

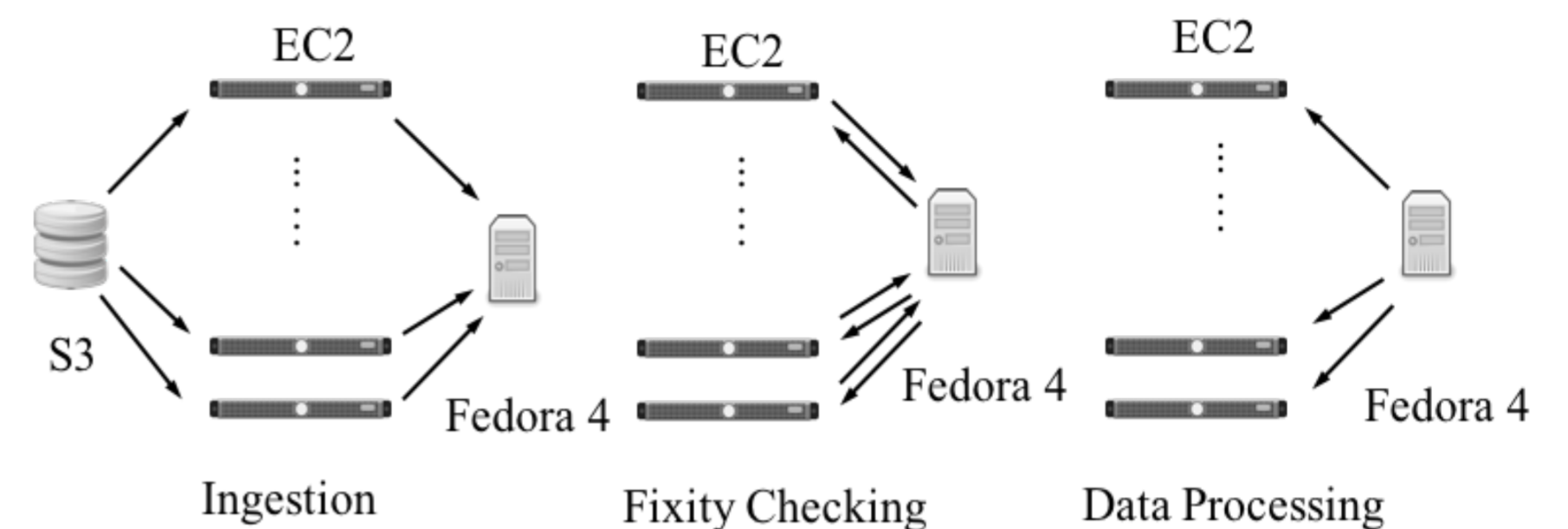
Evaluating Cost of Cloud Execution in a Data Repository

Zhiwu Xie, Yinlin Chen, Julie Speer, and Tyler Walters
University Libraries, Virginia Tech, Blacksburg, VA, USA

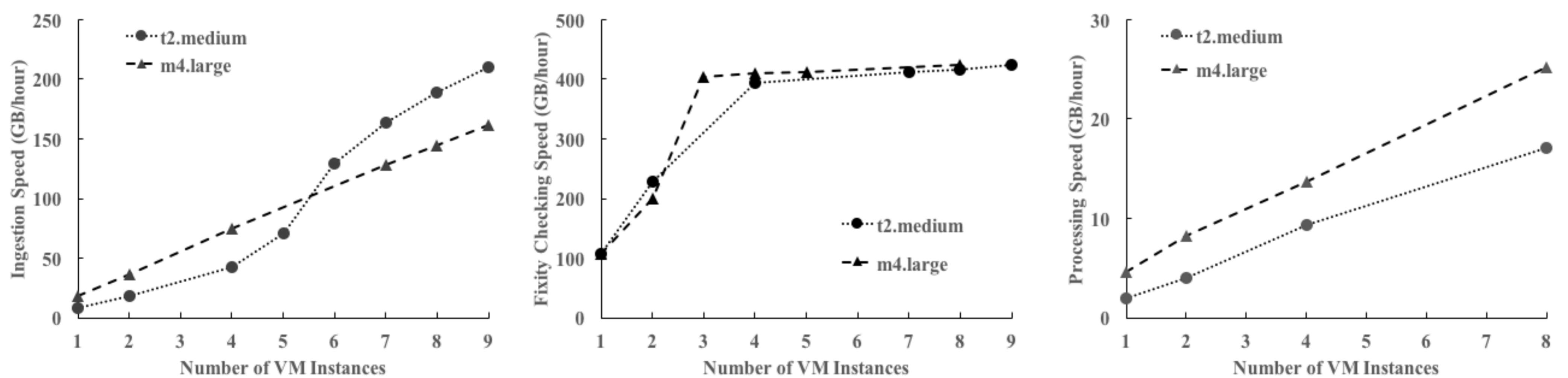
Costs of Cloud Execution

- A component of total cost
- Different from storage and bandwidth costs
- Dependent on computing complexity, compute node capacity, and bandwidth
- Should be measured independently

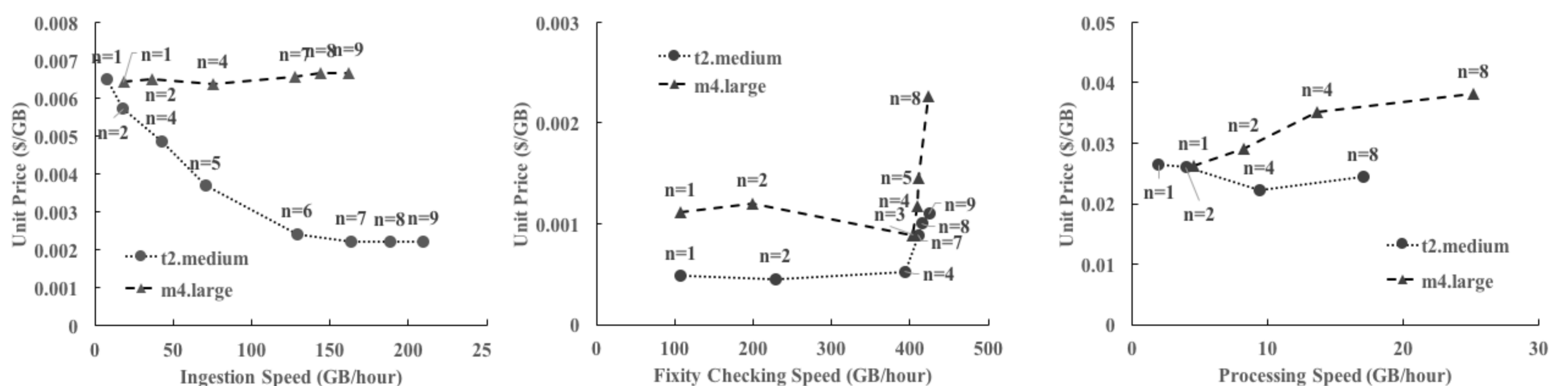
What Did We Measure



Results



Speedup



Unit Cost

Takeaways

- Distributing computing tasks on cheaper, slower instances tends to be more cost effective
- Bandwidth forms the scalability bottleneck and throttles cost effectiveness
- In certain cases, superlinearity may be observed.