SCC:
Cluster Storage Provisioning Informed by Application Characteristics and SLAs

Harsha V. Madhyastha*, **John C. McCullough**, George Porter,
Rishi Kapoor, Stefan Savage, Alex C. Snoeren, and Amin Vahdat
UC Riverside* and UC San Diego

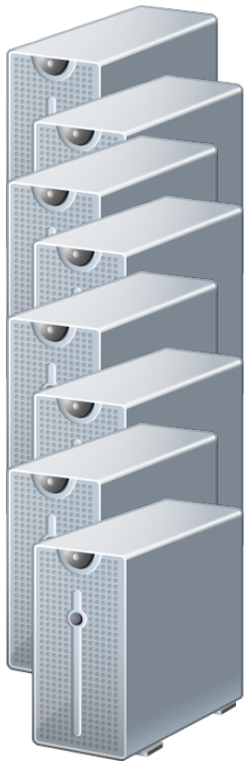
Provisioning Hardware for Cluster Applications

Provisioning Hardware for Cluster Applications

"Need 100 Queries/s"

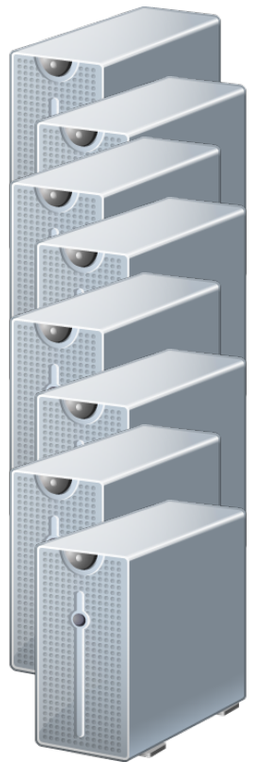
Provisioning Hardware for Cluster Applications

"Need 100 Queries/s"



Provisioning Hardware for Cluster Applications

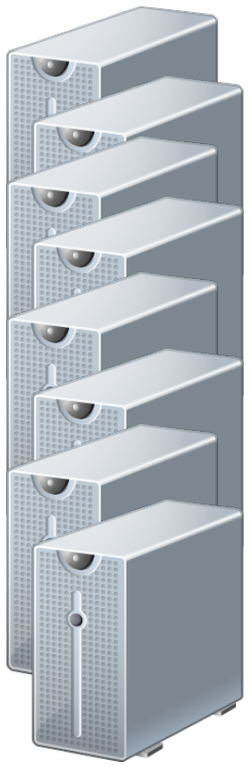
"Need 100 Queries/s"



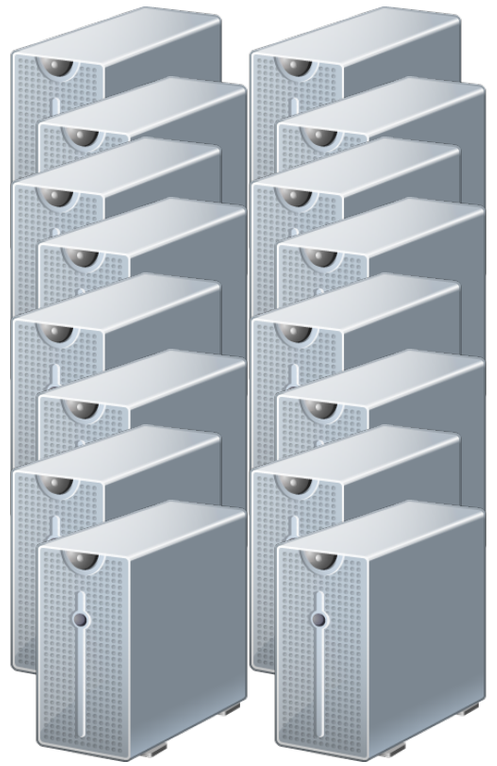
"Need 1000 Views/s"

Provisioning Hardware for Cluster Applications

"Need 100 Queries/s"



"Need 1000 Views/s"



Goals for Provisioning at Low-cost

- High performance
- Redundancy
- Supporting multi-tenancy
- High availability
- ...

Our focus: reach performance goal and minimize cost for a single application, emphasizing storage

Challenge: Large Configuration Space

- Diverse server enclosures/architectures
- Diverse storage options

	Size	MB/s (r/w)	IOPS	Cost
7.2k-rpm	500GB	90/90	125/125	$213
15k-rpm	146GB	150/150	285/285	$296
SSD	32GB	250/80	2500/1000	$496
DRAM	1GB	13k/13k	1.6B/1.6B	$36

Challenge: Large Configuration Space

- Diverse server enclosures/architectures
- Diverse storage options

	Size	MB/s (r/w)	IOPS	Cost
7.2k-rpm	500GB	90/90	125/125	$213
15k-rpm	146GB	150/150	285/285	$296
SSD	32GB	250/80	2500/1000	$496
DRAM	1GB	13k/13k	1.6B/1.6B	$36

- Current state-of-the-art:

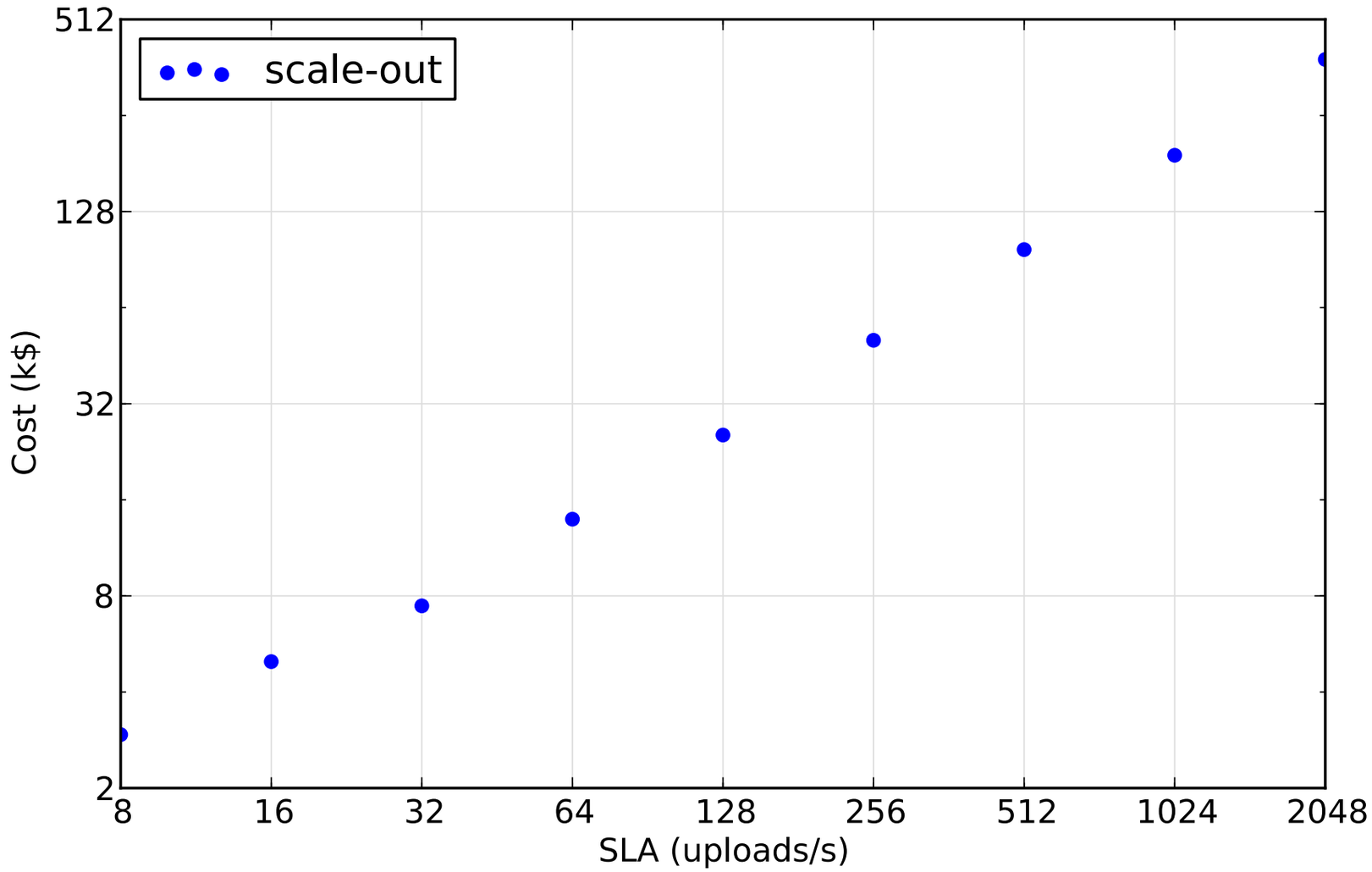
Challenge: Large Configuration Space

- Diverse server enclosures/architectures
- Diverse storage options

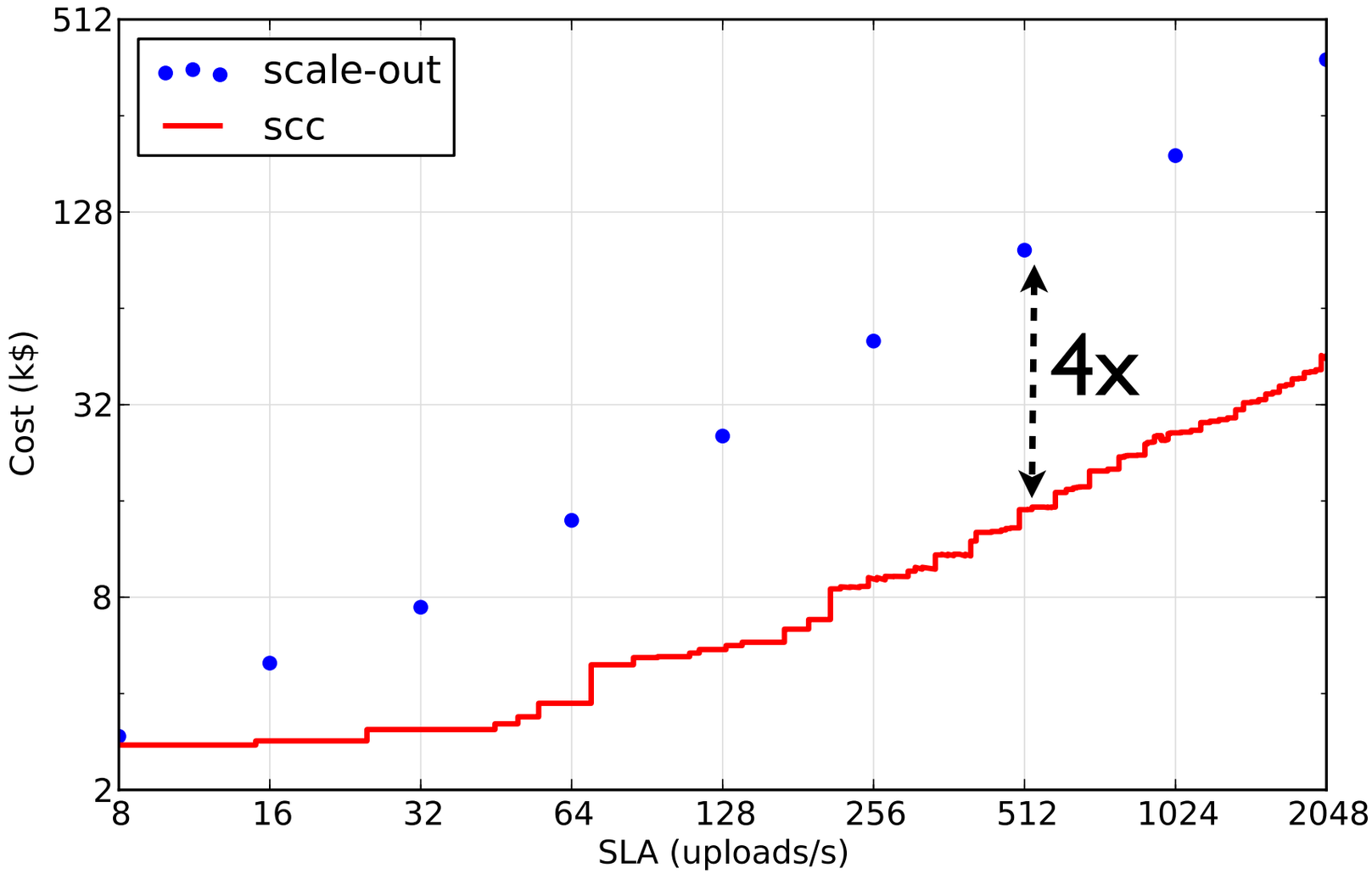
	Size	MB/s (r/w)	IOPS	Cost
7.2k-rpm	500GB	90/90	125/125	$213
15k-rpm	146GB	150/150	285/285	$296
SSD	32GB	250/80	2500/1000	$496
DRAM	1GB	13k/13k	1.6B/1.6B	$36

- Current state-of-the-art:
  - Apply **rules-of-thumb** from experience
  - **Trial-and-error** with various configurations
  - Configuration duplicated to scale-out

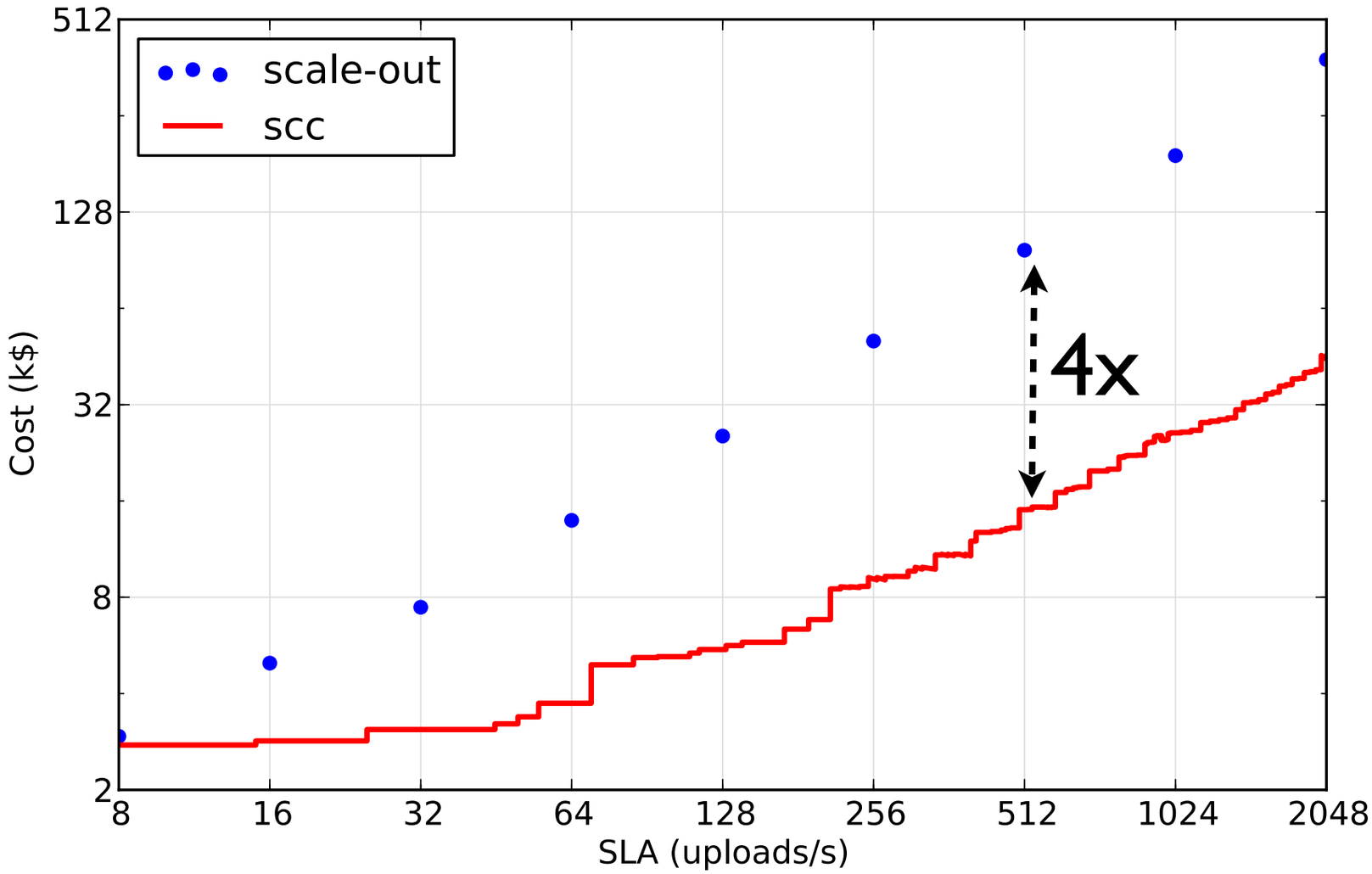
Scale-out Shortcomings for Photo-sharing Application



Scale-out Shortcomings for Photo-sharing Application



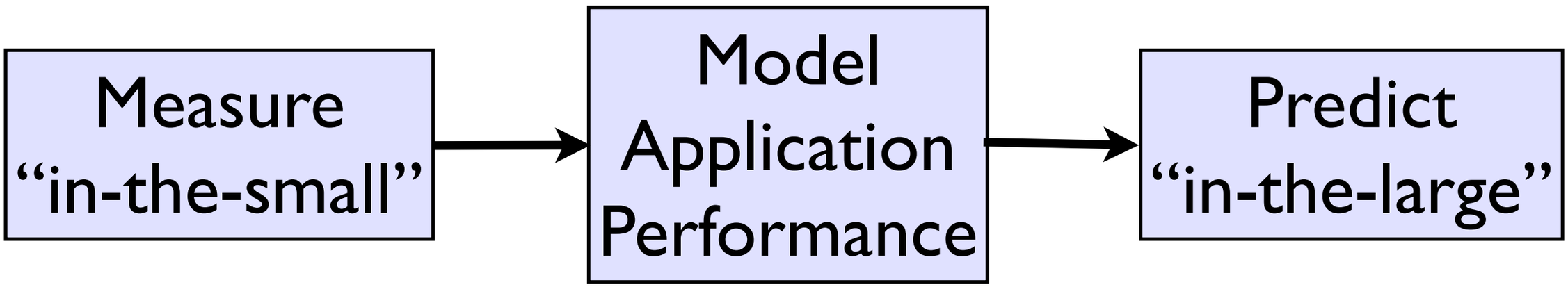
Scale-out Shortcomings for Photo-sharing Application



Insight: Match storage to workload, leverage heterogeneity

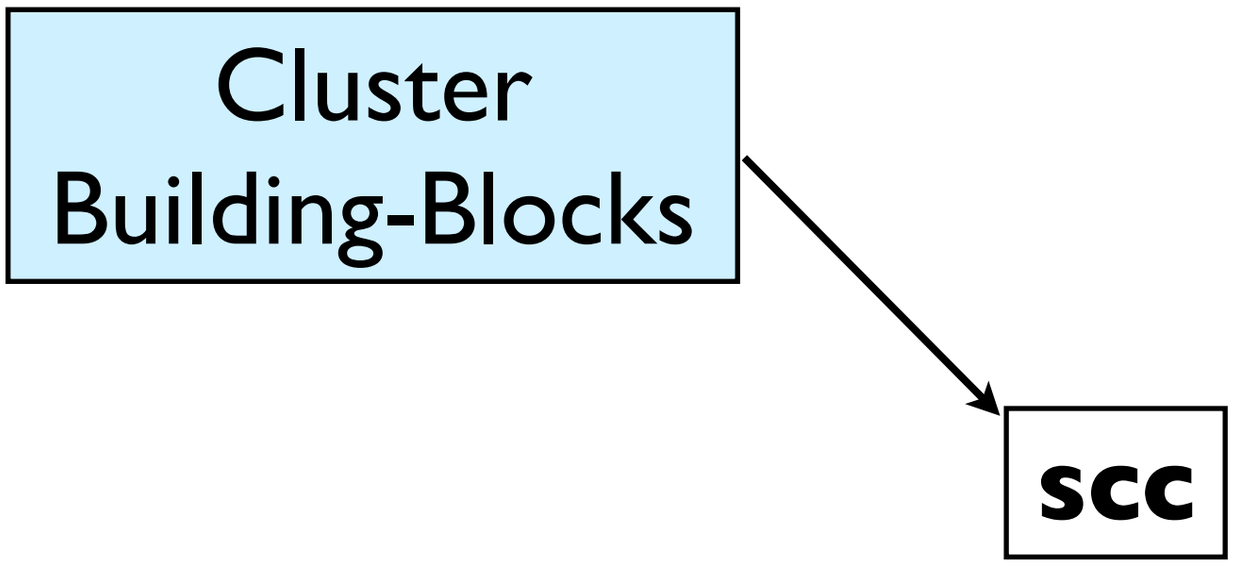
Goal: Understand Configuration Space

- What is a low-cost configuration now?
- What will low-cost configurations look like in the future?

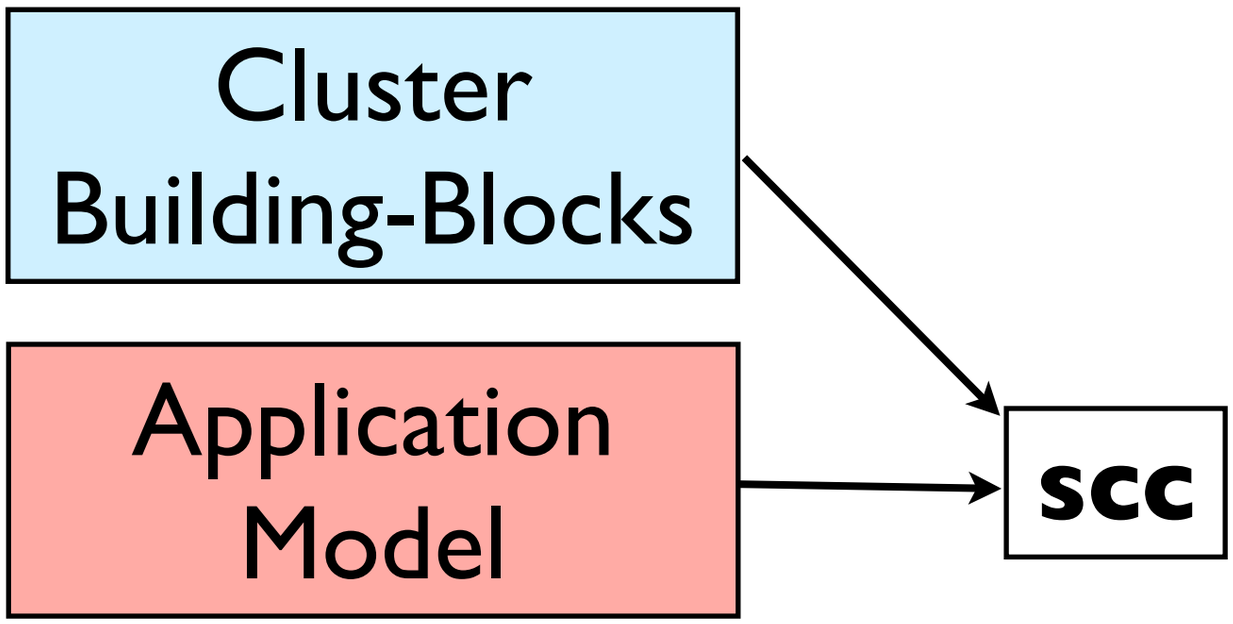


scc: Storage Configuration Compiler

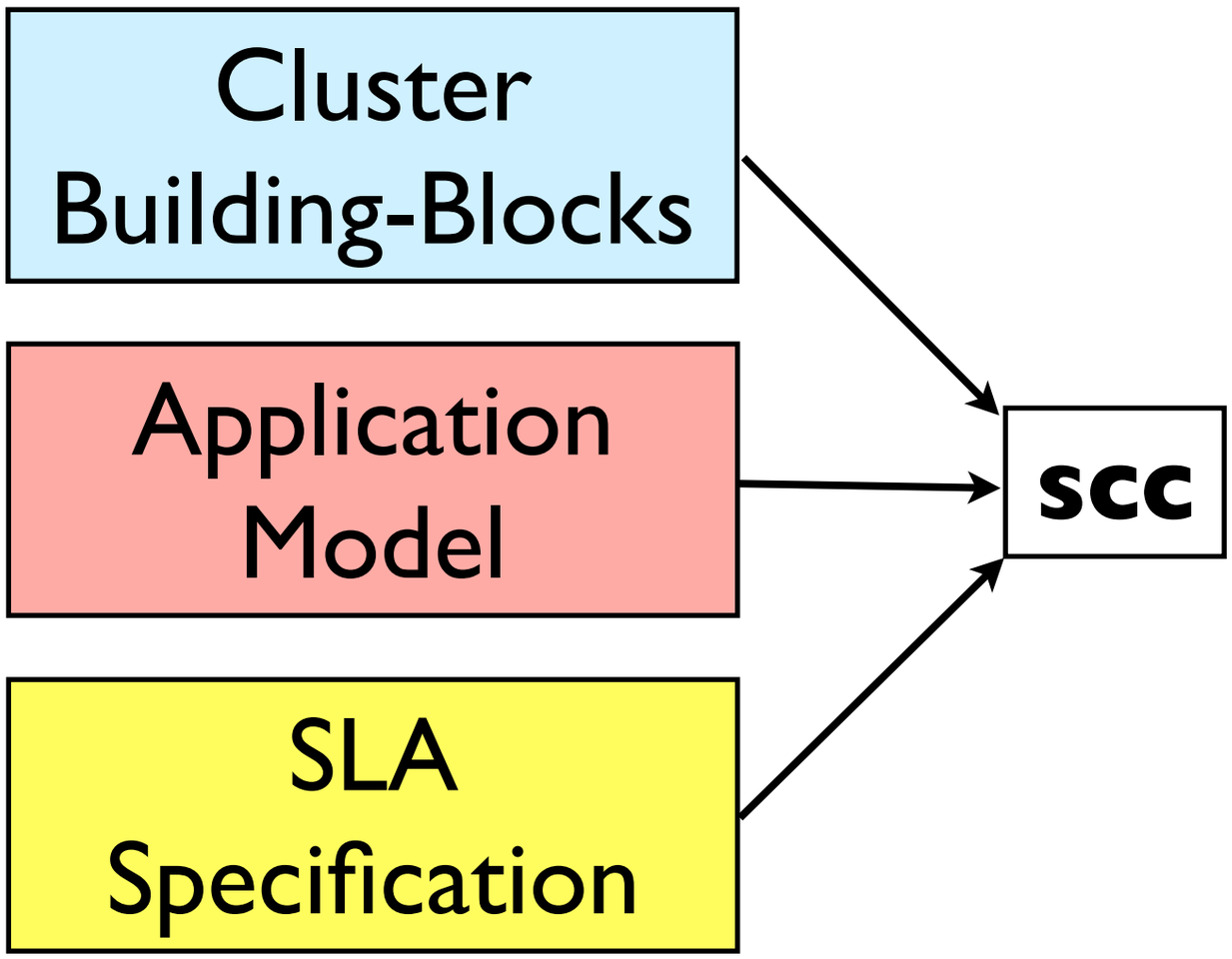
scc: Storage Configuration Compiler



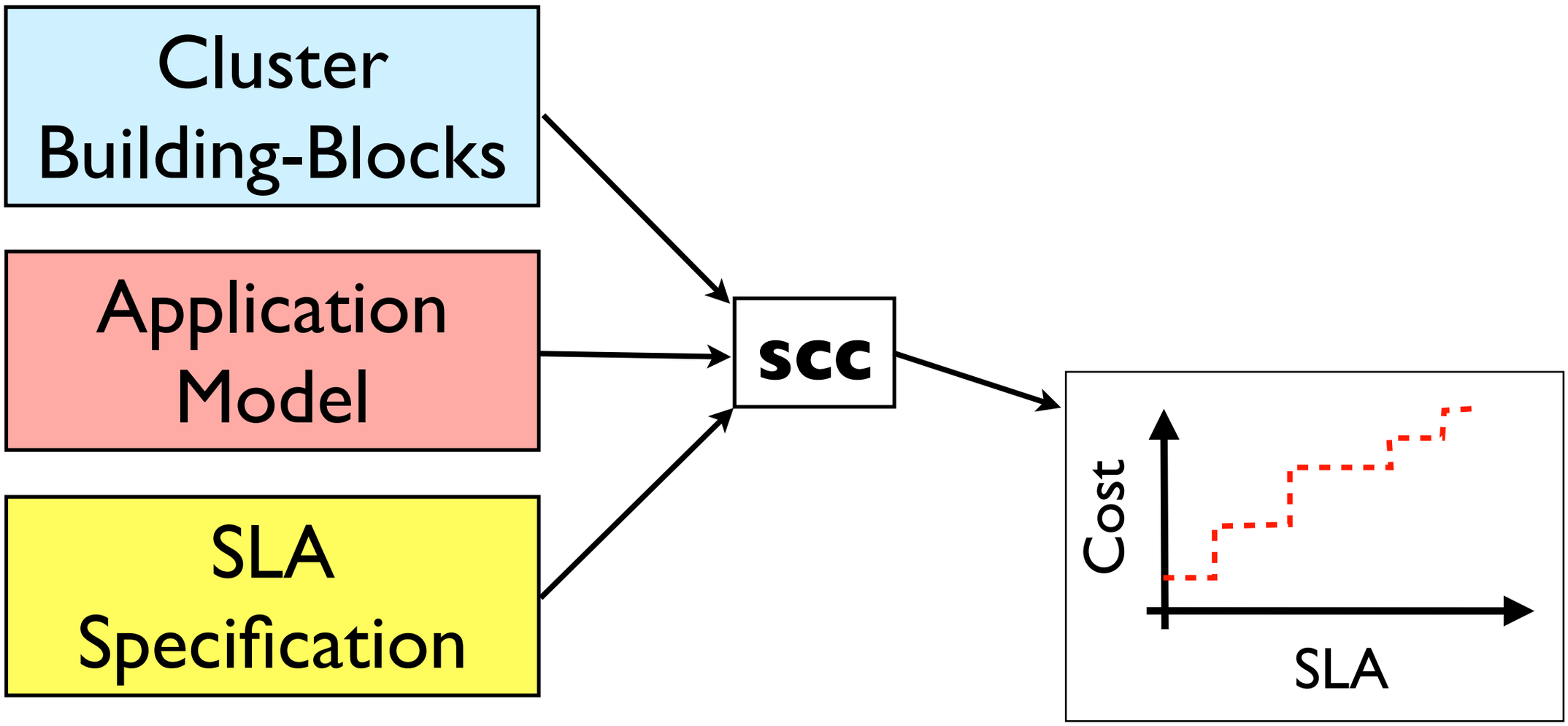
scc: Storage Configuration Compiler



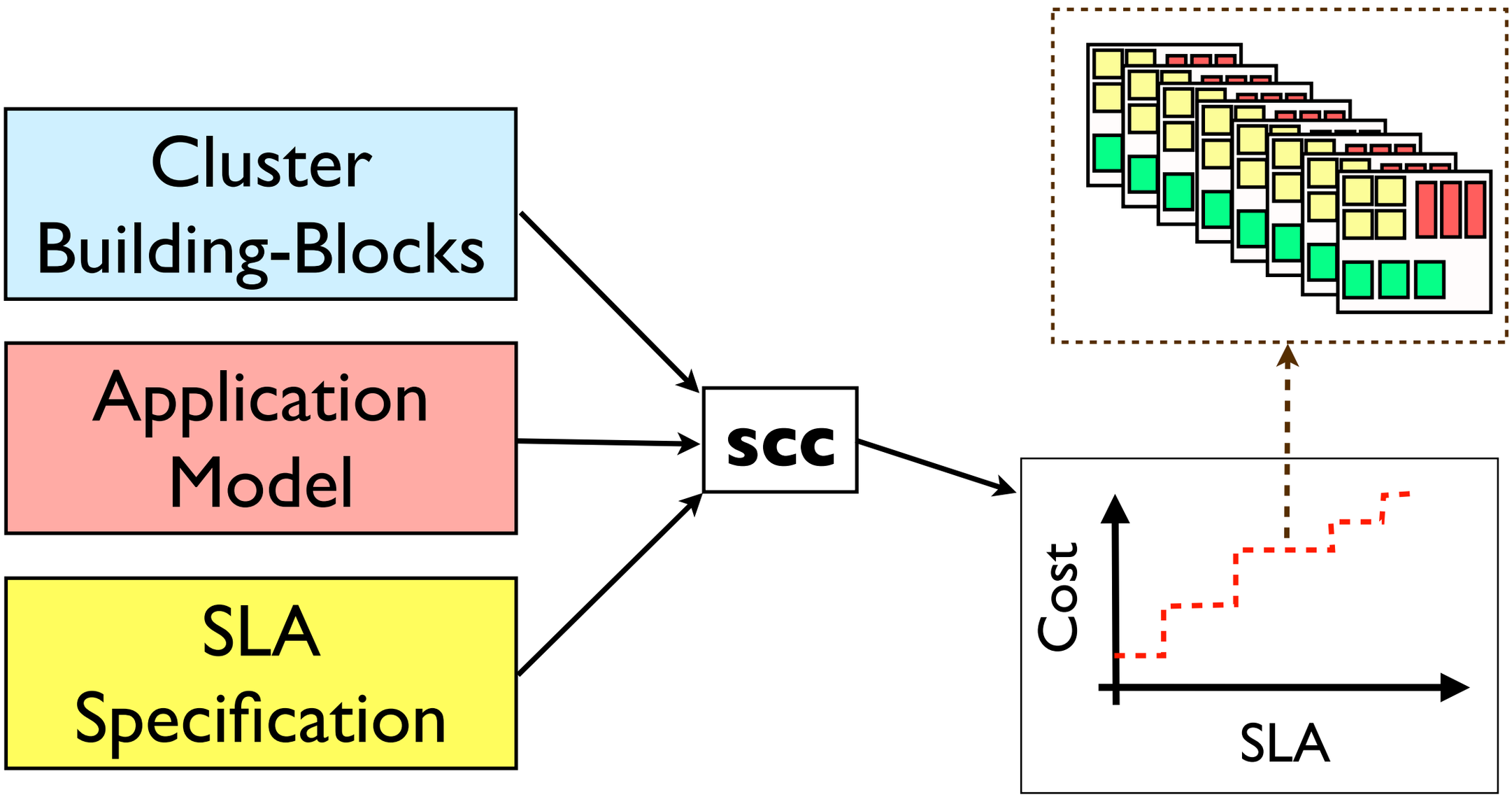
scc: Storage Configuration Compiler



scc: Storage Configuration Compiler



scc: Storage Configuration Compiler



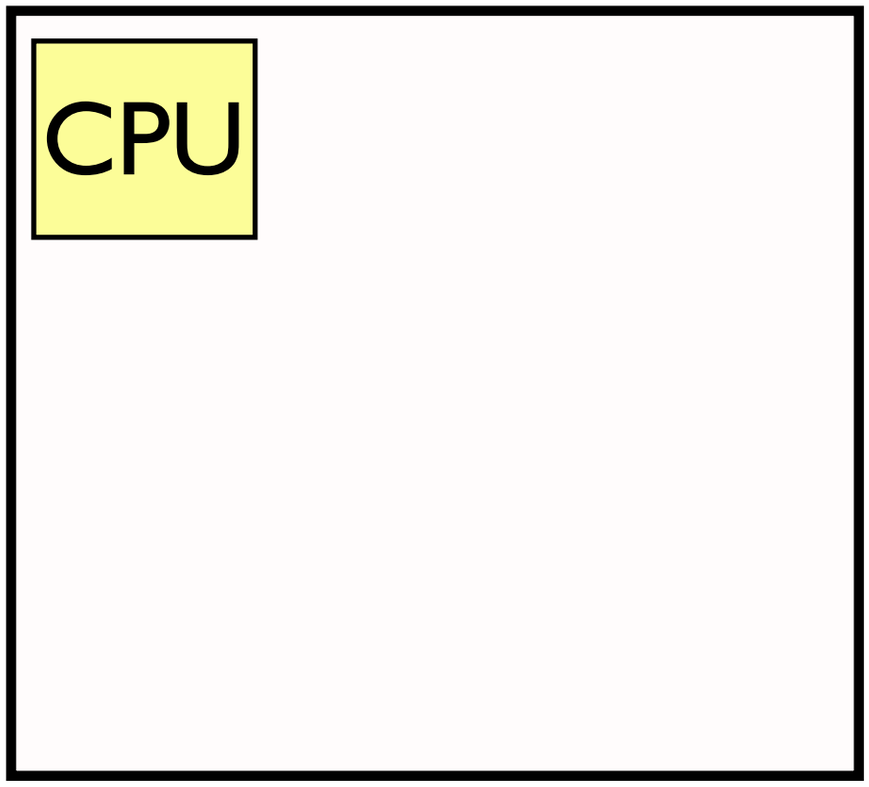
Outline

- Modeling Applications and Hardware
- Computing low-cost configurations
- Example
- Validation
- Applications of scc

Cluster Building Blocks

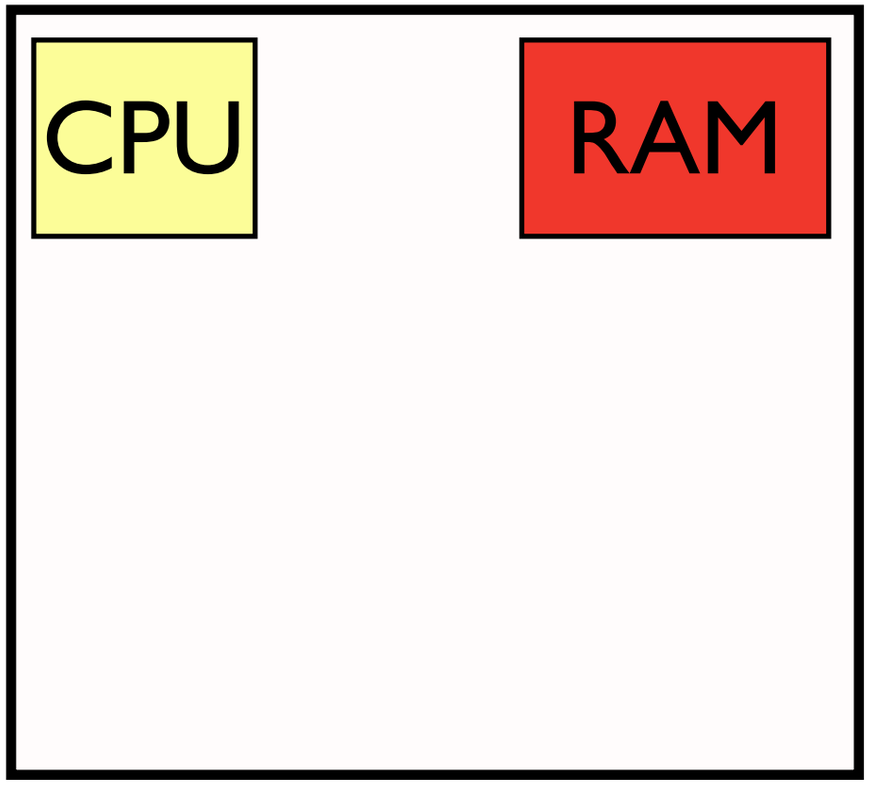
- Many types of servers

Cluster Building Blocks



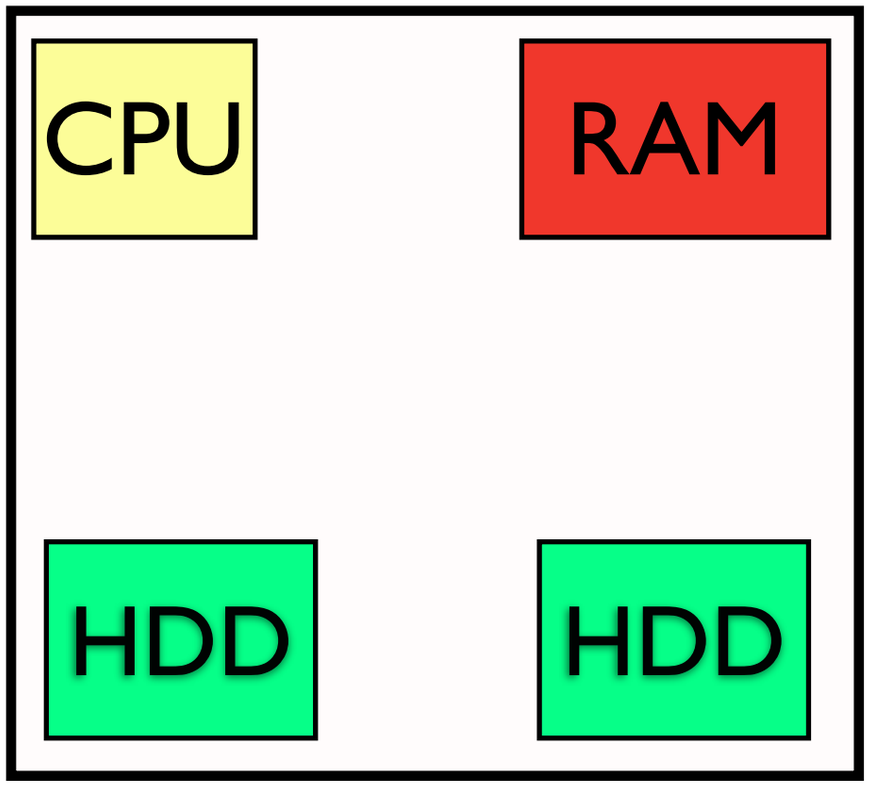
- Many types of servers
  - Cores

Cluster Building Blocks



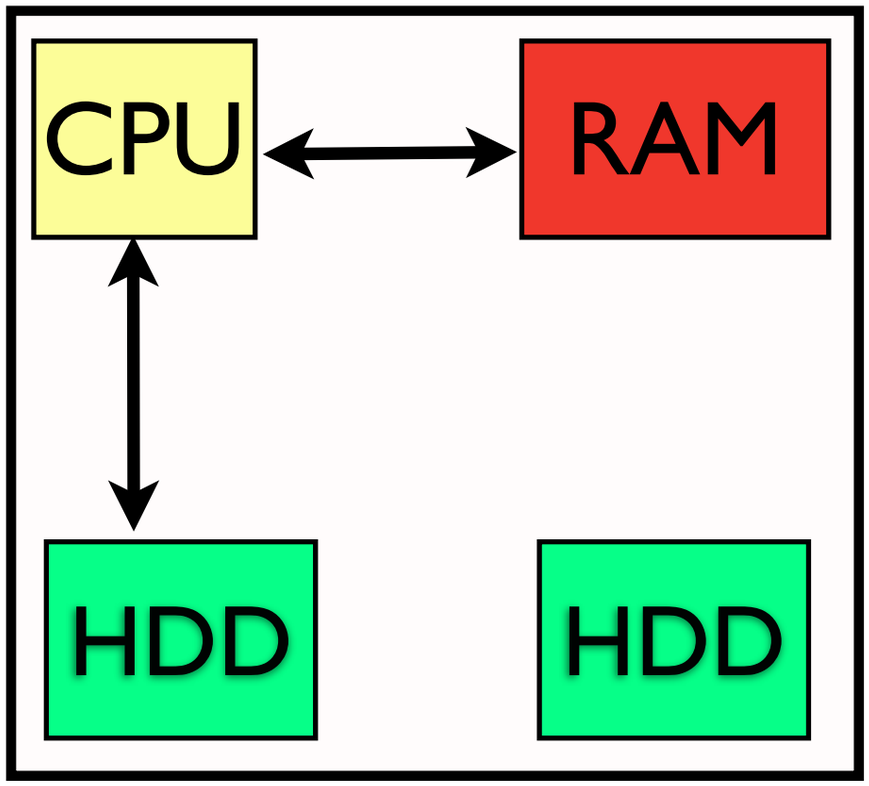
- Many types of servers
  - Cores
  - RAM

Cluster Building Blocks



- Many types of servers
  - Cores
  - RAM
  - Storage

Cluster Building Blocks



- Many types of servers
  - Cores
  - RAM
  - Storage
  - I/O & Network

Application Model

- Breakdown application into:
  - Tasks (Computation)
  - Datasets (Storage)
  - Edges between Tasks and Datasets (I/O)
  - Edges among Tasks (dependencies)

Example Model: Photo-Sharing

Tasks Datasets

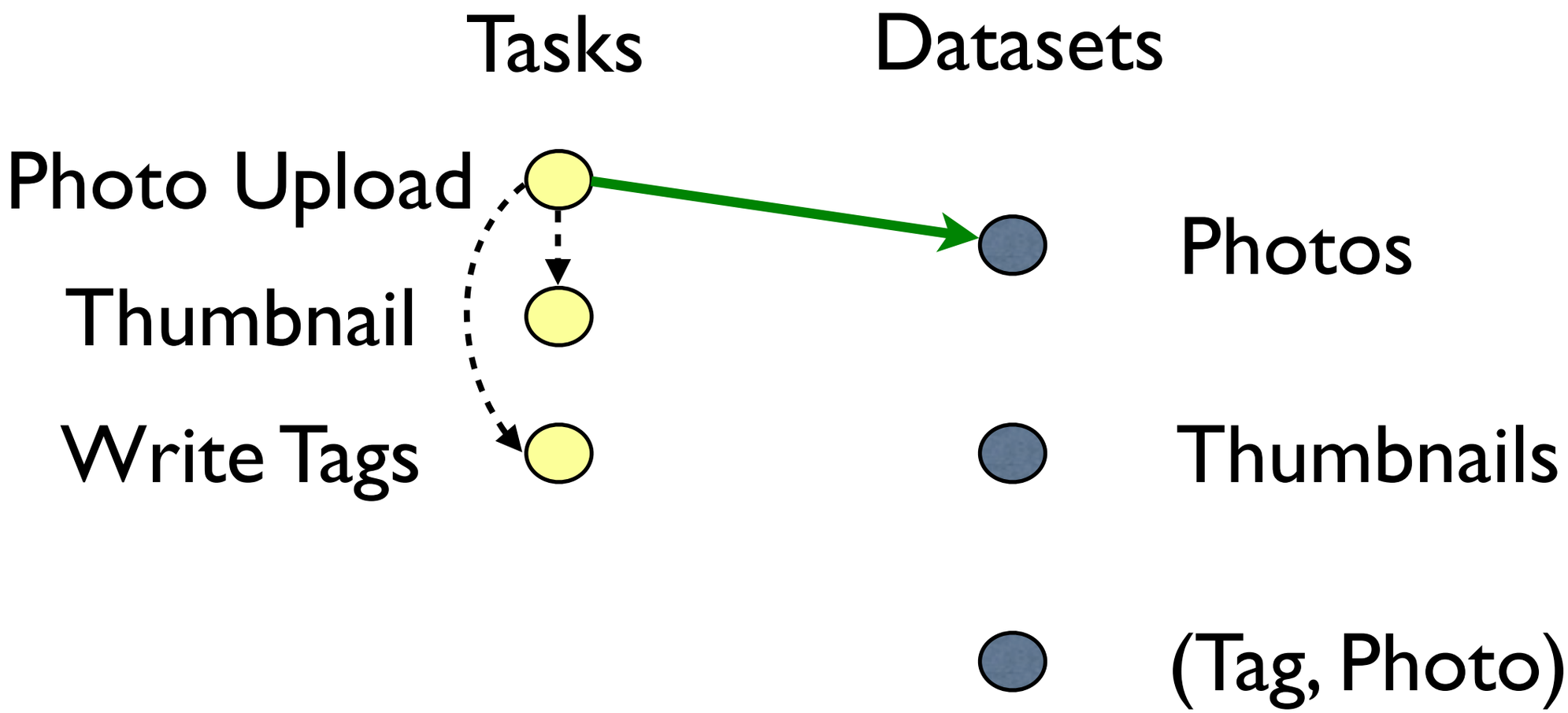
Example Model: Photo-Sharing

Tasks

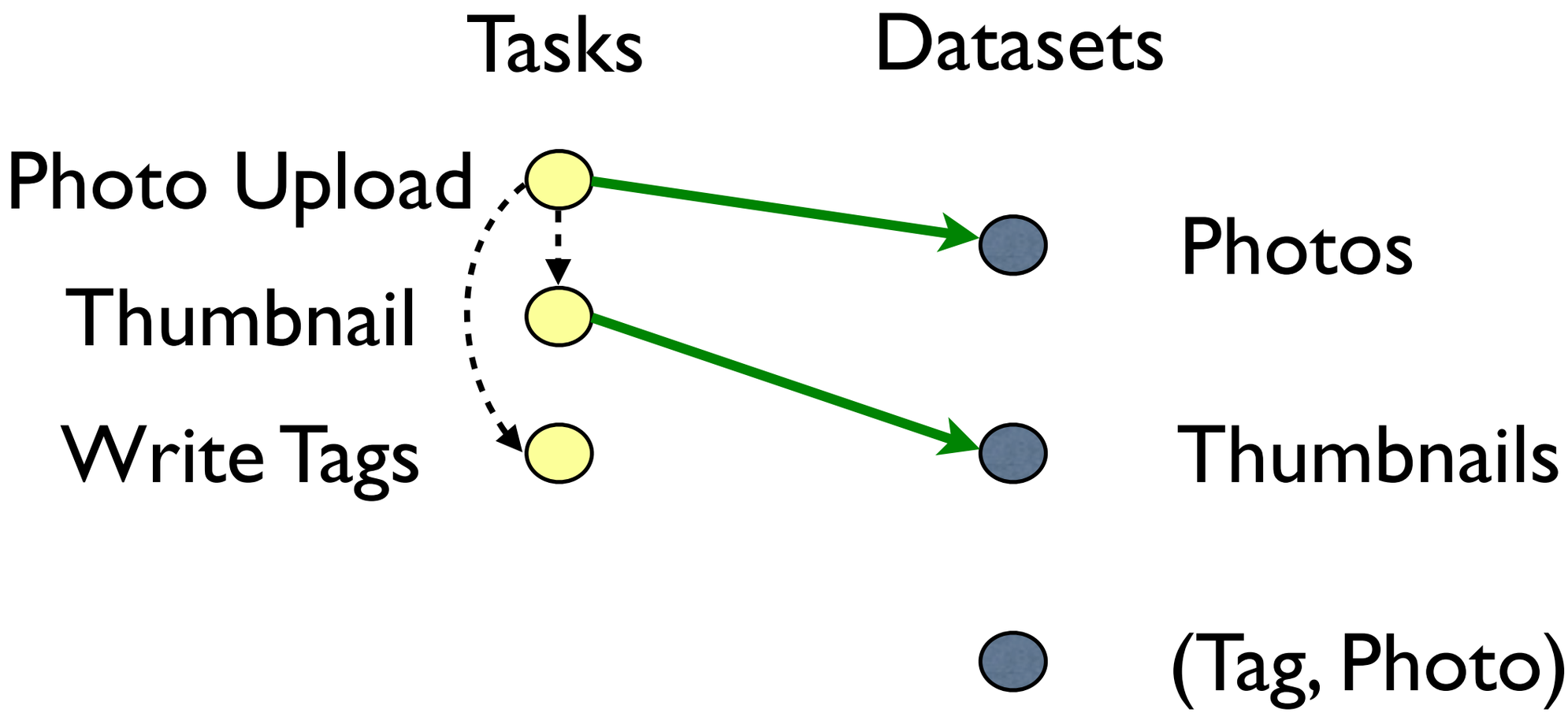
Datasets

- Photos
- Thumbnails
- (Tag, Photo)

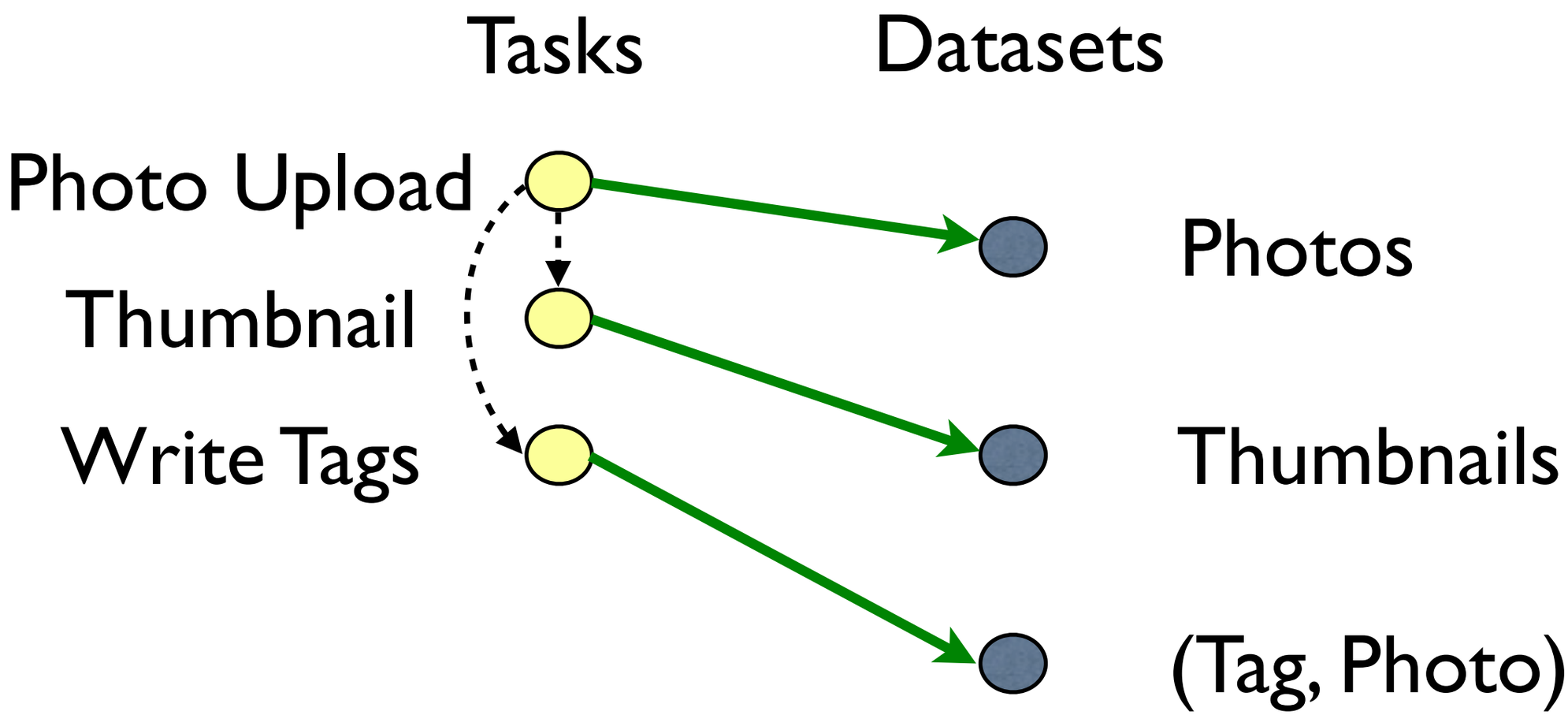
Example Model: Photo-Sharing



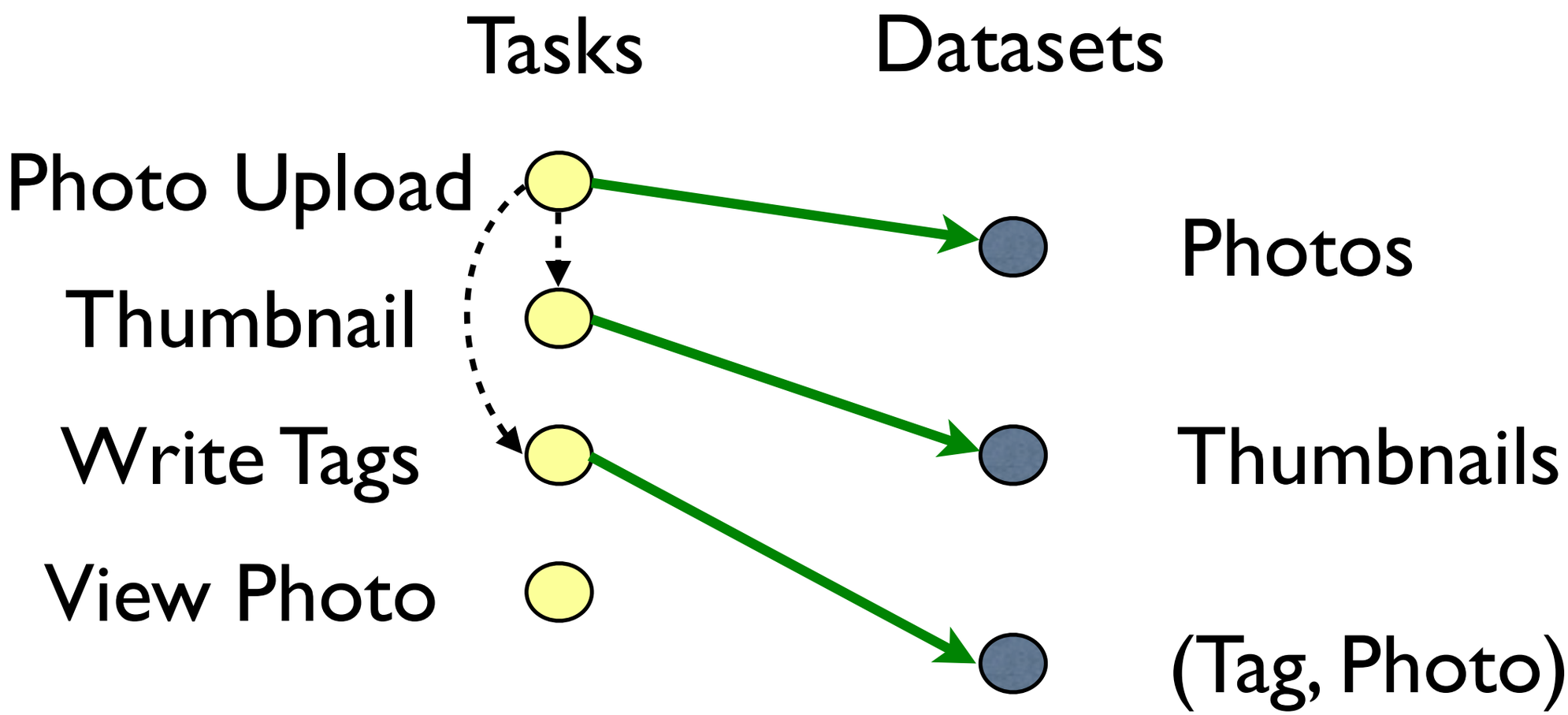
Example Model: Photo-Sharing



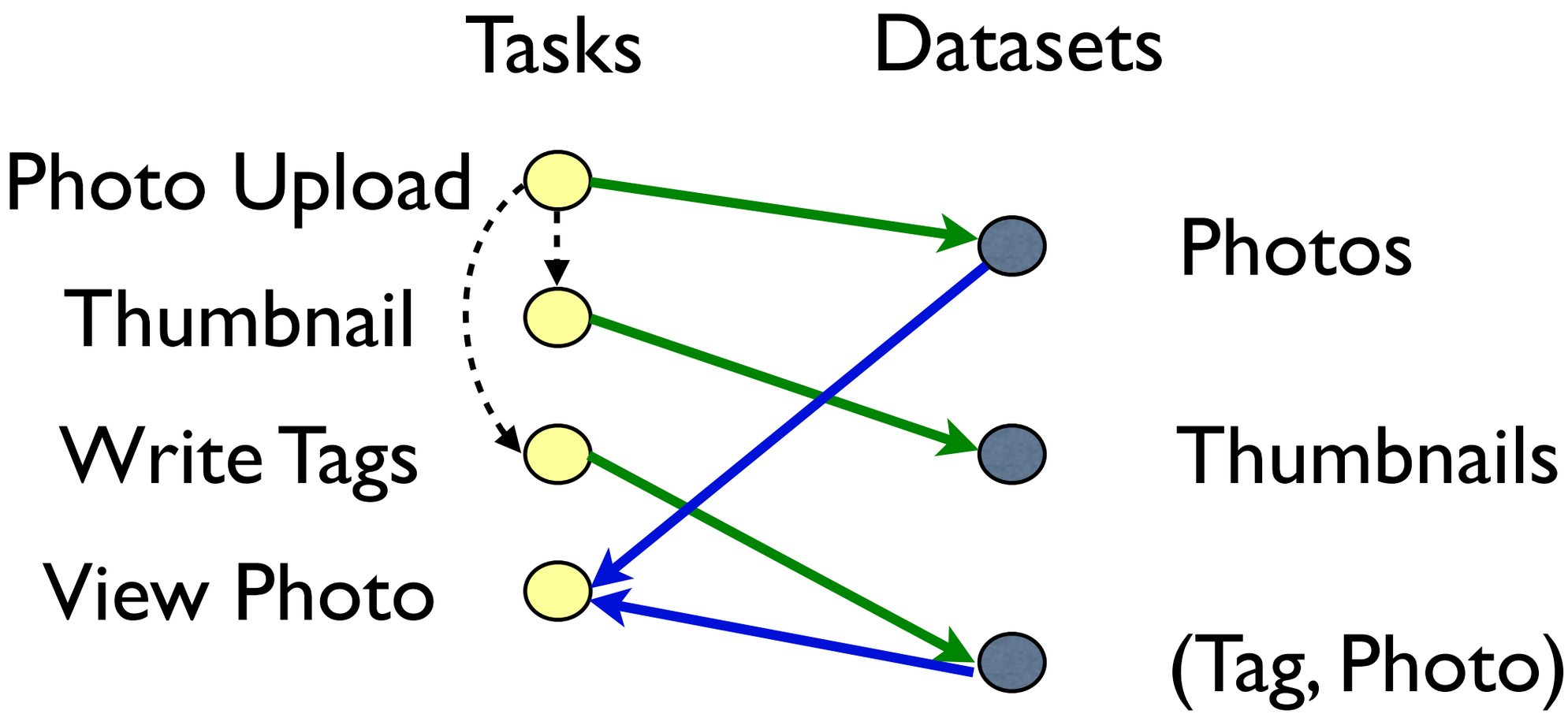
Example Model: Photo-Sharing



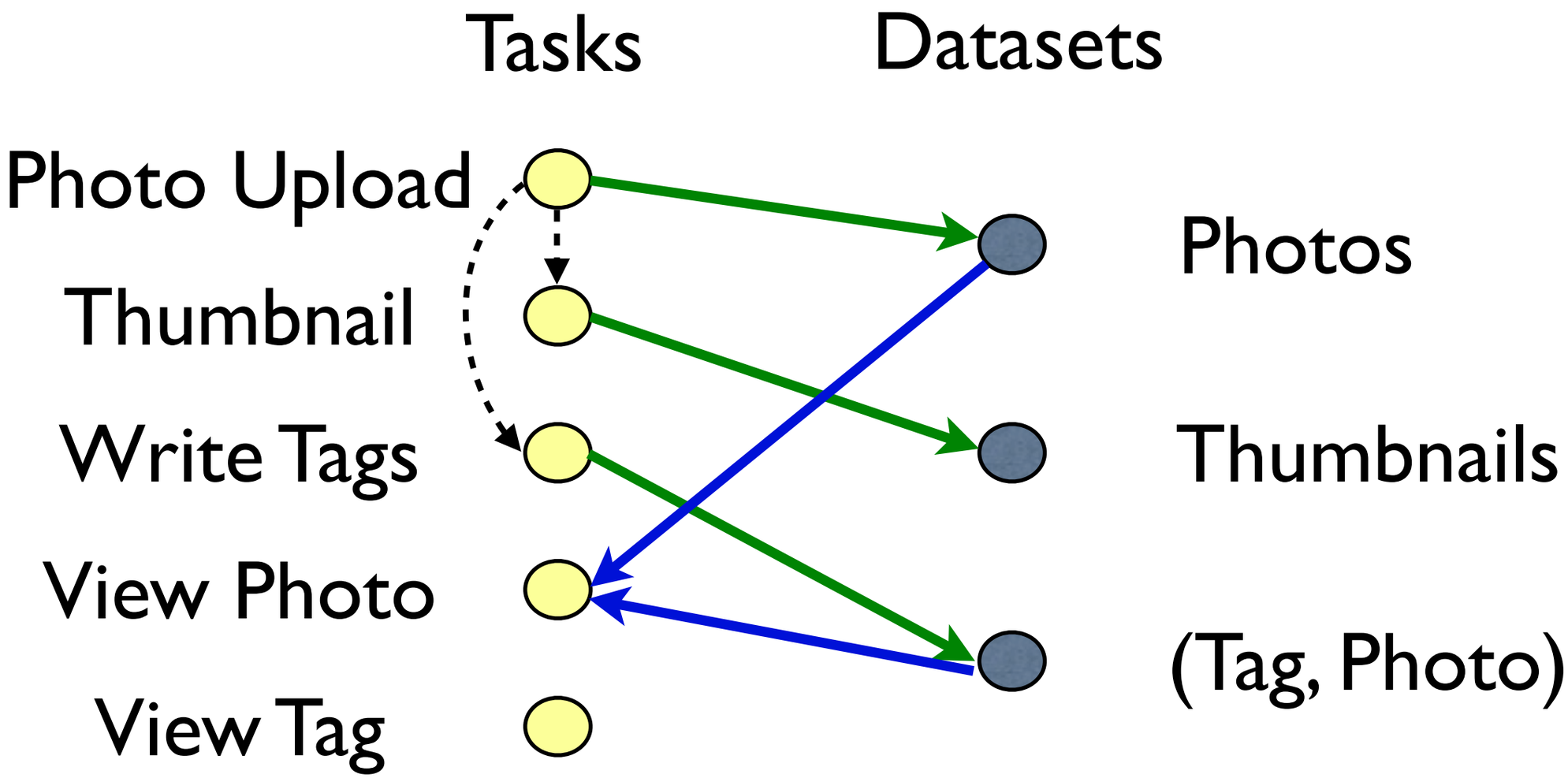
Example Model: Photo-Sharing



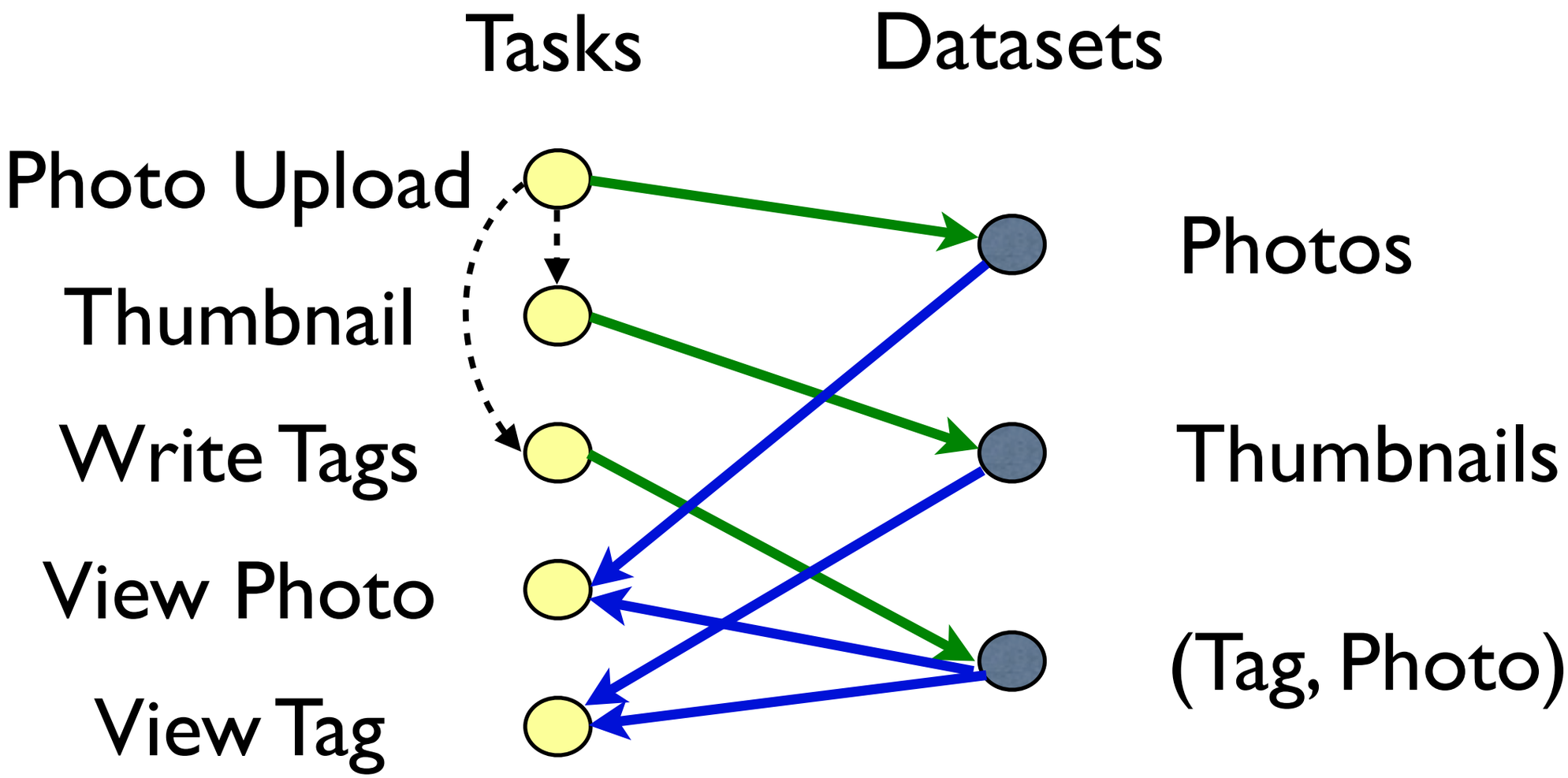
Example Model: Photo-Sharing



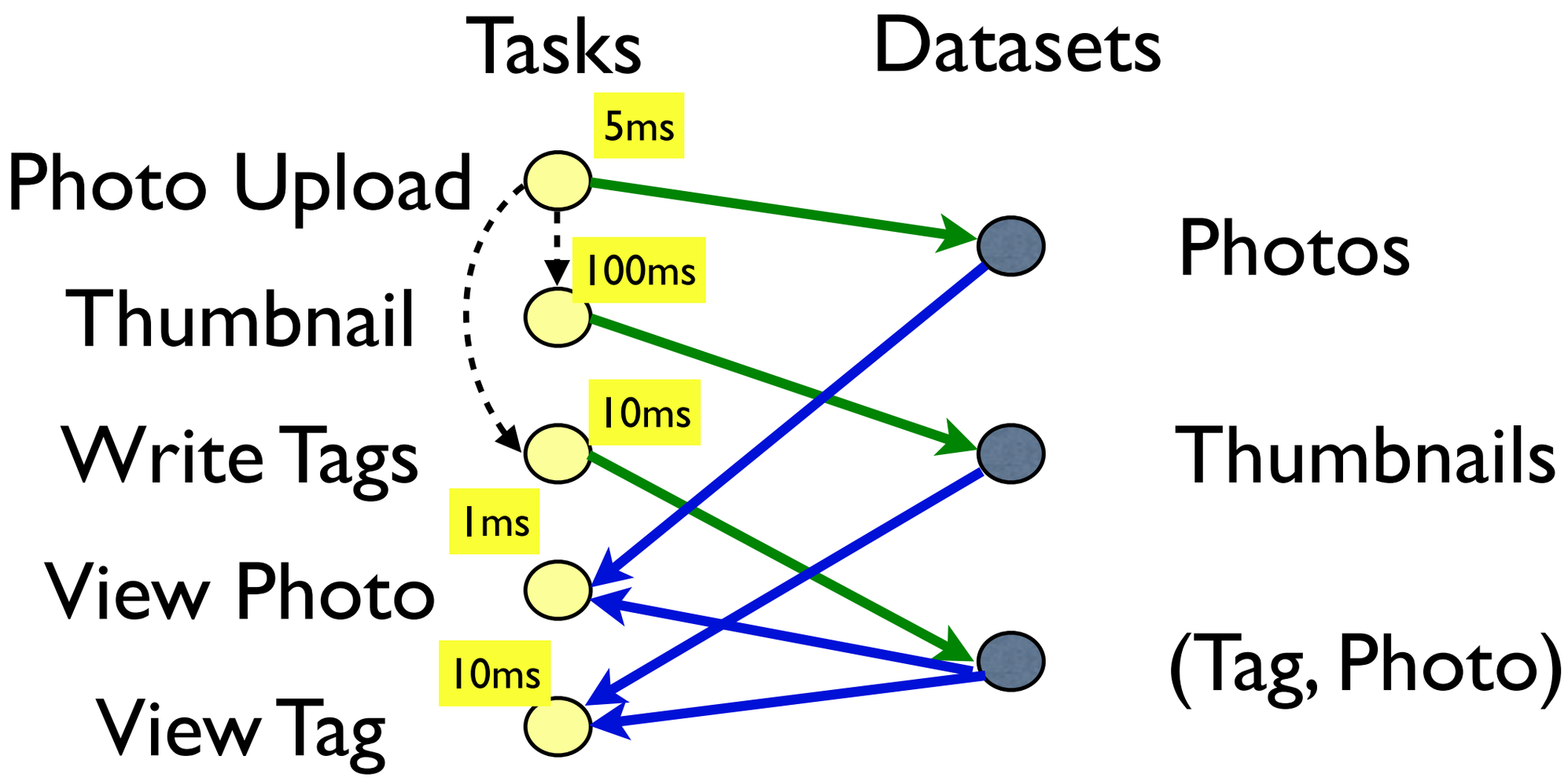
Example Model: Photo-Sharing



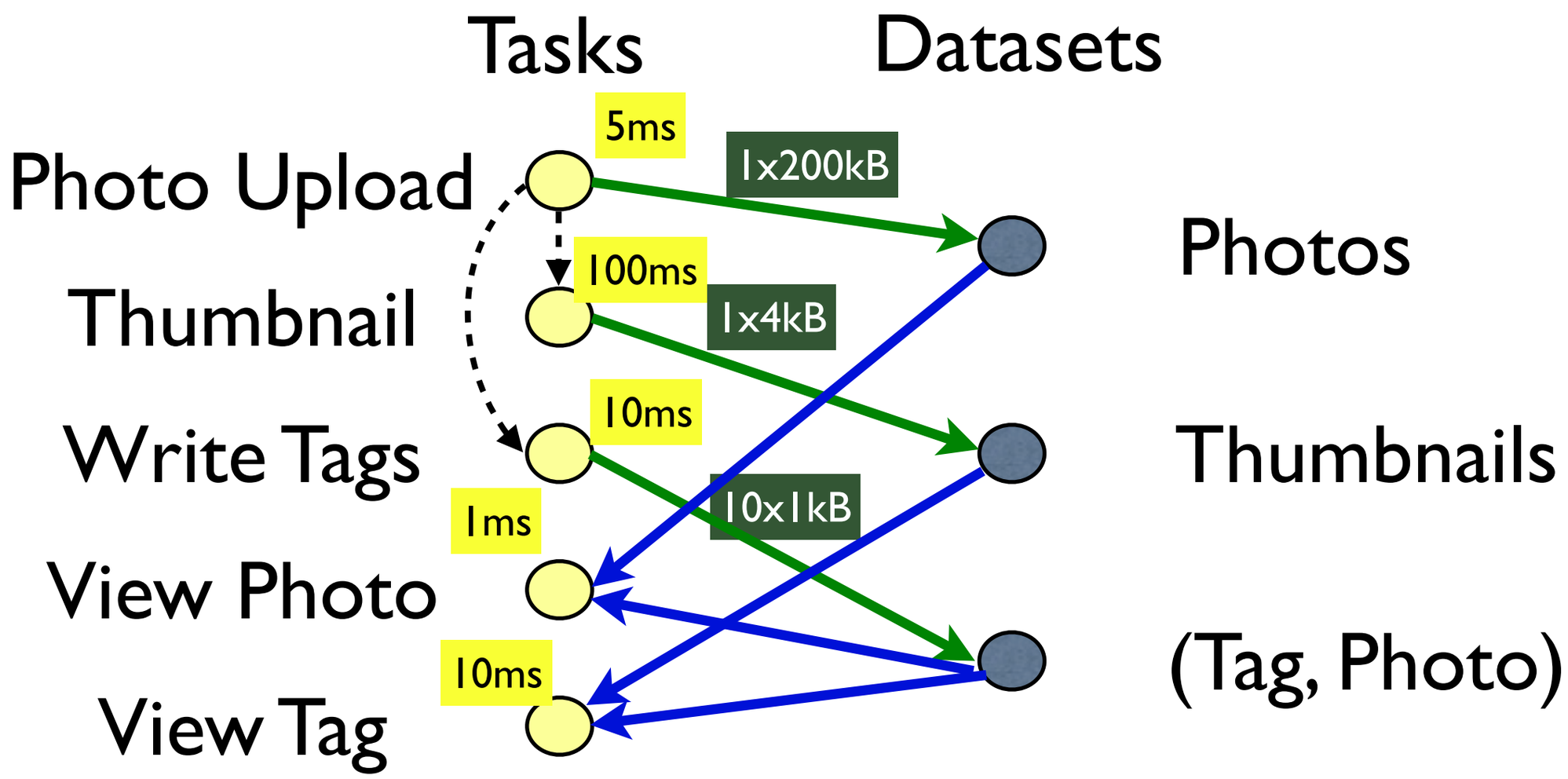
Example Model: Photo-Sharing



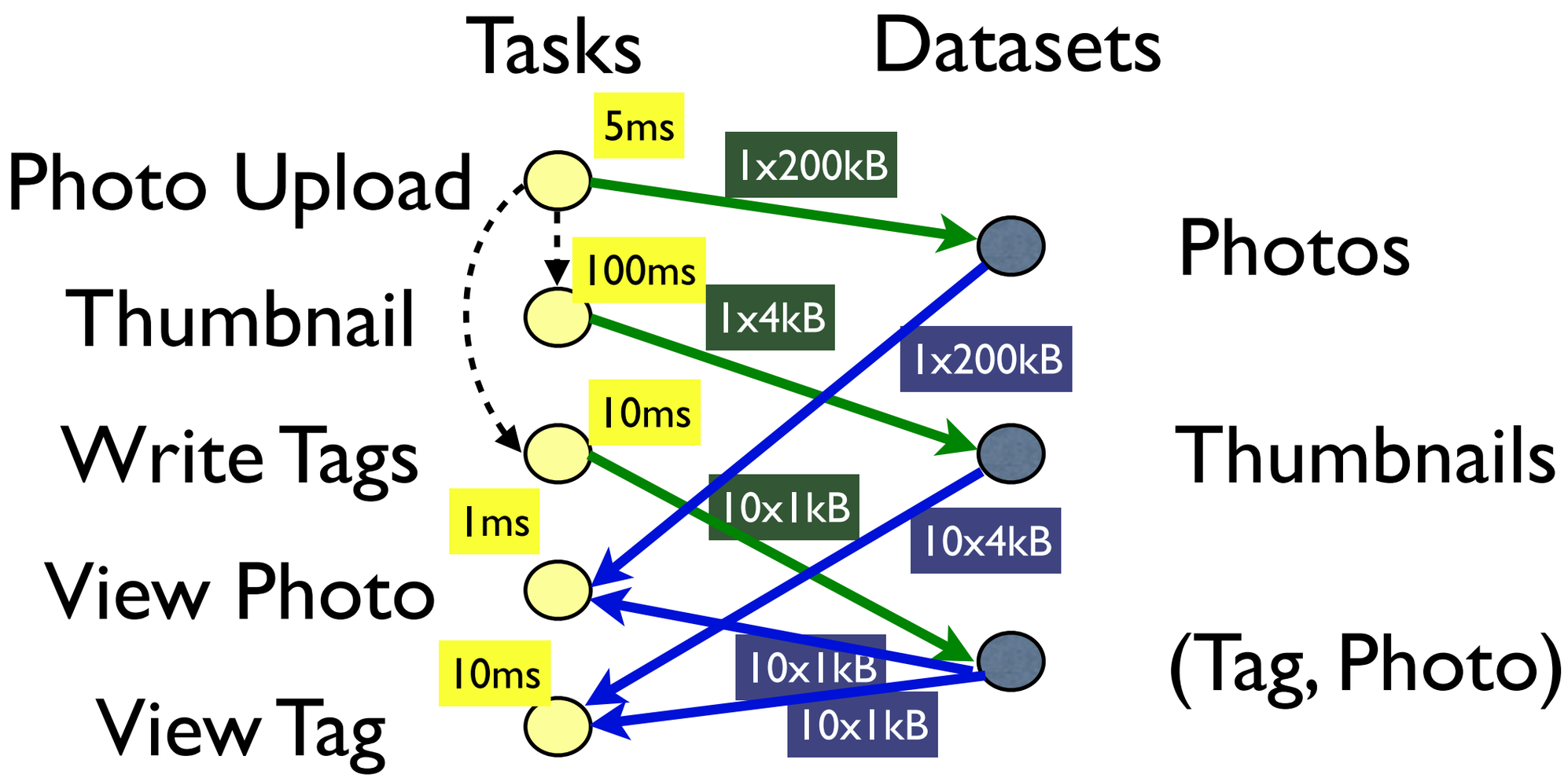
Example Model: Photo-Sharing



Example Model: Photo-Sharing



Example Model: Photo-Sharing



Example Model: Photo-Sharing

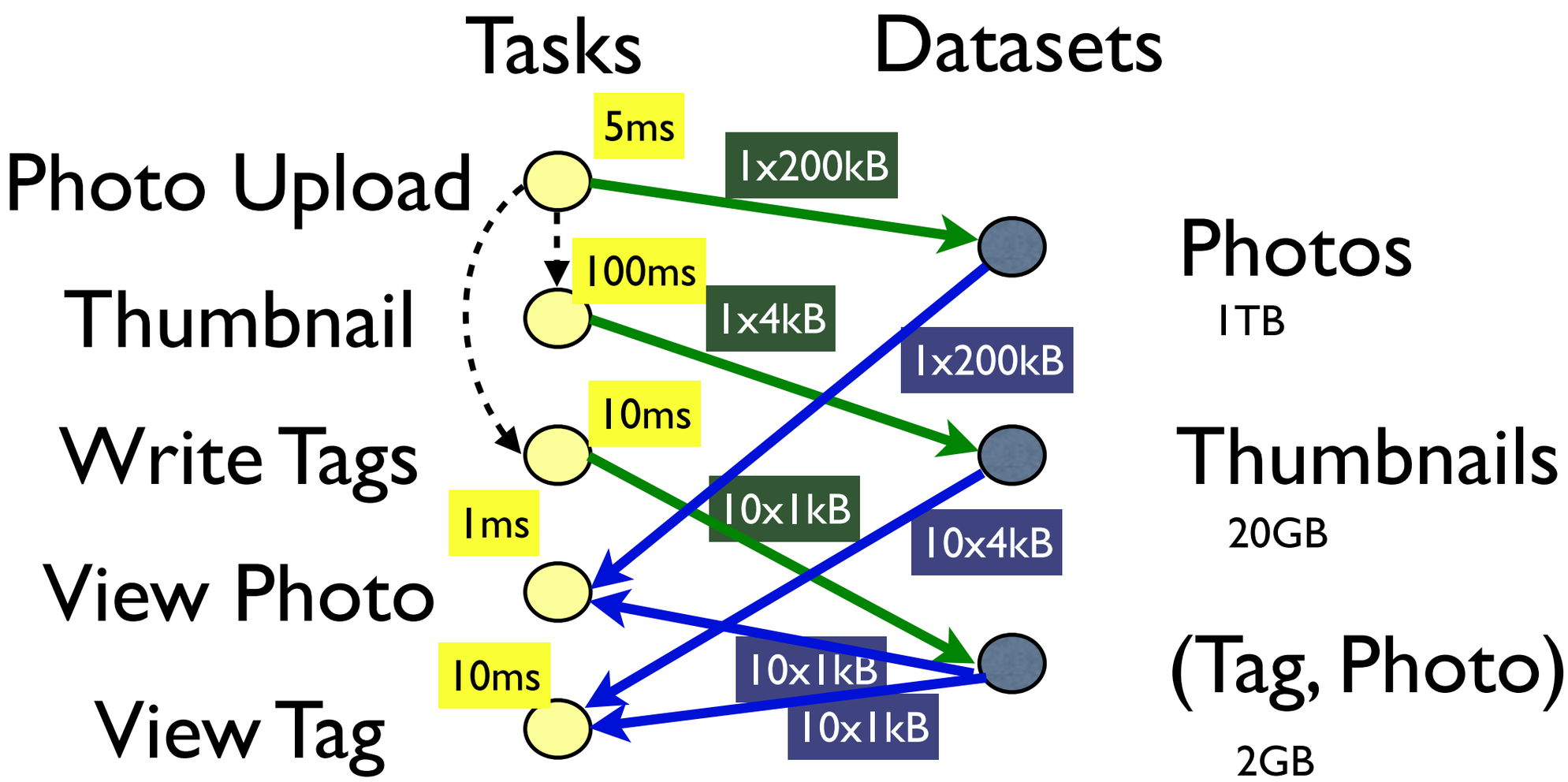


Photo-Sharing SLA

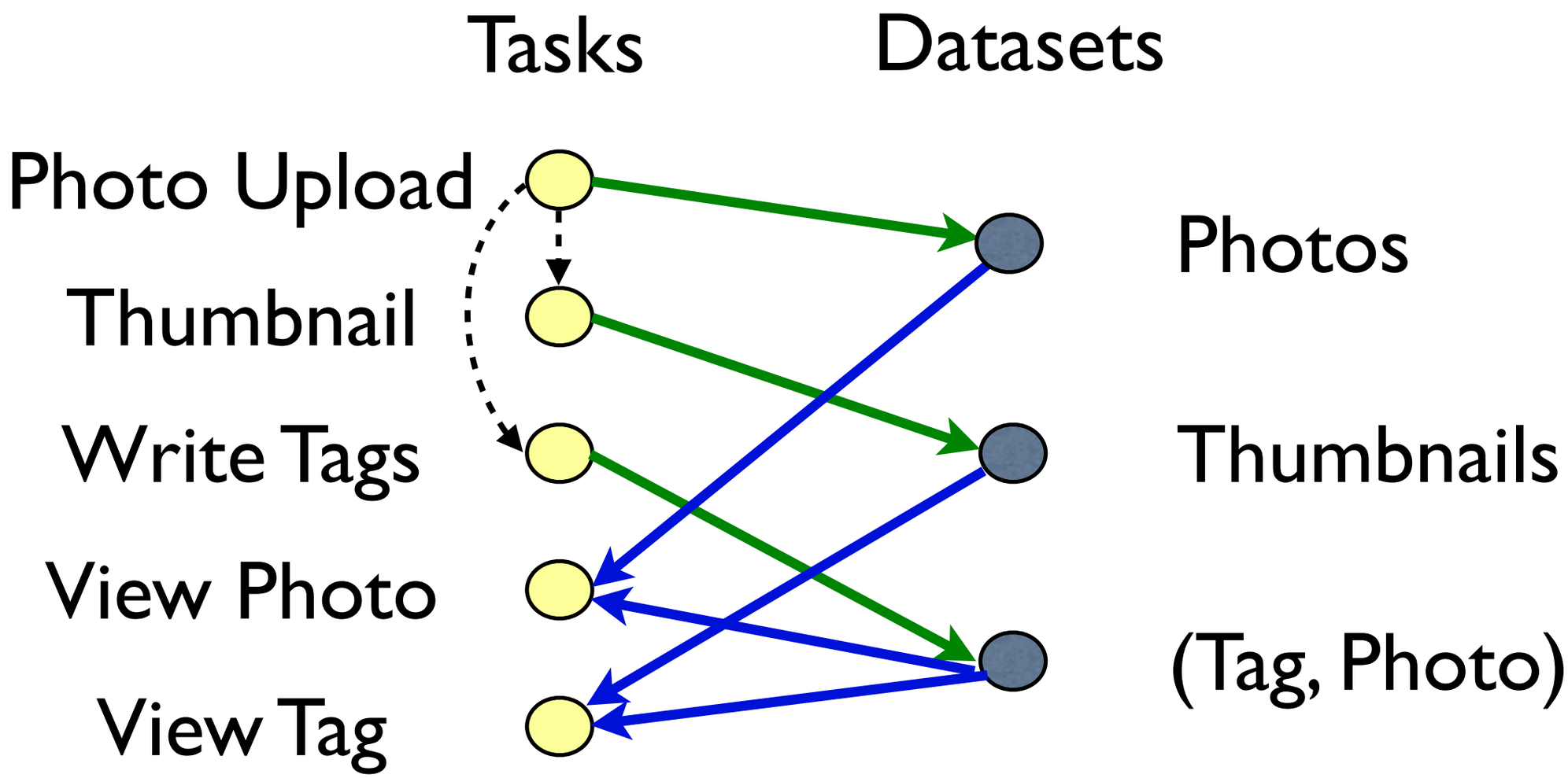
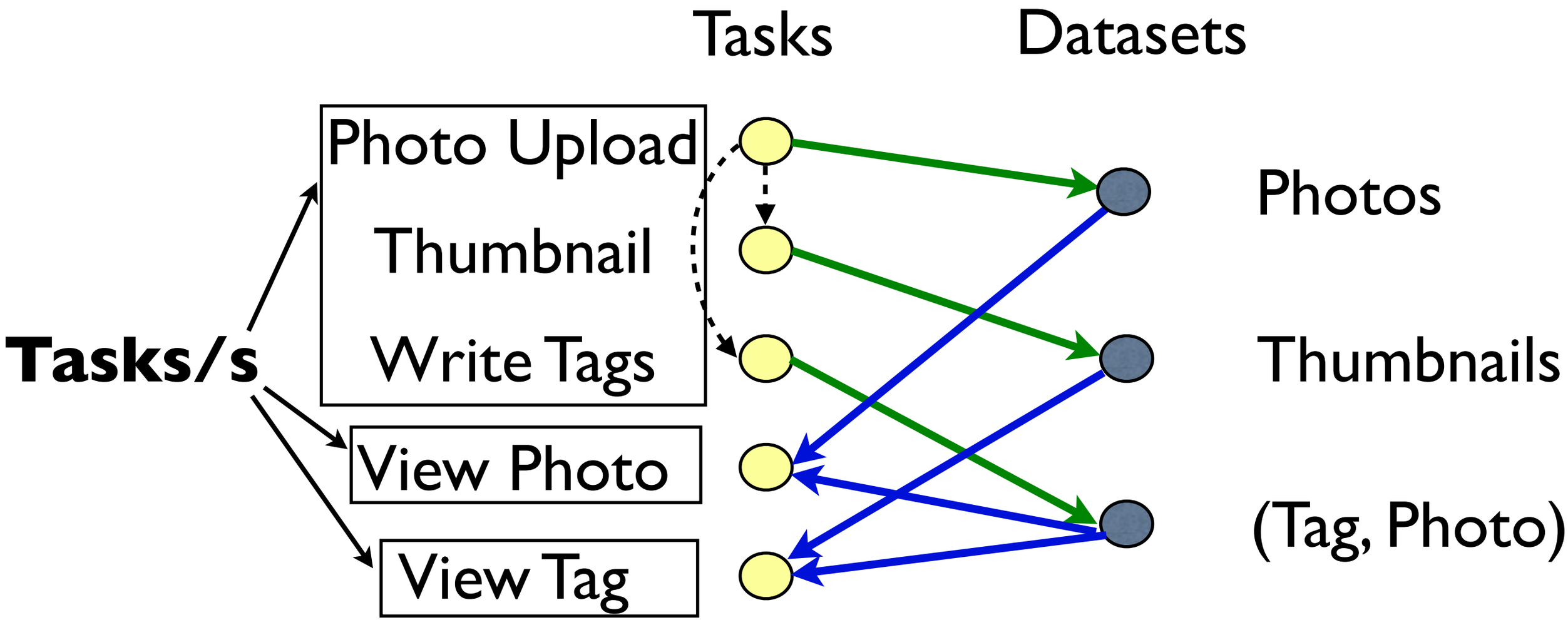


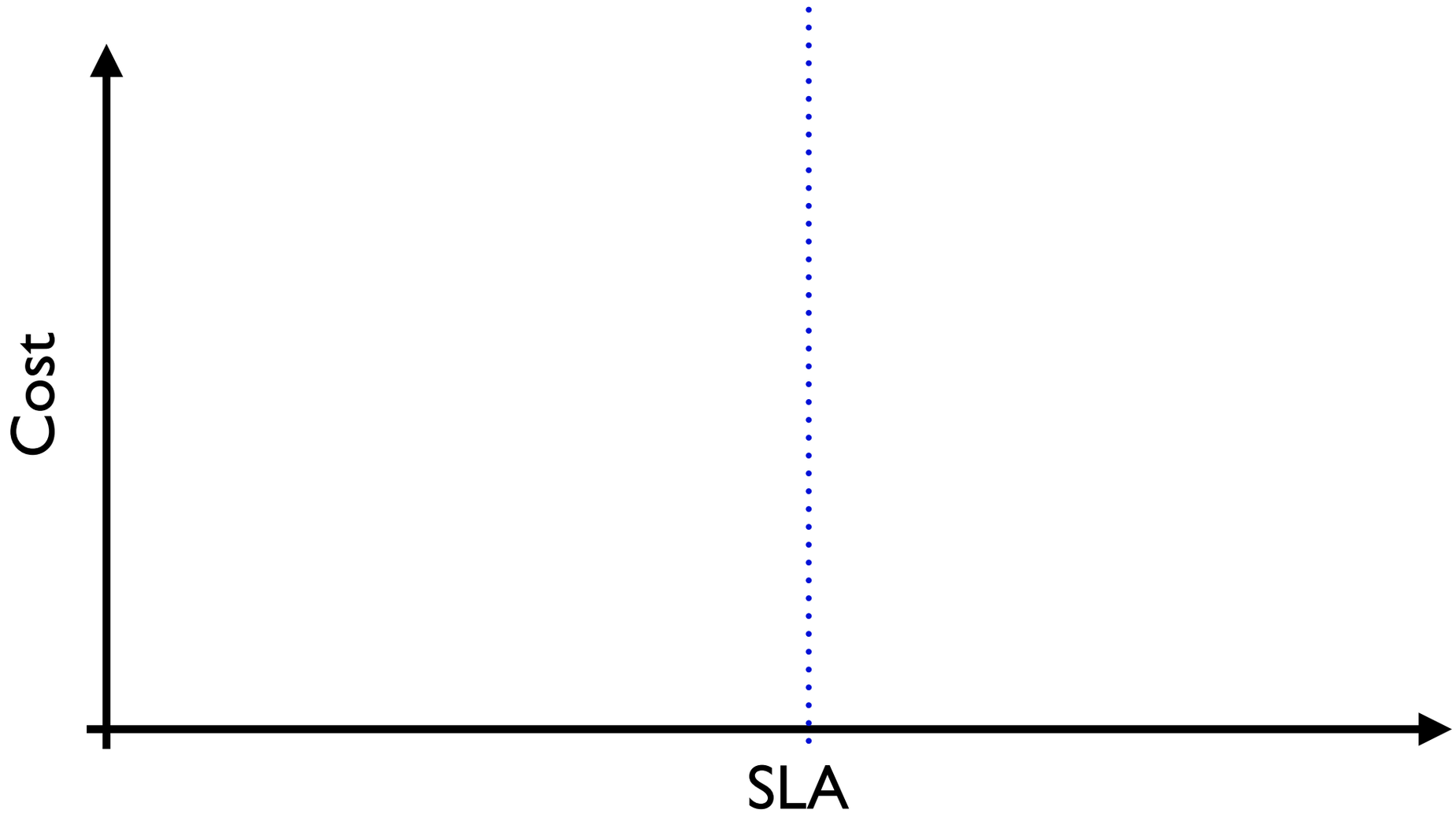
Photo-Sharing SLA



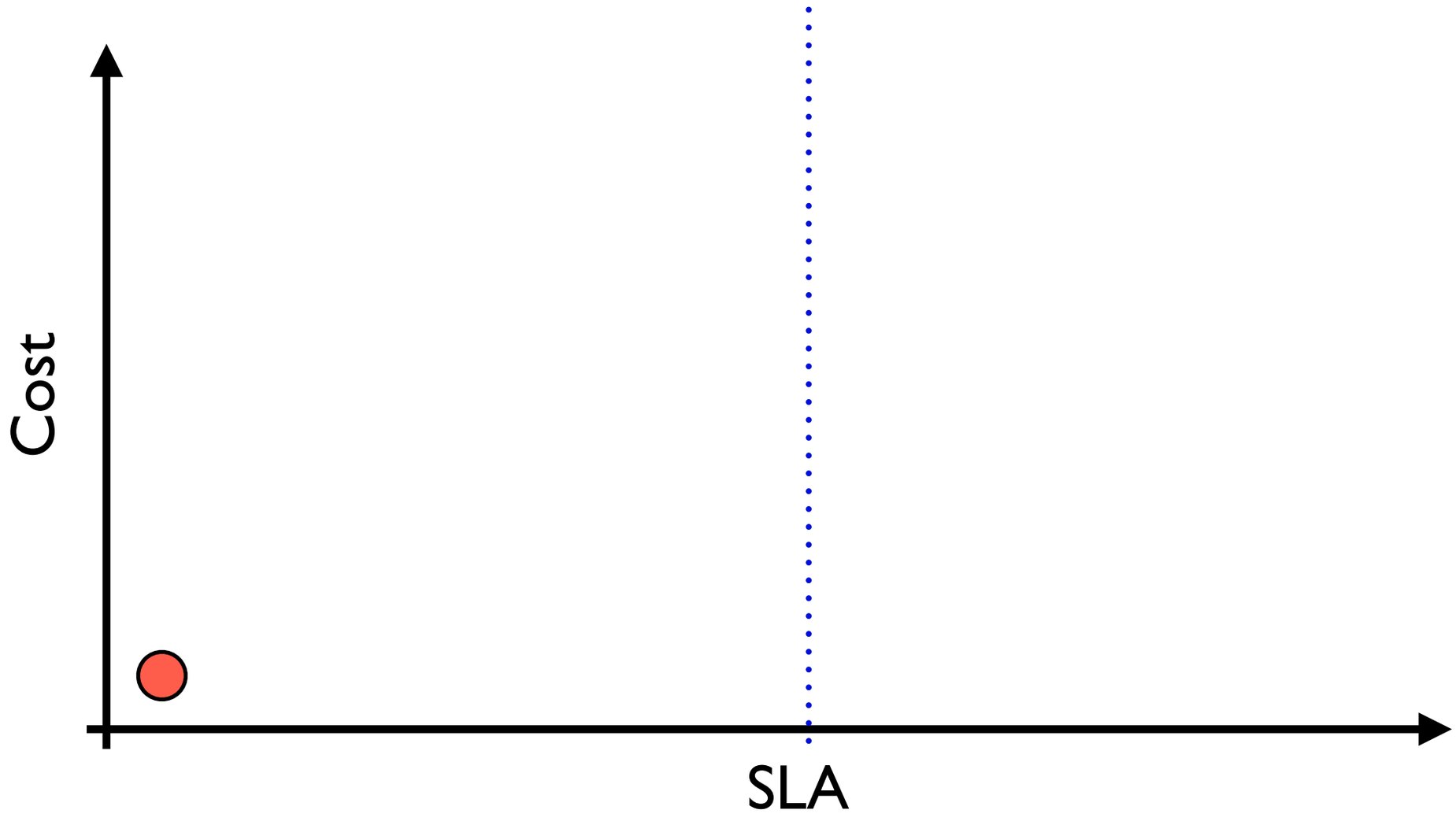
Outline

- Modeling Applications and Hardware
- Computing low-cost configurations
- Example
- Validation
- Applications of scc

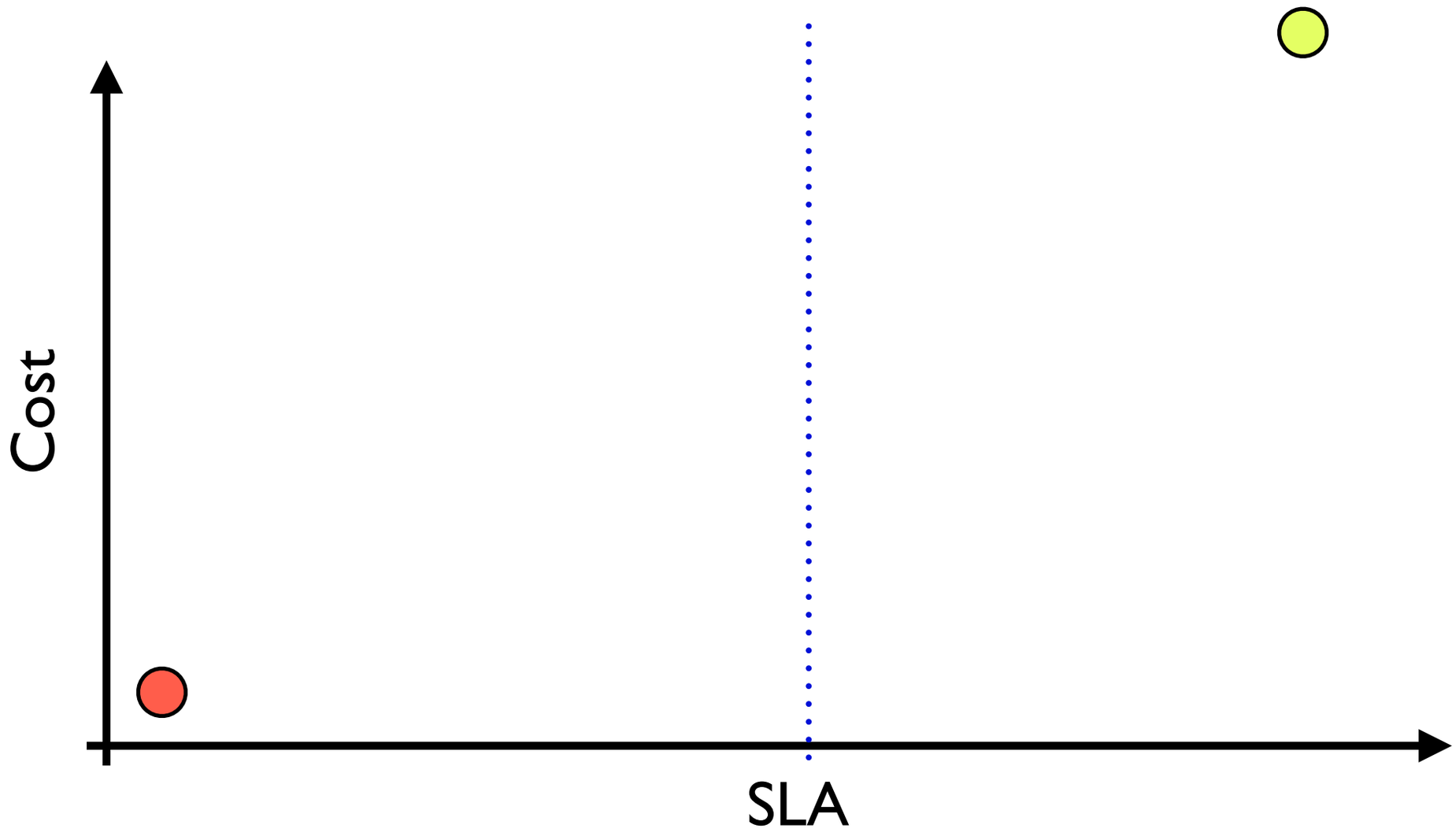
Navigating the Configuration Space



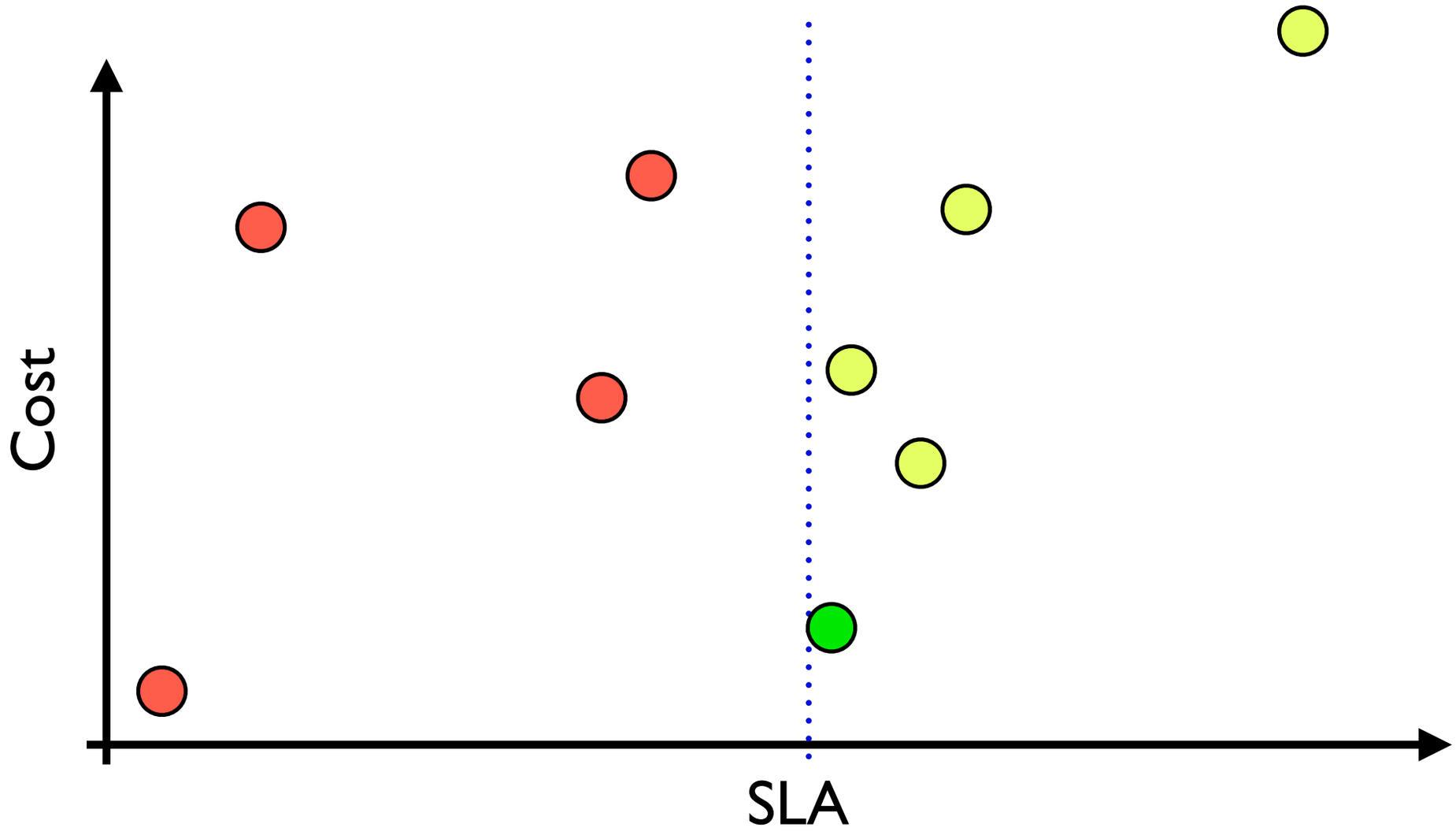
Navigating the Configuration Space



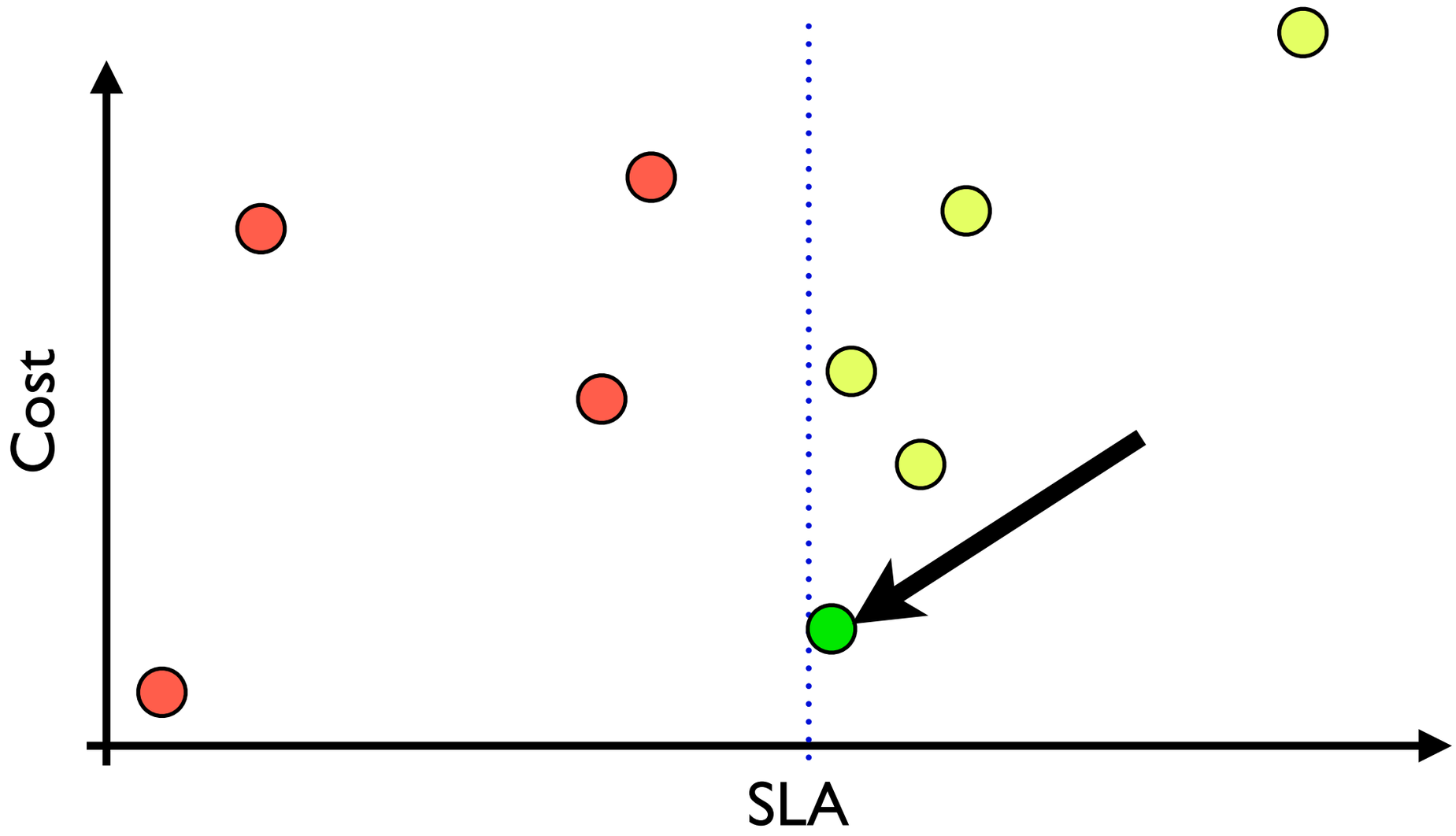
Navigating the Configuration Space



Navigating the Configuration Space



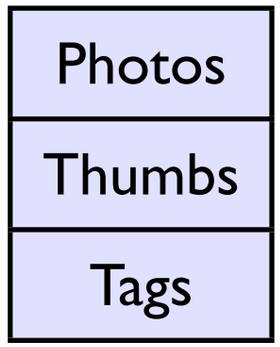
Navigating the Configuration Space



Guiding Principle to Meet SLA

- Complex interaction across storage-type and dataset assignments
- Need to consider costs of meeting SLA for each permutation
- Our configuration space is:
  - **D** datasets, **S** storage-types → $\mathbf{D}^{\mathbf{S}}$ configs

Meeting the SLA at Low-cost

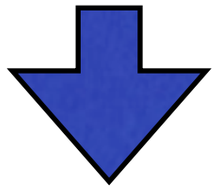


Meeting the SLA at Low-cost

Photos	7.2krpm
Thumbs	7.2krpm
Tags	7.2krpm

Meeting the SLA at Low-cost

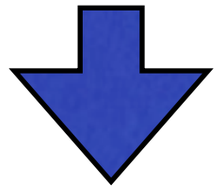
Photos	7.2krpm
Thumbs	7.2krpm
Tags	7.2krpm



storage units per-dataset

Meeting the SLA at Low-cost

Photos	7.2krpm
Thumbs	7.2krpm
Tags	7.2krpm

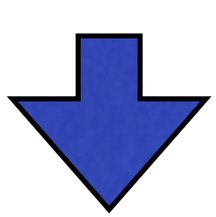


storage units per-dataset

Assign CPUs for computation

Meeting the SLA at Low-cost

Photos	7.2krpm
Thumbs	7.2krpm
Tags	7.2krpm



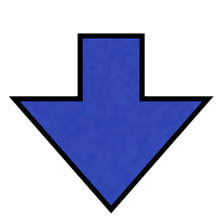
storage units per-dataset

Assign CPUs for computation

Pack into servers

Meeting the SLA at Low-cost

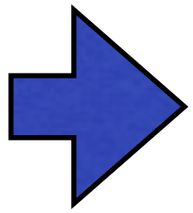
Photos	7.2krpm
Thumbs	7.2krpm
Tags	7.2krpm



storage units per-dataset

Assign CPUs for computation

Pack into servers

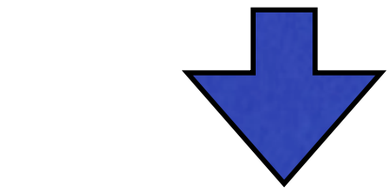


Compute cost

Meeting the SLA at Low-cost

Photos	7.2krpm	7.2krpm	7.2krpm	7.2krpm	7.2krpm
Thumbs	7.2krpm	7.2krpm	7.2krpm	7.2krpm	7.2krpm
Tags	7.2krpm	15krpm	SSD	7.2k+DRAM	15k+DRAM

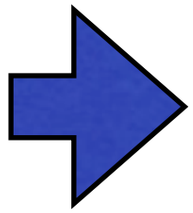
...



storage units per-dataset

Assign CPUs for computation

Pack into servers

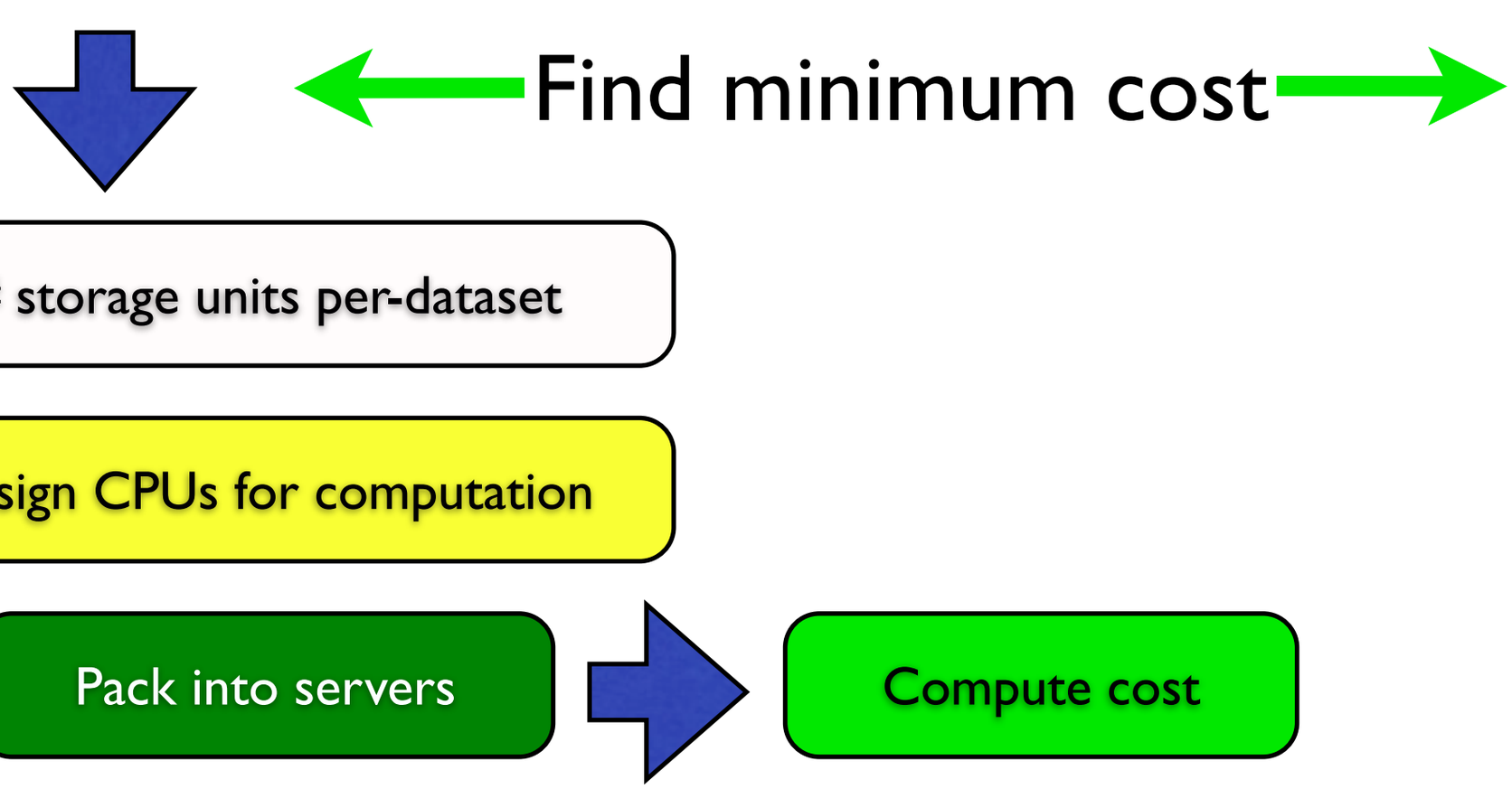


Compute cost

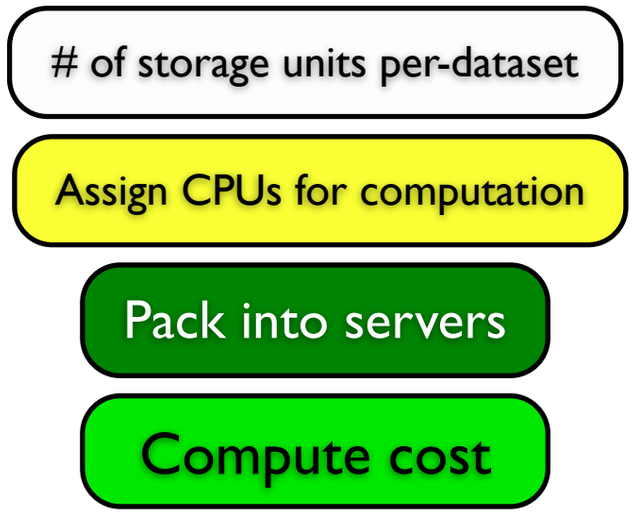
Meeting the SLA at Low-cost

Photos	7.2krpm	7.2krpm	7.2krpm	7.2krpm	7.2krpm
Thumbs	7.2krpm	7.2krpm	7.2krpm	7.2krpm	7.2krpm
Tags	7.2krpm	15krpm	SSD	7.2k+DRAM	15k+DRAM

...

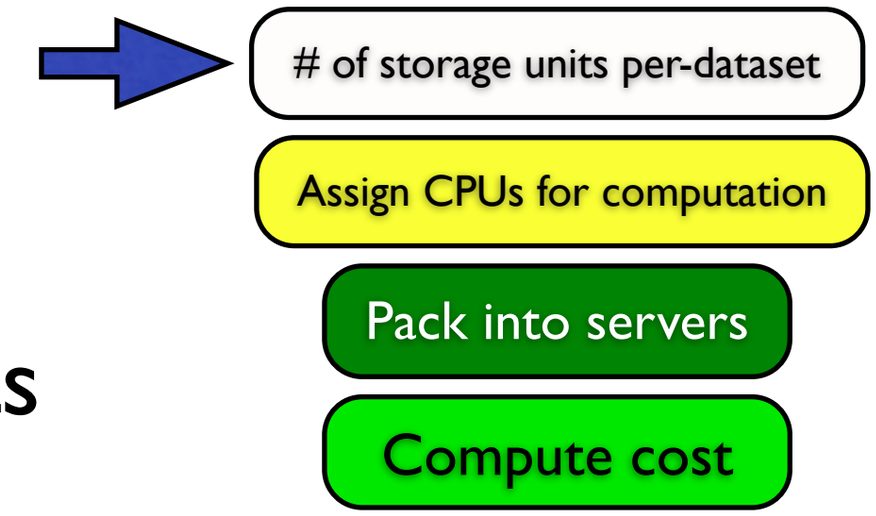


Detail: How Many Storage Units?

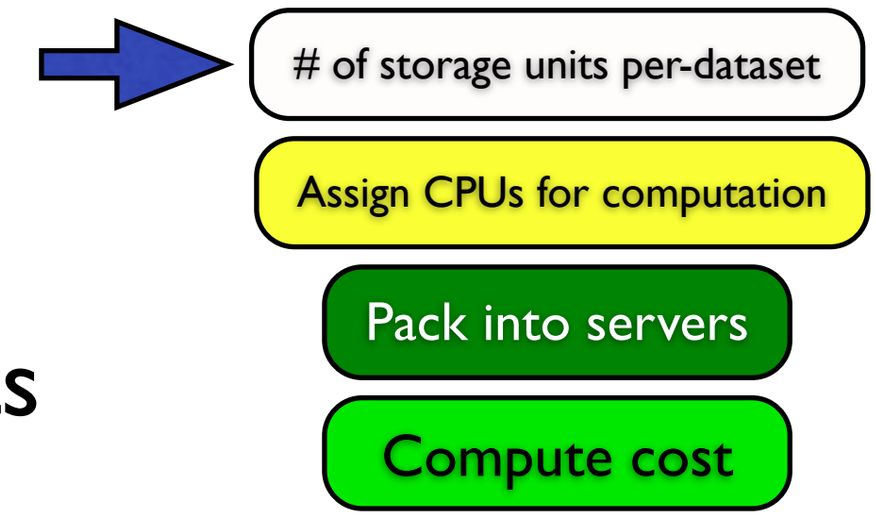


Detail: How Many Storage Units?

- SLA+Model=Requirements
- Need to match requirements to storage units

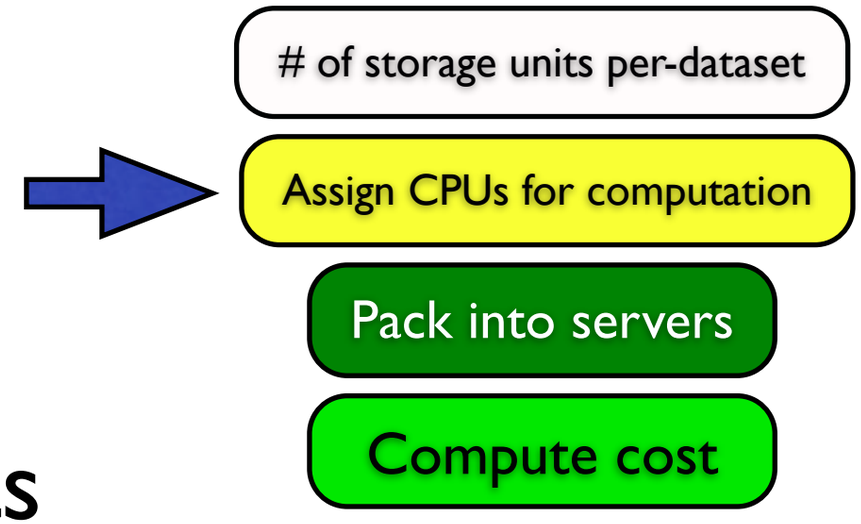


Detail: How Many Storage Units?



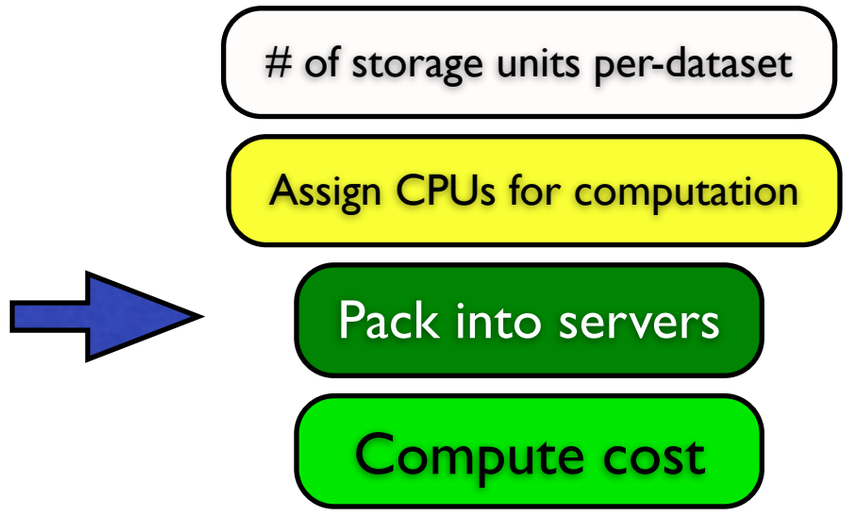
- SLA+Model=Requirements
- Need to match requirements to storage units
  - Capacity
  - IOPS
    - Write-heavy: Short-stroke disks
    - Read-heavy: Consider RAM for caching

Detail: How Many CPUs?



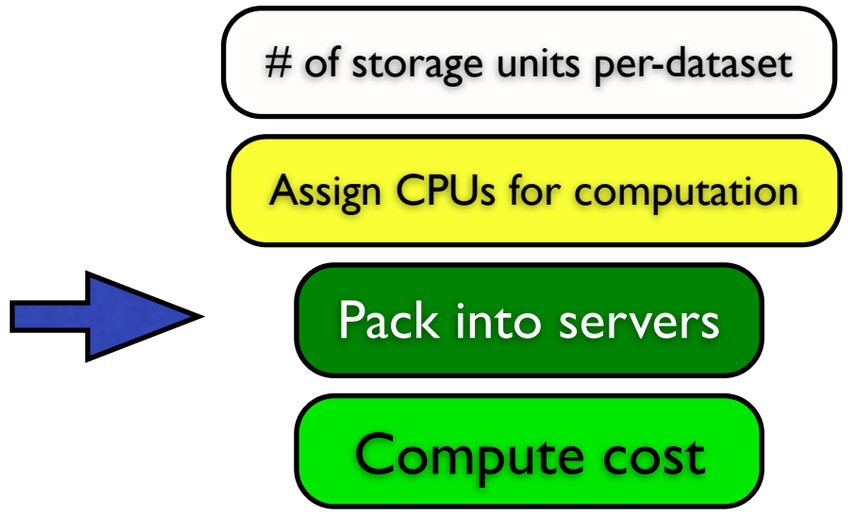
- Enough CPUs to satisfy computation requirements
  - Linearly extrapolate computation time
- Minimum one core/server
- (Details in paper)

Detail: Fits in how many servers?



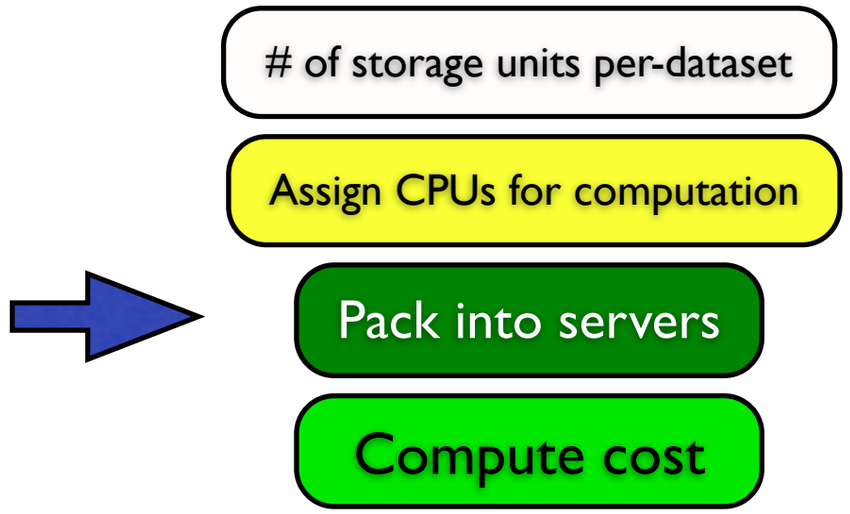
Detail: Fits in how many servers?

- Integer Linear Programming
- Multiple server types, each with different constraints:
  - I/O Bus Bandwidth
  - I/O Slots
  - CPU Cores
  - Network Bandwidth



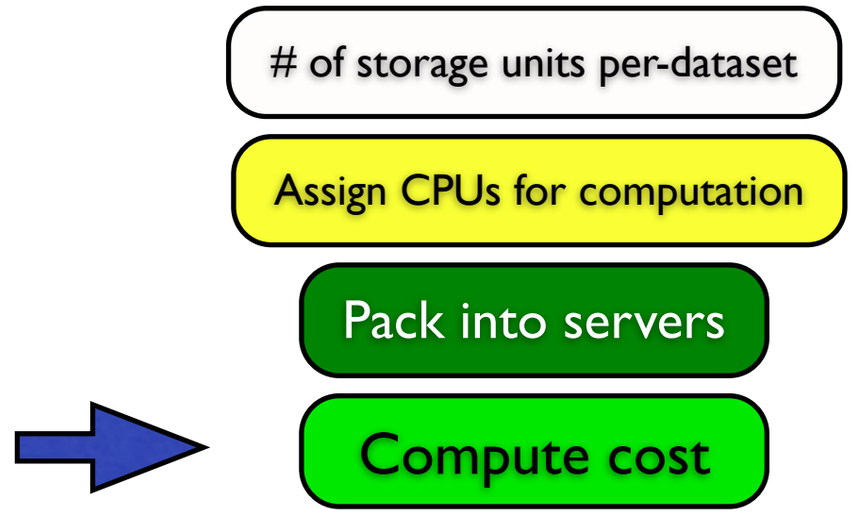
Detail: Fits in how many servers?

- Integer Linear Programming
- Multiple server types, each with different constraints:
  - I/O Bus Bandwidth
  - I/O Slots
  - CPU Cores
  - Network Bandwidth
- Sequential-workload → Bus/Network constrained
- IOPS-workload → Slot capacity constrained



Detail: How much does it cost?

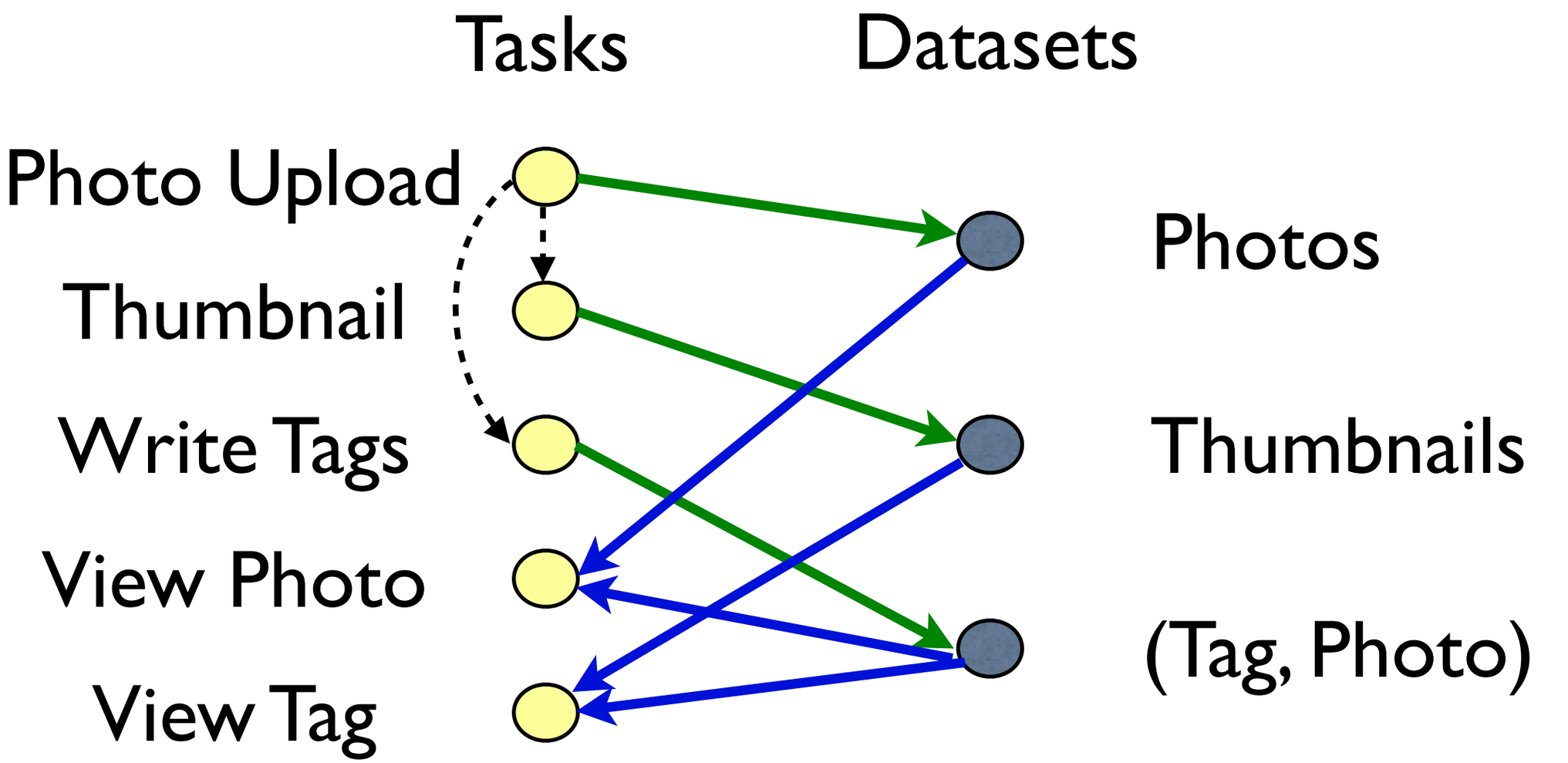
- ILP minimizes cost:
  - Capital expenses
    - Each component
  - Operation expenses
    - Power & cooling



Outline

- Modeling Applications and Hardware
- Computing low-cost configurations
- Example
- Validation
- Applications of scc

Example Application: Photo-Sharing



Example: Building Blocks

- Building blocks
  - 3GHZ Intel Xeon (max 4 cores)
  - 146GB 15k-rpm Disks
  - 500GB 7.2k-rpm Disks } (max 4 per machine)
  - 32GB SSD
  - 1 GB RAM (max 15 per machine)
  - Gigabit Ethernet

Example: SLA

- SLA Requirement
  - 100 image uploads/s
  - 300 photo views/s
  - 100 tags views/s

Example: Final Stages for Photo-sharing

Photos

15k-rpm

15k-rpm

15k-rpm

Thumbnails

SSD

Tags

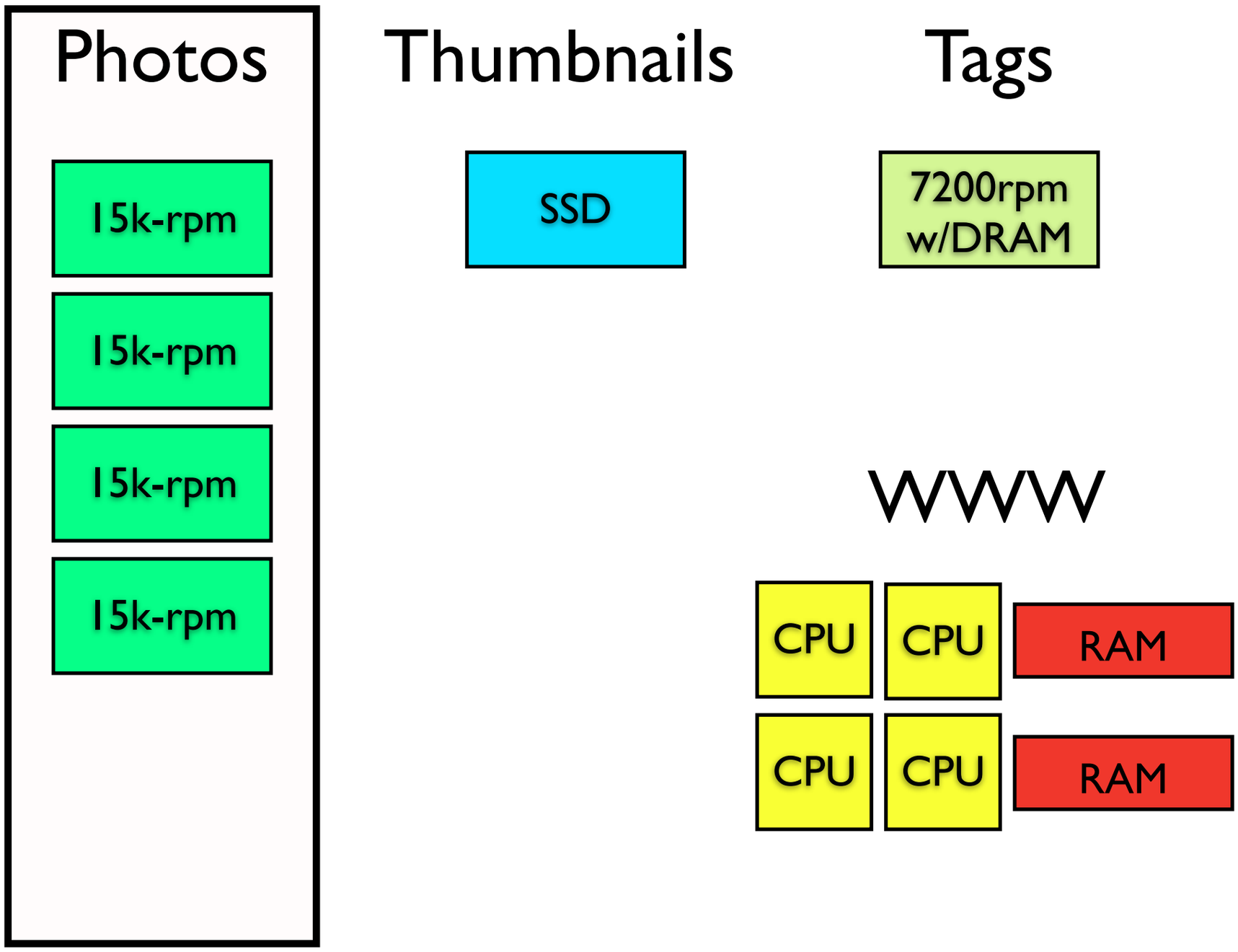
7200rpm w/DRAM

WWW

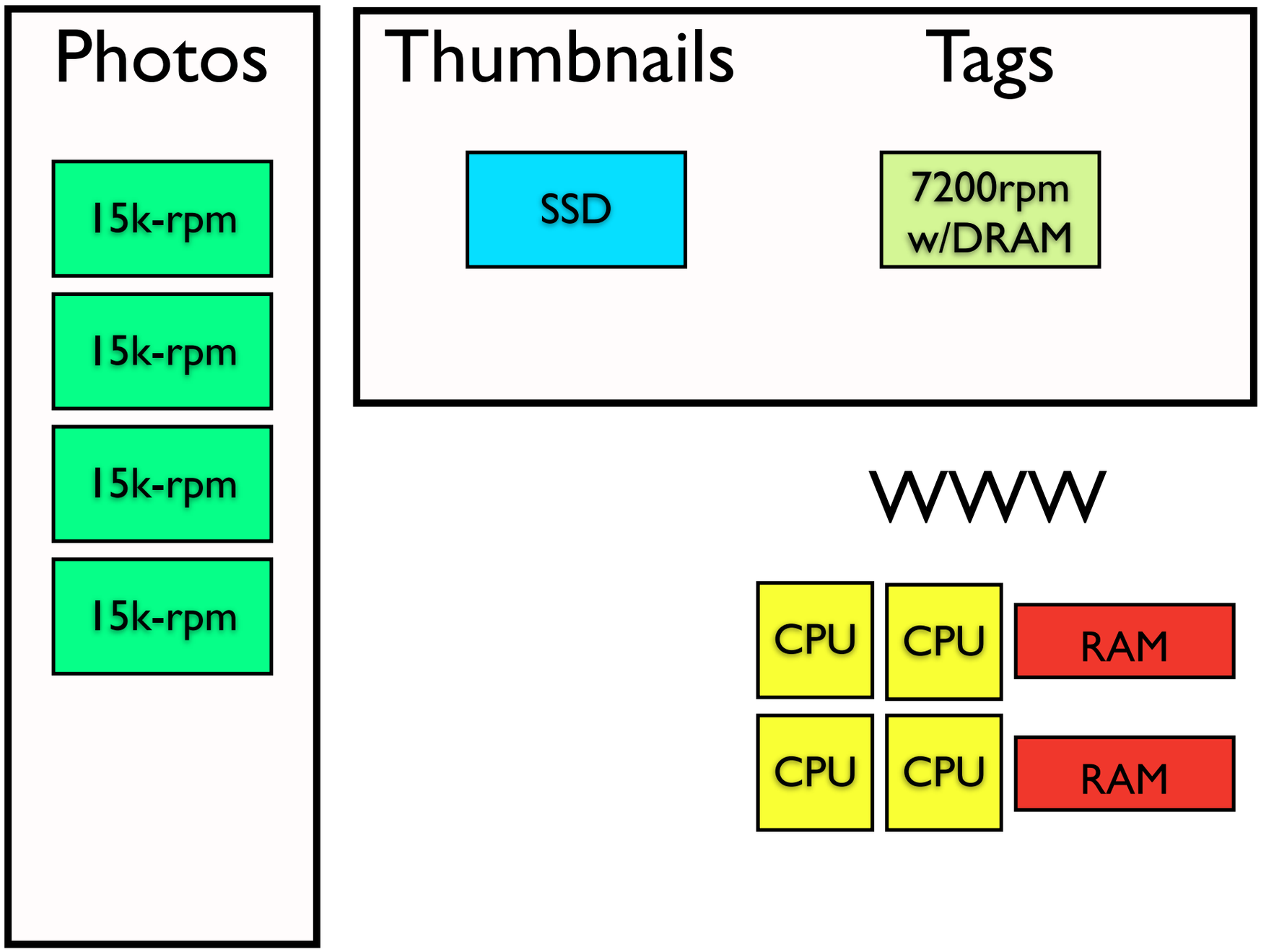
CPU CPU RAM

CPU CPU RAM

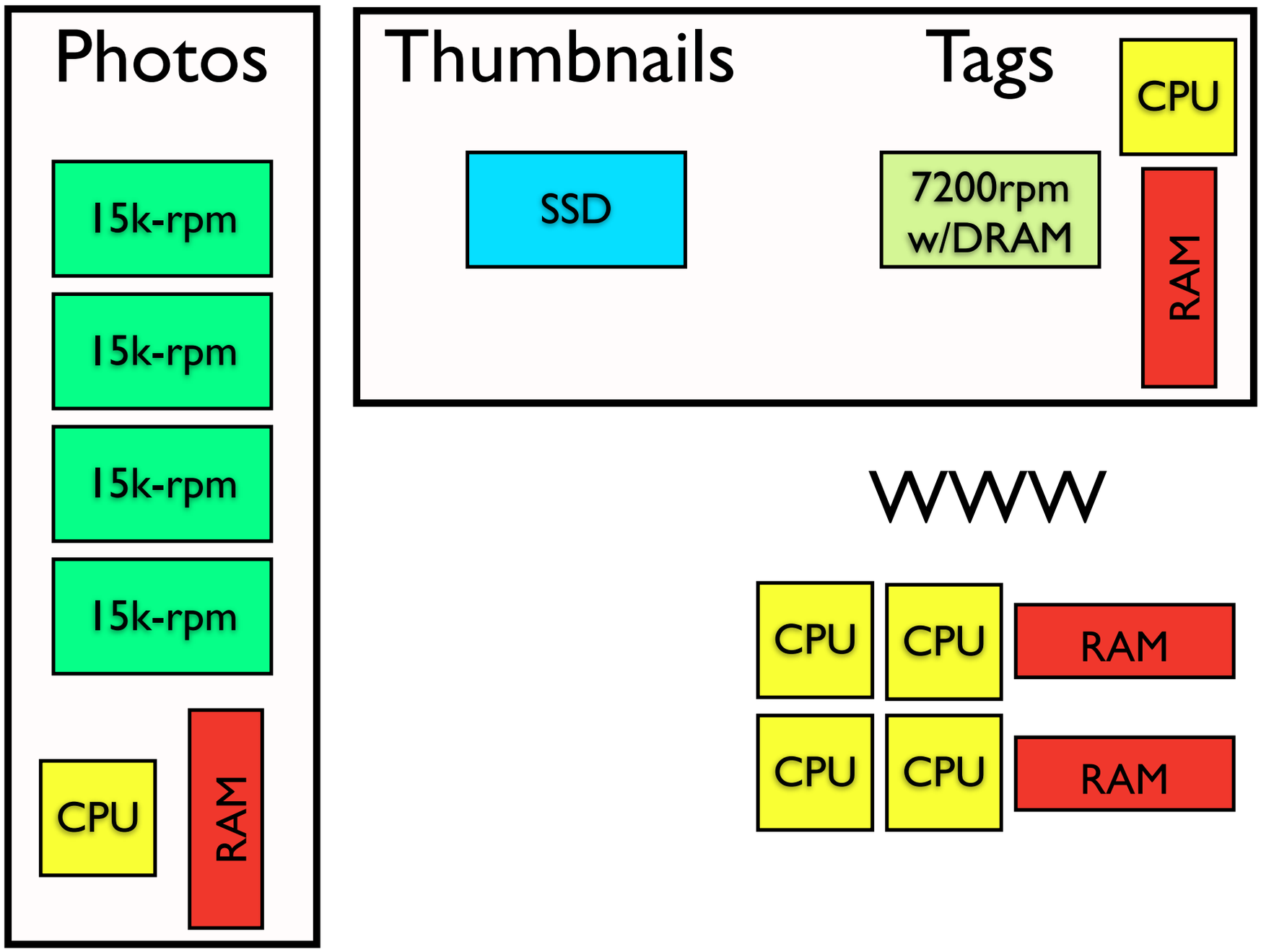
Example: Final Stages for Photo-sharing



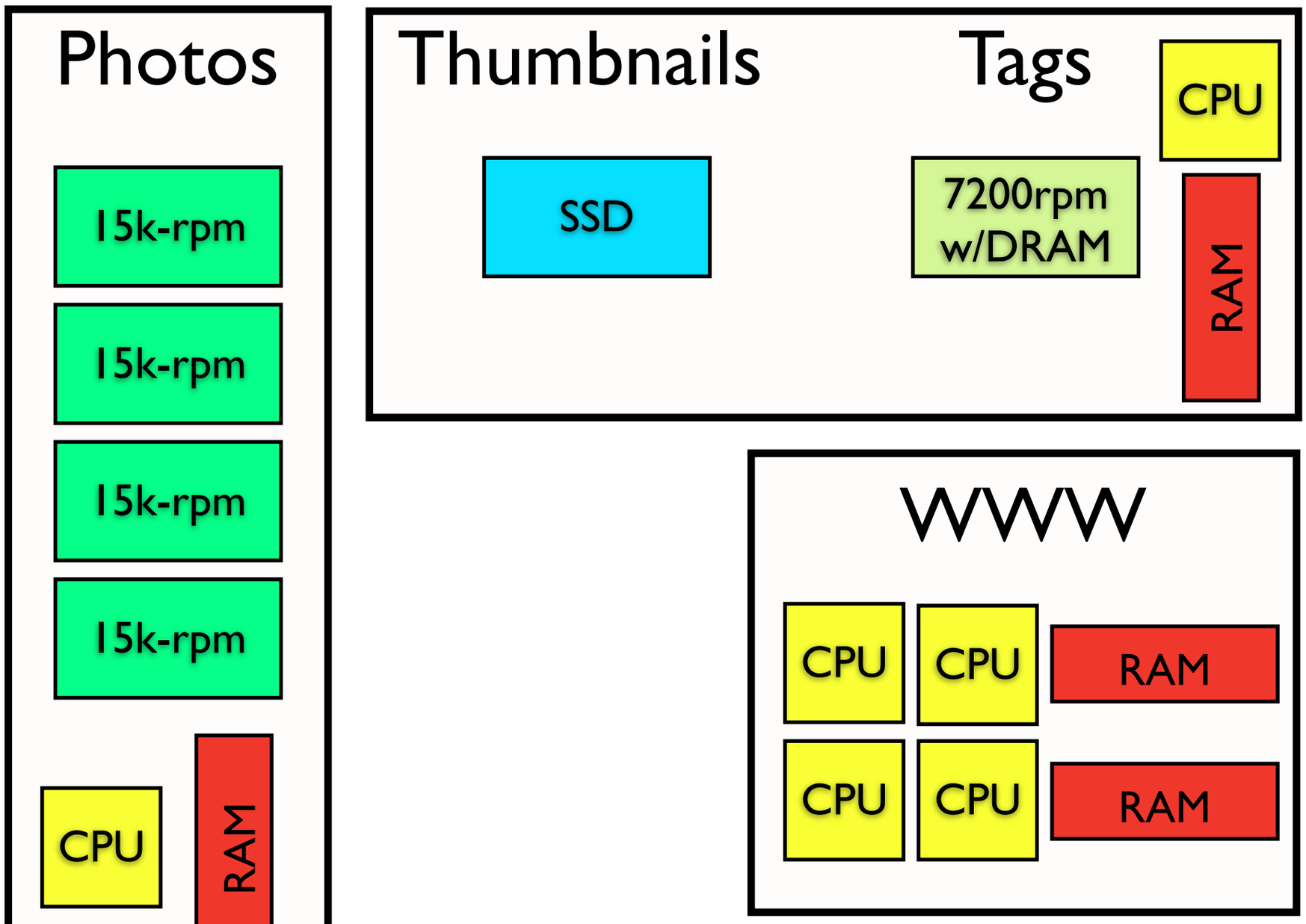
Example: Final Stages for Photo-sharing



Example: Final Stages for Photo-sharing



Example: Final Stages for Photo-sharing



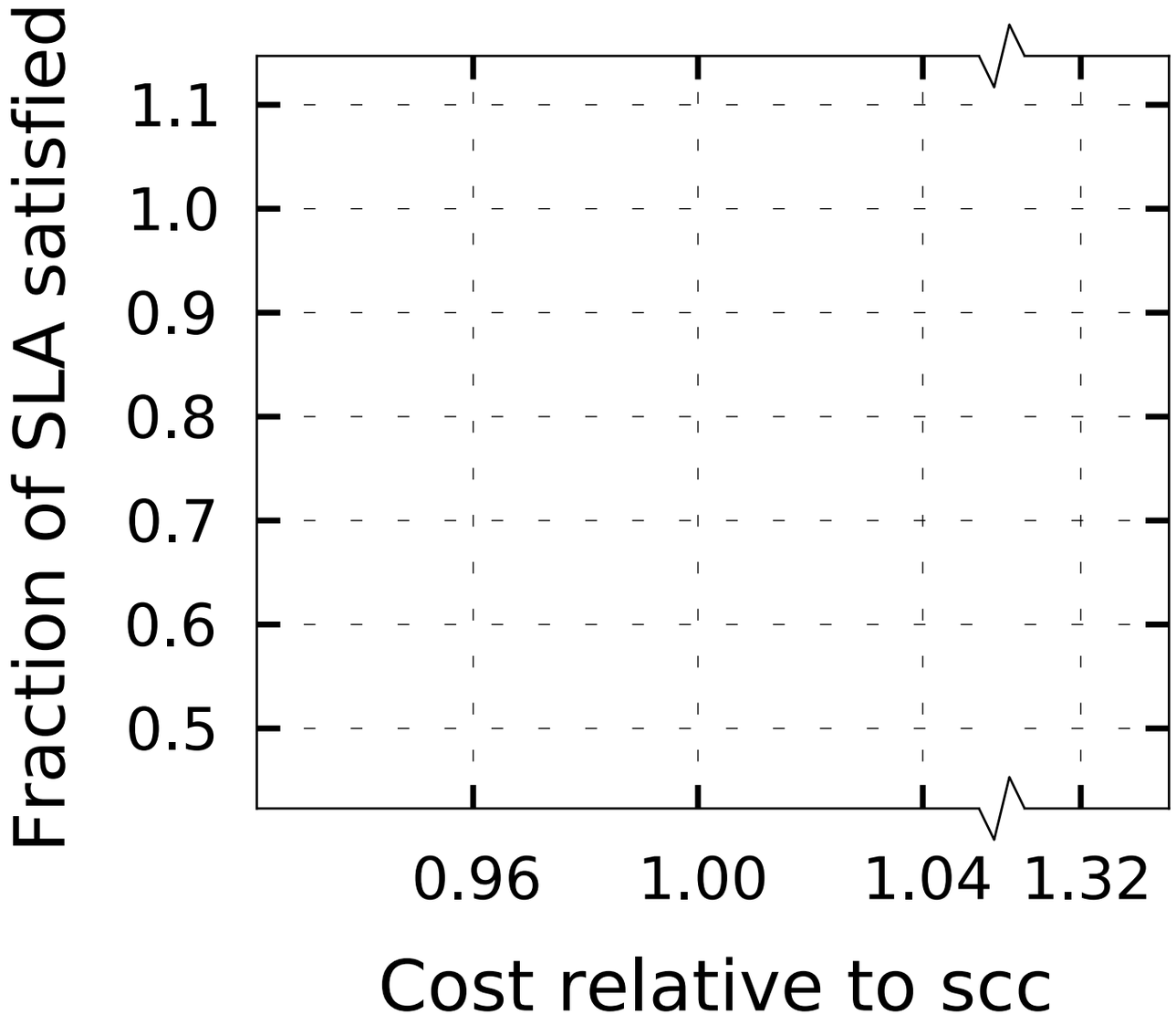
Outline

- Modeling Applications and Hardware
- Computing low-cost configurations
- Example
- Validation
- Applications of scc

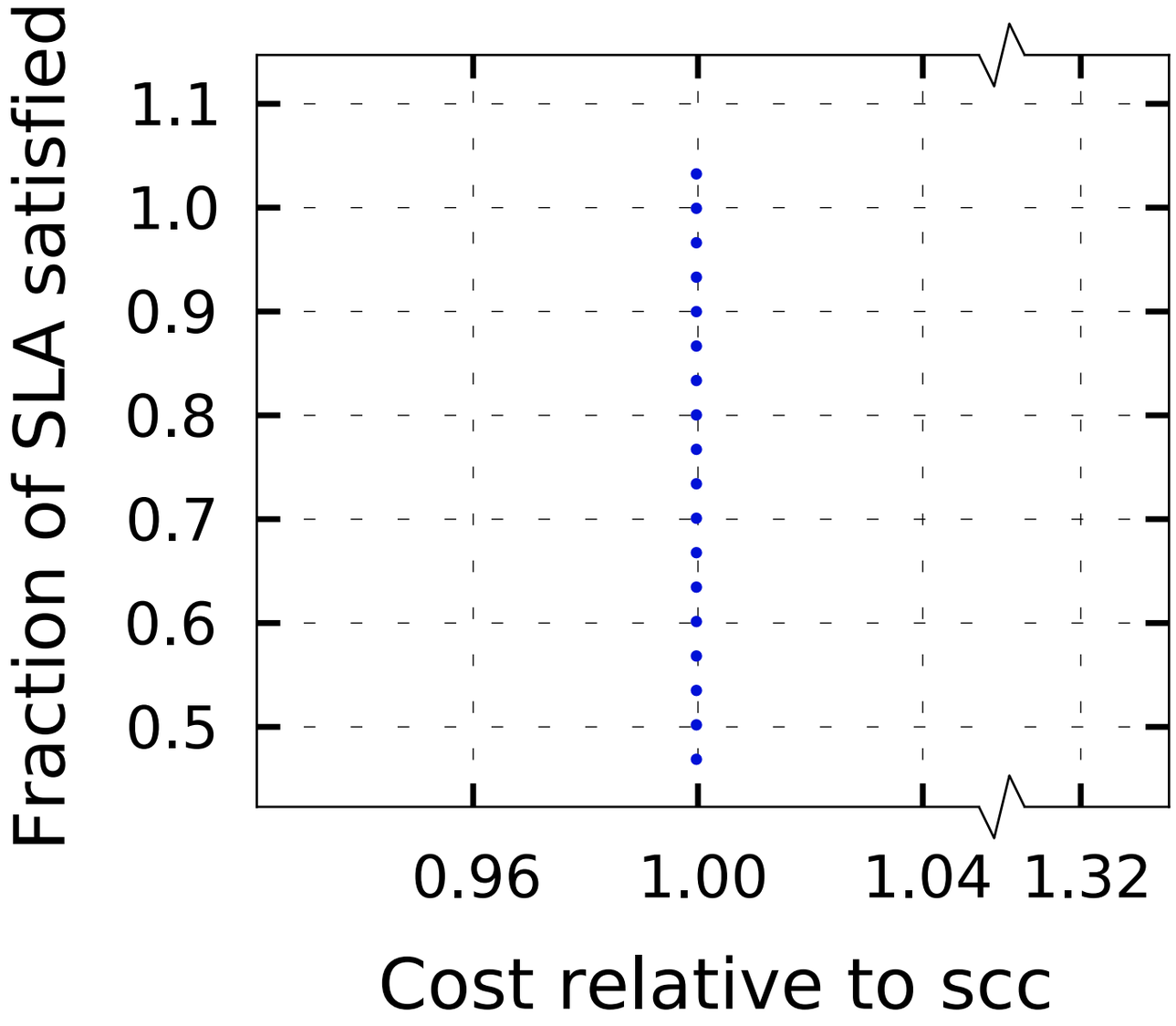
Experimental Methodology

- Built three applications
  - Photo-sharing
  - Product Search
  - Terasort
- Micro-benchmarked each to create model
- Deployed scc output to cluster of machines

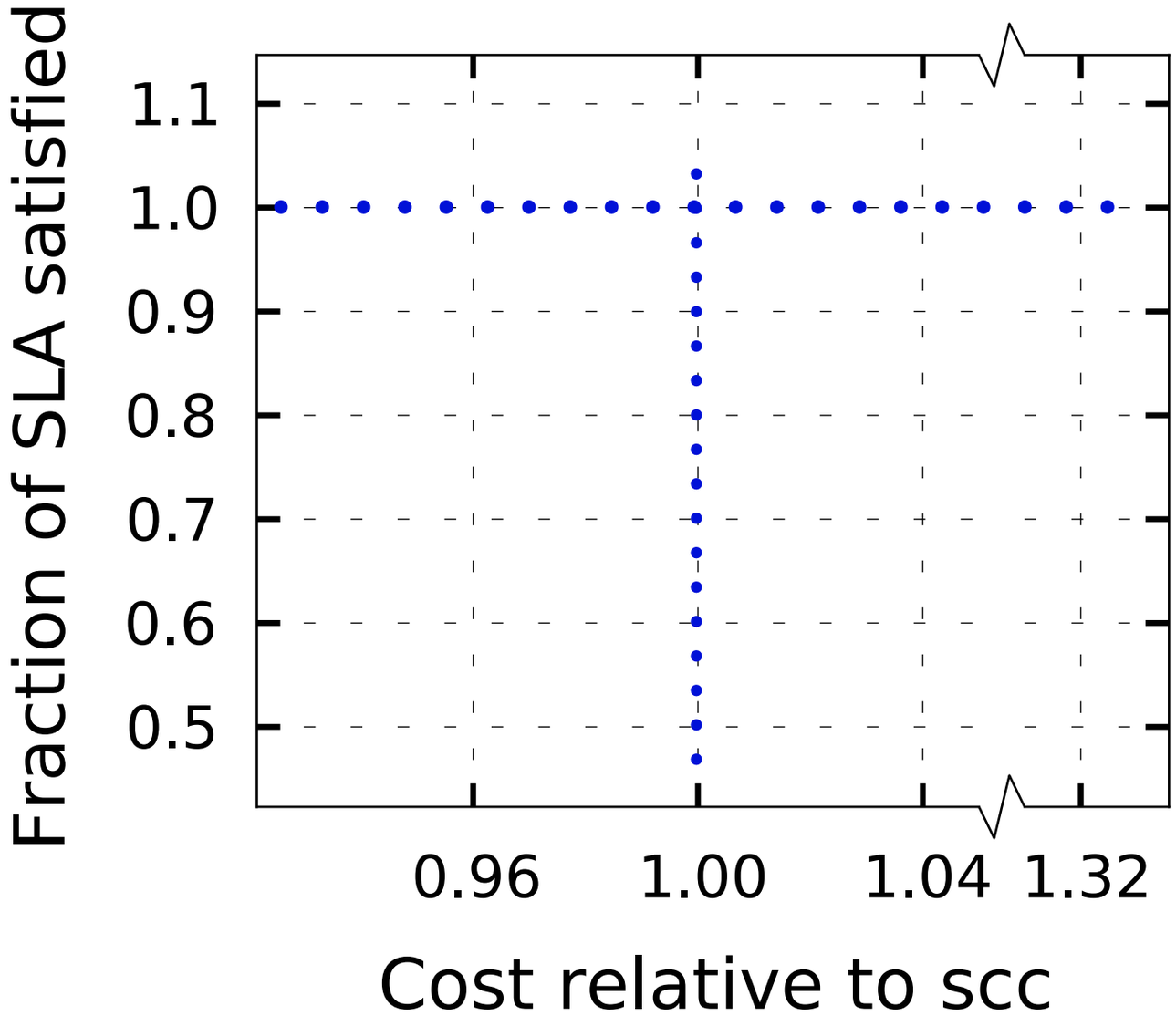
Validation: scc meets SLA at lower cost for Photo-sharing



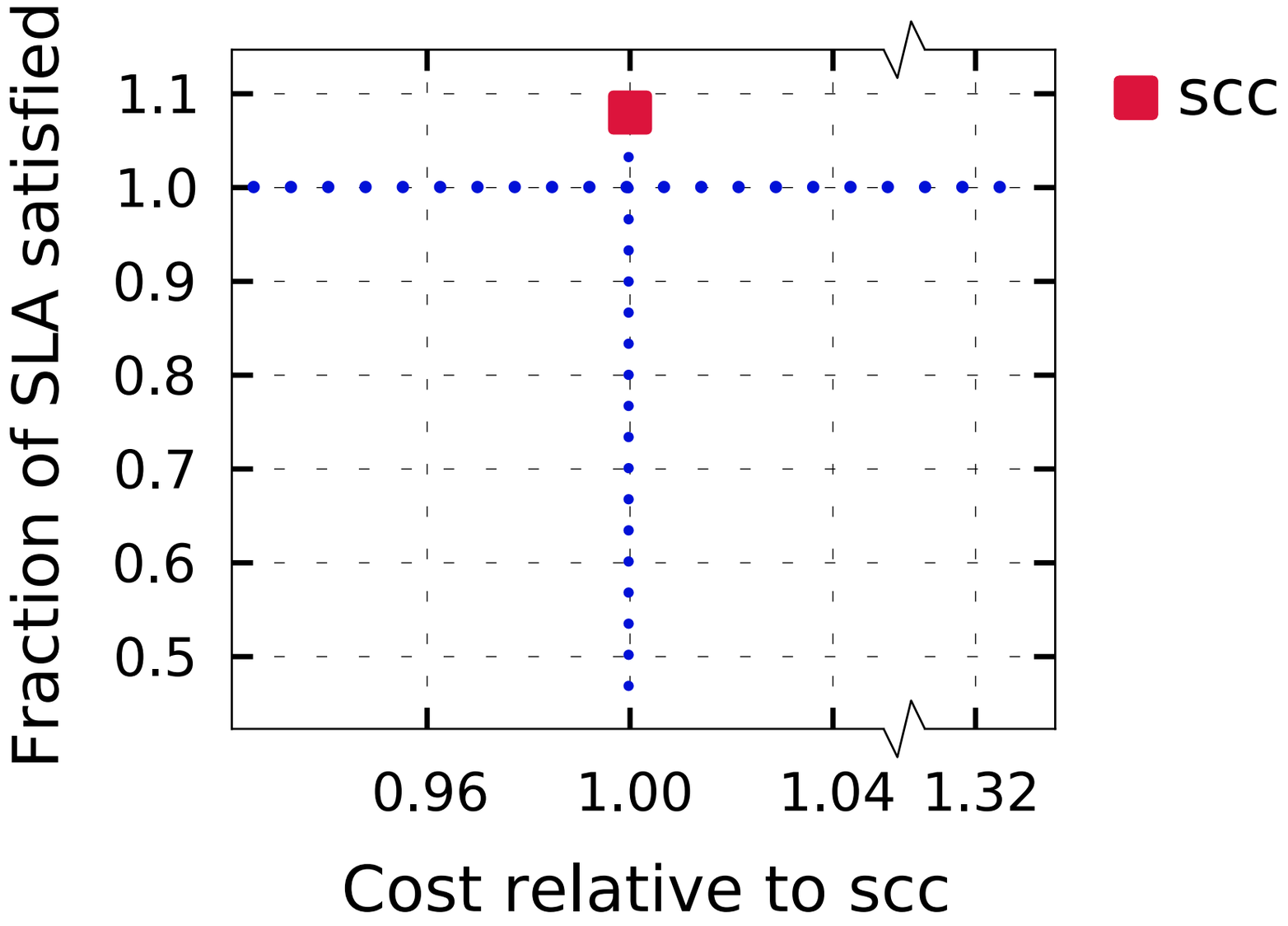
Validation: scc meets SLA at lower cost for Photo-sharing



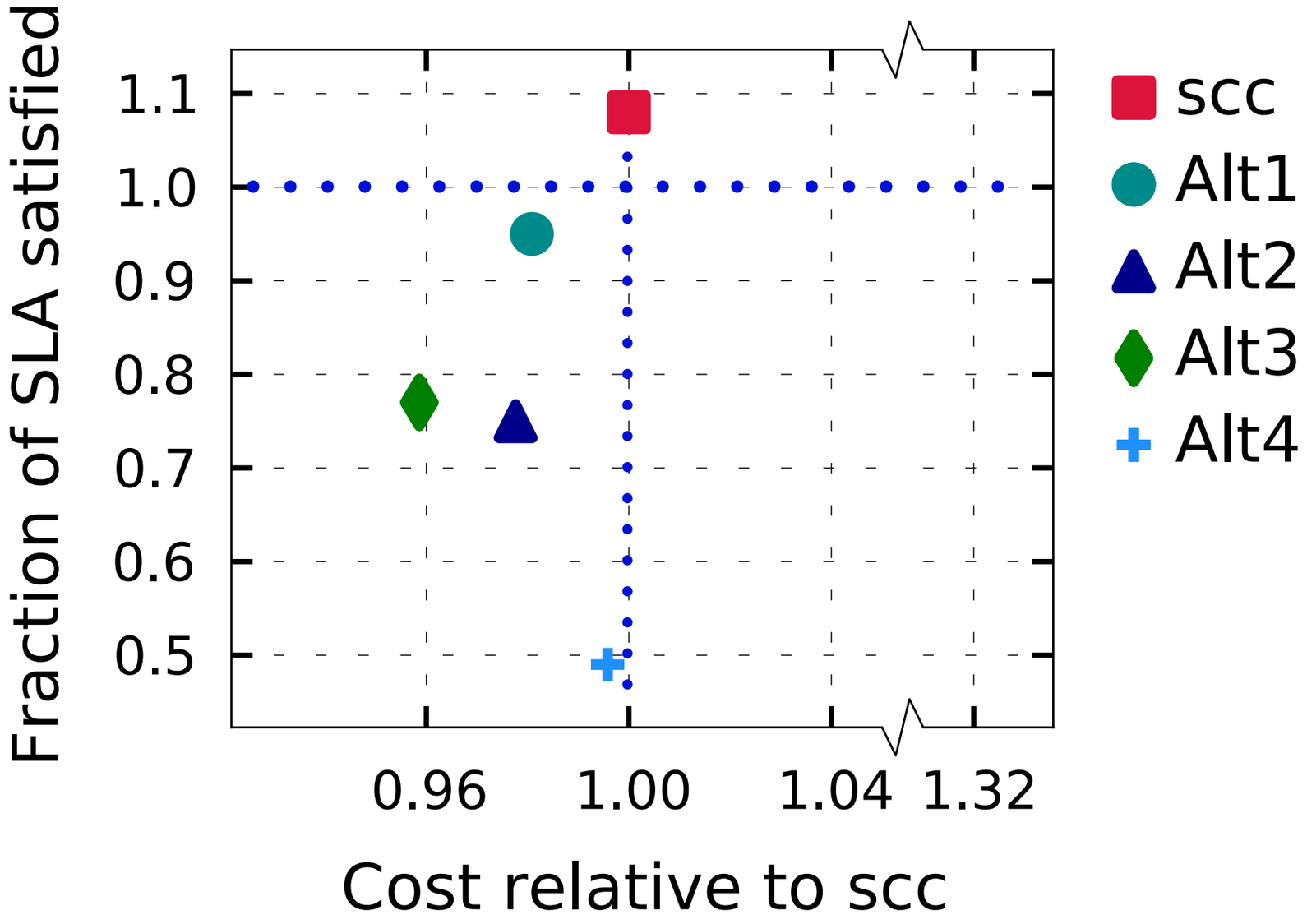
Validation: scc meets SLA at lower cost for Photo-sharing



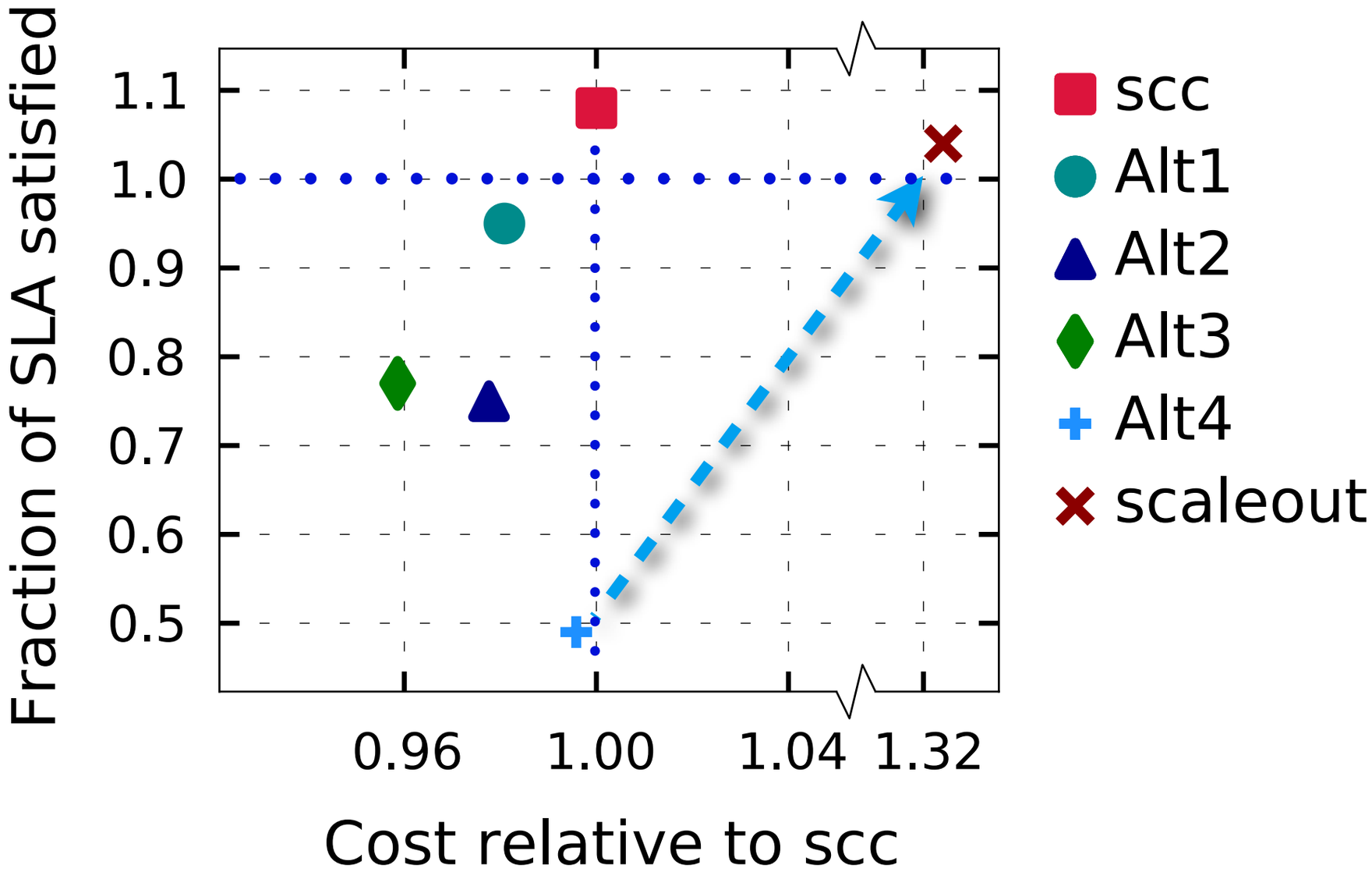
Validation: scc meets SLA at lower cost for Photo-sharing



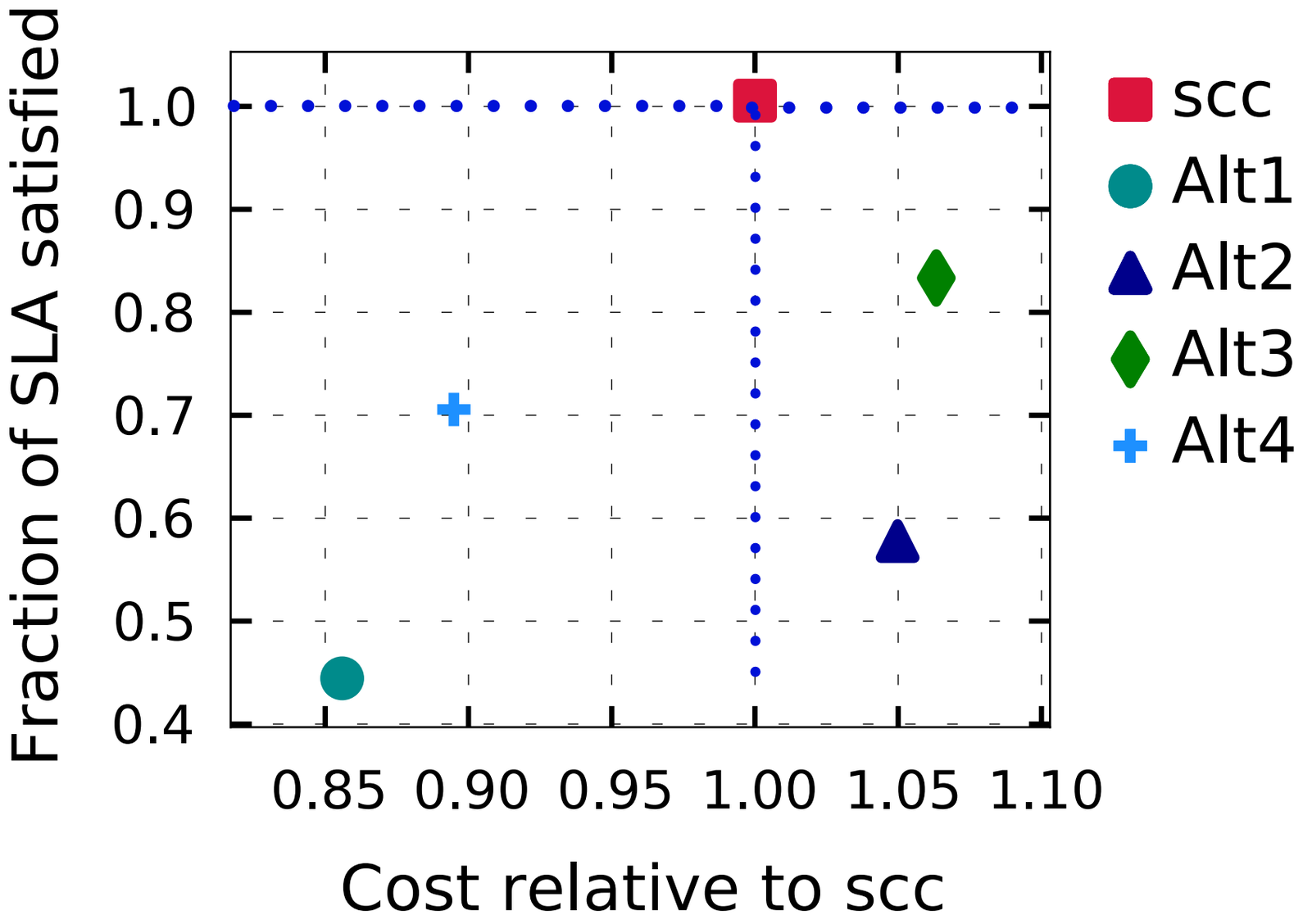
Validation: scc meets SLA at lower cost for Photo-sharing



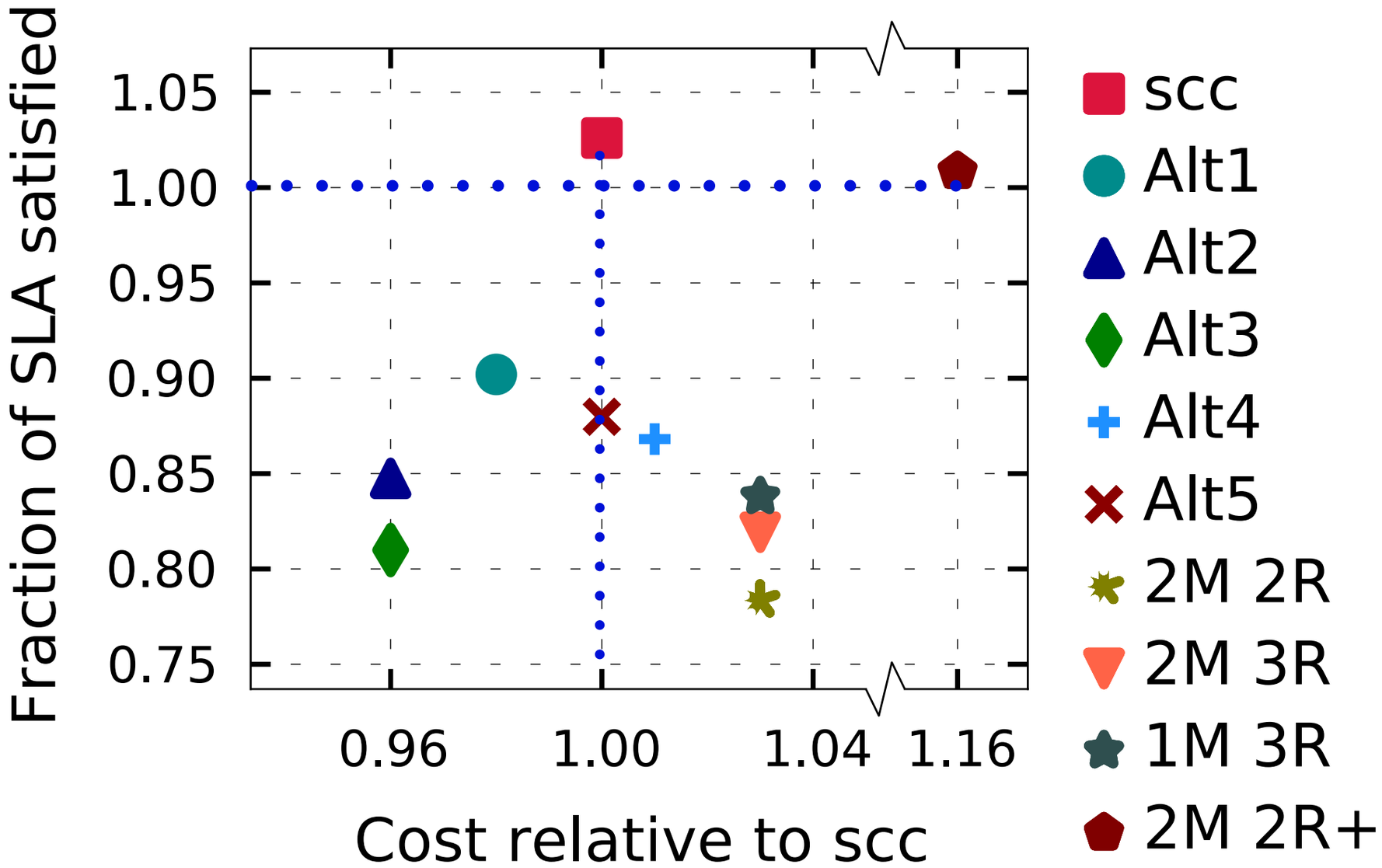
Validation: scc meets SLA at lower cost for Photo-sharing



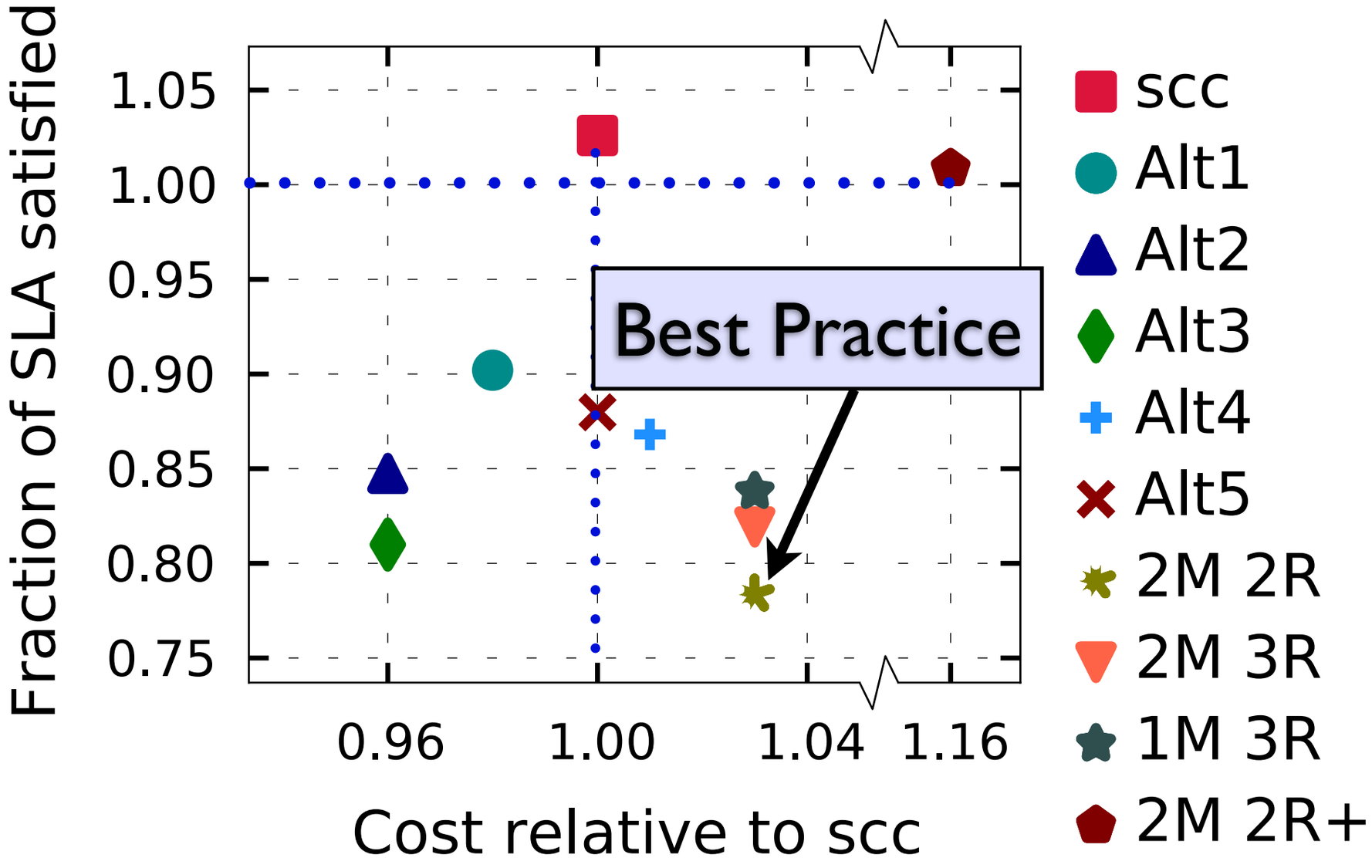
Validation: scc meets SLA at lower cost for Product Search



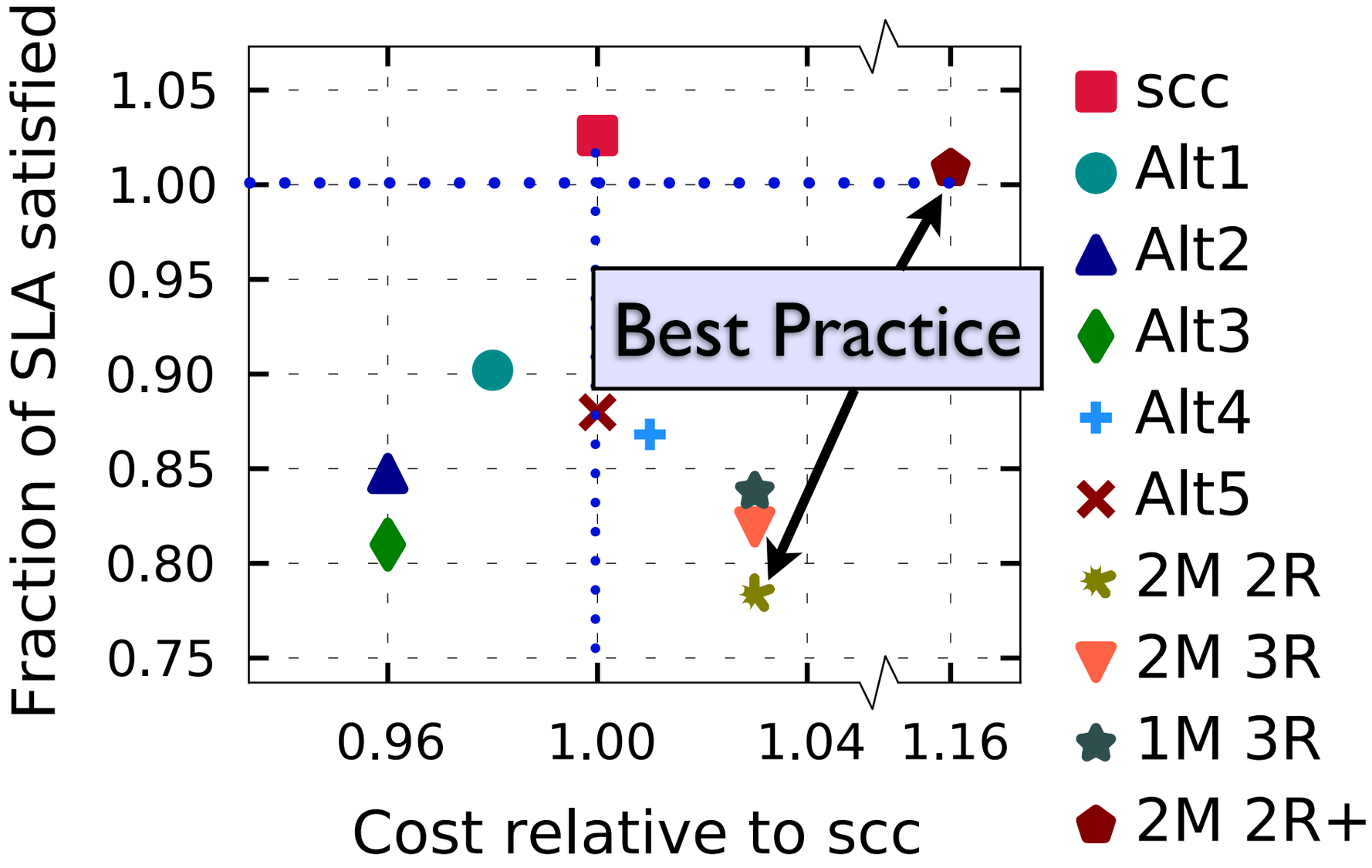
Validation: scc meets SLA at lower cost for Terasort



Validation: scc meets SLA at lower cost for Terasort



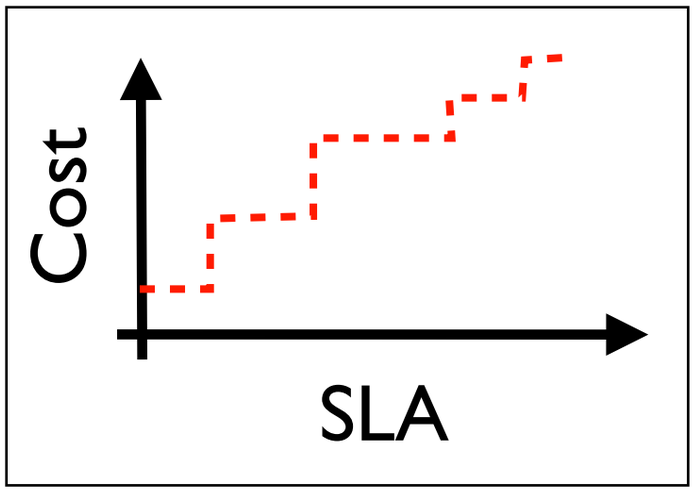
Validation: scc meets SLA at lower cost for Terasort



Outline

- Modeling Applications and Hardware
- Computing low-cost configurations
- Experimental methodology
- Validation
- Applications of scc

Storage Type Regimes



Uploads/s	Photos	Thumbnails	Tags
<5	Disk	Disk	Disk
5-25	Disk	Disk	Disk w/ DRAM
25-330	Disk	SSD	Disk w/ DRAM
330-930	SSD	Disk w/ DRAM	Disk w/ DRAM
930-10k	Disk w/ DRAM	Disk w/ DRAM	Disk w/ DRAM

Output Sensitivity to Model Parameters

Attribute	Range with same architecture
Photo size	50kB← 200kB→ 850kB
Thumbnail size	1kB← 4kB→ 30kB
SSD unit price	$200← $450→ $900

Output Sensitivity to Model Parameters

Attribute	Range with same architecture
Photo size	50kB← 200kB→ 850kB
Thumbnail size	1kB← 4kB→ 30kB
SSD unit price	$200← $450→ $900

Dataset	Sensitive to what hardware cost?
Photos	20% drop in 7.2k-rpm drive price
Thumbnails	92% drop in DRAM price
Tags	31% drop in 15k-rpm drive price

Conclusion

Conclusion

- Our scc tool finds low-cost cluster configurations
  - Saves 2-4.5x over simple scale-out for Photo-share and Product Search
  - 16% lower cost for Terasort vs. Map Reduce best-practice configuration

Conclusion

- Our scc tool finds low-cost cluster configurations
  - Saves 2-4.5x over simple scale-out for Photo-share and Product Search
  - 16% lower cost for Terasort vs. Map Reduce best-practice configuration
- Better matches hardware to application requirements

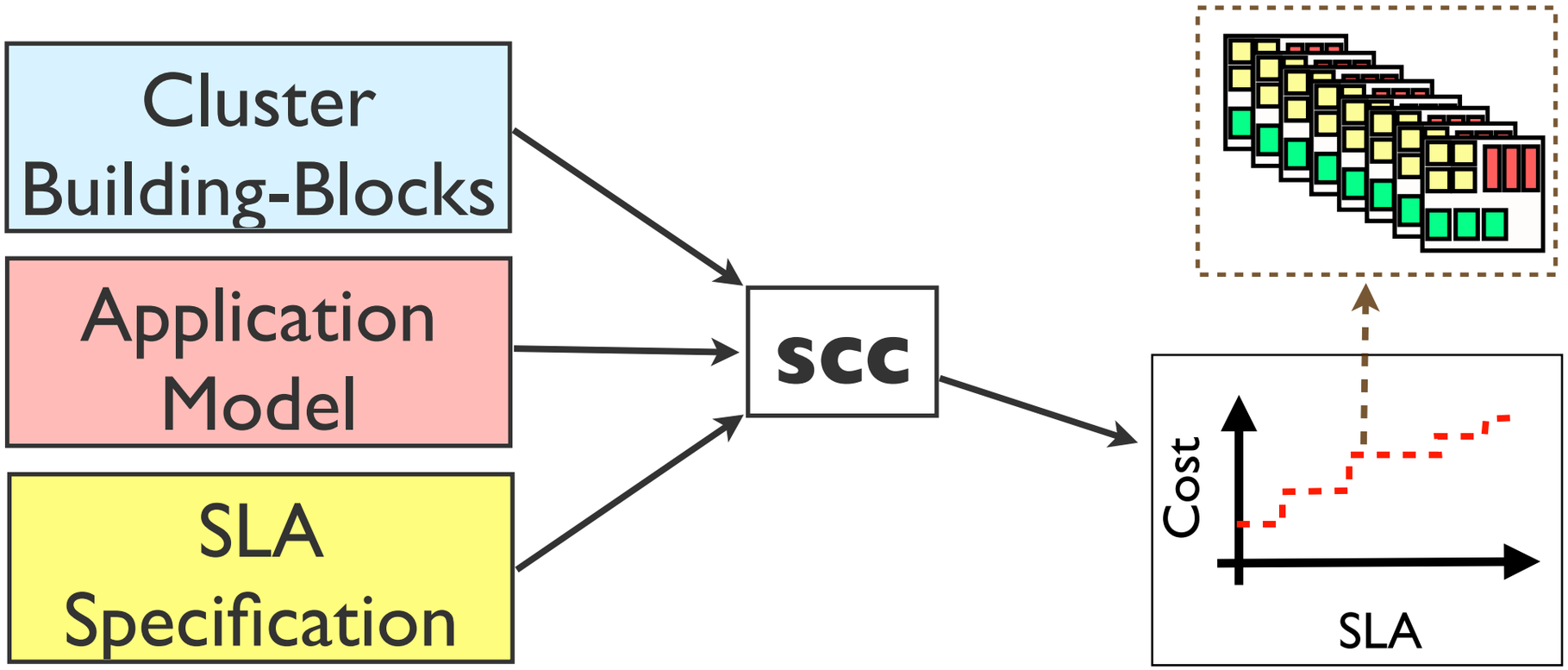
Conclusion

- Our scc tool finds low-cost cluster configurations
  - Saves 2-4.5x over simple scale-out for Photo-share and Product Search
  - 16% lower cost for Terasort vs. Map Reduce best-practice configuration
- Better matches hardware to application requirements
- Useful for predicting output sensitivity and future needs

Conclusion

- Our scc tool finds low-cost cluster configurations
  - Saves 2-4.5x over simple scale-out for Photo-share and Product Search
  - 16% lower cost for Terasort vs. Map Reduce best-practice configuration
- Better matches hardware to application requirements
- Useful for predicting output sensitivity and future needs
- Future work
  - More precise network models
  - Cloud deployment

Questions?



UC RIVERSIDE UNIVERSITY OF CALIFORNIA | Bourns College of Engineering

UCSDCSE Computer Science and Engineering