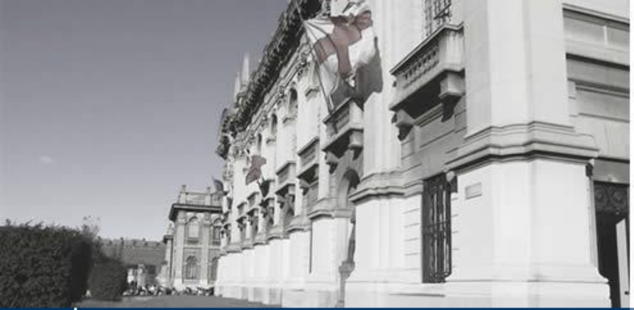


15th International Conference on the Design of Reliable Communication Networks (DRCN 2019)

Session: *Disaster Resilience of Communication Networks*
Supported by COST CA15127-RECODIS Project

POLITECNICO DI MILANO

Efficient Online Virtual Machines Migration for Alert-Based Disaster Resilience

O. Ayoub, O. Huamani, F. Musumeci, **Massimo Tornatore**

Department of Electronics, Informatics and Bioengineering, Politecnico di Milano

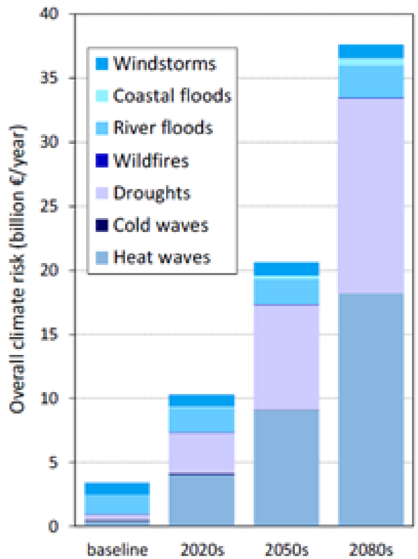
University of Coimbra, Thursday, March 21, 2019

Outline

- Background
  - Weather-Based Disasters
  - Data Center Virtualization
  - Virtual Machines (VMs) Migration
- Disaster-Resilient Online VM Migration
  - Motivation
  - Problem Statement
  - Integer Linear Programming Model
  - Evaluation Settings
  - Results
- Conclusion

Background
Weather-Based Disasters

- Welcome to Anthropocene!
  - Geological era when/where earth system processes (e.g., atmospheric) are altered by humans
- Extreme weather conditions will become much more common
  - Europe is expected to see a progressive and very strong climate change



2017 was the most costly year ever due to weather-based disasters (€3.4 billion)

Source: "Escalating impacts of climate extremes on critical infrastructures in Europe", *Global Environmental Change, 2017*

Background
Weather-Based Disasters: impact on data centers

- 10% of data center disruptions are caused by natural disasters*
- Hurricane Sandy, New York, 2012: Sudden flooding caused extensive data center disruption
  - Backup systems were located in same geographic area!!

- Data evacuation as a prime solution.. But is it enough in the five 9's availability era?

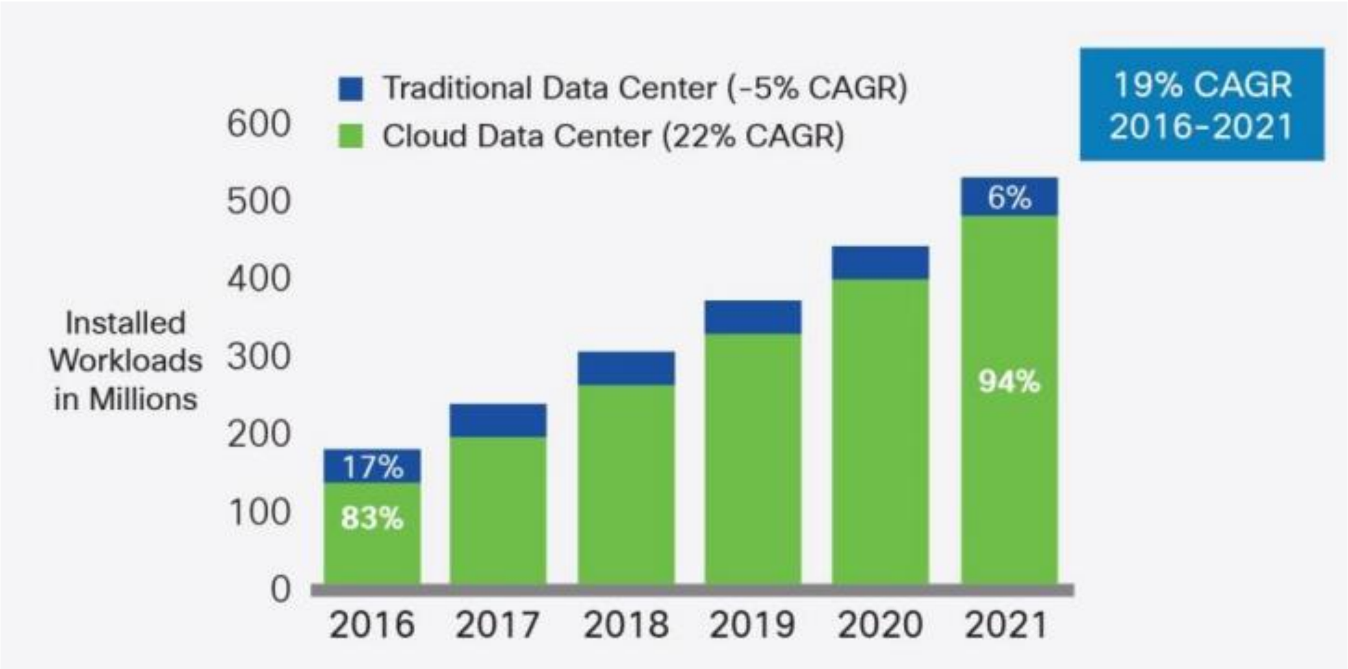
Rethinking disaster resiliency

Not only avoid data-loss but also avoid service-disruptions thanks to data center virtualization

*Source: https://www.365datacenters.com

Background
Data Center Virtualization

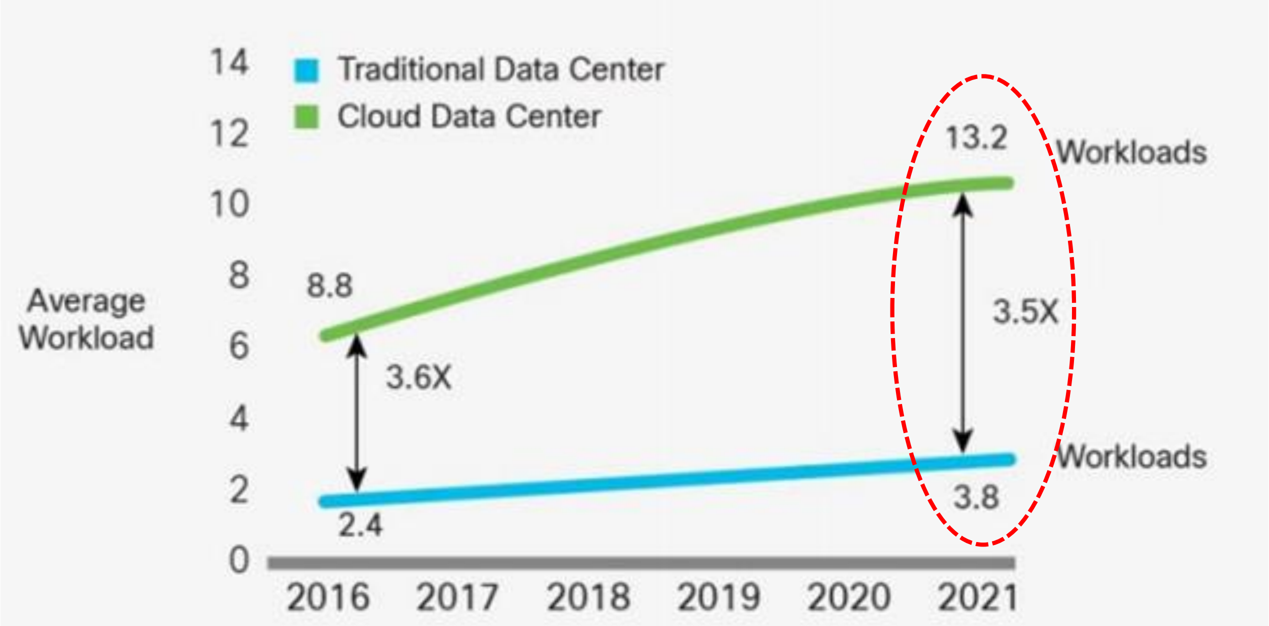
- Continued global data center virtualization
  - 83% of all DC workload were in Virtual Machines (VMs) in 2016
  - Expected to reach 94% in 2021



Source: Cisco Global Cloud Index, 2016-2021

Background
Data Center Virtualization

- Main factors to DC virtualization:
  - Higher physical resource utilization *(3.5X with virtualization)*

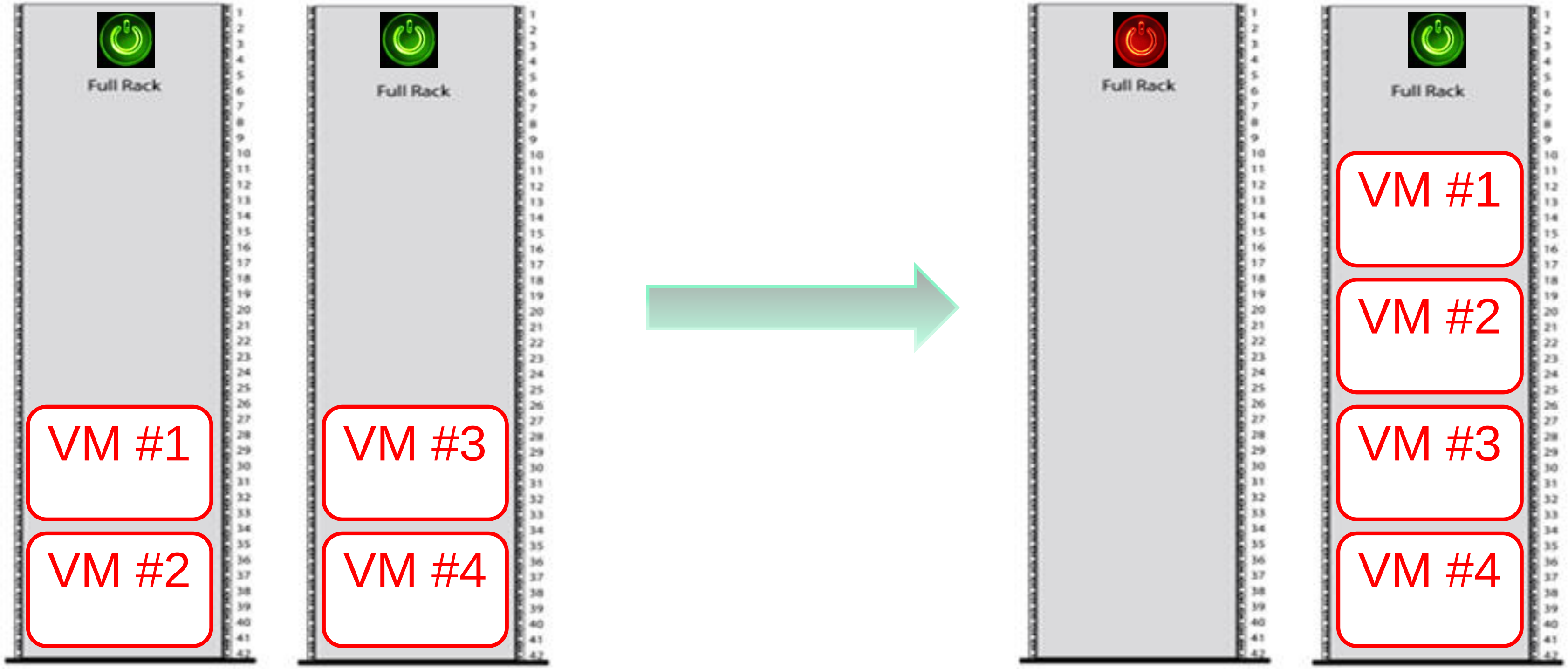


Source: Cisco Global Cloud Index, 2016-2021

  - Dynamic migration of VM/workloads to meet dynamic demands
  - VMs Migration: Transferring all VM data, i.e., disk memory and processors states from one server to another

Background
VM Migration: Intra-Data Center

- Consolidating VMs into servers and servers into racks
- Objective: power (..hence cost) savings

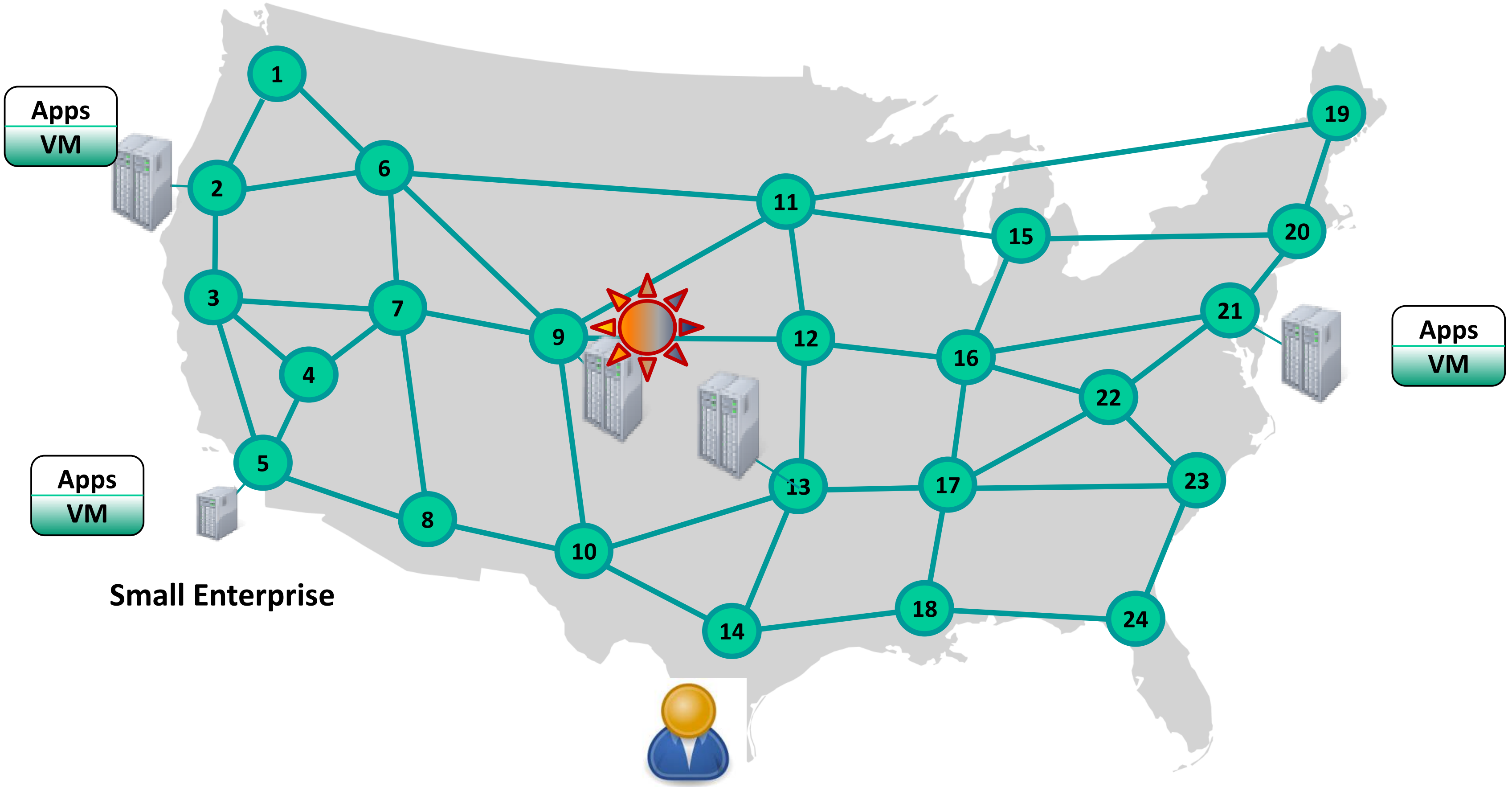


Each rack has its own power and cooling

Background

VM Migration: Inter-Data Center

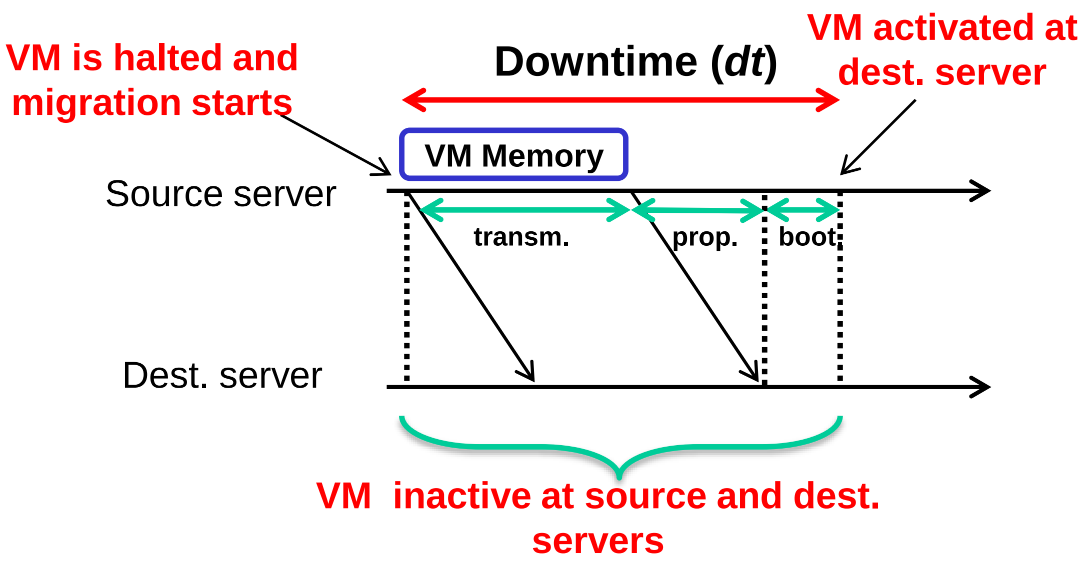
- Still consolidating VMs into servers and servers into racks, but <u>across geographically distributed DCs</u>
- Objectives: quality of service, cloud bursting, use of renewable energy



Background

VMs Migration: Offline vs. Online

- Offline
  - VM is halted at the source
  - All data (we refer to this by *VM size*) is transferred
  - VM is set up again at destination



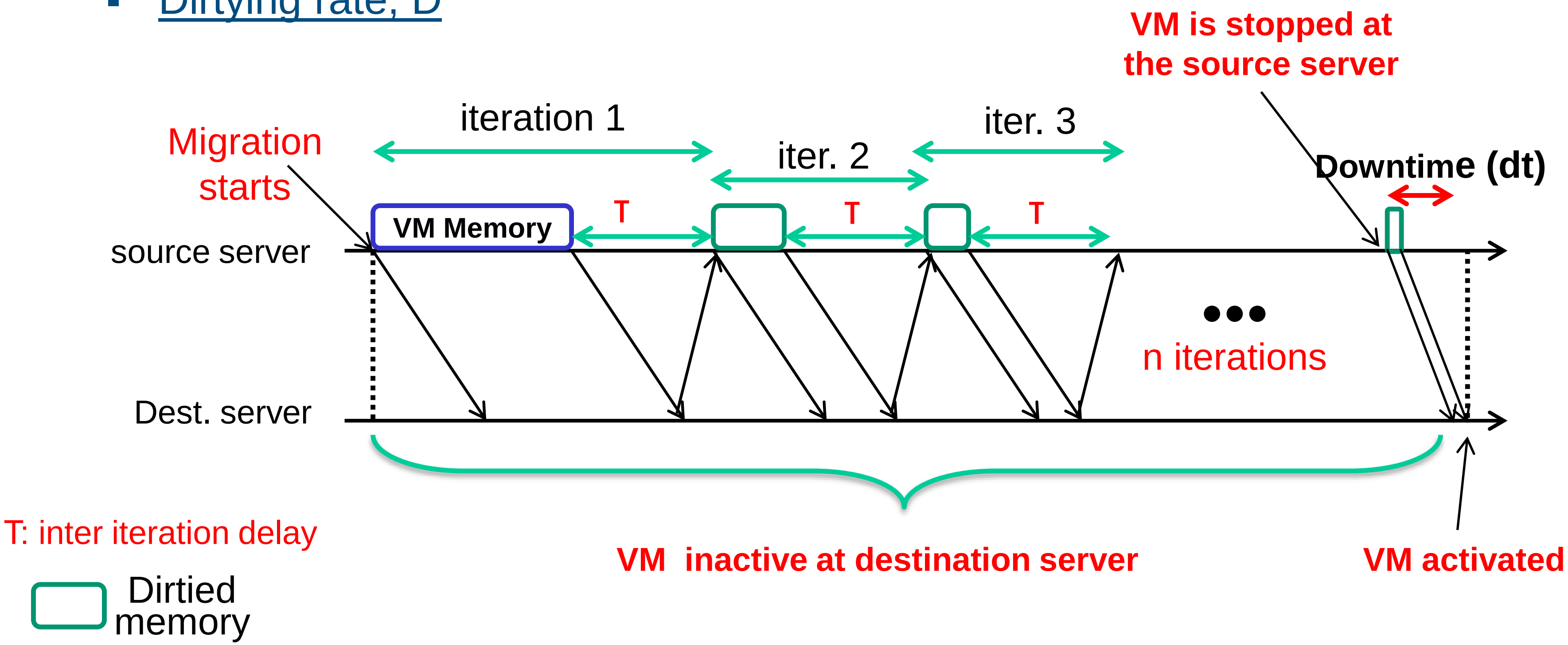
$$T_{mig,off} = \frac{VM\ size}{B}$$

Untoleratable Service Downtime?

Migration time (almost) coincides with downtime

Background
VMs Migration: Offline vs. Online

- Online
  - Iterative data transfer (while VM is running)
  - Users still access the server and modify VM memory (dirty memory)
  - Dirtying rate, D



$$T_{mig,on} = \frac{VM\,size}{B} + F\{D, B, VM\,size\}$$

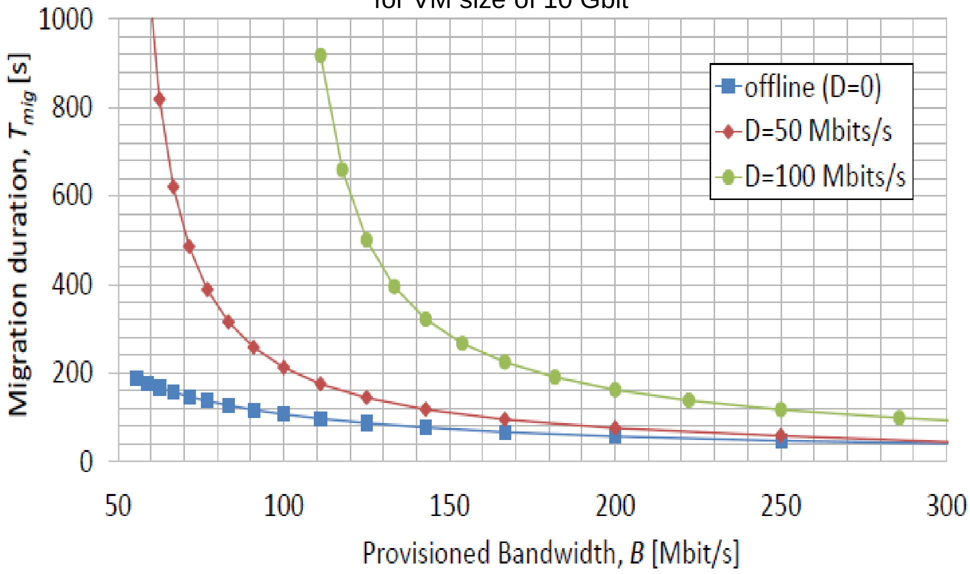
VMs Migration Duration: Offline vs. Online

- Migration Duration (T_{mig}) vs. Assigned Bandwidth (B_{mig})
- Online VM migration → Significant Network *Resource Occupation*
  $RO = B_{mig} \cdot T_{mig} \cdot Hops$ *[Mbit]*

	Offline	Online
B_{mig}	200 Mbps	200 Mbps
T_{mig}	50 sec.	180 sec.
Hops	4	4
RO	40,000	144,000
dt	50 sec.	≈0 sec.

Function-points curve

for VM size of 10 Gbit



NB: Selection of migration bandwidth to minimize RO is not trivial

Outline

- Background
  - Weather-Based Disasters
  - Data Center Virtualization
  - Virtual Machines (VMs) Migration
- Disaster-Resilient Online VM Migration
  - Motivation
  - Problem Statement
  - Integer Linear Programming Model
  - Evaluation Settings
  - Results
- Conclusion

Disaster-Resilient Online VM Migration
Motivation

- A baseline approach in case of a disaster warning ("alert time")

1 — Stop all VMs (terminate services)

2 — Migrate VMs offline (data evacuation)

3 — Restart VMs

- Five 9's availability requirements demand service providers to always keep services alive
- Do we have enough time to keep VMs alive?

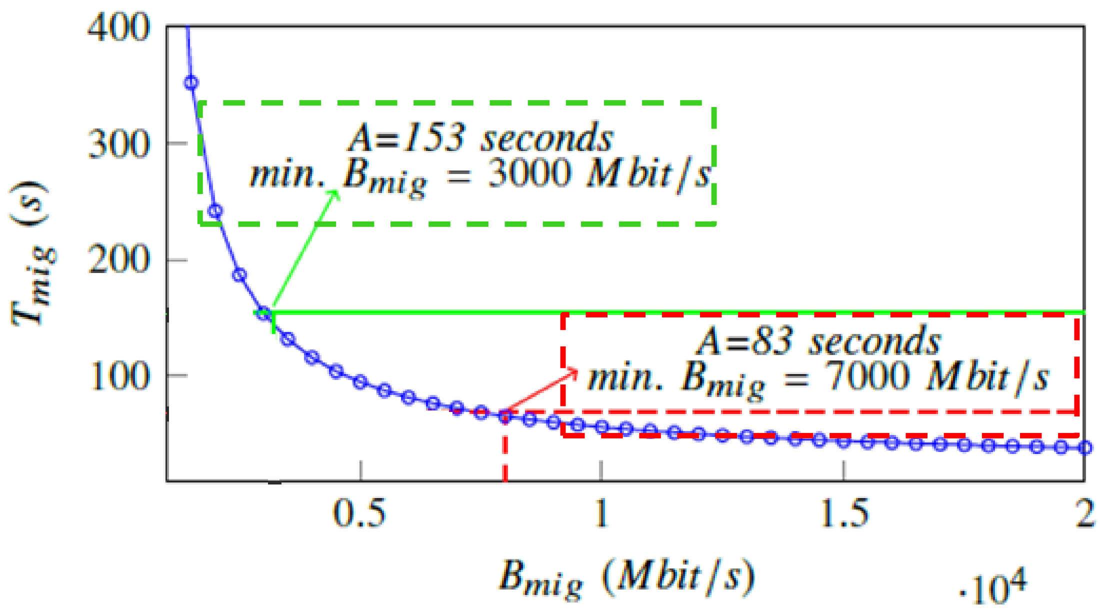
<u>Online VM migration emerges as a solution</u>

Performing efficient evacuation (i.e., selecting destination DC, route, migration bandwidth) is not trivial

- Alert time (*A*) limits the range of bandwidth values that may be utilized for VM migration

+ Many VMs are involved
+ Limited network capacity
+ Limited available time

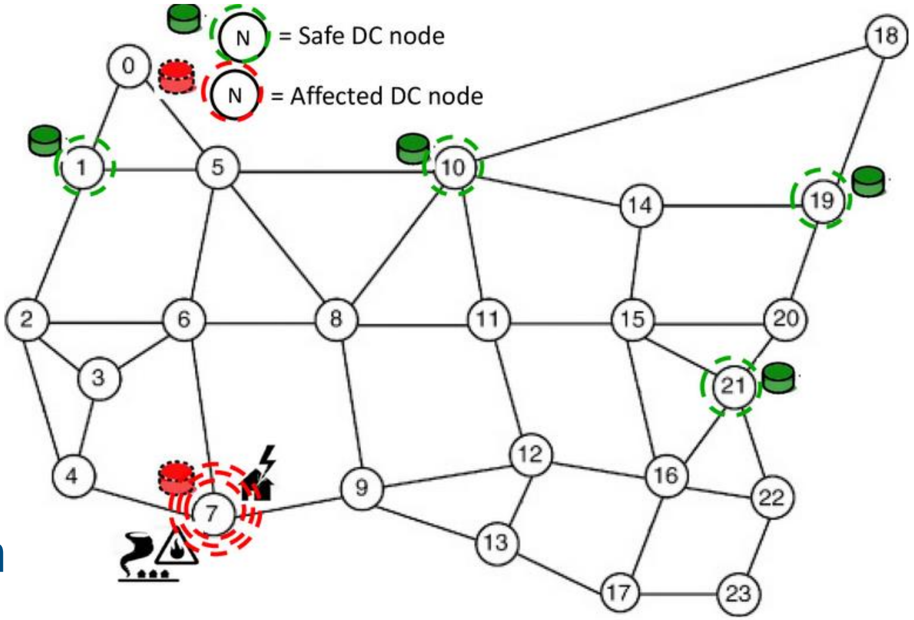
Also scheduling is decisive



Minimum acceptable B_{mig} for two different values of alert *A*

Disaster-Resilient Online VM Migration
Problem Statement

- Given
  - Network with a risky DC and safe DC locations
  - VMs to be migrated (size, dirtying rate, function-points curve)
  - Alert time/evacuation deadline (seconds)



- Decide, for each VM:
  - If migrated, online or offline
  - Destination DC
  - Route and migration bandwidth
  - Starting (and ending) time of migration (scheduling)
    - Note that VM migration duration is not known a-priori

- Multi-objective:
  1) Maximize number of VMs migrated (either offline of online)
  2) Minimize downtime (maximize online migration)
  3) Minimize Resource Occupation *(RO)*
  4) Minimize migration duration (T_{mig})

$$\min\left(\sum_{v\in V}\left(-\alpha(x_v+y_v)+\beta\psi_v+\sum_{(i,j)\in E}\sum_{t\in T}\gamma z_{v,t}^{i,j}\omega_v+\delta\lambda_v\right)\right)$$

# of Migrated VMs

$x_v = 1$ *iff* VM v was migrated online

$y_v = 1$ *iff* VM v was migrated offline

RO*: sum of all bandwidth occupied on all links at all time instances for all VMs*

***Ψ_v* = overall downtime**

Migration duration *of VM v*

α, β, γ and δ are positive constants used to set the priority of different terms

- Subject to:
  - Migration process and bandwidth constraints
  - Scheduling constraints
  - Flow and capacity constraints

- Migration process and bandwidth constraints

$$x_v + y_v \leq 1, \quad \forall v \in V$$

VM v may be migrate either online or offline

If online, a function-point is chosen and migration duration is found accordingly

$$\sum_{i \in I} l_{v,i} = x_v, \quad \forall v \in V$$

$$b_v^{on} = \sum_{i \in I} b_{i,v}^{on} \cdot l_{v,i}, \quad \forall v \in V$$

$$d_v^{on} = \sum_{i \in I} d_{i,v}^{on} \cdot l_{v,i}, \quad \forall v \in V$$

If offline, a migration bandwidth value is selected and migration duration is simply calculated

$$b_v^{off} \leq M \cdot y_v, \quad \forall v \in V$$

$$d_v^{off} = \frac{S_v}{b_v^{off}}, \quad \forall v \in V$$

linearized

Disaster-Resilient Online VM Migration

ILP Model (3/3)

- Scheduling constraints

$$e_v \leq A, \quad \forall v \in V$$

$$e_v = r_v + \lambda_v, \quad \forall v \in V$$

Migration ends before the alert time (evacuation deadline)

This scheduling problem is different than traditional ones. Migration duration is not known a-priori and is a decision variable in the problem

$$k_t - r_v + 1 \leq M \cdot n_{v,t}, \quad \forall v \in V, t \in T$$

$$e_v - k_t \leq M \cdot m_{v,t}, \quad \forall v \in V, t \in T$$

$$A - r_v = \sum_{t \in T} n_{v,t}, \quad \forall v \in V$$

$$e_v = \sum_{t \in T} m_{v,t}, \quad \forall v \in V$$

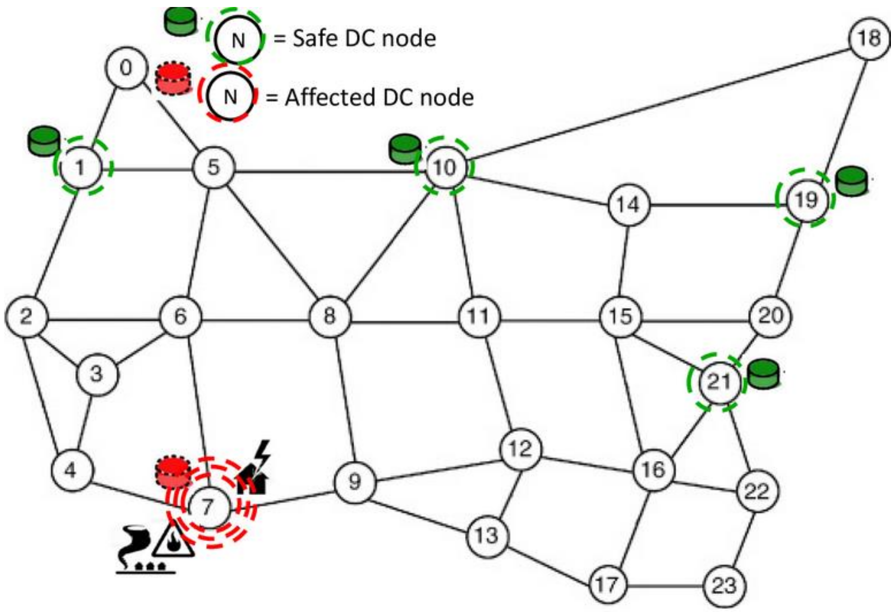
$$w_{v,t} \leq n_{v,t} \cdot m_{v,t}, \quad \forall v \in V, t \in T$$

Set of constraints that capture scheduling

$$z_{v,t}^{i,j} \geq w_{v,t} \cdot h_v^{i,j}, \quad \forall v \in V, t \in T, (i,j) \in E$$

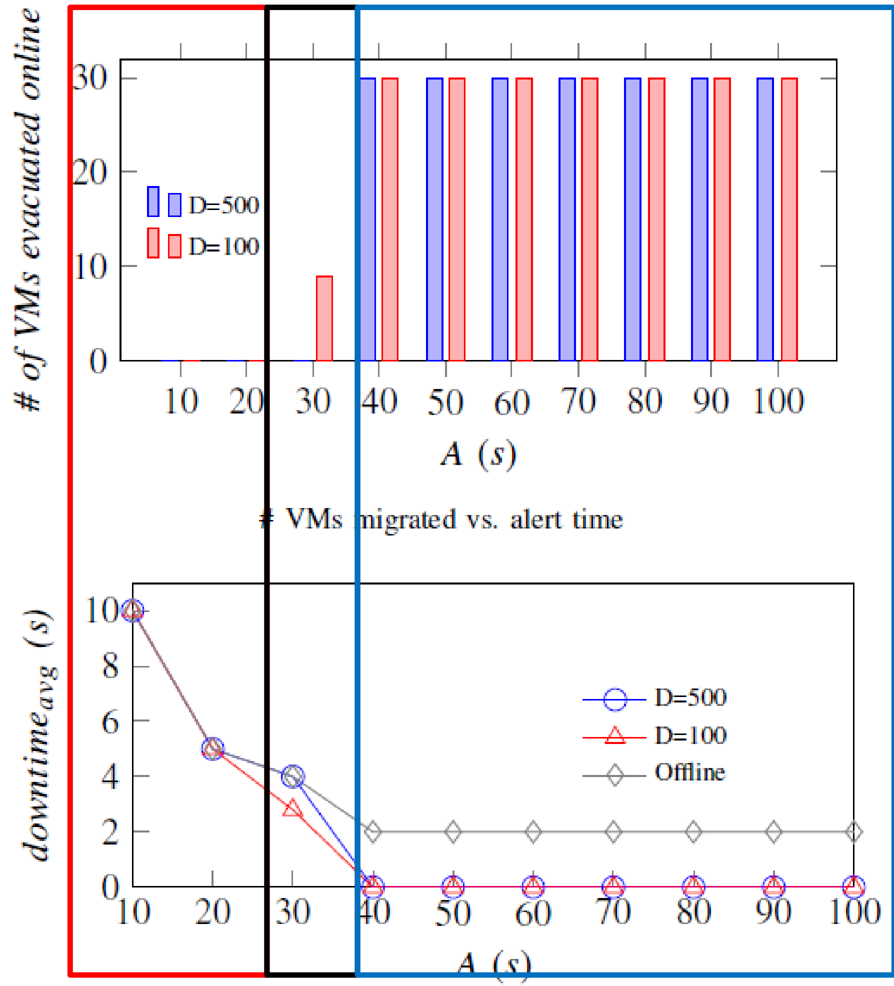
Disaster-Resilient Online VM Migration
Evaluation Settings

- USA topology
  - 24 nodes (5 are DC locations)
  - 43 links@100 Gbps
- 30 VMs (size 40 GB)
- *Dirtying rate D:*
  100 Mbps and 500 Mbps
- Alert time *A*: (10s -- 100s)
  (short alerts)
- 3 strategies



Strategies		RO-minimized	T-minimized	Offline
Multi-objective priority	1	Number of VMs migrated (offline + online)		
	2	Service downtime		
	3	Average *RO*	Average T_{mig}	
	4	Average T_{mig}	Average *RO*	

- On # of online migration
  - For A ≤ 20 sec. no VMs are migrated online
  - For A = 30 sec., 10 VMs are migrated online (D = 100 Mbps)
  - For A ≥ 40 sec., all VMs are migrated online
- On downtime
  - Downtime decreases for higher A as it allows utilizing higher bandwidth values and scheduling

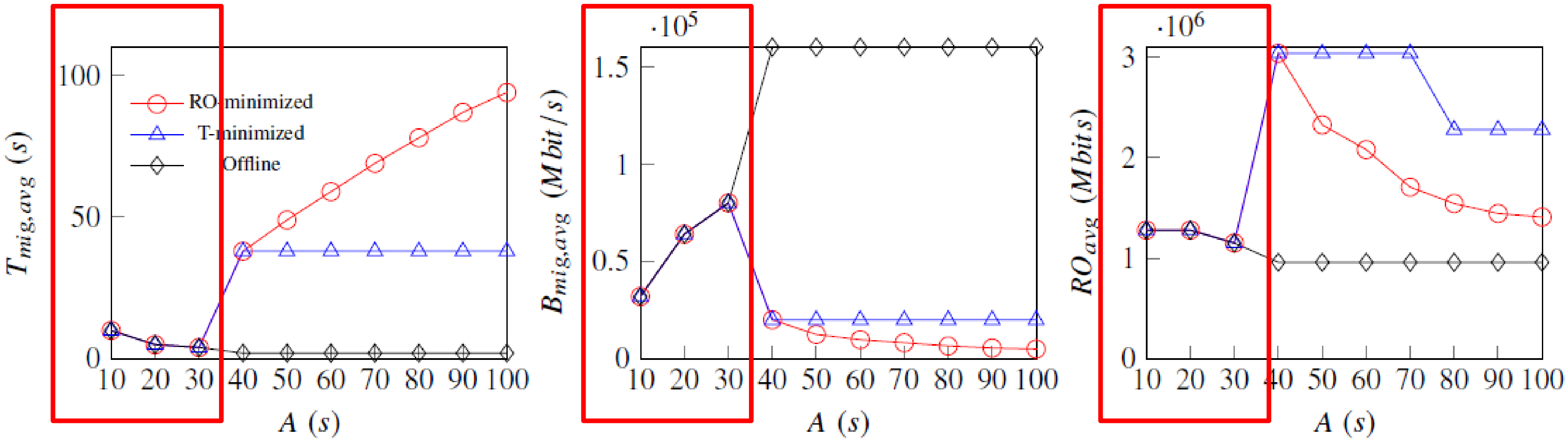


VMs migrated vs. alert time

Downtime vs. alert time

Disaster-Resilient Online VM Migration

Results (2/4): RO-minimized vs T-minimized vs Offline



$T_{mig,avg}$ vs. alert time

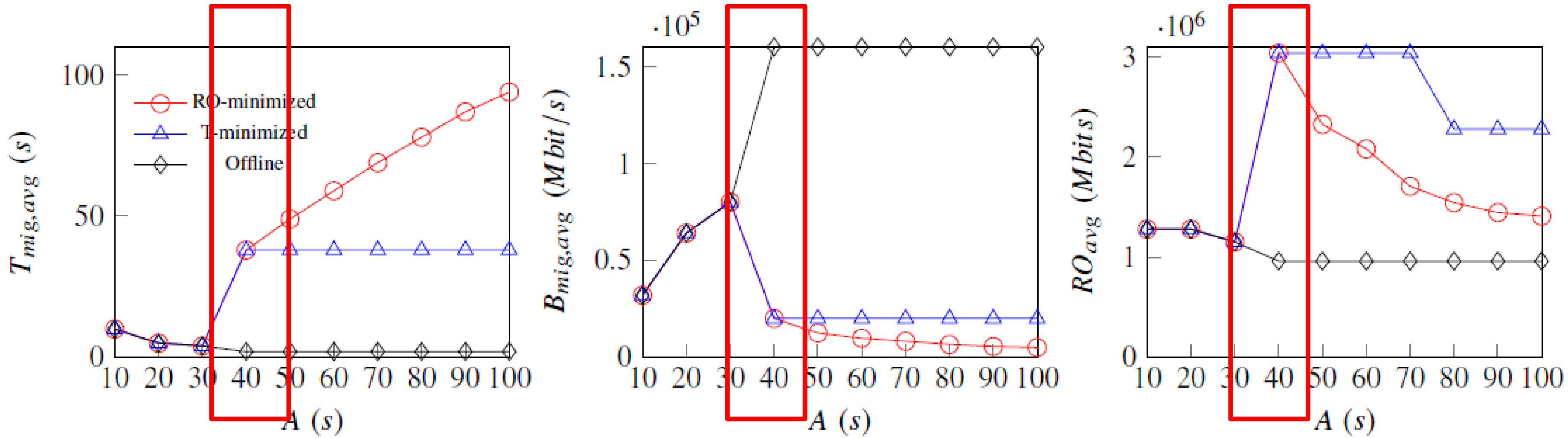
$B_{mig,avg}$ vs. alert time

RO_{avg} vs. alert time

- For 10 ≤ A ≤ 30 sec.
  - All VMs are migrated offline
  - All strategies show exact same behavior
  - B_{mig} (T_{mig} *and RO*) increases (decreased) with *A*, as higher A permits utilizing higher bandwidth

Disaster-Resilient Online VM Migration

Results (3/4): RO-minimized vs T-minimized vs Offline

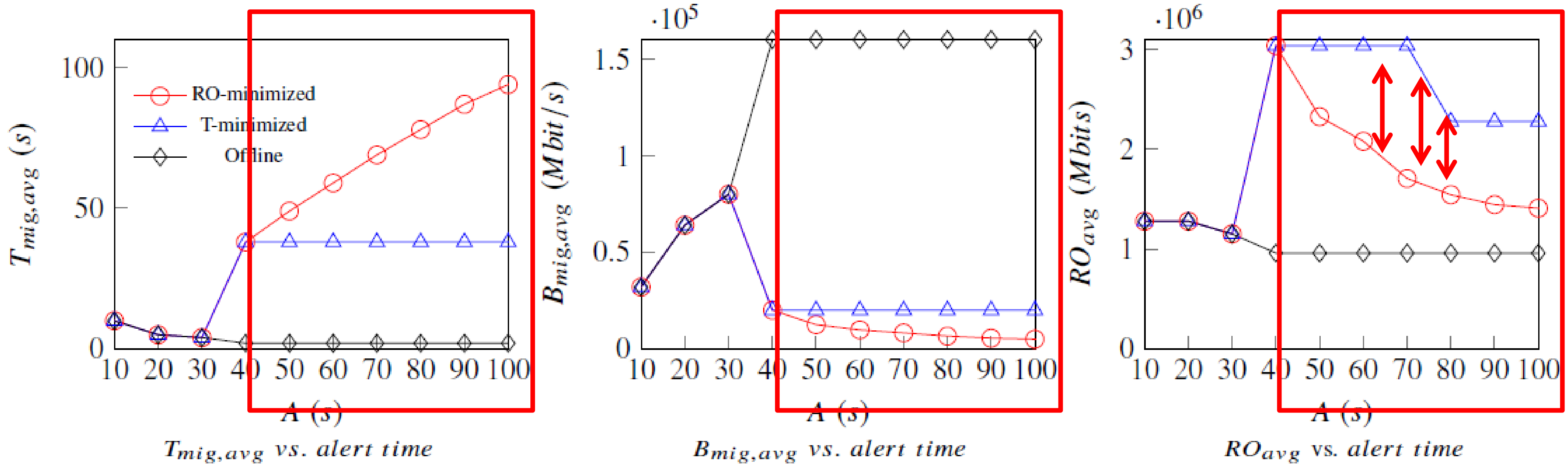


$T_{mig,avg}$ vs. alert time | $B_{mig,avg}$ vs. alert time | RO_{avg} vs. alert time

- For A = 40 sec., *RO-minimized* & *T-minimized* migrate all VMs online
  - Same behavior as only one function point can be used (maximum bandwidth value for online migration)
  - T_{mig} increases as online migration is performed
  - RO_{avg} increases to its maximum value meaning there was no room to choose efficient function-points to lower RO
- *Offline* utilizes maximum bandwidth value available ➔ min. downtime

Disaster-Resilient Online VM Migration

Results (4/4): : RO-minimized vs T-minimized vs Offline



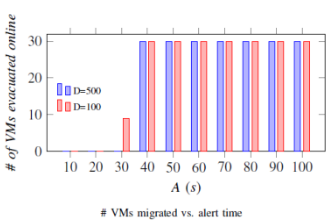
- For A ≥ 40 sec.
  - RO-minimized keeps decreasing B_{mig} → leading to higher T_{mig} but to more efficient network RO_{avg}
  - *T-minimized* continues to utilize maximum bandwidth value possible as the objective is to minimize T_{mig}

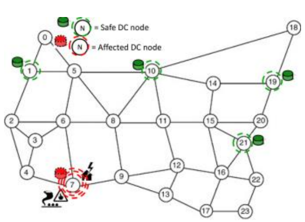
Conclusion

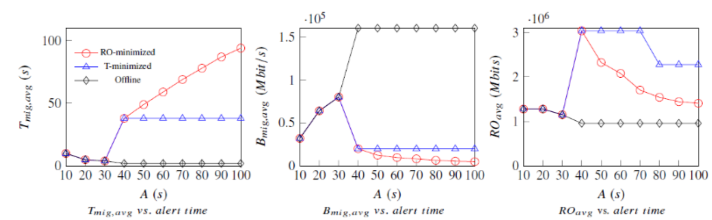
- ILP model for online VM migration for alert-based disaster resilience
  - Objective: Evacuate VMs online from risky to safe DC locations

- Thanks to our model, an DC operator can optimally decide
  - exact maximum number of VM to be migrated online
  - bandwidth to be assigned to VMs
  - destination DCs
  - routes to reach the destination DC
  - Migration scheduling

- Future work: a heuristic shall be devised for larger problem instances







Thank you!

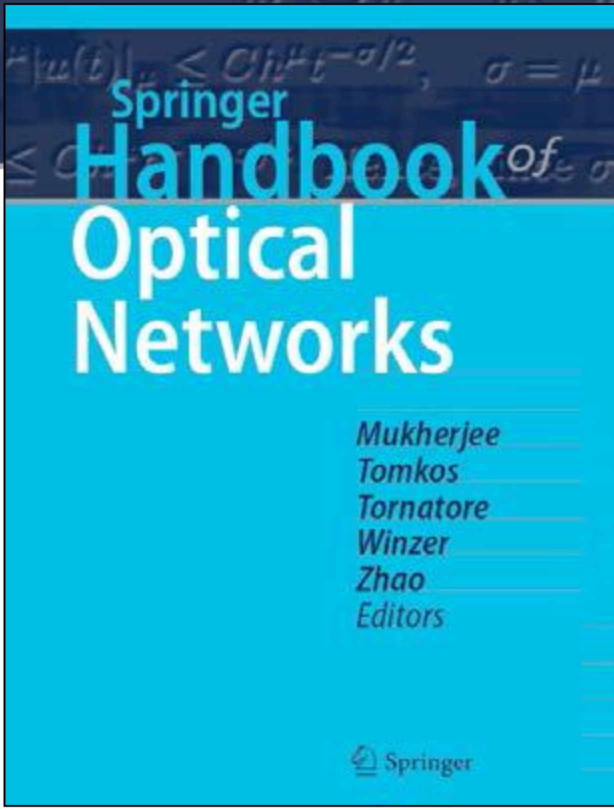
..and thanks to them!

COST
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

RECODIS
Resilient communication services protecting end-user applications from disaster-based failures

National Science Foundation
WHERE DISCOVERIES BEGIN

Disaster-Resiliency Strategies for Next-Generation Metro Optical Networks ATN: 1818972



Springer Handbook of Optical Networks

Edited by:

- **Biswanath Mukherjee**, UC Davis, USA
- **Ioannis Tomkos**, Athens Information Technology Center, Greece
- **Massimo Tornatore**, Politecnico di Milano, Italy
- **Peter Winzer**, Nokia Bell Labs, USA
- **Yongli Zhao**, Beijing University of Posts and Telecommunications, China

- **Offers a definitive reference for practitioners, researchers, and students in optical networks**
- **Represents a collective effort of over 100 top-level scientists from around the world**
- **Comprehensively treats the ever-growing field that represents the backbone of the internet**

Available in 2020!

- If
- If possible,
- If not, how many VMs can be migrated

1. Higher dirtying rate (i.e., higher user activity) affects the online migration process requiring higher bandwidth and longer migration duration

2. There exist trade-offs between minimizing the downtime and network resource occupation on one side and minimizing the VM migration duration and the network resource occupation on another side