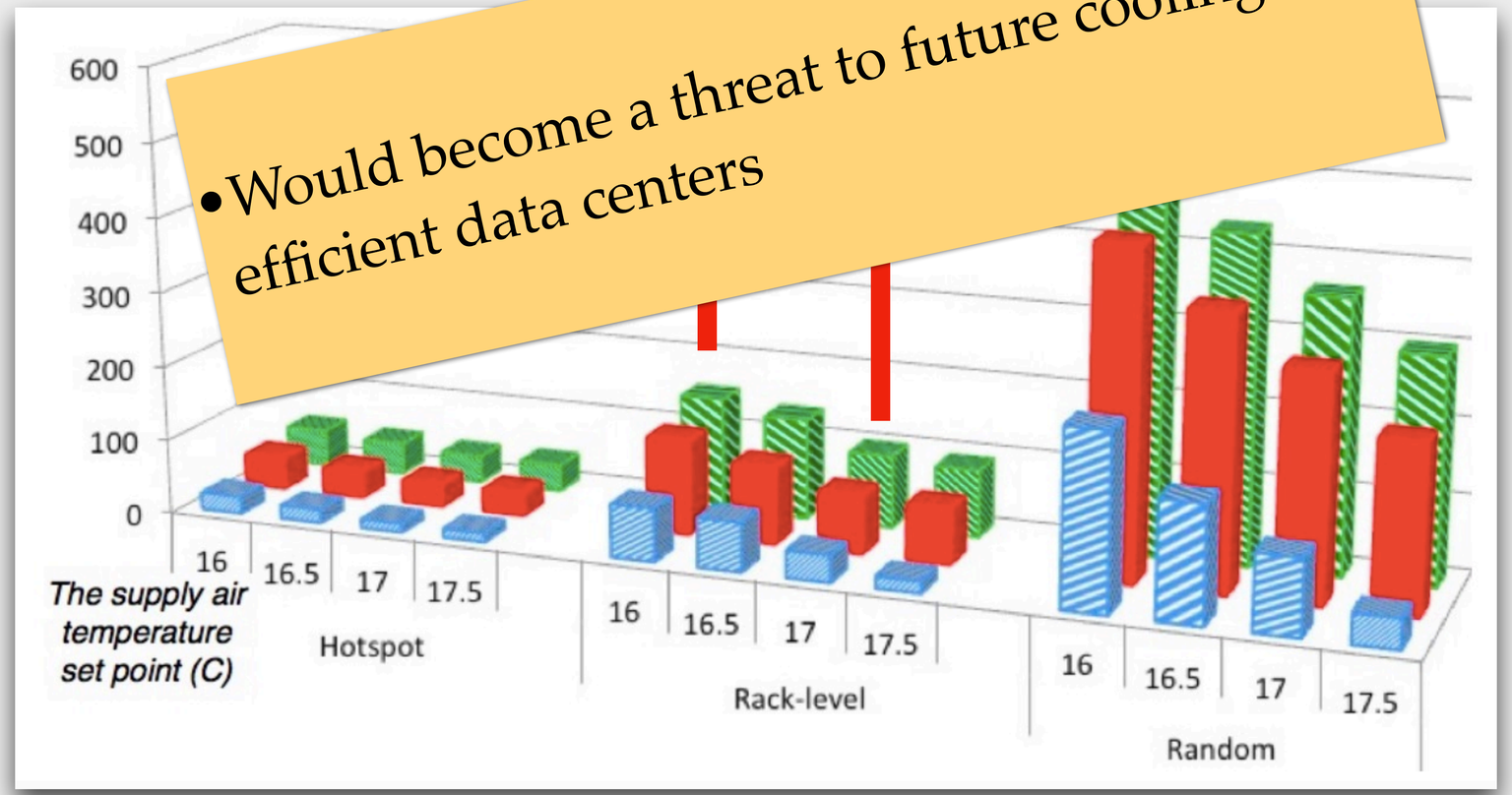
Attacking Efforts



Attacking Efforts



Attacking Efforts



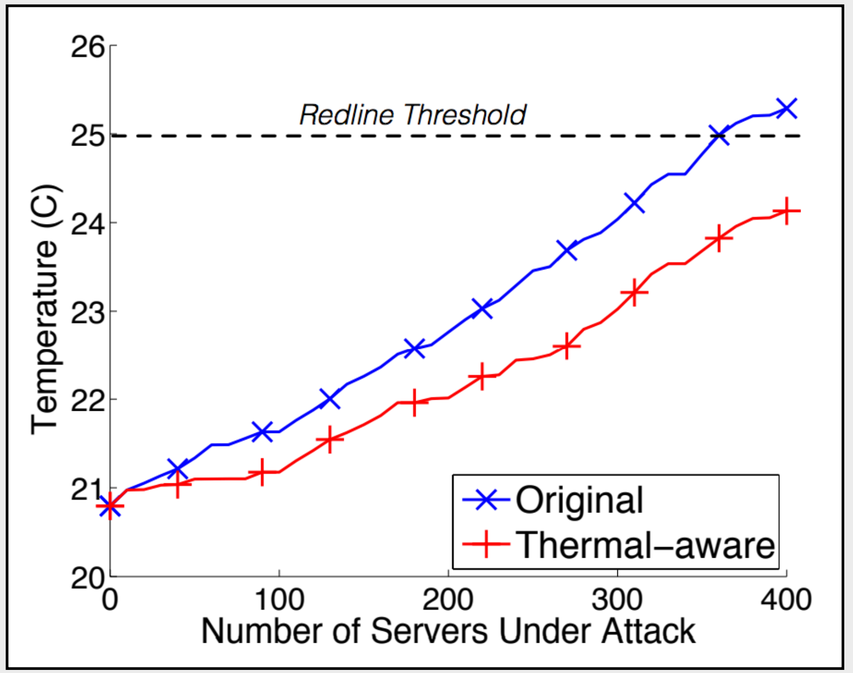
- Would become a threat to future cooling-efficient data centers

Preliminary Mitigating Method

- Dynamic thermal-aware load balancing
  - Consider the thermal conditions and physical locations of servers
- Robust anomaly detection at chip and server levels
- Cost-effective sensing solutions
- Proactive thermal management on rack and data center levels

Preliminary Mitigating Method

- Dynamic thermal-aware load balancing
  - Consider the thermal conditions and physical locations of servers
- Robust anomaly detection at chip and server levels
- Cost-effective sensing solutions
- Proactive thermal management on rack and data center levels



Summary

- Reveal a new vulnerability of reduced cooling redundancy.
  - Root cause: aggressive cooling saving policy and power-oversubscription
- Conduct testbed experiments and data center level simulation.
- Discuss mitigating methods.

Thank you!