Neutrality in Future Public Clouds: Implications and Challenges

George Kesidis

School of EECS
The Pennsylvania State University
kesidis@gmail.com

in collaboration with
Bhuvan Urgaonkar, Neda Nasiriani and Cheng Wang

USENIX HotCloud, Denver, Tues. June 21, 2016

This research supported in part by NSF CNS grants 1228717 and 1526133
and by NSF CAREER grant 0953541.

Introduction

- We analogize the evolution the public-cloud marketplace with that of the Internet (post web):
  - Initially, cloud over-engineered while experiencing dramatic growth
  - As the marketplace matured

Introduction

- We analogize the evolution the public-cloud marketplace with that of the Internet (post web):
  - Initially, cloud over-engineered while experiencing dramatic growth
  - As the marketplace matured
    - it has become critical infrastructure for many price-sensitive customers/tenants, and
    - subsidiaries/affiliates of the few public cloud providers compete with their own tenants, *e.g.*, both Amazon Prime and Netflix use Amazon EC2.

Introduction

- We analogize the evolution the public-cloud marketplace with that of the Internet (post web):
  - Initially, cloud over-engineered while experiencing dramatic growth
  - As the marketplace matured
    - it has become critical infrastructure for many price-sensitive customers/tenants, and
    - subsidiaries/affiliates of the few public cloud providers compete with their own tenants, *e.g.*, both Amazon Prime and Netflix use Amazon EC2.
- We expect antitrust concerns to lead to neutrality regulations in the public-cloud marketplace as they did in the commercialized, public Internet marketplace.
  - The FCC deemed the Internet a utility in 2015 in a move to shore-up neutrality rules.

Introduction

- We analogize the evolution the public-cloud marketplace with that of the Internet (post web):
  - Initially, cloud over-engineered while experiencing dramatic growth
  - As the marketplace matured
    - it has become critical infrastructure for many price-sensitive customers/tenants, and
    - subsidiaries/affiliates of the few public cloud providers compete with their own tenants, *e.g.*, both Amazon Prime and Netflix use Amazon EC2.
- We expect antitrust concerns to lead to neutrality regulations in the public-cloud marketplace as they did in the commercialized, public Internet marketplace.
  - The FCC deemed the Internet a utility in 2015 in a move to shore-up neutrality rules.
- There is relatively scant existing consideration of neutrality issues re. virtualized services in the public cloud, beyond how *network* neutrality relates to the cloud.

Neutrality of the Public Cloud

- "A truly neutral data centre provider is one that is independent of the companies colocating in the data centre, does not compete with them in any way, and offers no packaged services as part of colocation. Customers are free to contract directly with the providers of their choice" [Interxion, a colo]

Neutrality of the Public Cloud

- "A truly neutral data centre provider is one that is independent of the companies colocating in the data centre, does not compete with them in any way, and offers no packaged services as part of colocation. Customers are free to contract directly with the providers of their choice" [Interxion, a colo]
- Public cloud-computing services range from colo facilities & bare-metal IaaS to much more popular and cost-effective virtualized services (including PaaS,SaaS).

Neutrality of the Public Cloud

- "A truly neutral data centre provider is one that is independent of the companies colocating in the data centre, does not compete with them in any way, and offers no packaged services as part of colocation. Customers are free to contract directly with the providers of their choice" [Interxion, a colo]
- Public cloud-computing services range from colo facilities & bare-metal IaaS to much more popular and cost-effective virtualized services (including PaaS,SaaS).
- Virtualized services have diverse service-level agreements (SLAs) and associated price-performance trade-offs, spanning more than just network IO.

Neutrality of the Public Cloud

- “A truly neutral data centre provider is one that is independent of the companies colocating in the data centre, does not compete with them in any way, and offers no packaged services as part of colocation. Customers are free to contract directly with the providers of their choice” [Interxion, a colo]
- Public cloud-computing services range from colo facilities & bare-metal IaaS to much more popular and cost-effective virtualized services (including PaaS,SaaS).
- Virtualized services have diverse service-level agreements (SLAs) and associated price-performance trade-offs, spanning more than just network IO.
- So a more challenging auditing/enforcement setting.

Auditing and verification of neutral treatment

- Whereas network bandwidth allocated by an ISP can be effectively monitored and audited, doing so for virtualized IT resources offered by a public cloud is inherently difficult.

Auditing and verification of neutral treatment

- Whereas network bandwidth allocated by an ISP can be effectively monitored and audited, doing so for virtualized IT resources offered by a public cloud is inherently difficult.
- Existing approaches for measuring/auditing virtualized IT resource usage implicitly assume trusted resource managers.

Auditing and verification of neutral treatment

- Whereas network bandwidth allocated by an ISP can be effectively monitored and audited, doing so for virtualized IT resources offered by a public cloud is inherently difficult.
- Existing approaches for measuring/auditing virtualized IT resource usage implicitly assume trusted resource managers.
- A hypervisor may over-commit CPUs to multiple VMs, which makes it difficult to detect/quantify each VM's share of CPU capacity (CPU cycles) from within the guest OS.

Auditing and verification of neutral treatment

- Whereas network bandwidth allocated by an ISP can be effectively monitored and audited, doing so for virtualized IT resources offered by a public cloud is inherently difficult.
- Existing approaches for measuring/auditing virtualized IT resource usage implicitly assume trusted resource managers.
- A hypervisor may over-commit CPUs to multiple VMs, which makes it difficult to detect/quantify each VM's share of CPU capacity (CPU cycles) from within the guest OS.
- Netflix apparently relies upon "stolen time" (a measure of competition for CPU) to detect when its procured VM instances are not getting adequate CPU capacity.

Auditing and verification of neutral treatment

- Whereas network bandwidth allocated by an ISP can be effectively monitored and audited, doing so for virtualized IT resources offered by a public cloud is inherently difficult.
- Existing approaches for measuring/auditing virtualized IT resource usage implicitly assume trusted resource managers.
- A hypervisor may over-commit CPUs to multiple VMs, which makes it difficult to detect/quantify each VM's share of CPU capacity (CPU cycles) from within the guest OS.
- Netflix apparently relies upon "stolen time" (a measure of competition for CPU) to detect when its procured VM instances are not getting adequate CPU capacity.
- But what if the hypervisor misreported this metric? A hypervisor may deliberately "dilate" a VM's notion of time by slowing its delivery of virtualized timer interrupts [Gupta et al.'05].

Auditing and verification of neutral treatment

- Whereas network bandwidth allocated by an ISP can be effectively monitored and audited, doing so for virtualized IT resources offered by a public cloud is inherently difficult.
- Existing approaches for measuring/auditing virtualized IT resource usage implicitly assume trusted resource managers.
- A hypervisor may over-commit CPUs to multiple VMs, which makes it difficult to detect/quantify each VM's share of CPU capacity (CPU cycles) from within the guest OS.
- Netflix apparently relies upon "stolen time" (a measure of competition for CPU) to detect when its procured VM instances are not getting adequate CPU capacity.
- But what if the hypervisor misreported this metric? A hypervisor may deliberately "dilate" a VM's notion of time by slowing its delivery of virtualized timer interrupts [Gupta et al.'05].

Lessons from network neutrality reinterpreted

Equal treatment modulo SLAs:

- Neutrality does not mean that every tenant is treated equally.

Lessons from network neutrality reinterpreted

Equal treatment modulo SLAs:

- Neutrality does not mean that every tenant is treated equally.
- Fair/equal treatment across tenants (including the cloud's own affiliates) may only relate to those with similar SLAs in how the cloud:
  - allocates any discretionary resources,
  - effectively underbooks resources when exploiting statistical multiplexing among tenants with similar SLAs (see paper), or
  - deals with congestion.

Lessons from network neutrality reinterpreted

Equal treatment modulo SLAs:

- Neutrality does not mean that every tenant is treated equally.
- Fair/equal treatment across tenants (including the cloud's own affiliates) may only relate to those with similar SLAs in how the cloud:
  - allocates any discretionary resources,
  - effectively underbooks resources when exploiting statistical multiplexing among tenants with similar SLAs (see paper), or
  - deals with congestion.
- Without a single (or equivalent lumped) resource type, it may be challenging to identify groups of similar SLAs in order to define fair treatment.

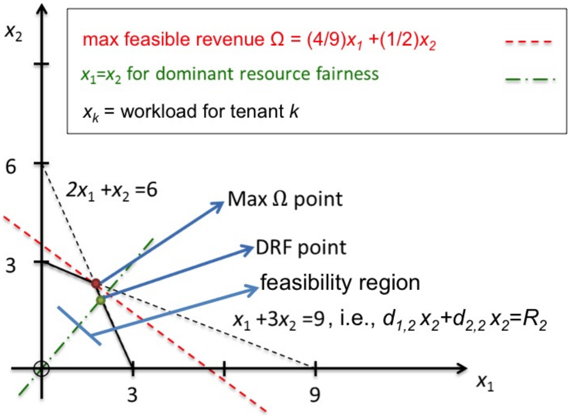
Lessons from network neutrality reinterpreted (cont)

- A neutral provider's behavior must not be based on "inside information" or preferences, *e.g.*,
  - [Mogul et al. '15] consider a network bandwidth allocation problem wherein a cloud provider reckons tenant sensitivity to network bandwidth underprovisioning and uses this to allocate bandwidth differentially to improve its profits.

Lessons from network neutrality reinterpreted (cont)

- A neutral provider's behavior must not be based on "inside information" or preferences, *e.g.*,
  - [Mogul et al. '15] consider a network bandwidth allocation problem wherein a cloud provider reckons tenant sensitivity to network bandwidth underprovisioning and uses this to allocate bandwidth differentially to improve its profits.
- Given resource-oriented SLAs (not performance oriented), if cloud reckons "sensitivity" based on:
  - precise knowledge of tenant's performance & requirements, then not neutral
  - measured resource usage, then may be neutral
  - Amazon "burstable" VM instances use token-bucket regulation to govern access to network IO and CPU, and so may be more aggressively consolidated without workload inferences (neutrally, see technical report)

Resource congestion, choice of boundary Nash equilibria



x = incident tenant workload, R=resource (CPU, memory) capacity, d = tenant resource demand per unit workload

Memcached case study

- We considered two Memcached tenants whose CPU and memory needs we manipulated to illustrate neutrality issues under different representative cloud resource management options.

Memcached case study

- We considered two Memcached tenants whose CPU and memory needs we manipulated to illustrate neutrality issues under different representative cloud resource management options.
- Each tenant workload was generated by YCSB in time-varying fashion according to i.i.d. Gaussian processes with identical mean demand but the variance of tenant 2 is double that of tenant 1.

Memcached case study

- We considered two Memcached tenants whose CPU and memory needs we manipulated to illustrate neutrality issues under different representative cloud resource management options.
- Each tenant workload was generated by YCSB in time-varying fashion according to i.i.d. Gaussian processes with identical mean demand but the variance of tenant 2 is double that of tenant 1.
- The tenants have the same SLAs corresponding to one guaranteed core, and the cloud has a discretionary third core that it can share between the two tenants.

Memcached case study

- We considered two Memcached tenants whose CPU and memory needs we manipulated to illustrate neutrality issues under different representative cloud resource management options.
- Each tenant workload was generated by YCSB in time-varying fashion according to i.i.d. Gaussian processes with identical mean demand but the variance of tenant 2 is double that of tenant 1.
- The tenants have the same SLAs corresponding to one guaranteed core, and the cloud has a discretionary third core that it can share between the two tenants.
- Empirically, we found that a single core can achieve throughput of 75k ops/s with satisfactory mean response times of 400μs.

Resv vs CFS based sharing of discretionary core - set-up

- The cloud may employ a weighted reservation (Resv) based approach or a work-conserving, proportional-share (CFS) scheduling for sharing the discretionary third core.

Resv vs CFS based sharing of discretionary core - set-up

- The cloud may employ a weighted reservation (Resv) based approach or a work-conserving, proportional-share (CFS) scheduling for sharing the discretionary third core.
- The Resv1:1 (equally weighted) and CFS1:1 schemes are arguably neutral as they are based only on parameters in the tenant SLAs, *i.e.*, the same mean demands in this example corresponding to a single core;

Resv vs CFS based sharing of discretionary core - set-up

- The cloud may employ a weighted reservation (Resv) based approach or a work-conserving, proportional-share (CFS) scheduling for sharing the discretionary third core.
- The Resv1:1 (equally weighted) and CFS1:1 schemes are arguably neutral as they are based only on parameters in the tenant SLAs, *i.e.*, the same mean demands in this example corresponding to a single core;
- while CFS3:7 is based on measured (conditional) demand-variation of the two tenants which is not part of their SLAs and so this scheme is arguably <u>not</u> neutral.

Resv vs CFS based sharing of discretionary core - set-up

- The cloud may employ a weighted reservation (Resv) based approach or a work-conserving, proportional-share (CFS) scheduling for sharing the discretionary third core.
- The Resv1:1 (equally weighted) and CFS1:1 schemes are arguably neutral as they are based only on parameters in the tenant SLAs, *i.e.*, the same mean demands in this example corresponding to a single core;
- while CFS3:7 is based on measured (conditional) demand-variation of the two tenants which is not part of their SLAs and so this scheme is arguably not neutral.
- If the intention is to prevent tenant 2 (the one with more demand variation) from defecting to other cloud-services providers, a neutral cloud may tend to adopt CFS1:1.

Resv vs CFS based sharing of discretionary core - set-up

- The cloud may employ a weighted reservation (Resv) based approach or a work-conserving, proportional-share (CFS) scheduling for sharing the discretionary third core.
- The Resv1:1 (equally weighted) and CFS1:1 schemes are arguably neutral as they are based only on parameters in the tenant SLAs, *i.e.*, the same mean demands in this example corresponding to a single core;
- while CFS3:7 is based on measured (conditional) demand-variation of the two tenants which is not part of their SLAs and so this scheme is arguably not neutral.
- If the intention is to prevent tenant 2 (the one with more demand variation) from defecting to other cloud-services providers, a neutral cloud may tend to adopt CFS1:1.
- On the other hand, a neutral cloud may want to entice tenant 2 to renegotiate a more costly SLA by adopting Resv1:1.

Resv vs CFS based sharing of discretionary core - set-up

- The cloud may employ a weighted reservation (Resv) based approach or a work-conserving, proportional-share (CFS) scheduling for sharing the discretionary third core.
- The Resv1:1 (equally weighted) and CFS1:1 schemes are arguably neutral as they are based only on parameters in the tenant SLAs, *i.e.*, the same mean demands in this example corresponding to a single core;
- while CFS3:7 is based on measured (conditional) demand-variation of the two tenants which is not part of their SLAs and so this scheme is arguably <u>not</u> neutral.
- If the intention is to prevent tenant 2 (the one with more demand variation) from defecting to other cloud-services providers, a neutral cloud may tend to adopt CFS1:1.
- On the other hand, a neutral cloud may want to entice tenant 2 to renegotiate a more costly SLA by adopting Resv1:1.
- The premise here is that with greater demand variability, tenant 2 will out-compete tenant 1 for the discretionary CPU core under CFS;

Resv vs CFS based sharing of discretionary core - set-up

- The cloud may employ a weighted reservation (Resv) based approach or a work-conserving, proportional-share (CFS) scheduling for sharing the discretionary third core.
- The Resv1:1 (equally weighted) and CFS1:1 schemes are arguably neutral as they are based only on parameters in the tenant SLAs, *i.e.*, the same mean demands in this example corresponding to a single core;
- while CFS3:7 is based on measured (conditional) demand-variation of the two tenants which is not part of their SLAs and so this scheme is arguably not neutral.
- If the intention is to prevent tenant 2 (the one with more demand variation) from defecting to other cloud-services providers, a neutral cloud may tend to adopt CFS1:1.
- On the other hand, a neutral cloud may want to entice tenant 2 to renegotiate a more costly SLA by adopting Resv1:1.
- The premise here is that with greater demand variability, tenant 2 will out-compete tenant 1 for the discretionary CPU core under CFS;
- but this may not be entirely the case when the tenant demands are positively correlated, *i.e.*, how workload is consolidated by the cloud will impact such "iso-neutral" decision-making.

Resv vs CFS based sharing of discretionary core - results

Table: Latency of tenants, uncorrelated demand scenario:

	CFS3:7		CFS1:1		Resv1:1	
	95-th	avg	95-th	avg	95-th	avg
Tenant 1	484	318	449	382	451	365
Tenant 2	418	254	435	284	444	299

- The more demand-variable tenant 2 has best latency performance (both in mean and 95 percentile) under non-neutral CFS3:7, irrespective of tenant demand correlation.

Resv vs CFS based sharing of discretionary core - results

Table: Latency of tenants, uncorrelated demand scenario:

	CFS3:7		CFS1:1		Resv1:1	
	95-th	avg	95-th	avg	95-th	avg
Tenant 1	484	318	449	382	451	365
Tenant 2	418	254	435	284	444	299

- The more demand-variable tenant 2 has best latency performance (both in mean and 95 percentile) under non-neutral CFS3:7, irrespective of tenant demand correlation.
- Also for simulated correlated tenant demand, improved 95-percentile latency performance for tenant 2 under CFS1:1 over Resv1:1; but not for mean latency (unlike uncorrelated scenario above).

Resv vs CFS based sharing of discretionary core - results

Table: Latency of tenants, uncorrelated demand scenario:

	CFS3:7		CFS1:1		Resv1:1	
	95-th	avg	95-th	avg	95-th	avg
Tenant 1	484	318	449	382	451	365
Tenant 2	418	254	435	284	444	299

- The more demand-variable tenant 2 has best latency performance (both in mean and 95 percentile) under non-neutral CFS3:7, irrespective of tenant demand correlation.
- Also for simulated correlated tenant demand, improved 95-percentile latency performance for tenant 2 under CFS1:1 over Resv1:1; but not for mean latency (unlike uncorrelated scenario above).
- See paper for an example involving memory allocation.

Summary challenges regarding cloud neutrality

- Tenant-side auditing/verification
  - select resource granularities
  - simplify by use of a lumped resource, *e.g.*, energy
  - use of hardware verification technology
  - role of third-party verifiers
  - account for "natural" performance impact of interference (through unvirtualized resources) and hardware heterogeneity, *i.e.*, effective capacity variation, that do not violate SLAs

Summary challenges regarding cloud neutrality

- Tenant-side auditing/verification
  - select resource granularities
  - simplify by use of a lumped resource, *e.g.*, energy
  - use of hardware verification technology
  - role of third-party verifiers
  - account for "natural" performance impact of interference (through unvirtualized resources) and hardware heterogeneity, *i.e.*, effective capacity variation, that do not violate SLAs
- Cloud-side regulation re. information limits, inferences it can make, and control actions it can take (*e.g.*, neutral scheduling)

Summary challenges regarding cloud neutrality

- Tenant-side auditing/verification
  - select resource granularities
  - simplify by use of a lumped resource, *e.g.*, energy
  - use of hardware verification technology
  - role of third-party verifiers
  - account for "natural" performance impact of interference (through unvirtualized resources) and hardware heterogeneity, *i.e.*, effective capacity variation, that do not violate SLAs
- Cloud-side regulation re. information limits, inferences it can make, and control actions it can take (*e.g.*, neutral scheduling)
- Relate to neutrality regulations in the Internet

Summary challenges regarding cloud neutrality

- Tenant-side auditing/verification
  - select resource granularities
  - simplify by use of a lumped resource, *e.g.*, energy
  - use of hardware verification technology
  - role of third-party verifiers
  - account for "natural" performance impact of interference (through unvirtualized resources) and hardware heterogeneity, *i.e.*, effective capacity variation, that do not violate SLAs
- Cloud-side regulation re. information limits, inferences it can make, and control actions it can take (*e.g.*, neutral scheduling)
- Relate to neutrality regulations in the Internet
- General scalability and accuracy issues