A Theory of Cloud Bandwidth Pricing for Video-on-Demand Providers

Di Niu, Chen Feng, Baochun Li

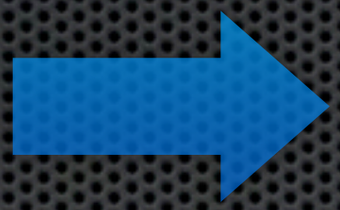
University of Toronto

A VoD Company Cares about

- Bandwidth guarantee
  - to maintain streaming quality
- Huge Bandwidth cost
  - “youku” pays > $ 38.6M/year for bandwidth

Can VoD leverage Cloud to reduce bandwidth cost with bandwidth guarantees

Current Cloud Platforms



Netflix moved to Amazon Web Services in 2010

- Problem: No bandwidth guarantee!

Extending Cloud-Tenant Interface

Bandwidth reservation for egress traffic of a VM

C. Guo *et al.*

SecondNet: a Data Center Network Virtualization Architecture with Bandwidth Guarantees

ACM **CoNEXT '10**

H. Ballani, P. Costa, T. Karagiannis, and A. Rowstron,

Towards Predictable Datacenter Networks

ACM **SIGCOMM '11**

How to Price Bandwidth Guarantees?

- VoD Companies want
  - bandwidth guarantee
  - reduced cost
- Cloud Providers want
  - economic incentives (profit)

How about Pay-as-You-Go?

- Charge based on usage
  - *number of bytes transferred in the past hour*
- Not suitable for pricing bandwidth guarantees

Example: Bursty vs. Constant

Our Contributions

- A New Model
  - to price bandwidth guarantees
- Theoretical Analysis
  - to set pricing policy properly and fairly
- Practical Performance
  - simulations based on workload traces

Assumptions

- Each tenant i:
  - random demand D_i
  - expectation $\mu_i = \mathrm{E}[D_i]$
  - standard deviation $\sigma_i^2 = \mathrm{var}[D_i]$
- Demand covariance matrix $\Sigma = [\sigma_{ij}]$

Probabilistic Bandwidth Guarantee

- With random demand D_i, tenant i wants to reserve bandwidth R_i such that

$$Pr(D_i > R_i) < \epsilon \longrightarrow \text{Small constant}$$

- If D_i is Gaussian, then

$$R_i = \underbrace{\mu_i}_{\text{Expected Demand}} + \underbrace{\theta(\epsilon)\sigma_i}_{\text{Risk Margin}}$$

A constant

Pricing Individual Reservation

- Assume:
  reserving 1 unit of bandwidth costs $1
- Tenant i should pay

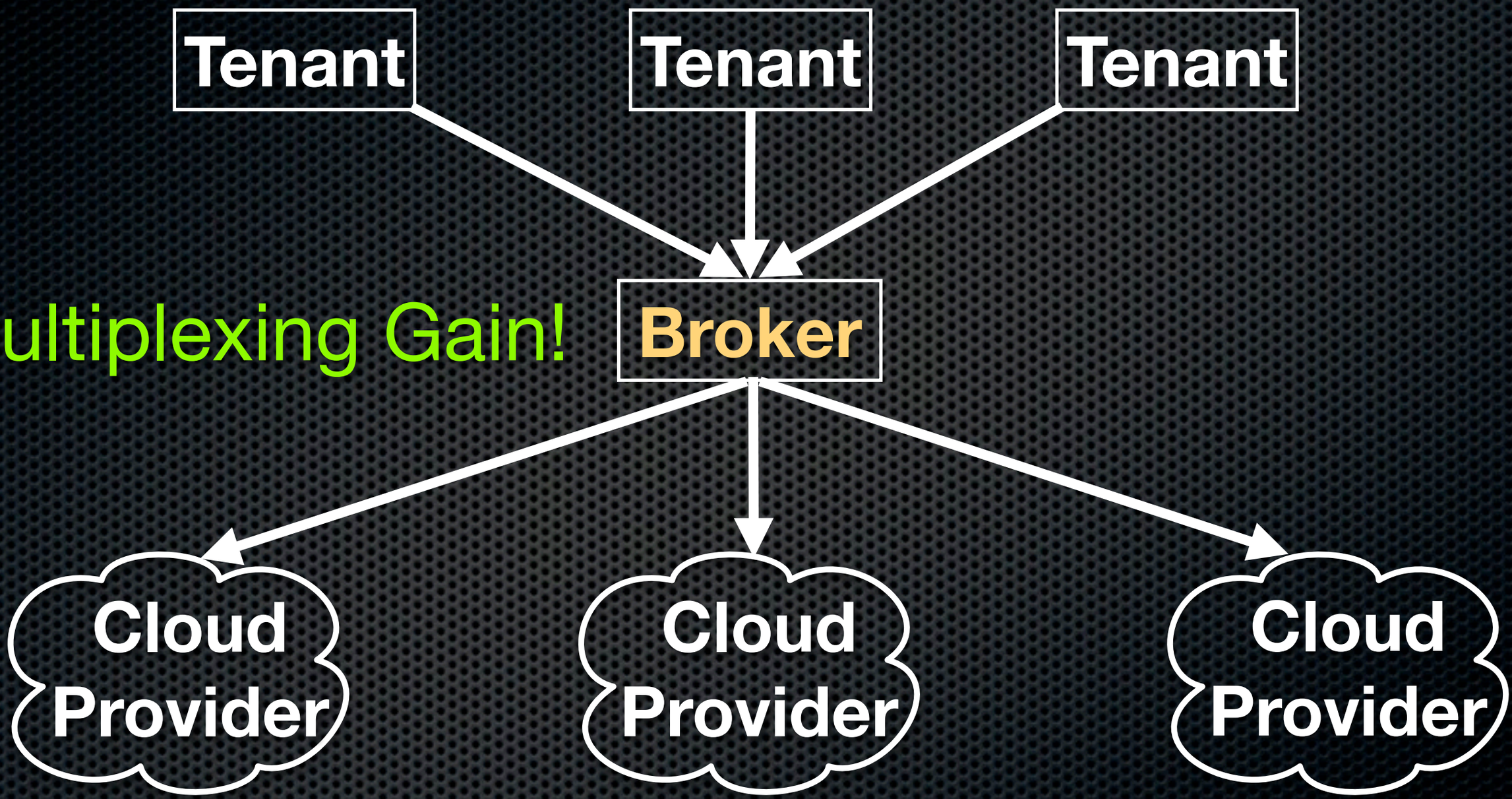
$$P_i = R_i = \mu_i + \theta(\epsilon)\sigma_i$$

Mean Demand

Burstiness

- Individual reservation is costly
- Most of the time, underutilization

An Economically Viable Broker



Reservation = $\sum \mu_i$ + Shared Risk Margin

Shared Risk Margin < $\sum$ Individual Risk Margin

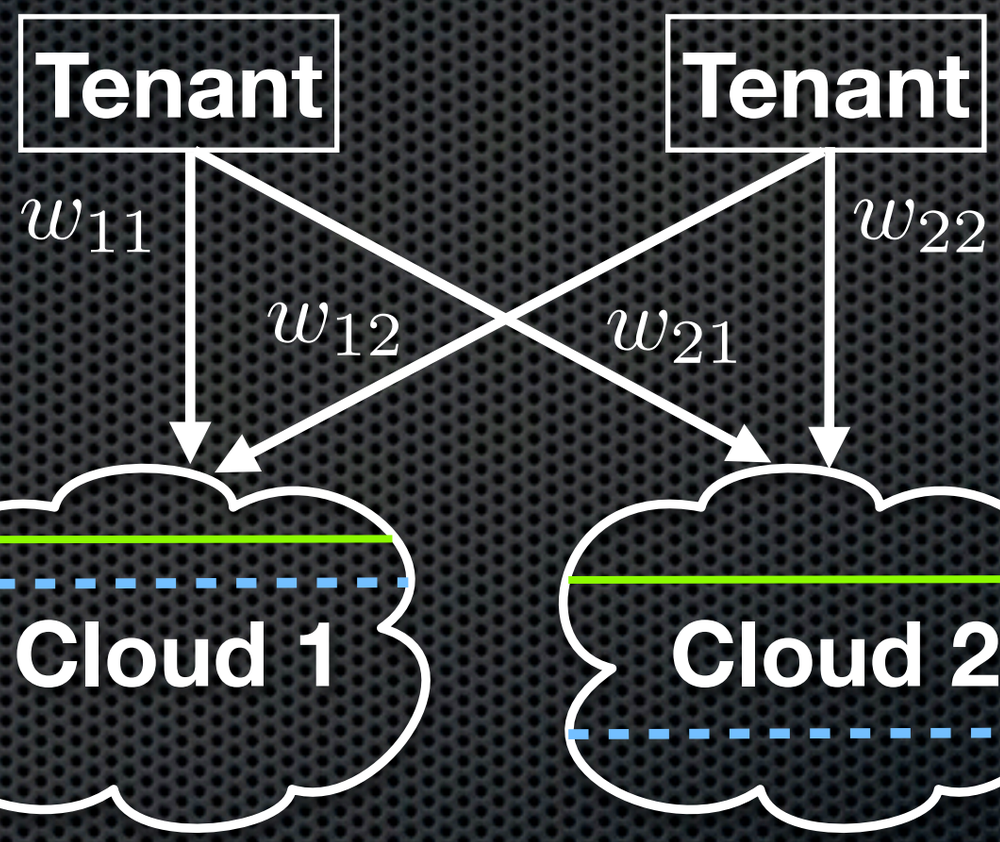
The Service Provided by the Broker

~~Individual Reservation R_i~~

~~Probabilistic Guarantee $\Pr(D_i > R_i) < \epsilon$~~

Random Demand D_1 D_2

Guaranteed Portion $w_i = \sum_s w_{si}$



Request Routing

Reservation A_1

A_2 Reservation

Load $L_1 = \sum_i w_{1i} D_i$

Load $L_2 = \sum_i w_{2i} D_i$

New Probabilistic Guarantee

$$\Pr(L_s > A_s) < \epsilon, \text{ where } L_s = \sum_i w_{si} D_i$$

Definition: *Pricing Policy*

- Guaranteed portion $w_i = \sum_s w_{si}$
- Pricing strategy $P_i(w_i, \mu_i, \sigma_i, \ldots)$
  - $P_i(w_i)$ *increasing, concave,* $P_i = 0$ if $w_i = 0$
- Pricing policy $\{P_i(\cdot)\}$
- Example: individual reservation

$$P_i = R_i = \mu_i + \theta(\epsilon)\sigma_i \longrightarrow P_i(w_i) = \big(\mu_i + \theta(\epsilon)\sigma_i\big)w_i$$

Main Result 1

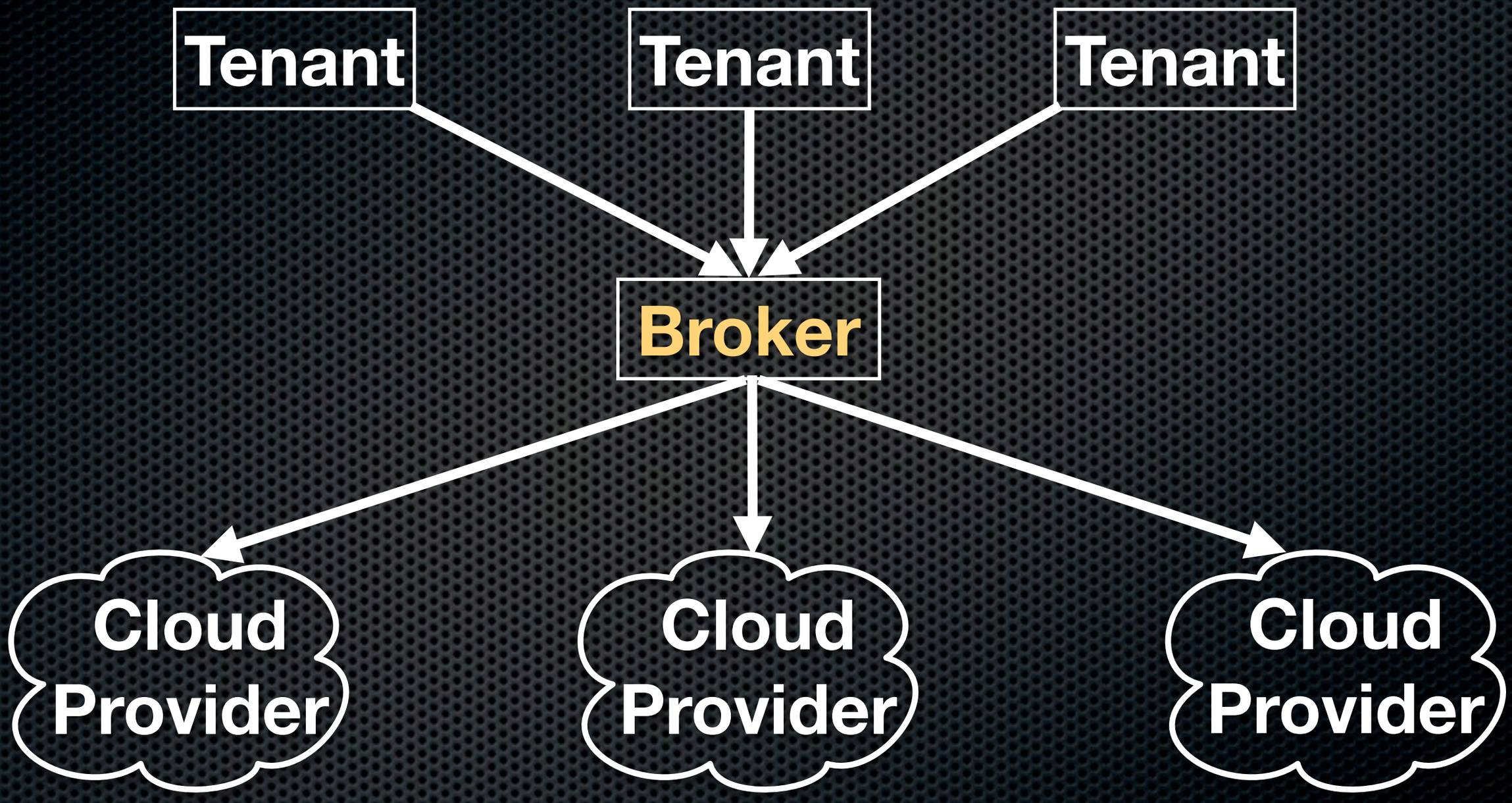
Pricing Region

Objectives:
Set prices such that
tenants enjoy discount
broker has incentive
all demands guaranteed

Good Pricing Region

- If and Only if $P_i(w_i) \leq \left(\mu_i + \theta(\epsilon)\sigma_i\right) w_i$
  - Tenant i enjoys some *price discount*
- If and Only if $P_i'(1) \geq \mu_i + \theta(\epsilon)\sigma_i\rho_{iM}$ → Correlation to the market
  - A *profit-driven* broker is willing to *guarantee* all the tenant demand ($w_i = 1, \forall i$)
- If both conditions are met:
  - Achieve Discount + Incentive + Guarantee!
  - $\mu_i + \theta(\epsilon)\sigma_i\rho_{iM} \leq P_i(1) \leq \mu_i + \theta(\epsilon)\sigma_i$

$$\mu_i + \theta(\epsilon)\sigma_i\rho_{iM} \leq P_i(1) \leq \mu_i + \theta(\epsilon)\sigma_i$$



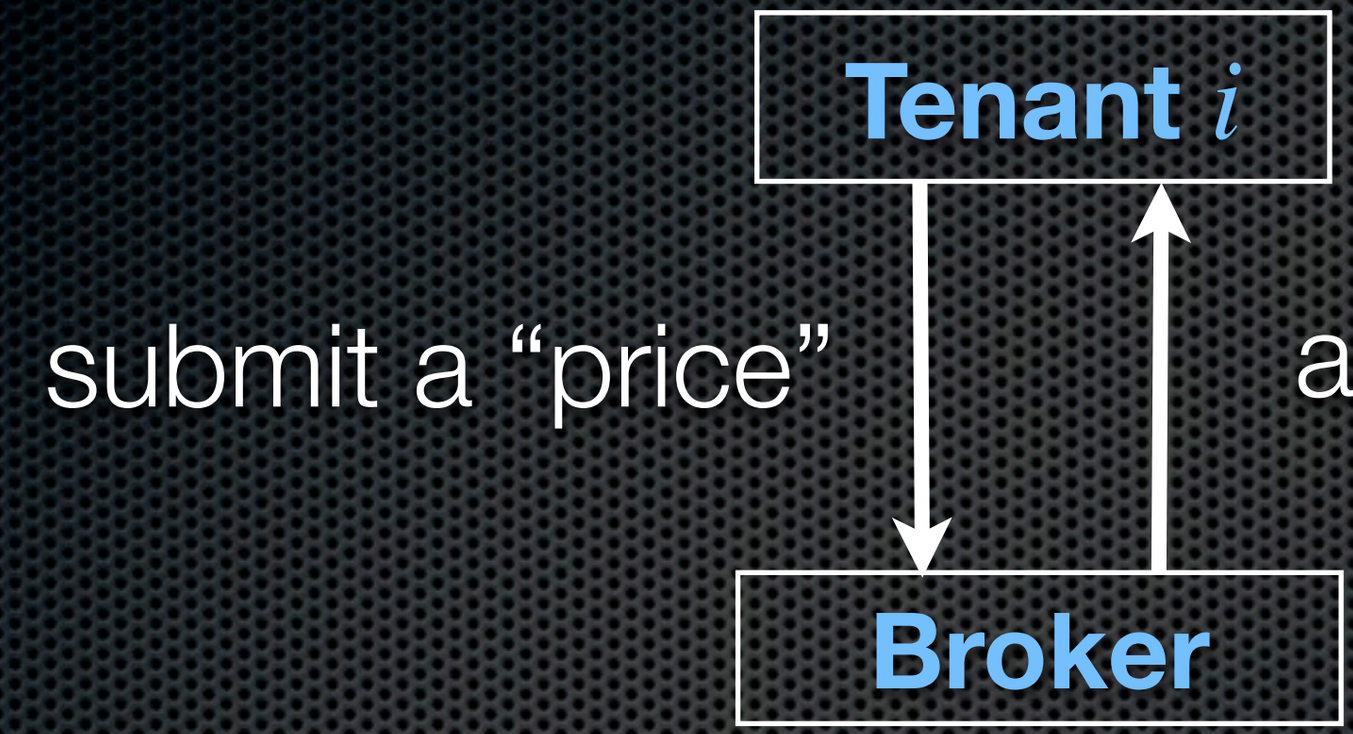
With a good pricing policy,
multiplexing gain is shared fairly.

How to enforce good pricing?

Enforcement by an Authority √

What about a Free Market?

Free Market: Allow Bargains



- w_i - guaranteed portion in $[0,1]$
- Submitted price is a function: $P_i(w_i)$
- If $P_i(\cdot)$ is too low, the broker will set $w_i < 1$.

A *Game* among all Tenants

$$max \text{ Utility} = \begin{cases} -P_i(1), & \text{if } w_i = 1 \\ -\infty, & \text{if } w_i < 1 \end{cases}$$

Wants to reduce $P_i(\cdot)$, but must have $w_i = 1$

Can prisoner's dilemma happen?

Players

Pricing function $P_i(\cdot)$

Guarantee $w_i D_i$

Strategy

Broker

Choose w_i to $\max_{w_1,\ldots,w_N} R = \sum_{i=1}^{N} P_i(w_i) - \sum_{s=1}^{S} A_s(w_1,\ldots,w_N)$

Profit

Total Revenue

Sum of Costs

Main Result 2

Equilibrium Pricing

Equilibrium of the Free Market

With multiplexing, the market has a *unique* Nash equilibrium:

$$P_i^*(1) = \mu_i + \theta(\epsilon)\sigma_i\rho_{iM}$$

- μ_i: Expected Demand
- $\theta(\epsilon)\sigma_i$: Demand Standard Deviation
- ρ_{iM}: Correlation to the market, in [-1, 1]

Recall the good pricing region

$$\underline{\mu_i + \theta(\epsilon)\sigma_i\rho_{iM}} \leq P_i(1) \leq \underline{\mu_i + \theta(\epsilon)\sigma_i}$$

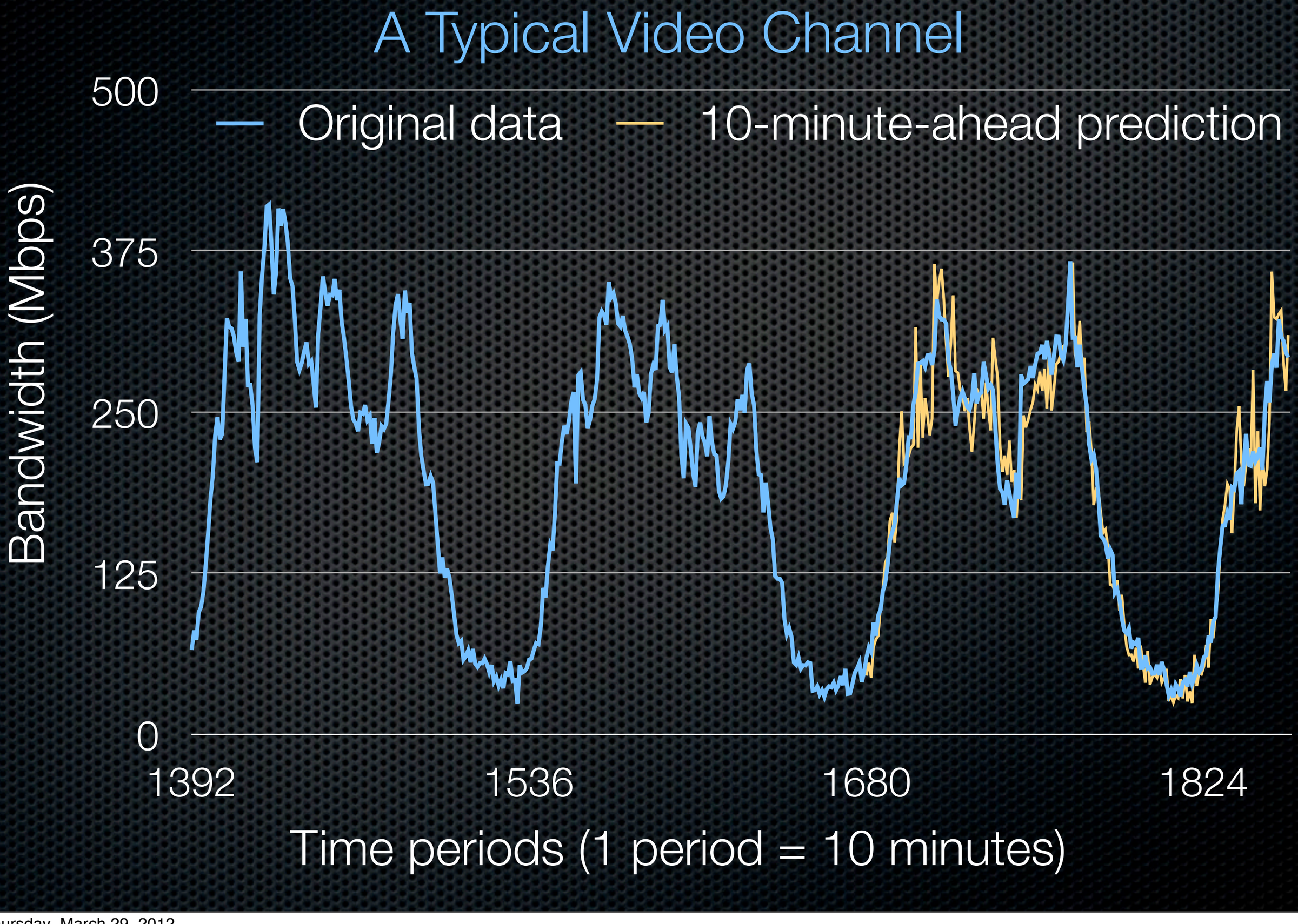
- $\mu_i + \theta(\epsilon)\sigma_i\rho_{iM}$: Nash Equilibrium
- $\mu_i + \theta(\epsilon)\sigma_i$: Individual Reservation

Trace-Driven Simulations

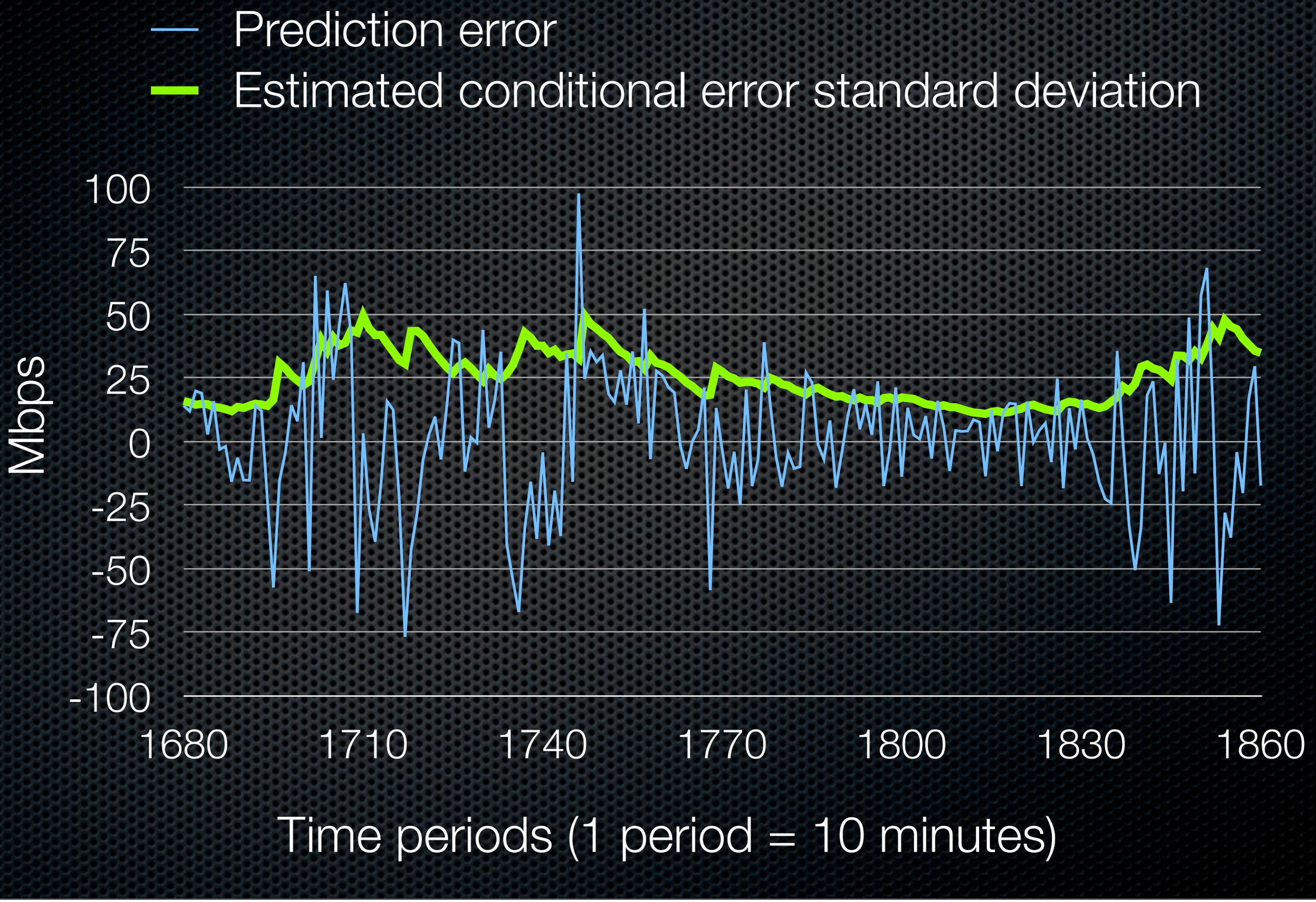
Data Mining

- UUSee: a VoD provider in China
- Aggregate BW demand in video channels

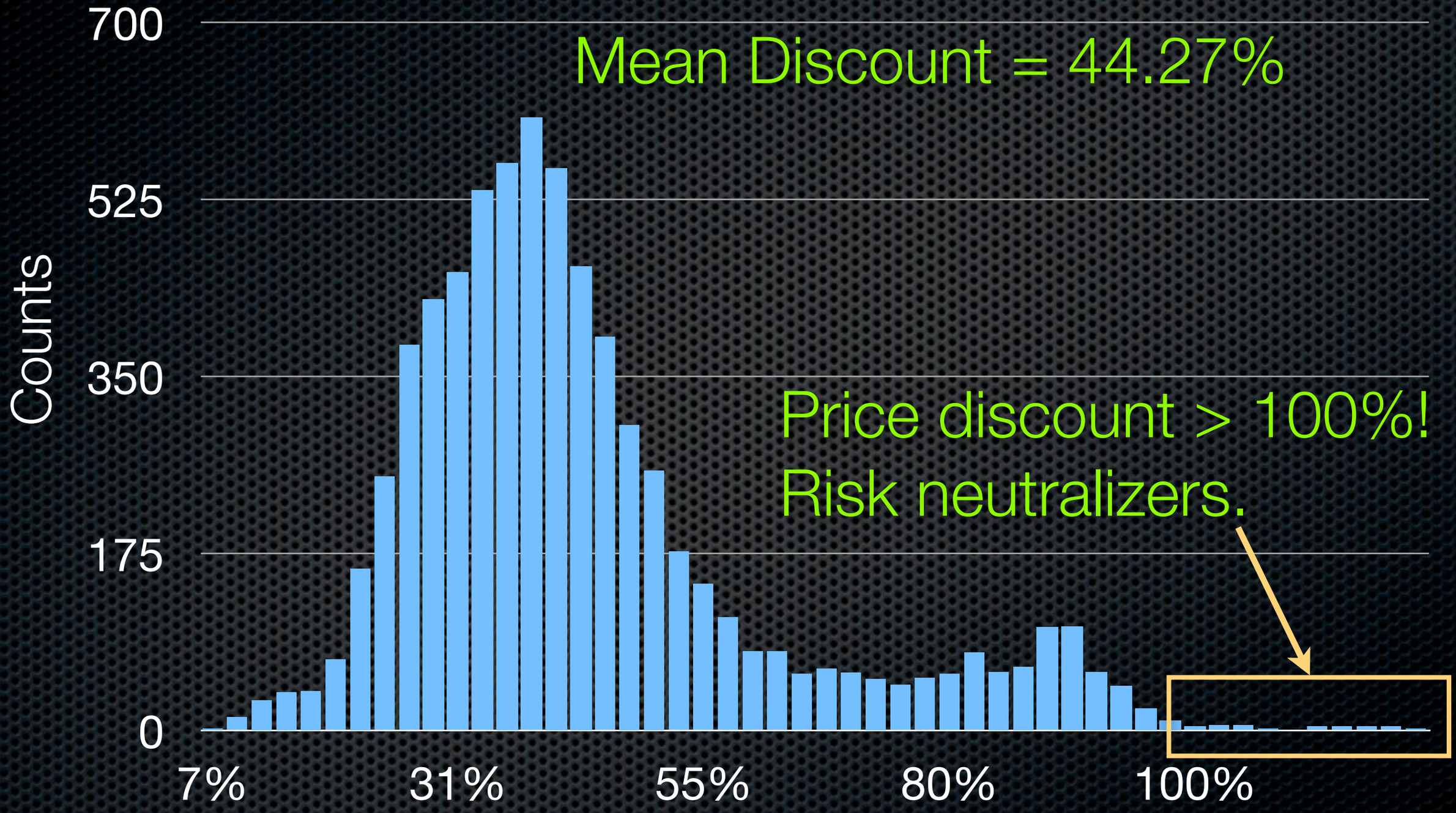
How to estimate expected demand?



How to estimate variance?



Histogram of Price Discounts



Discounts of All Tenants in All Test Periods

Conclusions

- Broker Service:
  - Bandwidth reservation via multiplexing
  - A theory to price bandwidth guarantees fairly and properly
  - Pricing based on burstiness and correlation

Thank you!

Google “Di Niu”

http://iqua.ece.toronto.edu/~dniu/