

# 基于超级基站的集中式接入网络架构的物理层虚拟化方法<sup>①</sup>

萧 放<sup>②\*\*\*</sup> 周一青<sup>\*\*</sup> 林江南<sup>\*\*\*</sup> 黄 姗<sup>\*\*\*</sup>

(<sup>\*</sup> 移动计算与新型终端北京市重点实验室 北京 100080)

(<sup>\*\*</sup> 中国科学院计算技术研究所 北京 100190)

(<sup>\*\*\*</sup> 中国科学院大学 北京 100049)

**摘要** 针对集中式接入网络基带物理层计算资源异构性、基带算法模块间依赖性和物理层通信协议的实时性的存在而导致传统虚拟化技术不再适用的问题,在中科院计算所的超级基站的架构基础上,按照 TD-LTE 协议提出了一种基于单元级和系统级的二层资源调度策略,从而达到物理层计算资源虚拟化和实现虚拟基站的负载均衡和负载聚合的目的。仿真结果显示,这种实现物理层虚拟化的资源调度策略能够为集中式接入网络的物理层带来 1.6 的资源复用增益和 3.725 的功率复用增益。本研究能够为集中式接入网物理层计算资源池虚拟化的实现提供参考。

**关键词** 集中式接入网络, 超级基站, 物理层, 计算资源池, 虚拟化

## 0 引言

集中式接入网络是面向下一代移动通信网络的高能效需求而提出的新型接入网络架构。与传统的离散分布的基站架构相比,集中式接入网络除了能够大幅降低建站和维护成本以外,还能够对各种资源进行集中和池化,更加利于物理资源的统计复用,从而带来通信能效的提升<sup>[1-3]</sup>。在集中式接入网络架构中,传统基站的功能集合表述为运行在特定硬件资源上的虚拟基站(virtual base station, VBS)软件,虚拟基站占用的硬件资源通过虚拟化技术在物理资源池中进行高效的调度分配,对虚拟基站软件本身屏蔽其底层硬件资源的管理细节。因此,虚拟基站具有硬件资源可伸缩配置、映射资源可迁移的特点。针对集中式接入网络的以上概念,很多机构提出了自己的架构设计,其中包括 IBM 提出的无线云网络架构<sup>[4]</sup>,中移动、华为、英特尔等分别提出的云接入网络架构<sup>[5]</sup>,以及中科院计算所提出的超级

基站架构<sup>[6-8]</sup>。

基带计算资源池的虚拟化是实现整个集中式接入网架构的关键技术。虚拟化技术的核心是在上层应用和底层硬件资源之间增加一个资源管理层,对底层硬件资源进行高效的组织分配,从而增加整个系统的灵活性和可扩展性,它承载了对池化物理资源进行统计复用的功能。针对通用处理器(general purpose processor, GPP)组成基带计算资源池,研究人员研究并提出了基于云计算操作系统的虚拟机技术。目前研究结果显示,这种技术面临以下挑战<sup>[9-11]</sup>: (1) 通用处理器因为其同构特性,对通信基带中物理层的计算任务的计算效率非常低下,能效低于专用指令集处理器(application specified instruction-set processor, ASIP); (2) 在虚拟机上运行物理层软件,实时性无法保证,通常需要提供过量的硬件资源来防止物理层任务的超时。于是,研究人员提出了在超级基站中采用专用指令集芯片进行基带物理层计算、采用通用处理器进行基带协议层处理的

<sup>①</sup> 863 计划(2014AA01A708)和国家自然科学基金(61431001)资助项目。

<sup>②</sup> 男,1987 年生,博士生;研究方向:集中式无线接入网络,移动通信基带芯片;联系人,E-mail: xiaofang@ict.ac.cn  
(收稿日期:2016-01-05)

基带计算资源池架构<sup>[6-8]</sup>。在此架构下,专用指令集处理器在物理层任务运行效率上与通用处理器相比具有较大的优势,但不能够采用基于操作系统的虚拟机技术进行计算资源的虚拟化。因此,本文针对超级基站架构中物理层计算资源的虚拟化需求,首先对物理层计算任务特点进行了分析和建模,然后结合计算任务模型提出了一种保证其实时性的资源虚拟化方法,最后通过系统仿真进行了验证,得出了这种方法带来的资源复用增益和功率复用增益。

## 1 物理层资源池虚拟化方法简析

### 1.1 问题的建模与表述

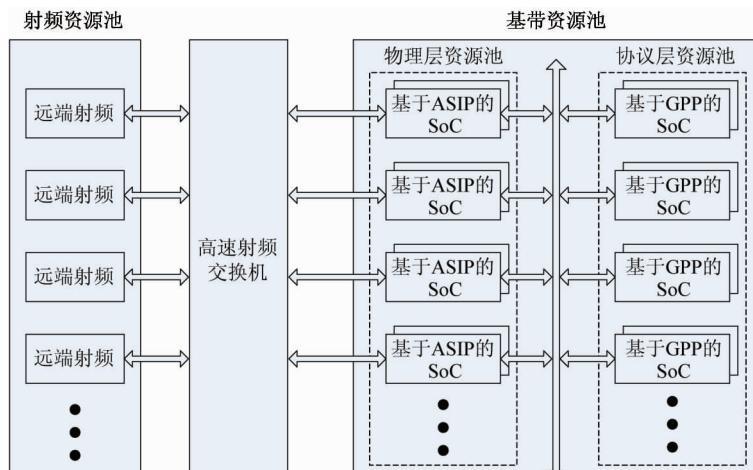


图1 超级基站资源池架构示意图

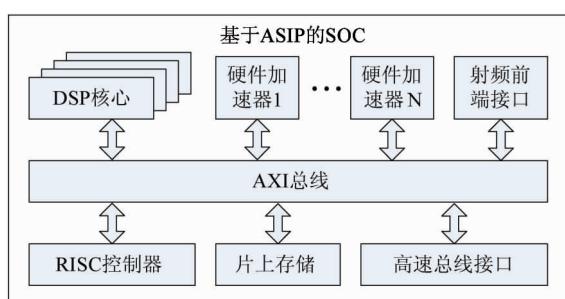


图2 物理层计算单元示意图

而运行在计算单元上的物理层虚拟基站(VBS)任务是一个复杂的包含上下行任务的无回路有向图(direct acyclic graph,DAG),具有因果性。每个算法模块是DAG中的一个节点,可简化表示为图3所

在超级基站架构中采用专用指令集处理器进行LTE(长期演进)物理层任务计算时,基带资源池可分为紧耦合的物理层资源池和协议层资源池<sup>[6]</sup>。其中,物理层资源池采用专用指令集处理器(ASIP)作为计算单元,协议层资源池采用通用处理器(GPP)作为计算单元,见图1。

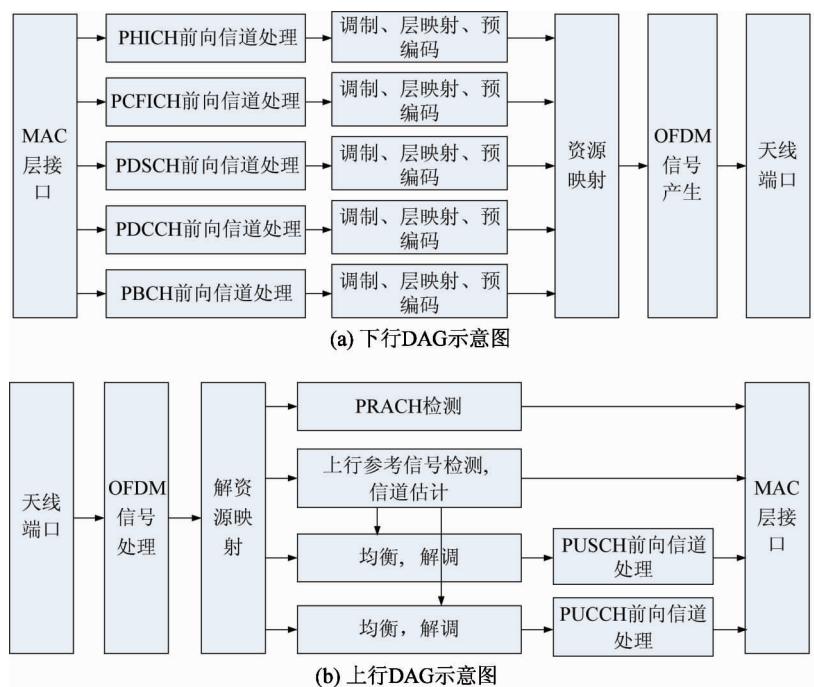
在物理资源池中,基于ASIP的片上系统(System on Chip,SoC)是基本的计算单元,内部包含矢量多核数字信号处理器(DSP)、物理层控制器、各种硬件加速器等多种硬件资源,见图2。

示。

所以,对VBS物理层的虚拟化可以等效为提供一种使所有VBS的DAG任务能够实时运行在物理层计算单元上的计算任务调度方法,并且得到较优的硬件资源利用效率,即在相同VBS负载的同时采用尽量少的计算单元。

### 1.2 现有研究背景

针对无回路有向图(DAG)任务调度,已经被证明为NP问题,不可在多项式时间内求解。基于DAG的任务调度方法主要分两类:(1)列表调度(list scheduling);(2)流水线调度(pipeline scheduling)。



PHICH:物理混合自动重传指示信道; PCFICH:物理控制格式指示信道; PDSCH:物理下行共享信道  
PDCCH:物理下行控制信道; PBCH:物理广播信道; MAC:媒体接入控制; OFDM:正交频分复用

图 3 分时长期演进(TD-LTE)物理层 DAG 简化示意图

列表调度<sup>[12,13]</sup>采用启发式算法 (heuristic algorithm) 计算 DAG 图中节点的优先级然后进行调度, 相对穷举法具有较优的复杂度。一般单处理核心资源上的调度复杂度为  $O(n \cdot e)$ , 其中  $n$  为节点数目,  $e$  连接节点的路径数目。

流水线调度<sup>[14,15]</sup>针对节点计算量固定的 DAG 子图, 通过生成算法使得 DAG 子图分为不同的阶段, 通过优化其流水间隔 (initial interval) 来提高系统的吞吐量。流水线调度在调度初始化时需要较高的运算量, 之后调度开销较小, 但仅适合每一个任务节点运算量固定的 DAG 图。

### 1.3 基于两层架构的资源调度方法

在上述已有研究中, 主要从针对通用 DAG 模型的调度来考虑进行调度策略的决策。在 LTE 物理层 DAG 调度时有很多特殊的情况:(1)1ms 实时任务运行截止周期;(2)每个节点计算量随时间波动;(3)物理层资源调度的目的是在相同负载下采用较少的硬件资源而不是传统的单纯缩小执行时间 (make span) 的需求。

针对 LTE 物理层 DAG 的特点带来的在 DAG 调度上的制约, 流水线调度显然并不适用。首先任务

到来的周期为固定的 1ms, 间隔不可变化。其次, 节点计算量变化较大, 已经划分好的流水线需要实时调整, 带来很大的调度开销。因此, 针对 LTE 物理层的 DAG 任务调度, 列表调度方法优于流水线调度方法。

但是列表调度并不能完全适用。首先, 很多列表调度的启发式算法考虑的是整个 DAG 的 make span 统计特性, 在面对节点计算量出现变化的时候, 会出现 make span 波动较大的情况, 不适合 1ms 固定截止时间的要求。其次, 列表调度的复杂度较高, 尤其是多个 DAG 任务在多个计算核心上进行列表调度, 其复杂度维度除了 DAG 的节点数, DAG 的边数以外, 还有 DAG 的个数  $m$ , 和异构计算核心的个数  $k$ 。

综合以上分析, LTE 物理层计算资源池中采用两层架构进行调度, 分别为单元级资源调度和系统级资源调度。其中, 在单元级采用列表调度进行, 限制单元级资源虚拟化的规模为 SoC 级。在单元级资源中 DAG 个数  $m$  和异构计算核心的个数  $k$  的规模都受到了限制, 便于合理控制调度算法的复杂度。而系统级资源调度的规模在资源池级, 通过与计算

单元的负载状态进行交互,对 DAG 到 SoC 的映射进行调度,从而实现对计算负载的均衡和聚合这一功能。

## 2 单元级计算资源的调度与负载表征

在 SoC 单元级,需要考虑的是 DAG 节点模型的建立、调度算法的选取设计和任务负载的表征。

### 2.1 DAG 节点模型的建立

DAG 节点模型的建立是整个单元级资源调度的基础。对 DAG 节点的调度除了需要建立 DAG 拓扑关系图,还需通过节点模型的建立预测每一个算法模块在相应计算资源上的执行时间。基于 TD-LTE 协议,算法模块节点的执行时间是可预测的,并且分为以下两类<sup>[16]</sup>:(1)静态节点,这类节点在虚拟基站(VBS)初始化以后,根据不同配置有固定的执行周期,例如 ODFM 信号产生和处理的算法模块;(2)动态节点,这类节点在 VBS 运行过程中,随着系统数据流量、占用资源块(resource block)数目和接入用户数目的改变有较大变化,例如信号编译码器等。

本文以 TD-LTE 接收端上行为例,对上行数据共享信道(physical uplink shared channel,PUSCH)算法模块进行基本的建模,建立节点执行时间和配置参数之间的关系,如表 1 所示。算法模块对应硬件资源的映射和执行周期统计基于中科院计算技术研究所研发的 DX-001-0.5E 芯片。在实际仿真环境中,还需要考虑相应的控制开销和数据传输开销的建模。实际建模为

$$P_K = P_{ck} + (P'_{ck} + P_{uk}) \times N_{uk} + P_{dk} \quad (1)$$

其中,  $P_K$  为算法模块 K 的总执行周期,  $P_{ck}$  为算法模块的固定调度开销,  $P'_{ck}$  为算法模块内资源参数有关的调度开销,  $P_{uk}$  为单位资源的数据处理开销,  $N_{uk}$  为资源参数的具体配置,  $P_{dk}$  为数据传输带来的开销。

### 2.2 调度算法设计

为了相对简化调度的复杂度,采用列表调度来对节点运行的顺序进行调度。调度对象为单元级计算资源上并行运行的多个 VBS,即多个 DAG(包括上行链路的 DAG 和下行链路的 DAG)。

表 1 PUSCH 算法模块建模示意分析

映射至 DSP	最大执行周期(符号)	建模情况
移除 CP	413	固定
解相位旋转 (Derot)	507	固定
FFT	2305	固定
解交叉(Dexplace)	416	与 RB 数正比
解资源映射 (RE-demap)	5280	与用户数正比
信道估计 CE	2005	与 RB 数正比
信道均衡 EQU	5851	与 RB 数正比
IDFT	9848	用户 IDFT 之和
(软解调) DeQAM	8024	与 PHY bit 正比
映射至 FEC 加速器	最大执行周期(子帧)	建模情况
解扰+解交织	13360	与 PHY bit 正比
解速率匹配	39078	与 PHY bit 正比
Turbo-Decoder	91182	与 TBS 正比
CRC	9452	与 TBS 正比

由于 TD-LTE 的 1ms 实时任务周期的要求,对任务的调度必须考虑 DAG 中执行时间最长的路径在截止时间内完成,所以采用基于最长完成时间路径具有最高优先级(highest level first with estimated time, HLFET)<sup>[17]</sup>的算法对单元级计算资源上的多个 DAG 进行调度,如图 4 所示。

- 1 针对计算单元上的每一个 DAG 生成节点关系拓扑表格
- 2 从后向前推算每一个节点的最长路径作为优先级等级
- 3 将已经可以被调度的节点加入到等待调度的队列
- 4 在等待调度队列中,依次调度优先级等级最高的节点
- 5 记录最后一个节点实际执行完成的时间

图 4 基于 HLFET 的调度算法

采用上述算法有两种实现方式,一种是静态调度,适用于对计算资源处理能力和资源间通信能力深度建模的情况,在实际执行调度之前可以通过基于模型的预调度完成对调度结果的评估。另一种是动态调度,由于对计算资源的建模存在一定的不确定性,可以在执行时动态的调度,在任务实际执行后对调度结果进行评估。本文中采用的是预调度的方式,在实际调度前完成预调度,得到相应的结果。

### 2.3 任务负载表征

相对于传统计算中以各个种类的硬件资源运行

时间占总时间的百分比表示该种资源的负载,单元级计算资源因为 DAG 任务节点间的依赖性需要采取一种新的负载表征方法,即计算单元上每个周期所有 DAG 的 make span 占规定时间(在 TD-LTE 协议中为 1ms)的百分比来表征单元级计算资源的负载。这种负载表征方法比较直观,但是会失去传统表征方式的线性特性。若  $\eta_a$ 、 $\eta_b$  分别为单元级计算资源  $a$  和  $b$  的负载,则他们的负载聚合到  $c$  上进行调度后,  $\eta_c < \eta_a + \eta_b$ 。

在这种表征方式的基础上,还是需要一种辅助的表征方法来确定在 make span 相近情况下的负载关系,本文提出了一种表示松弛度的方法,如下式所示:

$$\varphi = \frac{\sum_{j=1}^{N_{DAG}} (\sum_{i=1}^{M_{DAGj}} Node_{ij})}{Makespan \cdot K} \quad (2)$$

其中分子为运行在单元计算资源上的所有 DAG 的所有节点运行时间总和,分母为执行总时间与单元内计算资源数的积。在  $\eta$  相近时,  $\varphi$  的大小能够表示不同计算单元间的负载大小关系。

### 3 系统级计算资源调度

系统级计算资源调度建立在单元级计算资源的调度的基础上,根据单元级资源调度方法提供的对负载的表征方式,对分配给单元级资源的 DAG 负载进行决策。当负载发生变化导致该计算单元可能超负荷运行时,提前将该单元上映射的 DAG 迁移一部分至其他单元,完成负载均衡的功能;如果多个单元处于较低负载的状态,则将这多个低负载状态的单元上的 DAG 任务统一映射至一个或者少量几个计算单元上,并且关闭空闲的计算单元,完成负载聚合的功能。

由于每一个 DAG 的计算量会随着时间而有较大变化,基于全局的重映射方案会造成短时间内大量 DAG 的迁移,带来较大的开销,并且结合单元级预调度的全局寻优算法会带来高阶的计算开销。因此,采用一种步进迭代的思路,使得基于原有 DAG 映射的增量调整能够向理想的负载均衡和负载聚合

场景前进。最终,通过多次的迭代,使得 DAG 映射分布跟上其计算的变化,始终维持在较优的水平。

同时考虑到在实际情况下,DAG 计算量的变化并不是一个平滑的过程,除了具有统计变化规律以外,在小时间尺度内负载会有抖动。为了避免负载抖动对迁移判决带来的反复扰动,造成同一个单元来回进行负载均衡和负载聚合,判决的依据就不能够是单独的门限。因此,把单元的负载划分为了 4 个状态:休眠态、空闲态、正常态和警戒态。各状态间关系如图 5 所示。

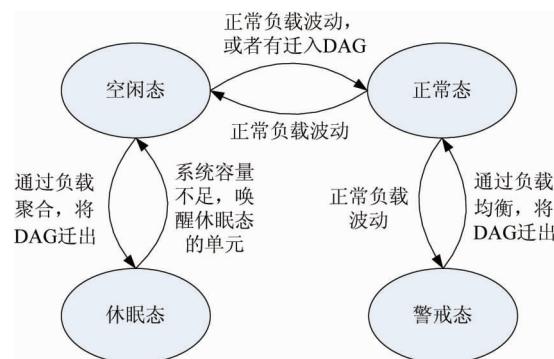


图 5 单元状态关系转化图

在此基础上,设计整个系统级调度的方案。在每个迭代周期中,首先解决超过警戒态门限的单元,通过负载均衡将 DAG 从警戒态单元迁出,直至变回正常态。其次,在没有进行负载均衡的迭代周期中,对多个处于空闲态的单元进行负载聚合,得到空载的单元使其进入休眠态。迭代周期内的整个过程如图 6 所示。

### 4 仿真实验

在单元级和系统级的资源调度模型的基础上,需要基于基站的流量负载模型进行系统仿真验证。依照中国移动提出的基站流量模型<sup>[5]</sup>,用户小区被主要划分为商务区和住宅区,其流量和时间的关系如图 7 所示。仿真参数见表 2。

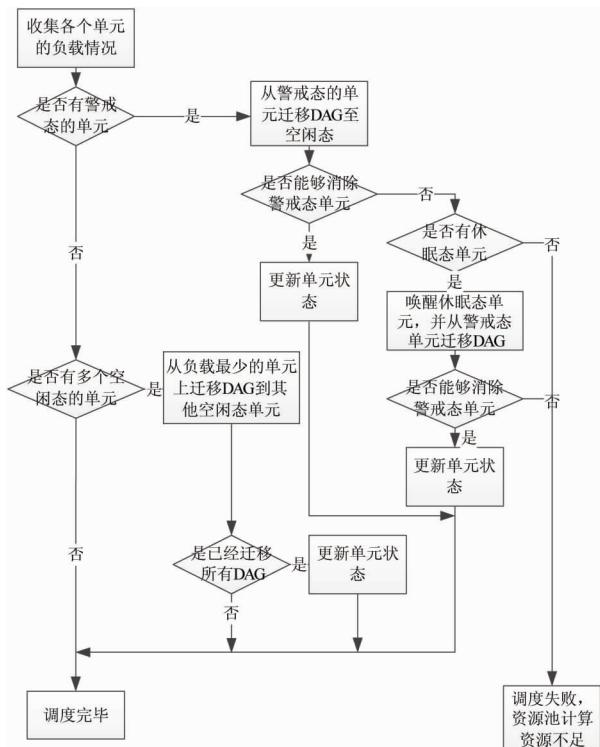


图 6 系统级资源调度流程图

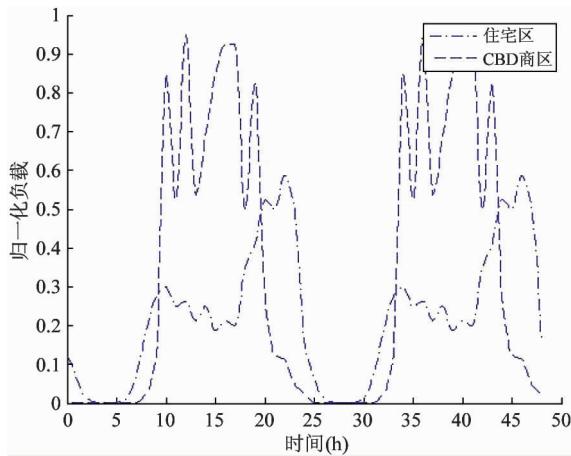


图 7 用户小区负载随时间变化示意图

表 2 仿真参数

参数	值
物理层资源池单元数	60
单元内异构计算资源数	DSP: X 7, ACC: X 4
负载变化分辨率	0.1h
单元级内部列表调度算法	HLFET
VBS 数量	住宅区 90, 商务区 90
VBS 流量变化期望	见图 7
VBS 流量 C. V	20%
单元状态转换门限设置	正常/警戒 0.85, 正常/空闲 0.7

在 VBS 负载随时间的变化模型符合上述统计规律的基础上,加上随机的负载波动模型,本文中为 20% 随机波动模型(归一化标准差为 0.2)。在计算单元数目为 60, 住宅区数目为 90、商务区数目为 90 的情况下,进行 48h 仿真得到物理层计算资源池中计算单元的活动规律,如图 8 所示。

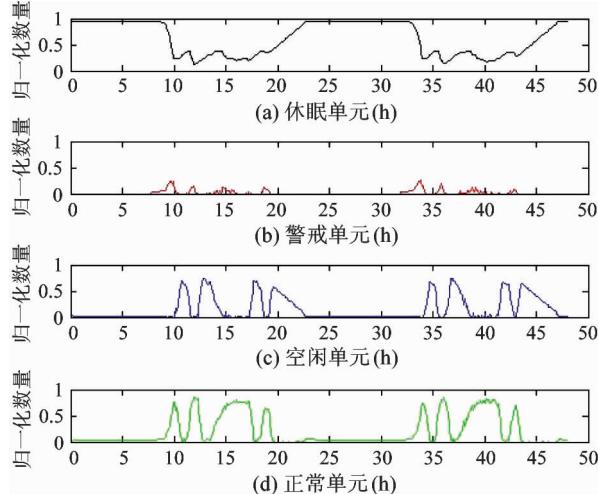


图 8 警戒态和休眠态单元统计图

从图中可以看出,在夜间低负载的情况下,大部分计算单元处于休眠关闭的状态。到了白天负载较高的状态时,随着警戒单元的增多,越来越多的计算单元从休眠状态中唤醒,以消除警戒态单元的出现。在负载较高时,大部分计算单元处于正常态,保证了每块硬件单元上的使用率。

针对不同数目的 VBS 所需求的计算单元个数,本文也进行了仿真,仿真结果如图 9 所示。

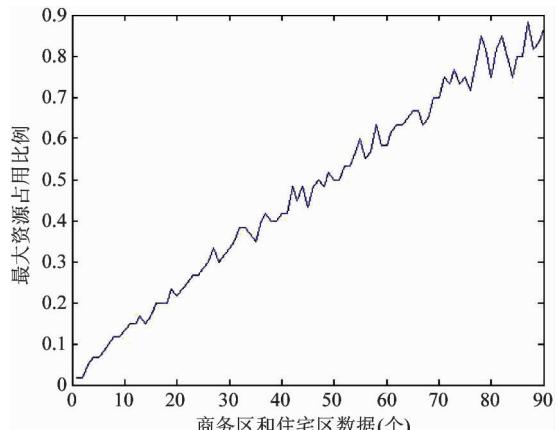


图 9 VBS 数目和最大计算单元数目关系图

从图 9 中可以看出,计算单元需求数目与 VBS 个数基本成线性关系,整个系统具有良好的可扩展性。其斜率的倒数反映了每个计算单元在统计复用情况的使用效率,数值与单个 VBS 固有的最大执行时间占总执行的比例有关。设  $R_o$  为单个 VBS 的 DAG 在负载最大时刻所占执行总时间的百分比,表示计算单元使用效率  $E_v$  和  $R_o$  的关系表达式如下:

$$E_v = \frac{N_{CBD} + N_{RES}}{N_{ASIP}} \cdot R_o \quad (3)$$

其中,  $N_{CBD}$  表示 CBD 商务小区的数目,  $N_{RES}$  表示住宅小区的数目,  $N_{ASIP}$  表示使用的 ASIP 计算单元数目。根据  $R_o$  的不同,分别得到了固定映射方式和本文中所述的虚拟化方法的计算单元效率,在计算单元数目为 60、商务区和住宅区的数目分别为 90 的情况下进行仿真,仿真结果如图 10 所示。

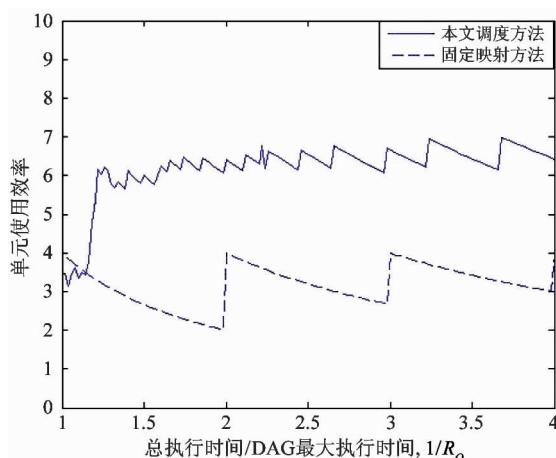


图 10 随  $1/R_o$  增长的单元效率变化情况

从图 10 中可以看出,当  $1/R_o$  稍大于 1.25 时,本文所述调度方法的单元使用效率  $E_v$  迅速上升到 6 到 7 之间,而固定映射方式单元使用效率最大只有 4。定义  $P_r$  为物理资源池架构的资源复用系数,  $P_r$  代表相对于固定映射方式,动态的调度算法所带来的资源使用效率的提升,如下:

$$P_r = \frac{E_v}{E_o} \quad (4)$$

在使用本文方法的情况下,资源复用系数  $P_r$  随着  $1/R_o$  的增大有小幅升高,平均约为 1.6 左右,也就是说采用此方法,同样数目的 VBS 负载情况下最大所需的计算单元数目为传统固定映射方式的  $1/P_r$ ,

也就是 62.5%。

除了资源复用系数带来的增益以外,功率复用带来的增益也是非常明显的。在本文的基于单元级和系统级的二层动态调度结构下,在空闲时间段完成了 VBS 的聚合,使得大部分计算单元处于休眠关闭的状态,带来的功率复用增益如下:

$$P_e = \frac{\text{Max}(N_{ACT}) \cdot T/\text{Resolution}}{\sum_{i=1}^{T/\text{Re}} N_{ACTi}} \cdot P_r \quad (5)$$

其中,  $\text{Max}(N_{ACT})$  表示在整个仿真周期内计算单元的最大激活数目,  $T/\text{Resolution}$  为仿真的采样点数,  $N_{ACTi}$  为每一个时刻实际激活的计算单元数目。在计算单元数目为 60,商务区和住宅区的数目分别为 90 的情况下进行仿真,  $P_r$  按照 1.6 来计算,  $P_e$  的统计值为 3.725。这个结果说明采用此方法,同样数目的 VBS 负载情况下,所消耗的平均功率为传统固定映射方式的  $1/P_e$ , 也就是 26.8%。

## 5 结论

本文重点研究了在以超级基站架构为基础的集中式接入网络物理层计算资源池虚拟化的实现途径。文中通过对通信系统计算任务的分析和对计算单元结构特点的分析,将物理层计算资源池虚拟化定义为一种将多个 VBS 的 DAG 动态运行在不同的计算单元上的资源调度方法,并且资源调度分成了单元级和系统级两个层面,每个层面设计了相应的资源调度方法完成相应功能。最后在仿真环境中进行了测试,验证了调度方法的功能性和可扩展性,最终通过仿真数据推导出了本文所述的资源调度方法能够带来 1.6 的资源复用增益和 3.725 的功率复用增益。

下一步的研究将会进一步考虑调度算法本身带来的开销和调度频率对系统流量产生的影响,综合更多的因素以得到一个更加接近实际情况的结果。

## 参考文献

- [1] Zhou Y Q, Liu H, Pan Z G, et al. Spectral and energy efficient two-stage cooperative multicast for LTE-A and beyond. *IEEE Wireless Magazine*, 2014, 21(2): 34-41
- [2] Zhou Y Q, Liu H, Pan Z G, et al. Two-stage cooperative

- multicast transmission with optimized power consumption and guaranteed coverage. *IEEE JSAC on SEED*, 2014, 32(2) : 274-284
- [ 3 ] Liu L, Zhou Y Q, Tian L, et al. CPC-based backward compatible network access for LTE cognitive radio cellular networks. *IEEE Communication Magazine*, 2015, 53 (7) : 93-99
- [ 4 ] Lin Y H, Shao L, Zhu Z B, et al. Wireless network cloud: Architecture and system requirements. *IBM Journal of Research and Development*, 2010, 54(1) : 1-12
- [ 5 ] China Mobile Research Institute. C-RAN The Road Towards Green RAN. <http://labs.chinamobile.com/cran>; China Mobile, 2011
- [ 6 ] Qian M L, Wang Y Y, Zhou Y Q, et al. A super base station based centralized network architecture for 5G mobile communication. *Digital Communications and Networks*, 2015, 54(2) : 152-159
- [ 7 ] 田霖,翟国伟,黄亮等. 基于集中式接入网络架构的异构无线网络及资源管理技术研究. 电信科学, 2013, 29(6):25-31
- [ 8 ] Zhai G W, Tian L, Zhou Y Q, et al. Load diversity based optimal processing resource allocation for super base stations in centralized radio access networks. *Science China, Information Sciences*, 2014, 57(4) : 1-12
- [ 9 ] Liu J C, Zhao T, Zhou S, et al. CONCERT: a cloud-based architecture for next-generation cellular systems. *IEEE Wireless Communications*, 2014, 21(6) : 14 - 22
- [ 10 ] Zhu Z B, Gupta P, Wang Q, et al. Virtual base station pool: towards a wireless network cloud for radio access networks. In: Proceedings of the 8th ACM International Conference on Computing Frontiers, Ischia, Italy, 2011. 1-10
- [ 11 ] Wubben D, Rost P, Bartelt J S, et al. Benefits and impact of cloud computing on 5G signal processing: Flexible centralization through cloud-RAN. *IEEE Signal Processing Magazine*, 2014, 31(6) : 35-44
- [ 12 ] Topcuoglu H, Hariri S, Wu M Y. Task scheduling algorithms for heterogeneous processors. In: Proceedings of the IPPS/SPDP Workshop on Heterogeneous Computing , San Juan, Puerto Rico, 1999. 3-14
- [ 13 ] Topcuoglu H, Hariri S, Wu M Y. Performance-effective and low-complexity task scheduling for heterogeneous computing, *IEEE Transactions*, 2002, 13(3) : 260-274
- [ 14 ] Manjunath K, Scott A M. Orchestrating the execution of stream programs on multicore platforms. In: Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation, Tucson, Arizona, 2008. 114-124
- [ 15 ] Carpenter P M, Ramírez A, Ayguadé E. Mapping stream programs onto heterogeneous multiprocessor systems. In: Proceedings of the Compilers, Architecture and Synthesis for Embedded Systems Grenoble, France, 2009. 57-66
- [ 16 ] Pelcat M, Nezan J F, Aridhi S. Adaptive multicore scheduling for the LTE uplink. In: Proceedings of the NASA/ESA Conference on Adaptive Hardware and Systems, Anaheim, USA, 2010. 36-43
- [ 17 ] Baxter J, Patel J H. The Last algorithm: A heuristic-based static task allocation algorithm. In: Proceedings of the International Conference on Parallel Processing, University Park, USA, 1989. 217-222

## A physical layer virtualization method for super base station based centralized radio access networks

Xiao Fang \* \* \* \* \*, Zhou Yiqing \* \* \*, Lin Jiangnan \* \* \* \* \*, Huang Shan \* \* \* \*

(\* Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing 100080)

(\*\* Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

(\*\*\* University of Chinese Academy of Sciences, Beijing 100049)

### Abstract

Considering that conventional virtualization techniques do not apply to a centralized radio access network due to the existence of the computing resource heterogeneity in baseband's physical layer, the dependency of baseband's algorithm modules, and the real-time ability of the communication protocol in the physical layer, a two-layer resource allocation strategy based on the unit level and the system level was presented on the basis of the super base station architecture given by the Institute of Computing Technology of Chinese Adacemy of Sciences according to the TD ( time division)-LTE ( long term evolution) protocol to achieve the computing resource virtualization in the physical layer and realize the load balance and load consolidation of virtual base stations. The simulation illustrated that the proposed method worked well, bringing a resource multiplexing gain of 1.6 and an energy multiplexing gain of 3.725 to the physical layer of the centralized radio access network. This study is of referential significance to realization of the virtualization of computing resource pools of centralized radio access networks.

**Key words:** centralized radio access network, super base station, physical layer, computing resource pool, virtualization