Approaches for Resilience Against Cascading Failures in Cloud Datacenters

Haoyu Wang, Haiying Shen and Zhuozhao Li

University of Virginia

2018 IEEE ICDCS, Vienna, Austria

Presented by Cole Miles

UNIVERSITY of VIRGINIA

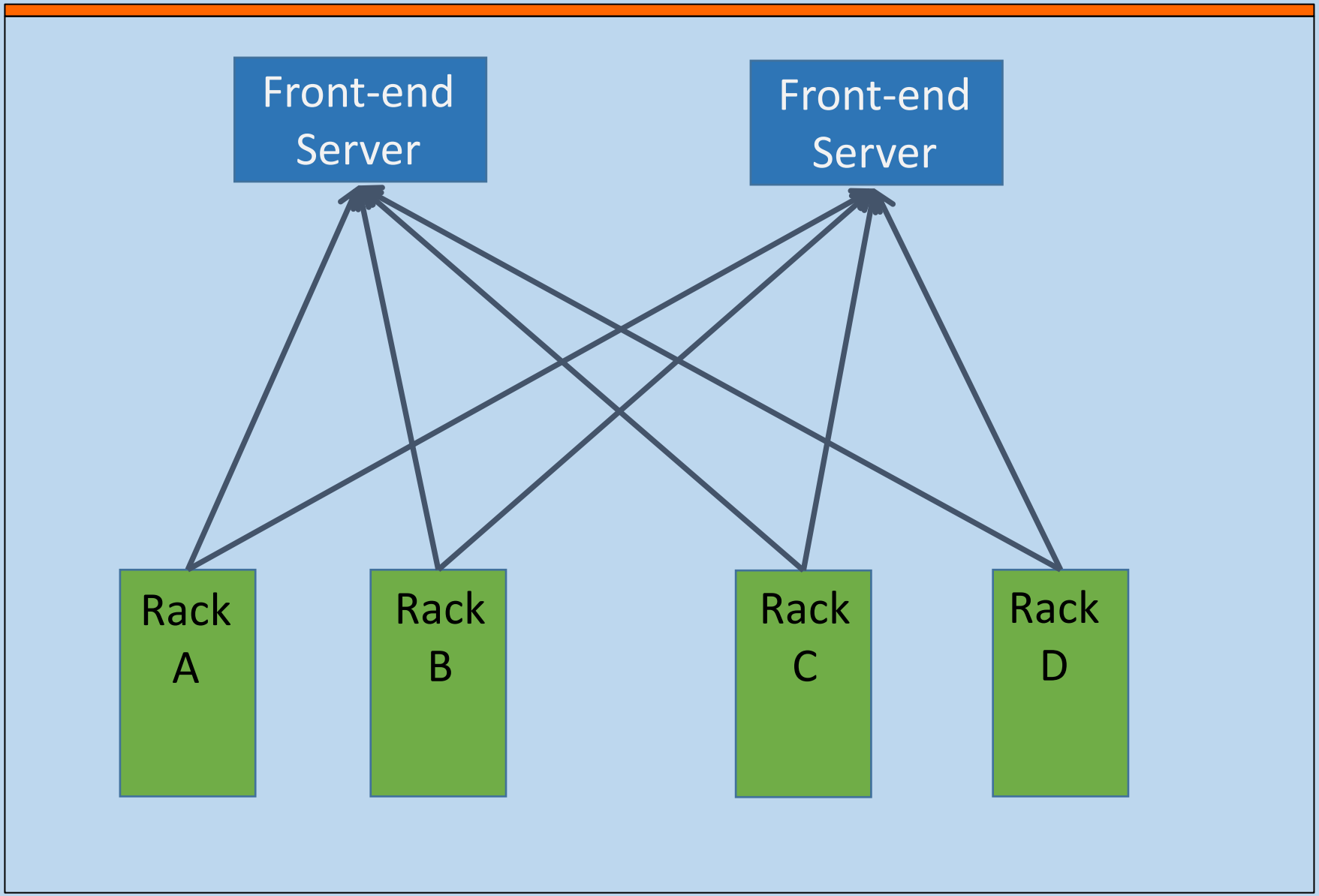
CLOUD

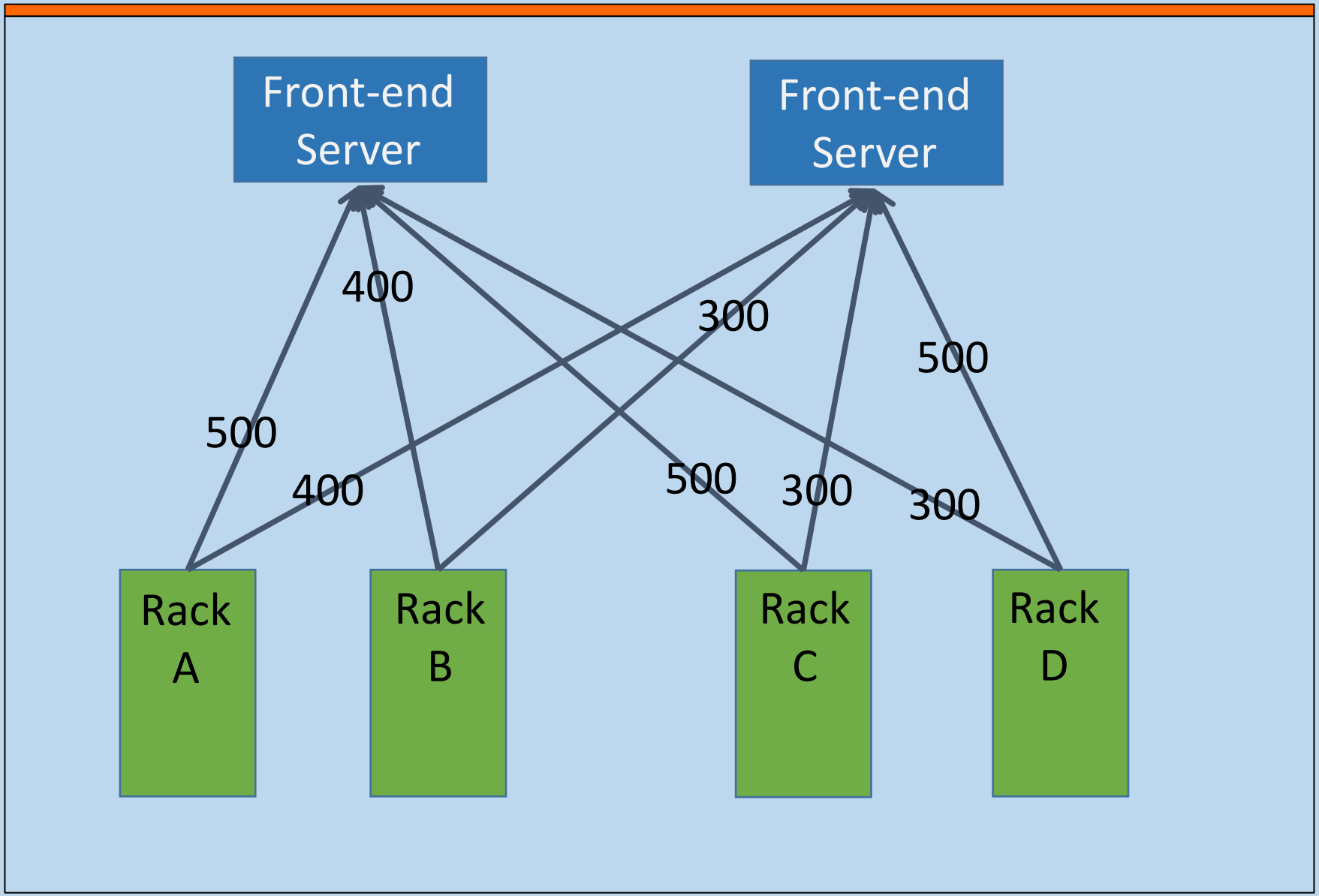
Software Bug, Cascading Failures Caused Amazon Outage

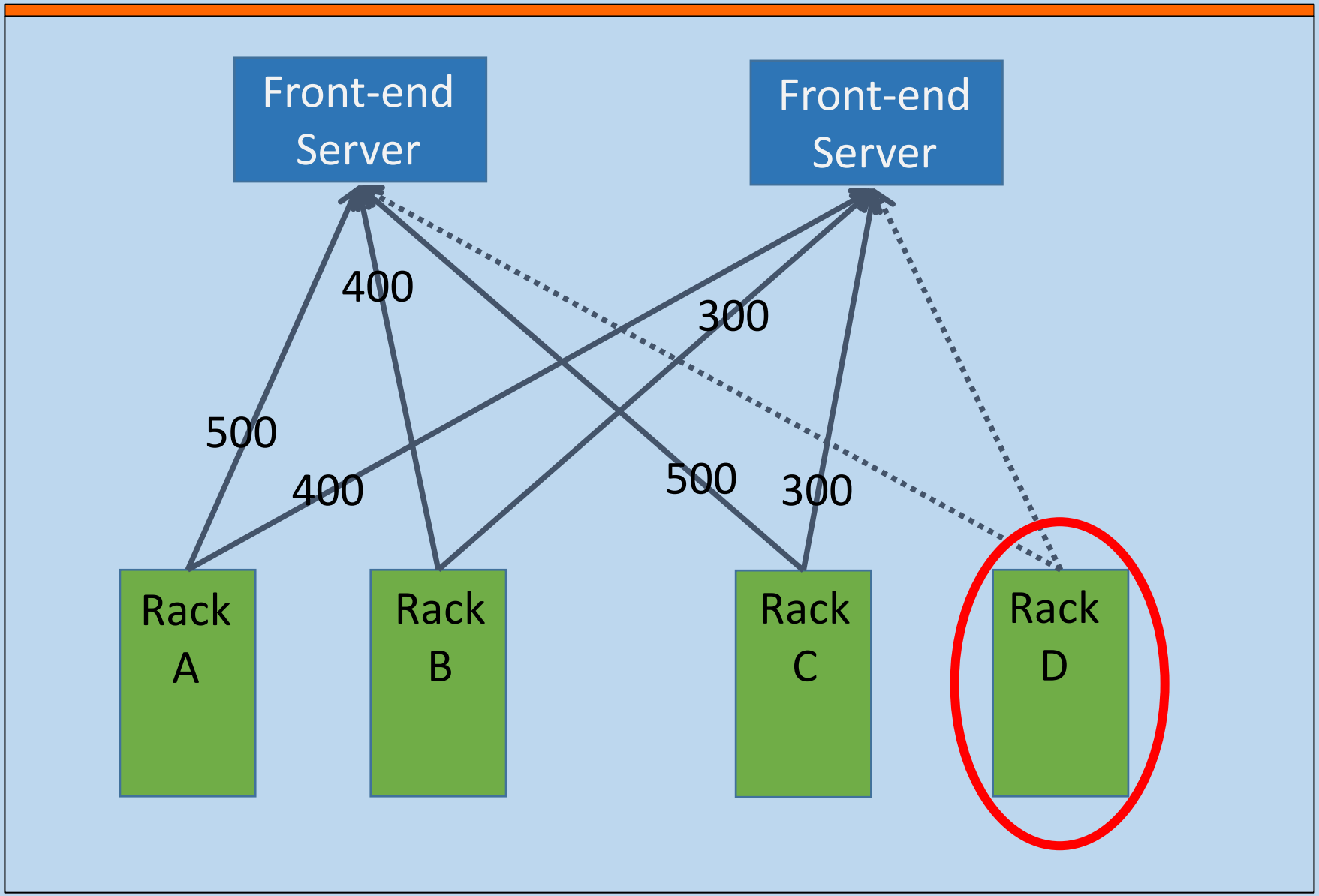
It was a stormy week in the cloud, as an outage at Amazon Web Services

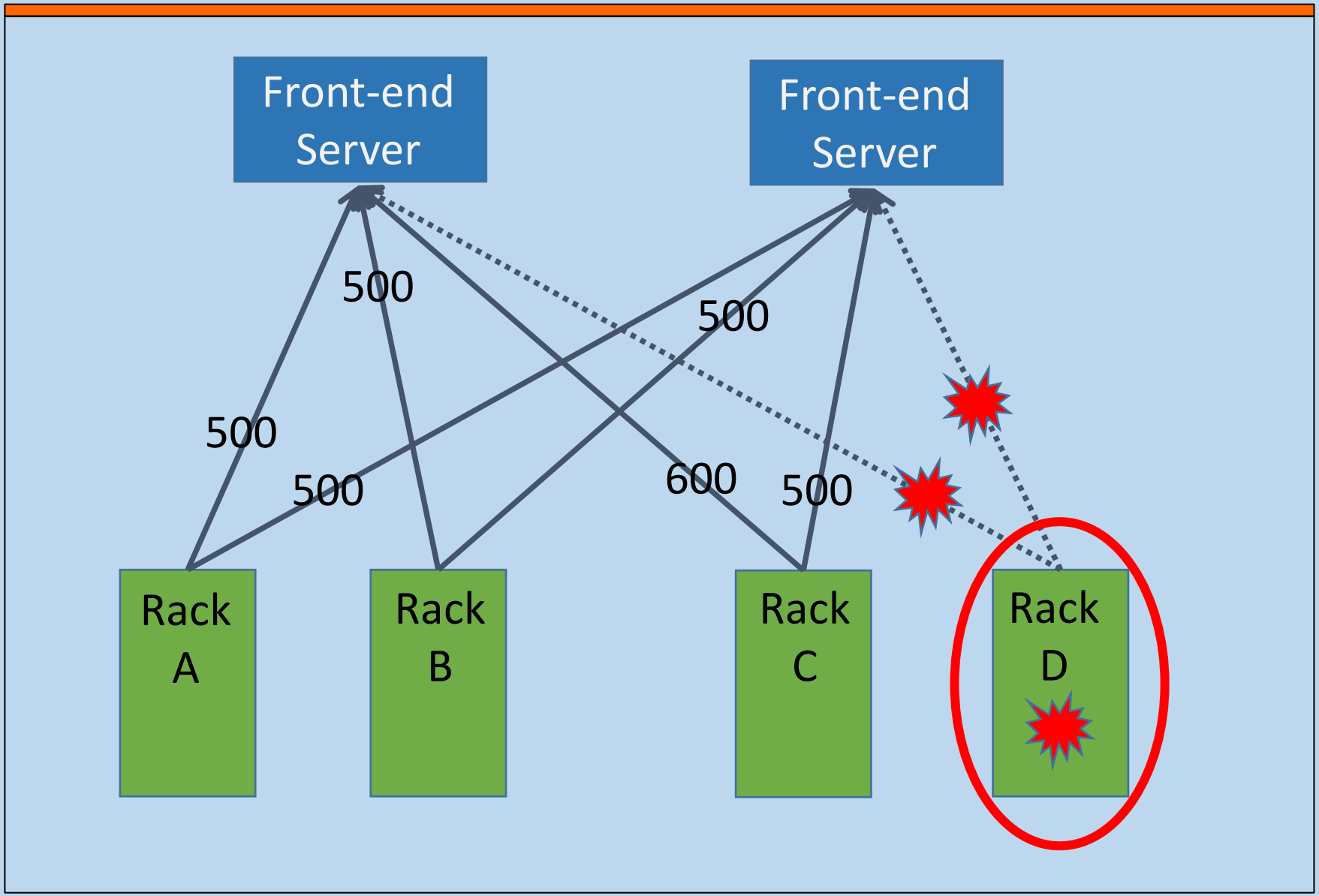
Outline

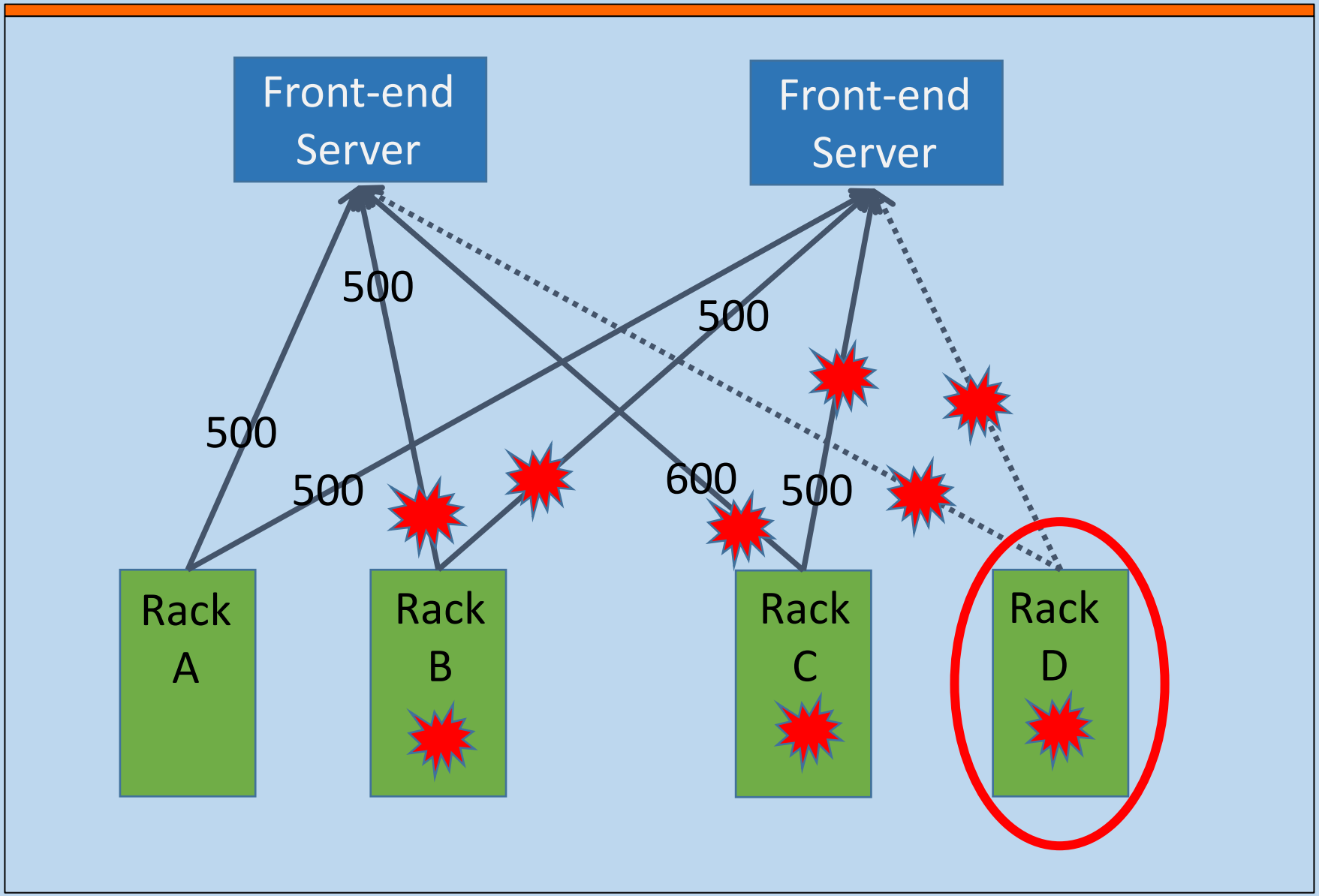
- How Cascading failures happen
- Previous work
- Main design of CFRS (Cascading Failure Resilience System)
- Evaluation of CFRS in simulation

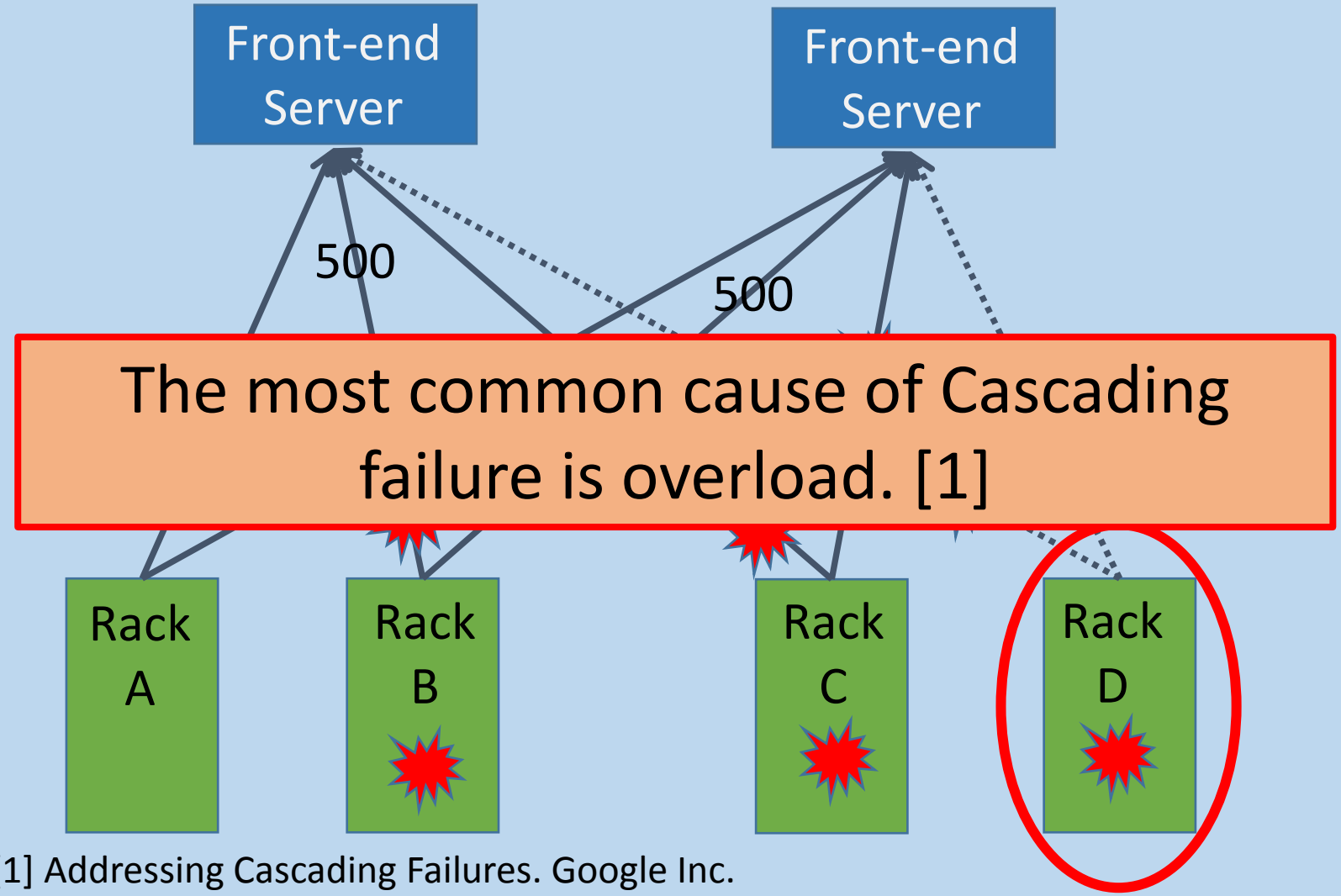

Front-end Server
Front-end Server
Rack A
Rack B
Rack C
Rack D


Front-end Server
Front-end Server
400
300
500
500
400
500
300
300
Rack A
Rack B
Rack C
Rack D


Front-end Server
Front-end Server
400
300
500
400
500
300
Rack A
Rack B
Rack C
Rack D


Front-end Server
Front-end Server
500
500
500
500
600
500
Rack A
Rack B
Rack C
Rack D


Front-end Server
Front-end Server
500
500
500
500
600
500
Rack A
Rack B
Rack C
Rack D



[1] Addressing Cascading Failures. Google Inc.
https://landing.google.com/sre/book/chapters/addressing-cascading-failures.html

Outline

- How Cascading failures happen
- Previous work
- Main design
- Evaluation in simulation

Previous work

VM migration:

- Zhang SIGCOMM'12, Bodik EuroSys'12, Bila INFOCOM'14
- Only consider a time point rather than a time period.

VM backup

- Yeow SIGCOMM'11
- Only for single point failure.

Failure mitigation

- R3 SIGCOMM'11, Netpilot SIGCOMM'12
- Cost of failure repair is very high.

Outline

- How Cascading failures happen
- Previous work
- Main design of CFRS
  - Overload-Avoidance VM Reassignment (OAVR)
  - VM Backup Set Placement (VMset)
  - Dynamic Oversubscription ratio Adjustment (DOA)
- Evaluation of CFRS in simulation

- Main design of CFRS
  - Overload-Avoidance VM Reassignment (OAVR)

Three rules:

1. VMs with higher workloads should be scheduled first.
2. Migrate VMs with the highest workload on some resource types to the most underloaded PMs on the resource types.
3. A VM should be migrated to its best-fit PM.

- Main design of CFRS
  - VM Backup Set Placement (VMset)

$$L_{q,e_k}^{P_i} = \sum_{v_j \in V_{P_i}} l_{q,e_k}^{v_j} \tag{1}$$

$$U_{q,e_k}^{P_i} = \begin{cases} L_{q,e_k}^{P_i} - C_q^{P_i} & \text{if } L_{q,e_k}^{P_i} \neq C_q^{P_i} \\ -c & \text{otherwise} \end{cases} \tag{2}$$

$$S_{P_i}^{v_j} = \sum_{e_k \in T} \vec{U}_{e_k}^{P_i} \bullet \vec{l}_{e_k}^{v_j}. \tag{3}$$

$$L^{v_j} = \sum_{e_k \in T} |\vec{l}_{e_k}^{v_j}|. \tag{4}$$

- Main design of CFRS
  - VM Backup Set Placement (VMset)

For instance, assume the datacenter has the following parameters: R = 3, N = 12, and W = N-1 = 11.

$$\frac{\#\ \ vmsets}{\binom{N}{R}} = \frac{220}{\binom{12}{3}} = 100\%. \qquad (6)$$

If W=4.

$$\frac{\#\ \ vmsets}{\binom{N}{R}} = \frac{72}{\binom{12}{3}} = 32.7\%. \qquad (8)$$

Using a lower spread width (W) can decrease the probability of VM backup loss from correlated failures.

- Main design of CFRS
  - VM Backup Set Placement (VMset)

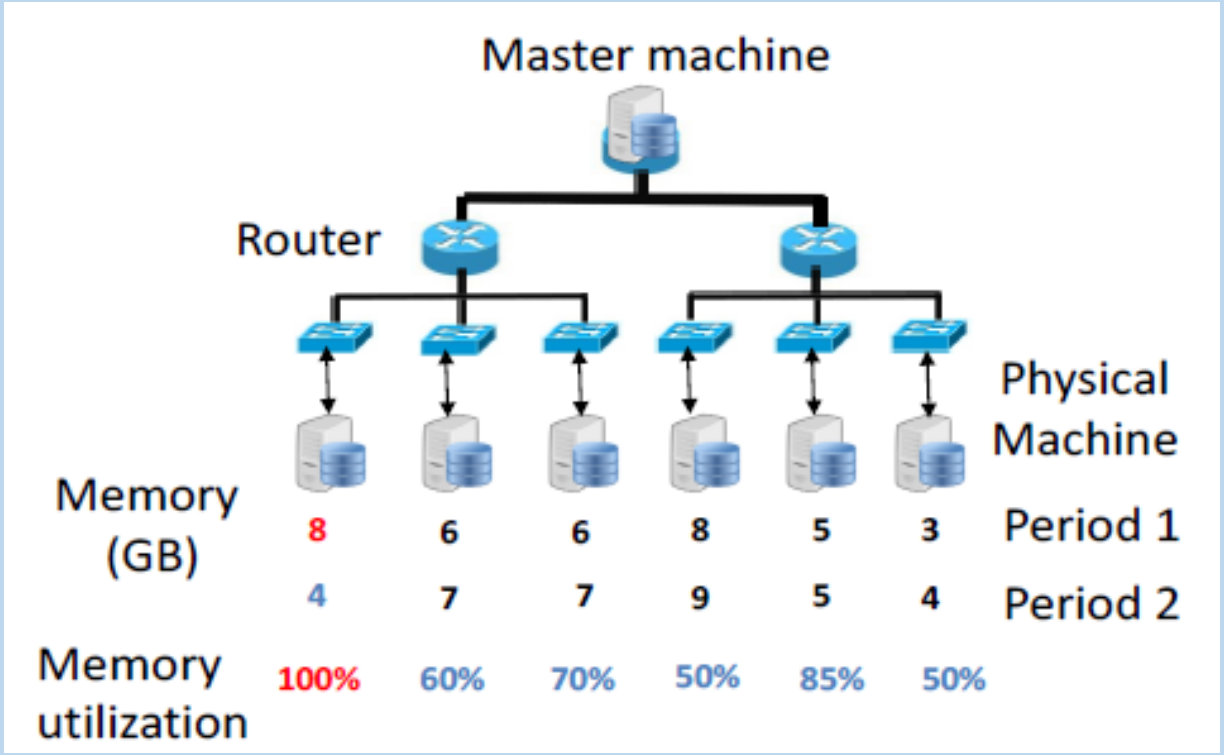
$$\{P_1, P_2, P_3\}\{P_4, P_5, P_6\}\{P_7, P_8, P_9\}\{P_{10}, P_{11}, P_{12}\},$$
$$\{P_1, P_5, P_9\}\{P_2, P_6, P_{10}\}\{P_3, P_7, P_{11}\}\{P_4, P_8, P_{12}\},$$

$$\frac{\#\ VMsets}{\binom{N}{R}} = \frac{8}{\binom{12}{3}} = 4\%, \tag{10}$$

For instance, assume the datacenter has the following parameters: N = 5000, R = 3, W = 10, when 1% of the PMs fail simultaneously.

$$\frac{\#\ VMsets}{\binom{N}{R}} = \frac{\frac{10}{3-1} \cdot \frac{5000}{3}}{\binom{5000}{3}} = \frac{8333}{2.01 \times 10^{10}} = 0.000042\%, \tag{11}$$

- Main design of CFRS
  - Dynamic Oversubscription Ratio Adjustment (DOA)



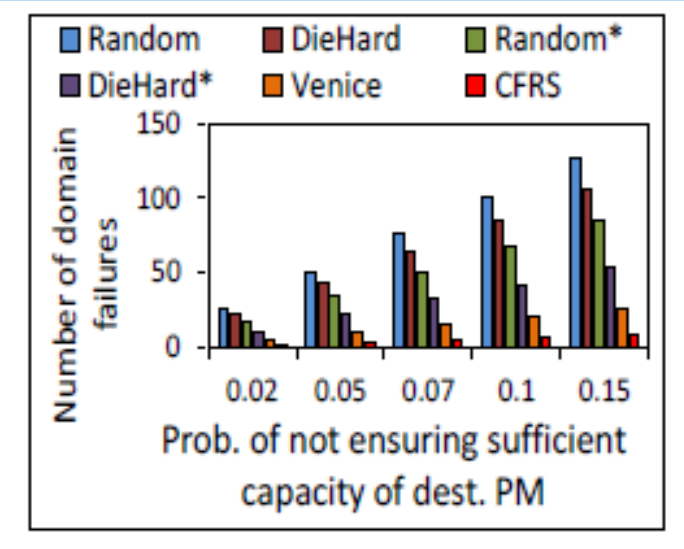
Outline

- How Cascading failures happen
- Previous work
- Main design
- Evaluation in simulation

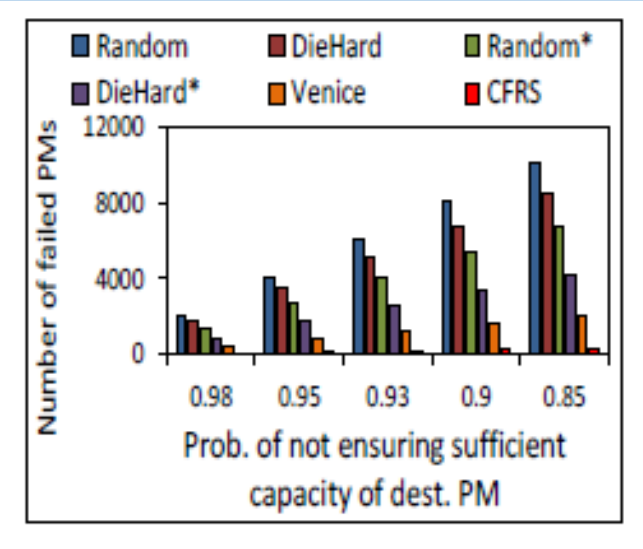
- Evaluation
  - Simulation Setup

1. Google Cluster trace
2. 19200 PMs are connected through 240 Top-of-Rack switches. 80 PMs are in one rack, each power station supplies 20 racks.
3. 240 network failure domains and 12 power failure domains.
4. The failure rate was randomly chosen from [0.000022, 0.000032] per hour for a network failure domain and 0.4*10e-6 per hour for a power failure domain. The failure rate of overloaded PM is 0.0001 per minute.

- Evaluation
  - Results

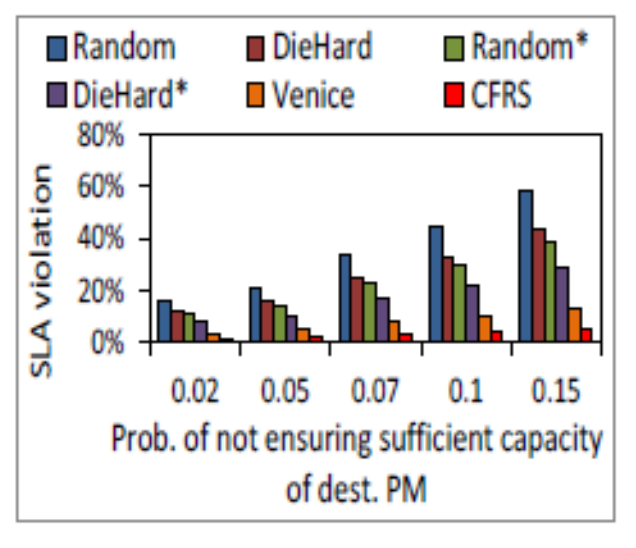


Number of domain failures

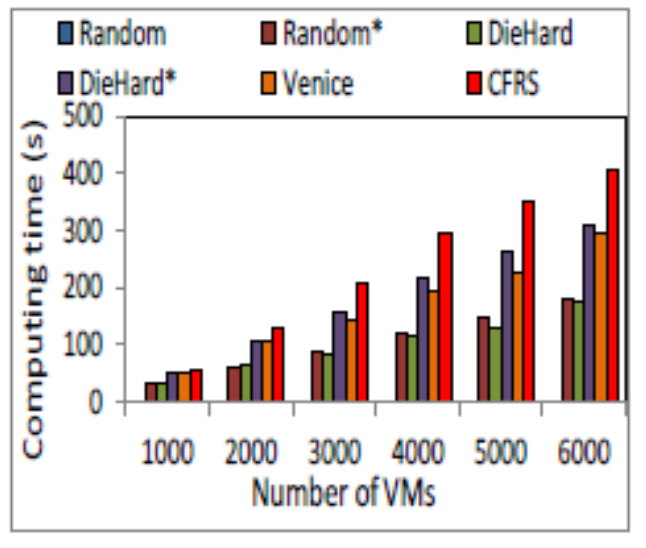


Number of failed PMs

- Evaluation
  - Results



SLO violations



Computing time

Conclusion

1. CFRS aims to achieve long-term load balance by VM migration, which can avoid cascading failures for long-term.
2. CFRS places VM backups to PMs to increase the backup reliability in failures.
3. CFRS dynamically adjusts oversubscription ratio.
4. The trace simulation shows the superior performance of CFRS in cascading failure avoidance.

Thank you!

Questions?