(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 103036817 A
(43)申请公布日 2013.04.10

(21)申请号 201210544600.5

(22)申请日 2012.12.14

(71)申请人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 彭华 萧晓晖

(74)专利代理机构 北京中博世达专利商标代理有限公司 11274
代理人 申健

(51)Int.Cl.
H04L 12/931(2013.01)
H04L 12/721(2013.01)

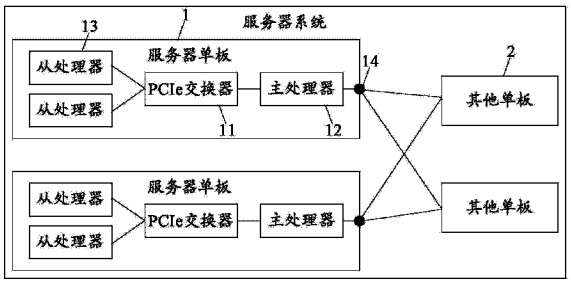
权利要求书 4 页 说明书 15 页 附图 5 页

(54)发明名称

一种服务器单板、服务器单板实现方法及主处理器

(57)摘要

本发明实施例提供一种服务器单板及其实现方法，涉及通信领域。该服务器单板可以降低服务器内部的延时，提高网络服务质量。包括：PCIe交换器，主处理器以及多个从处理器，PCIe交换器包括一个上行端口以及多个下行端口；主处理器与PCIe交换器的上行端口相连，多个从处理器分别与PCIe交换器的下行端口相连；主处理器用于通过网口接收其他单板发送的第一以太网数据包，将第一以太网数据包转成第一PCIe数据包，通过PCIe交换器将第一PCIe数据包转发给一个或多个从处理器；还用于接收一个或多个从处理器发送的第二PCIe数据包，将第二PCIe数据包转成第二以太网数据包，通过网口向其他单板转发第二以太网数据包。



CN 103036817 A

1. 一种服务器单板，应用于服务器系统，所述服务器系统还包括与所述服务器单板连接的至少一个其他单板，其特征在于，所述服务器单板包括：

PCIe 交换器，主处理器以及多个从处理器，所述 PCIe 交换器包括一个上行端口以及多个下行端口；

所述主处理器与所述 PCIe 交换器的上行端口相连，所述多个从处理器分别与所述 PCIe 交换器的下行端口相连；

所述主处理器用于通过网口接收所述其他单板发送的第一以太网数据包，将所述第一以太网数据包转成第一 PCIe 数据包，通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器；

所述主处理器还用于通过所述 PCIe 交换器接收所述从处理器发送的第二 PCIe 数据包，将所述第二 PCIe 数据包转成第二以太网数据包，通过所述网口向所述其他单板转发所述第二以太网数据包。

2. 根据权利要求 1 所述的服务器单板，其特征在于：

所述主处理器还用于，通过所述 PCIe 交换器接收其中一个所述从处理器发送的第三 PCIe 数据包，并通过所述 PCIe 交换器将所述第三 PCIe 数据包转发给另一个所述从处理器。

3. 根据权利要求 1 或 2 任一所述的服务器单板，其特征在于，所述主处理器用于将所述第二 PCIe 数据包转成第二以太网数据包具体为：

若所述第二 PCIe 数据包的净荷包括完整的以太网数据包，所述主处理器用于直接提取所述以太网数据包以得到第二以太网数据包；或，

若所述第二 PCIe 数据包的净荷包括以太网数据包的净荷，所述主处理器用于提取所述以太网数据包的净荷进行封装以得到第二以太网数据包。

4. 根据权利要求 1-3 任一所述的服务器单板，其特征在于，所述主处理器还用于，生成路径配置表，所述路径配置表包括从处理器标识以及与所述从处理器标识对应的路由信息；

所述主处理器用于将所述第一以太网数据包转成第一 PCIe 数据包，通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器具体为：

所述主处理器用于：

将所述第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者所述主处理器用于提取所述第一以太网数据包中的净荷，将所述第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中；根据所述路径配置表及所述从处理器的从处理器标识获取与所述从处理器对应的第一路由信息，将所述第一路由信息添加到所述第一 PCIe 数据包的头部；通过所述 PCIe 交换器的上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器根据所述第一 PCIe 数据包头部中的所述第一路由信息将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器。

5. 根据权利要求 1-4 任一所述的服务器单板，其特征在于，所述主处理器还用于：

通过所述 PCIe 交换器接收所述从处理器发送加载请求；

根据所述加载请求读取操作系统和应用程序镜像文件；

将读取到的所述操作系统镜像文件通过所述 PCIe 交换器发送至所述从服务器，使得所述从处理器根据所述操作系统镜像文件启动操作系统。

6. 根据权利要求 1-5 任一所述的服务器单板，其特征在于，所述主处理器还用于：

通过所述 PCIe 交换器向所述从处理器发送在线检测消息，以检测所述从处理器是否正常连接；

当确定一个从处理器连接失败时，根据预设的发送策略将发送至所述一个从处理器的数据包通过所述 PCIe 交换器发送至另一个从处理器。

7. 根据权利要求 6 所述的服务器单板，其特征在于，当确定所述一个从处理器恢复连接时，将发送至所述另一个从处理器的所述数据包通过所述 PCIe 交换器转发至所述一个从处理器。

8. 一种服务器单板实现方法，所述服务器单板应用于服务器系统，所述服务器系统还包括与所述服务器单板连接的至少一个其他单板，其特征在于，包括：

主处理器通过网口接收所述其他单板发送的第一以太网数据包；

所述主处理器将所述第一以太网数据包转成第一 PCIe 数据包；

所述主处理器通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器；

所述主处理器通过所述 PCIe 交换器接收所述从处理器发送的第二 PCIe 数据包；

所述主处理器将所述第二 PCIe 数据包转成第二以太网数据包；

所述主处理器通过所述网口向所述其他单板转发所述第二以太网数据包。

9. 根据权利要求 8 所述的方法，其特征在于，所述方法还包括：

所述主处理器通过所述 PCIe 交换器接收其中一个所述从处理器发送的第三 PCIe 数据包，并通过所述 PCIe 交换器将所述第三 PCIe 数据包转发给另一个所述从处理器。

10. 根据权利要求 8 或 9 任一所述的方法，其特征在于，所述主处理器将所述第二 PCIe 数据包转成第二以太网数据包包括：

若所述第二 PCIe 数据包的净荷包括完整的以太网数据包，所述主处理器直接提取所述以太网数据包以得到第二以太网数据包；或，

若所述第二 PCIe 数据包的净荷包括以太网数据包的净荷，所述主处理器提取所述以太网数据包的净荷进行封装以得到第二以太网数据包。

11. 根据权利要求 8-10 任一所述的方法，其特征在于，所述方法还包括：

生成路径配置表，所述路径配置表包括从处理器标识以及与所述从处理器标识对应的路由信息；

所述主处理器将所述第一以太网数据包转成第一 PCIe 数据包包括：

将所述第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者所述主处理器提取所述第一以太网数据包中的净荷，将所述第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中；根据所述路径配置表及所述从处理器的从处理器标识获取与所述从处理器对应的第一路由信息，将所述第一路由信息添加到所述第一 PCIe 数据包的头部；

所述主处理器通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连

的所述从处理器包括：

通过所述 PCIe 交换器的上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器根据所述第一 PCIe 数据包头部中的所述第一路由信息将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器。

12. 根据权利要求 8-11 任一所述的方法，其特征在于，所述方法还包括：

所述主处理器通过所述 PCIe 交换器接收所述从处理器发送加载请求；

根据所述加载请求读取操作系统和应用程序镜像文件；

将读取到的所述操作系统镜像文件通过所述 PCIe 交换器发送至所述从服务器，使得所述从处理器根据所述操作系统镜像文件启动操作系统。

13. 根据权利要求 8-12 任一所述的方法，其特征在于，所述方法还包括：

所述主处理器通过所述 PCIe 交换器向所述从处理器发送在线检测消息，以检测所述从处理器是否正常连接；

当确定一个从处理器连接失败时，所述主处理器根据预设的发送策略将发送至所述一个从处理器的数据包通过所述 PCIe 交换器发送至另一从处理器。

14. 根据权利要求 13 所述的方法，其特征在于，当确定所述一个从处理器恢复连接时，将发送至所述另一个从处理器的所述数据包通过所述 PCIe 交换器转发至所述一个从处理器。

15. 一种主处理器，应用于服务器单板，其特征在于，所述主处理器包括：

第一接收模块，用于通过网口接收其他单板发送的第一以太网数据包；

第一转换模块，用于将所述第一接收模块接收到的所述第一以太网数据包转成第一 PCIe 数据包；

第一发送模块，用于通过 PCIe 交换器将所述第一转换模块转换得到的所述第一 PCIe 数据包转发给与所述 PCIe 交换器相连的从处理器；

第二接收模块，用于通过所述 PCIe 交换器接收所述从处理器发送的第二 PCIe 数据包；

第二转换模块，用于将所述第二接收模块接收到的所述第二 PCIe 数据包转成第二以太网数据包；

第二发送模块，用于通过所述网口向所述其他单板转发所述第二转换模块转换得到的所述第二以太网数据包。

16. 根据权利要求 15 所述的主处理器，其特征在于：

所述第一发送模块，具体用于通过 PCIe 交换器的上行端口将所述第一 PCIe 数据包经由 PCIe 交换器发送给与所述 PCIe 交换器的下行端口相连的从处理器；

所述第二接收模块，具体用于从所述 PCIe 交换器的上行端口接收所述从处理器从所述 PCIe 交换器的下行端口发送的、经由所述 PCIe 交换器到达所述 PCIe 交换器的上行端口的第二 PCIe 数据包。

17. 根据权利要求 15-16 任一所述的主处理器，其特征在于，

所述第二接收模块还用于通过所述 PCIe 交换器接收其中一个所述从处理器发送的第三 PCIe 数据包；

所述第一发送模块还用于通过所述 PCIe 交换器将所述第二接收模块接收到的所述第

三 PCIe 数据包转发给另一个所述从处理器。

18. 根据权利要求 15-17 任一所述的主处理器，其特征在于，所述主处理器还包括：

配置表生成模块，用于生成路径配置表，所述路径配置表包括从处理器标识以及与所述从处理器标识对应的路由信息；

所述第一转换模块包括：

净荷处理模块，用于将所述第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者用于提取所述第一以太网数据包中的净荷，将所述第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中；

头部处理模块，用于根据所述路径配置表及所述从处理器的从处理器标识获取与所述从处理器对应的第一路由信息，将所述第一路由信息添加到所述第一 PCIe 数据包的头部，使得所述 PCIe 交换器收到后所述第一发送模块发送的所述第一 PCIe 数据包后，根据所述第一 PCIe 数据包头部中的所述第一路由信息将所述第一 PCIe 数据包转发给所述从处理器。

19. 根据权利要求 15-18 任一所述的主处理器，其特征在于，还包括：

第三接收模块，用于接收所述从处理器通过所述 PCIe 交换器发送的加载请求，根据所述加载请求读取操作系统和应用程序镜像文件；

第三发送模块，用于将读取到的所述操作系统镜像文件通过所述 PCIe 交换器发送至所述从处理器，使得所述从处理器根据所述操作系统镜像文件启动操作系统。

20. 根据权利要求 15-19 任一所述的主处理器，其特征在于，所述主处理器还包括：

检测模块，用于通过所述 PCIe 交换器向所述从处理器发送在线检测消息，以检测所述从处理器是否正常连接；

当所述检测模块确定一个从处理器连接失败时，所述第一发送模块还用于：

根据预设的发送策略将发送至所述从处理器的数据包通过所述 PCIe 交换器发送至另一从处理器；当确定所述一个从处理器恢复连接时，将发送至所述另一个从处理器的所述数据包通过所述 PCIe 交换器转发至所述一个从处理器。

一种服务器单板、服务器单板实现方法及主处理器

技术领域

[0001] 本发明涉及通信领域，尤其涉及一种服务器单板、服务器单板实现方法及主处理器。

背景技术

[0002] 随着通信技术的不断发展，向用户提供核心计算、信息资源管理、信息资源服务等功能的电信设备的数量也在不断增长。过高的能耗不仅影响运营商的运营成本，还将造成能源的严重浪费，因此，电信设备的节能已逐渐成为业界关注的焦点。

[0003] 电信机房一般由大量的服务器、交换设备、接入设备、存储设备以及网络设备等部件所组成。其中，以多种形态呈现的服务器（如机架式、刀片式、插卡式）往往是电信机房中能耗最高的部件。在众多的服务器节能措施中，采用低功耗处理器芯片构建低功耗的服务器就是一个非常有效的措施。

[0004] 现有技术中包括由多个低功耗处理器芯片构成的服务器，在此系统中，为了使数据能够在不同的处理器芯片之间进行合理配置并调度，处理器芯片之间可以通过高速以太网接口进行互连。由于每一个处理器芯片的以太网接口数量有限，处理器芯片之间通常采用树形结构进行互连，当处理器芯片数量较多时，则树形结构的深度需要很大才能够实现每个处理器芯片之间的全互连，但这样一来，对于处在叶子节点的处理器芯片来说，如果两个叶子节点的处理器要实现数据交互，必须要经过多级以太网交换，这将使得延时加大，从而严重影响网络服务的质量。

发明内容

[0005] 本发明的实施例提供一种服务器单板、服务器单板实现方法及主处理器，用于解决现有技术存在着的服务器内部延时大，从而严重影响网络服务质量的问题。

[0006] 为达到上述目的，本发明的实施例采用如下技术方案：

[0007] 第一方面，本发明实施例第一种实现方式提供了一种服务器单板，应用于服务器系统，所述服务器系统还包括与所述服务器单板连接的至少一个其他单板，所述服务器单板包括：

[0008] PCIe 交换器，主处理器以及多个从处理器，所述 PCIe 交换器包括一个上行端口以及多个下行端口；

[0009] 所述主处理器与所述 PCIe 交换器的上行端口相连，所述多个从处理器分别与所述 PCIe 交换器的下行端口相连；

[0010] 所述主处理器用于通过网口接收所述其他单板发送的第一以太网数据包，将所述第一以太网数据包转成第一 PCIe 数据包，通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器；

[0011] 所述主处理器还用于通过所述 PCIe 交换器接收所述从处理器发送的第二 PCIe 数

据包，将所述第二 PCIe 数据包转成第二以太网数据包，通过所述网口向所述其他单板转发所述第二以太网数据包。

[0012] 结合第一方面第一种可能实现方式，第二种可能的实现方式中，所述主处理器还用于，通过所述 PCIe 交换器接收其中一个所述从处理器发送的第三 PCIe 数据包，并通过所述 PCIe 交换器将所述第三 PCIe 数据包转发给另一个所述从处理器。

[0013] 结合第一方面第一种或第二种可能的实现方式，第三种可能的实现方式中，所述主处理器用于将所述第二 PCIe 数据包转成第二以太网数据包具体为：

[0014] 若所述第二 PCIe 数据包的净荷包括完整的以太网数据包，所述主处理器用于直接提取所述以太网数据包以得到第二以太网数据包；或，

[0015] 若所述第二 PCIe 数据包的净荷包括以太网数据包的净荷，所述主处理器用于提取所述以太网数据包的净荷进行封装以得到第二以太网数据包。

[0016] 结合第一方面第一到第三种任一可能的实现方式，第四种可能的实现方式中，所述主处理器还用于，生成路径配置表，所述路径配置表包括从处理器标识以及与所述从处理器标识对应的路由信息；

[0017] 所述主处理器用于将所述第一以太网数据包转成第一 PCIe 数据包，通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器具体为：

[0018] 所述主处理器用于：

[0019] 将所述第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者所述主处理器用于提取所述第一以太网数据包中的净荷，将所述第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中；根据所述路径配置表及所述从处理器的从处理器标识获取与所述从处理器对应的第一路由信息，将所述第一路由信息添加到所述第一 PCIe 数据包的头部；通过所述 PCIe 交换器的上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器根据所述第一 PCIe 数据包头部中的所述第一路由信息将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器。

[0020] 结合第一方面第一到第四种任一可能的实现方式，第五种可能的实现方式中，所述主处理器还用于：

[0021] 通过所述 PCIe 交换器接收所述从处理器发送加载请求；

[0022] 根据所述加载请求读取操作系统和应用程序镜像文件；

[0023] 将读取到的所述操作系统镜像文件通过所述 PCIe 交换器发送至所述从服务器，使得所述从处理器根据所述操作系统镜像文件启动操作系统。

[0024] 结合第一方面第一到第五种任一可能的实现方式，第六种可能的实现方式中，

[0025] 所述主处理器还用于：

[0026] 通过所述 PCIe 交换器向所述从处理器发送在线检测消息，以检测所述从处理器是否正常连接；

[0027] 当确定一个从处理器连接失败时，根据预设的发送策略将发送至所述一个从处理器的数据包通过所述 PCIe 交换器发送至另一个从处理器。

[0028] 结合第一方面第六种实现方式，第七种可能的实现方式中，当确定所述一个从处理器恢复连接时，将发送至所述另一个从处理器的所述数据包通过所述 PCIe 交换器转发

至所述一个从处理器。

[0029] 第二方面，在第一种可能的实现方式中，本实施例公开了一种服务器单板实现方法，所述服务器单板应用于服务器系统，所述服务器系统还包括与所述服务器单板连接的至少一个其他单板，包括：

[0030] 主处理器通过网口接收所述其他单板发送的第一以太网数据包；

[0031] 所述主处理器将所述第一以太网数据包转成第一 PCIe 数据包；

[0032] 所述主处理器通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器；

[0033] 所述主处理器通过所述 PCIe 交换器接收所述从处理器发送的第二 PCIe 数据包；

[0034] 所述主处理器将所述第二 PCIe 数据包转成第二以太网数据包；

[0035] 所述主处理器通过所述网口向所述其他单板转发所述第二以太网数据包。

[0036] 结合第二方面第一种可能的实现方式，第二种可能的实现方式中，所述方法还包括：

[0037] 所述主处理器通过所述 PCIe 交换器接收其中一个所述从处理器发送的第三 PCIe 数据包，并通过所述 PCIe 交换器将所述第三 PCIe 数据包转发给另一个所述从处理器。

[0038] 结合第二方面第一到第二种任一可能的实现方式，第三种可能的实现方式中，所述主处理器将所述第二 PCIe 数据包转成第二以太网数据包包括：

[0039] 若所述第二 PCIe 数据包的净荷包括完整的以太网数据包，所述主处理器直接提取所述以太网数据包以得到第二以太网数据包；或，

[0040] 若所述第二 PCIe 数据包的净荷包括以太网数据包的净荷，所述主处理器提取所述以太网数据包的净荷进行封装以得到第二以太网数据包。

[0041] 结合第二方面第一到第三种任一可能的实现方式，第四种可能的实现方式中，所述方法还包括：

[0042] 生成路径配置表，所述路径配置表包括从处理器标识以及与所述从处理器标识对应的路由信息；

[0043] 所述主处理器将所述第一以太网数据包转成第一 PCIe 数据包包括：

[0044] 将所述第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者所述主处理器提取所述第一以太网数据包中的净荷，将所述第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中；根据所述路径配置表及所述从处理器的从处理器标识获取与所述从处理器对应的第一路由信息，将所述第一路由信息添加到所述第一 PCIe 数据包的头部；

[0045] 所述主处理器通过所述 PCIe 交换器的所述上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器包括：

[0046] 通过所述 PCIe 交换器的上行端口将所述第一 PCIe 数据包转发给所述 PCIe 交换器，使得所述 PCIe 交换器根据所述第一 PCIe 数据包头部中的所述第一路由信息将所述第一 PCIe 数据包转发给与所述下行端口相连的所述从处理器。

[0047] 结合第二方面第一到第四种任一可能的实现方式，第五种可能的实现方式中，所述方法还包括：

[0048] 所述主处理器通过所述 PCIe 交换器接收所述从处理器发送加载请求；

[0049] 根据所述加载请求读取操作系统和应用程序镜像文件；

[0050] 将读取到的所述操作系统镜像文件通过所述 PCIe 交换器发送至所述从服务器，使得所述从处理器根据所述操作系统镜像文件启动操作系统。

[0051] 结合第二方面第一到第五种任一可能的实现方式，第六种可能的实现方式中，所述主处理器通过所述 PCIe 交换器向所述从处理器发送在线检测消息，以检测所述从处理器是否正常连接；

[0052] 当确定一个从处理器连接失败时，所述主处理器根据预设的发送策略将发送至所述一个从处理器的数据包通过所述 PCIe 交换器发送至另一从处理器。

[0053] 结合第二方面第六种可能的实现方式，第七种可能的实现方式中，当确定所述一个从处理器恢复连接时，将发送至所述另一个从处理器的所述数据包通过所述 PCIe 交换器转发至所述一个从处理器。

[0054] 第三方面，第一种可能的方式中，本实施例公开了一种处理器，应用于服务器单板，所述主处理器包括：

[0055] 第一接收模块，用于通过网口接收其他单板发送的第一以太网数据包；

[0056] 第一转换模块，用于将所述第一接收模块接收到的所述第一以太网数据包转成第一 PCIe 数据包；

[0057] 第一发送模块，用于通过 PCIe 交换器将所述第一转换模块转换得到的所述第一 PCIe 数据包转发给与所述 PCIe 交换器相连的从处理器；

[0058] 第二接收模块，用于通过所述 PCIe 交换器接收所述从处理器发送的第二 PCIe 数据包；

[0059] 第二转换模块，用于将所述第二接收模块接收到的所述第二 PCIe 数据包转成第二以太网数据包；

[0060] 第二发送模块，用于通过所述网口向所述其他单板转发所述第二转换模块转换得到的所述第二以太网数据包。

[0061] 结合第三方面第一可能的实现方式，第二种可能的实现方式中，所述第一发送模块，具体用于通过 PCIe 交换器的上行端口将所述第一 PCIe 数据包经由 PCIe 交换器发送给与所述 PCIe 交换器的下行端口相连的从处理器；

[0062] 所述第二接收模块，具体用于从所述 PCIe 交换器的上行端口接收所述从处理器从所述 PCIe 交换器的下行端口发送的、经由所述 PCIe 交换器到达所述 PCIe 交换器的上行端口的第二 PCIe 数据包。

[0063] 结合第三方面第一到第二种任一可能的实现方式，第三种可能的实现方式中，所述第二接收模块还用于通过所述 PCIe 交换器接收其中一个所述从处理器发送的第三 PCIe 数据包；

[0064] 所述第一发送模块还用于通过所述 PCIe 交换器将所述第二接收模块接收到的所述第三 PCIe 数据包转发给另一个所述从处理器。

[0065] 结合第三方面第一到第三种任一可能的实现方式，第四种可能的实现方式中，所述主处理器还包括：

[0066] 配置表生成模块，用于生成路径配置表，所述路径配置表包括从处理器标识以及

与所述从处理器标识对应的路由信息；

[0067] 所述第一转换模块包括：

[0068] 净荷处理模块，用于将所述第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者用于提取所述第一以太网数据包中的净荷，将所述第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中；

[0069] 头部处理模块，用于根据所述路径配置表及所述从处理器的从处理器标识获取与所述从处理器对应的第一路由信息，将所述第一路由信息添加到所述第一 PCIe 数据包的头部，使得所述 PCIe 交换器收到后所述第一发送模块发送的所述第一 PCIe 数据包后，根据所述第一 PCIe 数据包头部中的所述第一路由信息将所述第一 PCIe 数据包转发给所述从处理器。

[0070] 结合第三方面第一到第四种任一可能的实现方式，第五种可能的实现方式中，第三接收模块，用于接收所述从处理器通过所述 PCIe 交换器发送的加载请求，根据所述加载请求读取操作系统和应用程序镜像文件；

[0071] 第三发送模块，用于将读取到的所述操作系统镜像文件通过所述 PCIe 交换器发送至所述从处理器，使得所述从处理器根据所述操作系统镜像文件启动操作系统。

[0072] 结合第二方面第一到第五种任一可能的实现方式，第六种可能的实现方式中，所述主处理器还包括：

[0073] 检测模块，用于通过所述 PCIe 交换器向所述从处理器发送在线检测消息，以检测所述从处理器是否正常连接；

[0074] 当所述检测模块确定一个从处理器连接失败时，所述第一发送模块还用于：

[0075] 根据预设的发送策略将发送至所述从处理器的数据包通过所述 PCIe 交换器发送至另一从处理器；当确定所述一个从处理器恢复连接时，将发送至所述另一个从处理器的所述数据包通过所述 PCIe 交换器转发至所述一个从处理器。

[0076] 本发明实施例提供的服务器单板、服务器单板实现方法及主处理器，该服务器单板应用于服务器系统，该服务器系统还包括与该服务器单板连接的至少一个其他单板，该服务器单板具体包括 PCIe 交换器，主处理器以及多个从处理器，其中，PCIe 交换器包括一个上行端口以及多个下行端口；主处理器与该 PCIe 的上行端口相连，多个从处理器分别与该 PCIe 交换器的下行端口相连。采用这样一种结构的服务器单板，主处理器可以通过网口与其他单板进行数据的收发，该主处理器又通过 PCIe 交换器与各个从处理器相连接，因此，从处理器数据的收发以及各个从处理器之间数据的交换均可以通过主处理器完成。这样一来，无需采用多级树形结构就可以实现全处理器之间的互连，从而避免了数据在多级处理器之间转发，降低了服务器单板的延时，提高了网络服务的质量。

附图说明

[0077] 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

[0078] 图 1 为本发明实施例提供的一种服务器单板的结构示意图；

[0079] 图 2a 为本发明实施例提供的一种主处理器的结构示意图；

[0080] 图 2b 为本发明实施例提供的另一主处理器的结构示意图；

[0081] 图 3 为本发明实施例提供的一种服务器单板的连接结构示意图；

[0082] 图 4 为本发明实施例提供的一种服务器单板实现方法的流程示意图；

[0083] 图 5 为本发明实施例提供的另一服务器单板实现方法的流程示意图。

具体实施方式

[0084] 下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0085] 本发明实施例提供的服务器单板 1，可以应用于服务器系统，该服务器系统还可以包括与服务器单板 1 连接的至少一个其他单板 2，如图 1 所示，服务器单板 1 可以包括：

[0086] PCIe 交换器 11，主处理器 12 以及多个从处理器 13，该 PCIe 交换器 11 可以包括一个上行端口以及多个下行端口。

[0087] 该主处理器 12 与 PCIe 交换器 11 的上行端口相连，多个从处理器 13 分别与该 PCIe 交换器 11 的下行端口相连。

[0088] 该主处理器 12 用于通过网口 14 接收其他单板 2 发送的第一以太网数据包，将该第一以太网数据包转成第一 PCIe 数据包，通过 PCIe 交换器 11 的上行端口将该第一 PCIe 数据包转发给 PCIe 交换器 11，使得该 PCIe 交换器 11 将该第一 PCIe 数据包转发给与下行端口相连的从处理器 13。

[0089] 例如，主处理器 12 将第一以太网数据包转成第一 PCIe 数据包的过程具体可以包括：主处理器 12 将该第一以太网数据包整体封装到第一 PCIe 数据包的净荷中，采用这样一种整体封装的方式可以高效快速地实现数据包格式的转换；或者，主处理器 12 还可以提取该第一以太网数据包中的净荷，并将该第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中，这样一来，封装的 PCIe 数据包中不包含包括以太网数据包的包头等在内的非关键数据，大大减小了 PCIe 数据包的容量，提高了 PCIe 数据包的封装效率。

[0090] 该主处理器 12 还用于通过 PCIe 交换器 11 接收从处理器 13 发送的第二 PCIe 数据包，将该第二 PCIe 数据包转成第二以太网数据包，通过网口 14 向其他单板 2 转发该第二以太网数据包。

[0091] 相应的，主处理器 12 将第二 PCIe 数据包转成第二以太网数据包的过程具体可以包括：若该第二 PCIe 数据包的净荷包括完整的以太网数据包，主处理器 12 可以直接提取该以太网数据包以得到第二以太网数据包；或者，若该第二 PCIe 数据包的净荷包括以太网数据包的净荷，主处理器 12 可以提取该以太网数据包的净荷进行封装以得到第二以太网数据包。

[0092] 在本发明实施例中，处理器可以包括 x86 处理器或 ARM 处理器。其中，x86 处理器具体是指基于 x86 的低功耗的处理器。在实际应用中，由于基于 ARM 架构的处理器具有低成本、低功耗等特点，通过将多个较低性能的 ARM 处理器进行集群可以获得较高性能的服务器，在保证了服务器性能的同时大大降低了生产成本，降低了服务器的功耗。

[0093] 本发明实施例提供的服务器单板，应用于服务器系统，该服务器系统还包括与该服务器单板连接的至少一个其他单板，该服务器单板具体包括 PCIe 交换器，主处理器以及多个从处理器，其中，PCIe 交换器包括一个上行端口以及多个下行端口；主处理器与该 PCIe 的上行端口相连，多个从处理器分别与该 PCIe 交换器的下行端口相连。采用这样一种结构的服务器单板，主处理器可以通过网口与其他单板进行数据的收发，该主处理器又通过 PCIe 交换器与各个从处理器相连接，因此，从处理器数据的收发以及各个从处理器之间数据的交换均可以通过主处理器完成。这样一来，无需采用多级树形结构就可以实现全处理器之间的互连，从而避免了数据在多级处理器之间转发，降低了服务器单板的延时，提高了网络服务的质量。

[0094] 进一步地，主处理器 12 还可以用于，通过 PCIe 交换器 11 接收其中一个从处理器 13 发送的第三 PCIe 数据包，并通过 PCIe 交换器 11 将该第三 PCIe 数据包转发给另一个从处理器 13。

[0095] 这样一来，多个从处理器 13 之间可以通过主处理器 12 控制 PCIe 交换器 11 进行数据交换，从而避免了数据在多级处理器之间转发，降低了服务器单板的延时，提高了网络服务的质量。

[0096] 进一步地，主处理器 12 还可以用于：

[0097] 初始化 PCIe 交换器 11，通过配置 PCIe 交换器 11 的寄存器以使得该 PCIe 交换器 11 能够正常工作。

[0098] 其中，主处理器 12 初始化 PCIe 交换器 11 具体可以包括：为与该 PCIe 交换器 11 相连的各个从处理器 13 分配路由信息。

[0099] 需要说明的是，在本发明实施例中，路由信息具体可以包括 BDF(Bus（总线）、Device（设备）、Function（功能）号的统称）信息或者 BAR(Base Address Register，基地址寄存器）空间地址。

[0100] 该主处理器 12 还可以用于，生成路径配置表，该路径配置表包括从处理器 13 标识以及与该从处理器 13 标识对应的路由信息，其中，与该从处理器 13 标识对应的路由信息与为该处理器标识所指示的从处理器 13 分配的路由信息相同。

[0101] 进一步地，主处理器 12 将第一以太网数据包转成第一 PCIe 数据包，通过 PCIe 交换器 11 的上行端口将第一 PCIe 数据包转发给 PCIe 交换器 11，使得 PCIe 交换器 11 将第一 PCIe 数据包转发给与下行端口相连的从处理器 13 的具体过程可以如下：

[0102] 首先，主处理器 12 可以将第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者主处理器 12 可以提取第一以太网数据包中的净荷，将该第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中。

[0103] 在完成第一 PCIe 数据包的封装之后，主处理器可以根据路径配置表及从处理器 13 的从处理器标识获取与该从处理器 13 对应的第一路由信息，将该第一路由信息添加到第一 PCIe 数据包的头部；通过 PCIe 交换器 11 的上行端口将该第一 PCIe 数据包转发给 PCIe 交换器 11，使得 PCIe 交换器 11 可以根据该第一 PCIe 数据包头部中的第一路由信息将第一 PCIe 数据包转发给与下行端口相连的从处理器 13。

[0104] 例如，路径配置表中记录有每一个从处理器的身份信息，该身份信息可以是从处理器所处的总线 ID 或者是该从处理器的 MAC(Medium Access Control，介质访问控制层）

地址，每一个从处理器的身份信息均对应一个预设的路由信息，主处理器 12 在通过 PCIe 交换器 11 转发第一 PCIe 数据包之前，可以根据路径配置表及从处理器 13 的从处理器标识获取与该从处理器 13 对应的路由信息，将该路由信息添加到第一 PCIe 数据包的头部，PCIe 交换器 11 可以通过第一 PCIe 数据包头部记录的路由信息查找路径配置表，根据该路径配置表获得一个或多个从处理器 13 的路由信息，从而将第一 PCIe 数据包转发给一个或多个从处理器 13。

[0105] 当然，多个从处理器 13 之间同样可以根据该路径配置表进行数据的交换。例如，在主处理器 12 完成 PCIe 交换器 11 的初始化和从处理器 13 的扫描后生成路径配置表，并将该路径配置表下发到各个从处理器 13，当第一从处理器需要向第二从处理器转发数据时，可以将该数据包发送至 PCIe 交换器 11，该转发的数据包中添加有该第二从处理器的身份信息，PCIe 交换器 11 可以查询路径配置表，从而根据第二从处理器的身份信息确定该第二从处理器的路由信息以便完成数据的转发。

[0106] 在主处理器 12 初始化 PCIe 交换器 11 之后，从处理器 13 还可以用于：

[0107] 通过 PCIe 交换器 11 向主处理器 12 发送加载请求，以请求加载必要的操作系统和应用程序。

[0108] 相应的，在接收到该加载请求之后，主处理器 12 还可以用于：根据该加载请求读取操作系统和应用程序镜像文件；将读取到的操作系统镜像文件通过 PCIe 交换器 11 发送至从服务器 13，使得该从处理器 13 根据该操作系统镜像文件启动操作系统。

[0109] 这样一来，可以预先将操作系统和应用程序的镜像文件放置在主处理器 12 的存储介质中，当从处理器 13 需要加载时，只需要将该镜像文件从主处理器 12 的存储介质中取出发送至从处理器 13 即可，从而可以使得各个从处理器 13 无需自备操作系统和应用程序，节省了从处理器 13 的内存，提高了从处理器 13 的处理效率。

[0110] 进一步地，主处理器 12 还可以用于：

[0111] 通过 PCIe 交换器 11 向从处理器 13 发送在线检测消息，以检测从处理器 13 是否正常连接。

[0112] 具体的，当确定一个从处理器连接失败时，可以根据预设的发送策略将发送至该从处理器的数据包通过 PCIe 交换器 11 发送至另一个从处理器；当确定该从处理器恢复连接时，将发送至另一个从处理器的数据包通过 PCIe 交换器 11 转发至该从处理器。当连接失败时，说明从处理器出现了故障（或者两个主处理器之间的链路故障），因此需要将数据转发到其他的从处理器以对数据进行处理。这里的“预设的发送策略”具体并不限定，例如，可以事先定义好某个从处理器用于处理其他从处理器（或链路）出故障时的数据，如果其他从处理器（或链路）出现故障，则主处理器将数据转给给这个专用的从处理器来处理；或者“预设的发送策略”也可以是预设的“根据各个从处理器负载情况进行发送”的策略（如将数据包发送到负载少的从处理器）。

[0113] 例如，主处理器 12 与各从处理器 13 之间具有心跳检测机制，主处理器 12 将定时向各从处理器 13 发送信条检测报文以检测从处理器 13 是否正常连接。当一个从处理器的心跳丢失，主处理器 12 将该从处理器上的业务迁移到另一个从处理器或其他备用的从处理器上。

[0114] 本发明实施例提供的主处理器 12，可以应用于上述服务器单板，如图 2a 所示，主

处理器 12 包括 :

[0115] 第一接收模块 21,用于通过网口接收其他单板发送的第一以太网数据包。

[0116] 第一转换模块 22,用于将第一接收模块 21 接收到的第一以太网数据包转成第一 PCIe 数据包。

[0117] 第一发送模块 23,用于通过 PCIe 交换器将第一转换模块 22 转换得到的第一 PCIe 数据包转发给与该 PCIe 交换器相连的从处理器。

[0118] 第二接收模块 24,用于通过 PCIe 交换器接收从处理器发送的第二 PCIe 数据包。

[0119] 第二转换模块 25,用于将第二接收模块 24 接收到的第二 PCIe 数据包转成第二以太网数据包。

[0120] 第二发送模块 26,用于通过网口向其他单板转发第二转换模块 25 转换得到的第二以太网数据包。

[0121] 本发明实施例提供的主处理器,应用于服务器单板,该服务器单板应用于服务器系统,该服务器系统还包括与该服务器单板连接的至少一个其他单板,该服务器单板具体包括 PCIe 交换器,主处理器以及多个从处理器,其中,PCIe 交换器包括一个上行端口以及多个下行端口 ;主处理器与该 PCIe 的上行端口相连,多个从处理器分别与该 PCIe 交换器的下行端口相连。采用这样一种结构的服务器单板,主处理器可以通过网口与其他单板进行数据的收发,该主处理器又通过 PCIe 交换器与各个从处理器相连接,因此,从处理器数据的收发以及各个从处理器之间数据的交换均可以通过主处理器完成。这样一来,无需采用多级树形结构就可以实现全处理器之间的互连,从而避免了数据在多级处理器之间转发,降低了服务器单板的延时,提高了网络服务的质量。

[0122] 其中,第一发送模块 23,具体用于通过 PCIe 交换器的上行端口将第一 PCIe 数据包经由 PCIe 交换器发送给与该 PCIe 交换器的下行端口相连的从处理器。

[0123] 第二接收模块 24,具体用于从 PCIe 交换器的上行端口接收从处理器从该 PCIe 交换器的下行端口发送的、经由该 PCIe 交换器到达 PCIe 交换器的上行端口的第二 PCIe 数据包。

[0124] 进一步地,第二接收模块 24 还可以用于通过 PCIe 交换器接收其中一个从处理器发送的第三 PCIe 数据包 ;第一发送模块 23 还可以用于通过 PCIe 交换器将第二接收模块接收到的第三 PCIe 数据包转发给另一个从处理器。

[0125] 这样一来,多个从处理器之间可以通过主处理器 12 控制 PCIe 交换器进行数据交换,从而避免了数据在多级处理器之间转发,降低了服务器单板的延时,提高了网络服务的质量。

[0126] 在本发明实施例中,第一转换模块 22 还可以用于 :

[0127] 将第一以太网数据包整体封装到第一 PCIe 数据包的净荷中,或者提取第一以太网数据包中的净荷,将该第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中 ;

[0128] 相应的,第二转换模块 25 还可以用于 :

[0129] 若第二 PCIe 数据包的净荷包括完整的以太网数据包,直接提取该以太网数据包以得到第二以太网数据包,或者若第二 PCIe 数据包的净荷包括以太网数据包的净荷,提取该以太网数据包的净荷进行封装以得到第二以太网数据包。

[0130] 进一步地,如图 2b 所示,主处理器 12 还可以包括 :

[0131] 配置表生成模块 27，用于生成路径配置表，该路径配置表包括从处理器标识以及与该从处理器标识对应的路由信息。

[0132] 第一转换模块 22 还可以包括：

[0133] 净荷处理模块 221，用于将第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者用于提取第一以太网数据包中的净荷，将第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中。

[0134] 头部处理模块 222，用于根据该路径配置表及从处理器的从处理器标识获取与该从处理器对应的第一路由信息，将该第一路由信息添加到第一 PCIe 数据包的头部，使得 PCIe 交换器收到后第一发送模块 23 发送的第一 PCIe 数据包后，根据该第一 PCIe 数据包头部中的第一路由信息将该第一 PCIe 数据包转发给从处理器。

[0135] 需要说明的是，在本发明实施例中，路由信息具体可以包括 BDF(Bus（总线）、Device（设备）、Function（功能）号的统称）信息或者 BAR(Base Address Register，基地址寄存器）空间地址。

[0136] 如图 2b 所示，主处理器 12 还可以包括：

[0137] 第三接收模块 28，用于接收从处理器通过 PCIe 交换器发送的加载请求，根据该加载请求读取操作系统和应用程序镜像文件。

[0138] 第三发送模块 29，用于将读取到的操作系统镜像文件通过 PCIe 交换器发送至从处理器，使得该从处理器根据该操作系统镜像文件启动操作系统。

[0139] 这样一来，可以预先将操作系统和应用程序的镜像文件存储在主处理器 12 的存储介质中，当从处理器需要加载时，只需要将该镜像文件从主处理器 12 的存储介质中取出发送至从处理器即可，从而可以使得各个从处理器无需自备操作系统和应用程序，节省了从处理器的内存，提高了从处理器的处理效率。

[0140] 进一步地，主处理器 12 还可以包括：

[0141] 检测模块 20，用于通过 PCIe 交换器向从处理器发送在线检测消息，以检测该从处理器是否正常连接；

[0142] 当检测模块 20 确定一个从处理器连接失败时，第一发送模块 23 还可以用于：

[0143] 根据预设的发送策略将发送至从处理器的数据包通过 PCIe 交换器发送至另一从处理器；当确定一个从处理器恢复连接时，将发送至另一个从处理器的数据包通过 PCIe 交换器转发至该一个从处理器。当连接失败时，说明从处理器出现了故障（或者两个主处理器之间的链路故障），因此需要将数据转发到其他的从处理器以对数据进行处理。这里的“预设的发送策略”具体并不限定，例如，可以事先定义好某个从处理器用于处理其他从处理器（或链路）出故障时的数据，如果其他从处理器（或链路）出现故障，则主处理器将数据转给给这个专用的从处理器来处理；或者“预设的发送策略”也可以是预设的“根据各个从处理器负载情况进行发送”的策略（如将数据包发送到负载少的从处理器）。

[0144] 例如，主处理器 12 与各从处理器之间具有心跳检测机制，主处理器 12 将定时向各从处理器发送信条检测报文以检测从处理器是否正常连接。当一个从处理器的心跳丢失，主处理器 12 将该从处理器上的业务迁移到另一个从处理器或其他备用的从处理器上。

[0145] 需要说明的是，在本发明实施例中，主处理器 12 还可以包括内存单元和存储介质，主处理器 12 和从处理器 13 的结构可以相同。其中，内存单元具体可以包括 DRAM(dynamic

random access memory,动态随机存取存储器）或SRAM(static random access memory,静态随机存取存储器）;存储介质具体可以包括Flash（闪存）、SATA(serial advanced technology attachment,串行高级技术附件）、SAS(Serial Attached SCSI,串行连接接口）硬盘或SSD(solid state disk,固态硬盘),该存储介质可以用于存储操作系统和应用程序的镜像文件。

[0146] 图3为本发明实施例提供的一种低功耗的服务器单板1的连接结构示意图。服务器单板1包括服务器底板31,该服务器底板31具体是由电源311、单板管理控制器312和散热装置（图中未示出）所组成。一主处理器12和六个从处理器13之间通过PCIe交换器11相互连接,在本发明实施例中,主处理器12和从处理器13均采用结构相同的多功能子卡,该多功能子卡主要包括ARM处理器、内存单元和存储介质。

[0147] 其中,服务器底板31上具有多个插槽,作为主处理器12和从处理器13的多功能子卡就是插在这些插槽中。插槽又分为主插槽和从插槽,主插槽的插槽地址和从插槽的插槽地址是不同的。当各多功能子卡插到插槽中通电启动后,插到主插槽的多功能子卡上的ARM处理器通过识别插槽地址就自动地将自身配置为主处理器,插到从插槽的多功能子卡上的ARM处理器通过识别插槽地址就自动地将自身配置为从处理器。

[0148] 主处理器12的ARM处理器连接到PCIe交换器11的RC(rootcomplex,根组件）接口,从处理器13的ARM处理器连接到PCIe交换器11的EP(end point,终端设备）接口。

[0149] 在本发明实施例中,主处理器12包括处理单元、内存和存储介质,其中,处理单元可以包括接收模块、转换模块、发送模块、初始化模块、查表模块、掉电保护模块、XOR异或加速模块和安全加密模块。

[0150] 主处理器12连接服务器单板1的网口14,该主处理器12是从处理器13的业务分发点和业务汇聚点,作为主处理器12的多功能子卡是一个智能网卡。该智能网卡主要负责PCIe数据报文的收发、以太网报文的收发;PCIe数据报文的地址（包括地址路由、ID路由、隐式路由等）解析,并根据该地址查询路径配置表;以太网报文及上层报文的解析,并根据以太网及上层报文的特征查询路径配置表;PCIe数据报文与以太网报文之间的格式转换。除上述功能以外,主处理器12还可以完成PCIe交换器11的初始化、对从处理器13进行扫描、对PCIe交换器11进行管理等等。

[0151] 当该服务器单板通电启动后,主处理器12和从处理器13各自进行初始化,当主处理器12完成PCIe交换器11的初始化后,系统即进入加载阶段。从处理器13通过PCIe交换器11,向主处理器42请求加载操作系统和应用程序。主处理器12收到从处理器13的请求后,从下挂的存储介质（如:Flash、SATA、SAS硬盘、SSD硬盘等）中读取相关操作系统和应用程序,通过PCIe交换器11发送给从处理器13。

[0152] 主处理器12完成PCIe交换器11的初始化和从处理器13的扫描后生成路径配置表,并将该表下发到各从处理器13。从处理器13可以根据该路径配置表,通过PCIe交换器11与主处理器12通信,也可以和该服务器单板上的其它从处理器13进行通信。

[0153] 在本发明实施例中,每个从处理器13上运行的应用程序可以是动态部署的,用户可针对不同的应用,通过主处理器12对各从处理器13的工作模式进行配置。从处理器13之间是集群关系,各从处理器13上运行各自的操作系统和应用程序。主处理器12上也运行自己的操作系统和应用程序。比如:第一个从处理器131运行的是计算功能,第二个从处

理器 132 运行的是存储功能，这两个从处理器可以相互配合完成一个特定的业务功能。当负责计算的从处理器 131 需要向负责存储的从处理器 132 取数据时，可以直接通过 PCIe 交换器 11 完成数据交换，即本地完成数据交换，这样既可以降低延时又对外部网络的带宽没有影响。

[0154] 进一步地，主处理器 12 还可以用于：

[0155] 通过 PCIe 交换器 11 向从处理器 13 发送在线检测消息，以检测从处理器 13 是否正常连接。

[0156] 具体的，如图 3 所示，当确定第一从处理器 131 连接失败时，将发送至该第一从处理器 131 的数据包通过 PCIe 交换器 11 发送至第二从处理器 132。

[0157] 例如，主处理器 12 与各从处理器 13 之间具有心跳检测机制，主处理器 12 将定时向各从处理器 13 发送信条检测报文以检测从处理器 13 是否正常连接。当第一从处理器 131 的心跳丢失，主处理器 12 将第一从处理器 131 上的业务迁移到第二从处理器 132 或其他备用的从处理器上。

[0158] 需要说明的是，在本发明实施例中，主处理器与从处理器之间，从处理器与从处理器之间，均可以通过 Linux 内核态下自定义的以太网报文收发接口进行通信。在具体的物理形态上，本发明实施例提供的服务器单板可以是机架式的服务器，也可以是刀片式的服务器。

[0159] 这样一种结构的服务器单板，无需采用多级树形结构就可以实现全处理器之间的互连，从而避免了数据在多级处理器之间转发，降低了服务器单板的延时，提高了网络服务质量。

[0160] 本发明实施例提供的服务器单板实现方法，如图 4 所示，该服务器单板可以应用于服务器系统，该服务器系统还包括与该服务器单板连接的至少一个其他单板，该方法包括：

[0161] S401、主处理器通过网口接收其他单板发送的第一以太网数据包。

[0162] 需要说明的是，在本发明实施例中，主处理器和从处理器的结构可以相同。

[0163] 具体的，主处理器可以通过第一接收模块 21 从网口接收其他单板发送的第一以太网数据包。

[0164] S402、主处理器将该第一以太网数据包转成第一 PCIe 数据包。

[0165] 具体的，主处理器可以通过第一转换模块 22 将该第一以太网数据包转成第一 PCIe 数据包。例如，主处理器可以将该第一以太网数据包整体封装到第一 PCIe 数据包的净荷中，采用这样一种整体封装的方式可以高效快速地实现数据包格式的转换；或者，主处理器还可以提取该第一以太网数据包中的净荷，并将该第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中，这样一来，封装的 PCIe 数据包中不包含包括以太网数据包的包头等在内的非关键数据，大大减小了 PCIe 数据包的容量，提高了 PCIe 数据包的封装效率。

[0166] S403、主处理器通过 PCIe 交换器的上行端口将该第一 PCIe 数据包转发给 PCIe 交换器，使得该 PCIe 交换器将该第一 PCIe 数据包转发给与下行端口相连的从处理器。

[0167] 具体的，主处理器可以通过第一发送模块 23 将该第一 PCIe 数据包转发给一个或多个从处理器。

[0168] S404、主处理器通过 PCIe 交换器接收从处理器发送的第二 PCIe 数据包。

[0169] S405、主处理器将该第二 PCIe 数据包转成第二以太网数据包。

[0170] 相应的，若该第二 PCIe 数据包的净荷包括完整的以太网数据包，主处理器可以直接提取该以太网数据包以得到第二以太网数据包；或者，若该第二 PCIe 数据包的净荷包括以太网数据包的净荷，主处理器可以提取该以太网数据包的净荷进行封装以得到第二以太网数据包。

[0171] S406、主处理器通过网口向其他单板转发该第二以太网数据包。

[0172] 在本发明实施例中，处理器可以包括 x86 处理器或 ARM 处理器。其中，x86 处理器具体是指基于 x86 的低功耗的处理器。在实际应用中，由于基于 ARM 架构的处理器具有低成本、低功耗等特点，通过将多个较低性能的 ARM 处理器进行集群可以获得较高性能的服务器，在保证了服务器性能的同时大大降低了生产成本，降低了服务器的功耗。

[0173] 本发明实施例提供的服务器单板实现方法，该服务器单板应用于服务器系统，该服务器系统还包括与该服务器单板连接的至少一个其他单板，该服务器单板具体包括 PCIe 交换器，主处理器以及多个从处理器，其中，PCIe 交换器包括一个上行端口以及多个下行端口；主处理器与该 PCIe 的上行端口相连，多个从处理器分别与该 PCIe 交换器的下行端口相连。采用这样一种结构的服务器单板，主处理器可以通过网口与其他单板进行数据的收发，该主处理器又通过 PCIe 交换器与各个从处理器相连接，因此，从处理器数据的收发以及各个从处理器之间数据的交换均可以通过主处理器完成。这样一来，无需采用多级树形结构就可以实现全处理器之间的互连，从而避免了数据在多级处理器之间转发，降低了服务器单板的延时，提高了网络服务的质量。

[0174] 进一步地，如图 5 所示，本发明实施例提供的服务器单板实现方法包括：

[0175] S501、主处理器初始化 PCIe 交换器，使得该 PCIe 交换器能够正常工作。

[0176] 具体的，在 PCIe 交换器初始化的过程中，主处理器可以为与该 PCIe 交换器相连的各个从处理器分配路由信息。

[0177] S502、主处理器生成路径配置表，该路径配置表包括从处理器标识以及与该从处理器标识对应的路由信息，其中，与该从处理器标识对应的路由信息与为该处理器标识所指示的从处理器分配的路由信息相同。

[0178] S503、主处理器通过网口接收其他单板发送的第一以太网数据包。

[0179] S504、主处理器将该第一以太网数据包转成第一 PCIe 数据包。

[0180] 具体的，主处理器可以将该第一以太网数据包整体封装到第一 PCIe 数据包的净荷中或者主处理器可以提取该第一以太网数据包中的净荷，将该第一以太网数据包中的净荷封装到第一 PCIe 数据包的净荷中；根据该路径配置表及从处理器的从处理器标识获取与该从处理器对应的第一路由信息，将该第一路由信息添加到第一 PCIe 数据包的头部。

[0181] S505、主处理器通过 PCIe 交换器的上行端口将该第一 PCIe 数据包转发给 PCIe 交换器，使得该 PCIe 交换器根据第一 PCIe 数据包头部中的第一路由信息将第一 PCIe 数据包转发给与下行端口相连的从处理器。

[0182] 具体的，主处理器可以通过查表模块 25 查询路径配置表，根据该路径配置表获得一个或多个所述从处理器的路由信息。

[0183] 例如，路径配置表中记录有每一个从处理器的身份信息，该身份信息可以是从处理器所处的总线 ID 或者是该从处理器的 MAC(Medium Access Control，介质访问控制层)

地址，每一个从处理器的身份信息均对应一个预设的路由信息，主处理器在通过 PCIe 交换器转发第一 PCIe 数据包之前，可以通过第一 PCIe 数据包记录的路由信息查找路径配置表，根据该路径配置表获得一个或多个从处理器的路由信息，通过 PCIe 交换器将第一 PCIe 数据包转发给一个或多个从处理器。

[0184] 当然，多个从处理器之间同样可以根据该路径配置表进行数据的交换。例如，在主处理器完成 PCIe 交换器的初始化和从处理器的扫描后生成路径配置表，并将该路径配置表下发到各个从处理器，当第一从处理器需要向第二从处理器转发数据时，可以将该数据包发送至 PCIe 交换器，该转发的数据包中添加有该第二从处理器的身份信息，PCIe 交换器可以查询路径配置表，从而根据第二从处理器的身份信息确定该第二从处理器的路由信息以便完成数据的转发。

[0185] S506、主处理器通过 PCIe 交换器接收从处理器发送加载请求。

[0186] S507、主处理器根据该加载请求读取操作系统和应用程序镜像文件。

[0187] 具体的，当主处理器接收到加载请求之后，主处理器将从存储介质中读取操作系统和应用程序镜像文件。存储介质具体可以包括 Flash（闪存）、SATA(serial advanced technology attachment，串行高级技术附件）、SAS(Serial Attached SCSI，串行连接接口）硬盘或 SSD(solid state disk，固态硬盘），该存储介质可以用于存储操作系统和应用程序的镜像文件

[0188] S508、主处理器将读取到的操作系统镜像文件通过 PCIe 交换器发送至从处理器，使得该从处理器根据操作系统镜像文件启动操作系统。

[0189] 在本发明实施例中，每个从处理器上运行的应用程序可以是动态部署的，用户可针对不同的应用，通过主处理器对各从处理器的工作模式进行配置。从处理器之间是集群关系，各从处理器上运行各自的操作系统和应用程序。主处理器上也运行自己的操作系统和应用程序。比如 :第一个从处理器运行的是计算功能，第二个从处理器运行的是存储功能，这两个从处理器可以相互配合完成一个特定的业务功能。当负责计算的从处理器需要向负责存储的从处理器取数据时，可以直接通过 PCIe 交换器完成数据交换，即本地完成数据交换，这样既可以降低延时又对外部网络的带宽没有影响。

[0190] 这样一来，可以预先将操作系统和应用程序的镜像文件放置在主处理器的存储介质中，当从处理器需要加载时，只需要将该镜像文件从主处理器的存储介质中取出发送至从处理器即可，从而可以使得各个从处理器无需自备操作系统和应用程序，节省了从处理器的内存，提高了从处理器的处理效率。

[0191] S509、主处理器通过 PCIe 交换器向从处理器发送在线检测消息，以检测从处理器是否正常连接。

[0192] S510、当确定一个从处理器连接失败时，主处理器根据预设的发送策略将发送至该从处理器的数据包通过 PCIe 交换器发送至另一个从处理器 ；当确定该从处理器恢复连接时，将发送至另一个从处理器的数据包通过所述 PCIe 交换器转发至所述从处理器。

[0193] 当连接失败时，说明从处理器出现了故障（或者两个主处理器之间的链路故障），因此需要将数据转发到其他的从处理器以对数据进行处理。这里的“预设的发送策略”具体并不限定，例如，可以事先定义好某个从处理器用于处理其他从处理器（或链路）出故障时的数据，如果其他从处理器（或链路）出现故障，则主处理器将数据转给给这个专用的从

处理器来处理；或者“预设的发送策略”也可以是预设的“根据各个从处理器负载情况进行发送”的策略（如将数据包发送到负载少的从处理器）。

[0194] 例如，主处理器与各从处理器之间具有心跳检测机制，主处理器将定时向各从处理器发送信条检测报文以检测从处理器是否正常连接。当第一从处理器的心跳丢失，主处理器将第一从处理器上的业务迁移到第二从处理器或其他备用的从处理器上。

[0195] S511、主处理器通过 PCIe 交换器接收从处理器发送的第二 PCIe 数据包。

[0196] S512、主处理器将该第二 PCIe 数据包转成第二以太网数据包。

[0197] S513、主处理器通过网口向其他单板转发该第二以太网数据包。

[0198] S514、主处理器通过 PCIe 交换器接收其中一个从处理器发送的第三 PCIe 数据包，并通过 PCIe 交换器将该第三 PCIe 数据包转发给另一个从处理器。

[0199] 这样一种服务器单板实现方法，无需采用多级树形结构就可以实现全处理器之间的互连，从而避免了数据在多级处理器之间转发，降低了服务器单板的延时，提高了网络服务质量。

[0200] 本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读取存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0201] 以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以所述权利要求的保护范围为准。

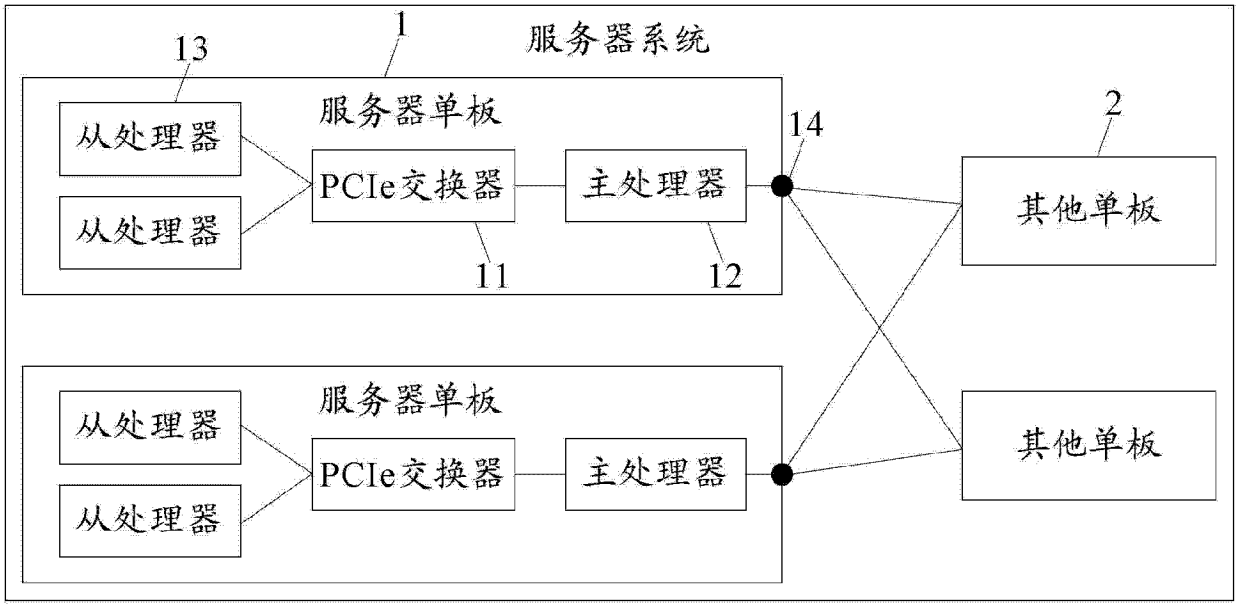

1
服务器系统
13
服务器单板
从处理器
从处理器
PCIe交换器
主处理器
11
12
14
2
其他单板
服务器单板
从处理器
从处理器
PCIe交换器
主处理器
其他单板

图 1

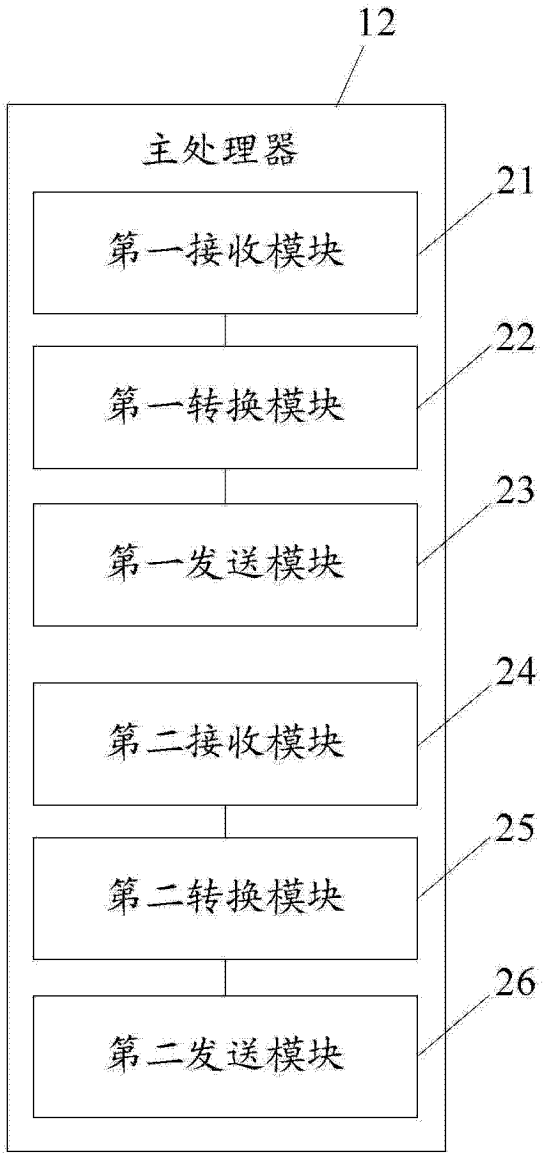

12
主处理器
21
第一接收模块
22
第一转换模块
23
第一发送模块
24
第二接收模块
25
第二转换模块
26
第二发送模块

图 2a

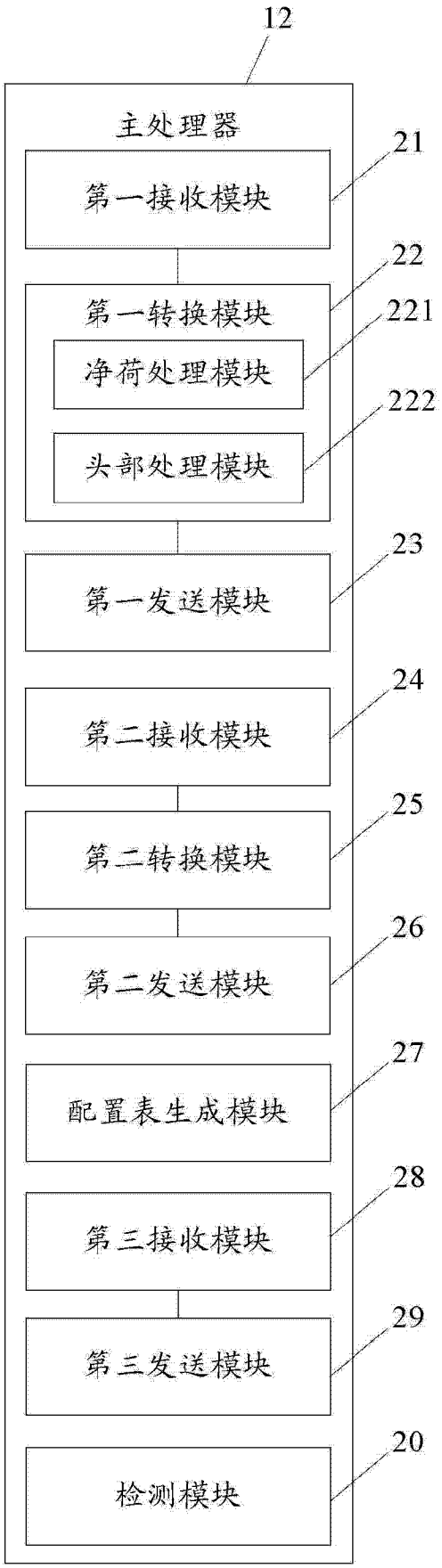

12
主处理器
21
第一接收模块
22
第一转换模块
221
净荷处理模块
222
头部处理模块
23
第一发送模块
24
第二接收模块
25
第二转换模块
26
第二发送模块
27
配置表生成模块
28
第三接收模块
29
第三发送模块
20
检测模块

图 2b

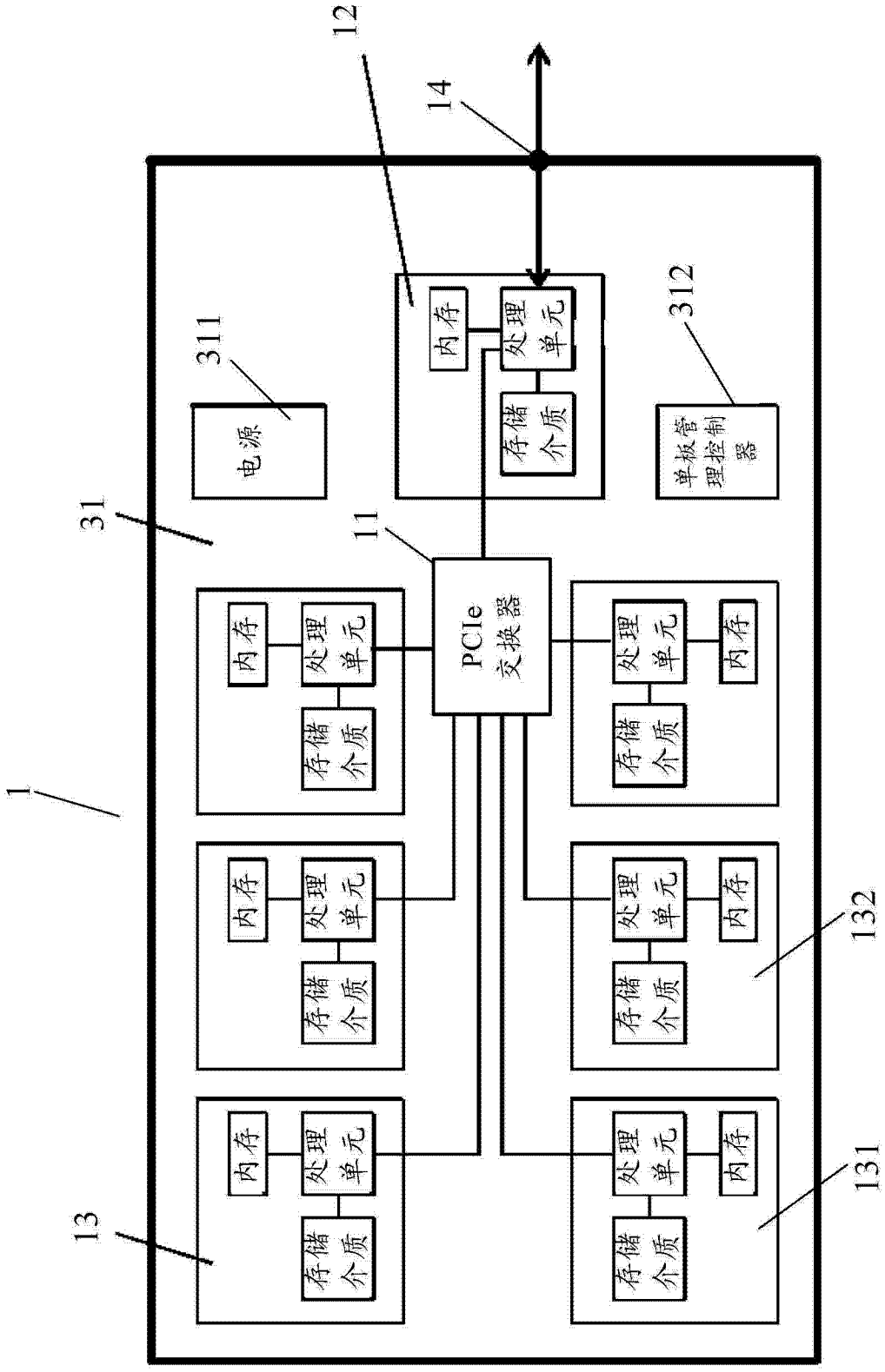


图 3

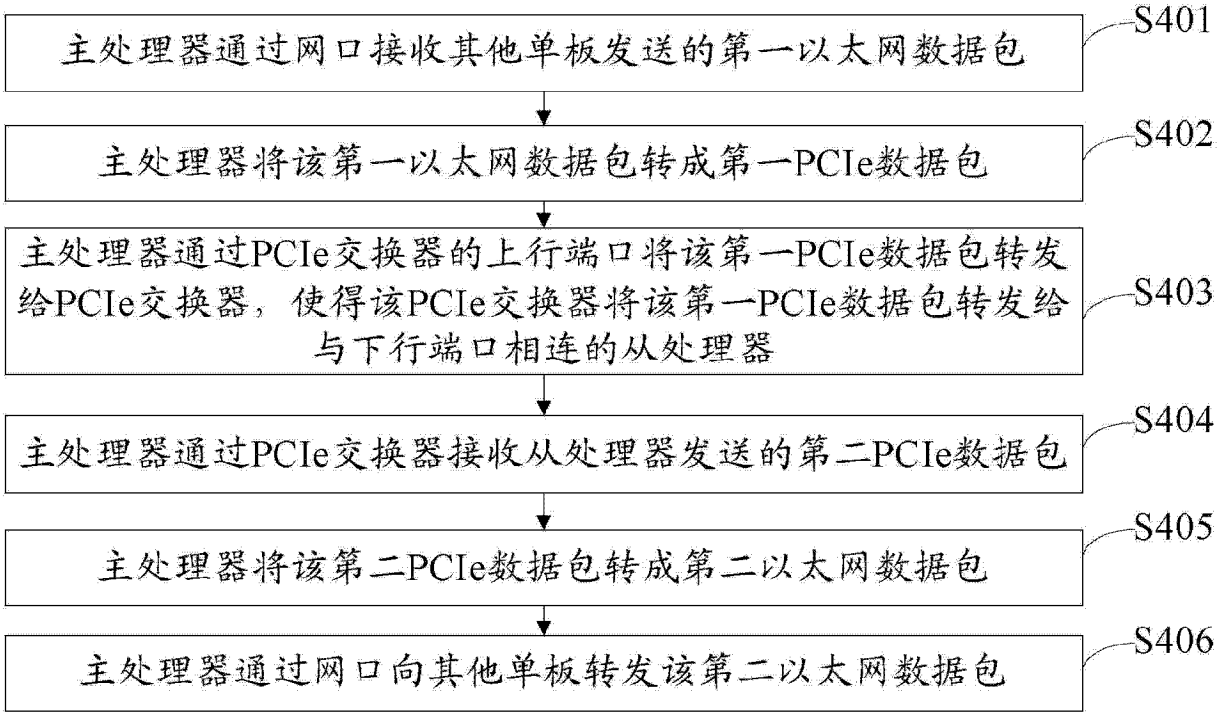

主处理器通过网口接收其他单板发送的第一以太网数据包
S401
主处理器将该第一以太网数据包转成第一PCIe数据包
S402
主处理器通过PCIe交换器的上行端口将该第一PCIe数据包转发给PCIe交换器，使得该PCIe交换器将该第一PCIe数据包转发给与下行端口相连的从处理器
S403
主处理器通过PCIe交换器接收从处理器发送的第二PCIe数据包
S404
主处理器将该第二PCIe数据包转成第二以太网数据包
S405
主处理器通过网口向其他单板转发该第二以太网数据包
S406

图 4

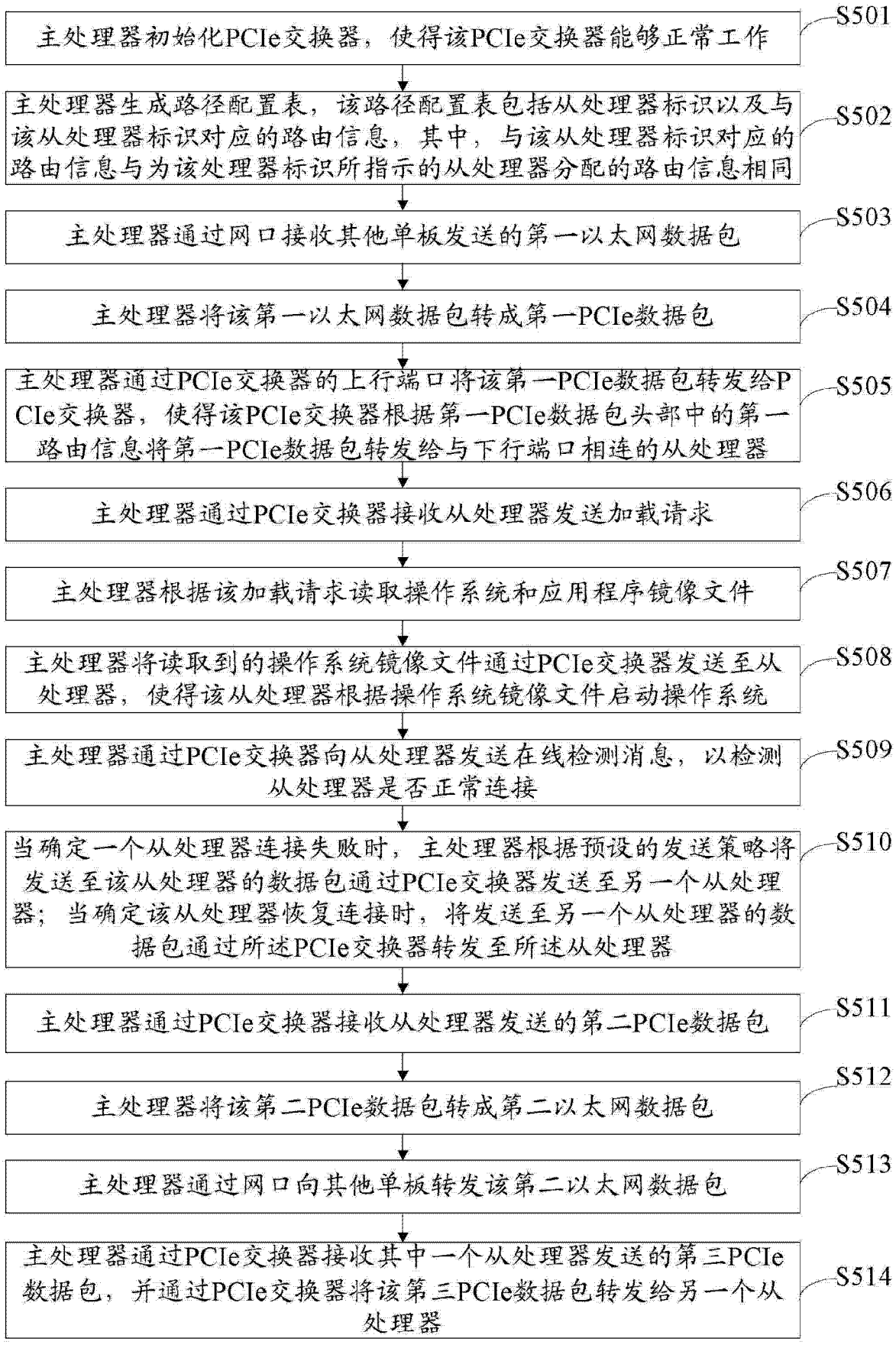

主处理器初始化PCIe交换器，使得该PCIe交换器能够正常工作
S501
主处理器生成路径配置表，该路径配置表包括从处理器标识以及与该从处理器标识对应的路由信息，其中，与该从处理器标识对应的路由信息与为该处理器标识所指示的从处理器分配的路由信息相同
S502
主处理器通过网口接收其他单板发送的第一以太网数据包
S503
主处理器将该第一以太网数据包转成第一PCIe数据包
S504
主处理器通过PCIe交换器的上行端口将该第一PCIe数据包转发给PCIe交换器，使得该PCIe交换器根据第一PCIe数据包头部中的第一路由信息将第一PCIe数据包转发给与下行端口相连的从处理器
S505
主处理器通过PCIe交换器接收从处理器发送加载请求
S506
主处理器根据该加载请求读取操作系统和应用程序镜像文件
S507
主处理器将读取到的操作系统镜像文件通过PCIe交换器发送至从处理器，使得该从处理器根据操作系统镜像文件启动操作系统
S508
主处理器通过PCIe交换器向从处理器发送在线检测消息，以检测从处理器是否正常连接
S509
当确定一个从处理器连接失败时，主处理器根据预设的发送策略将发送至该从处理器的数据包通过PCIe交换器发送至另一个从处理器；当确定该从处理器恢复连接时，将发送至另一个从处理器的数据包通过所述PCIe交换器转发至所述从处理器
S510
主处理器通过PCIe交换器接收从处理器发送的第二PCIe数据包
S511
主处理器将该第二PCIe数据包转成第二以太网数据包
S512
主处理器通过网口向其他单板转发该第二以太网数据包
S513
主处理器通过PCIe交换器接收其中一个从处理器发送的第三PCIe数据包，并通过PCIe交换器将该第三PCIe数据包转发给另一个从处理器
S514

图 5