

Frederick L. Cottrell, III Director 302-651-7509 Cottrell@rlf.com

December 9, 2008

REDACTED PUBLIC VERSION

VIA ELECTRONIC FILING AND HAND DELIVERY

The Honorable Vincent J. Poppiti Special Master Blank Rome LLP Chase Manhattan Centre, Suite 800 1201 North Market Street Wilmington, DE 19801-4226

> Re: Advanced Micro Devices, Inc., et al. v. Intel Corporation, et al., C.A. 05-441-JJF; C.A. 05-485-JJF; MDL No. 05-1717-JJF

Dear Judge Poppiti:

Pursuant to the Stipulation and Order Regarding Intel Discovery Into AMD Evidence Preservation, entered by Your Honor on November 25, 2008, AMD submits this Statement Re Status of Intel's "Histograms."

I. Introduction

On October 9, 2008, Intel sent a letter asserting that AMD has "systemic" evidence preservation problems. Intel has yet to identify what these purported "systemic" issues are. Instead, Intel has embarked on repeated fishing expeditions, casting about for any possible anomaly, real or imagined, on which to base an assertion of "systemic" error.

Intel's latest gambit is the generation of dozens of "histograms," which are bar charts that Intel contends demonstrate email preservation problems by AMD custodians. On November 14, 2008, Intel provided AMD and Mr. Friedberg with histograms for 79 AMD custodians.¹ Intel has also announced that it is preparing and intends to submit histograms over the next few weeks for every other AMD custodian, 179 in all. Intel maintains that AMD should assume the burden of analyzing all of these histograms and rebutting whatever it is Intel purports them to show.

¹ These 79 histograms included replacements of 35 histograms Intel had previously submitted on October 9, 2008, all of which Intel later admitted to be erroneous.

One Rodney Square = 920 North King Street = Wilmington, DE 19801 = Phone: 302-651-7700 = Fax: 302-651-7701

As Your Honor suggested during a hearing on November 20th, and with the subsequent assistance of the Special Master's advisors, Mr. Friedberg and Ms. Martin, the parties have settled upon a sample set of 21 histograms (slightly more than 25% of the total provided thus far), which AMD agreed to analyze and report on in the first instance. Intel and AMD each picked 10 sample AMD custodians, and Ms. Martin added one additional custodian of interest who was not already on the parties' lists. Following the joint selection of these sample histograms, AMD provided Mr. Friedberg and Intel with its preliminary analysis in writing, following which Mr. Friedberg conducted a December 5, 2008 telephone conference to discuss AMD's preliminary findings and thereby to enable him to assist Your Honor.

This Statement sets forth AMD's analysis concerning 20 of the 21 sample custodians selected,² and our suggestions about how the process for assessing Intel's histograms should proceed going forward.

AMD's Statement consists of this letter brief and attachments, and an Appendix that consists of custodian-specific analysis for each of the sample custodians. The Appendix also contains "counter-histograms" that AMD has itself generated to accurately depict the relevant data.

II. Background Regarding Intel's Histograms

Intel's histograms are bar charts that purport to show a custodian's monthly volume of email produced by AMD from that custodian's own files, as well as what are known as "OCFs"---that is, unique emails that the custodian purportedly sent or received that were produced from <u>o</u>ther <u>c</u>ustodians' <u>files</u>, but not from the custodian's own. Intel's histograms contain "yellow shading" which, according to Intel, is meant to depict and quantify "the precise number of emails that should have been, but were not, produced in the custodian's data." (*See* Intel's letter dated October 9, 2008, at p. 1.) Intel's histograms were accompanied by lists of DCNs (document control numbers) corresponding to each of the email files Intel claims to constitute a unique OCF for each custodian.

Since the time these histograms were submitted by Intel just under four weeks ago, a large team of AMD counsel and vendor personnel have spent literally hundreds of hours -- including over the Thanksgiving holiday weekend -- analyzing these histograms and the accompanying lists of roughly 120,000 DCNs. That effort has already cost AMD a tremendous amount of money, and has also diverted resources from many other important case tasks.

As we will show, Intel's histograms grossly exaggerate the presence of OCFs, and do not demonstrate any "systemic" issues. Based on AMD's analysis thus far, which encompasses 20 of the 21 sample custodians, Intel has overstated OCFs by at least 50%, and by almost 100% for certain individual custodians. The number of incorrectly-attributed OCFs in Intel's histograms

² As explained below, AMD has not been able in the time allotted to comprehensively assess Intel's histogram with regard to one AMD custodian,

will likely grow as further work is performed. To this point, AMD's analysis has been accomplished with the aid of some automation; what lies ahead is a laborious, manual effort that will require weeks to complete. Because Intel is equally capable of running the manual search strings on the remaining OCFs, we will conclude this Statement with the suggestion that if Intel cares to pursue its claim of systemic issues with regard to these histograms, it should be required to conduct the manual search itself, certify to AMD and Your Honor that it has properly completed it, and then generate new, corrected histograms that *accurately* portray true OCFs and eliminate all of the "false positives."

III. Observations About OCFs

The presence of OCFs in a large document production is unsurprising and to be expected. To be sure, Intel's own production includes massive quantities of OCFs. Indeed, the centerpiece of Intel's remediation plan is its reliance on OCFs to supplement its own custodians' productions that were decimated by the auto-delete function it neglected to switch off.

Nor is the presence of OCFs within AMD's production a new issue. In September 2007, Intel raised this very issue with AMD. At that time, Intel identified a number of custodians with OCFs -- that is, custodians whose productions did not include items apparently authored or received by them, but which were found in the production of other custodians' files. At some considerable expense, AMD thoroughly investigated those allegations with respect to the very first AMD custodian on Intel's list, **but and and every** one of the 593 supposedly missing emails, or OCFs, that Intel had attributed to him. AMD communicated this to Intel. (A copy of AMD's September 14, 2007 letter setting forth this analysis is attached hereto as **Exhibit** A.)

As Intel knows through this September 2007 exchange, through discussions between the parties, and through discovery -- including the informal discovery in which the parties are now engaged -- there are many reasons that OCFs may exist.

<u>First and foremost</u>, OCFs will inevitably occur whenever human beings are required to make individual judgments. Every custodian must necessarily make personal, on-the-fly decisions -- in some cases, perhaps a thousand or more of them each month -- about whether a given email is or is not within the scope of the preservation instructions given to him/her. In a production of this magnitude, it is to be expected that one custodian may judge the responsiveness of a given email differently than another custodian looking at the same item.

<u>Second</u>, OCFs will often result from the exercise of different relevance judgments by reviewing attorneys looking at the same document. In short, different reviewers looking at the same items in different custodians' data sometimes come to different judgments about relevance and responsiveness. Therefore, the fact that an email was produced from the files of one custodian does not necessarily mean that a second custodian who was also party to that email communication did not also preserve it.

<u>Third</u>, OCFs will occur as a by-product of the fact that AMD's processing vendor (like many e-discovery vendors using state-of-the-art processing) applies deduplicating and near-deduplicating protocols to email collections as part of routine data processing. The culling of "near-duplicates" is fully explained in the document attached hereto as **Exhibit B**, which was provided to Intel counsel on October 15, 2007. The effect of this is that only the longest, unique string of an email is produced; the identical email "fragments" of that longer email chain that may exist separately within the custodial collection are culled, and not produced as separate email items. Thus, as fully disclosed and explained to Intel more than a year ago, Intel may identify what it contends is a unique OCF when it is, in fact, wholly contained within a longer email string that was both preserved and produced by the subject custodian.

<u>Fourth</u>, OCFs will occur in those presumably unusual cases in which different attorney reviewer decisions may be made about whether a document is privileged, such that a document deemed privileged when reviewed in one custodian's files may not be so viewed by another reviewer looking at another custodian's files, with the result that it is produced as part of one custodian's data but not the other's.

The presence of an OCF does not necessarily mean, therefore, that the subject custodian did not retain that very same document. To understand the reason for an OCF, each must be examined individually. While some OCF analysis can be performed electronically with the aid of programming (which itself imposes substantial programming and processing time and expense), finding all of the duplicate email "fragments" within a longer, deduplicated email chain cannot be. Instead, this requires manual review that, depending on volume, can entail very significant and costly work.

III. AMD's OCF Assessment Method and the Burden Intel Improperly Inflicted

In this part, we describe the method AMD and its processing vendor, Forensics Consulting Solutions ("FCS") used to assess Intel's purported OCFs and to identify falselyattributed OCFs. We also summarize the burden this exercise has already imposed on AMD -- a burden that AMD believes Intel could easily have reduced substantially by undertaking a proper analysis, using data available to it, before firing off dozens of erroneous histograms.

Intel claims that, over the time period from March 2005 through November 2006, there are 120,300 OCFs attributable to the sample AMD custodians (other than **March 2005**). Intel provided DCNs for these files. To assess this, FCS developed electronic programming that allowed it to compare Intel's purported OCFs to the sample custodians' email collections. Generally described, FCS first aggregated associated metadata for Intel's purported OCFs, and assembled the set of emails from the custodian's population where the custodian was either a sender or a recipient of the email. Through processes of "exact matching" and "ThreadHash" matching, FCS was able to identify falsely-attributed emails and track the results. It is the figures so derived that are reflected in this letter brief, and in the written summaries and AMD's "counter-histograms" that are attached in the Appendix.

As set forth in more detail below, FCS has thus far determined through this process that Intel falsely attributed more than 50% of the 120,300 OCFs. For some custodians, the error rate approaches 100%.

The electronic methodology outlined above has required substantial effort. AMD's vendor, FCS, was required to spend significant time developing, programming and executing this process. This method does, however, utilize attributes of the electronic email files that facilitate assessment in a semi-automated sense. AMD believes that the remaining OCFs will be further reduced by the manual, document-by document review and comparison process described above, which might perhaps be accompanied by some text-searching methodology. Specifically, AMD believes based on its experience last year chasing down OCFs that many of the remaining Intel-characterized OCFs are part of larger email strings maintained or produced by the subject custodian that the electronic process can not identify. Top level metadata from these files simply does not permit ruling out all false OCFs electronically. To execute this part of the OCF review exercise, Intel's false OCFs would have to be identified through creation and assignment of comparative email collections to document review attorneys, who necessarily would have to review and compare each document and track results by hand. Additionally, AMD believes that some of the remaining OCFs may be falsely-attributed but are part of the data repository maintained by AMD's alternate processing vendor, Stratify, Inc. FCS has collected that data but is encountering some difficulty in manipulating it in the electronic process described above. This work remains in progress. It is this issue that precluded AMD's ability to provide a full analysis and assessment in relation to AMD's

Thus far, AMD has incurred hundreds of thousands of dollars of expense, and has spent hundreds of hours of its lawyers' and consultants' valuable time, to conduct these analyses and respond in an expedited fashion to Intel's OCF assertions. We have already determined that Intel's attribution of OCFs is wildly inaccurate and exaggerated. And while we have now begun the laborious manual review needed to attack the remaining OCFs, AMD does not believe it should be its sole burden to complete it.

Intel has had in hand the data needed to eliminate many, if not most, false OCFs. For at least two of these 21 sample custodians **sector and the sector and**

³ Intel's errors do not appear to be limited to **second and the second and analysis** continues, AMD believes that there are at least 7 other custodians within the sample set of custodians whose productions contain exact matches of at least some of Intel's purported OCFs.

simply made its overblown OCF assertions and attempted to put AMD to the task and expense of debunking them.

IV. AMD's Assessment of Intel's Histograms

A. <u>Overall Statistics Applicable to the Sample AMD Custodians</u>

Statistical observations relevant to the overall population of sample AMD custodians are worth making. As noted already -- and as Intel should have predicted -- Intel's histograms vastly overstate the number of actual OCFs. For the 20 sample AMD custodians analyzed, Intel asserted that a total of 120,300 unique OCFs exist for the time period from March 2005 through November 2006. AMD's analysis thus far shows that 62,910 of Intel's purported OCFs were, in fact, retained by the subject custodians and/or produced from their files.⁴ Actual OCFs are thus at least 52% lower than Intel has claimed. Put another way, Intel has overstated actual OCF figures by at least 110%.

Intel has also calculated that AMD produced 308,320 emails from the actual files of these 20 sample AMD custodians. As such, Intel is contending that OCFs represent over 28% of the global production for these custodians. The data actually reveal that OCFs comprise less than 16% of the total production. Intel, of course, has yet to take a position as to what this type of data shows or means about AMD's production -- or, indeed, about Intel's own production.

Interestingly, during the parties' conference with Mr. Friedberg and Ms. Martin on December 5, 2008, Intel stated for the first time that its concerns are limited to alleged OCFs through May 2006. This was a curious comment, since Intel's lists of DCNs included more than 22,000 alleged OCFs during the period from June 2006 through November 2006.⁵ At any rate, even excluding the period following May 2006, the results are not materially different. Over that somewhat shorter time frame, Intel identified 97,916 purported OCFs. Of that total, we have thus far determined that 48,602 are falsely attributed to the sample AMD custodians. This represents an error rate -- thus far -- of 50%, substantially the same as the 52% error rate in the period including June-November 2006.

B. <u>Custodian-Specific Analyses</u>

Assessment of OCFs, data retention and productions patterns, and generating the statistics applicable to both is mostly a custodian-specific inquiry and exercise. Indeed, each of the sample AMD custodians had varying levels of OCFs attributed by Intel, and differing

⁴ The total number of false OCFs that FCS has identified has increased from 62,871 to 62,910 since AMD's December 5, 2008 report to Mr. Friedberg.

⁵ Because Intel included those post-May 2006 DCNs with its histograms, AMD went to the trouble and expense of analyzing them. Regrettably, this appears to be a part of Intel's overbroad and inappropriate effort to saddle AMD with burdensome tasks, to which Intel now responds, "Never mind."

production patterns and file counts. The results of AMD's analyses thus naturally vary depending on these and other custodian-specific traits, such as idiosyncratic emailing and preservation habits.

AMD has set forth the bulk of these custodian-specific analyses in its written summaries and "counter-histograms" attached in the Appendix. As noted, the written summaries provide data, statistics and AMD's observations of certain relevant patterns. These summaries are most easily understood when viewed along with AMD's counter-histograms.

AMD's counter-histograms supplement the written summaries by depicting numerically and pictorially several things in three different charts. The first chart is titled "Total Sent and Received Items." An example chart appears below.



This chart provides a single, multi-colored bar representing the entire production for each custodian, including OCFs, for each month of the time period identified. The "red" shading depicts production from the custodian's own files and provides a file count number (which AMD took from Intel's histograms except for the state of the state of

OCFs that we have thus far determined Intel to have erroneously attributed to that custodian for that month. The charts also show, with vertical dotted lines, the date on which AMD delivered a litigation hold notice to the custodian, and the date (if applicable) that the custodian's email account was migrated to AMD's automated archiving systems. In addition, in the upper right hand corner is a box in which we have set forth the total number of OCFs Intel alleged for the custodian; the OCFs from that population that FCS has thus far located within the custodian's collection; and the resulting percentage reduction made thus far in the OCF total Intel has alleged.

The second and third charts are the "Total Sent Items" and "Total Received Items," respectively. The same bars, shading and dotted lines appear as described above.

C. Analysis of Custodian Categories and Apparent Data Trends

Comments can also be made about data trends we thus far seem to be finding with different groups of custodians.

1. <u>Analysis of Custodian Categories</u>

The sample AMD custodian histograms appear to fall into three general categories. The first is those custodians for whom OCFs are low by any reasonable standard and, therefore, appear to us to raise no issue and to warrant no further analysis. One example is **a standard** the subject of the September 2007 exchange with Intel on the topic of OCFs. Intel has attributed a total of 474 sent and received item OCFs to **a standard** over the time period from May 2005 through April 2006 -- which covers both the pre-archiving and post-archiving time periods. AMD's analysis shows that, at most, only 24 OCFs actually exist over this entire timeframe. In other words, **b** himself retained over 99.6% of the subject email files that were ultimately produced by AMD on **b** behalf. Other custodians -- like **b** and **b** appear to be part of this group.

The second category of custodians is those whose OCFs or email file counts exhibit anomalous patterns, for example, relatively high numbers of OCFs in a certain window of time with lower numbers at other times, or a high number of sent items relative to received items during some time frames. Whether such an unexpected pattern is the result of a custodian's idiosyncratic preservation habits, a corrupted .pst file or other similar issue, a failure by Intel to fully take into account the custodian's production, or something else, requires a custodian-bycustodian assessment.

AMD's **AMD's as a sample custodian selected by Intel, is one example.** In the selected at deposition about his idiosyncratic method of preserving data, which included using his sent items folder as his primary preservation repository.

AMD's **MAD**'s **is another example**. Intel apparently contends that **mathematical** failed to retain sent items prior to the automated archiving of his email account in November 2005, and points to the existence of sent item OCFs. Our analysis thus far demonstrates that Intel has

overstated total sent item OCFs in this time period by almost 50%. Indeed, the data show that himself preserved, and AMD produced, 6,795 emails from April through October 2005, an average of over 970 emails per month. However, when all is said and done, for a single month (April 2005), production does appear to contain a high number of OCFs relative to files produced from his own collection. AMD has not yet completed its ongoing investigation on this point.

The third category is those custodians whose actual OCFs are consistently high relative to the number of files produced from their own collection. Perhaps the best example is AMD's In the time period from March 2005 through May 2006, **Sector** total monthly OCFs appear to average about 400. The total number of files produced, both from collection and actual OCFs, is very close to this average every month. In addition, there are no "gap" months of obviously low total file counts, and there are no sharp discontinuities in total produced email volumes month-to-month.

So far, AMD has not discovered anything to indicate that the relatively low number of email files which himself appears to have retained is a result of a data collection anomaly. In addition, which we retain the relatively small number of files compared to OCFs is not the result of a failure by AMD to impose a proper litigation hold or to monitor it. Instead, as detailed in which we written summary, AMD issued a comprehensive litigation hold to him within a week after this lawsuit began. AMD thereafter provided with numerous written reminders about preservation. Instead, selectivity in deciding which files were relevant and which were not does not reflect what AMD would have preferred. That is not, however, the consequence of some failure by AMD to exercise reasonable efforts to secure compliance by with AMD's preservation instructions.

2. <u>Trends Apparent in the Data</u>

Several trends throughout the sample custodian data also are apparent.

As to a number of custodians, Intel asserted that a large number of OCFs existed in the post-archiving time period⁶ (which varies by custodian, but is often November 2005). In fact, FCS' analysis shows across the board that actual OCFs in the post-archiving period are negligible.⁷ Indeed, OCF totals in the post-archiving time period often reduce to zero or only a few emails, as they did, for example, for **Example**. The absence of material

⁶ The "post-archiving time period" represents the period following the custodian's migration to an automated system for preserving email; such a system, which AMD began implementing in the fall of 2005, is a "passive" preservation system in that it does not require or depend upon the exercise of custodian judgment.

⁷ AMD's current working supposition is that the OCFs present in the post-archiving time period may be the product of the Stratify data repository on which AMD has not yet been able to run analyses. Even were that not the case, the actual OCF numbers in that time period are *de minimus* by any standard.

numbers of OCFs in the post-archiving time period suggests that AMD's archiving systems are working effectively to capture custodian email, and that AMD was able to successfully extract that data from its archives.

We also observe that there is reason to question the utility of any analysis that compares volumes of sent items to volumes of received items and draws conclusions from that comparison, as Intel seems to have done. First, some custodians simply do not send much email, at least email that is deemed responsive to Intel's document requests. (*See*, for example, AMD's counter-histograms for **Example and Context and Text and Te**

Second, disparities between a custodian's sent and received email volumes may also be the product of idiosyncratic, but perfectly appropriate, emailing habits. Consider, for example, AMD's **management** testified that he used his sent items folder as his email preservation repository. AMD's analysis shows that Intel's purported sent item OCFs reduce to virtually zero in almost every month, while reductions in received items OCFs number in the hundreds almost every month. Indeed, total OCFs are 48% lower than Intel alleged, with an overall reduction of 3,318 of Intel's alleged OCFs. This type of result is found with other custodians, and particularly those who saved more sent items relative to received items, or *vice versa*. (*See, e.g.*, AMD's counter-histograms for

among others.) There is good indication that for these and other custodians the purportedly missing "sent" OCF was actually within the "received" collection for the subject custodian, or the other way around. It thus follows that disparities between the volume of sent versus received email does not by itself have any particular significance.

Assessing the differences between sent and received items is both difficult and often meaningless. Consider this example: Custodian A sends an email to Custodian B and copies Custodian C. Custodian B receives the email and replies only to Custodian A. Custodian A now has two email items: A shorter sent item, and a longer received item. Custodian B also has two email items: A shorter received item and a longer sent item. Custodian C, on the other hand, was not a party to the reply and thus has only one email item: the shorter received item. AMD will produce the shorter received item for Custodian C. As a result of near-deduplication, however, AMD will only produce the longer received item for Custodian A and the longer sent item for Custodian B. Of course, both items incorporate a complete copy of the shorter email. Yet under this scenario, Intel's OCF identification method would have improperly counted the shorter item produced by Custodian C as two OCFs: For Custodian A, the shorter item is a missing sent item even though it was produced within a longer received item; for Custodian B, the shorter item is a missing received item even though it was produced as part of a larger sent item. The distinction between sent and received items in this example thus has no probative value. In both cases, Intel would have received a complete copy of the shorter item in Custodian A and B's production and could have located it within the text of the larger items prior to carelessly asserting that Custodians A and B failed to retain it.

Many other permutations are possible, such as an instance where Custodian B is a "cc" on Custodian A's email, and Custodian B "replies to all." When he does so, Custodian B will receive his own reply as a received item, and will have the identical item in his sent items. Whether Custodian B's practice is to preserve the item in his sent mail, or in his inbox, is a matter of personal habit, although that personal habit could substantially affect the relative size of the custodian's sent item vs. received item collection.

V. Conclusions and AMD's Suggestions For Future Analysis

AMD has shown that Intel's OCF analysis, and analysis overall, is fraught with error. Intel's overstated OCF allegations are the most obvious. AMD considers it a serious problem that Intel knew the probable explanations for OCFs from the September 2007 experience involving **but, in its desire to inflict massive, unnecessary cost and burden on** AMD, made material assertions that are simply untrue -- and which could have been avoided had Intel assumed the burden, as it should have, to rigorously analyze and test its assertions first.

AMD's analysis also allows us to draw this conclusion with certainty: Contrary to Intel's repeated assertions, the issues and so-called anomalies that Intel has raised with respect to the sample AMD custodians, and generally, are in no sense "systemic." Quite to the contrary, the issues here are unique and custodian-specific, and the explanation for them depends, and will continue to depend, on the characteristics of such things as each AMD custodian's emailing habits and tendencies, preservation practices, individually-retained file counts and actual OCF totals. The anomalies encountered are simply not, as in Intel's case, the consequence of a failure to disable an aggressive, systematic auto delete, a systematic failure to migrate custodians to backed-up servers; or a systematic failure to notify custodians of their preservation obligations.

For all the reasons set forth in these materials, AMD believes that the sampling and analysis it has done so far is more than sufficient to demonstrate the absence of any systemic issue or error. Intel should not be permitted to continue to force AMD to toil away on dozens more pointless OCF hunts, especially at AMD's expense.

Instead, AMD submits that Intel should be required to do two things before any further proceedings about "histograms" take place. First, Intel should be required to do the work it should have done at the outset to identify all false OCFs and reduce its allegations to those OCFs that are truly unique. It has the data it needs and the capacity to do so. Before any additional histograms are presented, or further response is required of AMD, Intel ought to be required to certify that it has carried this burden that properly is placed on it.

Second, with regard to each and every histogram Intel presents to AMD and the Court -whether now-existing or later-generated -- Intel ought to be ordered to state and define with particularity and in writing precisely what "anomaly" or other issue it believes the histogram shows, and what Intel contends should be done about it. Indeed, so far, Intel has done nothing specific whatsoever to identify its complaint about each custodian or what justifies its complaint. It is too late for that. And without this kind of specificity and clarity -- where Intel lays all its

cards on the table in an open, direct manner -- AMD is being forced to play a very expensive and distracting game of pin the tail on the OCF. That is not a game AMD should be forced to play.

AMD looks forward to discussing these issues with Your Honor during the conference call scheduled for 2:00 p.m. EST on December 12.

Respectfully,

/s/ Frederick L. Cottrell, III

Frederick L. Cottrell, III (#2555) Cottrell@rlf.com

FLC,III/III Enclosures

cc: Clerk of the Court Eric Friedberg, Esquire (w/e) (By Electronic Mail) Jennifer Martin, Esquire (w/e) (By Electronic Mail) Donn Pickett, Esquire (w/e) (By Electronic Mail) Richard L. Horwitz, Esquire (w/e) (By Electronic Filing) James L. Holzman, Esquire (w/e) (By Electronic Filing)