·体系结构与软件技术· 文章编号：1000-3428(2017)08-0074-08 文献标志码：A 中图分类号：TP391

基于大数据的高能效数据中心服务器研究

张俊卫[1a,1b]，王 晶[1a,1b]，张伟功[1b,2]，邱柯妮[1c]

（1. 首都师范大学 a. 信息工程学院；b. 高可靠嵌入式系统北京市工程技术研究中心；
c. 电子系统可靠性技术北京市重点实验室，北京 100048；2. 北京数学与信息交叉科学 2011 协同创新中心，北京 100048）

摘 要： 分析 ARM 服务器的发展以及应用现状，结合数据中心的特点，设计基于 ARMv8 架构的高能效数据中心服务器，选取目前工业界具有代表性的 ARMv8 服务器参数，使用典型的数据中心负载，对 X86 和 ARMv8 两种架构的微体系结构级指标、性能和功耗进行定量评估。实验结果表明，与 X86 架构服务器相比，ARMv8 架构服务器在数据中心领域具有更好的能效比，能够有效降低能耗和成本。

关键词： 数据中心；服务器架构；ARMv8 架构；评测模型；能效

中文引用格式： 张俊卫，王 晶，张伟功，等. 基于大数据的高能效数据中心服务器研究[J]. 计算机工程，2017，43(8)：74-81.

英文引用格式： Zhang Junwei，Wang Jing，Zhang Weigong，et al. Research of High Energy-efficient Data Center Server Based on Big Data[J]. Computer Engineering，2017，43(8)：74-81.

Research of High Energy-efficient Data Center Server Based on Big Data

ZHANG Junwei[1a,1b]，WANG Jing[1a,1b]，ZHANG Weigong[1b,2]，QIU Keni[1c]

（1a. College of Information Engineering；1b. Beijing Engineering Technology Research Center of High Reliable Embedded System；
1c. Beijing Key Laboratory of Electronic System Reliability Technology，Capital Normal University，Beijing 100048，China；
2. Beijing Center for Mathematics and Information Interdisciplinary Sciences，Beijing 100048，China）

【Abstract】 This paper analyzes the development and present application situation of ARM server. Combines the characteristics of data center，it designs an energy-efficient data center server based on ARMv8 architecture. It selects present representative ARMv8 server parameters，quantitatively evaluates the micro architecture index，performance and power consumption of the ARMv8 and X86 architectures using typical data center loads. Experimental results show that，compared with X86 architecture，ARMv8 architecture server is more energy-efficient in the field of the data center，which can effectively reduce energy consumption and cost.

【Key words】 data center；server architecture；ARMv8 architecture；evaluation model；energy-efficiency

DOI： 10.3969/j.issn.1000-3428.2017.08.013

0 概述

随着互联网技术的飞速发展，互联网所提供的服务从最初的静态网页已经发展到提供全方位的服务，人们在应用的过程中积累了大量的数据，据专家统计，2013 年的数据总量已经超过 4 万亿 GB，且每年的数据增加仍在加速[1]，因而具有强大处理能力和存储功能的数据中心得到了迅猛发展。然而，快速发展的数据中心面临着高成本、高能耗、低效率等诸多挑战[2]。大型数据中心有成千上万台服务器，每台服务器的空闲功耗一般在 100 W 以上[3]，数据中心的功耗等于服务器运行功耗乘以服务器数量，再加上其他配套设施的能源消耗。有数据表明，一个拥有 10 万台服务器的公司，每年用于服务器的能量消耗大概在 2 亿美元左右[4]。

绿色数据中心已经成为国内外持续关注的重要话题之一。当前针对于数据中心的高能耗问题，主要的解决方案包括充分利用自然环境、低功耗的程序设计以及低功耗硬件[4]。其中开发低功耗硬件是最直接也是最有效的方法。处理器的能源消耗量[5]取决于处理器架构以及制造工艺。然而，制造工艺的进步已经越来越难，从制造工艺的角度来降低功耗难度较大，而从处理器架构方面来考虑制造低功耗硬件则是行之有效的方案。

计算机架构的设计从 RISC 诞生开始，CISC 和 RISC 的优劣就是学术界和工业界争论的焦点。在

基金项目： 国家自然科学基金(61402302，61170009，61472260)。

作者简介： 张俊卫(1988—)，男，硕士研究生，主研方向为计算机体系结构；王 晶，讲师；张伟功，研究员；邱柯妮，副研究员。

收稿日期： 2016-07-19 **修回日期：** 2016-08-25 **E-mail：** jwang@cnu.edu.cn

过去的十多年中,以 ARM 为代表的 RISC 架构占据了低端嵌入式计算市场,而 CISC 架构的 X86 则占据了桌面计算和服务器市场。然而,当计算机系统的设计从对芯片面积和设计复杂度的关注,转移到功耗和节能时,研究人员关于 RISC 和 CISC 的讨论逐渐升温。数据中心对能效性的追求,使得基于 ARM 处理器的服务器能够发挥重要作用。

ARM 处理器一直在移动嵌入式等领域占有主导地位,而近年发布的 64 位 ARMv8 架构则在数据中心服务器领域得到了学术界和工业界的普遍关注,其原因在于 ARMv8 架构的处理器核(小核)相比于传统 X86 核(大核)微体系结构简单得多,因此,其在功耗方面具有绝对的优势,同时在相同的芯片面积之下,ARMv8 架构可以集成更多数目的处理器核,具有更高的系统吞吐量,这可以满足典型数据中心应用对单线程性能要求不高而更加注重系统吞吐量的特征。另外,在移动市场上的长期积累使得 ARM 架构具有相对完善的生态环境,在 ARM 平台上进行开发的程序员为数众多,这些都将为 ARMv8 服务器生态系统的建立奠定坚实的基础。

综上,ARMv8 架构目前已被认为在未来的数据中心服务器市场上具有强大的竞争力,甚至有可能取代传统的 X86 服务器。然而,对于 X86 和 ARM 的对比分析大多针对于性能指标,现有工作缺乏基于云计算应用的能耗和能效性的定量评测。而数据中心场景下的云计算应用和传统的桌面应用从自身特性到关注的性能指标都有很大的不同,基于 SPEC 等基准测试程序的评测工作无法反映云计算应用下 ARM 和 X86 架构的优劣,更无法定量地给出 ARM 服务器相对于 X86 的能效优势。针对上述问题,本文分析了 ARM 架构的发展历程,结合当今工业界的 ARMv8 服务器,系统地总结了具有代表性的 ARMv8 服务器芯片的体系结构特征。在仿真环境下使用典型的数据中心负载对 ARMv8 小核在不同结构(变换核数和最后一级高速缓存的容量)下的性能和功耗进行了评估,并与传统的 X86 大核进行了对比分析。

1　相关工作

目前,通用服务器处理器的架构大多是 CISC 架构的 X86 处理器,性能非常强大,但缺点是功耗也比较大。基于 RISC 架构的 IBM 的 Power PC 和 Oracle 的 SPARC 一般应用在大型机领域,这类计算机体系结构封闭,性能强大,主要应用在高端行业的核心节点,采用该芯片的服务器价格昂贵,因而不适合应用在具有大量服务器的数据中心领域。而同为 RISC 架构的 ARM 处理器则具有不断发展的架构以及独特的商业模式,使得可以考虑基于 ARM 架构来设计数据中心服务器的处理器。

2013 年 1 月,百度公司开放南京数据中心,该数据中心的服务器是全球第 1 个商用 ARM 架构的服务器。该服务器采用 Marvell 公司推出的一站式解决方案,目前正在被用于支持名为“百度云”的个人云存储应用服务。随着云服务和云应用以及网站托管和社交站点所带来的“高存储-轻计算”任务的快速增长,当前的数据中心对服务器提出了全新的要求,ARM 低功耗、低成本的优势明显。根据百度公司提供的数据来看,新的存储器功耗降低 50%,存储密度提升 70%,总体应用成本降低 25%,接入成本比同类商用方案下降 83%[6]。Marvell ARM 服务器解决方案在实际应用效果虽好,但处理器的位宽是 32 位,寻址空间有限,而数据中心的典型负载的数据集都比较大,这就要求处理器的位宽要大,因此需要 64 位架构来满足需求。

2011 年 11 月,ARM 公司发布了首款支持 64 位指令集的处理器架构 ARMv8,并于 2012 年率先发布了两款 ARMv8 架构:高端版本 Cortex-A57 和低端版本 Cortex-A53,其中 A53 主要应用于智能手机市场。Cortex A57 的 SPEC 2006 得分大约在1 200 分上下,Cortex A9 的得分在 1 000 分左右,Cortex A53 的得分与 Cortex A9 相当,但是功耗更低[7]。2015 年 2 月,ARM 正式发布了基于 ARMv8-A 架构的 Cortex-A72,A72 在性能方面最多可以达到 A57 性能的 1.8 倍,功耗比 A57 有 18% ~30% 的下降,频率有了 20% 的提升。与 X86 架构的 Intel Core M 5Y10C 处理器相比,4 核 A72 处理器在单线程方面 SPECint,SPECfp 均为 Core M 的 0.8 倍,多线程方面其 SPECintRate 得分是 Core M 的 1.7 倍,在内存基准测试程序 STREAM 的测试中,A72 仅在 copy 方面比 Core M 略低,在 add,scale,triad 方面均高出 Core M 处理器,功耗方面 A72 为 1 W,而 Core M 则为 4 W。在与 Intel 服务器系列芯片 Xeon 对比中,频率 2.7 GHz、20 核 20 线程的 A72 在 SPEC 2006 测试下性能超越了 Xeon E5 2650 v3,接近更高级的 E5 2660 v3,功耗方面 A72 只有 30 W,而 Xeon 系列则有 105 W[8]。在制造商方面,2012 年 6 月,Applied Micro Circuits Corporation 宣布了全球首款与 ARMv8 架构相兼容的处理器“X-Gene”[9],该处理器采用自行设计与 ARMv8 架构相兼容的指令集,主要面向下一代云计算、无线基础架构、企业级网络、存储和安全等应用[10]。2014 年 6 月,Applied Micro 公布了其 X-Gene 2 和 X-Gene 3 ARM 服务器芯片细节,X-Gene 2 和 X-Gene 3 相比 X-Gene 采用了更低的工艺和更多的核心,在 SPEC 2006 基准测试程序下性能比 X-Gene 提高了 0.5 倍[11]。2014 年 12 月,Cavium 公司正式发售 4 款 ThunderX 芯片,其主要应用领域分别是云计算工作、云计算存储、网络应用和计算安全。2014 年 1 月,AMD 公司发布代号为“西雅图”的 Opteron A1100 处理器[12]。该芯片的定位是需要低功

耗计算、大容量内存/存储的服务器领域。在国内方面，广州飞腾信息技术有限公司在顶级处理器技术研讨大会上发布了基于 ARMv8 指令集的火星(Mars)处理器，其浮点性能仅略低于 Power 处理器，大幅领先于 Intel 处理器，在整数测试方面低于 Intel 系列处理器而高于 Power 处理器。

著名研究机构 The Linley Group 评测了包括 Intel，Cavium，AMD，APM 在内的现有服务器处理器的 SPEC 整点性能，其评测结果如图 1 所示。其中所有处理器运行的 SPEC 程序都使用 GCC 编译器编译，而 Intel 处理器在 GCC 编译器下性能相比 ICC 编译器要低 10%，从图 1 中可以看出，Cavium 公司的 48 核处理器性能已经接近 Intel Xeon E5 处理器，Cavium 16 核处理器和 APM X-Gene 处理器的性能要强于 Atom 8 核处理器，AMD A1100 相对 Atom 处理器的最强性能具有一些差距[13]。

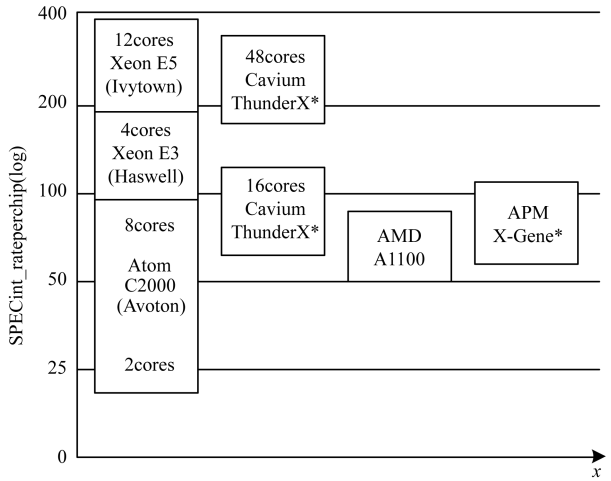


图 1 服务器处理器性能

当前基于 ARMv8 架构的服务器产品已经逐步走向市场，但是尚未真正应用在数据中心，想要部署在实际的生产环境中，需要综合考虑性能、功耗、价格等方面因素。

本文通过定量实验可以发现，在吞吐量 IPC 方面，ARM 处理器低于 X86，差别大小与具体程序有关，最大的差别是 1.56 倍，最小的差别是1.07 倍。而在功耗方面，ARM 处理器也低于 X86 架构，同样，功耗差别与具体程序有关，最大差别是1.97 倍，最小的差距是 1.52 倍。综合考虑两者的能效性方面，ARM 架构处理器的能效性是 X86 架构的1.25 倍。实验结果证明，ARM 架构在数据中心中具有高能效低成本的优势。

本文通过评测典型云应用基准测试程序，得出了更适宜 ARM 处理器的先后顺序：streaming，classification，nutch，cassandra，cloud9。其中，streaming 最适应使用 ARM 服务器；cloud9 最不适用 ARM 服务器。另外还评测了 8 核情况下典型云应用程序的最佳 LLC 配置，其中 cassandra，classification，cloud9，cloudstone 和 streaming 基准测试程序的最佳 LLC 配置是 8 MB，nutch 最佳 LLC 配置是 4 MB，该评测数据可以用来为定制化处理器提供参考。

2 评测模型

对于不同处理器架构在数据中心中的应用研究，有 2 个关键的方面，首先要分析当前工业市场主流的服务器处理器，结合体系结构设计原则，选取典型的微体系结构的参数，准确地模拟处理器特征；其次要选取区分于传统桌面应用基准测试程序，如 SPEC 等不具有典型云计算大数据引用特征的基准测试程序。本节将分别对处理器模型抽象和云计算应用负载进行讨论。

2.1 服务器的处理器模型

从 ARM 的发展历史分析，影响 ARM 处理器性能和功耗的重要参数包括发射宽度、频率和工艺等。而对不同制造商集成的服务器进行了分析和对比，可以看到影响服务器的因素还包含的处理器核数、Cache 的大小等参数。因此，综合上述因素，本文的处理器模型选择了 2 个方面的特征：1)代表处理器自身性能的核心内部的设计，包括处理器的工作频率、每个周期最大的指令发射宽度 Issue width、重排序缓存的大小 ROB 和存取缓存的大小 Load-store-buffer 和 LSB。2)服务器板级特征，包括一级指令 Cache 大小、一级数据 Cache 的大小、二级 Cache 的大小和总的核数。典型的 ARM 处理器参数如表 1 所示[13]。

表 1 典型 ARM 处理器主要参数

参数	Cavium ThunderX	Mars[14]	AMD A1100	APM X-Gene
核心	ThunderX	xiaomi	Cortex-A57	X-Gene
发射宽度	2	4	3	4
频率/GHz	2.5	2	2.0≠	2.4
核数/线程数	48/48	64/64	8/8	8/8
SPECint(2006)	350≠	672	80	110≠
多路	2S	1S	1S	1S
功耗(TDP)/W	80≠	120	25	35≠
工艺	28 nm HKMG	28 nm HKMG	28 nm HKMG	40 nm G

为了有效地对比 ARM 架构和 X86 架构，分别设置了 2 种配置，用大核代表处理能力强但功耗高的 X86 架构，用小核代表处理能力和功耗较弱的 ARM 架构。因此，小核是一个简单的、乱序发射的核，拥有两级 Cache，本文选择了 Applied micro 公司的 APM 883408 服务器；大核比小核拥有复杂的内

部结构,如更深的发射宽度、更大的队列深度、更大的 Cache 容量等,本文借鉴了 Intel Xeon 2690 处理器的设置。具体的参数选择如表 2 所示。

表 2　处理器微体系结构参数

类型	小核	大核
主频/GHz	2.4	2.9
L1I	每核拥有 32 KB L1I	每核拥有 32 KB L1I
L1D	每核拥有 32 KB L1D	每核拥有 32 KB L1D
L2	所有核共享 16 MB L2	所有核共享 32 MB L2
核数	8	8
Issue	2	8
ROB	32	192
LSB	16	64

2.2　云计算应用基准测试程序

云计算是一种新兴的计算平台,主要用来向全球客户端提供可扩展的在线服务。现在流行的在线服务主要包括网络搜索、社交网络、视频共享和数据存储等[15]。这些应用通常数据集较大,高并行度,且具有一定的实时性要求。云计算负载应包含如下的特性:按需服务,既能够根据负载的需求自动地提供计算能力;广泛的网络接入服务、云服务可以通过标准网络访问机制获得;资源池、物理和虚拟资源都能够根据多租户模型动态分配给多个用户使用;资源能够快速地分配和释放;测量服务、云应用能够计算当前的资源利用率,并控制资源的使用,实现按用途支付。因此,数据中心的应用程序跟桌面程序以及传统服务器程序不同[15-16],本文需要一种新型的测试基准程序来评测云计算中的程序性能。

典型的数据中心应用包括以下类形:

使用 MapReduce 框架对海量数据集通过机器学习实现数据分析的 Hadoop 应用,例如排序、词频统计和文件系统,Hibench 就是这种应用的典型基准测试程序。

基于雅虎云服务标准测试程序(YCSB)的数据服务,例如对于随机数据集的随机操作。YCSB 是一个测试数据存储系统的框架,这个框架为多种流行的数据服务系统提供测试接口。

在所有基于互联网的服务中,Web 服务是基础应用程序。基于 Faban 框架的网络服务 cloudstone,用来测试 Web2.0 应用程序的一个 Web2.0 的 social-events 应用程序和一个由 Faban 负载生成器实现的客户端,可以被用来测试各种 Web 技术和 Web 服务器软件栈。

搜索引擎应用,典型的 Nutch 搜索引擎由 3 个部分组成:一个客户端机器模拟真实环境下的客户;一个索引处理节点和一个前端服务器接收客户端请求并将其发送给索引处理节点。

基准测试程序 Cloudsuite 覆盖了前面提到的 4 种典型云计算复杂,同时还包括了流媒体和软件测试应用[16]。

Cloudsuite 的媒体流基准测试程序(Streaming)包括 2 个主要组件:一个客户端和一个服务器端。客户端模拟真实环境中的客户,通过发送请求来增加媒体流服务器端的负载。在默认配置下,客户端请求的视频长度和质量可以通过客户端提供的接口进行控制,从而可以通过控制使得客户端发出的请求与实际观察到的真实环境的相同。

Cloudsuite 的软件测试基准测试程序(Cloud9)则是一个既消耗资源又花费时间的任务,可以给云计算增加负载,它把应用程序的执行并行化并扩展到常规硬件配置的集群上。

本文所选择的 Cloudsuite 中的 6 种云计算基准测试程序,其描述如表 3 所示[17]。

表 3　经典云应用描述

cloudsuite 应用	描述
Data Serving	NoSQL 数据存储使用多种多样和快速发展的存储策略,可以提供快速和可扩展的存储,目前已经广泛地用来作为大规模 Web 应用的备份存储
Data Analysis	map-reduce 模式可以自动化分析处理大规模人为信息。它转发请求到集群节点先执行过滤和转换数据(map),然后聚合结果(reduce)
Media Streaming	流媒体服务使用大型服务器集群分包并传输 MB 到 GB 大小的媒体文件,文件被预先编码成不同格式和不同比特率的文件,以适应高带宽的智能家居和移动设备上多种客户需求
Sortware Testing	软件测试是一个大规模的仿真任务,云计算可以提供基于 IP 网络连接的动态异构的资源,因此软件测试负载在工作者排队模型中使用集中的负载均衡机制
Web Search	大规模的数据中心应用是用来提供访问公共互联网索引(如谷歌、必应、百度)。TB 级别的索引被分区,与每个索引节点(ISN)负责处理自己的分区的服务请求。前端节点并行地向所有 ISN 发送搜索请求,并收集整理 ISN 的回复,最后将格式化的回复发回请求客户端

3　评测指标

程序的性能和功耗是普遍被关注的指标,因此本文的评测从这两方面开展。

3.1　性能指标

对系统性能的衡量方式多种多样,本文的评测分别对核的处理性能和 Cache 的性能进行了统计分析。

1)吞吐量 IPC

传统的处理器采用处理器在每一时钟周期内执行的指令数 IPC 作为衡量 CPU 性能的指标。而在

高性能计算和数据中心场景下,单个处理器核的性能往往不是关注的焦点,取而代之的是整个系统的吞吐量。因此,针对于数据中心场景,本文选用吞吐量 IPC 来表达整个数据中心的性能,其定义如式(1)所示,其中,IPC_i 表示每个处理器核的 IPC。

$$Throughput_{\mathrm{IPC}} = \sum IPC_i \tag{1}$$

2)指令数

针对 IPC 的差别,本文进一步分析了指令的执行情况,并对比了处于空闲态、系统态、用户态和陷阱态的指令分布。

3)Cache 失效率

影响系统性能的因素不仅包括处理器的执行效率,Cache 的容量也是影响系统性能的关键因素,而 Cache 也恰恰是服务器处理器模型中的重要特性之一,因此,Cache 容量的变化对于基准测试程序的影响也是需要考虑的性能指标。而反映高速缓存性能的典型指标是每千条指令的失效率 MPKI,因此本文的性能指标选择了统计末级高速缓存 LLC 的 MPKI。

3.2 功耗指标

低功耗是体系结构设计者追求的目标。能耗问题也是引发 X86 服务器还是 ARM 服务器讨论的根本因素。能耗通常可以用功率 power 或者能耗 energy 来衡量。

然而单纯的低功耗已经不能满足绿色数据中心的需求。Intel 提出使用每条指令的能耗(Energy Per Instruction,EPI)作为处理器能效性的评测指标。

$$EPI = Runtime\ Dynamic \times Runtime/Instructions$$

其中,*Runtime Dynamic* 是由 McPAT 测量出来的运行时动态功率;*Runtime* 是基准测试程序在指定运行周期数下的运行时长;*Instructions* 是所执行的指令数。

功耗评测使用 McPAT 软件[18]。McPAT 是惠普实验室开发的一个结构级的模型框架,它集成了多核处理器的功耗、面积和时序模型。Mcpat 在功耗评测方面具有以下特点:

1)McPAT 的功耗模型不仅考虑动态功耗,还考虑静态功耗和短路功耗。

2)McPAT 为复杂的多核处理器部件提供功耗模型。

3)McPAT 使用 ITRS 提供的对动态功耗、静态功耗、短路功耗的技术预测,相比其他功耗工具,McPAT 的功耗建模比较准确。

4)McPAT 使用 XML 接口,该特性使 McPAT 能独立作为一种功耗模拟工具使用。

4 实验结果与分析

由于目前尚没有发现基于 ARMv8 架构服务器投入商用,因此在模拟环境中完成对比评测。Cloudsuite 基准测试程序提供了在 SimFlex 上运行的镜像,这些镜像允许 SimFlex 在 1,2,4,8,16,32,64 核下对 Cloudsuite 进行仿真。SimFlex 由 Simics[19] 和 Flexus 模拟器组成。Flexus 模拟器不是一个能够独立运行的模拟器,需要建立在 Simics 全系统模拟器的微体系结构接口(MAI)上。Flexus 模拟器由多个独立的模块连接在一起。其中模块通常相当于一个硬件结构。例如,一个分支预测或缓存。Flexus 的关键点是其隔离组件:一个特定的模块可以替换为一个不同的实现,而无需修改其他模块,这对于进行本次评测以及后续的实验提供了良好的基础,这也是选用 Flexus 模拟器的原因。

评测实验主要分为 2 个部分:第 1 个部分是评测小核和大核在不同核数下的性能和功耗,统计指标有 IPC 和 EPI,总计运行 15M 个时钟周期;第 2 部分评测 LLC 容量在 ARM 架构下对于 Cloudsuite 基准测试程序的影响情况,统计指标有 IPC 和 MPKI。实验运行周期设定为 20M 个时钟周期。

4.1 性能评测

图 2 是 5 种基准测试程序在不同核数下的指令数。其中横轴是基准测试程序,纵轴是处理器运行 15M 个时钟周期后的指令数。其中,x 核(n)代表 x 架构下 n 核,如小核(8)代表 ARM 架构下的 8 核数据。由于 cloudstone 部分镜像缺失,因此没有统计 cloudstone 数据。从核数变化范围可以看出,除了 streaming 外的其他基准测试程序的指令数都是随着核数的增加而增大,但是不同的基准测试程序增长幅度不同,cassandra 的指令数在 8 核到 16 核增长幅度大,从 16 核增长到 32 核的变化幅度反而比较小,而 classification、cloud9 和 nutch 基准测试程序的变化幅度大的出现在 16 核到 32 核之间,streaming 的变化有异常,指令数在 8 核到 16 核时性能出现了下降。从图 2中可以得到不同应用程序的最佳核数配置。

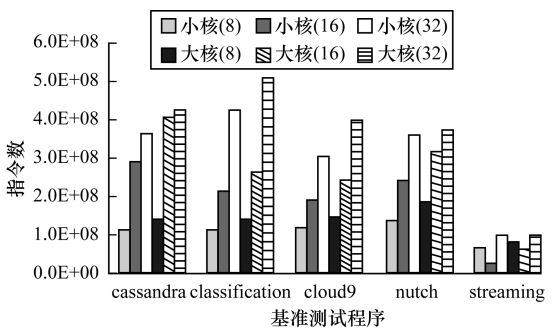


图 2 不同核数下基准测试程序的指令数

从 ARM 架构和 X86 架构方面可以看出,ARM 架构的指令数全部小于 X86 架构的指令数,但是差距不同。从这里可以得出不同的应用程序的最佳架构,除了指令数之外,还统计了吞吐量 IPC。

图 3 是 5 种基准测试程序在不同核数下的吞吐量 IPC。其中横轴是基准测试程序,纵轴是 IPC 的值。从

核数变化范围看,除了 nutch 与 streaming 2 种基准测试程序外,其余基准测试程序的 IPC 趋势与图 2 指令数变化趋势相似。nutch 的吞吐量 IPC 在 32 核时出现了下降,streaming 的吞吐量 IPC 在 16 核时出现了上升。

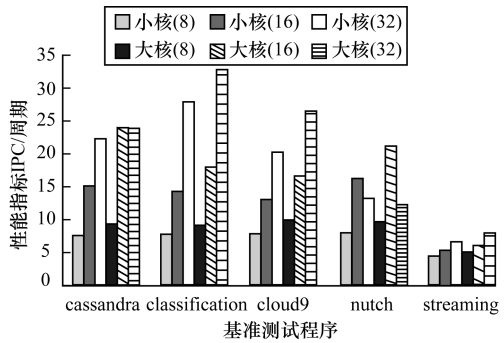


图 3　不同核数下基准测试程序的 IPC

从 ARM 架构和 X86 架构的角度来看,在核数相同的情况下 ARM 的 IPC 弱于 X86,不同程序从 ARM 到 X86 的性能提升幅度也不一样,以 8 核为例,提升效果从大到小的顺序是:cloud9,cassandra,nutch,classification,streaming。cloud9 基准测试程序对应的是 software test,这个应用对应于云计算中的自动化软件测试,计算量比较大且费时,大核的强大性能在这里得到体现。cassandra 基准测试程序对应于 data serving,该程序除了计算之外还有一定的文件操作,I/O 访问量比较大,因而大核架构对于性能的提升比 cloud9 小。nutch 基准测试程序对应于 Web search,这个程序主要是对索引进行处理,计算量不如上述 2 个,因而提升幅度排在中间。classification 基准测试程序对应于 data analysis,是典型地把大数据切分后分配给单个核心来计算,每个核心的计算量就不会很大,所以大核在这个应用上提升效果就比较小。streaming 基准测试程序对应于 media streaming,这种应用主要是网络要求较高,因此大核提升的幅度最小。

为了解释异常趋势,统计了 nutch 基准测试程序的具体指令数,指令分类数如图 4 所示。从图 4 中可以看出,在统计吞吐量 IPC 时没有统计 Idle 指令和 Trap 指令,nutch 程序的 Idle 指令数在 16 核时是最低的,User 指令是最大的,这也就解释了图 3 中 nutch 程序在 16 核时的 IPC 最高的原因。

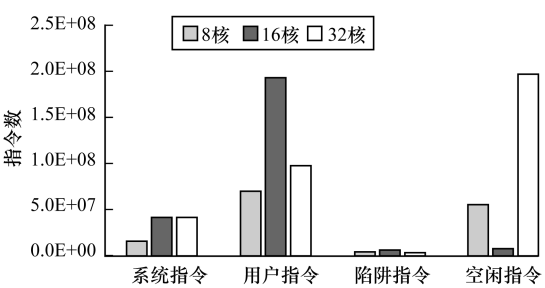


图 4　nutch 基准测试程序分类指令数

从图 3 中可以得出某些基准测试程序更适宜使用小核架构。另外,还统计了 2 种架构下的程序运行时间,具体运行时间如图 5 所示。从图 5 中可以看出,ARM 架构的运行时间大于 X86 架构。对于一些实时性需求比较大的应用,ARM 架构还需要进行一些改进。

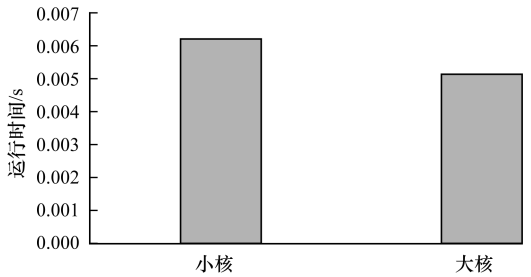


图 5　2 种架构下的程序运行时间

4.2　功耗评测

图 6 是 5 种基准测试程序在不同核数下的功耗。从核数变化方面可以看出只有 classification 是随着核数的增加而增大,其他基准测试程序的功耗并没有随着核数的增加而线性增长,产生上述现象的原因是不同程序的并行度不同,以 cassandra 为例,功耗在 16 核下是最高的,此时最大程度发掘了程序的并行性,32 核功耗相比 16 核并没有增加则是由于 32 核空闲较多,而 8 核最低则是硬件资源占用的少从而导致功耗更低。从处理器架构方面看,所有基准测试程序在 ARM 架构下的功耗都低于 X86 架构。因为 ARM 架构结构相比于 X86 架构简单得多,所以 ARM 架构的功耗小于 X86。

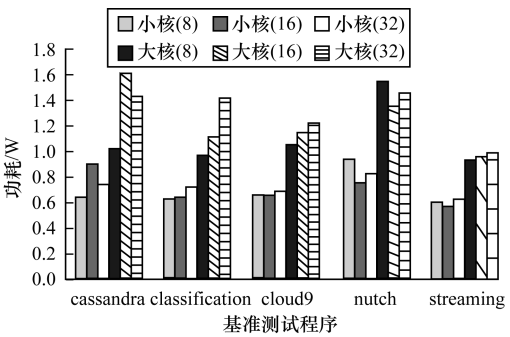


图 6　不同核数下基准测试程序的功耗

在处理器内部,还对 ARM 架构和 X86 架构在相同核数情况下对处理器内部不同部件进行了功耗测试,具体功耗情况如图 7 所示。图 7 展示了大小核处理器内部不同部件的功耗。从图中可以看出 2 种架构下不同部件的功耗。由于 ARM 架构的发射宽度、ROB 缓存和 LSB 缓存均少于 X86 架构,因此 ARM 处理器自身功耗远远低于 X86 处理器功耗。从总的功耗来看,综合图 3 中的 8 核 IPC,ARM 架构功耗不到 X86 架构的 3/5 却达到了大核 4/5 的性能,假如把 ARM 架构的功耗与 X86 架构的功耗

设置成一样时,那么 ARM 架构的能效比是 X86 架构的 1.25 倍。从图 7 可以看出,ARM 架构处理器应用在数据中心里具有更好的能耗比。

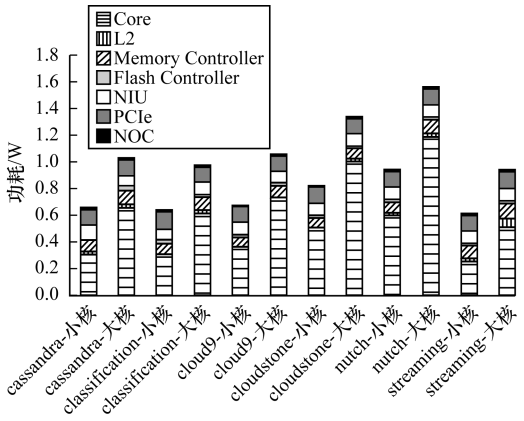


图 7 不同架构下处理器部件功耗

4.3 LLC 容量敏感度评测

LLC 的容量是影响程序性能的关键因素,因此评测了LLC容量变化对于不同基准测试程序的影响,选择 8 核的情况评测 LLC 大小为 1 MB,2 MB,4 MB,8 MB,16 MB 和 32 MB 下的 IPC 和 MPKI。

在实际系统中,Cache 访问延迟会随着容量的增大而增加,然而在模拟器中的 LLC 访问延迟需要计算并配置。访问 Cache 不同块的访问延迟是不同的,在实验中采用随机延迟来代表整个 Cache 的访问速度。随机延迟指程序随机访问 Cache 中的一个块所花费的周期数。在计算 Cache 随机延迟方面使用 Cacti 6.0 仿真模型。根据要测试的 LLC 容量大小配置 Cacti 文件,编译并运行 Cacti 获取各种容量下的随机延迟,根据同一容量下随机延迟与模拟器中的延迟建立对应比例关系,并依此比例计算出其他容量下的模拟器延迟,最后把 LLC 容量和延迟配置到模拟器中并进行测试,测试结果图 8 所示,左侧纵轴是基准测试程序的 IPC,右侧纵轴是基准测试程序的 MPKI。6 个基准测试程序的 IPC 和 MPKI 的取值范围是相同的,可以看出不同基准测试程序在相同的配置下具有不同的 IPC 和 MPKI。

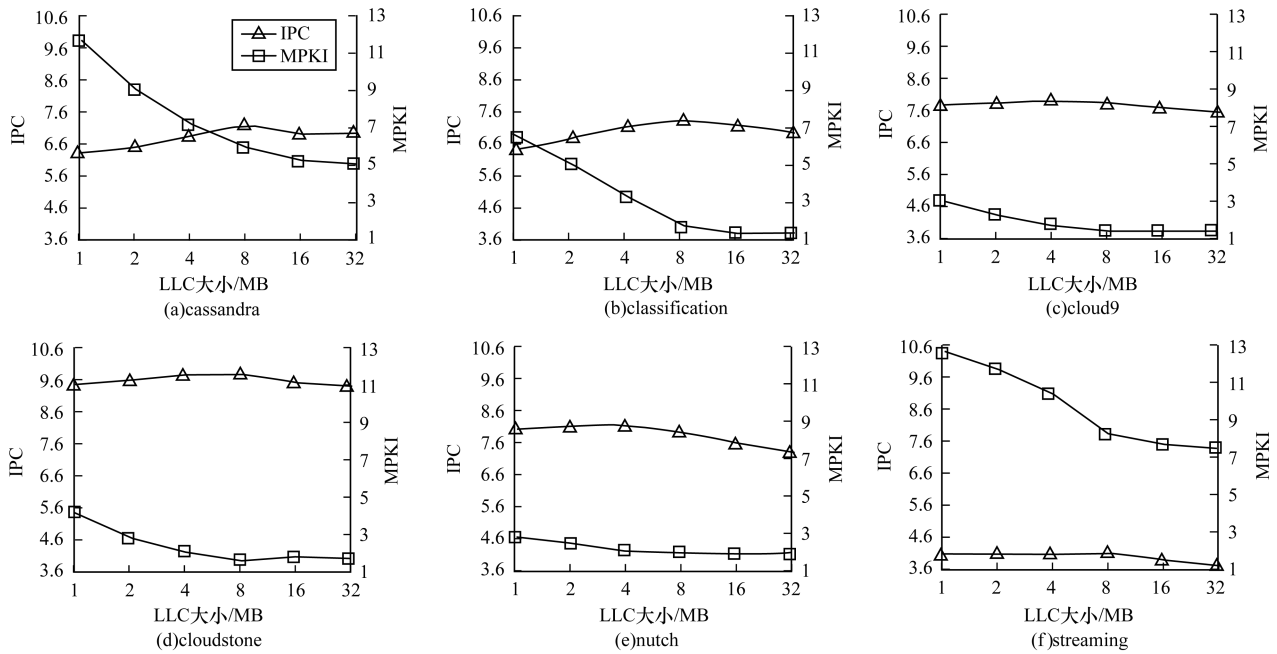


图 8 不同 LLC 下基准测试程序的 IPC 和 MPKI

从 MPKI 方面来看,随着 LLC 的增大,MPKI 会逐渐减小,但是当 LLC 容量达到基准测试程序的工作集,MPKI 就会保持不变。

从 IPC 方面来看,所有基准测试程序的趋势是先增大后减小。这是因为 Cache 命中率会随着 LLC 容量增长而增加,命中率增加会降低平均访问时间,因而程序性能会有一定程序的提升。然而,当 LLC 的容量增加到一定程度后,LLC 容量会大于基准测试程序的工作集,额外增加的容量不仅不会带来命中率的提升,反而会由于容量的增加导致 LLC 的访问延迟增加,这样导致了基准测试程序的性能降低,这也就是 IPC 先增长后降低的原因。

把 MPKI 值达到最低而 IPC 值达到最高时 LLC 大小的值称为 LLC 的最佳配置,从图 8 中可以看出不同基准测试程序的 LLC 最佳配置不一致。可以观察到不同基准测试程序的最佳 LLC 配置,为 ARMv8 服务器芯片的设计提供一些参考。

4.4 能效性评测

能效性是能耗和效率的比值,数据中心在追求低功耗的同时也要兼顾性能,能效性是两者的折中,

是当前数据中心设计者关注的重要指标,也是数据中心服务器处理器设计和选择的重要参考。

本文用每条指令消耗的能量代表能效性,因此实验中测试了平均每条指令在执行过程中消耗的能量,单位是焦耳,不同基准测试程序在不同核数下的 EPI 如图 9所示。图 9 展示了基准测试程序在不同核数、不同架构下的 EPI。在核数相同的情况下,ARM 架构的 EPI 都小于 X86 架构,这说明 ARM 架构平均每条指令执行消耗的能量更少,能效性更高。在同一种架构下,除了 nutch 程序外其他基准测试程序的 EPI 随着核数的增大而减少,多核服务器充分利用了数据中心并行度高的特点,但由于不同程序的数据特点使得每个应用程序对应的能效性最高的核数也不尽相同。

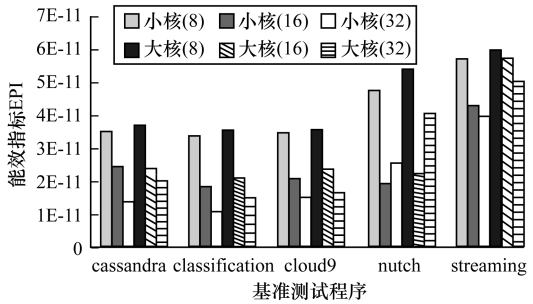


图 9　不同核数下基准测试程序的 EPI

综上所述,使用小核架构的服务器具有高性能、低功耗的特点,在某些应用的数据中心上可以使用 ARM 服务器替代现有的 X86 服务器。

5　结束语

本文综述了 ARM 服务器的发展以及应用现状,结合数据中心的特点,设计基于 ARMv8 架构的高能效数据中心服务器。针对不同的架构建立了评测模型,对 2 种不同架构的处理器进行抽象并评测性能和功耗,此外,评测了不同容量 LLC 在 ARMv8 架构中对于不同基准测试程序的影响情况。实验结果表明,ARMv8 架构在数据中心服务器方面相比现有的 X86 架构具有更好的能效,是未来数据中心服务器架构的发展趋势。在 LLC 的容量方面,不同的应用对于 LLC 的需求也不相同,每个基准测试程序都拥有各自最佳的 LLC 配置,通过评测,可以为以后 ARM 服务器处理器的发展以及应用提供借鉴。

参考文献

[1] 贾　同. 大数据对高等教育发展的推动研究[D]. 重庆:西南大学,2015.

[2] 孙　鑫. 面向云环境数据中心的高效资源调度机制研究[D]. 北京:北京邮电大学,2012.

[3] Govindan S, Sivasubramaniam A, Urgaonkar B. Benefits and Limitations of Tapping into Stored Energy for Data Centers[J]. ACM SIGARCH Computer Architecture News, 2011, 39(3): 341-351.

[4] 张桂刚,李　超,邢春晓. 一种云环境下的绿色计算模型[J]. 小型微型计算机系统,2013,34(5):1016-1020.

[5] Zhong Benjamin, Feng Ming, Lung Chung-horng. A Green Computing Based Architecture Comparison and Analysis[C]// Proceedings of IEEE/ACM International Conference on Green Computing. Washington D. C., USA: IEEE Press, 2010: 386-391.

[6] 赵艳秋. ARM 楔入数据中心[J]. IT 经理世界,2013(9): 32-33.

[7] ARM 公司. ARMv8 架构[EB/OL]. [2016-06-21]. http://www. icpcw. com/Smartphone/iphone/iphonedg/3209/320921_all. htm.

[8] ARM 公司. Cortex-A72 架构[EB/OL]. [2016-06-21]. http://news. mydrivers. com/1/423/423236. htm.

[9] Applied Micro Circuits Corporation. X-Gene[EB/OL]. [2016-06-21]. http://dy. qq. com/article. htm? id=20150706A00LTH00.

[10] Applied Micro Circuits Corporation. X-Gene[EB/OL]. [2016-06-21]. https://myapm. apm. com/technical_documents/download/apm883408-product-brief1.

[11] Applied Micro. X-Gene2, X-Gene3[EB/OL]. [2016-06-21]. http://www. hotchips. org/wp-content/uploads/hc_ archives/hc26/HC26-11-day1-epub/HC26. 11-4-ARM-Servers-epub/HC26. 11. 430-X-Gene-Singh-AppMicro-HotChips-2014-v5. pdf.

[12] AMD 公司. Opteron A1100 处理器[EB/OL]. [2016-10-21]. http://www. tomsitpro. com/articles/amd-opteron-seattle-arm-soc,1-2105. html.

[13] The Linley Group. SPEC[EB/OL]. [2016-06-21]. http://www. cavium. com/pdfFiles/ThunderX_Rattles_Server_Market. pdf.

[14] Mars[EB/OL]. [2016-06-21]. http://www. expreview. com/42643. html.

[15] Ferdman M, Adileh A, Kocberber O. Clearing the Clouds: A Study of Emerging Scale-out Workloads on Modern Hardware[J]. ACM SIGARCH Computer Architecture News, 2012, 40(1): 37-48.

[16] CloudSuite 1.0[EB/OL]. [2016-06-21]. http://parsa. epfl. ch/cloudsuite/cloudsuite. html.

[17] Sui Xiufeng, Sun Tao, Li Tao. Understanding the Implication of Virtual Machine Managemant on Processor Microarchitecture Design[C]//Proceedings of IEEE International Symposium on Performance Analysis of Systems & Software. Washington D. C., USA: IEEE Press, 2013: 45-53.

[18] 胡　妍. 结合结构级和门级的多核处理器功耗评估方法[D]. 长沙:湖南大学,2013.

[19] 冯云贺. 基于 Simics 全系统仿真环境的嵌入式系统的研究与开发[D]. 北京:北京工业大学,2014.

编辑　索书志