



## 美洲狮的怒吼

# AMD PUMA+低功耗APU技术与性能分析

不要以为AMD与英特尔的战场只有酷睿与Kaveri, AMD与英特尔早已展开一场全面的战争。从英特尔推出Atom, 进入超低功耗处理器市场开始, AMD就针锋相对地推出了相应的产品和英特尔竞争。之前AMD已推出过代号分别为Brazos、Kabini和Temash的多代超低功耗产品, 不过市场反响不算出色。而在今年四月底, AMD则发布了它们全新一代、代号为Beema和Mullins的超低功耗APU。这一次, AMD又有怎样的改变? Beema和Mullins又有哪些绝活呢?

文/图 张平

事实上, AMD在超低功耗CPU市场早已耕耘多年, 去年还发布了每瓦特性非常出色的Kabini和Temash APU, 但市场反应并不理想。在这些产品中, 尽管低功耗APU的TDP最低能够下探至15W, 而其中超低功耗的1GHz双核心产品功耗更是降低至3.9W, 但相比英特尔已经将超极本的Haswell处理器TDP功耗降低到15W以内, 同时能

提供更强悍的CPU性能而言, AMD的这些产品除了价格以外就没有太多亮点。尤其是英特尔发布了Bay Trail架构的产品后, 英特尔展示了自己是如何使得类似的产品运行在更高的性能上, 同时TDP又保持在8W以内的。这样一来, AMD必须更新自己的产品线, 才能尽可能多地保持自己的竞争力。

终于在2014年, AMD发布了

代号为Beema和Mullins的超低功耗APU。虽然有两个代号, 但这两款产品实际上采用了完全相同的Puma+(美洲狮)架构。其中Beema面向的是入门级或者超便携的笔记本电脑, 而Mullins则更偏向于平板电脑。这两款产品都是为Windows市场而设计的。

从架构角度来说, AMD在Beema和Mullins上也并没有彻底



更新架构设计。Beema和Mullins的最根本改变在于将制程从TSMC 28nm转移到了GlobalFoundries的28nm, 同时在架构和设计上进行了一些调整。作为从AMD拆分出去的晶圆工厂, GlobalFoundries一直以来都在新工艺研发上困难重重, 这次的超低功耗处理器能够顺利在GlobalFoundries工厂投产, 证明GlobalFoundries已经从之前的技术泥潭中抽身而出。所以在了解AMD的新品之前, 让我们先来看看有关GlobalFoundries 28nm工艺的情况。

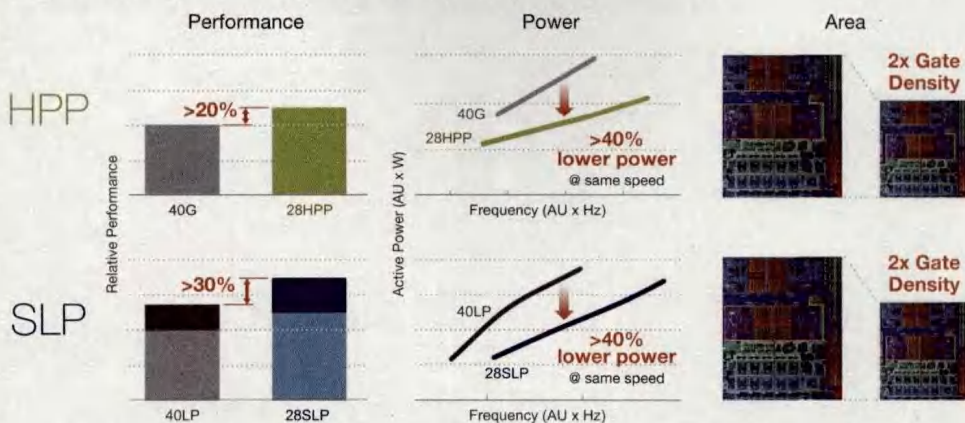
### 最后的Gate First? GlobalFoundries 28nm 工艺解析

说起GlobalFoundries在工艺上的跌跌撞撞, 不得不提及Gate First和Gate Last。Gate First和Gate Last实际上是在进入45nm以后, 以IBM为核心的阵营和以英特尔为核心的阵营在未来工艺发展上的一次分歧。最终以英特尔大胜、IBM失败告终。

Gate First和Gate Last都是半导体制造的其中一步, 其中Gate First是指先为晶圆生成金属栅极, 再进行漏区、源区的离子注入, 最后进行退火操作; Gate Last则完全相反, 其生产流程是先注入离子, 再退火, 最后才生成金属栅极结构。这两种方法在当时看起来各有优劣, Gate First工艺简单, 但是随后的高温退火可能影响栅极尤其是PMOS的性能, 并且对栅极金属性能要求较高。而Gate Last虽然对栅极金属性能要求不那么高, 但是工艺更为复杂。

IBM在Gate First上已经研究了10年之久, 并且IBM认为Gate First是进入45nm时代后最应该选择的工艺, 还成功拉拢了三星、TSMC、AMD作为自己的盟友, 而英特尔一

28nm-HPP and 28nm-SLP versus comparable 40/45nm platforms:



■ 目前GlobalFoundries主力推荐的28nm生产工艺有HPP和SLP两种。

方坚持Gate Last才符合未来发展的需求。不过随着研究进展的深入, IBM和盟友们发现, Gate First最多只能坚持到28nm, 就会由于材料和高温的问题而无法使用, 反倒英特尔是正确的, Gate Last虽然短期内看起来麻烦一些, 但是一直在28nm以后的工艺都可以继续使用。

当发现问题后, 三星虽然嘴上会说提供Gate First的产品, 但是却发布了新的论文, 将支持Gate Last的研究。TSMC也宣布在28nm以后全面转向新工艺, 不再考虑Gate First。不过AMD就有点麻烦了, 受制于财务问题再加上企业并购拆分, AMD的晶圆厂在工艺上的进展一直很不顺利, 随后拆分出去的GlobalFoundries在32nm工艺和28nm工艺上都摔了跟头, 直到2013年才彻底搞定了28nm Gate First High-k Metal Gate的全部技术问题, 得到了和预想中一样的产品。但是这个时候, 英特尔的FinFET 22nm已经投产多时, 14nm也已经箭在弦上了。

目前GlobalFoundries提供三

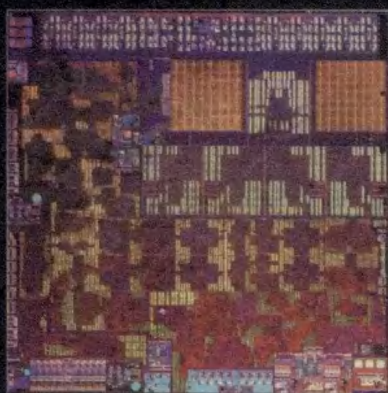
种28nm工艺供用户选择, 其中有专门为超低功耗设备设计的28nm SLP (Super Low Power)、为高性能设备设计的28nm HPP (High Performance-Plus) 以及兼顾高性能和低功耗的28nm LPH (Low Power, High Performance)。其中28nm SLP最为便宜, 28nm HPP价格最为昂贵, 中间的则是28nm LPH工艺。不过在比较新的宣传内容中, 已经看不到28nm LPH工艺的内容了, 可能是GlobalFoundries考虑到市场和生产的问题, 取消了LPH工艺。

根据GlobalFoundries的官方资料, 28nm HPP工艺和SLP工艺对比40/45nm时代的类似工艺, 有显著的优势。比如28nm的HPP工艺比40G工艺性能提高20%, 同时功耗降低大约40%; 28nm SLP工艺相比低功耗的40LP工艺速度提升高达30%, 功耗也最多降低了40%。无论是哪种28nm工艺, 芯片面积都只有40/45nm工艺的一半左右, 大大节省了成本。

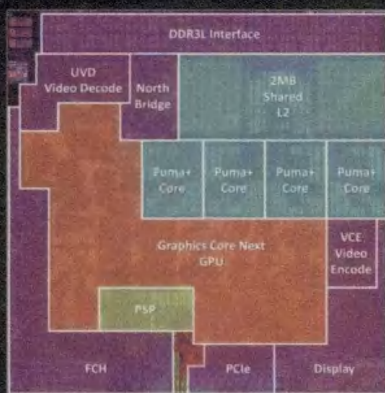
目前没有资料表示AMD的新



## MULLINS



AMD

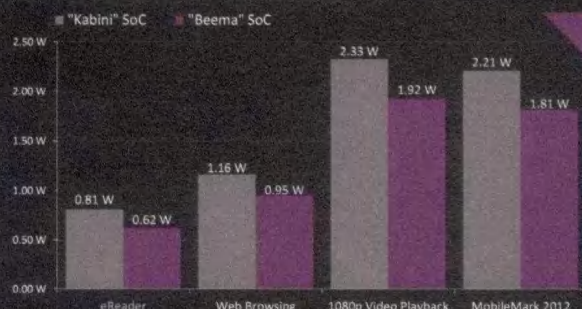


■ AMD公布的Mullins核心架构图,可以看到,它拥有四颗Puma+ CPU核心。

## REDUCED POWER CONSUMPTION, MORE BATTERY LIFE GENERATION OVER GENERATION

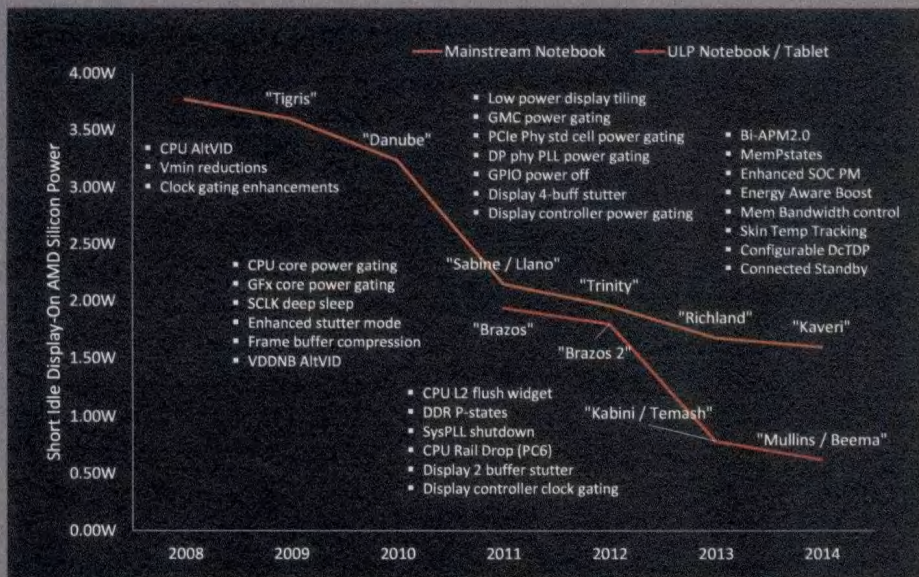
AMD

### "BEEMA" POWER IMPROVEMENTS



**Up to 20%**  
POWER REDUCTION<sup>21</sup>

■ 采用Beema架构的产品比上一代Kabini APU的能耗比综合提升了约20%。



■ AMD在功耗控制技术上获得了巨大的发展,处理器待机功耗逐年得到降低。

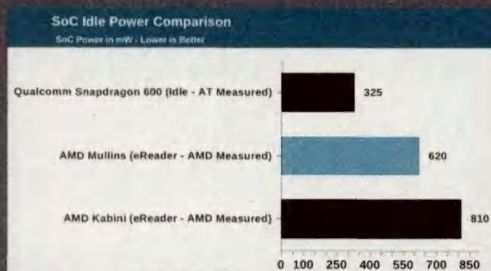
APU使用了哪种工艺,不过据推测使用28nm HPP的可能性更高一些,这个工艺本身就是面向高性能设备使用的,核心电压又比较低,能够达到更高的性能功耗比。至于28nm SLP,面向的是超低功耗领域,比如手机、蓝牙模块等,本身频率最高只能达到1.8GHz。当然不排除AMD分别使用这两种工艺的可能。总的来说,目前在AMD的Beema和Mullins上所使用的28nm工艺,很可能是最后一代使用Gate First的产品了,未来GlobalFoundries也将转向Gate Last工艺,并加入FinFET大军中去。

## 更低的功耗指标

前面我们已经说过, Beema和Mullins实际上并没有包含太多的架构方面的调整, AMD在转换工艺的同时, 加入了大量频率控制和功耗控制技术, 显著提高了产品的性能功耗比并降低了TDP指标。根据AMD给出的功能模块图, 一个典型的Beema或Mullins芯片拥有四个Puma+ CPU核心, 含有128个流处理单元的GCN架构GPU核心, 以及显示模块、PCI-E总线模块、UVD影像模块、DDR3L内存控制器、北桥、2MB共享L2缓存、VCE视频编码模块、FCH (Fusion controller hub) 南桥功能模块以及平台安全处理器PSP模块, 这所有的模块组合在一起组成了复杂的Beema或Mullins APU芯片。

从CPU架构来说, Beema或Mullins所使用的Puma+架构和之前的Jaguar是一样的, 它依旧是一个双发射的乱序执行架构, 内存位宽也维持为64bit。不过AMD宣称Beema和Mullins受益于新工艺, 在1.2V电压下核心漏电比前代产品降低了19%, GPU部分的漏电流流量低了38%, 这也是Beema和Mullins拥有更低的电压和更高能耗比的主





■ 测试表明AMD的Mullins APU在电子阅读应用下, 功耗相对上一代产品有明显降低, 不过依旧赶不上ARM架构的移动SOC。

表1: AMD和英特尔近几代产品工艺和核心面积情况

SOC芯片名称	工艺	晶体管数量	核心面积(平方毫米)
AMD Zacate	TSMC 40nm	0.450亿	75
AMD Kabini/Temash	TSMC 28nm	0.914亿	约107
AMD Beema/Mullins	GF 28nm	0.93亿	约107
AMD Llano	GF 32nm SOI	11.8亿	228
AMD Trinity/Richland	GF 32nm SOI	13.0亿	246
AMD Kaveri	GF 28nm SHP	24.1亿	245
Intel Haswell (4C/GT2)	Intel 22nm	14亿	177

要原因。

不仅如此, AMD还调整了Beema和Mullins的内存接口。这一代产品使用的是DDR3L这种低电压的产品, 而上一代产品则使用的是DDR3。相比DDR3L来说, DDR3由于要兼容各种类型不同的内存模块, 因此设计更为复杂。AMD从移动SOC那里得到了设计灵感, 使用了较为专用的DDR3L内存界面, 这样的改进使得运行在低功耗状态时, 内存的功耗降低了500mW。不过, 虽然目前的Beema APU已经基本完成了SOC化, 但是考虑到其定位并非手机这样空间狭小的地方, 因此AMD并没有考虑将内存芯片和APU芯片采用堆叠封装以减少面积, 依旧使用了传统的分离式架构。

从测试数据来看, 在电子书阅读、网页浏览、高清播放以及MobileMark 2012的测试中, Beema都展示出了相对于Kabini更为出色的功耗表现, 平均降低了约20%。此外, 从AMD历年来的产品来看, 2008年AMD推出的代号为“Tigris”的Turion处理器的待机功耗高达3.5W~4W, 随后代号为“Danube”的处理器将待机功耗降低到了3W左右。最近一次最显著的变化是第一代代号为“Llano”的APU中, 待机功耗大幅度降低到了2W左右, 最新的代号为“Kaveri”

的APU的待机功耗已经来到了1.5W附近。在超低功耗处理器方面, 第一代超低功耗处理器代号为“Brazos”的产品待机功耗甚至高达2W以上, 而最新的Beema和Mullins待机功耗降低到了0.5W左右, 着实令人惊讶。

为了达到这样的成果, AMD使用大约30项技术: 包括各种功耗门控技术、各种频率控制技术、电压控制技术、显示控制技术等, 甚至一些接口诸如PCI-E、DisplayPort的接口功耗控制都被纳入其中。AMD总结自己使用了四个方面的技术, 包括智能动态功耗管理、进一步整合系统元器件、电路功耗优化、生产工艺升级改进。这四个方面的技术配合一些特色设计一起努力, 最终实现了AMD目前在移动处理器上的低功耗表现。同时AMD还给出了有关电子书阅读模式下的功耗情况, AMD新的Mullins APU在电子书阅读状态下相比上一代产品功耗降低了大约20%~25%, 不过依旧无法和目前比较主流的ARM架构移动SOC芯片相提并论。不过AMD目前也不打算染指安卓设备, 这一点也是完全可以容忍的。

### 更高的频率提升空间

一般来说, 同时降低功耗并提

升频率是几乎不能完成的事情。不过在Beema和Mullins上, AMD在降低了功耗的同时还提升了频率。出现这样的情况, 主要原因是AMD使用了更为出色的频率控制技术, 并且调整了温度控制的阈值。首先请让我们来看看表2的数据, 表中展示的是AMD新的Mullins对比上一代产品Temash在最高频率和TDP功耗方面的对比情况。可以看到, A10 Micro-6700T的TDP功耗为4.5W, 最高频率为2.2GHz; 相对应的上一代Temash架构的A6-1450 TDP功耗为8W, 最高频率仅为1.4GHz。此外, 其他两款Mullins APU产品也分别提升了60%和40%的最高频率。而在表3中, 我们则对比了Beema与Kabini CPU核心的不同, 其中除了E2-6110对比E2-3000这一组产品略有降低外, 其余产品的CPU核心都获得了明显的频率增加, 并降低了TDP。

除了CPU部分的频率可以提升外, Beema与Mullins的GPU部分频率也获得了明显提升。如表4与表5所示。那么, AMD是怎么做到在生产工艺没有革命性改变的时候, 降低功耗并提高频率的呢? 原来AMD设计了一个名为STAPM的技术, Skin Temperature Aware Power Management即表面温度感应功耗管理。简单来说, 这个技术的本质实



表2 Mullins和Temash 架构CPU频率和功耗对比

Mullins架构产品型号	TDP功耗	最高CPU频率	Temash架构的对应产品	TDP功耗	最高CPU频率	Mullins CPU频率提升幅度
A10 Micro-6700T	4.5W	2.2GHz	A6-1450	8W	1.4GHz	57%
A4 Micro-6400T	4.5W	1.6GHz	A4-1250	9W	1.0GHz	60%
E1 Micro-6200T	3.95W	1.4GHz	A4-1200	3.9W	1.0GHz	40%

表3: Beema和Kabini架构CPU频率和功耗对比

Beema架构产品型号	TDP功耗	最高CPU频率	Kabini架构的对应产品	TDP功耗	最高CPU频率	Beema CPU频率提升幅度
A6-6310	15W	2.4GHz	A6-5200	25W	2.0GHz	20%
A4-6210	15W	1.8GHz	A4-5000	15W	1.5GHz	20%
E2-6110	15W	1.5GHz	E2-3000/E1-2500	15W	1.65GHz/1.4GHz	-10%/7%
E1-6010	10W	1.35GHz	E1-2100	9W	1.0GHz	35%

表4: Mullins和Temash架构GPU频率对比

Mullins架构产品型号	TDP功耗	最高GPU频率	Temash 架构对应产品型号	TDP功耗	最高GPU频率	Mullins GPU频率提升幅度
A10 Micro-6700T	4.5W	500MHz	A6-1450	8W	400MHz	25%
A4 Micro-6400T	4.5W	350MHz	A4-1250	9W	300MHz	16%
E1 Micro-6200T	3.95W	300MHz	A4-1200	3.9W	225MHz	33%

表5: Beema 和Kabini架构GPU频率对比

Beema架构产品型号	TDP功耗	最高GPU频率	Kabini架构对应产品型号	TDP功耗	最高GPU频率	Beema GPU频率提升幅度
A6-6310	15W	800MHz	A6-5200	25W	600MHz	33%
A4-6210	15W	600MHz	A4-5000	15W	500MHz	20%
E2-6110	15W	500MHz	E2-3000/E1-2500	15W	450/400MHz	11%/25%
E1-6010	10W	350MHz	E1-2100	9W	300MHz	16%

表6: Mullins系列主要产品规格

型号	图形核心	SDP场景设计功耗	TDP	CPU核心数量	最高CPU频率	L2缓存	流处理器数量	最高GPU频率	最高内存频率
A10 Micro-6700T	R6	2.8W	4.5W	4	2.2GHz	2MB	128	500MHz	DDR3L 1333
A10 Micro-6400T	R3	2.8W	4.5W	4	1.6GHz	2MB	128	350MHz	DDR3L 1333
E1 Micro-6200T	R2	2.8W	3.95W	2	1.4GHz	1MB	128	300MHz	DDR3L 1066

表7: Beema系列主要产品规格

型号	图形核心	TDP	CPU核心数量	最高CPU频率	L2缓存	流处理器数量	最高GPU频率	最高内存频率
A6-6310	R4	15W	4	2.4GHz	2MB	128	800MHz	DDR3L 1866
A6-6210	R3	15W	4	1.8GHz	2MB	128	600MHz	DDR3L 1600
A6-6110	R2	15W	4	1.5GHz	2MB	128	500MHz	DDR3L 1600
A6-6010	R2	10W	2	1.35GHz	1MB	128	350MHz	DDR3L 1333

际上还是通过提高芯片允许温度上限来实现更高的频率——这和英特尔在Bay Trail上做的事情差不多。

同时，平板电脑本身是拥有一

定散热能力的，为了衡量并更好地使用这个能力，AMD还设计了一个称之为TSP功耗的值。这个值是通过在安装了四核心Mullins APU的

11.6英寸平板电脑上运行3DMARK 06时，设备表面温度升高至用户可以接受的极限状态来确定的。对一个4.5W TDP的Mullins处理器来



说, TSP瓦数为3.5W。换句话说, TSP就是设备能够容忍的热量极限, 在实际运行中, 虽然芯片在不停地散发热量, 但是设备拥有一定的热存储和热迟滞的能力, 使得即使芯片温度达到比较高的程度, 但设备依旧处于安全、舒适的温度范围内。

而上一代AMD低功耗产品仅控制芯片温度, 只要芯片温度超过60℃, 那么立刻会进入降频状态, 无论CPU还是GPU都是这样。但是设备温度的上升速度总是远远落后于芯片温度, 并且设备温度的上升过程是缓慢的, 甚至在芯片满载状态下数分钟之内都不会导致设备过热。所以, AMD新的STAPM设计更为优秀: 只要设备的温度控制在一定的范围内(用户不会感觉到设备温度太高), 同时提高芯片温度上限, 那么芯片就可以更长时间运行在更高的频率上。

新的Beema和Mullins针对每一个设备增加了额外的温度测试接口, 专门用于测试设备表面温度, 只要这个温度依旧安全, 那么处理器就会稳定运行在高频率上——当然, AMD还设置了一个阈值, 芯片温度不能超过100℃。当设备的外部温度达到了设定上限或者芯片达到了100℃时, 处理器都会立刻降低频率以保证安全。AMD宣称这样的设计不需要改变处理器本身结构, 也不会带来额外的负担。据信AMD会联合设备OEM厂商对每款不同的设备外部设置相应的温度探头, 以保证在长期运行时设备的安全与稳定性。

此外, AMD还特别说明了新的动态频率调整是完全智能的, 频率调节会根据软件的情况来进行合理的加速。

## 加入ARM安全模块

早在2012年, AMD就宣布自己购买了ARM的相关授权, 准备开发基于ARM的产品。当时AMD宣称

和ARM合作会主要集中在AMD的Opteron处理器上, 这类服务器处理器将使用ARM的Cortex-A5架构用于执行相关的安全操作。实际上AMD需要类似的硬件安全平台才能更好地保持自己的竞争力, 相比之下, 英特尔有命名为TXT的安全模块, 并且没有授权给AMD, 所以AMD才开发出来了TrustZone。

事实上在上一代的产品中, AMD已经为Kabini和Temash APU集成了Cortex-A5核心, 不过当时由于种种原因他们并没有启用。在全新的Beema和Mullins上, AMD终于启用了安全模块并命名为平台安全处理器, 也就是Platform Security Processor, 简称为PSP。

AMD新加入的PSP模块拥有单独的处理器、ROM和SRAM, 支持目前比较常见的诸如ECC、SHA、RSA、AES、Zlib、TRNG等多种加密算法, 并且能够直接访问系统内存, 调用资源。这样的设计使得AMD的新产品在运行有关安全设置的应用时显得更为得心应手。

## 能耗比提升极高 基于AMD Mullins的平板性能测试

那么新一代低功耗APU的性能到底如何呢? AMD率先为我们展示了他们采用Mullins APU的平板性能。AMD的这款参考平板尺寸为11.6英寸, 显示屏幕分辨率为1080p, 安装了Windows 8.1操作系统, 处理器型号为Mullins中的最高端产品A10 Micro-6700T。

首先让我们来看看在JavaScript性能和网页浏览的测试结果, 在这些测试中, AMD的产品均获得了领先, 尤其是以较大幅度胜出了英特尔的Bay Trail和苹果A7。而在CPU单线程性能测试中, AMD的Mullins比英特尔的Silvermont Atom Z3770快了大约

### SKIN TEMPERATURE AWARE POWER MANAGEMENT WITH STAPM

- Boost aggressively until Tskin reaches user defined maximum
- Reduce power only when necessary to adhere to Tskin,max limit
- Most use-cases for mobile devices are short in duration → Result is higher performance most of the time
- All without using more power

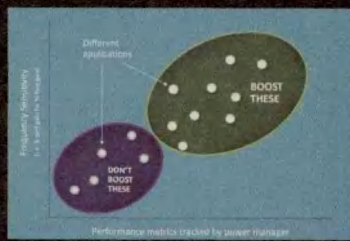


DYNAMIC SKIN-TEMPERATURE AWARE POWER MANAGEMENT CAN ENABLE UP TO 63% PERFORMANCE INCREASES IN KEY WORKLOADS

■ AMD的STAPM技术很好地解决了产品的温度和频率控制问题。

### AVOIDING POWER WASTE WITH INTELLIGENT BOOST CONTROL

- Intelligent Boost is designed to avoid power waste that results from boosting applications that benefit very little from higher frequency
- Enables long battery life and cool operation while maintaining great performance
- Power management micro-controller tracks application behavior real-time to determine frequency sensitivity
- Boost behavior is adjusted accordingly



■ AMD的加速技术会针对不同的应用采用不同的方案, 智能决定是否加速。

### AMD PLATFORM SECURITY PROCESSOR (PSP) HARDWARE

1" AND ONLY x86 PROCESSOR TO INTEGRATE AN ARM CORE

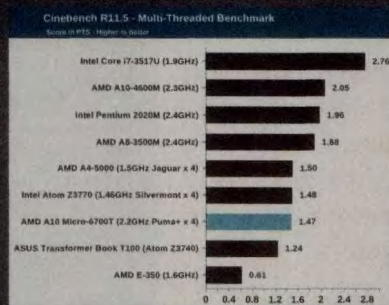
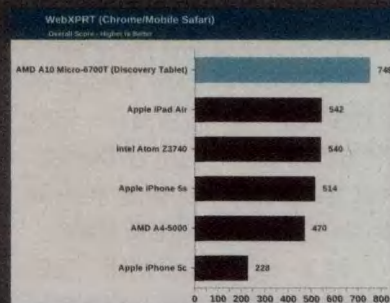
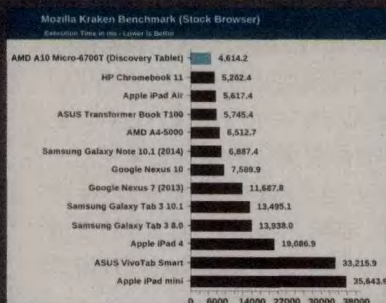
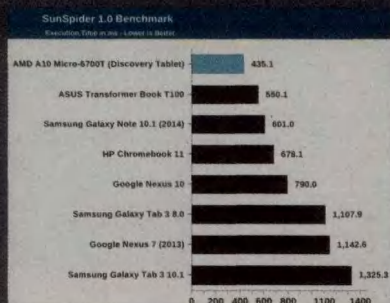


■ AMD在产品中加入了ARM架构的核心, 作为专门的安全处理器, 保障整个设备平台的安全。

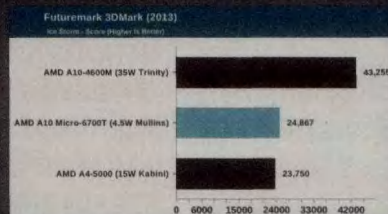
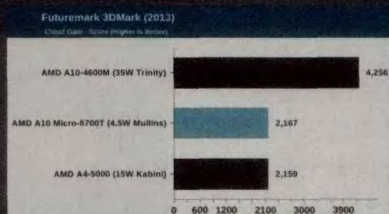
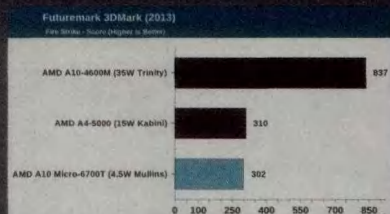
35%, 且能提供大约相当于80%的AMD桌面APU A10-4600M的性能, 这是非常令人惊讶的。和第一代APU Llano架构的A8-3500M相比, Mullins则大约能达到其85%的性能。而在CINEBENCH R11.5 CPU多线程性能测试中, Mullins和英特尔的对比产品性能基本相当, 甚至赶上了上代Kabini这样TDP高达15W的产品, 要知道这款A10 Micro-6700T的TDP功耗仅为



■ 在JavaScript性能测试中，AMD Mullins APU较竞争对手拥有明显的优势。



■ 在CPL性能测试中，Mullins APU已达到AMD桌面级处理器的80%，多线程性能则与英特尔同级产品基本相当。



■ 4.5W的Mullins在GPU性能上已经能够同15W级别的Kabini匹敌，能耗比极高。

4.5W。此外，相比AMD之前老架构的E-350，性能翻了一番还多。

最后是GPU性能测试，测试软件是3DMark，使用的场景是Fire Strike和Cloud Gate。在这项测试中，4.5W的Mullins依旧展示了和上一代15W级Kabini极为接近的性能。虽然没有直接对比英特尔的产品，不过考虑到Kabini有大约比英特尔Bay Trail强50%到200%的图形性能，Mullins应该也有类似的表现。另外一点是目前的Trinity移动处理器，它的功耗相比Mullins高了大约10倍，不过性能领先幅度却只有2倍多一点，这展示了Mullins极为优异的性能功耗比。

## 实际功耗将成成功关键因素

AMD新的Beema和Mullins使得AMD在超低功耗平台上终于拥有了一款优异的产品。根据性能测试和AMD官方功耗数据来看，AMD提供了一款和去年15W TDP的产品性能相似，但是实际TDP功耗只有4.5W的高能耗比处理器，同时还保持了非常强大的图形性能，这都给人留下了深刻的印象。

如果一切顺利的话，AMD的新产品在市场上有可能会具备比较强大的竞争力。不过目前最大的问题在于，AMD还没有公开Beema、Mullins与英特尔对应的Bay Trail系列产品进行功耗对比的实际测试，

究竟Beema和Mullins的功耗表现到底如何，电池续航时间有多长，还需要更多的测试才能确定。AMD预计Beema和Mullins在接下来的一到两个季度中，就会有实际产品上市了，其实目前联想、三星等厂商已经发布了基于AMD Beema和Mullins的产品。

另外，AMD是否有兴趣在安卓平台上开疆拓土呢？看起来Mullins非常适合高性能的安卓平板电脑，不过迄今为止AMD都只在Windows平台上发力。根据目前的市场情况来看，AMD很有可能出现在安卓市场上。未来的一切，只有等待时间来告诉我们答案了。■