



加微信：1716143665，领取配套福利课程



043 | 芯片7：芯片难题的短期解决方案（上）

卓克·科技参考2（年度日更）

18/小时前

043 | 芯片7：芯片难题的短期解决方案（上）

12分29秒

| 卓克亲述 |

众筹新课联系微信：**1716143665**，你好。

欢迎回到《科技参考》，我是卓克。

在上周的内容里，我们曾经说过，传统芯片设计发展到 2012 年之后，摩尔定律就开始失效。此后，实现设计的成本会越来越高，而不是再像之前那样，单位晶体管的成本每 18 个月减一半。

我们也说过，想要彻底解决这个问题，就得跳出当前的冯·诺依曼结构，比如我们曾经介绍过的存算一体。

此外，还有其它的解决方法，比如使用模拟芯片，或者寻找硅基以外的新材料。甚至是寄希望于今天遥不可及、今后肯定会普及的量子计算机。

但一切能彻底解决问题的方法都没法解决当前的困难，因为人们实在是等不起。

这些刚刚出现萌芽的技术，想要发展到今天硅基材料、冯·诺依曼结构的计算机的程度，至少还需要三四十年，甚至上百年的时间。这可是几代人职业生涯的跨度。

那么，中间这段过渡期，人们用什么解决芯片性能提升的困境呢？答案就是我们今天的《科技参考》要介绍的 Chiplet 技术。

Chiplet的本质是什么

Chiplet 的意思是“小芯片”，也有翻译成“芯粒”的。实际上，这是一种用搭积木的方式，把一堆小芯片组合成一块大芯片的技术路线。而小芯片，指的就是那些积木的基础模块。

为了更直观的介绍 Chiplet 的优势，我们打个比方，用“在一块玉上雕刻工艺品”的过程来说。

传统的芯片工艺和封装大约是这样的：任务就好比是把一栋建筑物的内外结构通过工匠精湛的手艺，立体地雕刻在一块玉中。

这类工艺，我们其实在故宫也见到过，有的是用木头雕的，有的是用一块玉雕的，有的甚至是用一整块象牙雕的。

既然是雕刻建筑物，那就有简单和复杂之分了。过去的年代，雕刻一间两居室就可以了。但随着雕刻工艺的进展和客户需求的提高，现在已经要求我们在一块玉里雕刻出一栋别墅了。

到了这一步，工艺实在没法继续大幅提高了。但客户又提要求了，给我刻一套更复杂的北京地铁结构图。

如果继续按从前的思路，这么复杂的结构需要的空间太大了，世界上压根就没有足够大的一整块玉来刻，没办法，那就放弃吧。

但现在，正是因为有了 Chiplet 技术，这个任务还是能勉强完成的。具体方法是这样：

首先，不再用一整块玉了，而是用 N 块，比如说 3000 块玉。把每块玉作为单独的模块，分别完成雕刻。比如说，西直门站用一块玉，东直门站用一块玉。如果有些地方结构太复杂，像北京南站那样的地铁站，那么一层用一块玉，二层一块，三层一块。

当这些站点都雕刻好了以后，中间的连线也不必用一整块玉去雕刻模拟，就简单的用细铜丝连接。然后，再把这 3000 块玉和互联的铜丝全部镶嵌在一个立体的、像蜂窝一样的木质结构中。这样一套北京地铁的雕塑，照样是精美的、满足客户要求的。

对比来看，从前非要用一块整玉来实现这个任务，那是不可能完成的，但自从使用了 Chiplet 技术，任务勉强完成了。

Chiplet有什么优势

当然，刚才咱们只是打个比方，现在问题来了，这个比喻分别对应了芯片设计、制造、封装的哪些环节？这么做又有什么好处呢？

首先的好处就是刚才说的，太过于复杂的芯片，原本用过去的技术没法制造的，现在可以使用 **Chiplet** 技术分模块完成了。

这里的关键词就是那个“一块整玉”，它对应半导体技术中的掩膜版（mask）。掩膜版是什么呢？

你可以简单理解成是照相的底片。但和照相不同的是，一个完整设计的芯片，并不是由一张底片冲印出来的，而是由十几张到几十张底片按照特定的顺序，叠加后冲印出来的。之所以张数特别多，和工艺步骤有关。

比如说，在制造芯片的过程中，首先，需要让 A 区域被光照射到，照射半小时后，A 区域的顶层就被那种光晒没了。接着，需要把 A 区域挡住，用另外一种光围绕 A 区域的边缘一圈照射。

这个过程就需要两张像幻灯片一样的掩膜版：一张 A 区域是透光的、其余部分不透光；另一张沿着 A 区域边缘一圈是透光的，其余部分不透光。

但掩膜版的尺寸是有限制的。今天，最大的掩膜版大约也只有 850 平方毫米。再大的话，特定的光透过掩膜版后，波长就会不那么精确，就会导致该融化的地方不融化。这就会直接导致芯片制造失败。

而一旦要求太多的晶体管容纳在一起，比如说 1000 亿个晶体管在一个芯片里，需要的掩膜版就超过了最大限度。这就相当于要找出一块能把整个北京地铁都雕刻进去的玉，是不可能完成的任务。

而现在，用了 **Chiplet** 技术，不用再找一块那么大的玉了，分模块就能完成。

其次的好处就是，这样做便宜。

什么意思呢？其实，哪怕客户提出的要求并没有超出单块整玉能雕刻的指标上限，但只要临近这个上限，制造成本就会很高。

比如，掩膜版的尺寸上限是 850 平方毫米，在这个尺寸下，生产出的芯片的合格率只有 15%。但如果掩膜尺寸是 250 平方毫米，这时生产出的芯片合格率就高达 90%。

所以，如果能用 250 平方毫米的掩膜版制造出 N 个合格的小模块，再把它们合理的镶嵌在一起，这么做的成本，远比用 850 平方毫米的掩膜版一次成型要低得多。

而除了良率高能降低成本之外，还因为可以使用不同工艺处理芯片不同的部分，于是也能降低成本。

这是什么意思呢？我们都见过 CPU 的样子，就是一个 PCB 基板上扣着一个铜盖。但盖子下面是什么呢？

从前，盖子下面就是一颗晶莹剔透但又非常脆弱的核心（die），但今天，如果你揭开一些 CPU 的铜盖，就会发现，里面有若干块晶莹剔透的核心。

这就是传说中的多核心吗？不是的，有些部分确实是计算核心，但有些部分是做其它事情用的。比如说，那些负责逻辑运算的部分，必须用 5 nm 的工艺做，这样才能塞进更多的晶体管。

但还有其它部分，比如内存控制器，或者一些处理输入、输出接口的控制器，就不是非要用 5 nm 的工艺做了，用 14 nm 的工艺一样没问题。

以前没有 Chiplet 设计时，只有那么一颗晶莹剔透的核心，于是大家就都要用 5 nm 工艺来实现。但现在，其中一部分用 14 nm 这种很老旧的工艺就可以了，成本自然就能降下来。

Chiplet的技术演进

Chiplet 的发展，经历过几个阶段：

早在 2012 年，就有一些公司把同样性质的小核心单独制造，最后再把几个拼成一组、整体封装起来。在第一阶段期间，小核心的设计都完全相同，相当于把一个营的步兵升级成一个团的步兵了。

第二阶段中，拼在一起的小核心就具有不同的功能和不同的工艺了。就像我刚刚说的，一个铜盖子下既有 5 nm 的核心，也有 14 nm 的核心，它们分别承担不同的任务。这个叫作 Chiplet 的异构。

第三阶段中，拼在一起的小核心不但功能和工艺不同，甚至使用的半导体材料都不一样。

在第一和第二阶段里，那些小核心都是硅基材料，但到了第三阶段，有些小核心会采用氮化镓（GaN），有的采用碳化硅（SiC），有的采用磷化铟（InP）。之所以用不同的材料，是因为芯片里集成了太多从前根本不可能容纳进来的设备，比如射频芯片、光电芯片、功率芯片等等。

第四阶段是今天正在进行当中的。之前三个阶段，芯片功能的丰富、性能的增加都是在二维平面展开的。你需要多一些计算性能，就多加一个计算核心的模块进去；需要光电芯片，就多加一个磷化铟为基底的模块就行。

但是，第四阶段开始往三维空间发展，也就是垂直方向上也开始堆芯片了。

最初让人惊讶的二维异构 Chiplet 芯片就是苹果手表里的那个 S1 芯片。这里除了有传统的 APL0778 处理器，负责逻辑计算；还有博通的 BCM4334 通信芯片，负责蓝牙和 WiFi；还有来自 AMS 公司的 NFC 信号放大器；8 GB 的闪存与尔必达的 512 MB 的运存；意法半导体的陀螺仪；ADI 的触控模组和 IDT 的无线充电芯片……

所有这些都封装在一个铜盖子下。苹果的 S1 芯片让二维异构芯片的概念一下子普及开来。这种创新直接点燃了智能手表行业。

而三维的异构 Chiplet，最早受启发于 SSD 硬盘里存储单元的开发。

比如，三星已经做出了把 176 层存储单元堆叠在一个单位上的存储器，这让单位容量的 SSD 硬盘的价格大幅下降。而且，这项技术也能用来堆叠处理器内部的缓存。从前，一个 CPU 里有几 MB 的缓存就已经不得了了，而今天，我们能见到超过 100 MB 的缓存。

这其实就是靠这种垂直方向上堆积木的方式实现的。比如，AMD 在 2022 年将要推出的处理器，内部缓存就要这样堆。其中，最高端型号能堆进 768 MB 的缓存，简直骇人听闻。

而如果能把逻辑运算部分也立体地堆起来，就能成倍的增加处理器的性能。英特尔在 2018 年开发的 Foveros 3D Chiplet 就是这样的结构。而第一个能在处理器里把逻辑运算部分立体堆叠起来的企业，也正是英特尔。

Chiplet 的局限

不过说到底，这依然是冯·诺依曼结构走到末尾时的修补。**Chiplet** 技术路线并不是没有障碍，比如业界讨论得最多、困难也最大的就是热管理技术跟不上。

这个道理很好理解。如果仅仅是平铺，2 平方厘米的面积，就能产生 300 瓦的峰值功率。稍微算一下，一平方厘米就是 150 瓦。而 150 瓦/平方厘米的发热，已经超过了典型的核反应堆的功率密度了，今天的散热设施还能勉强应付。

假如再堆上 6 - 7 层。如果那时候，还是只能从 CPU 顶部铜盖这一个面散热的话，这个面的功率密度就要超过火箭发动机的喷口了。这么大的功率，现在的散热设施是搞不定的。

而现在看来，人们好像打算在芯片内部嵌入一些像毛细血管的、立体的液体冷却通道，用液冷方式把热量带出来。

总之，在真正摆脱冯·诺依曼结构的处理器普及之前，**Chiplet** 就是芯片结构发展的必由之路。这条路，可能还要走 30 年或者更久。

好，这就是今天的内容。我是卓克，我们明天再见。

划重点

1. 在芯片性能困境彻底解决之前，**Chiplet** 技术是一种短期解决方案。所谓**Chiplet**，就是一种用搭积木的方式把一堆小芯片组合成一块大芯片的技术路线。
2. 使用**Chiplet** 技术，不仅能制造以前无法完成的复杂芯片，而且能显著降低芯片的成本，不过，也存在热管理技术跟不上的问题。
3. 在真正摆脱冯·诺依曼结构的处理器普及之前，**Chiplet** 就是芯片结构发展的必由之路。





众筹新课联系微信：1716143665



642945106 “ ” “2”

0 / 5000



公开

仅限群内使用！严禁商业！

默认 最新 只看作者回复



钰箭风
9小时前

关注

芯片上每个独立的大功能的实现一般称为一个 IP 核。
IP 核对应描述功能行为的不同分为三类，即软核（Soft IP Core）、固核（Firm IP Core）和硬核（Hard IP Core）。
这些称呼都是站在芯片设计者角度去看的。

★★★

软核就是直接给了源码的功能单元，芯片设计者拿到之后，可以修改功能，最为灵活，使用难度也最大。

★★★

硬核就是芯片的物理实现也给出了，一个个硬核连接起来之后形成的单一 SoC 芯片，这种对于芯片设计者来说最简单，风险最小，但最不灵活，因为 IP 核的功能实现和制造工艺都确定了，只能照着来。
当 IP 硬核是以硅片的形式提供时，就变成了 Chiplet。就是今天卓老板讲到的 Chiplet 技术实现的芯片，这可以化繁为简，化大为小，极大提升芯片制造的...
功率。

★★★

固核就是介于软盒和硬核之间的产物，这个形式的 IP 功能确定了，源码不可见，像软件预编译好的链接文件，不能修改功能了，但物理实现未确定，芯片工艺可修改，灵活性和设计难度都介于软核和硬核之间。

↓↓↓↓↓↓

使用 chiplet 实现的芯片并不少，苹果手表是 S1 是代表，也有国产的 2021 年 11 月 3 日，寒武纪推出的自研第三代云端 AI 芯片思元 370。该芯片是寒武纪首款采用 chiplet（芯粒）技术的 AI 芯片，应该也是国内首颗 chiplet AI 芯片。

基于台积电 7nm 制程工艺，整体集成了 390 亿个晶体管，最大算力达到 25... TOPS（INT8）。

●●●

要是功耗 散热可控，大有可为啊。

展开

☐ 2

☐ 1

☐ 64

☐ 分享



左星星

14小时前

对于 Chiplet 来说，最为关键还是在于先进封装技术，使得每个 “Chiplet” 高速互联在一起，整合成一个系统级芯片。

☐ 关注

☐

虽然目前应用广泛的 SiP (System in Packaging, 系统级封装) 技术也是通过不同元件间的整合与封装，但是 Chiplet 对于封装技术的要求更高，因为每颗芯粒之间需要高密度的互联，才能实现类高速的互联，达到类似原来单个大芯片中各个功能模块间的信号传输速度。

其次，Chiplet 是将原有的系统单芯片打散成多个独立的芯粒，而要把这些芯粒通过先进封装技术整合到一起之后，还需要能够高速互联起来，而怎么去实现各个芯粒之间高速互联，则是需要解决的难题。

目前头部的芯片设计公司都在推动自己的高速互联协议标准，而且主要也是用在自家的芯片上。但是，随着 Chiplet 逐步发展，未来来自不同厂商的芯粒之间的互联需求，必然会爆发。因此，近年来也有不少行业组织、研究机构以及企业在积极推动 Chiplet 的标准。

展开

- ☐ 3
- ☐ 6
- ☐ 58
- ☐ 分享



佛祖门徒
10小时前

chiplet 的商业动机是什么？

☐ 关注 ☐

最初 chiplet 是为了打破芯片面积的物理上限，而近期是为了解决最新工艺节点的高昂 IC 设计费。例如，28nm 需要 \$51.3 million，7nm 芯片需要 \$297.8 million (International Business Strategies (IBS))。理想的情况是一个开放的数量巨大的 chiplet 市场，客户可以自由的 mix-and-match, 不同的逻辑使用不同的工艺节点，IP 可以重用，研发费用在多个设计之间平摊，创新度剧增。那么谁受益，谁有可能受损？...

展开

- ☐ 1
- ☐ 评论
- ☐ 20
- ☐ 分享



伪装
11小时前

整体一块的大芯片制造困难，成本高昂，那就把制造对象拆解开来，分成大量的小芯片，然后用它们再重新组合成一个符合要求的大芯片。这个思路比较直观，就像是我们日常解决问题不好解决，就把问题拆解成一个个小问题来解决一样。但是这个做法也不是没有缺点的，比如用系统来取代整体，系统必然是臃肿的，其运转的速度肯定是受限的。

☐ 关注 ☐

- ☐ 转发
- ☐ 评论
- ☐ 16
- ☐ 分享



耿旭52944981

17小时前

关于散热，火箭发动机也面临同样的问题，如何在高温中保持发动机部件不融化。常见的主动冷却的方案有以下几种：烧蚀冷却，再生冷却，薄膜冷却，发汗冷却。

关注



其中的发汗冷却，是将发动机部件的表面做成海绵式的细密多孔结构，让冷却液流经这些多孔结构，在进行热交换的同时，冷却液从小孔中喷出形成薄膜，将发动机中的高温气体与发动机部件表面隔离开。这与芯片内部嵌入通道的...

法不谋而合，
补充其他三种冷却方式：

烧蚀冷却，发动机钟型喷管表面涂上烧蚀材料，烧蚀材料在高温中蒸发吸热，起到冷却作用。

再生冷却，将发动机钟型喷管做成套层结构，用冷燃料作为冷却剂流过外层带走热量。

薄膜冷却，在套层结构的发动机钟型喷管上开槽，让冷燃料作为冷却剂喷出...

展开

2

评论

14

分享



AI架构师易筋

11小时前

Chiplet 是一种微型集成电路 (IC)，其中包含明确定义的功能子集。它旨在与单个封装中的插入器上的其它 chiplet 组合。一组 chiplet 可以在混合搭配的“乐高积木”组件中实现。与传统的片上系统 (SoC) 相比，这提供了几个优势：

关注



1. 可重复使用的 IP：相同的 chiplet 可用于许多不同的设备
2. 异构集成：chiplet 可以用不同的工艺、材料和节点制造，每个都针对其特定功能进行了优化
3. 提前挑选好的 die：可以在组装前测试 cheplets，提高最终器件的良率

在单个集成电路中协同工作的多个 chiplet 可以称为多芯片模块 (MCM)、混合 IC 或 2.5D IC。

-- 维基百科

展开

- ☐ 1
- ☐ 评论
- ☐ 13
- ☐ 分享



Charles

11小时前

发展接近极限换赛道就是最好的选择方向！

☐ 关注



铁路提速大概经历过几个阶段：1、煤改电改变驱动方式；2、铁轨截弯取直等；3、改变列车自重。相关因素改变完了，慢慢就开始修建客运专线，比方目前的京广高铁专线、京沪高铁专线等，在这样的新建线路上，全套都是高标准，自然速度提速也最明显，基本可以实现 380km / 小时。

铁路轨道和列车本身，摩擦减少，自重减少都是有极限的，再优化投入产出比就变差了，所以换赛道也就成了必然的选择。磁悬浮和超级高铁就是两种不同的路径，但是两者都是侧重解决阻力的问题！

我爱问卓克

- ☐ 转发
- ☐ 评论
- ☐ 13
- ☐ 分享



囚之

17小时前

Chiplet 技术简单来讲就是把整体的问题模块化，从正面硬刚改成逐个突破，再通过链接重新整合再一起，成本大大降低。

☐ 关注



只是这种方式面临的问题也不少，最基础的各芯粒之间如何实现高速互联仍是一大难题，未来的不同厂商一同制造芯片肯定会发生，它们之间能不能很好地兼容？还有随着芯片越来越复杂，测试工作将会越来越沉重，同时测试多个芯粒可比单测一个要难得多。但是在新技术出来之前，还是只能往 Chiplet 技...

- ☐ 转发

- ☐ 评论
- ☐ 11
- ☐ 分享



戚志光

7小时前

真想不到一个 **cpu** 的功率密度，已经超过核反应堆了。更没想到的是，火箭发动机的喷口，功率密度远大于核反应堆。很多东西不量化，就不能很好的理解。科学还是得从搞清楚单位做起啊。

☐ 关注



- ☐ 1
- ☐ 评论
- ☐ 10
- ☐ 分享



星安理得

11小时前

我有一个特别幼稚的问题
为什么不把 **CPU** 做大一点呢？目前的大小放大 10 倍，不也才主板那么大小嘛。

☐ 关注



可能之前有讲到过，我忘记了，这是今天早上跑步时听这一讲最大的疑惑。

- ☐ 转发
- ☐ 2
- ☐ 9
- ☐ 分享

作者 回复：
发热解决不了



WvkvW

11小时前

这一讲的知识量不小难度挺大，其实卓老板可以考虑重拾那个电磁学课，反正

☐ 关注



得到线上流量基本稳定，不用太考虑拉新。。。

- ☐ 转发
- ☐ 评论
- ☐ 6
- ☐ 分享



你当像飞往你的山
11小时前

Chiplet，AMD 的 7000 号显卡好像就是这个意思，再不够还能再回到双路四路
我现在用的 5800X，都 32M 缓存，早年系统内存好像才 32M

☐ 关注 ☐

- ☐ 转发
- ☐ 评论
- ☐ 5
- ☐ 分享



用户64267763
7小时前

感谢卓老板行云流水般的梳理，受益良多！台积电（TSMC）在 chiplet 的先进封装这块好像也是不遗余力在推进，他们从十年前的 CoWoS 到后来的 InFo 以及各种衍生版本，到近两年的 SoIC，一路狂奔的感觉，最近他们联席 CEO 还表示以后 TSMC 每两年就要把芯片互联的 pitch 缩小一半。结合最近新闻漫天飞舞的宣传 TSMC、Intel 动辄数十数百亿美元的 Fab 建厂计划，您觉得 Intel、TSMC 和三星在这个后摩尔时代生态圈里各自优劣和定位会在哪？咱...

☐ 关注 ☐

- ☐ 转发
- ☐ 评论
- ☐ 4
- ☐ 分享



十三
10小时前

☐ 关注 ☐

这么看高通 8gen1, 不完全是三星工艺的锅了
看来, 未来十年, 对于手机行业来说, 散热是一个很大的挑战呀

- ☐ 转发
- ☐ 评论
- ☐ 4
- ☐ 分享



得到用户

7小时前 编辑

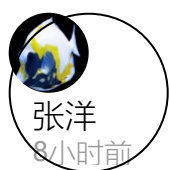
欣喜 **Chiplet** 技术的出现, 属于厉害的受益匪浅的解决问题的思路 ~ 想起薄世宁教授课程里说的“代偿”, “人体的代偿能够让器官在持续损伤的状态下, 基本上能够满足功能, 也就是凑合着用, 所以才不会出现明显的症状。... 代偿是人体不得已而为之的机制, 是妥协。... 代偿就是人类进化的另一种智慧。”前天看到丹尼尔·利伯曼在《人体的故事: 进化、健康与疾病》里强调我们的身体充满了妥协, 不由得迁连到人类社会其实也像生物群落, 符合演化的而不是...

☐ 关注



展开

- ☐ 转发
- ☐ 评论
- ☐ 3
- ☐ 分享



张洋

8小时前

Chiplet 技术 (小芯片) 对于没有什么冗余的大芯片来说是最具成本效益的, 例如在 5nm 工艺中, 晶圆成本高达 17000 美元, 而如果采用小芯片就可以节省相当的成本 (预计 20-30%), 而且随着尺寸的缩小 (3nm) 还可以节省更多 (针对于基础价格而言)。

☐ 关注



但是要注意, 小芯片并不适合所有设计。例如 **PC** 和智能手机处理器的尺寸通常为 150mm² 或更小, 使用此项技术, 收益就会非常有限。

当然, 小芯片的优势不仅在于节省成本, 还包括可以构建比任何单片芯片都更大的设计、通过解耦新技术降低进度风险以及提供更灵活的产品配置。

展开

- ☐ 转发
- ☐ 评论
- ☐ 3
- ☐ 分享



Aming

11小时前

复杂芯片面临最大挑战就是散热。最早 AMD 台式芯片就是散热不佳，所以不太受待见。手机由于很难加载散热风扇，骁龙处理器在发展到后期，高端机往往变成暖手宝，甚至不时采用降频手段减少热量。相比来说，苹果的 M1 型片和 A 系列芯片表现可谓优秀。芯片速度提升是一方面，散热也是不容忽视的特性。不能顾此失彼。

☐ 关注



- ☐ 转发
- ☐ 评论
- ☐ 3
- ☐ 分享



牧云昶

10小时前

懂了，这一波利好乐高！

☐ 关注



- ☐ 转发
- ☐ 评论
- ☐ 1
- ☐ 分享



潘飏

12小时前

早期计算机的部件，CPU 的功能比较单一，主要就是负责算术和逻辑运算，浮点运算有专门的浮点运算部件，内存访问也有专门的芯片，后来芯片集成度高了，开始把这些都放到一个 CPU 里了。现在 chiplet 又开始把 CPU 的功能往外拆了。技术也在不断轮回，或者叫“螺旋式发展”

☐ 关注



转发



评论

1

分享

加微信：642945106 发送“赠送”领取赠送精品课程 发数字“2”获取众筹列表

