



Ginseng: Market-Driven LLC Allocation

Liran Funaro, Orna Agmon Ben-Yehuda, Assaf Schuster
 {funaro,ladypine,assaf}@cs.technion.ac.il USENIX ATC '16



What is the Problem?

How should cache ways be allocated among virtual machines to optimize the aggregate benefit?

Resource-as-a-Service

- Fine resource granularity
- Fine time granularity
- Market-driven resource pricing

Cache Layout Reminder

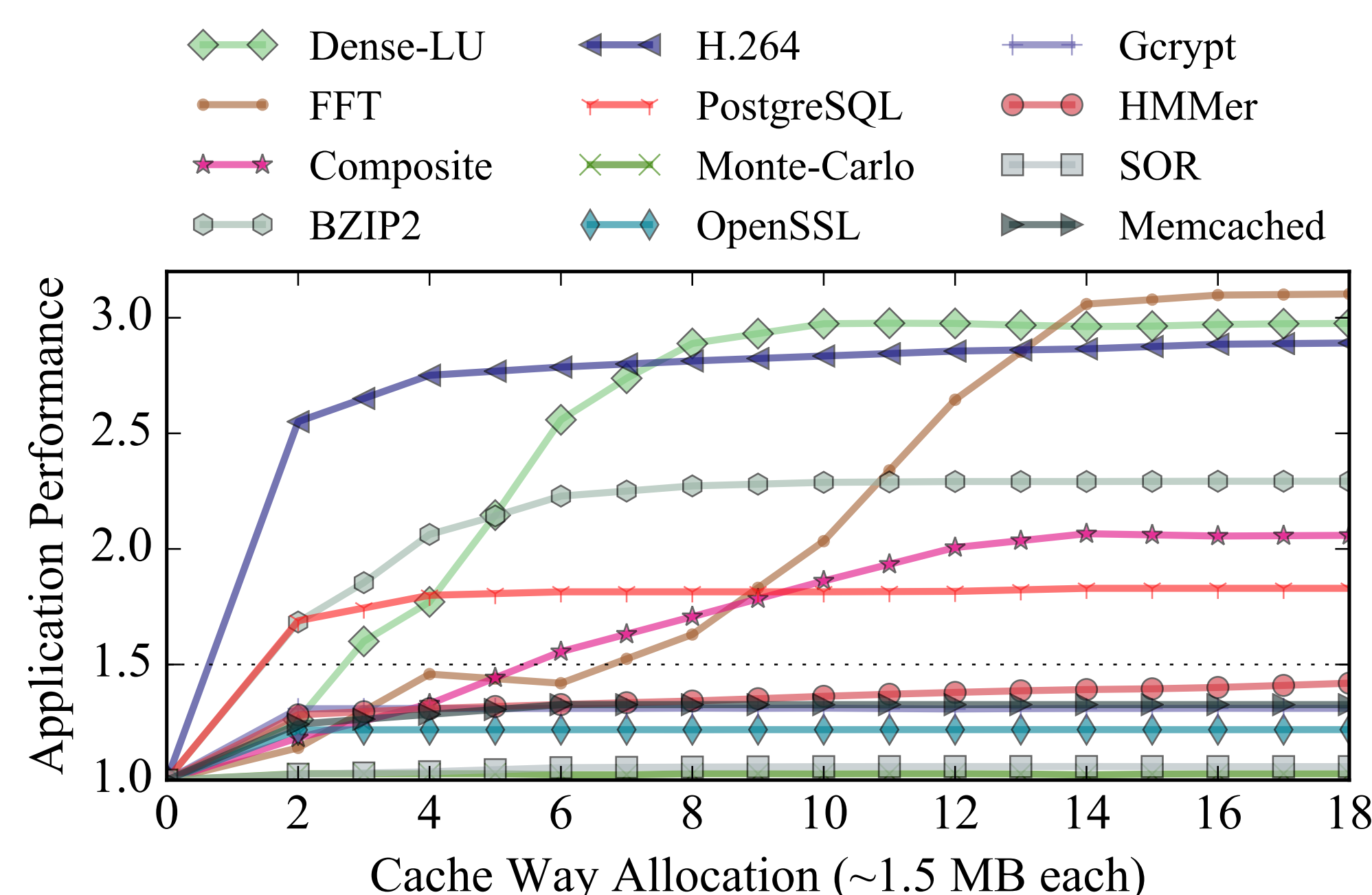
	Way 1	Way 2	Way 3	Way 4
Set 1	line	line	line	line
Set 2	line	line	line	line
⋮	⋮	⋮	⋮	⋮
Set n	line	line	line	line

Cache Allocation Technology

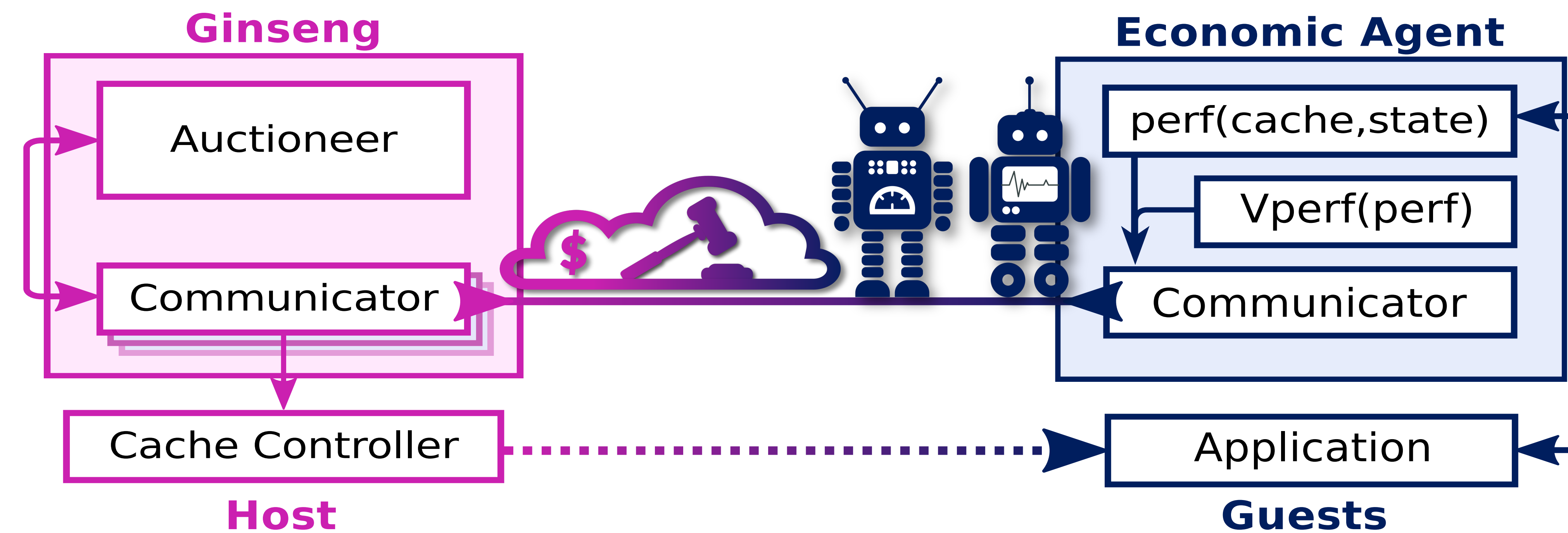
	Way 1	Way 2	Way 3	Way 4
COS 1				
COS 2				
COS 3				
COS 4				

A thread that is assigned to a COS is only allowed to store new cache lines in the ways determined by the COS. There is no reading limitation.

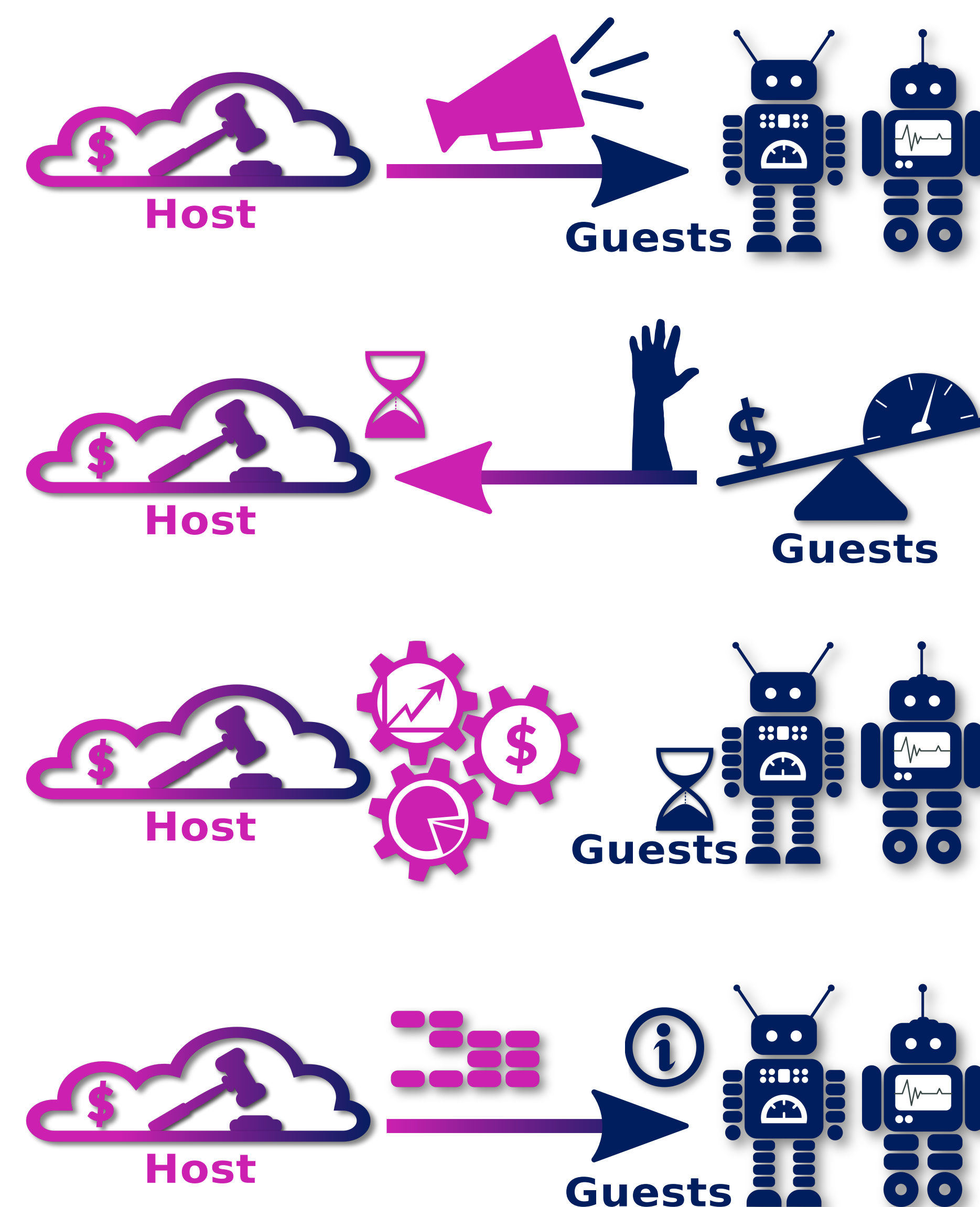
Cache Utilizers vs. Neutrals



System Architecture



Auction Protocol



The host announces an auction **every 10 seconds**

Each guest bids with a valuation for each quantity of cache ways

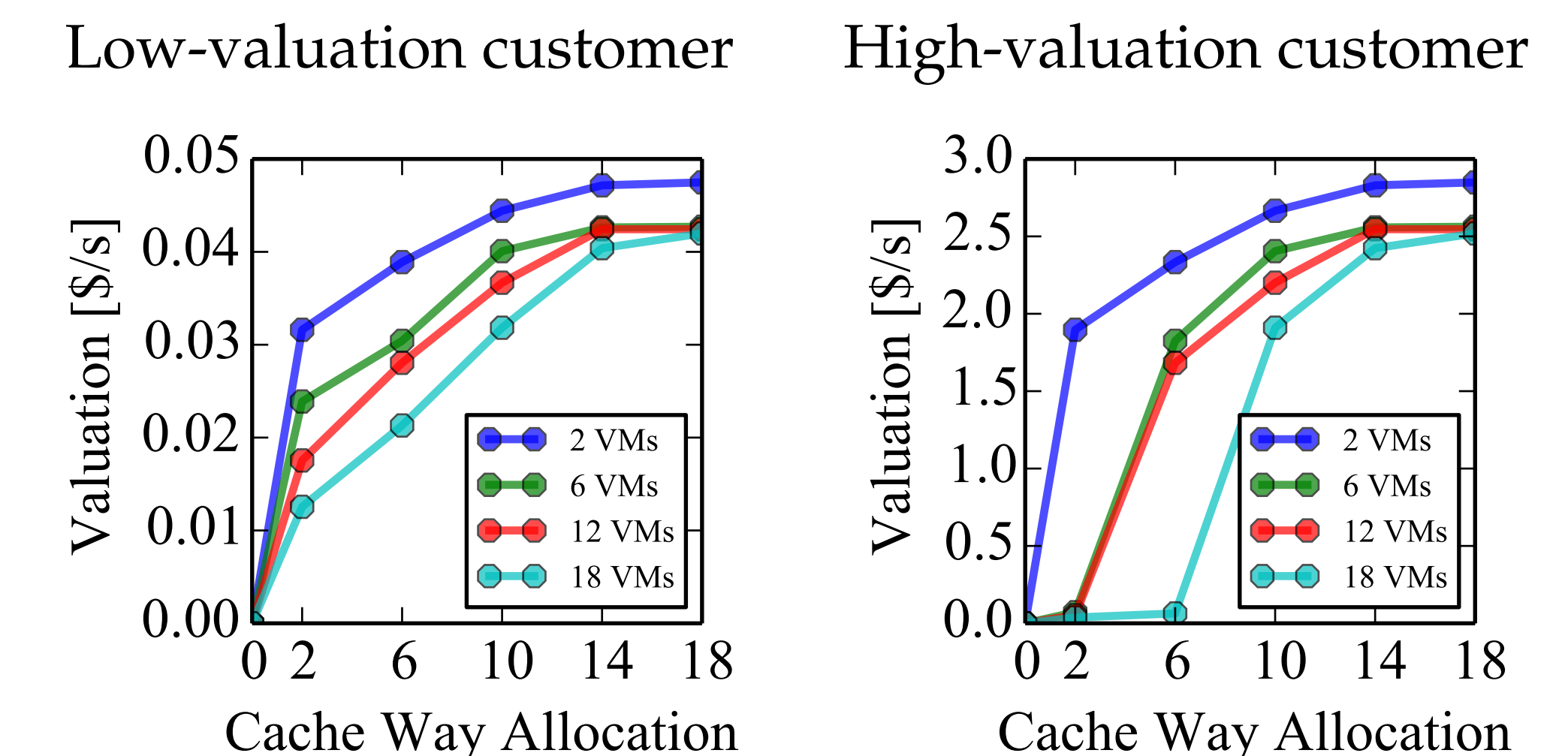
The auctioneer finds the allocation that maximizes the **social welfare** (sum of guest valuations)

The auctioneer informs the guests of their allocation and charges them according to the **exclusion-compensation principle**

Conclusions

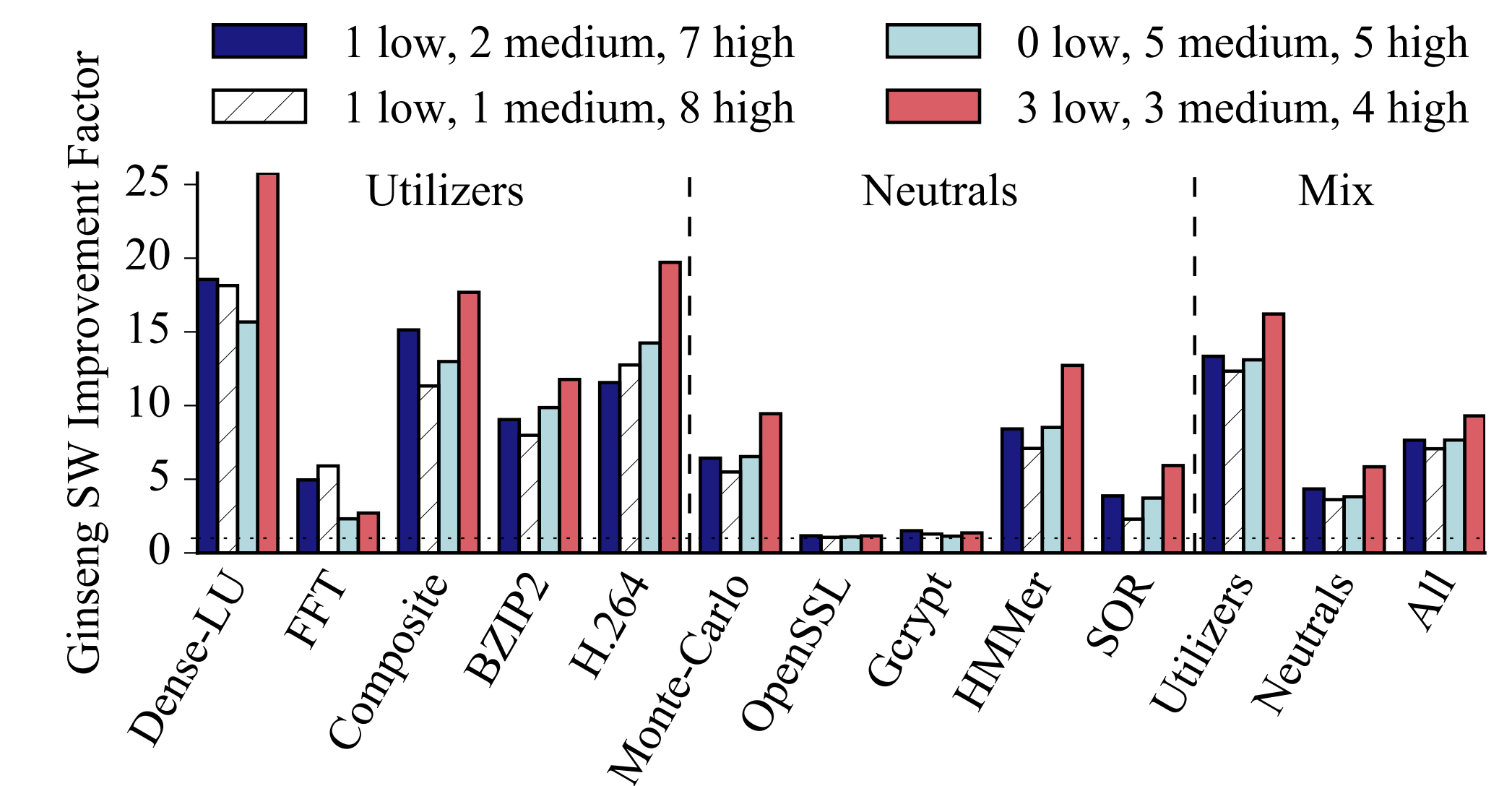
- *Ginseng* efficiently allocates cache to selfish guests while maximizing their social welfare.
- The guests utilize their cache fast enough to allow such rapid changes in the allocation without any substantial effect on their performance.

Valuations



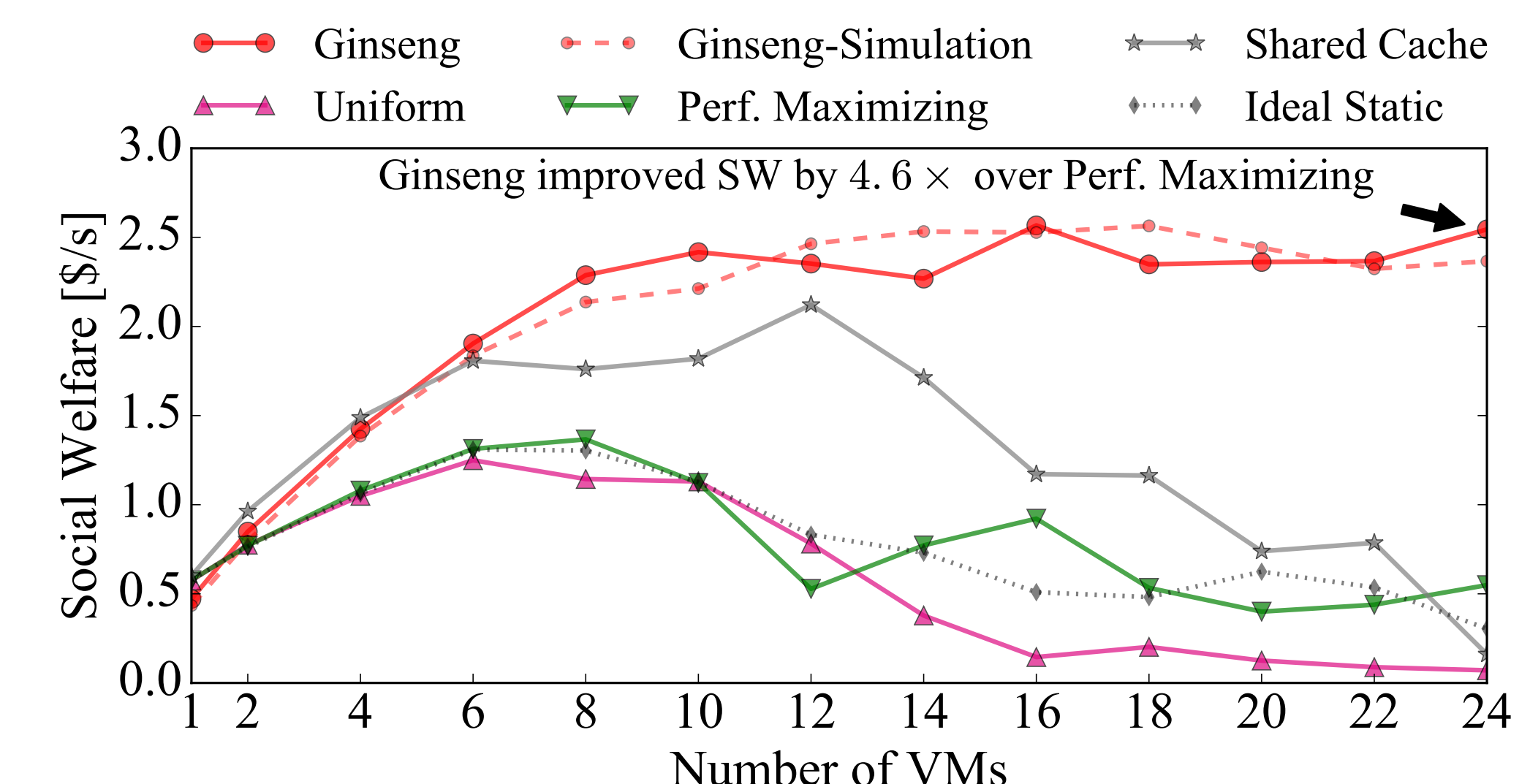
Composite-Scimark valuation function for different server-loads (i.e., active VMs) and when allocated 2 hardware threads.

Improvement over Perf. Max.



Maximum improvement factor of *Ginseng* compared to the performance-maximizing method with different assumptions on the number of high, medium, and low valuation customers. The maximum is over any number of guests with the application, or mixture of applications.

Growing Number of VMs



All guests run *Fast Fourier Transform* with 1 high-valuation customer, 1 medium-valuation customers and 8 low-valuation customers.