Fixed and Market Pricing for Cloud Services

Vineet Abhishek

ECE, University of Illinois at Urbana-Champaign

NetEcon 2012

Joint work with Ian A. Kash and Peter Key

Microsoft Research Cambridge

Cloud Computing

- Cloud computing:
  - Selling units of computation time on virtual machines.
  - Windows Azure, Amazon EC2, etc.
- How to price cloud services?

The Basic Problem

Two common forms of payment

Pay as you go (PAYG)

Fixed price per unit time per instance

Windows Azure,
Amazon EC2 On-Demand Instances

Spot Market

Auction based pricing, priority to higher bids

Amazon EC2 Spot Instances

Should a provider operate both PAYG and spot market?

PAYG and Spot Market: Is Coexistence Best?

Trade-offs of operating a spot market

Gain

Loss

PAYG and Spot Market: Is Coexistence Best?

Trade-offs of operating a spot market

Gain

- ✓ Low value jobs can compete
- ✓ Revenue gain from jobs who find PAYG price too high

Loss

PAYG and Spot Market: Is Coexistence Best?

Trade-offs of operating a spot market

Gain

✓ Low value jobs can compete

✓ Revenue gain from jobs who find PAYG price too high

Loss

✗ Cheaper alternative to PAYG.

✗ Revenue loss from jobs who would have paid PAYG price

PAYG and Spot Market: Is Coexistence Best?

Trade-offs of operating a spot market

Gain

- ✓ Low value jobs can compete
- ✓ Revenue gain from jobs who find PAYG price too high

Loss

- ✗ Cheaper alternative to PAYG.
- ✗ Revenue loss from jobs who would have paid PAYG price

Net gain from operating a spot market is not clear!

PAYG and Spot Market: Questions

- Resulting equilibrium if jobs can choose b/w the two?
- Revenue comparison with PAYG in isolation?
- Effect of:
  - Spot pricing mechanism?
  - Demand process?

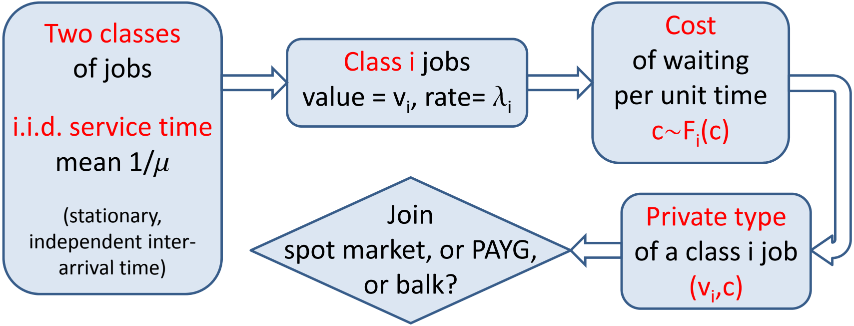
Focus of this talk.

Model

- Jobs: unit demand, associated with a unique user.

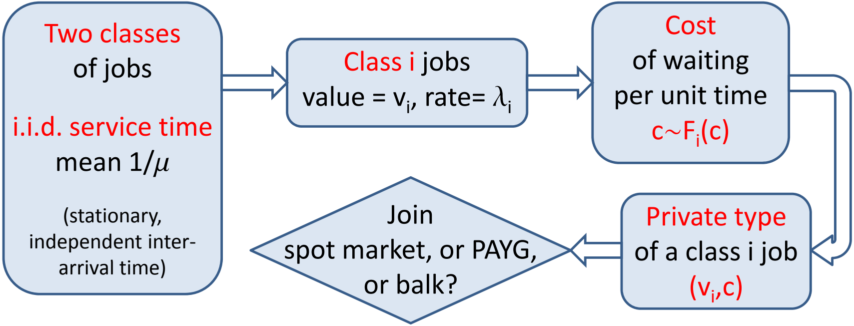
Model

- Jobs: unit demand, associated with a unique user.



Model

- Jobs: unit demand, associated with a unique user.

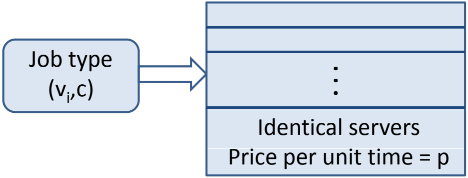


- Payoff = $v_i - cw - m$.

  (Type (v_i, c), waiting time w, payment m) .

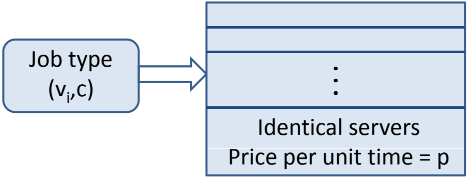
Pay as you go (PAYG)

- $GI/GI/\infty$ system, service rate μ.



Pay as you go (PAYG)

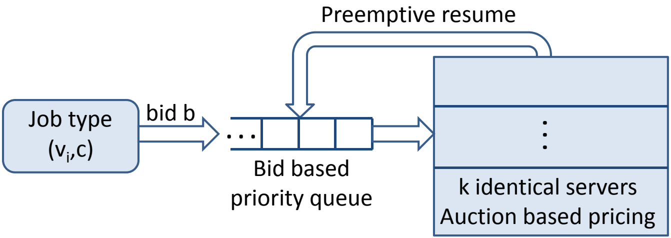
- $GI/GI/\infty$ system, service rate μ.



- Waiting time = service time.
- $\mathbb{E}\,[\text{waiting time}] = 1/\mu$, $\mathbb{E}\,[\text{payment}] = p/\mu$.

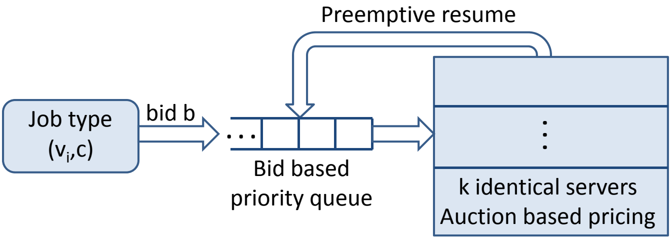
Spot Market

- $GI/GI/k$ system, service rate μ.



Spot Market

- $GI/GI/k$ system, service rate μ.



- Waiting time = queuing delay + service time.
- Assume unobservable queue state.

Characterizing Spot Market

- Spot market characterization:
  - $\mathbb{E}$ [waiting time] and $\mathbb{E}$ [payment] for any bid?
- Difficulties:
  - Pricing mechanism? Bidding strategies?
  - Multiple priority classes, multiples servers.

Characterizing Spot Market

- Spot market characterization:
  - $\mathbb{E}$ [waiting time] and $\mathbb{E}$ [payment] for any bid?
- Difficulties:
  - Pricing mechanism? Bidding strategies?
  - Multiple priority classes, multiples servers.
- Our approach:
  - Truthful reporting of private types as equilibrium bids.
  - Compute $\mathbb{E}$ [payment] in any pricing mechanism.

Key Simplifying Steps

✓ Class independent quantities, given cost c:

- $\mathbb{E}$ [waiting time in spot market].
- $\mathbb{E}$ [payment in spot market].

$\boxed{\text{Payoff of type } (v_i, c) = v_i - cw - m.}$

Key Simplifying Steps

✓ Class independent quantities, given cost c:

- $\mathbb{E}$ [waiting time in spot market].
- $\mathbb{E}$ [payment in spot market].

✓ Threshold strategy for joining the spot market.

- $\overline{c}_i \triangleq$ spot market cutoff for class i.
- Join if cost $c < \overline{c}_i$.

Key Simplifying Steps

✓ Class independent quantities, given cost c:

- $\mathbb{E}$ [waiting time in spot market].
- $\mathbb{E}$ [payment in spot market].

✓ Threshold strategy for joining the spot market.

- $\overline{c}_i \triangleq$ spot market cutoff for class i.
- Join if cost $c < \overline{c}_i$.

➥ $\mathbb{E}$ [waiting time] and $\mathbb{E}$ [payment] are functions of:

- Cost c, cutoffs $(\overline{c}_1, \overline{c}_2)$.

Spot Market: Waiting Time and Payments

- Given cutoffs $\overline{\mathbf{c}} \triangleq (\overline{c}_1, \overline{c}_2)$:
  - $w(c; \overline{\mathbf{c}}) \triangleq \mathbb{E}$ [waiting time in spot market if cost is c].
  - $m(c; \overline{\mathbf{c}}) \triangleq \mathbb{E}$ [payment in spot market for if cost is c].
  - Defined for any $\overline{\mathbf{c}}$ for which the queue is stable.

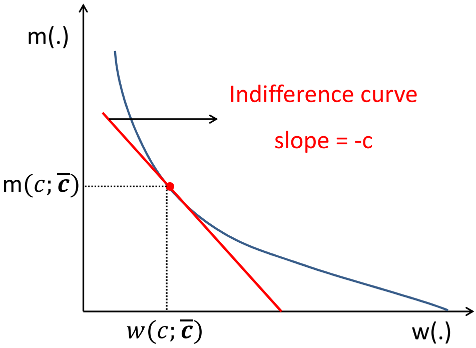
Spot Market: Waiting Time and Payments

- Given cutoffs $\overline{\mathbf{c}} \triangleq (\overline{c}_1, \overline{c}_2)$:
  - $w(c; \overline{\mathbf{c}}) \triangleq \mathbb{E}$ [waiting time in spot market if cost is c].
  - $m(c; \overline{\mathbf{c}}) \triangleq \mathbb{E}$ [payment in spot market for if cost is c].
  - Defined for any $\overline{\mathbf{c}}$ for which the queue is stable.
- Truth-telling as equilibrium:
  - $v_i - cw(c; \overline{\mathbf{c}}) - m(c; \overline{\mathbf{c}}) \geq v_i - cw(\widehat{c}; \overline{\mathbf{c}}) - m(\widehat{c}; \overline{\mathbf{c}}) \quad \forall c, \widehat{c}.$

Spot Market: Waiting Time and Payments

- Given cutoffs $\overline{\mathbf{c}} \triangleq (\overline{c}_1, \overline{c}_2)$:
  - $w(c; \overline{\mathbf{c}}) \triangleq \mathbb{E}\,[\text{waiting time in spot market if cost is } c]$.
  - $m(c; \overline{\mathbf{c}}) \triangleq \mathbb{E}\,[\text{payment in spot market for if cost is } c]$.
  - Defined for any $\overline{\mathbf{c}}$ for which the queue is stable.
- Truth-telling as equilibrium:
  - $A(c) \triangleq (w(c; \overline{\mathbf{c}}), m(c; \overline{\mathbf{c}}))$.
  - $A(\widehat{c})$ must lie above or on the $\underbrace{\text{line through } A(c) \text{ with slope } -c}_{\text{indifference curve for type } (v_i,\, c)}$.

Truth-telling as Equilibrium



- Tangent at point $(w(c;\bar{\mathbf{c}}), m(c;\bar{\mathbf{c}}))$ has slope $-c$.
- $m(\cdot)$ is a unique convex function of $w(\cdot)$.

Spot Market: Waiting Time and Payments

- $w(c; \overline{\mathbf{c}})$ is decreasing in c, increasing in $\overline{\mathbf{c}}$.
- $m(c; \overline{\mathbf{c}}) = \int_0^c w(t; \overline{\mathbf{c}})dt - cw(c; \overline{\mathbf{c}})$.
  - One the lines of Myerson [1981].
  - $m(c; \overline{\mathbf{c}})$ is increasing in c.

Spot Market: Waiting Time and Payments

- $w(c; \overline{\mathbf{c}})$ is decreasing in c, increasing in $\overline{\mathbf{c}}$.
- $m(c; \overline{\mathbf{c}}) = \int_0^c w(t; \overline{\mathbf{c}})dt - cw(c; \overline{\mathbf{c}})$.
  - One the lines of Myerson [1981].
  - $m(c; \overline{\mathbf{c}})$ is increasing in c.
- Same expected payment in any pricing mechanism!
- Spot market is fully characterized by $w(c; \overline{\mathbf{c}})$.

Spot Market: Waiting Time and Payments

- $w(c; \overline{\mathbf{c}})$ is decreasing in c, increasing in $\overline{\mathbf{c}}$.
- $m(c; \overline{\mathbf{c}}) = \int_0^c w(t; \overline{\mathbf{c}})dt - cw(c; \overline{\mathbf{c}})$.
  - One the lines of Myerson [1981].
  - $m(c; \overline{\mathbf{c}})$ is increasing in c.
- Same expected payment in any pricing mechanism!
- Spot market is fully characterized by $w(c; \overline{\mathbf{c}})$.

Find cutoffs $\overline{\mathbf{c}}$ that constitute an equilibrium?

Results: Equilibrium

- Equilibrium with cutoffs $(\overline{c}_1(p), \overline{c}_2(p))$ if each job type (v_i, c):
  - Reports its type truthfully.
  - Joins spot market if $c < \overline{c}_i(p)$.
  - Joins PAYG if $\overline{c}_i(p) \leq c \leq \mu v_i - p$.
  - Gets $\mathbb{E}\left[\text{payoff}\right] \geq 0$ from using the service.

Results: Equilibrium

- Equilibrium with cutoffs $(\overline{c}_1(p), \overline{c}_2(p))$ if each job type (v_i, c):
  - Reports its type truthfully.
  - Joins spot market if $c < \overline{c}_i(p)$.
  - Joins PAYG if $\overline{c}_i(p) \leq c \leq \mu v_i - p$.
  - Gets $\mathbb{E}\,[\text{payoff}] \geq 0$ from using the service.

Theorem

- For any PAYG price, equilibrium cutoffs uniquely exist.
- Equilibrium cutoffs:
  - Depend on PAYG price p, values (v_1, v_2).
  - Coupled with each other.

Results: Revenue Consequences

- Revenue per unit time:
  - From PAYG = $\frac{p}{\mu}\left(\sum_{i=1,2}\lambda_i\left[F_i(\mu v_i - p) - F_i(\bar{c}_i(p))\right]^+\right)$.
  - From spot market = $\sum_{i=1,2}\lambda_i\int_0^{\bar{c}_i(p)} m(t;\bar{\mathbf{c}}(p))f_i(t)dt$.

Results: Revenue Consequences

- Revenue per unit time:
  - From PAYG = $\frac{p}{\mu}\left(\sum_{i=1,2}\lambda_i\left[F_i(\mu v_i - p) - F_i(\overline{c}_i(p))\right]^+\right)$.
  - From spot market = $\sum_{i=1,2}\lambda_i\int_0^{\overline{c}_i(p)} m(t;\overline{\mathbf{c}}(p))f_i(t)dt$.

Theorem

If price maximizing the total revenue is not too large, then optimal revenue from PAYG in isolation > optimal revenue from PAYG + spot market.

Results: Revenue Consequences

- Revenue per unit time:
  - From PAYG = $\frac{p}{\mu}\left(\sum_{i=1,2} \lambda_i \left[F_i(\mu v_i - p) - F_i(\overline{c}_i(p))\right]^+ \right)$.
  - From spot market = $\sum_{i=1,2} \lambda_i \int_0^{\overline{c}_i(p)} m(t; \overline{\mathbf{c}}(p)) f_i(t) dt$.

Theorem

If price maximizing the total revenue is not too large, then optimal revenue from PAYG in isolation > optimal revenue from PAYG + spot market.

Conjecture

PAYG in isolation always revenue dominates PAYG + spot market.

Simulations: Setup

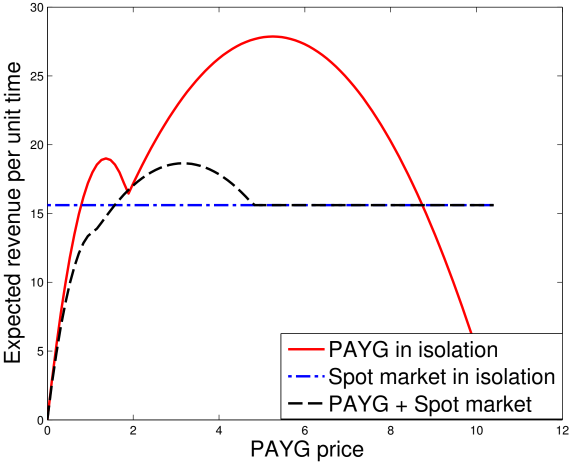
- Spot market:
  - k parallel $M/M/1$ queues.
  - An arriving job is uniformly sent to one of the k queues.
  - First price auction, high bid priority.
- Randomly generated parameters (v_i's, λ_i's, k); service rate $\mu = 1$.
- $F_i(c)$ is uniformly distributed in $[0, \mu v_i]$.

Simulations: Results

- In each random configurations, PAYG in isolation dominates PAYG + spot market.

Simulations: Results

- In each random configurations, PAYG in isolation dominates PAYG + spot market.



Conclusions and Future Work

- Analysis of fixed and market based pricing.
    - Uses a very general model.
- Evidence suggesting that fixed price is better.

Conclusions and Future Work

- Analysis of fixed and market based pricing.
  - Uses a very general model.
- Evidence suggesting that fixed price is better.
- Next steps:
  - Effect of finite capacity?
  - Effect of competition among providers?

Conclusions and Future Work

- Analysis of fixed and market based pricing.
  - Uses a very general model.
- Evidence suggesting that fixed price is better.
- Next steps:
  - Effect of finite capacity?
  - Effect of competition among providers?

Thank you. Questions?