

Sharing Cloud Networks

Lucian Popa, Gautam Kumar, **Mosharaf Chowdhury**
Arvind Krishnamurthy, Sylvia Ratnasamy, Ion Stoica



State of the Cloud

Request Instances Wizard Cancel X

CHOOSE AN AMI **INSTANCE DETAILS** CREATE KEY PAIR CONFIGURE FIREWALL REVIEW

Provide the details for your instance(s). You may also decide whether you want to launch your instances as "on-demand" or "spot" instances.

Number of Instances: Instance Type:

Type	CPU Units	CPU Cores	Memory
Micro (t1.micro)	Up to 2 ECUs	1 Core	613 MB
Large (m1.large)	4 ECUs	2 Cores	7.5 GB
Extra Large (m1.xlarge)	8 ECUs	4 Cores	15 GB
High-Memory Extra Large (m2.xlarge)	6.5 ECUs	2 Cores	17.1 GB
High-Memory Double Extra Large (m2.2xlarge)	13 ECUs	4 Cores	34.2 GB
High-Memory Quadruple Extra Large (m2.4xlarge)	26 ECUs	8 Cores	68.4 GB
High-CPU Extra Large (c1.xlarge)	20 ECUs	8 Cores	7 GB

Network???

< Back Continue ▶

Guess the Share

TCP Per flow

3 : 1

Per Source

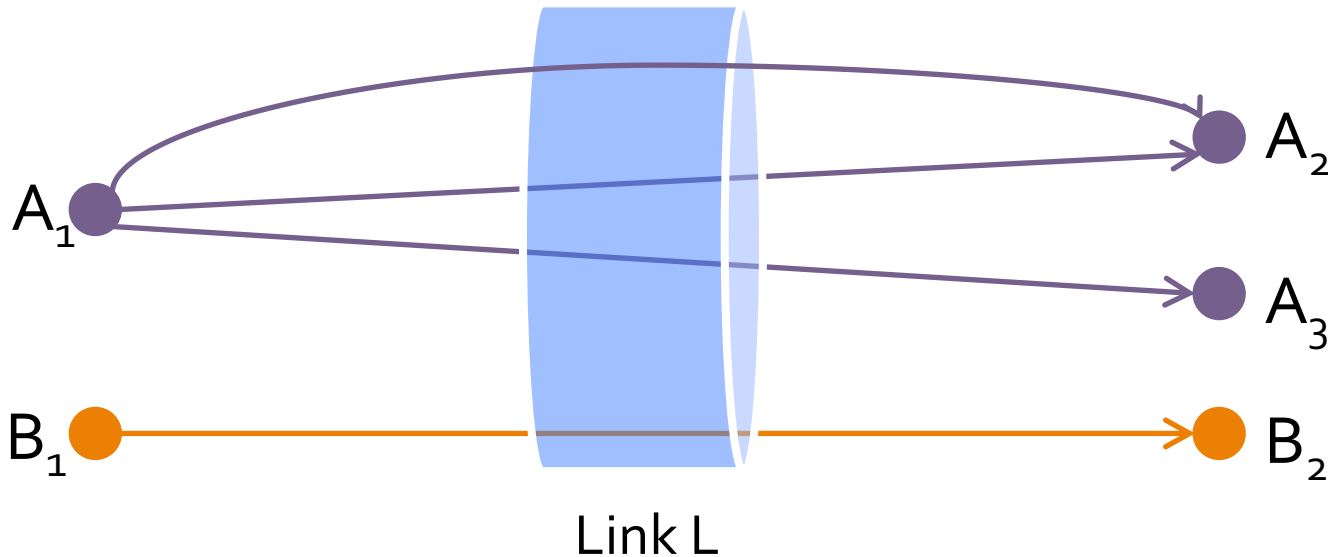
1 : 1

Per Destination

2 : 1

Per-VM
Proportional

3 : 2



Alice : Bob = ? : ?

Challenges

Network share of a virtual machine (VM) V depends on

- » Collocated VMs,
- » Placement of destination VMs, and
- » Cross-traffic on each link used by V

Network differs from CPU or RAM

- » Distributed resource
- » Usage attribution (source, destination, or both?)

Traditional link sharing concepts needs rethinking

Requirements

Min
Bandwidth
Guarantee

Introduce
performance
predictability

Aggregate
Proportionality

Network shares
proportional to
the number of
VMs

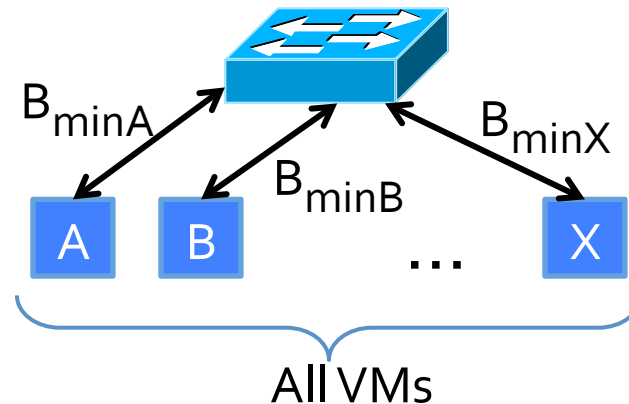
High
Utilization

Do not leave
bandwidth
unused if there
is demand

Requirement 1: Guaranteed Minimum B/W

Provides a minimum b/w guarantee for each VM

Captures the desire of tenants to get performance isolation for their applications



Requirement 2: Aggregate Proportionality

Shares network resources across tenants in proportion to the number of their VMs

Captures payment-proportionality

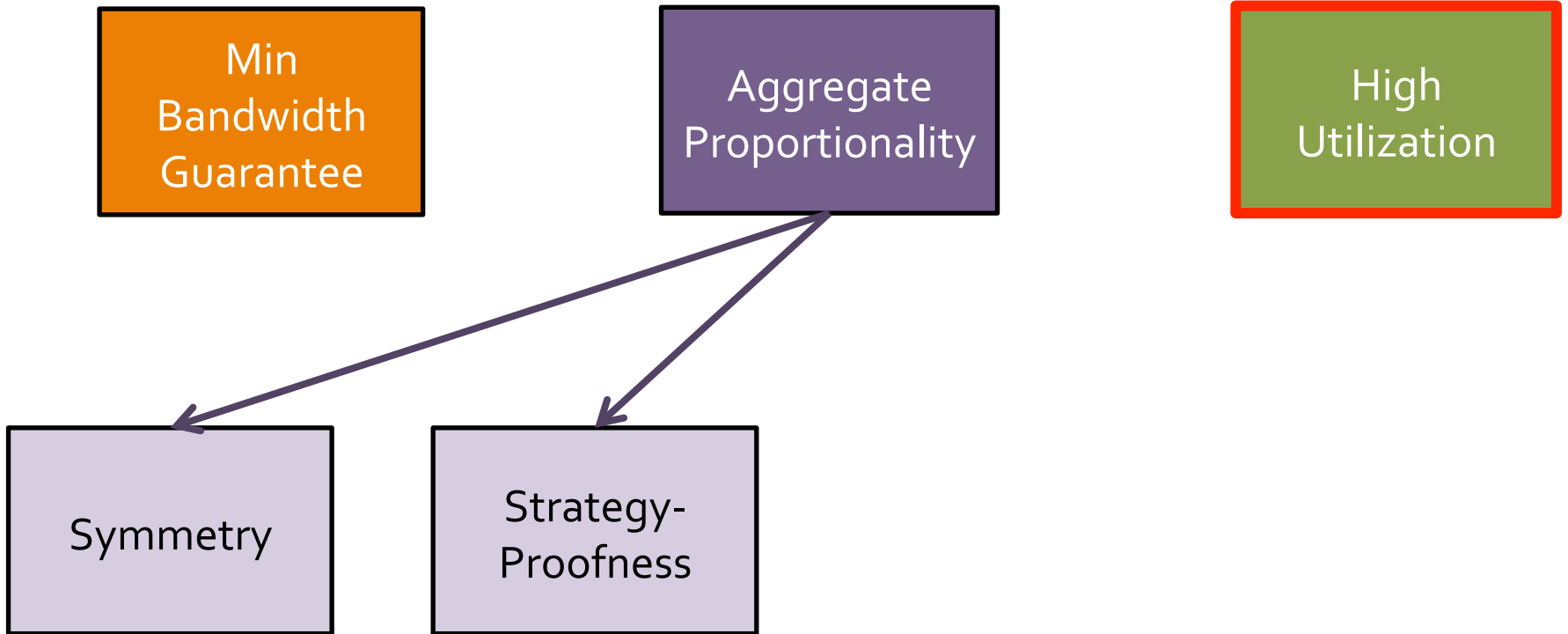
» Similar to other resources like CPU, RAM etc.

Desirable properties

» **Strategy-proofness:** Allocations cannot be gamed

» **Symmetry:** Reversing directions of flows does not change allocation

Design Space



Requirement 3: High Utilization

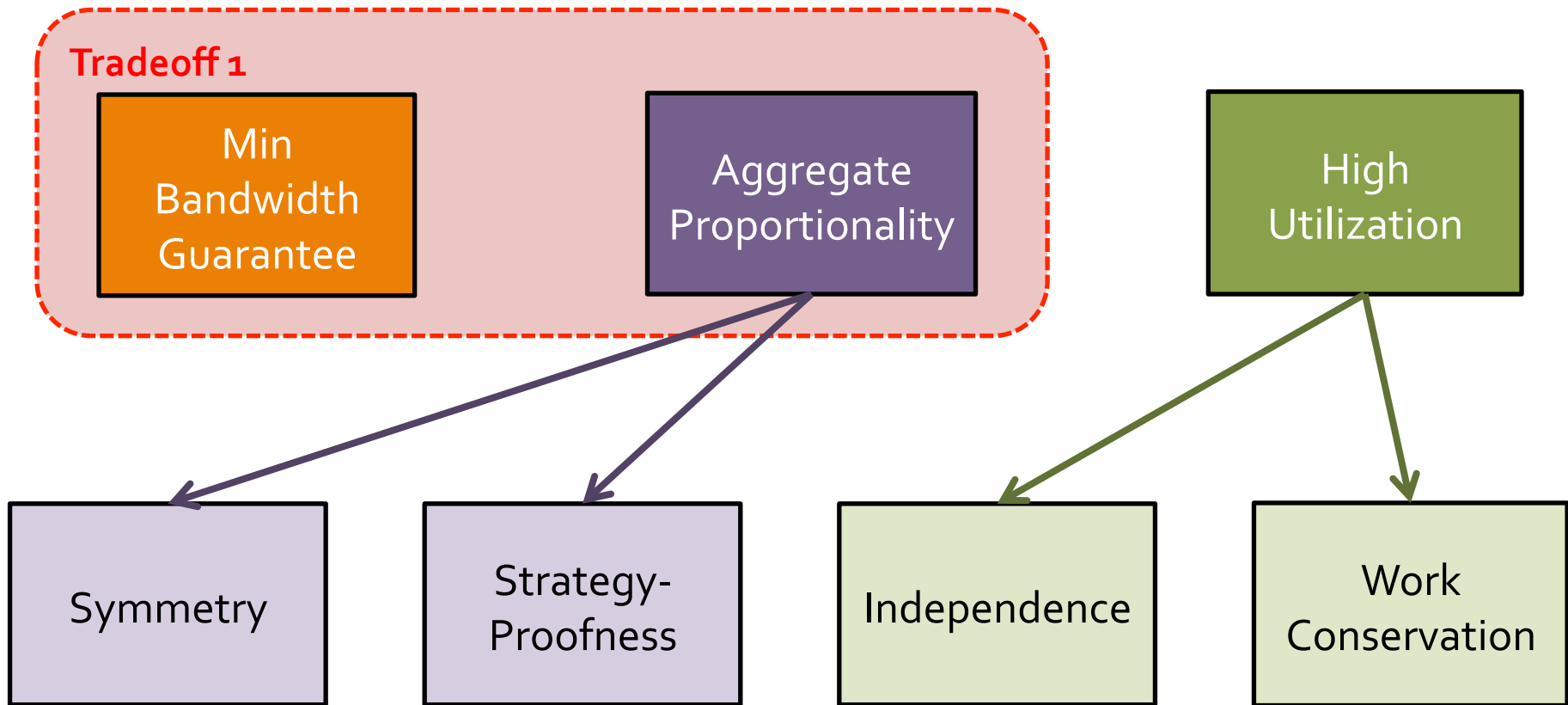
Provides incentives such that throughput is only constrained by the network capacity

- » Not by the inefficiency of the allocation or by disincentivizing users to send traffic

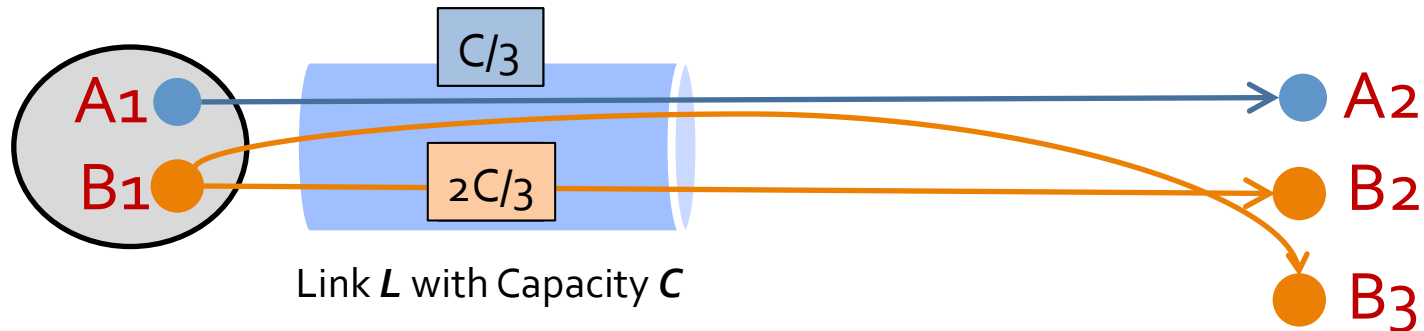
Desirable properties

- » **Work Conservation:** Full utilization of bottleneck links
- » **Independence:** Independent allocation of one VM's traffic across independent paths

Design Space

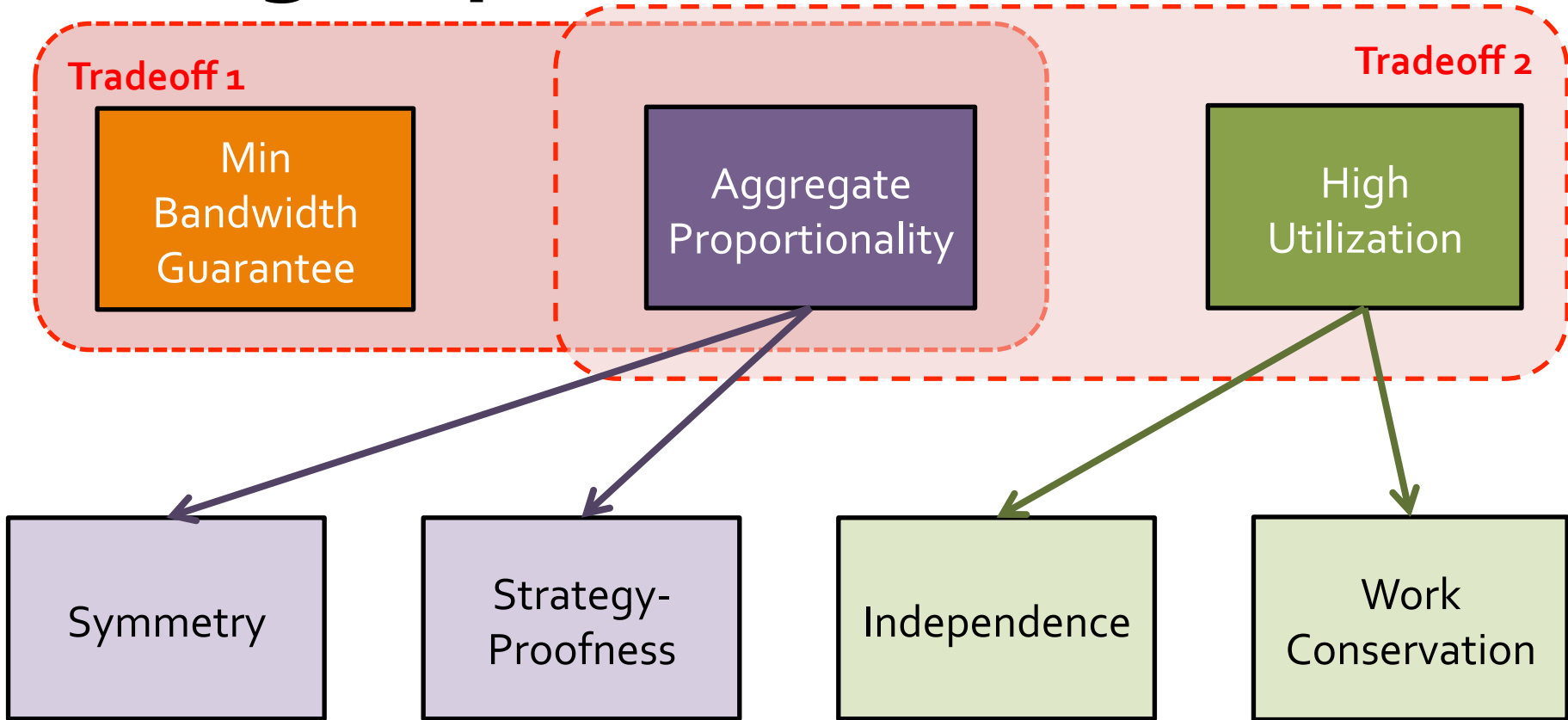


Tradeoff 1: Min B/W vs. Proportionality

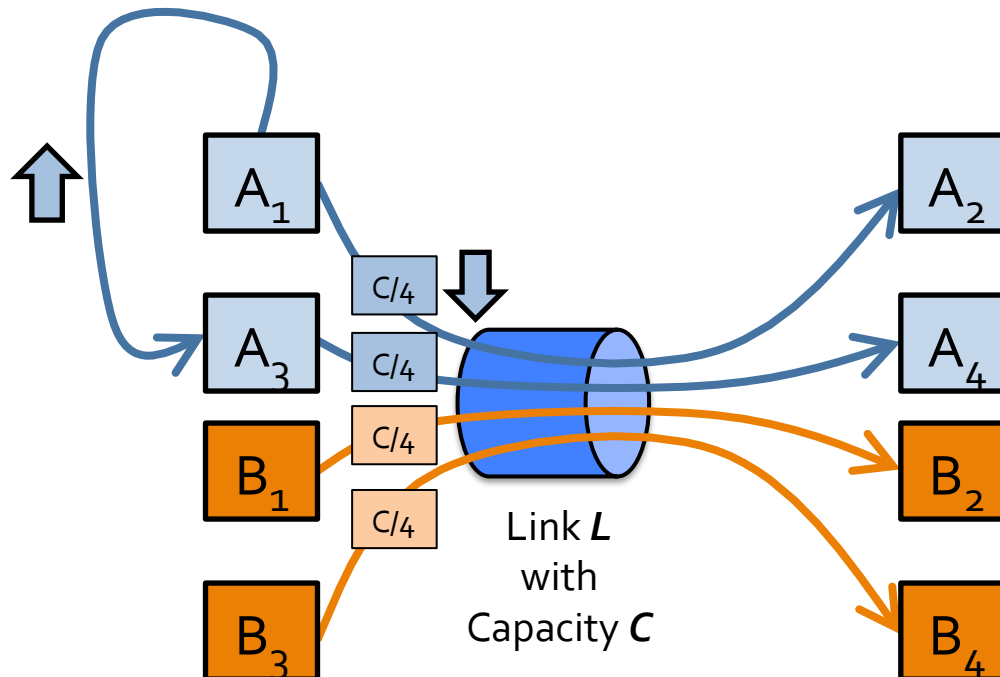


Share of Tenant A can decrease arbitrarily!

Design Space



Tradeoff 2: Proportionality vs. Utilization



To maintain proportionality, equal amount of traffic must be moved from A_1 - A_2 to A_1 - A_3 => **Underutilization of A_1 - A_3**

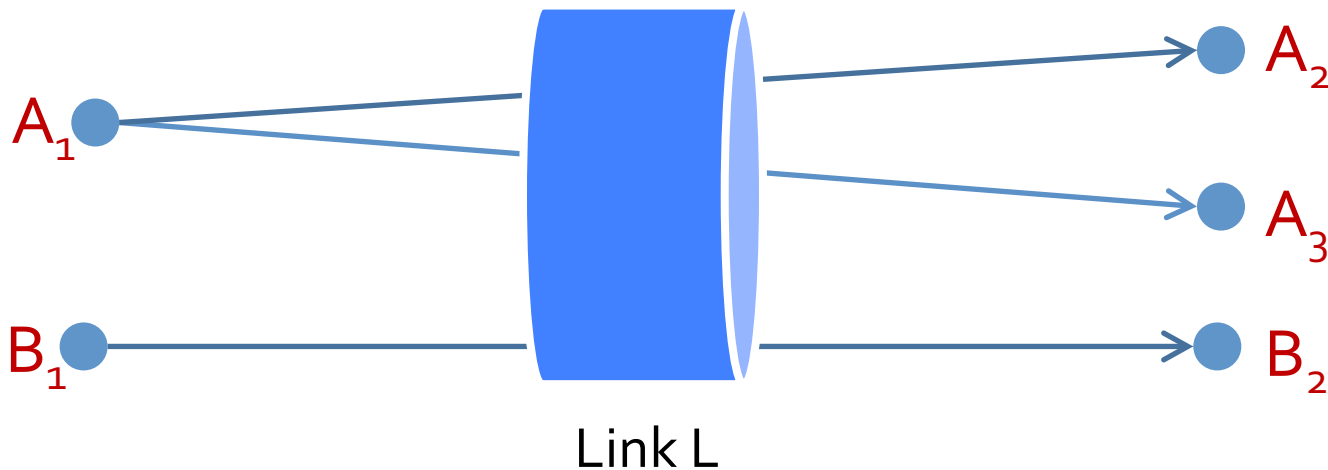
Per-link Proportionality

Restrict to congested links only

Share of a tenant on a congested link is proportional to the number of its VMs sending traffic on that link

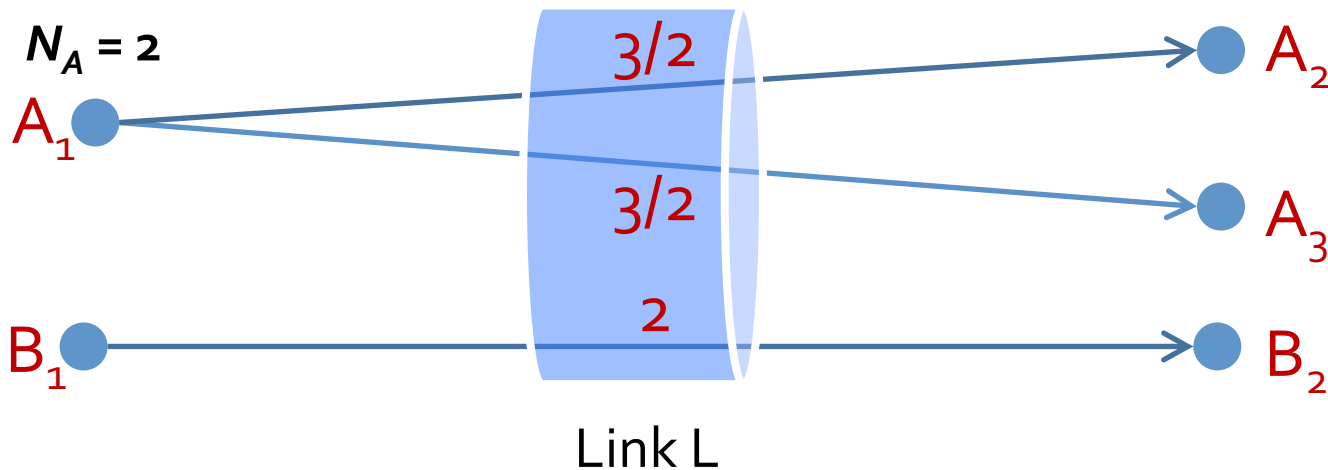
Per Endpoint Sharing (PES)

Five identical VMs (with unit weights) sharing a Link L



Per Endpoint Sharing (PES)

Resulting weights of the three flows:



To generalize, weight of a flow A-B on link L is $W_{A-B} = \frac{W_A}{N_A} + \frac{W_B}{N_B}$.

Per Endpoint Sharing (PES)

Symmetric

$$W_{A-B} = \frac{W_A}{N_A} + \frac{W_B}{N_B} = W_{B-A}$$

Proportional

» sum of weights of flows of a tenant on a link L =
sum of weights of its VMs communicating on that link

Work Conserving

Independent

Strategy-proof on congested links

Generalized PES



Scale weight of A by α

Scale weight of B by β

$$W_{A-B} = W_{B-A} = \alpha \frac{W_A}{N_A} + \beta \frac{W_B}{N_B}$$

$\alpha > \beta$ if L is more important to A than to B (e.g., A's access link)

One-Sided PES (OSPES)



Scale weight of A by α

Scale weight of B by β

$$W_{A-B} = W_{B-A} = \alpha \frac{W_A}{N_A} + \beta \frac{W_B}{N_B}$$

$$\alpha = 1, \beta = 0$$
$$\alpha = 0, \beta = 1$$

Highest B/W Guarantee

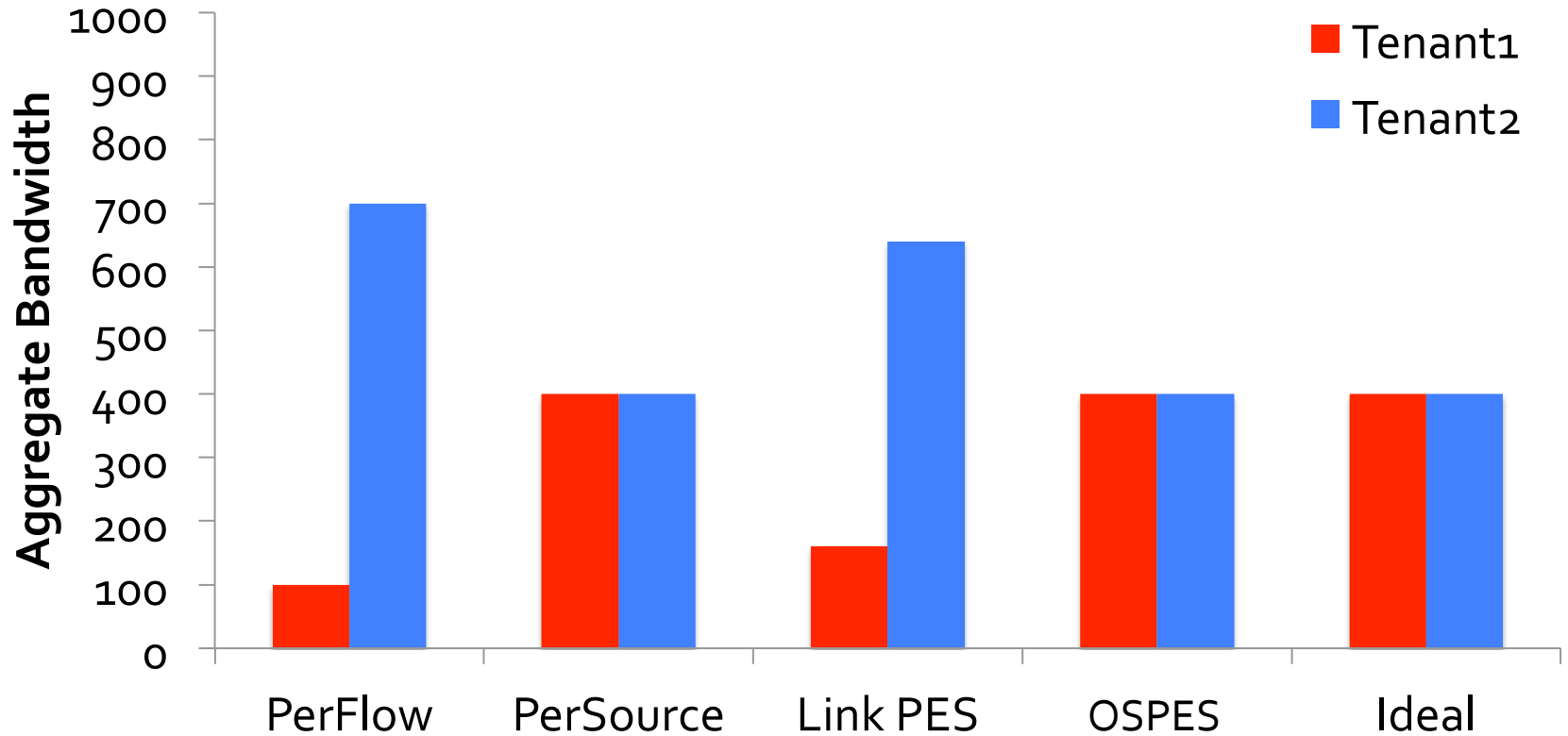
*In the Hose Model

to source
to destination

Comparison

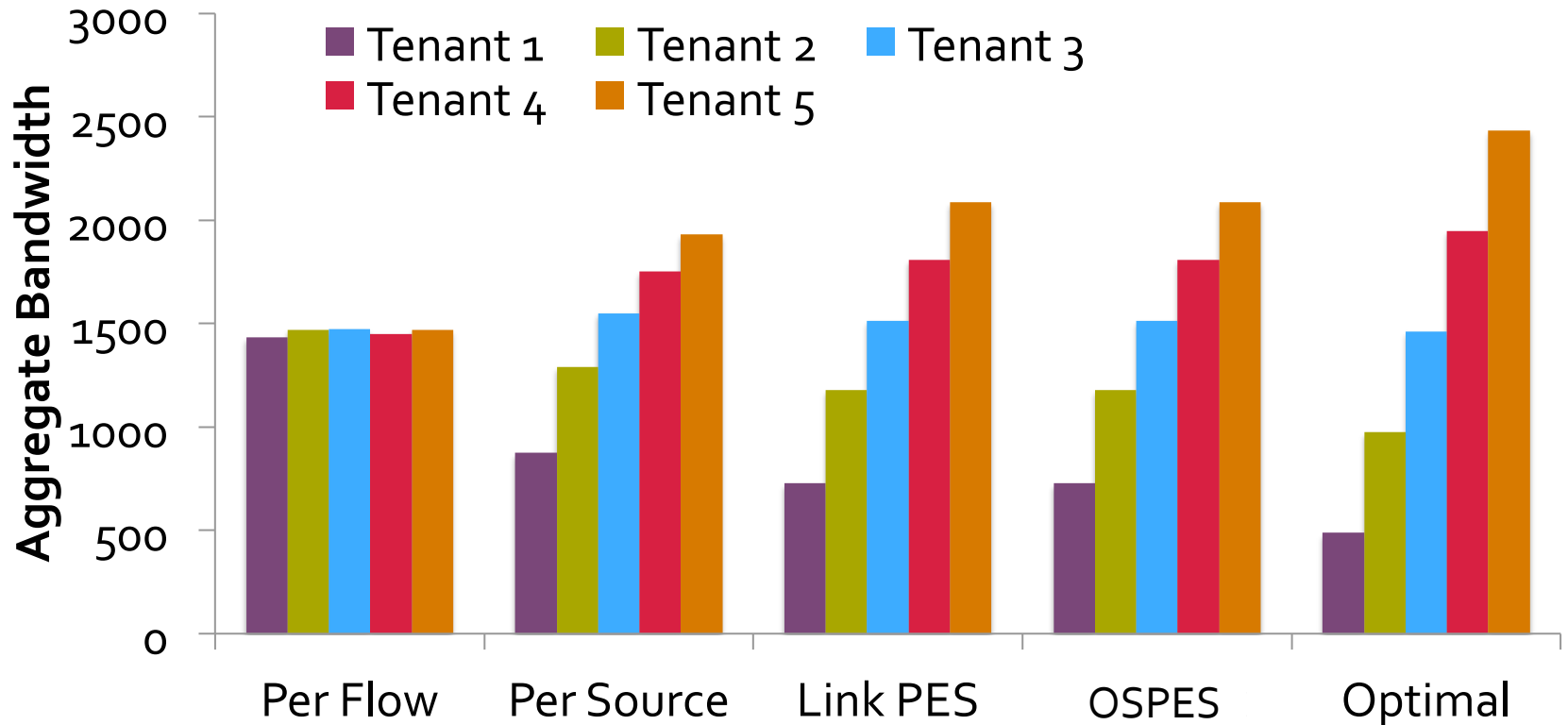
	Per Flow	Per Source	Static Reservation	Link PES	OSPES
Link Proportionality	☹️	☹️	😊	😊	☹️
▪ Symmetry	✓	✗	✓	✓	✓
▪ Strategy-Proofness	✗	✓	✓	✓	✓
Utilization	😊	😊	☹️	😊	😊
▪ Independence	✓	✓	✓	✓	✓
▪ Work Conservation	✓	✓	✗	✓	✓
B/W Guarantee	☹️	☹️	😊	☹️	☹️

Full-bisection B/W Network



Tenant 1 has one-to-one communication pattern
Tenant 2 has all-to-all communication pattern

MapReduce Workload



$$W_1:W_2:W_3:W_4:W_5 = 1:2:3:4:5$$

Summary

Sharing cloud networks is all about making tradeoffs

- » Min b/w guarantee **VS** Proportionality
- » Proportionality **VS** Utilization

Desired solution is not obvious

- » Depends on several conflicting requirements and properties
- » Influenced by the end goal