A Scalable Master-Worker Architecture for PaaS Clouds

Vibhor Aggarwal, Shubhashis Sengupta, Vibhu Saujanya Sharma, Aravindan Santharam

Accenture Technology Labs
Bangalore, India

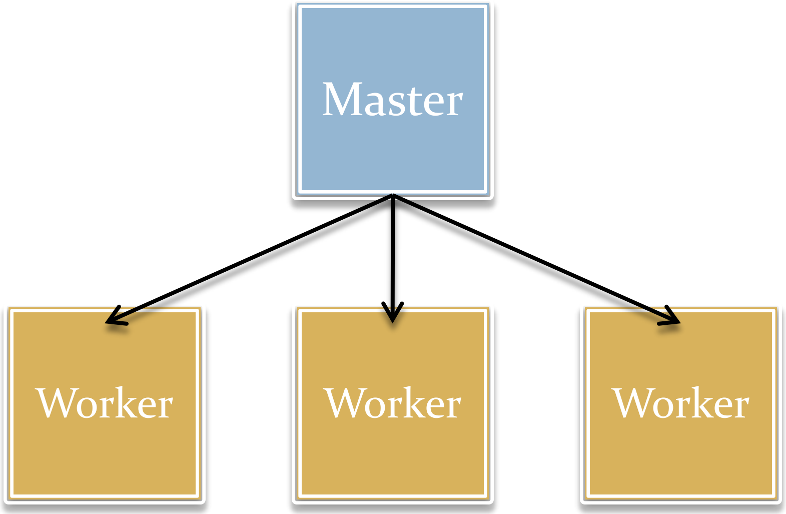
Outline

- Introduction
  - Master-Worker
  - PaaS Cloud
  - Motivation
- Methodology
  - Agent-based Architecture
  - Application: Animation Rendering
  - Platform: Heroku
- Results
- Conclusions and Future Work

Introduction

Introduction

- Master-Worker Architecture
  - Loosely-coupled parallel tasks
  - Centralized queue
  - MTC Applications, Batch Systems, Parallel Workflows, Map Reduce
  - Limitations:
    - Single point of failure
    - Scalability



Introduction

- Platform as a Service (PaaS) Clouds
  - Basic software stack – OS, run-time, DB, Logging
  - Host you own application in Container
  - On demand scaling
  - Pay-as-you-go
  - Limitations:
    - Ephemeral storage
    - Volatitlity

Motivation

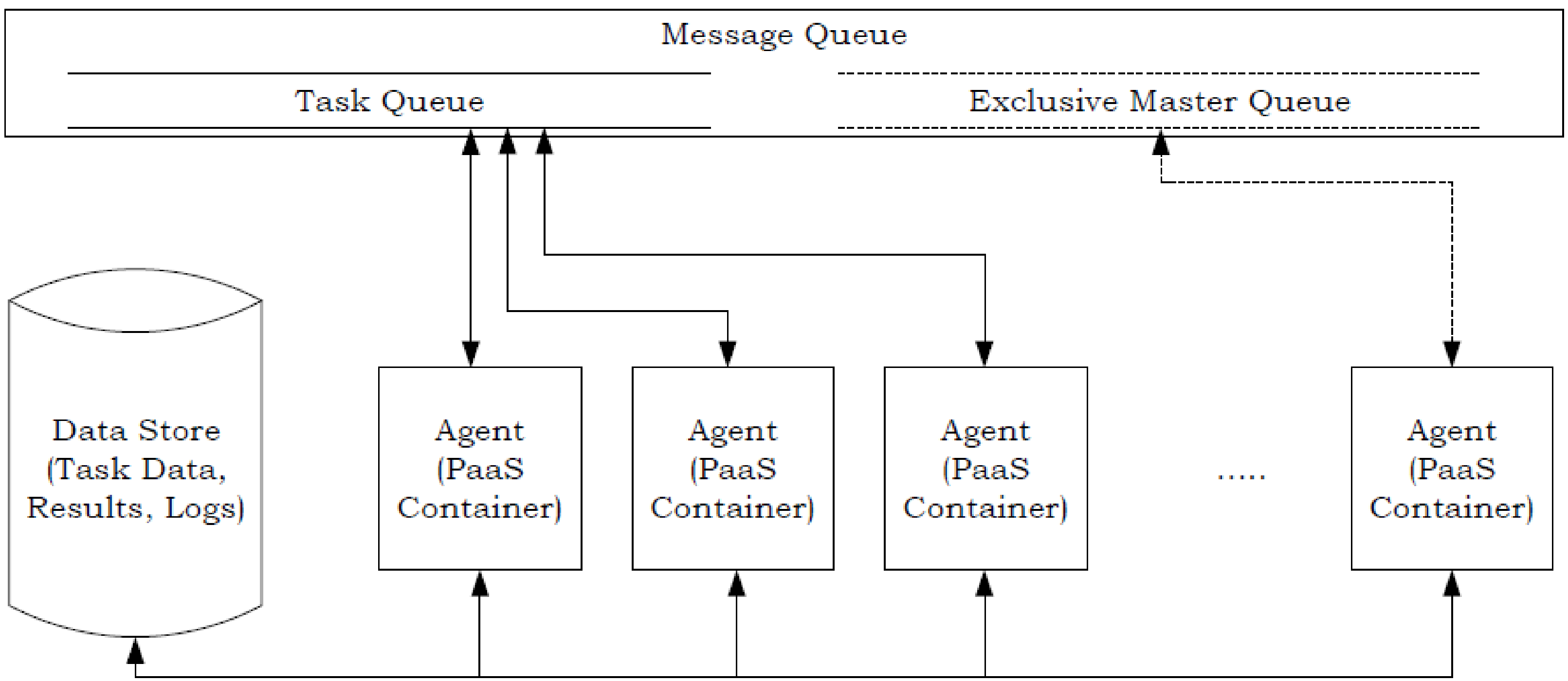
- Migrate Master-Worker Applications to PaaS
  - Indirect communication
    - Message Queues
  - Move global state information outside container

Methodology

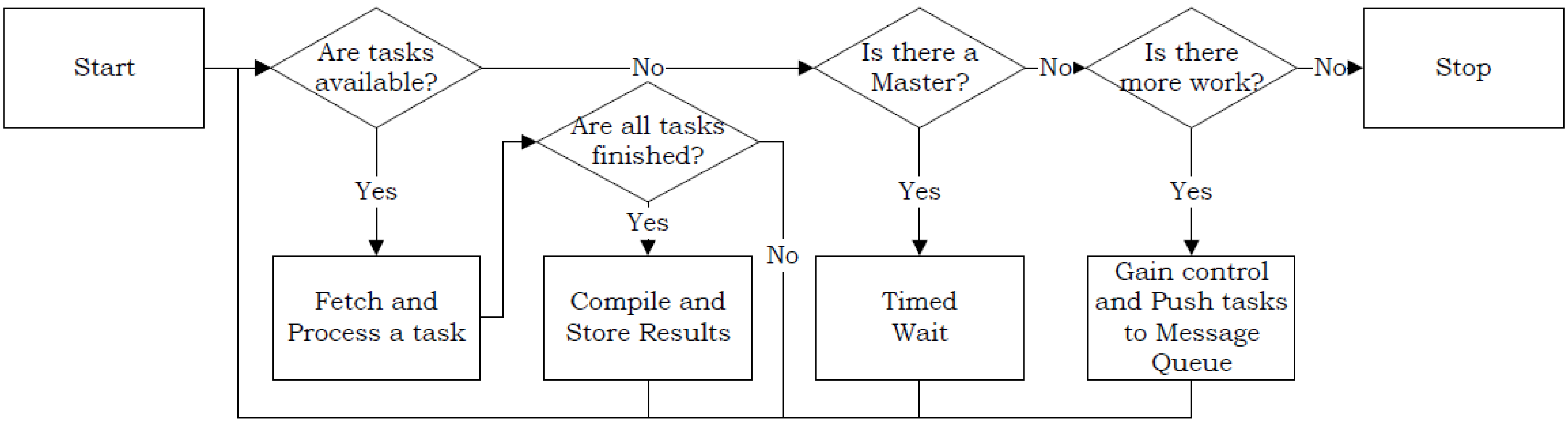
Agent-based Master-Worker

- Global Information in Message Queue
- Simple Reflex Agent
- Dynamically switch between Master/Worker

Master-Worker on PaaS



Agent-based Master-Worker



Application: Animation Rendering

- High-fidelity Rendering of Animations
  - Image generation from physically-based quantities
- Computationally Intensive
  - Monte-Carlo Process
  - Loosely-coupled Parallel Tasks
- Many-Task Computing Application


Mobil 1
KÄRCHER
D2
DUNLOP
DUNLOP
Mobil 1
mannesmann
Mobil 1

Platform: Heroku

- Production Environment
- Support for long asynchronous jobs
- Application container – Dyno
  - Cost - $0.05/Hour
- Addon services
  - MongoDB and RabbitMQ
- Hosted on Amazon EC2

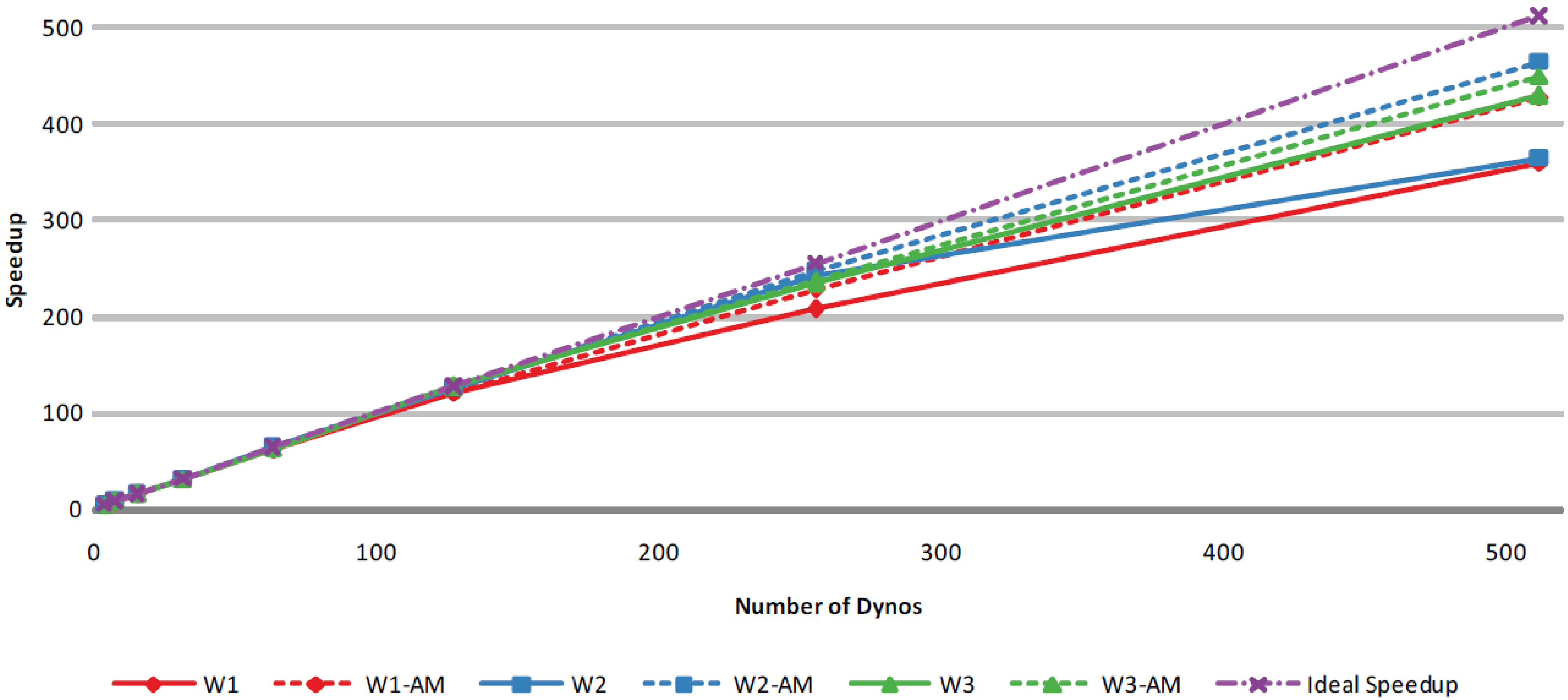
Workloads

- Animation Rendering (W1, W2)
  - 800 x 600 Frame
  - 200 Tasks per Frame
  - Path Tracing
- Nth Prime Number Calculation (W3)
  - Five types

	Min Task Time (in s)	Max Task Time (in s)	Total Time (in Hr)	Total Tasks
w1	1	460	240	24,000
w2	20	220	400	18,000
w3	20	60	190	18,000

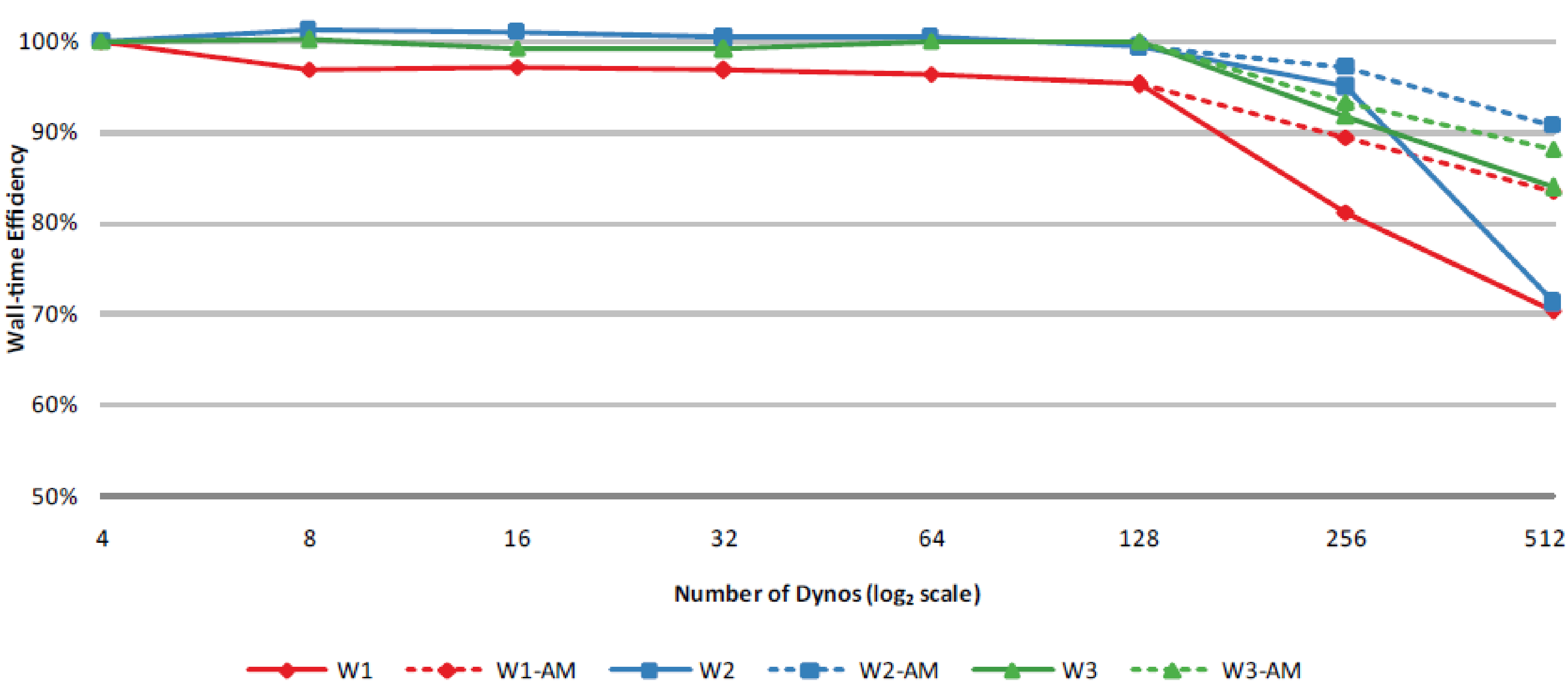
Results

Speed-up

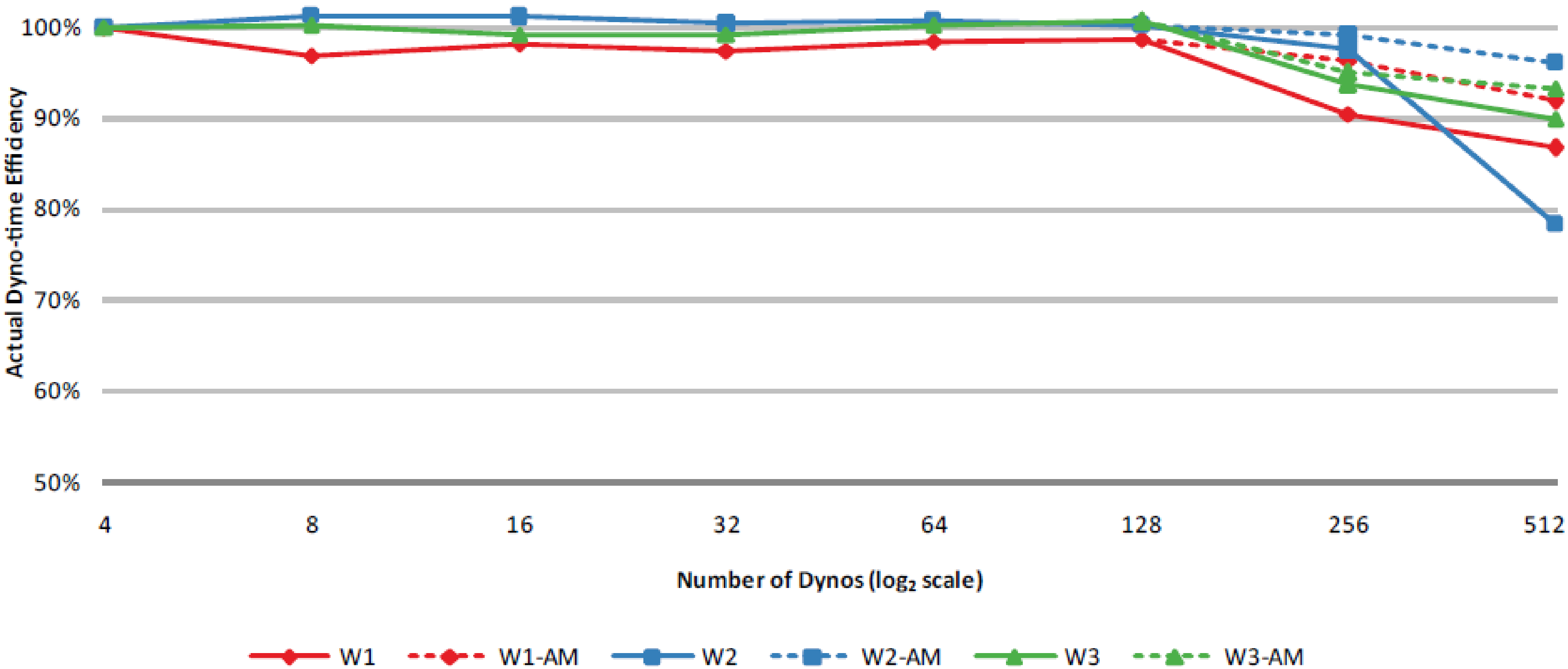


Speedup
500
400
300
200
100
0
0
100
200
300
400
500
Number of Dynos
W1
W1-AM
W2
W2-AM
W3
W3-AM
Ideal Speedup

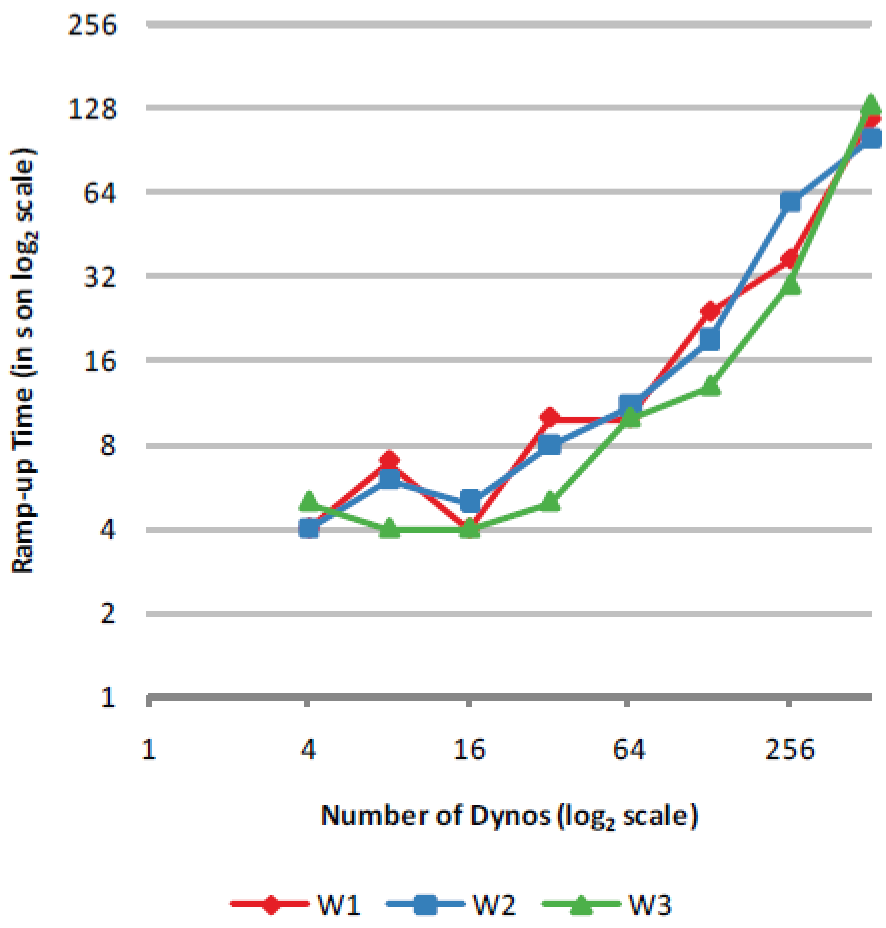
Efficiency: Wall-time



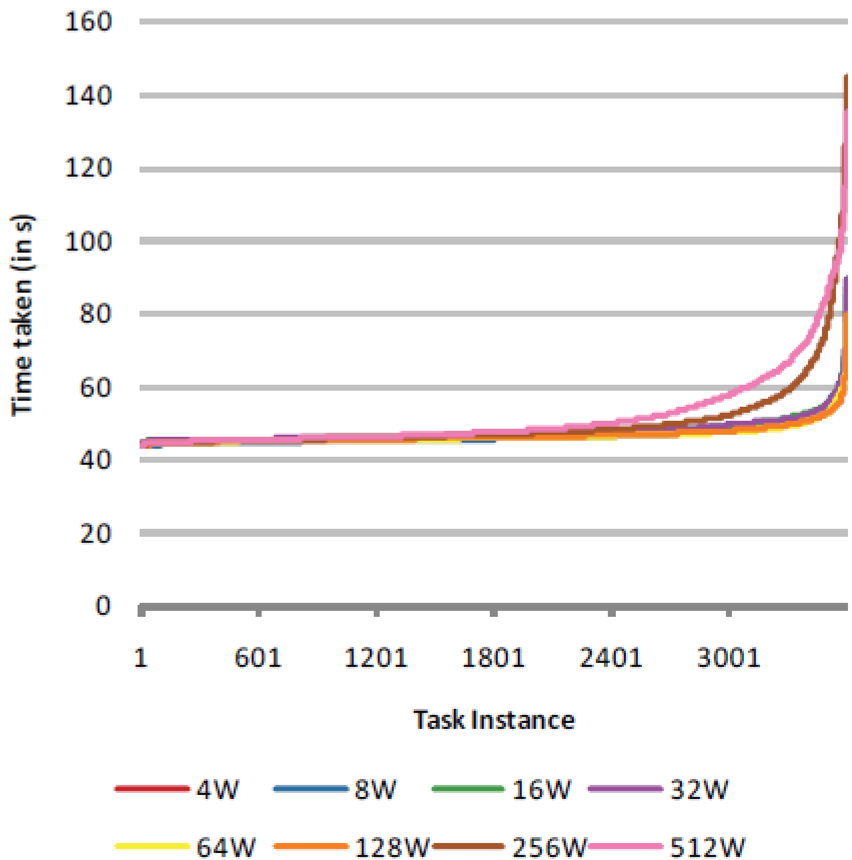
Efficiency: Actual Dyno-time

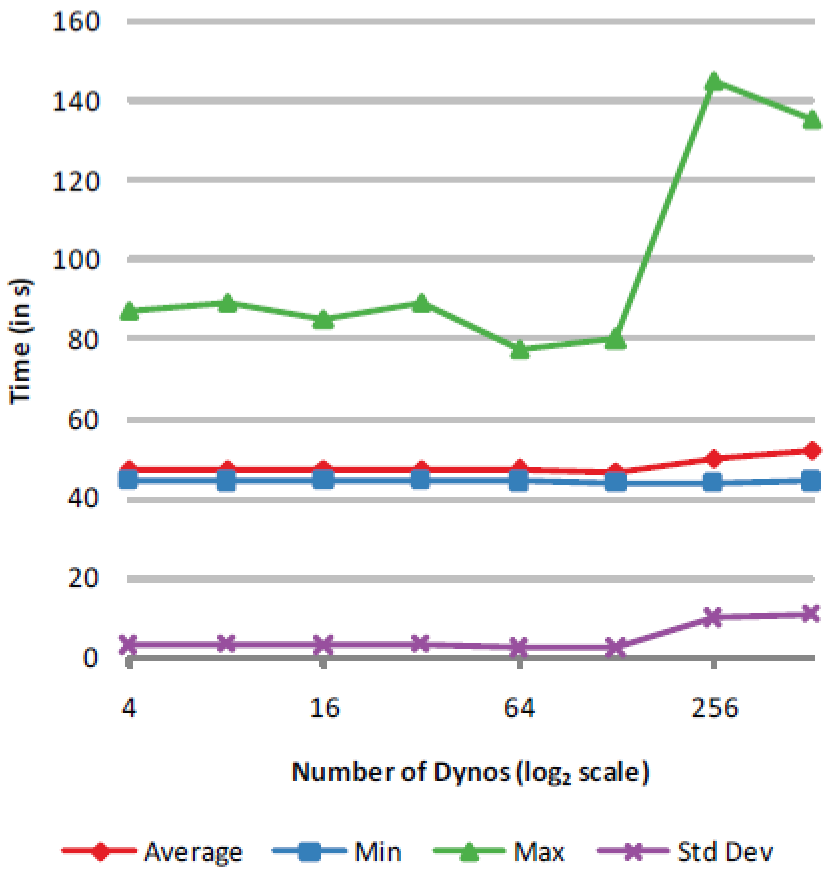


Ramp-up Times

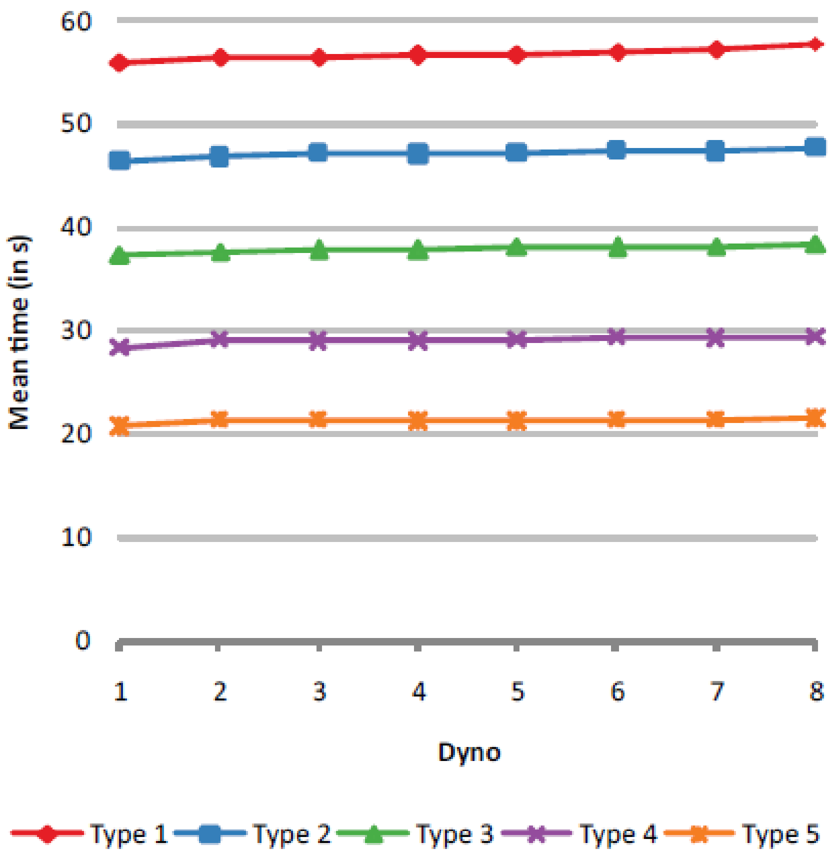


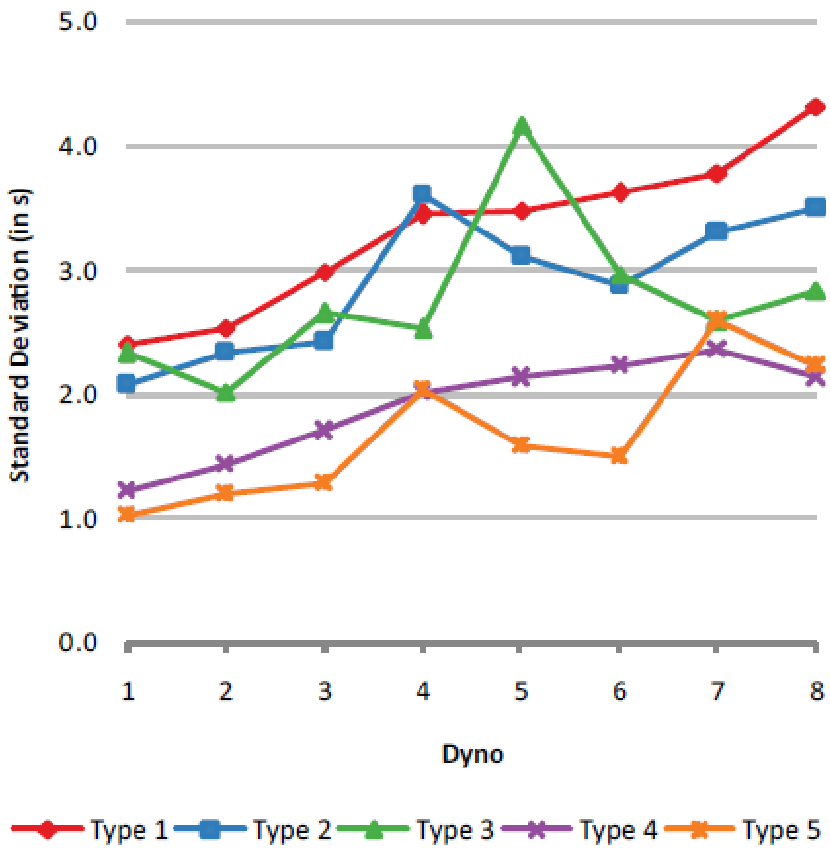
Task Execution Times: W3 Type 2





Task Execution Times: W3





Conclusions and Future Work

Conclusion and Future Work

- Novel Master-Worker Architecture for PaaS
- 400 Hr workload in 52 minutes
  - 90% Wall-time Efficiency
- Effects of Multi-tenant PaaS infrastructure
- PaaS suitability for Master-Worker

- Alternate Message–brokers/Platforms
- Extension for Map Reduce

Thank You

- Contact:
  - Vibhor.Aggarwal@accenture.com

- Acknowledgements
  - Marko Dabrovic for Sponza Model
  - Heroku Engineers