Media Rating Council, Inc.

420 Lexington Avenue
Suite 343
New York, NY 10170

Tel: (212) 972-0300
Fax: (212) 972-2786
www.mediaratingcouncil.org

February 8, 2017

To: Digital Measurement Vendors Subject to MRC Audit

From: George Ivie, David Gunzerath and Ron Pinelli

Re: Invalid Data-Center Traffic

Abstract: Related to GIVT filtration and use of non-industry invalid data-center traffic lists:

- ***Invalid data-center traffic associated to the 3 largest hosting entities should be filtered;***
- ***Lists should be maintained by vendors and subject to continuous evaluation and update;***
- ***Vendors should supplement these lists with other information, sources or techniques;***
- ***Use of proprietary/outsourced lists requires support;***
- ***Proprietary/outsourced lists must be subject to Business Partner Qualification;***
- ***GIVT data-center sources should be documented and communicated to industry leads; and***
- ***SIVT data-center sources are subject only to the limited communication requirements.***

Background:

On October 27, 2015 the Media Rating Council (MRC) issued the final *Invalid Traffic (IVT) Detection and Filtration Guidelines, Version 1.0*. The guidelines can be found here:

http://mediaratingcouncil.org/101515_IVT%20Addendum%20FINAL%20(Version%201.0).pdf

The *IVT Guidelines* provide examples of General Invalid Traffic (GIVT) and Sophisticated Invalid Traffic (SIVT), including Section 1.1.2, which states:

GIVT: Known data-center traffic (determined to be a consistent source of non-human traffic; not including routing artifacts of legitimate users or virtual machine legitimate browsing).

The *IVT Guidelines* also state:

For a start, we are proposing that the TAG (with MRC staff assistance, where necessary) maintain lists of findings (identified IVT sources) – where applicable, IAB and/or TAG filtration lists will continue to be facilitated, as expanded by the requirements of this addendum.

The filtration of invalid data-center traffic contemplates the availability and use of industry lists in order to promote consistency amongst vendors. While measurement vendors are strongly encouraged to utilize available industry lists, in lieu of the use of such industry lists, measurement vendors must seek alternate means to develop filtration rules for this type of invalid traffic. While impression-level granularity in filtration is preferred, as a starting point, the MRC is recommending filtration of ***invalid*** data-center traffic originating from IPs associated to the three largest known hosting entities: Amazon AWS, Google and Microsoft. To be clear, this

means filtration of IPs within those of known hosting entities determined to be a consistent source of non-human traffic; not including routing artifacts of legitimate users or virtual machine legitimate browsing.

We believe that this list could be readily compiled based on information currently and commonly available for immediate use by measurement vendors (with the use of exception strings to reduce false positive filtration of legitimate traffic as discussed in the *IVT Guidelines*). This list should be maintained by vendors and subject to continuous evaluation and update.

Vendors should supplement these lists with other information, third-party sources or other techniques in order to meet the objective of effective filtration of this type of invalid traffic. However, as discussed in the *IVT Guidelines*, vendors are required to empirically support identification of invalid traffic or filters as well as to analyze and minimize material false positives resulting from them. The use of proprietary or outsourced lists requires auditable support and is subject to Business Partner Qualification requirements outlined in the *IVT Guidelines*.

The presence of proprietary or multiple commonly available lists utilized to filter GIVT runs counter to a primary objective of this category of IVT in that it potentially detracts from consistency in application across vendors. By design, GIVT filtration is intended to be consistent between vendors. For this reason the MRC intends to work with TAG to develop a periodic function to routinely ingest and compare such lists with results made available on a controlled basis.

Finally, measurement vendors will also routinely identify invalid data-center traffic through analysis of traffic patterns or more sophisticated techniques beyond use of known data-center sources. Section 3.5 of the *IVT Guidelines* states:

Each measurement organization should have functions devoted specifically to communications related to IVT matters. ***Routine communication functions are limited to General Invalid Traffic detection and processes. Sophisticated Invalid Traffic detection and processes should be closely controlled and subject to limited communication to staff of industry oversight bodies as required, etc., these should only occur on a broad generalized basis when major new issues [of new methodologies for creating and monetization of invalid traffic types] are discovered, and information about such discoveries should be communicated in a manner that maximizes the effectiveness of reducing IVT.*** *Communication processes should encompass: (1) ensuring internal notifications are provided as necessary to foster awareness and clues to detecting invalid traffic (referred to as "internal communications"), (2) communication with industry leads in this area – specifically IAB staff, IAB TAG and MRC staff (referred to as "industry communications"), (3) communication of learning and best practices in a facilitated manner to other industry practitioners to encourage ecosystem improvements (referred to as "outside practitioner communications"), and (4), as necessary, communication to law enforcement and/or measurement service legal counsel on significant invalid traffic matters (referred to as "legal communications").*

To the extent that data-center sources are identified via general techniques such that they can be filtered under the definition of GIVT, such sources should be routinely documented and communicated to industry leads in a timely manner as required above. Measurement vendors may choose to classify invalid data-center sources as SIVT. Such classification should be

documented and supported by auditable evidence, but is subject only to the limited communication requirements discussed above.

The MRC has produced this interim guidance based on input from an IVT Update working group and until such time as there is a formal standards update that incorporates it, this interim guidance is considered authoritative and should be applied by measurement services in the MRC accreditation process.

Please contact Ron Pinelli at MRC (rpinelli@mediaratingcouncil.org) with any questions.