



嵌入式云计算与视频大数据 基于TI嵌入式处理器

西安电子科技大学
张亮



报告内容

1 视频大数据背景

2 大数据、云计算及嵌入式系统

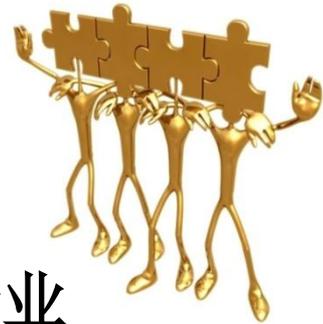
3 TI Keystone Cloud平台

4 如何用嵌入式构建云计算平台

5 基于TI多核的并行模型



视频大数据背景



- ◆ 智慧地球，智慧平安城市
- ◆ 城市海量视频监控摄像头的安装，覆盖了各个行业，尤其是城市交通，高速公路，小区，各种重要场所，等等
- ◆ 从2010年到2014年，中国的视频监控摄像头数目将以每年20%的速度增加，据统计，仅仅2010年，中国新增摄像头数目为超过1000万个摄像头



视频大数据背景

- ◆ 世界范围内，英国的摄像头比例为1:15
- ◆ 我国的摄像头数目将持续增加，重庆计划在2014年前城市的摄像头数目将达到50万台。据统计北京目前摄像头数目达到60万台。
- ◆ 在这些官方统计数据中，还不包括民用的各种摄像头



视频大数据背景

如何进行海量视频数据的传输及存储？

如何从海量视频数据中快速进行有效视频数据的提取？

海量非结构化数据？？？

如何统一化各种不同类型的摄像头视频数据？

分析问题、传输问题、存储问题



视频大数据背景

1、如何利用通用的大数据处理方法对非结构化海量异构视频数据进行处理？（比如视频分析、视频转码、实时处理、离线处理等）

2、对于结构化的视频数据如何利用进行存储？比如HDFS，HBASE等方式？

因此，从纯粹的视频结构化大数据而言，利用传统的大数据处理方法存在很多可优化、可研究的点



报告内容

1 大数据背景

2 云计算与嵌入式系统

3 TI Keystone Cloud平台

4 如何用嵌入式构建云计算平台

5 基于TI多核的并行模型



必要性

- 1: 大的计算资源
- 2: 功耗消耗
- 3: 智慧城市和智慧海洋
- 4: WSN (Ad-hoc)
- 5: 嵌入式云计算平台已经开始了……





云计算与嵌入式系统

很多处理器厂商已经设计了类型的嵌入式处理器，其特点在于体积小，功耗低，针对性强



ELECTRONICS

Exynos Multi-core processors



TEXAS INSTRUMENTS

Multi Arm processors + Multi DSP processors



Multi-core signal processors



Zynq (Arm and FPGA processor)

Tilera many cores



云计算与嵌入式

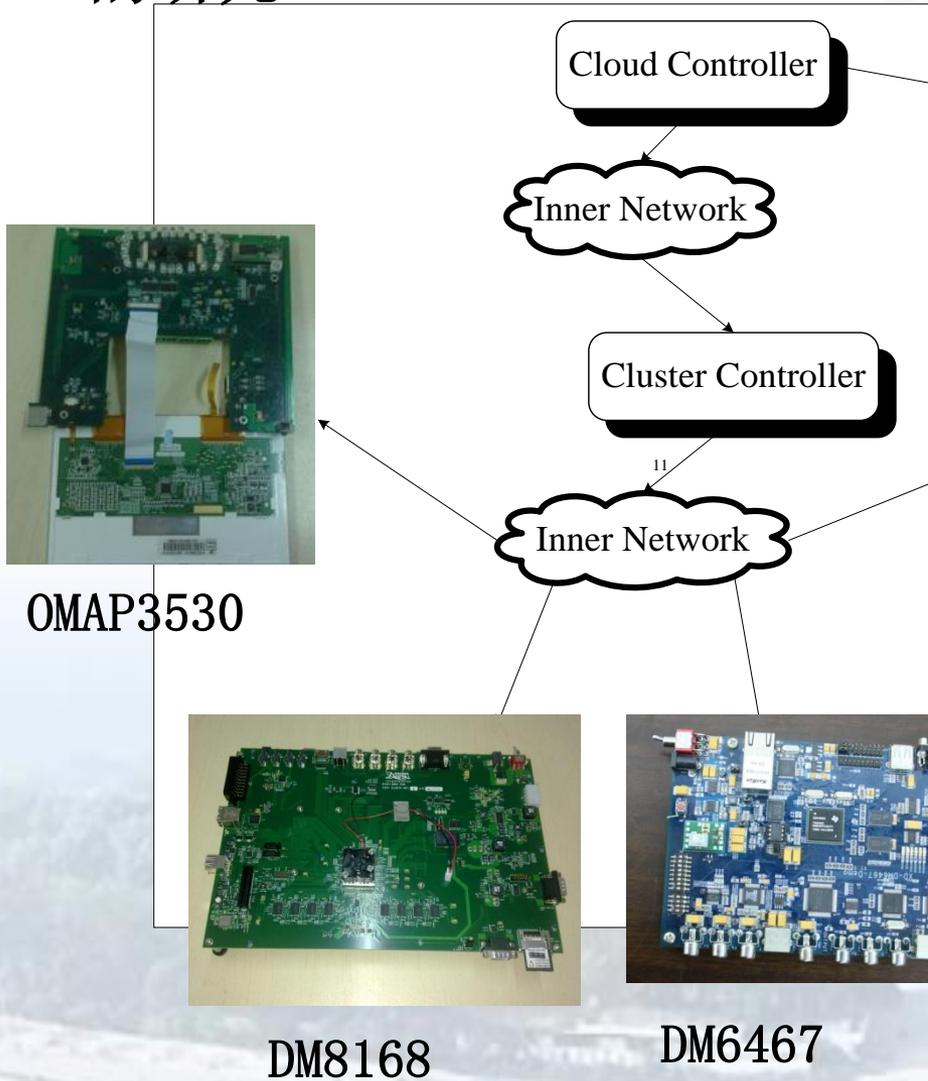
主要研究：

- 1、以多核、众核嵌入式处理器为节点的云计算系统技术的研究（OMAP3530/3730，DM8168，DM6467/6446，KeystoneI/II 等）
- 2、研究适合于嵌入式多核处理器及嵌入式云计算平台的轻量级并行编程模型
- 3、云计算平台下，利用嵌入式多核众核平台进行并行视频分析处理技术



云计算与嵌入式

1、以多核、众核嵌入式处理器的研究。

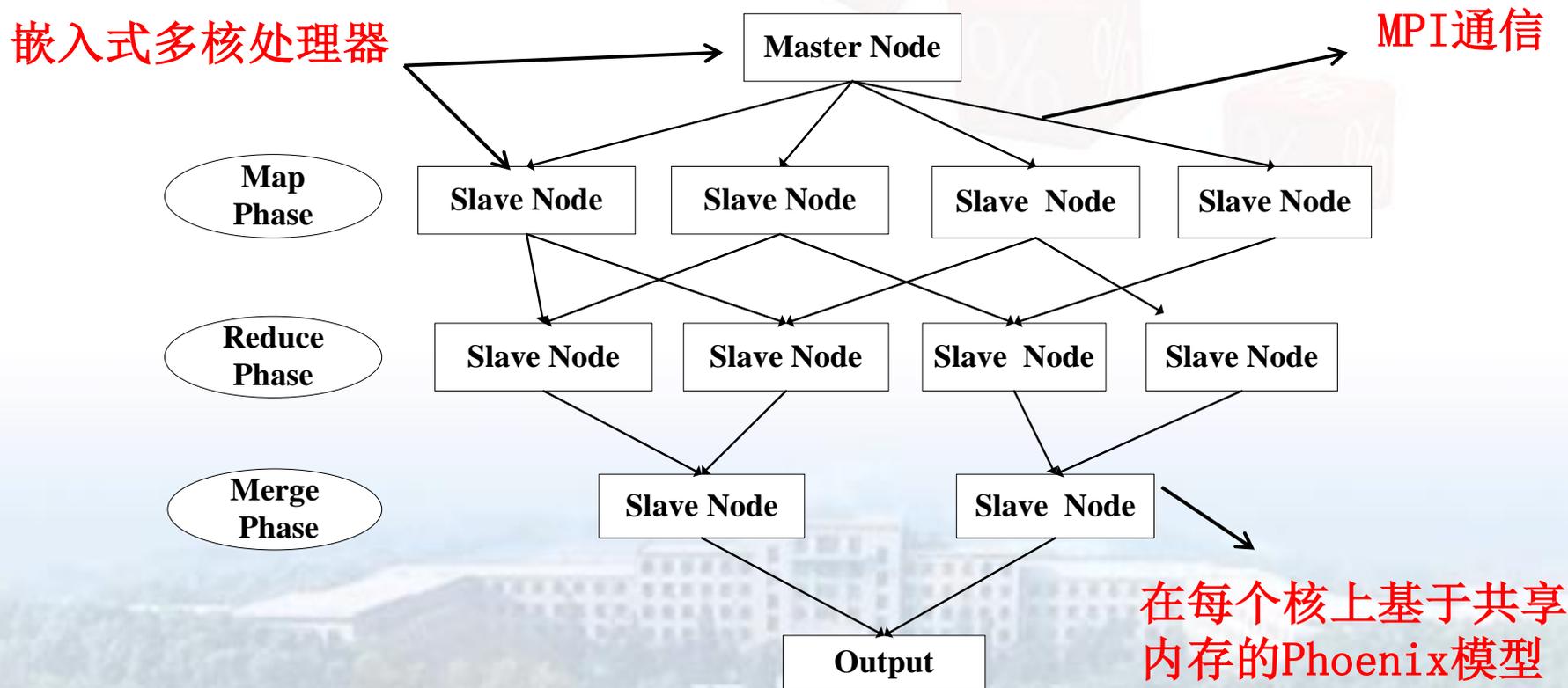


ss
e



云计算与嵌入式

2、研究基于MapReduce+MPI, MPI+Phoenix等适合于嵌入式多核处理器及嵌入式云计算平台的并行编程模型





Message Passing Interface (MPI)

- 一个跨语言的通讯协议，用于**编写并行计算机**。支持**点对点**和**广播**。
- 一个**信息传递应用程序接口**，包括**协议和语义**说明，指明其如何在各种实现中发挥其特性。
- 目标是高性能，大规模性，和可移植性。
- MPI在今天仍为**高性能计算的主要模型**。
- 大多数信息传递接口的实现为库，亦不需要编译器支持。
- 从语法上说，它遵守所有对库函数/过程的调用规则，和一般的函数/过程没有什么区别。



Phoenix并行模型

- 适用于基于单芯片多CPU的轻量级MapReduce框架
- 基于C和C++语言实现的框架
- 基于共享内存的MapReduce实现
- 适用于单节点，不适用于多节点
- 可以应用于多核以及众核平台



Phoenix并行模型

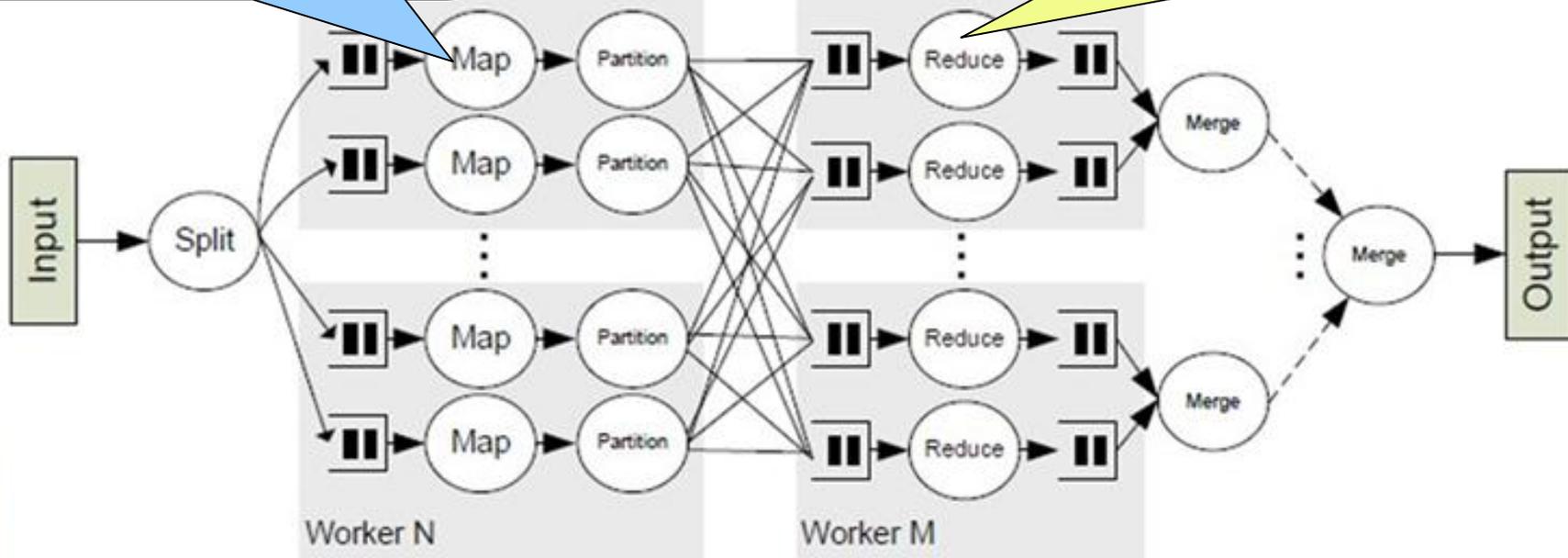
- 一个基于共享内存系统的MapReduce模型的实现。它包括一个编程API和一个高效的运行时系统。
- Phoenix运行时自动管理线程创建、动态任务调度、数据划分和跨处理器节点的错误容忍。
- Phoenix使用线程来大量生产Map和Reduce任务。使用共享内存, 无需大量数据拷贝的通信。
- 运行时在可用的处理器上动态调度任务, 从而实现负载平衡和任务吞吐量最大化。
- 运行时在任务执行时自动从暂时和长久的错误中恢复, 主要通过重复或重新布置任务实现, 并将他们的输出和其余的计算合并起来。



Phoenix并

Map函数应用于输入数据之上，并产生一系列<key, value>对的中间值。

Reduce函数应用于所有相同key的中间值。通常它执行一些合并操作并产生零对或多对输出。最后输出对按其key值进行排序。





报告内容

1 大数据背景

2 大数据、云计算及嵌入式系统

3 TI Keystone Cloud平台

4 如何用嵌入式构建云计算平台

5 基于TI多核的并行模型



TI Keystone Cloud平台

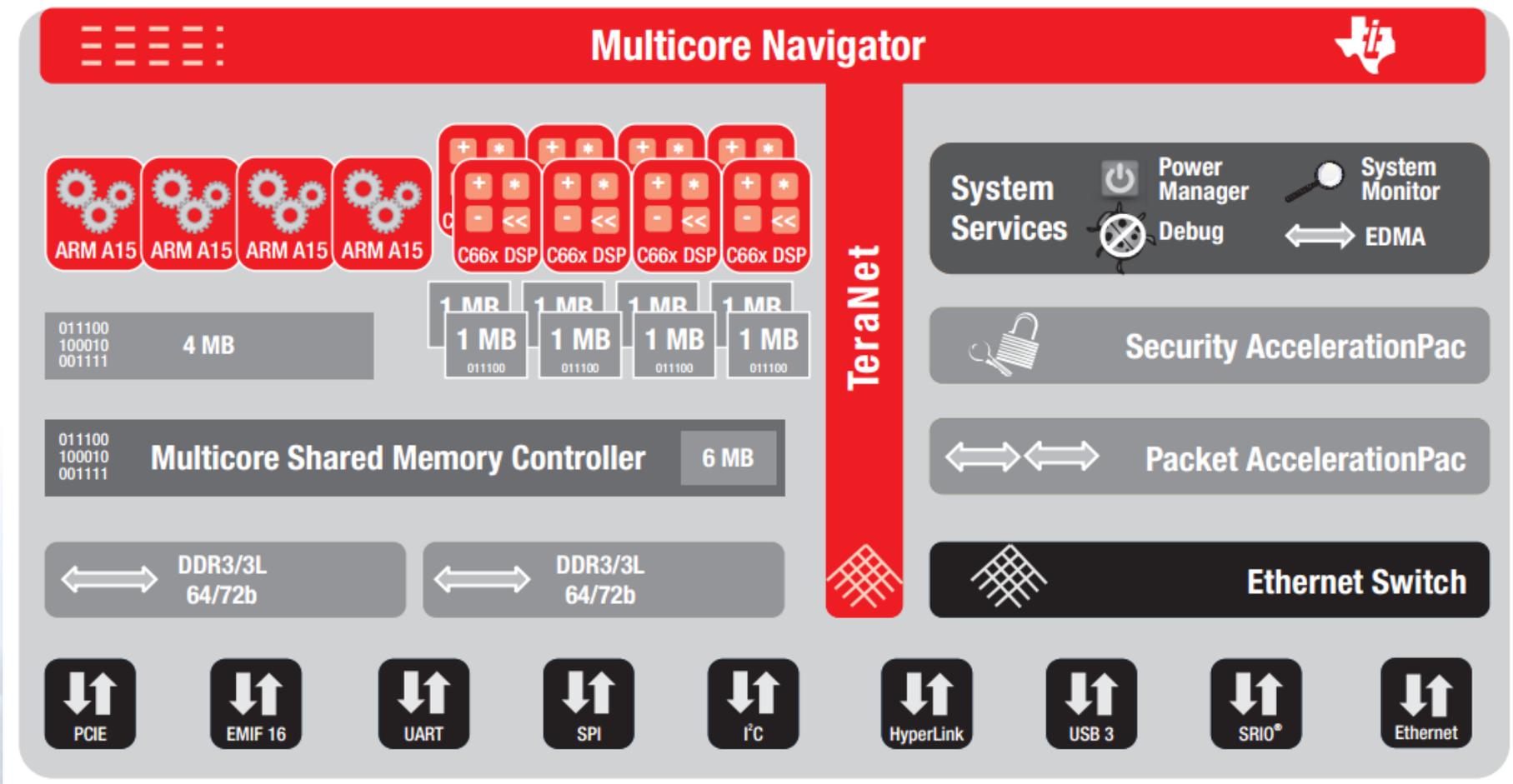
- TI推出新一代 KeyStoneII架构，为未来异构“绿色环保”无线网络实现高度集成的可扩展型低功耗解决方案。
- KeyStoneII集成四个ARM Cortex-A15内核的群集，与传统精简指令集计算（RISC）内核相比，可在提供高性能的同时功耗锐降 50%。这是未来绿色环保网络基础设施设备的一大要素。





TI Keystone Cloud平台

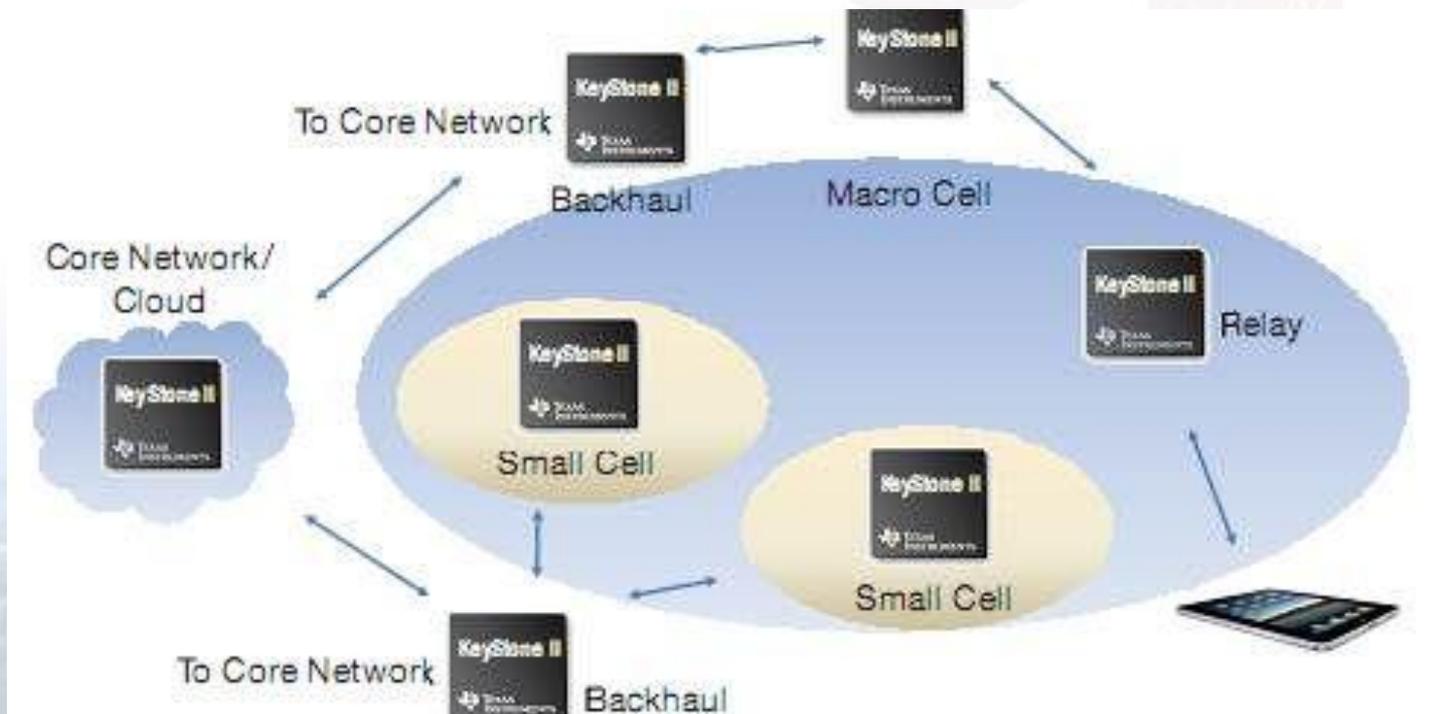
TI 基于KeyStone II的66AK2H12 SoC





TI Keystone Cloud平台

- KeyStoneII架构中DSP内核、ARM内核、AccelerationPacs以及 I/O 的高灵活配置，可为创建 SoC 器件提供丰富的功能。
- KeyStoneII可提供高性能与低成本，充分满足各种计算需求。





报告内容

1 大数据背景

2 大数据、云计算及嵌入式系统

3 TI Keystone Cloud平台

4 如何用嵌入式构建云计算平台

5 基于TI多核的并行模型



构建嵌入式云计算平台

两种方法:

1、在嵌入式平台上运行Hadoop集群

该方法主要是针对特定类型的嵌入式平台，移植Hadoop运行所必须的工具软件，以及针对Hadoop的部分在嵌入式平台上运行有问题的代码，进行一定程度的修改。

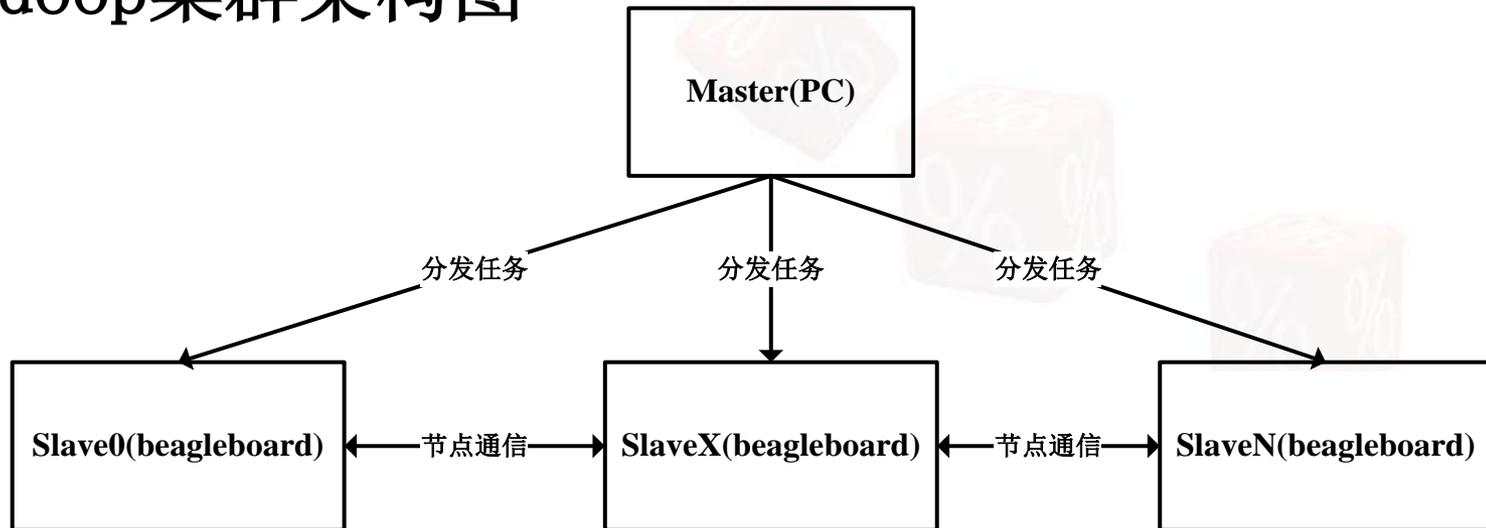
2、重新设计轻量级的MapReduce框架

比如Phoenix、MPI-MapReduce。其中每个都有其各自的特点，Phoenix是基于共享内存的一个实现，仅适用于单节点。MPI-MapReduce是基于MPI来进行通信的，适用于多个节点之间。



构建嵌入式云计算平台

Hadoop集群架构图

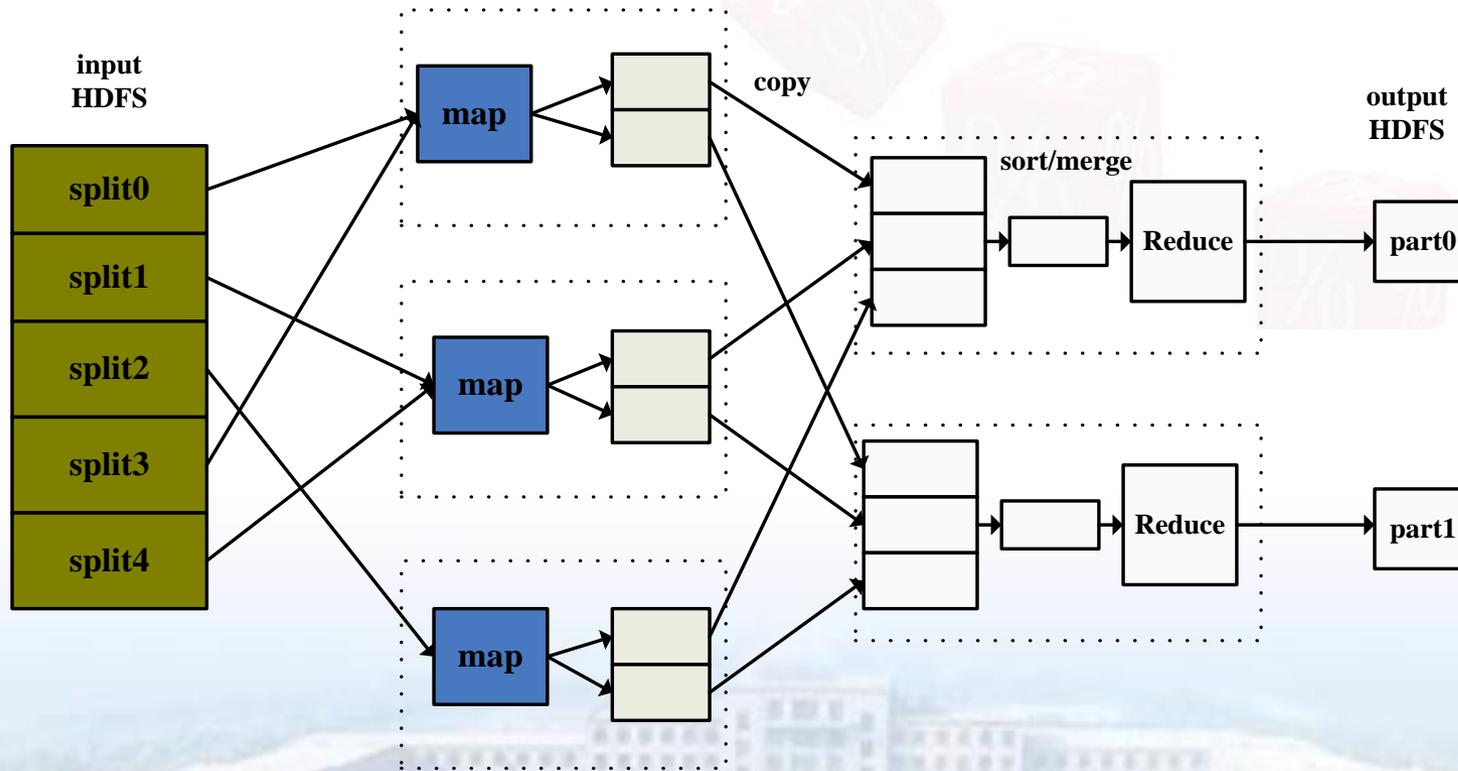


- Hadoop集群的Master节点用一台PC充当
- 若干BeagleBoard板卡等嵌入式处理器作为集群的Slave节点
- Master节点分发任务给各个Slave节点，Slave节点上运行Hadoop任务的map工作，然后将生成的key-value在节点之间分发，将key Hash发送到对应节点上进行reduce处理，最终将处理结果存储到HDFS中。



构建嵌入式云计算平台

Hadoop工作流程图



Master节点负责input输入数据的分割，然后将其分发给对应的map处理任务，各个节点的map处理之后会根据keyvalue进行reduce处理，然后分别将处理结果保存到hdfs，map和reduce任务分别运行在不同的slave节点中。



构建嵌入式云计算平台

➤Hadoop任务中DSP资源的使用

- 1、TI Codec Engine框架中，ARM调用DSP端C语言代码的封装。
- 2、jni接口对调用DSP的C语言函数接口进行封装。
- 3、Hadoop任务的map或reduce函数中调用相应jni接口来对DSP进行调用。

➤Hadoop任务中FPGA资源的使用

- 1、Zynq平台下同样能够用C语言接口调用FPGA逻辑。
- 2、jni接口对调用FPGA的C语言接口进行封装。
- 3、Hadoop任务的map或reduce函数中调用相应jni接口来对FPGA进行调用。



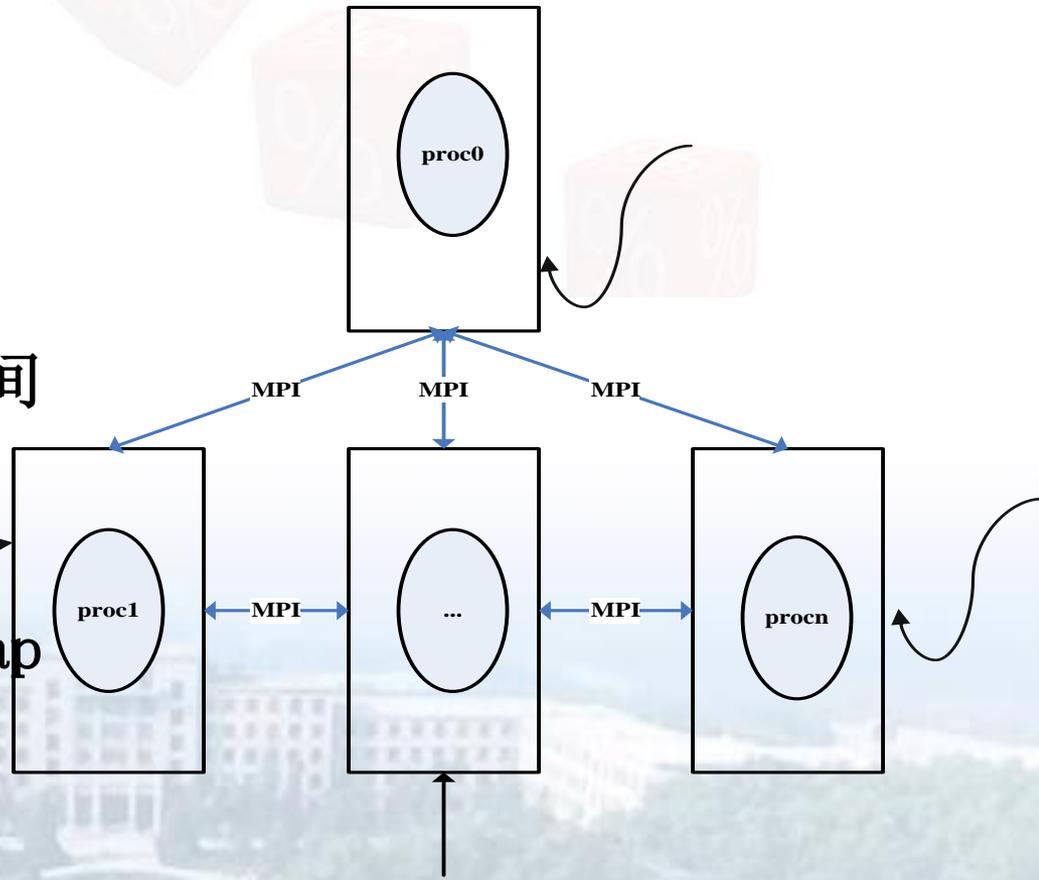
构建嵌入式云计算平台

➤ MPI-MapReduce框架

1、基于多节点的MapReduce
C++框架实现。

2、基于MPI库来进行节点之间的
通信和数据交换。

3、每个节点上执行相应的map
任务和reduce任务。





构建嵌入式云计算平台

➤ Phoenix + MPI-MapReduce 结合

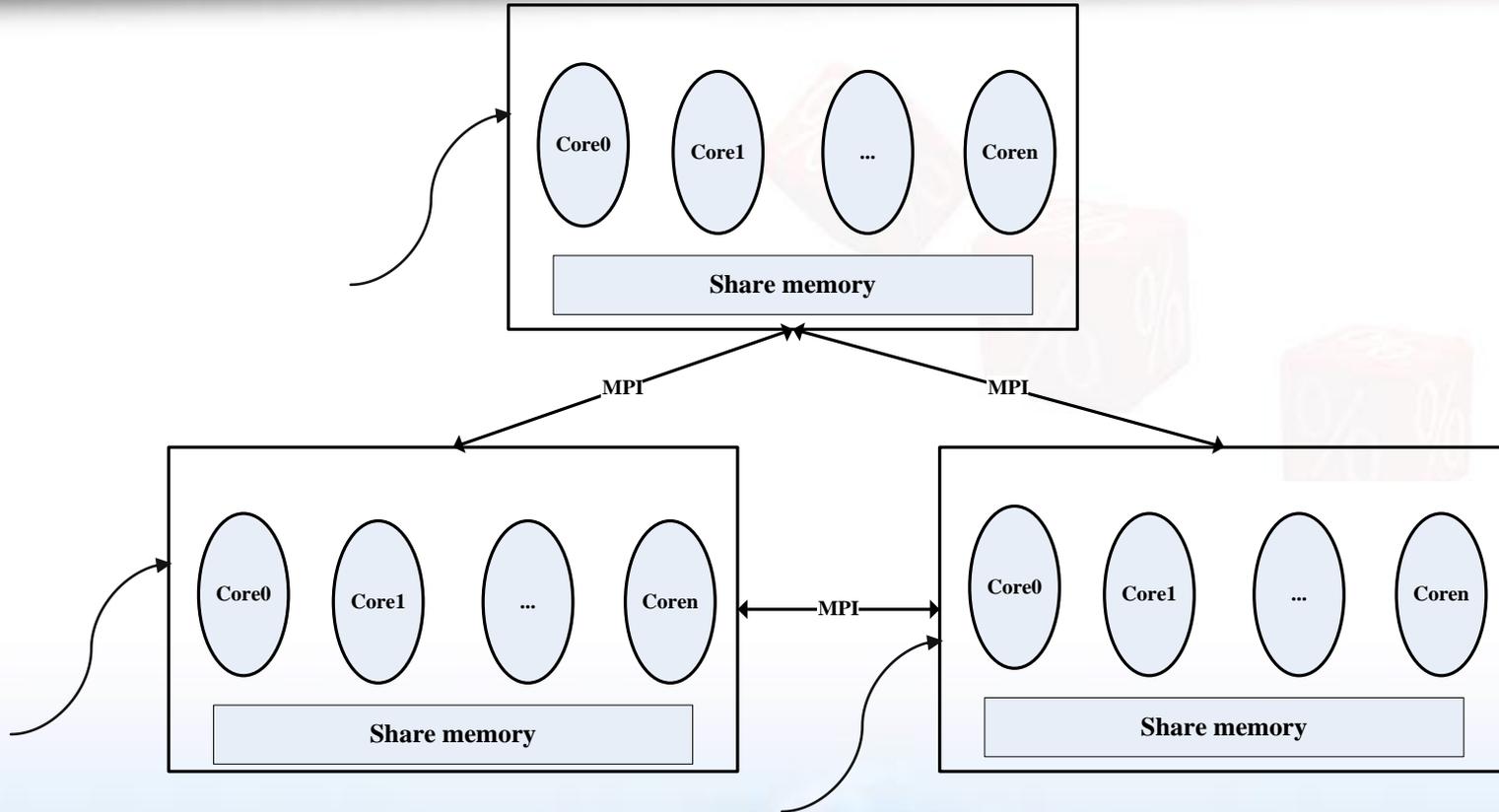
由于两者的实现各自有其优点，并且使用范围也不同，因此我们将两个框架的思想进行了整合，实现基于多节点的MapReduce框架，并且在各个节点内部充分利用Phoenix的思想来充分利用各个核心。

➤ 该框架的优点

- 1、由于两个框架基于C、C++实现，在嵌入式平台上能够更高效的运行。
- 2、该框架能够在利用多个节点进行大数据量的计算，同时能够充分利用嵌入式平台下每个节点特有的硬件资源。



构建嵌入式云计算平台



Node 0节点利用MPI分发任务，各个节点接收到任务后充分利用各个核心和其它硬件资源进行处理，在本节点内部先进行一次reduce，进而减小节点之间交换的数据量，提高传输速度。节点内部做完之后在节点之间通过MPI进行分发，使相同key映射到同一个node，然后在node内部在一次进行map-reduce，加速执行过程。



报告内容

1 大数据背景

2 大数据、云计算及嵌入式系统

3 TI Keystone Cloud平台

4 如何用嵌入式构建云计算平台

5 基于TI多核的并行模型



基于TI多核的并行 模型

- 如何将TI的多核处理器应用到云计算平台中？
- 怎样利用多核处理器提供的通信方式、处理器的体系架构进行多核节点上的通用并行模型，该并行模型要考虑到多个节点之间的通用性



基于TI多核的并行 模型

- 6678开发板
- CCS v5.3

DC 12V

USB mini-B

Gigabit Ethernet





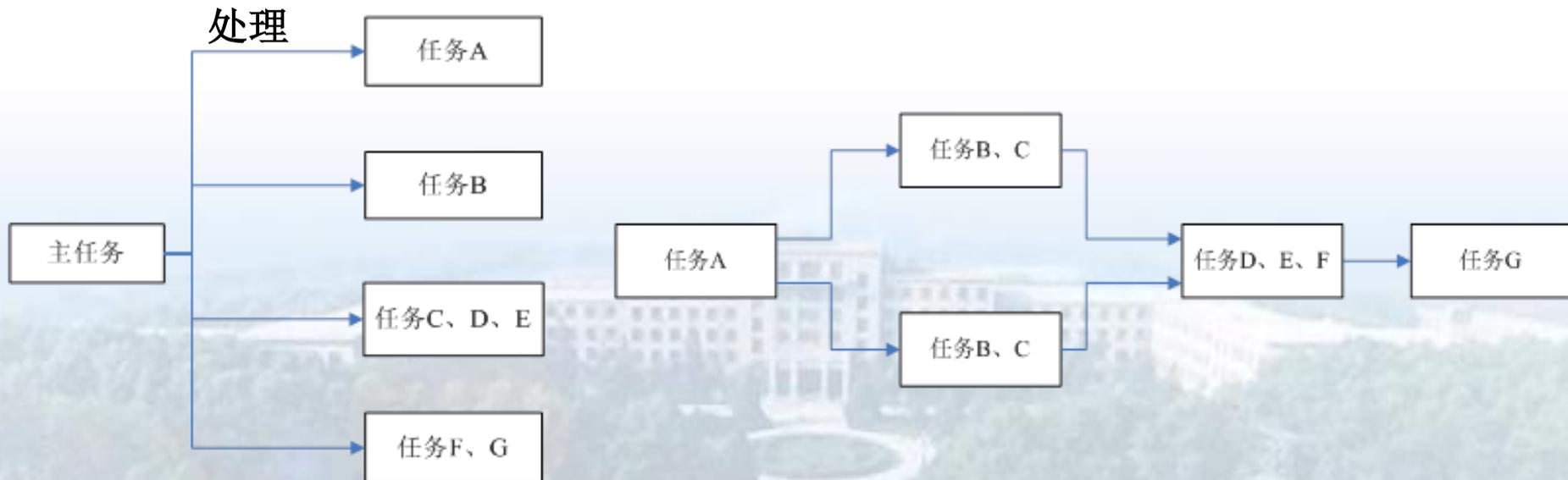
基于TI多核的并行 模型

● 主/从模型

- 主/从模型表示集中控制，分布执行
- 在这种模式下，OS主导调度

● 数据流模型

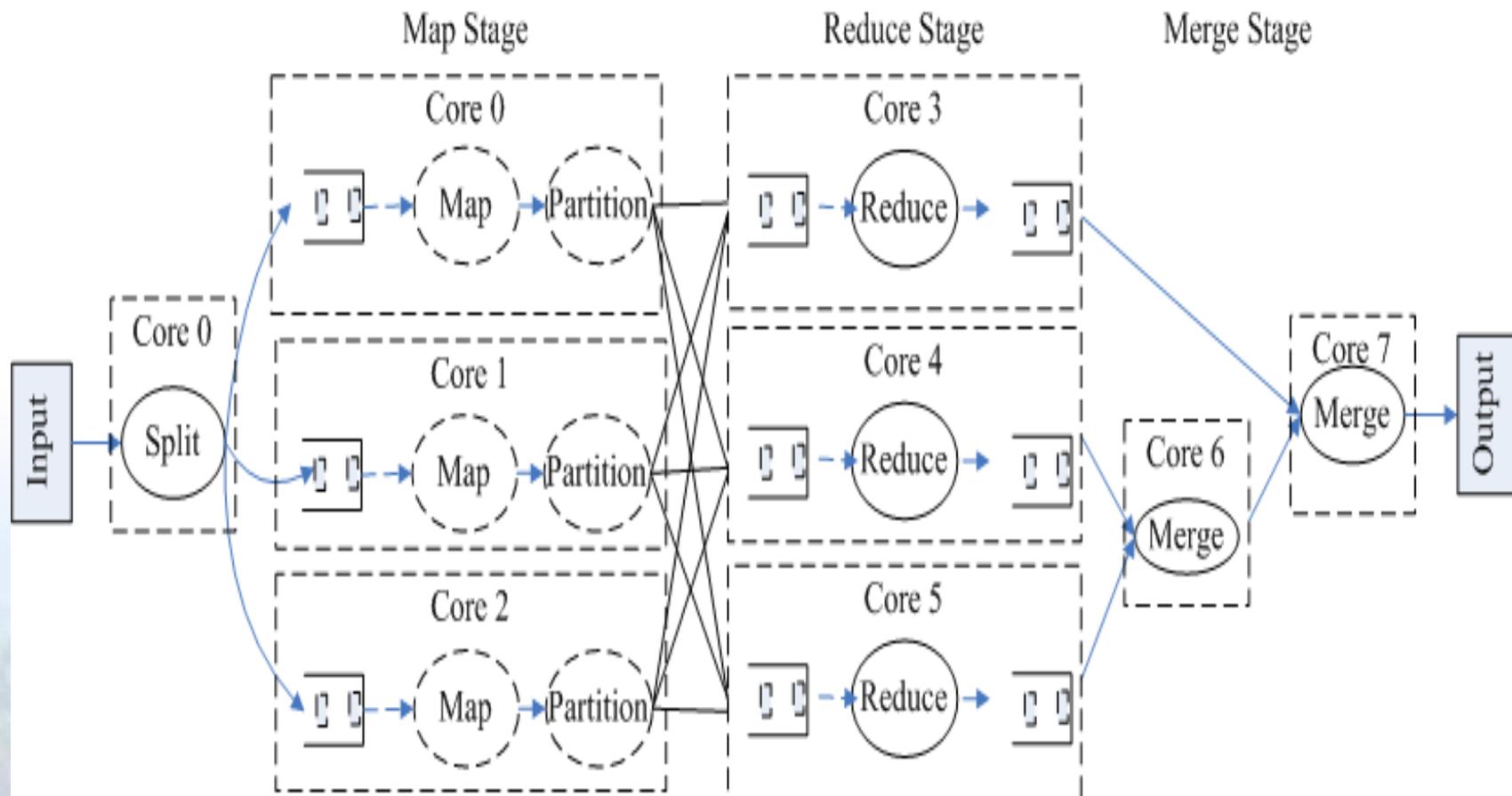
- 数据流模型表示分布式控制和执行
- 每个核使用各种算法处理一块数据，然后传递数据到下一个核做下一步





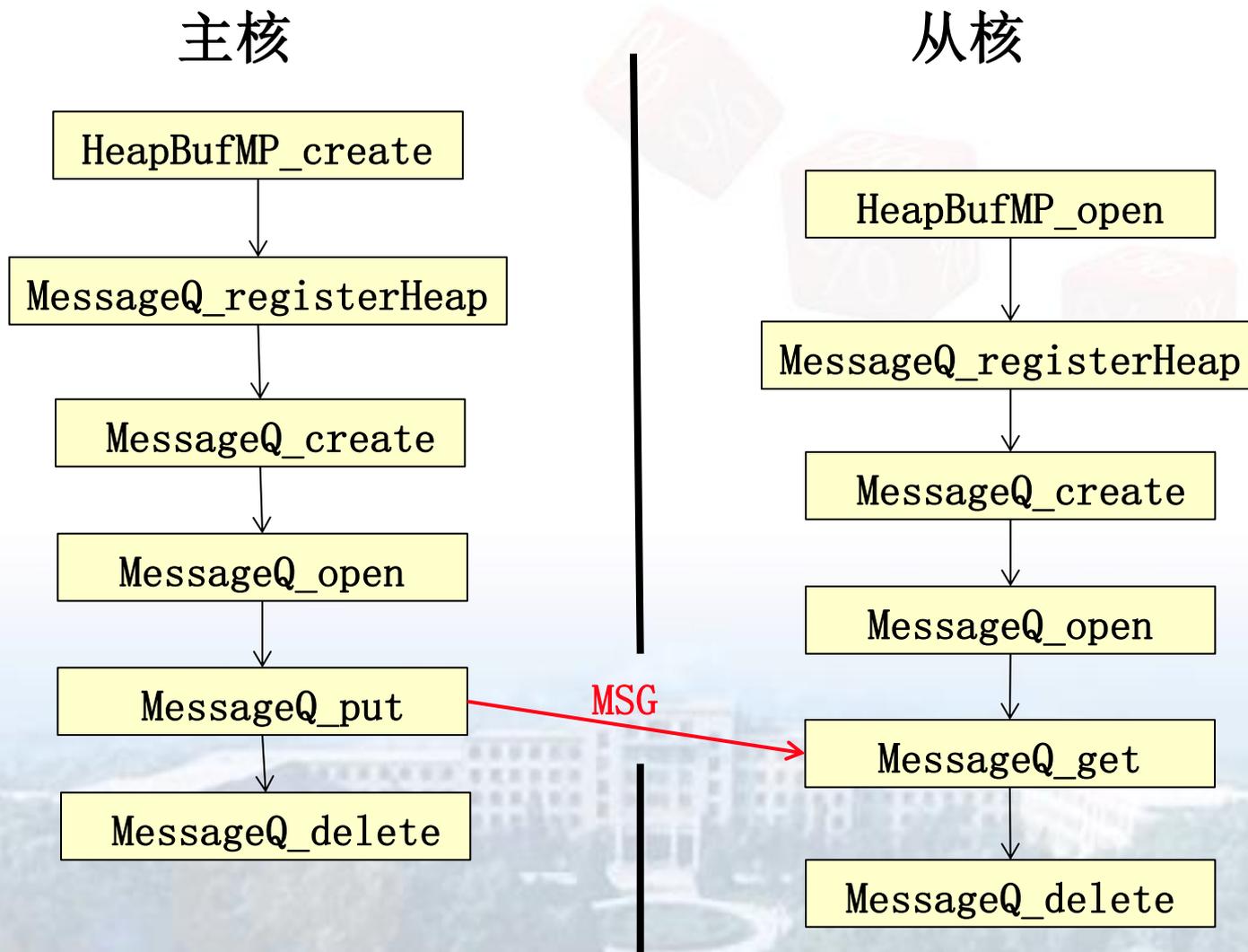
基于TI多核的并行模型

基于TI 6678的Phoenix架构





基于TI多核的并行 模型





总结

- 1、嵌入式云计算平台中资源的虚拟化问题
- 2、适合于TI异构多核处理器的并行计算模型
- 3、云计算平台中适合于嵌入式处理器的数据存储方法研究
- 4、云计算平台中TI嵌入式多核异构处理器的动态调试方法，动态资源监控与管理方法……



谢谢

liangzhang@xidian.edu.cn