



OPEN DATA CENTER ALLIANCESM USAGE: INPUT/OUTPUT (IO) CONTROLS

LEGAL NOTICE

© 2011 Open Data Center Alliance, Inc. ALL RIGHTS RESERVED.

This “Open Data Center AllianceSM Usage: Input/Output (IO) Cotrols” is proprietary to the Open Data Center Alliance, Inc.

NOTICE TO USERS WHO ARE NOT OPEN DATA CENTER ALLIANCE PARTICIPANTS: Non-Open Data Center Alliance Participants only have the right to review, and make reference or cite, this document. Any such references or citations to this document must give the Open Data Center Alliance, Inc. full attribution and must acknowledge the Open Data Center Alliance, Inc.’s copyright in this document. Such users are not permitted to revise, alter, modify, make any derivatives of, or otherwise amend this document in any way.

NOTICE TO USERS WHO ARE OPEN DATA CENTER ALLIANCE PARTICIPANTS: Use of this document by Open Data Center Alliance Participants is subject to the Open Data Center Alliance’s bylaws and its other policies and procedures.

OPEN CENTER DATA ALLIANCESM, ODCASM, and the OPEN DATA CENTER ALLIANCE logoSM are service marks owned by Open Data Center Alliance, Inc. and all rights are reserved therein. Unauthorized use is strictly prohibited.

NOTICE

This document and its contents are provided “AS IS” and are to be used subject to all of the limitation set forth herein.

Users of this document should not reference any initial or recommended methodology, metric, requirements, or other criteria that may be contained in this document or in any other document distributed by the Alliance (“**Initial Models**”) in any way that implies the user and/or its products or services are in compliance with, or have undergone any testing or certification to demonstrate compliance with, any of these Initial Models.

Any proposals or recommendations contained in this document including, without limitation, the scope and content of any proposed methodology, metric, requirements, or other criteria does not mean the Alliance will necessarily be required in the future to develop any certification or compliance or testing programs to verify any future implementation or compliance with such proposals or recommendations.

This document does not grant any user of this document any rights to use any of the Alliance’s trademarks.



OPEN DATA CENTER ALLIANCESM USAGE: INPUT/OUTPUT (IO) CONTROLS

EXECUTIVE SUMMARY

Cloud computing's potential is all about providing an IT service that is elastic to a large range of demand fluctuations. But cloud services are not yet immune to unanticipated downtime, making capacity management a top priority.

Capacity management in most enterprise IT environments has primarily been focused on CPU, memory and storage capacity, without significant focus on IO capacity. But as multi-tenancy and sufficient service level agreements (SLAs) for quality of service (QoS) of cloud environments are established, the impact of properly monitoring and controlling network and storage IO at various levels of the environment will become increasingly important.

The Open Data Center AllianceSM recognizes the need for bandwidth control in order to remedy inefficient use of physical systems caused by IO bottlenecks. This Usage Model is aimed at ensuring organizations can create and launch virtual machine (VM) workloads that meet their storage and network IO performance requirements and effectively manage IO performance requirements. At the same time, it seeks to ensure that providers of cloud services have the technical capability to efficiently and effectively manage IO demands from multiple running VM images.

PURPOSE

As host hardware becomes increasingly more powerful, the ability to increase VM density also increases. This can lead to potential bottlenecks in the IO performance of the system, making it hard to realize any gains. Part of the problem is that capacity management in most enterprise IT environments has not given significant consideration of IO capacity. What's more, controls over IO do not generally allow management of IO on a per VM (or more granular) basis. This can result in uncontrolled contention between VMs (i.e., a noisy neighbor) and failure to meet QoS targets for an application or workload. The ability to see potential contention, the cause of existing contention and the ability to control contention needs to be strictly managed to minimize the consequence of decreased performance for applications running on multi-tenant infrastructure.

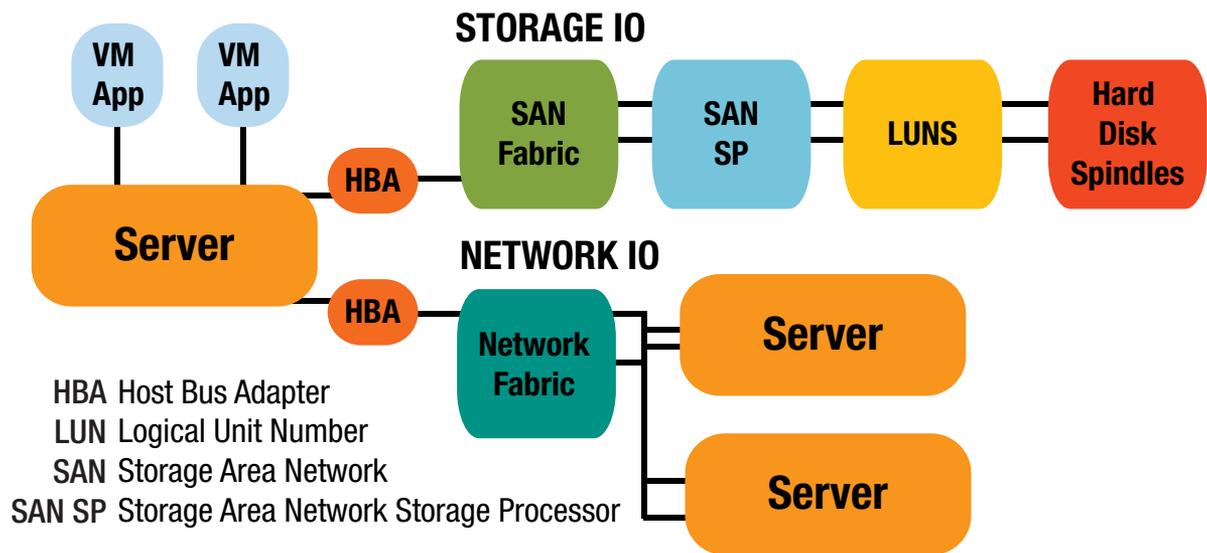
To encourage potential solutions to these IO issues and drive implementation requirements, this Usage Model focuses on making network and storage resources fully manageable. This means making aspects like guarantees, limits, and QoS manageable by the infrastructure or provider of the cloud service and exposed to the customer through appropriate controls and monitors. The intent is to provide bandwidth partitioning to bandwidth allocation by task, VM or time of day, using priority scheduling and bandwidth throttling. The key here is development of performance targets by type of IO and being able to guarantee these performance targets. The ability to manage this level of control requires an understanding of all potential thresholds or bottlenecks of the underlying infrastructure, and even more importantly the IO requirements of the workload in idle and peak scenarios. Knowing application/workload requirements in relation to infrastructure capabilities will help ensure appropriate mapping happens at provision time, as well as help enable automated control with dynamic changes during runtime of the workload. It is important to note that many workloads will require high IO at peak times, and latency and bandwidth controls for IO should be focused on ensuring that QoS can be met for the workloads providing mission critical and other important services.

TAXONOMY

Actor	Description
Cloud-Subscriber	A person or organization that has been authenticated to a cloud and maintains a business relationship with a cloud.
Cloud-Provider	An organization providing network services and charging Cloud-Subscribers . A (public) Cloud-Provider provides services over the Internet.

Through the use of sufficient monitors, **Cloud-Subscriber** consumption of IO and **Cloud-Provider** ability to provide IO can be balanced appropriately. Ideally, the applications and workloads that a **Cloud-Subscriber** submits to the cloud would be closely matched to the appropriate multi-tenant environment where the impact of the workload would not cause issues for other tenants, and where other tenants' workloads would not constrain the throughput and latency requirements of their cloud neighbors.

CONTEXT DIAGRAM



IO CONTROLS USAGE MODEL

Goals:

1. To ensure that **Cloud-Subscribers** have the capability to create and launch VM workloads that meet their storage and network IO performance requirements.
2. To ensure that the **Cloud-Provider** has the technical capability to manage IO demands from multiple running VM images in a way that ensures that an individual VM workload's IO cannot adversely and unexpectedly impact service to another running VM workload.

Considerations:

1. **Cloud-Providers** may provide varying levels of assurance based on price, and would therefore provide the appropriate level of instrumentation to the **Cloud-Subscriber** to ensure that the VM workloads are matched to the appropriate underlying infrastructure.
2. In order to allow the **Cloud-Subscribers** to better understand their workload requirements, the **Cloud-Provider** will need to provide appropriate visibility (through monitors), allowing the **Cloud-Subscriber** to see workload characteristics in idle and peak scenarios. This will enable both **Cloud-Provider** and **Cloud-Subscriber** to improve long-term mapping of IO requirements to infrastructure capabilities.

Assumptions:

Assumes the [National Institute of Standards and Technology](#) (NIST) Usage Model “[Allocate VM Instance.](#)”

Success Scenario 1 (instrumented):

1. The **Cloud-Provider** shall be able to provision fully instrumented VMs and hosting infrastructure to obtain complete visibility of IO activity on a per component (physical and virtual) basis.
2. The **Cloud-Provider** is therefore able to monitor IO consumption and take necessary steps (e.g., provision additional IO capacity) to ensure that **Cloud-Subscriber** requirements are met over the mid/long term.
3. The **Cloud-Subscriber** is notified of the level of IO performance assurance available and, when appropriate, informed of the need to upgrade to a higher level of assurance based on the **Cloud-Subscriber's** workload characteristics.
4. The **Cloud-Subscriber** is able to utilize on-demand monitors and reports to see its workload/application characteristics and understand the peaks and valleys of utilization of IO.

Failure Conditions 1:

1. The **Cloud-Provider** is unable to: 1) identify the sources of IO traffic through the hosting infrastructure, and/or 2) quantify the volume and rate of IO from each VM through the infrastructure.
2. The **Cloud-Subscriber** is not able to acquire necessary IO to ensure performance and throughput of the running workload/application. In a multi-tenant environment this leads to constraints across the entire infrastructure that is utilizing the IO paths.

Success Scenario 2 (partial):

1. The **Cloud-Provider** shall be able to provision fully instrumented VMs and hosting infrastructure to provide complete visibility of IO activity on a per component (physical and virtual) basis.
2. The **Cloud-Provider** is able to assign each VM image relative weightings/performance shares that are then used by the hosting infrastructure to determine the share of available IO that each **Cloud-Subscriber** VM may make use of at any one time.
3. These shares are enforced deterministically by the hosting infrastructure to manage contention between individual VM IO demands.
4. Correct provisioning is verified and confirmed to the **Cloud-Provider**.
5. The **Cloud-Subscriber** is notified of the level of IO performance assurance available and the allocated performance shares relative to the overall capacity.
6. The **Cloud-Subscriber** is able to utilize instrumentation to determine when the workload characteristics are impacted by throttles implemented through the performance shares.

Failure Conditions 2:

Same as Failure Conditions 1, but in addition, the infrastructure is unable to allocate available IO capacity based on individual VM weightings/performance shares. VMs may then exceed the assigned relative performance constraints and negatively and non-deterministically impact the available IO capacity available for other VMs.

Success Scenario 3 (full):

1. Same as Success Scenario 2 (partial), but in addition the **Cloud-Provider** is able to configure full end-to-end IO QoS management, enabling storage and network specific SLAs (e.g., target storage service times) with assured deterministic QoS to be assigned to **Cloud-Subscriber** VMs (either individually or in predefined groupings).
2. Correct provisioning is verified and confirmed to the **Cloud-Provider**.
3. The **Cloud-Subscriber** is notified of the level of IO performance assurance available and the allocated IO characteristics.

Failure Conditions 3:

Same as Failure Conditions 2, but in addition the infrastructure is unable to allocate available specific IO capacity to an individual VM. VMs may then exceed the desired performance constraints and negatively and non-deterministically impact the available IO capacity available for other VMs.

Failure Handling:

1. For all failure conditions, both the **Cloud-Provider** and the **Cloud-Subscriber** should be notified of the inability to provide IO assurance.
2. Failure Conditions 3 should result in an automatic attempt to achieve Success Scenario 2 (partial).
3. Failure Conditions 2 should result in an automatic attempt to achieve Success Scenario 1 (instrumented).

IO CONTROLS USAGE REQUIREMENTS

At a fundamental level it is expected that all usage requirements are multi-vendor and open. Key requirements need to be met for each of the infrastructure component vendors and hypervisors.

Usage Requirements	Description
Monitoring	<p>For Cloud-Provider:</p> <ul style="list-style-type: none"> • Monitor network and storage at individual VMs • Monitor network storage IO: throughput, latency • Monitor latency and throughput at individual component level and aggregate level • Monitor aggregate network IO capacity and bandwidth • Monitor aggregate storage IO capacity and bandwidth <p>For Cloud-Subscriber:</p> <ul style="list-style-type: none"> • Network and storage IO reservations – per VM • Aggregate workload IO consumption – by hour, day, week, etc.
SLA Metrics	<p>For Cloud-Subscriber:</p> <p>Per VM:</p> <ul style="list-style-type: none"> • Average latency/time-period, max latency/time-period, min latency/time-period • Average throughput/time-period, max throughput/time-period, min throughput/time-period
APIs	<p>Representational State Transfer (REST) and Web Service (WS) APIs for SLA definition, monitoring, reporting, etc.</p>
Timeslice Monitoring and Control	<ul style="list-style-type: none"> • Time granularity over which reservations are met to ensure definable latency • Throttling Thresholds: 1-Sigma, 2-Sigma, 3-Sigma deviations from mean
IO Reservations	<p>For Cloud-Subscriber, these reservation attributes: Min, max, average – network IO shares/VM</p> <ul style="list-style-type: none"> • Min, max, average – storage IO shares, IO Operations per Second (IOPS)/VM • Latency and throughput (see SLA Metrics section) <p>For Cloud-Provider:</p> <ul style="list-style-type: none"> • Definition of IO share that is independent of machine • Allocation of different IO capacity to different VM sharing the same pool of IO resource

Cloud-Providers would be equipped with the necessary monitoring tools to allow viewing visually and through APIs levels for all component and aggregate IO constraints. The monitoring tools would enable threshold views for each component and aggregate levels for the entire cloud platform to allow both proactive remediation and reactive monitoring, as well as expose issues with the cloud platform that would lead to SLA misses for the **Cloud-Subscriber**. We would expect that the solution implementation would allow connection through APIs and standard connection methods to enable interoperability between both existing manageability solutions and new standards-based solutions. The **Cloud-Provider** will be able to set thresholds at each appropriate level of the cloud platform to automate monitoring and throttling or migrating **Cloud-Subscriber** workloads to ensure the SLA is being met.

Cloud-Subscribers will be able to view how the **Cloud-Provider** is meeting their SLA requirements at an aggregate level for their landed compute, storage and network capacity. Using tools provided by the **Cloud-Provider**, **Cloud-Subscribers** will be able to analyze their workloads/applications to find issues with IO (e.g., too much IO for the workload) that would allow them to tune their workload and meet performance costs at a potentially lower tier of cloud platform.

SUMMARY OF INDUSTRY ACTIONS REQUIRED

In the interest of giving guidance on how to create and deploy solutions that are open, multi-vendor and interoperable, we have identified specific areas where the Alliance believes there should be open specifications, formal or de facto standards, or common intellectual property-free (IP-free) implementations. Where the Alliance has a specific recommendation on the specification, standard or open implementation, it is called out in this Usage Model. In other cases, we will be working with the industry to evaluate and recommend specifications in future releases of this document.

The following industry action is required to refine this usage model:

- The Open Data Center Alliance needs to engage with the DMTF in defining standards around resource unit for consumption, allocation and enabling platform support for consistent use across all ecosystems.