Memory Errors in Modern Systems

The Good, The Bad, and The Ugly

1140310323 蒲毅

## 论文类容简介

最近的一些出版物表明，在内存子系统中的硬件故障如同家常便饭，并且预测这些故障在将来的包含多级DRAM和SRAM的系统中将更加的常见。本文指出在这些内存子系统中，需要提供容错的弹性技术，特别是在高性能计算和含有数万数据节点的数据中心。因此，了解他们的功效是非常重要的，以确定它们是否将适合于未来的系统。

在本文中，作者使用两个数据领导级的高性能计算机系统分析硬件韧性方案在当前的系统的可靠性影响。他们的研究有几个关于当前部署的可靠性技术，如DRAM ECC，DDR的效率，主要结果地址/命令奇偶校验，和SRAM ECC和奇偶校验。作者还进行方法学考察，发现计数错误不是错误，研究人员的普遍做法和数据中心运营商，可导致不正确的关于系统的可靠性的结论。最后，作者用数据预测未来大规模系统的需求，发现该SRAM的故障不可能构成显著较大在未来可靠性威胁，而DRAM故障会有一个主要问题，更强的韧性DRAM方案将需要保持类似于可接受今天的系统中发现的那些故障率。

然后论文分了11各部分进行了讨论。

**第一部分 Introduction：**

**介绍了论文所讨论问题的背景，同时叙述了作者所做工作与千人的不同与创新之处，最后还提出了三种分类：**

**(1) The Good : 对错误表现以及系统稳定新的正确理解。比如：DDR明林的地址奇偶校验，粒子碰撞对SRAM故障的影响，大多数未修正的SRAM的错误是由于单比特错误造成的。**

**(2) The Bad：有些领域还有很多研究和深入的理解。比如：海拔的高度增加了一些DRAM设备的故障率，虽然这种效应因供应商的不同而存在差异，和片内的SRAM不同，未来系统的external memory（DRAM）将需要一个更加强大的容错技术，甚至比当进的超级计算机的都要强。**

**(3) The Ugly : 存在潜在问题的并且当前广泛采用的技术手段。比如：执行实地研究是困难的，SEC-DED ECC。**

**第二部分 Terminology**

**定义了文章所用的基本术语。其中：**

**Fault 由会引发error和不会引发error 分为显式和隐式**

**Error 有能否被检测和能否被改正分类。**

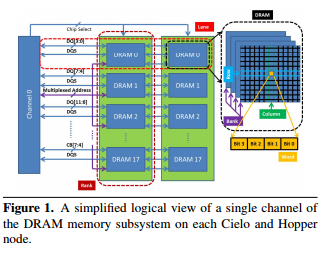
**第三部分 dicusses**

**介绍了相关的研究领域以及介绍和作者所做研究的差异，及其研究和方法。**

**其中指出的前人所做研究的不足之处：往往没有较长的时间间隔和较大的规模，并且强调了作者所做的研究规模较大，并且间隔时间比大多数研究都要长，使其能够呈现对故障和错误代表性的数据。**

**第四部分 介绍了Cielo和Hopper 的特点**

**Hopper包含大约6000个计算节点。每个节点包含两个12核AMD Opteron TMprocessors，每个有十二个32KB L1数据缓存，十二个512KB二级高速缓存和一个12MB的L3缓存。每个节点有八个4GBDDR-3内存，DIMM共计DRAM的32GB的。**

**Cielo包含大约8500个计算节点。 每节点包含两个8核心AMD OpteronTMprocessors，每八32KB L1数据缓存，八512KB L2缓存，和一个12MB的L3缓存。每个节点具有八个4GB的DDR-3内存，DIMM共计DRAM的32GB的。**

**第五部分 描述实验方法和细节**

1. **介绍了实验所用的三种数据集:** corrected error messages from console logs, uncorrected error messages from event logs, and hardware inventory logs.
2. **数据指出至多会有2个DRAM出现多重错误，概率极低忽略不计**
3. **随机选择单元进行测试，以排除偶然误差。**

**第六部分 列出在Hopper实验中的数据并对其进行分析**

**第七部分 证明对失误率计算得到的稳定性比对错误率计算的到的稳定性更佳的准确**

**第八部分 分析的结果**

1. **SEC—DED ECC 系统不适合于现代DRAM子系统，在大范围使用的时候，失误率非常高。**
2. **DDR命令和奇偶地址校验的使用价值非常高，并且非常可靠**
3. **SRAM中如果加入单位校验部件如奇偶校验那么将会成为非常可靠的方法。**

**第九部分 实验中总结的经验**

1. **失误是不可预测的**
2. **对于任何设备，仔细的分析不可或缺**
3. **分析能力很关键并且很难**
4. **不同数据的规模会带来完全不同的实验结果**

**第十部份 根据实验结果分析未来系统的趋势**

**SRAM误差率会在未来逐渐稳定下来，但是DRAM的错误率会不断升高，所以需要有新的技术来弥补DRAM的缺点。**

**第十一部分 总结**

**在未来纪念中，可靠性将继续是显著挑战。了解故障所经历的性质在实践中可以受益所有利益相关者，包括处理器和系统架构师，数据中心运营商，甚至应用开发者，在追求设计出更大规模的弹性数据中心和系统。在本文中，我们提出了DRAM和SRAM数据断层，量化几个硬件韧性的影响技术和可靠的有关制度设计汲取经验。我们的研究结果表明，尽管系统多年来取得显著进步（例如，从移动SEC-DED ECC来CHIPKILL在DRAM子系统），但依然需要更多的工作，来提供一个强大的平台作为未来的大规模计算系统。**

## 论文审阅

大数据是一个趋势，所以大数据的储存带来的问题，比如本文提到的文件失误将会是分布式研究专家不可避免的一个问题，本文就此问题做了实验。

个人认为本论文好的地方如下：

1. 通过在两个不同的分布式系统上进行实验，一个是hopper，一个是ciego，而不是在单一的平台上进行实验，从而排除了实验结论可能会因平台不同而带来的不同，从而使结果更加的精确，并且也更加有说服力。
2. 提出了设备的表现和海拔有一定的关系，虽然不同厂商的芯片在同一海拔的性能不同，并且同一厂商的芯片在不同海拔上也会有不同的性能。思维独特，有创新！
3. 分段多，适合慢读，条理比较清楚。

不足之处：

1. 说了很多当前的很多不足之处，但并没有提出一个相对有效且简单的方案来弥补当前DRAM的不足之处
2. 不够通俗，不够易懂。