基于解耦 De-skew PLL 的处理器低功耗 同步间歇时钟系统设计

杨丽琼",2 吴瑞阳"杨梁"王焕东"

¹⁾(中国科学院计算技术研究所计算机体系结构国家重点实验室 北京 100190)
 ²⁾(中国科学院大学计算机科学与技术学院 北京 100049)
 ³⁾(龙芯中科技术股份有限公司 北京 100095)

随着高性能处理器集成度、面积以及工作频率的不断增加,时钟动态功耗呈指数级增加,时钟分布不均导 摘 要 致跨时钟域的同步开销显著增大,这些问题逐渐成为制约处理器能效提升的瓶颈.通常处理器核的功耗占多核处 理器整体功耗超过 70%,而时钟功耗是处理器核功耗的主要组成部分.数字方式的系统动态调频 DFS(Dynamic Frequency Scaling)降频的方法需要触发时钟中断例外重新配置时钟生成模块锁相环的相关寄存器,由此带来系统 超过毫秒级等待时间开销;而模拟方式连续自适应调节 AFS(Adaptive Frequency Scaling)频率变化过程中存在频 率过冲响应会增加物理时序设计压力.与此同时功耗的调节降低要以高性能为前提.片上时钟分布长延时随 PVT (Process Voltage Temperature)变化产生的不确定时钟相位偏差,为此物理设计增加时序冗余补偿会直接影响到 处理器性能.本文提出了新的基于解耦去偏斜锁相环 De-skew PLL (De-skew Phase Locked Loop)的同步间歇时 钟系统,采用12nm CMOS工艺实现了去偏斜锁相环的设计,并对整个系统进行了时序性能和时钟功耗的评估.该 系统一方面可以利用去偏斜锁相环的远端时钟反馈技术实现不同时钟域之间的实时相位对齐,同时也可以抵抗反 馈环内时钟分布延时随 PVT 的变化;另一方面可以利用新增加的解耦模块,无频率过冲地响应处理器核内产生的 时钟间歇控制(时钟脉冲间断性停拍)信号降频,从而实现亚纳秒级时钟动态功耗控制.以12nm工艺同步级联结构 为例,每层时钟分布校准后同步偏差小于 10 ps. 使用 16 核 LS3C5000 处理器 RTL 在仿真加速平台上运行 SPEC CPU 2000 测试集来评估本方案对处理器核时钟功耗的影响,并进一步通过 PTPX 后仿真验证,结果表明,定点及 浮点程序平均功耗节约分别大于 4.5% 和 20.3%.

关键词 多核处理器;同步间歇时钟系统;解耦去偏斜锁相环;低功耗设计中图法分类号 TP301 **DOI**号 10.11897/SP.J.1016.2022.02207

Synchronized Intermittent Clock System Based on Decouple De-skew Phase Locked Loop Used in Low Power Processor Design

YANG Li-Qiong^{1),2)} WU Rui-Yang³⁾ YANG Liang³⁾ WANG Huan-Dong³⁾

¹⁾ (State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190) ²⁾ (School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049)

³⁾ (Loongson Technology Corporation Limited, Beijing 100095)

Abstract With the increasing of processor's integration, area and working frequency, clock power consumption is increasing exponentially, and the cost of synchronization across different clock domains becomes serious due to distribution's non-uniform. Both of these issues have already become the bottleneck that restricts the energy efficiency of the processor. Normally, the processor core's power consumption accounted for more than 70% of the total power of the multi-

收稿日期:2020-12-06;在线发布日期:2021-08-19.本课题得到中国科学院战略性先导科技专项(C类)课题(XDC05020100)项目资助.杨 丽琼,博士研究生,副高级工程师,主要研究方向为计算机体系结构、高速互连多核处理器时钟系统及片上传感器. E-mail: yangliqiong@ ict. ac. cn. 吴瑞阳,博士,主要研究方向为处理器核微架构设计. 杨 梁,博士,主要研究方向为高性能处理器物理设计. 王焕东,博士,主要研究方向为多核处理器架构设计.

core processor, and clock power is the main component of the process core's power consumption. DFS (Dynamic Frequency Scaling) requires to trigger clock interrupt exceptions and reconfigure PLL (Phase Locked Loop)'s relevant registers, but state shift results in milliseconds level system waiting time. AFS (Adaptive Frequency Scaling) without system control continually adjusts operating frequency by tracking power supply level's change. While frequency overshoot could not be avoided during the tuning process, which brings extra physical timing constraints. Clock system's low power design could not be at the expense of processors' performance. Clock distribution's delay deviates with PVT (Process Voltage Temperature)'s variation. Increasing timing margin to compensate for clock phase differences will directly affect the timing performance of critical paths. In this paper, a new synchronized intermittent clock system based on decoupled De-skew PLL is proposed firstly; subsequently a De-skew PLL which supporting stable phase error calibration is realized in 12 nm CMOS process; finally, the timing performance and clock power consumption are evaluated for the whole system. On the one hand, this new clock system structure not only can realize real-time phase alignment between different clock domains by Deskew PLL's remote feedback, but also can immune clock tree delay's PVT variation by real-time in loop tracking; on the other hand, De-skew PLL's decoupling module can decouple the relationship between clock tree frequency and PLL's loop configuration without loss lock, which can be used to realize the dynamic power consumption control in sub 1ns response time without frequency over shoot. Taking 12 nm cascade clock system as an example, each stage clock distribution synchronization deviation can be achieved less than 10 ps after software phase calibration. LS3C5000 16-core processor's RTL is adopted to run the SPEC CPU 2000 test set on the simulation acceleration platform to evaluate the effect of this new structure on the processor core's clock power consumption, and further through PTPX post-simulation verification, the results showed that the average power saving of fixed-point program and the floating-point program is greater than 4.5% and 20.3%, respectively.

Keywords multi-core processor; synchronized intermittent frequency system; decoupled De-skew phase locked loop; low power design

1 引 言

21世纪之初,微处理器的发展一直遵循着摩尔 定律^[1]法则,集成度和主频性能随着晶体管尺寸的 缩小而不断增加.目前硅晶体管的制程已进入纳米 时代,处理器芯片性能和集成度不断提高的背后,功 耗引发的问题越来越严重.处理器主频的提升带来 芯片的动态耗电指数性增加,功耗密度过大引发温 度持续升高.片内逻辑单元的漏电和供电网络的寄 生电阻随温升进一步加大,这样导致了更大的功耗 开销及不理想的供电压降,制约了处理器性能的进 一步提高.与此同时,芯片的功耗及温度对封装及散 热提出了更高要求.这些问题使得处理器设计在不 断提升处理器核主频性能的同时格外关注高性能下 的低功耗设计. 近些年处理器架构设计在能效设计上进行了不断的改进和创新. Sandy bridge^[2]的 TurboBoost 技术实现了高性能与低功耗之间的分时分类匹配管理; Haswell^[3]的 FIVR (Fully Integrated Voltage Regulator)实现了多相多级的区域性供电灵活控制; Skylake^[4]的 ISST (Intel Speed Shift Technology)的控制层面发生了根本性的变化,从顶层操作系统深入到底层处理器核模块内进行动态功耗调整,相比上一代获得 14%的能效提升; Cascade Lake^[5]和 Tiger Lake^[6]进一步细化动态功耗调节的效率控制. 功耗和性能的协同管理逐步深入到处理器核的微架构设计及物理电路设计中^[7:8],在采集物理层的相关信息上传到顶层系统控制的同时增加了本地反馈,从而可进行更多细致的快速调节,使得系统和物理的跨层次的融合设计越来越紧密.

处理器核功耗是处理器功耗的主要组成部

分^[9-10],其中时钟动态功耗占处理器核功耗超过 30%^[10].时钟系统的低功耗设计是处理器低功耗设 计的关键部分.流水线时钟的翻转、逻辑的求值及存 储的预充读写是动态功耗的主要来源,这些均与时 钟系统的配置控制息息相关.与此同时,处理器性能 的重要标志之一就是流水线的效率:一方面体现在 流水线的速度(时钟频率);另一方面则是流水线的 执行效率(非冗余空拍占比).转移猜测的失误引发 的流水重排、存储层次的有限带宽导致的读取指令 的等待写回以及总线异步通讯带来的时序冗余等 都会影响执行效率.在保证处理器效率的条件下降 低处理器功耗优化处理器能效是本文研究的主要 问题.

本文提出了一种基于解耦 De-skew PLL 的同 步时钟系统,在取得同步时钟系统性能提升的同时 通过加入本地反馈实时监测处理器核工作状态,实 现亚纳秒级的动态时钟间歇变频响应调节.文中论 述了该时钟系统的结构及重要模块设计,通过电路 级仿真验证以及在多核系统中运行真实多任务负载 来评估整体功耗节约控制的效果.以12 nm 数字工 艺下级联时钟结构为例,每层时钟通过解耦控制实 现时钟间歇可当拍快速响应,结构上保证了无频率 过冲风险;时钟分布在校准后可获得 10 ps 以内的稳 态相差.基于此时钟系统的处理器核动态功耗的评 估实验采用了 LS3C5000 16 核处理器 RTL 在仿真 加速平台运行 SPEC CPU 2000 的方法并辅以 PT-PX 后仿真验证,结果表明定点和浮点程序平均功 耗分别节约超过 4.5%和 20.3%.

2 相关工作

系统能耗控制通常采用动态调节各个处理器核 的工作频率进行主动降频来实现.系统通过检测各 处理器核程序执行的负载状态是否超过目标负载值 来动态调节各个处理器核的工作频率.为了满足独 立可调,划分时钟域并独立配置锁相环的倍频系数 得到目标时钟频率^[11-13]必不可少.随之需要解决的 问题就是各时钟域之间的时钟同步.这里所说的同 步主要针对的是相位的同步,即物理实现总线通讯 时需要满足的时钟相位关系一致(时钟相位对齐或 存在一定量的稳态相差).相位的不确定会带来跨时 钟域的异步冗余开销.由于集成在同一个芯片中各 功能模块的规模差异明显,模块内部的时钟负载和 分布延时参差不齐,带来时钟相位的同源不同步.龙

芯 3B^[10] 采用了被动去相差的方法. 根据模块之间 相位偏差检测结果,调节预先在各个模块内增加的 冗余数控延时链的延时长度,实现不同时钟树支端 时钟相位的对齐.与此同时在顶层时钟分布节点采 用门控方法对各个时钟域实现时钟功耗的独立控 制. 被动相差校准方法存在如下弊端:一方面数控延 时链的控制码确定后不易更改,这样该方法仅可实 现各时钟域的固态相差补偿,温度和电压变化导致 延时动态变化则无法实时调整,只能靠增加时序裕 量或是降低整体工作频率来适应;另一方面在于用 来补偿的冗余时钟链会引入新的功耗开销和噪声. Nehalem 采用去偏斜锁相环结构^[14-15]实现了时钟的 动态相差跟随.这种结构跟随数字电压变化可以实 现不同规模时钟树之间的相位对齐,同时可以自动 连续变频响应系统变频需求.具体跟随变化的幅度 与片内实际物理实现频率之间的对应关系需要严格 的磨合.另外,在无需变频时存在引入所跟随的数字 电源噪声的风险,对片内电源的设计提出更高的要 求.ARMv8^[16]采用更为数字化的降频方法直接降 到一半,并不支持连续降频.频率变化过程存在2% 的频率过冲,需要增加物理时序设计冗余.降频参数 的重新配置,需要系统为此触发时钟中断,刷新锁相 环配置寄存器,锁相环电路内部也需要时间重新锁 定频率,由此带来较长的系统调节反馈时间.Sandy bridge^[2]采用的 Turbo Boost 也面临系统调节时间 过长的问题,状态进出变化时需要系统调度时间超 过 10 ms. Skylake^[4]在 TurboBoost 的基础上采用了 新的能效管理,系统调度只限于基础的状态转换.针 对 P 态进行了更细致的划分,进出 P 态的调度增加了 处理器核内部指令响应,检测响应时间从10ms 缩短 到1ms,对于小于1ms的主动降频空间不可见.系 统检测响应时间限制了能效管理的最小颗粒度.

降低时钟功耗不仅可以通过改变时钟脉冲的宽 度,也可以通过改变时钟脉冲的个数.被反馈控制的 时钟脉冲变得不再连续,我们称之为间歇时钟.间歇 时钟相比同脉宽的普通时钟在一段时间内具有更少 的时钟翻转次数,可以降低这段时间内的平均时钟 动态功耗.与此同时,由于时钟脉冲宽度没有变化, 系统不需要重新配置锁相环倍频数寄存器,锁相环 也不需要重新锁定,系统不需要保存状态等待唤醒, 控制反馈时间明显减少,功耗管理的最小颗粒度降 到亚纳秒级.基于此理论,本文提出了新的基于解耦 去偏斜锁相环的间歇时钟系统结构,并采用 12 nm 工艺实现了锁相环的设计.这种结构一方面可以实 现不同时钟域之间的相位实时跟随补偿,同时可以 利用新增加的解耦功能当拍实现时钟间歇控制,支 持亚纳秒级的系统降频功能控制.

3 同步设计与变频能效设计之间的 矛盾

逻辑单元延时的静态时序计算和平衡同步的时 钟树结构设计是处理器流水线物理时序同步设计的 基础.基于多工艺角标准单元库综合时序的路径计 算保证了数据路径静态时序的准确性,并根据流片 经验可增加一定的冗余满足路径时序的动态变化. 而时钟分布的同步问题通常需要增加数字设计流程 之外的定制设计和系统控制来协助解决.这些额外 的设计及控制不免会带来功耗的增加,引发同步设 计与能效设计之间的矛盾.

物理布局层面,各高速时钟路径上通常增加定 制冗余的延时单元进行同步时钟相位补偿.时钟延 时的基本单元是反相器,反相器的动态功耗 Pare 计 算公式如下所示:

 $P_{dyn} = N_{inv} \times C_{load} \times V_{DD} \times f_{0>1}$ (1) 其种 N_{inv} 代表反相器个数, C_{load} 代表平均负载电容 V_{DD} 代表供电电压, $f_{0>1}$ 代表正沿翻转率正比于工 作时钟频率.单级传输距离越远负载电容 C_{load} 越 大,时钟树面积越大驱动时钟反相器数目 N_{inv} 越大, 工作频率越高 $f_{0>1}$ 越大.所以 P_{dyn} 正比于时钟网络 覆盖面积、供电电压及工作频率.对全速时钟信号采 用延时的方法进行时钟相位补偿会直接增加时钟功 耗.不仅如此,冗余的时钟延时还会引入更多的噪声 增加时钟抖动.

系统控制层面,随着芯片规模的增大,同步时钟 系统结构通常划分多个时钟域并配有独立 PLL.当 某处理器核进入低功耗状态需要降低供电电压及频 率时,需要触发时钟中断重新配置用于产生该时钟 域的锁相环;当处理器核重返高时钟频率工作状态 时,同样需要历经系统调控.这种系统的反馈控制需 要处理器耗费超过毫秒量级的等待时间,造成流水 线的空置,影响系统性能.

通过以上分析可以看到同步时钟设计不仅增加 时钟功耗,降低时钟性能,且无法避免由于变频调节 带来的较长的系统等待时间,导致流水线空置.本文 提出了基于解耦去偏斜锁相环的同步时钟结构可支 持本地时钟间歇动态调频,为快速升降频的能效管 理提供了基础.

4 处理器时钟系统的同步低功耗设计

时钟系统的同步低功耗设计主要有以下几个 方面:

首先是同步设计,时钟系统的同步设计基于物 理时序的计算,关注的是时钟的相位关系.多核处理 器中集成的各功能模块的频率通常不完全相同,可 分为同源同频跨域、同源倍频跨域以及不同源跨域. 不同源跨域通常需要跨域异步开销,具体依据系统 协议的规定.这里研究的同步问题主要针对同源跨 域的时序优化.

其次是低功耗设计,这里的低功耗设计主要针 对处理器核功耗中占比较大的动态时钟功耗部分. 动态功耗与时钟频率紧密相关,是时钟系统设计的 重点.静态功耗有电源控制的方法可以实现.如何更 精确的控制功耗与时钟的关系,降低时钟降频控制 过程开销是本文研究的重点.

最后是关键定制时钟电路锁相环设计.锁相环 作为时钟源模块负责产生稳定可配置的时钟.环路 稳态相差成为锁相环去相位偏差的性能瓶颈,如何 校准补偿稳态相差是本文设计的重点.另外,级联时 钟结构中后级锁相环面临高频输入高带宽环路跟随 的要求,输入不分频的情况下,输入时钟频率越高环 路运行越快,允许环路单次计算时间越短.远端反馈 增加了环路计算时间.环路响应速度是否足够快可 以在环路高速运行下正常锁定并输出稳定时钟成为 设计的关键.

4.1 相位同步设计

多核处理器无论是采用统一的主频时钟还是各 个处理器核独立时钟控制都涉及到相位同步的问 题.相位偏差过大会带来总线通讯时序的冗余开销, 传输效率随之降低.以单节点4核共享缓存系统为 例(由此扩展为8核、16核)见图1,时钟系统的同步 结构分全局和区域两个主要层次.

全局层次的时钟相位同步是多个模块间入口时 钟同步.由于各模块位置分散,节点时钟 mid_ck 先 通过主干线大驱动反相器到达各方向中部再进行时 钟树分布到各处理器核 Core、共享缓存 SC(Shared-Cache)对齐.这部分属于物理实现的顶层时钟网络 部分,物理设计需要对路径的平衡和时钟的快速传 递进行重点设计.顶层时钟跨越布局面积大,因此除 了采用远端反馈技术(时钟树路径包含在环路内)进 行相位对齐外,采用大驱动反相器进行长距离的驱动



图 1 基于解耦锁相环的级联同步时钟结构

可以明显减小时钟延时并易于实现网络负载平衡. 区域层次的时钟相位同步是在模块入口节点时 钟已全局同步的基础上各模块之间的子节点时钟同 步.比如处理器核 Core 内部触发器时钟和 SC 内部 触发器之间的时钟相位同步(处理器和高速缓存同 频). 处理器核与共享缓存之间的通讯路径是整个芯 片访存最繁忙的路径, LS3C5000 处理器核和共享 缓存物理面积相差超过 40%,内部的时钟树延时差 别明显,本文提出的基于解耦锁相环的同步结构利 用解耦锁相环远端反馈环路(模块的时钟树路径包 含在环路内)实时反馈补偿时钟树路径的延时变化, 免疫了时钟树长度不一致的问题,从而实现了相同 参考输入的不同模块内的叶子节点的相位同步.如 图1所示,各同步模块都具有自己的解耦去相位偏 差时钟环路. DFC(Decouple Frequency Control)为 解耦模块,每个分支通过解耦锁相环 生成时钟后送 入模块时钟树入口,经过时钟树分布后选择离入口 最近的叶子节点时钟作为反馈端送回锁相环的输入 进行相位比较调整,从而实现时钟树子节点与输入 参考时钟 mid_ck 对齐.

这部分属于物理实现的区域时钟网络层次.模 块负载重且规模大小参差不齐,物理设计将重点放 在时钟扇出的均匀和延时的一致.区域时钟网络主要 时钟结构为平衡树、时钟 SPINE 与时钟 MESH^[10] 结构,除此之外本文利用了去偏斜锁相环的反馈结 构将区域时钟分布延时缩短到一个锁相环的稳态相 差,明显降低了区域时钟分布延时.

通过上面两个层次的对齐所有高频模块的叶子 节点时钟通过反馈级联结构实现相位的基本对齐. 这里忽略了区域网格分布之下的本地时钟分布.由 于区域时钟分布采用网格结构进行相位偏差的消 除,时钟相位可以保持较好的对齐效果,所以对经过 聚类的本地时钟分布未增加额外的对齐调节.不同 频时钟域之间的同步(如 SC 与 DDR 控制器之间) 采用异步 FIFO 方式处理.以上所示基于解耦锁相 环级联的同步时钟结构对同源不同时钟域的时钟树 延时不敏感,明显降低了不同时钟域跨域异步冗余 开销,提高了处理器时序性能.

以LS3C5000处理器的模块设计为基础计算上 述时钟系统的相位误差.处理器核与SC之间通过 交叉开关XBAR1互连,SC与DDR之间通过 XBAR2互连,再通过RING与HT互连.时钟分布 如图2所示,main_ck在底部产生后沿主干线从底 部传到中心.位于中心的锁相环以main_ck为参考 产生mid_ck,mid_ck沿主干线高频时钟路径传到 XBAR四边中间再展开时钟树,选择中心PLL附近 的子节点作为 mid_ck 的反馈节点,各模块入口参考时钟 mid_ck 与 main_ck 对齐.模块内时钟树叶子节点再与 mid_ck 对齐.高频时钟路包含在两级解耦锁相环的环路内,通过环路进行实时跟随对齐,每级输出与输入相差一个锁相环的稳态相差. mid_ck 时钟树分布没有整体的 MESH 结构导致各模块入口时钟之间存额外 5 ps 分布偏差.



图 2 同步时钟物理分布示意图

以4核为基础,16核可以扩展为4个 mid_ck. 采用环形互连结构将 XBAR 连接起来,同样是所有 的 mid_ck 与 main_ck 对齐.时钟结构级数不变, main_ck 的分布延时偏差类似 mid_ck.

4.2 快速降频低功耗设计

低功耗控制可以分为顶层系统控制调节和本地 处理器核反馈信号控制调节.远端系统控制调节反 馈时间超过毫秒,通过调节锁相环的倍频系数降低 输出时钟频率.本地处理器核反馈信号调节采用本 文提出的时钟结构反馈时间小于纳秒,通过控制 DFC 解耦模块实现在控制信号有效范围内的处理 器核时钟间歇,从而降低输出时钟的平均频率.

系统调节方法反馈时间较长,主要来源于系统 变换时钟频率需要触发时钟中断并重新配置寄存 器.频率调节锁相环重新锁定过程中会出现超过目 标频率的高频时钟,导致物理时序出错系统宕机,因 此需等待锁相环重新锁定.毫秒级的反馈时间内,该 时钟域处于冗余的停流水状态,一定程度影响系统 性能.同理从低频状态恢复正常频率或超频同样有 如上的开销.因此,考虑到系统调节状态切换开销, 系统动态降频的应用中如果没有取得足够的功耗优势,整体能效或许并没有得到提升.Turbo boost 应用^[2]也会存在这样的问题.

相比系统调节方法,本地反馈调节的间歇时钟 方法从结构上避免了状态转换开销.因为时钟间歇 并不需要系统重新配置锁相环,所以既没有系统调 节的等待时间也无需等待锁相环重新入锁.整个过 程没有冗余的停流水时间.时钟间歇方法可以选择 不同的间歇方式.紧凑的间歇方式可对间歇拍数进 行阈值控制,在阈值控制范围内(例如几十拍时钟) 间歇控制信号保持有效为"1"则时钟一直间歇,超出 阈值拍数后短暂恢复时钟脉冲供应,然后再重新计 数并继续执行新的间歇.松弛的间歇方式可长时间 统一按照一定配比间歇(例如每5拍间歇1拍),只 要间歇控制信号为"1"则一直执行每5拍间歇一拍 的时钟间歇控制.

时钟间歇结构无法直接应用在已有的具有同步 功能的去偏斜锁相环上^[12].如图 3 所示,没有解耦 控制的锁相环与时钟树之间可选择增加时钟门控 CG(Clock Gating)单元来中断时钟供应. 间歇控制 信号 innerstop_ctl 为"1"时正常供应的参考时钟与 无沿时钟进行相位比较生成 up_nodfc/dn_nodfc 信号(代表参考时钟相位超前/落后),连续为高的 up_nodfc 信号产生正向调节的相差积累,不断增加 的相差 Δø_nodfc 导致输出频率一直升高直到环路 失锁.这段时间输出时钟不稳定,系统无法正常使 用. 解耦模块设计目标在于处理器核反馈信号在控 制时钟间歇过程中保持锁相环锁定,随时可提供稳定 可靠的间歇时钟.具体电路如图 3 所示,DFC 需要成 对串联出现在解耦锁相环的输入和输出通路上. DFC采用锁相环的锁定信号 lock 作为使能信号,采 用 innerstop_ctl 信号作为控制间歇信号. 当解耦锁 相环环路锁定后 lock 信号置高, DFC 进入可间歇状 态,否则输出时钟直通不受间歇信号控制.在 lock 信号置高条件下,当间歇控制信号 innerstop_ctl 由 '0'变为'1'时,间歇有效,输出时钟 frefp 和 fout 的时钟脉冲间歇.参考时钟和反馈时钟均没有时钟 参与相位比较所以 up/dn 信号为'0'代表没有超前 或者落后,相位差 $\Delta \phi$ 也就无变化.振荡器保持在原 来控制条件下稳定输出时钟 foutp,无频率变化;当 *innerstop_ctl* 从间歇控制有效状态'1'变为'0'时, 时钟恢复正常.解耦锁相环由解耦模块配合锁相环 工作的关键信号波形图如图 3 所示.同步间歇过程 中时钟相位仍保持对齐.



图 3 解耦模块电路结构、时序波形图及非解耦失锁状态关键信号波形

4.3 De-skew PLL 设计与实现

4.3.1 时钟环路设计的基本原理

锁相环的主要工作原理是:首先采用鉴相器 PFD(Phase Frequency Detector)对输入参考时钟 *clkref_in*和时钟振荡器输出时钟分频后的反馈时 钟*clkfb_in*进行相位比较产生相差Δø如图 4 所 示,再通过相位增益控制和滤波控制产生与相差线 性相关的控制电压 V_{ett},最后 V_{ett}控制振荡器产生 新的输出时钟,完成一次计算反馈.其中 *clkout* 与 *clkfb* 具有相同的频率只是分布位置不同.环路通 过若干拍调整最终达到锁定.输出时钟的频率满足 倍频公式(2),同时反馈时钟和参考时钟频率相同且 相位基本一致.相位基本一致是因为整个环路在锁 定后存在一个固定的稳态相差.本文设计的锁相 环通过相位校准机制补偿稳态相差的影响使输入参 考时钟和输出时钟之间的相差小于最小延时调节单 位 10 ps.

$$clkout = clkfb = \frac{vco_clkp}{2 \times M} = \frac{clkref \times P}{N} \quad (2)$$

去偏斜锁相环整体结构如图 4 所示,其中通过 配置分频系数 N(输入分频数)、M(输出分频数)和 P(环路分频数)来实现频率转换.振荡器的工作频 率 vco_clkp 限制了环路输出时钟的最高时钟频率. 图 4 中虚线箭头回路为环路本地反馈模式,去偏斜 锁相环打开了本地回路历经右半部时钟树后从时钟 树叶子节点返回到鉴相器进行相位比较.图4的左 半边是一个可选择开环的去偏斜锁相环,右半边是 物理数字综合生成时钟树分布.通常处理器核内时 钟树分布在区域模块内采用 MESH 结构,这种结构 有利于实现众多子节点之间的相位对齐.在设计中 选择离输出时钟 clkout 位置近的子节点 clk fb 作为 反馈点,这样反馈回路不需要额外增加路径延时. clk fb 送到环路分频器进行分频后再反馈到鉴相器 入口,实现环路闭合.通过环路多拍调整实现时钟树 的叶子节点时钟 clk fb 与输入时钟 clkref 之间的相 位对齐.本文设计的锁相环采用了自偏置双环路结 构[17-18] 如图 5 虑线框所示. 自偏置结构加上双环路 滤波可以很好的适应数字低电压供电条件并可适应 高输入参考频率产生高带宽满足相位跟随需求.为 实现时钟系统的相位同步,稳态相差的相位校准补 偿以及频率转换过程中分频器相位控制成为相位同 步控制的关键.



图 4 去偏斜锁相环结构原理图





4.3.2 稳态相差的相位补偿设计

相位补偿通过对 clkref_in 和 clkfb_in 进行相位比较得到稳态相差的偏差方向并进行软件补偿控制,从而逐步调整通路延时,缩小 clkref 与 clkbk 相差,达到相位同步.

*clkfb*与*clkref*是*clkfb_in*与*clkref_in*的相位外延.如图 5 所示 *clkfb*通过 Delay-Line-B(DLB)成为 *clkfb_in*的外延,*clkref*通过 Delay-Line-R(DLR)成为 *clkref_in*的外延.在不调整两延时链时,*clkfb*与*clkref_in*的外延.在不调整两延时链时,*clkfb*与*clkref_in*与*clkref_in*之间的相差 Δ*α*。等于 *clkfb_in*与*clkref_in*之间的相差 Δ*α*。每是去偏斜锁相环的固有相差,由于芯片生产的偏差,不同芯片内 PLL的固有相差存在差异,在 Δ*ø*基础上 Δ*α*可通过延时链调整而减小至最小调整步进,达到近似相位同步.

下面从延时链调节、补偿算法及控制逻辑三方 面重点展开论述.

首先可调延时链的精确程度直接关系到相位调

节的精确度. 在环路的输入路径和反馈路径分别串 行连接了用于延时调节的 DLR 和 DLB,两条路径 的延时链具有相同的延时单元电路结构及版图结 构. 延时链的基本延时单元为自负载反相器(延时小 于 5ps),且均为本地连接,驱动能力大于负载,最差 工艺角下可调延时步进为小于 10ps.

通路延时链 DLR 和 DLB 采用二进制控制码进 行对称延时控制的功能如表1所示.控制码最高位

表 1 通路延时控制逻辑时间功能表

checkcode	checkcode(3) = 1		checkcode(3) = 0				
<2:0>	输入 DLR	输出 DLB	输入 DLR	输出 DLB			
000	直通	直通	直通	直通			
001	$10\mathrm{ps}$	0	0	$10\mathrm{ps}$			
010	$20\mathrm{ps}$	0	0	$20\mathrm{ps}$			
011	$30\mathrm{ps}$	0	0	$30\mathrm{ps}$			
100	$40\mathrm{ps}$	0	0	$40\mathrm{ps}$			
101	$50\mathrm{ps}$	0	0	$50\mathrm{ps}$			
110	$60\mathrm{ps}$	0	0	$60\mathrm{ps}$			
111	$70\mathrm{ps}$	0	0	$70\mathrm{ps}$			

作为通道区分位,具体延时电路及版图均相同,保证 两通路具有相同的调节步进和范围.

其次,需要采用数字相位补偿算法来补偿工艺 生产带来的延时偏差.去偏斜锁相环中有两个相位 检测的功能模块,环路内相位检测(PFD)和环路外 相位补偿检测(Adjust Phase Check)如图 5 所示. PFD 用以构成环路内反馈,当环路稳定时 $clkref_in$ 和 $clkfb_in$ 之间存在稳态相差 $\Delta ø$. APC 包含相同 精度的鉴相器、DLR、DLB 以及控制逻辑. $clkref_in$ 和 $clkfb_in$ 分别经过 APC 中的 DLR 和 DLB 到达 鉴相器入口进行相位比较. APC 检测到 $\Delta ø$ 并配合 软件控制产生与 $\Delta ø$ 等量的补偿延时码字. 再将码 字高位取反应用到环内的输入输出链路,进行稳态 相差的实时补偿,实现输入输出的相位对齐. 输入时 钟 clkref 与 $clkref_in$ 之间输入路径相位延时计算 如式(3);反馈时钟 clkfb 与 $clkfb_in$ 之间反馈路径 相位延时计算如式(4).

$$P_{ref} = P_{ref_in} - D_{and} - D_{dlr} - D_{div}$$
(3)

 $P_{bk} = P_{bk_{in}} - D_{mux} - D_{div} - D_{dlb}$ (4) 其中, P_{ref} 、 P_{bk} 、 $P_{ref_{in}}$ 和 $P_{bk_{in}}$ 分别代表 claref、 clk fb、 $clkref_{in}$ 和 $clk fb_{in}$ 的相位, D_{dlr} 和 D_{dlb} 分 别代表输入延时链 DLR 和反馈延时链 DLB 的延 时, D_{div} 、 D_{and} 和 D_{mux} 分别代表分频器、与非门和选 择器的延时. 链路如图 5 所示,选择等延时的与非 门和选择器时 D_{and} 与 D_{mux} 取得相等. 初始状态两延 时链均为直通即 D_{dlr} 与 D_{dlb} 相等, 环路锁定后 P_{ref} 和 P_{bk} 之间的相位差等于 $P_{ref_{in}}$ 和 $P_{bk_{in}}$ 之间存在正 向稳态相差 $\Delta \phi$ ($\Delta \phi$ 代表 $clk fb_{in}$ 落后 $clkref_{in}$ 的 延时量). 为了实现 P_{ref} 和 P_{bk} 的相位对齐, 需要通过 调节 D_{dlr} 或 D_{dlb} 来抵消掉稳态相差, 实现 $P_{ref} = P_{bk}$.

环外相位补偿逼近过程如图 6 所示.鉴相信 号 early和 late分别代表 clkfb_in 提前和落后于 clkref_in.环路锁定后开始补偿检测. late:early 的初始复位值为"00".当相位检测出"10"时表明 clkfb_in 落后于 clkref_in,对环外需增加 DLR 使 环外 clkref_in 支路相位落后,经过逐步增加直到 检测出"11"后再延续增加控制码,如果仍然保持 "11"则补偿结束.由于延时链延时步进小于鉴相器 阈值相差,所以在检测出"11"后进一步调节延时控 制码使增加的延时与输入相差接近最小延时步进. 当检测出"01",表明 clkfb_in 提前于 clkref_in,对 环外需要增加 DLB 的延时使 clkfb_in 支路相位落 后.同样经过逐步调节直到连续检测出"11".当检测 出"11"软件控制同样需要增加延时码来逼近最小相



图 6 环外相位补偿软件控制

位误差.

4.3.3 频率转换中的相位控制

去偏斜锁相环通过配置输入、输出和环路分频器的分频数实现频率转换.频率转换过程易出现多相位输出的问题,比如 400 MHz 时钟信号经过 4 分频可随机产生 0 度、90 度、180 度和 270 度 4 种初始相位的 100 MHz 时钟,无法做到输入输出相位对齐.如图 5 所示,PFD 输入端的 clkref_in 和 clkfb_ in 在环路负反馈的作用下可以取得时钟相位对齐, clkref 与 clkref_in 之间以及 clkfb 与 clkfb_in 之 间都存在用于频率转换的分频器.暂不考虑分频器的通路延时问题,采用多相位输出分频器无法 保证在 clkref_in 与 clkfb_in 相位对齐的条件下 clkref、clkfb 分别与 clkref_in、clkfb_in 相位一致, 也就无法保证 clkref 与 clkfb 之间相位对齐.

针对多相位输出问题,本文设计了图 7 所示的 相位同步分频结构.通过初始相位的固定和多通路



相位延时的匹配实现分频器输入输出时钟的相位一 致.如图7所示d1、d2、d4和d8为输入信号的1、2、4、 8分频,分频触发器均进行了复位控制,复位信号在 开始工作的前半个周期取消复位控制进入正常可翻 转状态.另外,分频器在不同分频通路选择下输出相 对输入的相位延时均需要一致.采用的方法是对2、4、 8分频支路进行了高频重采样,保证这三条输出通路 的延时都等于一个触发器的采样延时,而1分频通 过延时单元进行补偿.最后4个分频后时钟信号经过 相同级数的逻辑选通生成输出时钟,从结构上保证了 输出时钟相位的固定和各通路相位延时的一致.

4.3.4 锁相环电路仿真

锁相环的环路仿真采用输入 1 分频条件下最高 输入 800 MHz 参考时钟频率,版图寄生后提取网表 以及最慢工艺角库进行稳态相差的瞬态仿真验证. 频率配置数 M 和 P 均为 2, N 为 1. 如图 8 所示,在 启动控制 pd 信号变为"0"后环路启动,经过约 0. 3 μ s 后 lock 信号和 phase_stable 信号先后置高, 标志环路锁定,时钟振荡器的控制电压稳定,输出时 钟频率 vco_ckp 符合式(2)的计算,达到 6.4 GHz. 时钟树延时接近 120 ps,环路锁定后 clkfb 与 clkref 之间的稳态相差小于 15 ps.



图 8 去偏斜锁相环远端反馈锁定波形图

5 系统实验及同步分析

5.1 同步时序分析

基于去偏斜锁相环的仿真分析结果,结合本文 提出的同步时钟结构进行时序性能评估.各主要模 块时钟域输入输出均为高频时钟如图 5 所示.时钟 域的入口具有各自的去偏斜锁相环,锁相环带宽随 输入参考时钟频率增加自适应增大,具有相位快速 跟随的功能.本文设计的去偏斜锁相环最高输入不 分频时钟频率可达 800 MHz.各模块内时钟叶子节 点相位通过去偏斜锁相环环路跟随输入参考时钟相 位.最慢工艺角下,锁相环远端反馈稳态相差小于 15 ps,软件补偿后小于 10 ps.

全局时钟分布实现各模块入口时钟 mid_ck 与 系统主时钟 main_ck 相位同步.以 16 核为例,4 个 mid_ck 可通过去偏斜锁相环与 main_ck 对齐,同样 在补偿后存在小于 10 ps 的稳态相差. 全局和区域两层时钟级联会对稳态相差产生叠加.另外 mid_ck和 main_ck 均存在物理分布,叠加后相位偏差小于 10 ps.数字校准后,级联整体相位误差小于 30 ps.

5.2 系统能耗实验

时钟级联结构通过锁相环的远端反馈实现了不同区域的时钟树子节点相位对齐.进一步展开下面的系统能耗实验.基于解耦 PLL 当拍响应反馈控制的结构特点,验证采用间歇时钟的功耗控制效果. 5.2.1 实验平台

本文采用仿真加速器对 16 核 LS3C5000 处理器 RTL 代码进行 CPU SPEC 2000 测试,其中包括 定点和浮点 test 规模测试、启动阶段测试和低负载 测试.

LS3C5000 处理器的处理器核采用四发射乱序 结构,最多每拍时钟取 8 条指令,各队列加起来共 10 个发射端口.处理器的高速缓存结构分三级,具 体各级缓存的描述及访问延时如表 2 所示.

表 2 高速缓存分级列表

级数	高速缓存描述	访问延时
L1	数据和指令的私有高速缓存	4 拍
L2	数据和指令共享的核内私有高速缓存	14 拍
L3	核外片上共享高速缓存LLC(Last Level Cache)	约 50 拍

如果以上所有高速缓存都未命中,访问内存的延 时通常超过100个时钟周期(处理器核时钟).可见高 速缓存访存失效会明显增加流水线阻塞时间.本文实 验针对 LLC 访存失效也进行了性能计数统计.

5.2.2 实验算法

解耦 De-skew PLL 的级联时钟结构结合处理 器核内部标志信号,可当拍响应处理器核动态功耗 的低功耗控制.本实验针对这个功耗优化方法进行 评估,并抓取了访存失效率作为参考.动态功耗的低 功耗控制选择合适的时机间歇处理器核时钟. 流水 线间歇反馈过程中锁相环保持锁定,不需要通过系 统降低锁相环倍频系数进行动态降频 控制回路发 生在处理器核与其所属的解耦锁相环之间,因此传 输路径短反馈速度快.与此同时,流水线间歇不能影 响系统运行性能,需要选择流水线已空或者已堵没 有可以运行的操作时进行时钟间歇.

判断流水线空转或阻塞的条件主要有以下几个 方面:访存或取指是否存在失效;功能部件可执行指 令是否为空;是否有新的指令进入流水线;是否有新 的指令待发射;以及在执行访存操作全局可见操作 时是否遇到较大的延迟.

根据以上几方面归纳成三类统计条件: (1) noins fetching. 取指时发生 cache miss,导 致流水线全空的周期数.

(2) sync_waiting. 执行多核同步需要触发存储 一致性全局可见且该同步指令已经是流水线中最老 的指令而导致的流水线停顿周期数.此时,除核内高 速缓存部件外其它功能和运算部件均无可执行的 操作.

(3) noissue. load 指令发生 cache miss,且该指 令已经是流水线中最老的指令,当拍没有新指令取 指进入流水线也没有指令从发射队列发射并执行的 周期数.相比针对单一访存失效空拍应用时钟间歇 的调控方法,如果发生连续空拍且采用时钟间歇,能 效提升的效果会更明显.因此细化访存失效拍数的 统计条件,观察失效空拍是连续空拍还是单一空拍. 对 noissue 进行多拍情况扩展,最近 N 拍发射执行 的指令数之和小于等于 M 的情况,其中 N 和 M 分 别小于等于3和4,共6种配置.最长的统计设置为 3le4_nomap,其含义为当拍无新指令取指进入,且 最近3拍发射执行的指令数之和小于等于4.

采用以上几类统计模式将多个计数器嵌入 LS3C5000 处理器的 RTL 代码中,在仿真加速器上 启动 Linux 系统并运行 SPEC CPU 2000 16 线程的 test 规模输入集. 通过统计各类可间歇拍数占总运 行拍数的百分比来评估采用本文提出的间歇时钟系 统的时钟能效提升空间.

5.2.3 实验结果及分析

实验采用 16 核 LS3C5000 的数据进行测试统 计.每4个核一个节点,各节点的数据结果分布类似. 这里选取了节点0下4个核的数据,统计结果如图9



图 9 CPU SPEC 2000 各处理器核内各类可间歇空拍占整体运行拍数比例统计

所示.通常访存密集型浮点测试用例会带来更多的 时钟能效提升空间.采用多种测试用例整体统计的 目的在于观察间歇时钟降频方法的普适性(时钟能 效提升的平均水平).

测试统计结果中浮点测试(ft)的流水线可间歇 计数结果占比明显大于定点测试(int).且浮点测试 中 noissue占比超过 24%,明显高于 sync_waiting 和 noins_fetching.noissue 扩展下连续空拍条件的 统计结果与单一空拍基本一致.这说明在绝大多数 情况下 load 指令发生 cache miss 且流水线无指令 可发射执行后,通常后面一段时间内也基本无指令 可发射.这段时间可进行时钟间歇,noissue 可作为 触发条件.当 load 指令从下面的存储层次得到数据 回填后,即可恢复正常时钟.除此以外,启动过程 (setup)和低负载(low_load)情况应用间歇调节作 用不明显,原因在于 LS3C5000 处理器中遇到系统 空闲则操作系统执行 wait 指令进入空闲态,此时已 有硬件关闭时钟的设计.

进一步地,本文统计了浮点与定点程序的片上 LLC访问失效率 MPKI(Misses Per thousand Instructions).处理器运行定点程序时 LLC的 MPKI 值为 0.3;运行浮点程序时 LLC的 MPKI 值小于 2.5.浮点程序的 LLC 访问失效率超过定点程序的 5 倍是导致定点程序可供时钟间歇的空间明显低于 浮点的主要原因.以此为基础,采用 PTPX 进行后 仿真功耗分析.对比本文提出的采用核内 noissue 控制条件触发解耦 De-skew PLL 当拍响应间歇时 钟的方法和正常连续时钟供应处理器运行的方法进 行功耗仿真计算.MAX 寄生提取方式下,两种不同 仿真条件的仿真结果如表 3 所示.定点和浮点程序 下,间歇流水线的功耗提升在最差工艺角下分别达 到 4.872%和 20.3%.

表 3 连续时钟间歇方式时钟功耗降低比例

工艺角	定点/%	浮点/%
TT 0.8V 85 度	5.352	23.0
FF 1.05V 125 度	4.872	20.3

解耦锁相环响应控制时钟间歇的逻辑路径短, 对工艺角不敏感,而 FF 1.05V 125 度的漏电功耗 是 TT 0.8V 85 度的 7.94 倍.可见后者是导致工艺 角差异的主要因素.

本实验受限于仿真速度仅运行了 SPEC CPU 2000 test 规模输入集.这个规模测试占用的内存空间小,缓存访问失效率也不高.当运行较大规模测试

程序如 SPEC CPU 2006 ref 规模输入集时,片上末 级缓存的 MPKI 随之增加,采用本文提出的方法可 以获得更高的能效收益.

6 总 结

时钟分布延时、相位同步性能及低功耗响应是 时钟系统结构设计的三个主要方面.本文提出了一 种基于解耦去偏斜锁相环的同步间歇时钟系统方 案.相比已有的动态变频的功耗调节方法,本时钟结 构用调相的方法实现调频,使用处理器核内部反馈 信号进行时钟的间歇控制,响应时间小于1ns,取得 数量级的性能提升.解耦 De-skew PLL 的远端反馈 机制一方面从结构上保证了时钟系统的相位高效同 步,节省了已有设计中采用数控延时链进行相位补 偿的冗余;另一方面显著降低时钟区域分布延时,区 域面积越大时钟分布延时越大,大规模区域时钟网络 的分布延时降低为时钟入口处的稳态相差(校准后 小于10ps).本方案在时钟同步、功耗细颗粒度控制 上相比其它实现方案具有的优势统计见表 4.

表 4 时钟系统特征参数对比

Ň	数控延时 链时钟树 补偿	时钟 过冲	低功耗 调节方式	低功耗 控制时间 开销	处理器核 与共享缓存 异步 FIFO
[4]	<i>F</i>	> -	处理器核 动态调整	控制时间 间隔1ms	_
[10]	需要	无过冲	系统控制 DFS调整	_	需要
[16]	_	2%过冲	系统控制 DFS调整	_	需要
本设计	不需要	无过冲	处理器核 动态调整	控制时间 小于1ns	不需要

本文采用从电路到系统不同层面对提出的系统进行性能评估.首先,针对方案的重点电路采用 12 nm 数字工艺实现了环路设计并完成物理版图, 通过对版图寄生提取进行远端时钟树反馈环路稳态 相差仿真.仿真结果显示最差工艺角下,时钟树远端 叶子节点与输入参考时钟稳态相差小于 15 ps,校准 后可小于 10 ps.基于此仿真结果,对以 4 核为单位 的级联时钟结构进行时钟同步偏差分析,每层时钟 分布校准后同步偏差小于 10 ps.其次开销增加方 面,去偏斜锁相环 De-skew PLL 的面积相比处理器 核的面积开销增加小于 0.55%.整个锁相环功耗小 于 10 mW,占处理器核功耗功不足 0.1%.两个 DFC 器件开销小于 10 个标准单元,在面积和功耗上可以 忽略.间歇控制的外围控制逻辑集成在原有的功耗 控制模块内部,面积及功耗开销也可忽略.最后系统 方面,使用 16 核 LS3C5000 处理器 RTL 的仿真加 速平台运行 SPEC CPU 2000 测试集,评估本方案 对处理器核时钟动态功耗的影响,并在此基础上进 行 PTPX 功耗仿真计算,仿真结果表明定点程序平 均功耗节约超过 4.5%,浮点程序平均功耗节约超 过 20.3%.

时钟系统将完整嵌入到下一代多核处理器中进 行流片验证.在已展开的针对处理器核进行本地快 速时钟间歇降频控制的基础上,未来将时钟间歇范 围扩展到前端总线,进一步降低总线中的冗余等待 功耗开销.

参考文献

- Moore G E. Cramming more components onto integrated circuits. Proceedings of the IEEE, Glasgow, 2002, 86(1): 82-85
- [2] Rotem E, Naveh A, Rajwan D, et al. Power-management architecture of the Intel microarchitecture code-named Sandy bridge. IEEE Micro, 2012, 32(2): 20-27
- [3] Hammarlund P, Martinez A J, Bajwa A A, et al. Haswell, The fourth-generation Intel core processor. IEEE Micro. 2014, 34(2): 6-20
- [4] Doweck J, Kao W, Lu A K, et al. Inside 6th-generation Intel core: New microarchitecture code-named Skylake. IEEE Micro, 2017, 37(2): 52-61
- [5] Arafa M, Fahim B, Kottapalli S, et al. Cascade Lake: Next generation Intel Xeon scalable processor. IEEE Micro, 2019, 39(2): 29-36
- [6] Vera X. Inside Tiger Lake: Intel's next generation mobile client CPU//Proceedings of the HOT CHIPS 32 Symposium. Palo Alto, USA, 2020: 1-26
- [7] Yang Liqiong, Wang Linfeng, Xiao Junhua, et al. A 1.2 V,
 3.1% 3σ-accuracy thermal sensor analog front-end circuit in
 12 nm CMOS process. Journal of Semiconductors, 2021,
 42(3): 1-6
- [8] Al-Obaidy F, Asad A, Mohammadi F. Power-management based on reconfigurable last-cache level on non-volatile



YANG Li-Qiong, Ph. D. candidate. Her research interests include computer architecture, high-speed link, multi-core processor clock systems, and sensors on chip. memories in chip-multi processors//Proceedings of the IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). Edmonton, Canada, 2019: 1-4

- [9] Tam S, Rusu S, Chang J, et al. A 65 nm 95 W dual-core multi-threaded Xeon processor with L3 cache//Proceedings of the IEEE Asian Solid-State Circuits Conference. Hangzhou, China, 2006: 15-18
- Hu W W, Zhang Y F, Yang L, et al. Godson-3B1500: A 32 nm 1.35 GHz 40 W 172.8 Gflops 8-core processor// Proceedings of the IEEE International Solid-State Circuits Conference. San Francisco, USA, 2013: 54-56
- [11] Yuffe M, Knoll E, Mehalel M, et al. A fully integrated multi-CPU, GPU and memory controller 32 nm processor// Proceedings of the IEEE International Solid-State Circuits Conference. San Francisco, USA, 2011: 264-265
- [12] Fayneh E, Yuffe M, Knoll E, et al. 14 nm 6th-generation core processor SoC with low power consumption and improved performance//Proceedings of the IEEE International Solid-State Circuits Conference. San Francisco, USA, 2016: 72-73
- [13] Singh T, Schaefer A, Rangarajan S, et al. Zen: An energyefficient high-performance x86 core. IEEE Journal of Solid-State Circuits, 2018, 53(1): 123-234
- [14] Fayneh E, Knoll E. Clock generation and distribution for Intel Banias mobile microprocessor//Proceedings of the Symposium on VLSI Circuits Digest of Technical Papers.
 Kyoto, Japan, 2003: 17-20
- [15] Kurd N, Mosalikanti P, Neidengard M, et al. Next
 Generation Intel core micro-architecture (Nehalem) clocking//
 Proceedings of the IEEE International Solid-State Circuits Conference. San Francisco, USA, 2009: 1121-1128
- Ravezzi L. Partovi H, Wang D, et al. Clock and synchronization networks for a 3 GHz 64 bit ARMv8 8-core SOC// Proceedings of the 40th European Solid State Circuits Conference. Venezia Lido, Italy, 2014: 247-250
- [17] Maneatis J G, Kim J, McClatchie I, et al. Self-biased highbandwidth low-jitter 1-to-4096 multiplier clock generator. IEEE Journal of Solid-State Circuits, 2003, 38(11): 1795-1803
- [18] Viswanathan B, Nair S. R. R., Viswam V, et al. 4 GHz 130 nm low voltage PLL based on self biased technique// Proceedings of the 23rd International Conference on VLSI Design. Honolulu, USA, 2010: 330-334

WU Rui-Yang, Ph. D. His main research interest is microarchitecture design of processor cores.

YANG Liang, Ph. D. His main research interest is physical design of high-performance processors.

WANG Huan-Dong, Ph. D. His main research interest is multi-core processor architecture design.

Background

At the beginning of the 21st century, the development of microprocessors' integration and main frequency performance increased with transistor sizes shrunk followed Moore's law. Presently, the process of silicon transistors has entered the nano time, and the integration degree of multi-core processors exceeds hundreds of millions of gates scale. With the improvement of processor's performance, power consumption is becoming more and more serious. Processor working frequency's increase not only accelerates clock path flips, but also speeds up the data evaluate path, resulting in the overall increase of power and temperature of the chip. In turn, the temperature rise inside the chip not only aggravates the leakage power of logic gates, but also increases the connection resistance of the power supply network, which leads to a vicious circle caused by greater current consumption and voltage drop of the power supply. Moreover, the power consumption and temperature of the chip directly determine the cost of packaging and cooling. These issues have led processor design to pay special attention to energyefficient designs that reduce power consumption while maintaining high performance. Intel's processor architecture design has continuously improved and innovated for energy efficiency design in recent years. Sandy Bridge adopted Turbo-Boost 2.0 technology to classify high-performance and low power state and shifted by operating system (OS) control. Haswell fully integrated voltage regulator, which enables multi-phase multi-stage control of regional power. Skylake's ISST (Intel Speed Shift Technology) control plane has changed radically, with dynamic power tuning from the top operating system to the underlying processor core module for a 14% energy efficiency boost compared to the previous generation, and Cascade Lake and Tiger Lake further refining the efficiency control of dynamic power regulation. The close relationships between performance and power consumption led to ever-closer co-design of system and physical. Clock power

consumption accounts for the main part of processor core power consumption, while clock power consumption is mainly derived from pipeline clock's flip and data's evaluation. The efficiency of the pipeline is the main indicator of processor performance, which includes both clock frequency and execution efficiency (non-redundant empty beat ratio). The reflow is caused by transfer guess error, the waiting cycles for load store instructions due to the limited bandwidth of the memory cells, and the timing redundancy caused by asynchronous communication, all of these above affect the execution efficiency. How to reduce processor core's clock power consumption while keeping high performance is the main focus of this paper. In this paper, a new synchronous intermittent clock system based on decoupling De-skew PLL is proposed, which can realize the dynamic clock control response at subnanosecond by adding hardware design to monitor the core operation status of the processor in real time. In this paper, the new clock system structure is presented and relevant important circuits are described, supplemented by circuit-level simulation and the operation of real multi-task load in the whole multi-core system to evaluate the effect of low power control. Taking 12 nm cascade clock system structure as an example, each stage clock distribution variation achieved less than 10 ps after phase compensation, and the decoupling frequency module guaranteed no overshooting frequency risk. Using 16 core LS3C5000 RTL to run SPEC CPU 2000 on the simulation acceleration platform to evaluate the impact of the intermittent clock system, and further through PTPX postsimulation verification, the test results show that the average power saving of fixed-point and floating-point program are greater than 4.5% and 20.3%, respectively.

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences, Grant No. XDC05020100.