

一种分布式分级时延提供的名字解析系统构建方法

相关申请

本申请要求名称为“一种分布式分级时延提供的名字解析系统构建方法”、于2021年6月18日提交的中国专利申请号为202110677688.7以及2021年9月27日提交的中国专利申请号为202111137585.8的优先权，在此通过引用包括上述申请。

技术领域

本发明涉及计算机网络技术和分布式系统领域，尤其涉及一种分布式分级时延提供的名字解析系统构建方法。

背景技术

近年来，随着互联网技术的飞速发展，网络设备和网络规模急剧地增长，加剧了网络各方面的复杂性，推动着各个组成部分如通信技术、网络设备技术的进步与革新。反之，随着第五代移动通信技术、信息中心网络等新兴网络技术的出现，网络设备处理器能力的不断提升，网络提供的功能更加丰富和强大。一些对网络性能有严格要求的应用和场景也开始出现，如对时延敏感的工业物联网、车联网和远程医疗等，给相应提供服务的分布式系统提出了所提供服务要有时延保障的新需求。与传统的分布式系统就近、尽力而为的服务不同，在时延保障的分布式系统中，要求每个分布式节点对服务区域内用户的服务请求在一个确定的时延上限内将服务提供，由于不同的应用场景有不同的时延敏感性，需要保障的时延也分为多个等级，分布式系统的组织结构相应地也进行层次型地组织。

在节点规模较大的时候，分布式系统会面临难以部署管理和结构生成的问题。现有的集群技术中分布式节点的结构由一个中心的根节点管理系统进行完整的管理，这样的方式简单直接，但相应地扩展性支持较差；去中心化的对等分布式系统不存在单点瓶颈，对扩展性支持较好，但在结构组织和管理上更加复杂和困难。此外，现有的结构生成方法难以适应分等级时延保障分布式系统特定结构约束的需求。

发明内容

本发明的目的在于克服现有技术缺陷，提出了一种分布式分级时延提供的

名字解析系统构建方法。

为了实现上述目的，本发明提出了一种分布式分级时延提供的名字解析系统构建方法，所述分布式分级时延提供的名字解析系统包括管理系统和多个服务提供节点；所述方法包括：

以服务提供节点间实际往返时延测量获得的节点距离作为依据，通过嵌套关系判定条件确定每个服务提供节点在森林结构中的位置及与其他服务提供节点的连接关系；通过找祖先节点过程确定待加入服务提供节点从根节点到加入位置的父节点之间的所有分支节点；通过找子节点过程将较低服务时延等级的节点加入到待加入服务提供节点下，并随之调整树结构；其中，

森林结构具体包括：森林中的每一棵树代表一个服务区域，一棵树包括的节点层数与具有的服务时延等级对应，在同一棵树中，根节点为具有最大服务时延等级的节点，位于森林中相同层级的节点具有相同的服务时延等级，较高层节点比较低层节点的服务时延等级高；所述节点包括实节点和虚拟节点，在树的构建过程中，采用虚拟节点代替未部署的实节点从而构成完整的树结构。

作为上述方法的一种改进，所述实节点对应实际部署的服务提供节点，为用户提供解析服务和系统结构管理功能；

所述虚拟节点对缺失部署的服务时延等级起占位作用，提供系统结构管理功能；在系统结构中不存在父节点的虚拟节点为虚拟根节点；虚拟节点由一个实节点在加入过程中生成，并由生成该虚拟节点的实节点所部署的服务器进行承载；

虚拟节点与实节点之间的节点距离，用实节点与以该虚拟节点为根的子树中最高层的所有实节点间节点距离的平均值表示。

作为上述方法的一种改进，所述服务时延等级对应一个具体的服务时延，用于表示位于该服务提供节点服务覆盖范围内的用户获取该服务所需时延的上限。

作为上述方法的一种改进，所述服务提供节点服务覆盖范围为与该服务提供节点的距离符合服务时延要求的所有用户节点、该节点的子树节点以及每个子树节点的所有用户节点，位于不同层级树节点的覆盖范围为嵌套关系，较高层节点的覆盖范围包含较低层节点的覆盖范围。

作为上述方法的一种改进，所述嵌套关系判定条件包括：

对于两个节点 p 和 r ，各自的服务时延等级分别为 L_p 和 L_r ，两个节点间距离

的测量时延为 $D(p,r)$, 节点 q 为以 r 为根的子树中存在的最高层的实节点;

当节点 p 和 r 均为实节点时, 如果 $D(p,r) < L_p + L_r$, 则满足实节点等级判定条件;

当节点 p 和 r 均为实节点时, 如果 $L_p > L_r$, 且 $D(p,r) < L_p - L_r$, 则满足实节点同树判定条件;

当节点 p 为实节点, r 为虚拟节点, 如果 $L_p = L_r$, 且对于所有节点 q , 都有 $D(p,q) < L_p - L_q$, 则满足虚拟节点确定性替代判定条件;

当节点 p 为实节点, r 为虚拟节点, 如果 $L_p < L_r$, 且对于所有节点 q , 都有 $D(p,q) < 2 * L_r + L_p - L_q$, 则满足虚拟节点同树判定条件; 或者如果 $L_p > L_r$, 且对于所有节点 q , 都有 $D(p,q) < L_p + L_q$, 则满足虚拟节点等级判定条件;

当节点 p 为实节点, r 为虚拟节点, 如果 $L_p < L_r$, 且存在一个节点 q , 使得 $D(p,q) < |L_q - L_p|$, 则满足虚拟节点同树判定条件; 或者如果 $L_p > L_r$, 且对于所有节点 q , 都有 $D(p,q) < L_p - L_q$, 则满足虚拟节点同树判定条件。

作为上述方法的一种改进, 所述通过找祖先节点过程确定待加入服务提供节点从根节点到加入位置的父节点之间的所有分支节点; 具体包括:

步骤 601) 比较待加入服务提供节点 s 的服务时延等级 L_s 与系统最高服务时延等级 L_{max} : 如果 $L_s = L_{max}$, 转步骤 602); 如果 $L_s < L_{max}$, 转步骤 603);

步骤 602) 根据待加入服务提供节点 s , 从管理系统查询所有虚拟根, 组成集合 VR , 对待加入服务提供节点 s 和集合 VR 中的每个虚拟节点分别进行虚拟节点确定性替代判定, 将满足的虚拟节点加入确定性替代集合 GR , 如果 GR 不为空, 则用待加入服务提供节点 s 替代确定性替代集合 GR 中的所有节点, 转步骤 609);

步骤 603) 根据待加入服务提供节点 s 的已知可服务用户集合 U , 查询已知可服务用户集合 U 中所有用户当前所属的根节点, 组成集合 UR , 将集合 UR 中与待加入服务提供节点 s 不满足实节点等级判定条件的节点删除, 选定其中对应用户的优先级最高的根节点 ur , 删除根节点不是 ur 的用户, 以 ur 为待加入服务提供节点 s 的根, 转步骤 606); 如果无法选出满足的根节点 ur , 转步骤 604);

步骤 604) 根据待加入服务提供节点 s , 从管理系统查询所有实根, 组成集合 R , 对待加入服务提供节点 s 和集合 R 中的每个实节点进行实节点同树判定, 如果存在满足判定条件的实节点 r , 则待加入服务提供节点 s 以实节点 r 作为根,

转步骤 606); 如果无满足判定条件的实节点 r , 则对待加入服务提供节点 s 与 R 中的每个节点分别进行实节点等级判定, 将不满足判定条件的节点从集合 R 中删除, 并按与待加入服务提供节点 s 间时延从小到大的顺序对集合 R 进行排序, 依序遍历集合 R 取出实节点 r , 如果实节点 r 的负载未超过预设的负载阈值, 则以该实节点 r 为待加入服务提供节点 s 的根, 记当前节点 V_{cur} 为 r , 转步骤 606); 如果无法选出满足判断条件的实节点 r , 转步骤 605);

步骤 605) 根据待加入服务提供节点 s , 从管理系统查询所有虚拟根, 组成集合 VR , 对待加入服务提供节点 s 和集合 VR 中的每个虚拟节点分别进行虚拟节点同树判定, 如果存在满足判定条件的虚拟节点 vr , 则以虚拟节点 vr 为待加入服务提供节点 s 的根, 转步骤 606); 如果无满足判定条件的虚拟节点 vr , 则对待加入服务提供节点 s 与集合 VR 中的每个虚拟节点分别进行虚拟节点等级判定, 删除不满足的虚拟节点, 从集合 VR 中选取与待加入服务提供节点 s 间时延最小的虚拟节点 vr 作为待加入服务提供节点 s 的根, 记当前节点 V_{cur} 为 vr , 转步骤 606); 如果无满足判定条件的虚拟节点 vr , 则待加入服务提供节点 s 生成一个新的虚拟根节点, 并生成中间服务时延等级对应的所有虚拟节点, 并令这些虚拟节点按等级建立对应的父子关系, 转步骤 609);

步骤 606) 待加入服务提供节点 s 记录当前节点 V_{cur} 为该层的祖先节点, 待加入服务提供节点 s 从 V_{cur} 查询其所有实子节点, 组成集合 RC , 对待加入服务提供节点 s 和 RC 中的每个实节点进行实节点同树判定, 如果存在满足判定条件的实节点 rc , 则令 V_{cur} 等于 rc , 转步骤 608); 如果无满足判定条件的实节点, 则对待加入服务提供节点 s 与 RC 中的每个节点分别进行实节点等级判定, 将不满足判定条件的节点从集合 RC 中删除, 并按与待加入服务提供节点 s 间时延从小到大的顺序对集合 RC 进行排序, 依序遍历集合 RC 取出实节点 rc , 如果实节点 rc 的负载未超过预设的负载阈值, 则令 V_{cur} 等于 rc , 转步骤 606); 如果无法选出满足判断条件的实节点 rc , 转步骤 607);

步骤 607) 待加入服务提供节点 s 从 V_{cur} 查询其所有虚拟子节点, 组成集合 VC , 对待加入服务提供节点 s 和集合 VC 中的每个虚拟节点分别进行虚拟节点同树判定, 如果存在满足判定条件的虚拟节点 vc , 则令 V_{cur} 等于 vc , 转步骤 608); 如果无满足判定条件的虚拟节点, 则对待加入服务提供节点 s 与集合 VC 中的每个虚拟节点分别进行虚拟节点等级判定, 删除不满足的虚拟节点, 从集合 VC 中选取与待加入服务提供节点 s 间时延最小的 vc , 令 V_{cur} 等于 vc ,

转步骤 608); 如果无满足判定条件的虚拟节点, 则待加入服务提供节点 s 生成一个新的虚拟节点, 并生成中间服务时延等级对应的所有虚拟节点, 并令这些虚拟节点按等级建立对应的父子关系, 转步骤 609);

步骤 608) 如果待加入服务提供节点 s 的服务时延等级比 V_{cur} 的服务时延等级低一级, 则转步骤 609); 否则转步骤 606);

步骤 609) 待加入服务提供节点 s 完成找祖先节点的过程。

作为上述方法的一种改进, 所述通过找子节点过程将较低服务时延等级的节点加入到待加入服务提供节点下, 并随之调整树结构; 具体包括:

步骤 701) 如果父节点 f 为实节点, 且父节点 f 无子节点, 则待加入服务提供节点 s 的子节点为空, s 为叶子节点, 转步骤 706); 否则转步骤 702);

步骤 702) 如果父节点 f 为虚拟节点, 且父节点 f 无子节点, 则待加入服务提供节点 s 的子节点为空, s 为叶子节点, 转步骤 706); 否则转步骤 703);

步骤 703) 根据待加入服务提供节点 s , 查询得到父节点 f 的所有虚拟子节点, 组成虚拟兄弟节点集合 SV , 如果 SV 为空, 则待加入服务提供节点 s 的子节点为空, 转步骤 706); 否则, 转步骤 704);

步骤 704) 根据待加入服务提供节点 s , 查询虚拟兄弟节点集合 SV 中每个虚拟兄弟节点的子节点, 组成待选子节点集合 CS , 对待加入服务提供节点 s 与待选子节点集合 CS 中的每个节点进行虚拟节点同树判定, 将满足判定条件的虚拟节点加入确定性子节点集合 GCS , 待加入服务提供节点 s 向确定性子节点集合 GSS 中的所有节点发起过继邀请请求, 对于回复拒绝的节点, 从确定性子节点集合 GSS 中删除, 如果 GSS 不为空, 则待加入服务提供节点 s 将确定性子节点集合 GSS 的所有节点过继到自身下, 转步骤 705);

步骤 705) 待加入服务提供节点 s 检测每个被过继子节点的原父节点, 如果其因子节点被过继而成为一个虚拟叶子节点, 则其向上循环删除虚拟叶子节点, 直至该分支不存在虚拟叶子节点, 转步骤 706);

步骤 706) 找子节点过程结束。

作为上述方法的一种改进, 所述待加入服务提供节点 s 将确定性子节点集合 GSS 的所有节点过继到自身下; 具体包括:

将确定性子节点集合 GSS 中选定的子节点的父节点改为待加入服务提供节点 s , 并通知选定子节点的原父节点, 从其子节点列表中删除该子节点。

作为上述方法的一种改进, 所述管理系统采用弱中心化的部署, 接受有限

节点上报信息，其中，作为根节点的服务提供节点上报的信息，包括但不限于节点的标识、网络地址、地理位置、节点虚实类型和节点状态；提供根节点查询服务，但不执行复杂的拓扑结构的计算和分析。

与现有技术相比，本发明的优势在于：

与现有技术相比，本发明公开的一种分布式分级时延提供的名字解析系统构建方法，采用去中心化的结构，分布式服务节点自组织地结构生成避免了结构管理的单点瓶颈，具备更好的扩展性；基于节点间实际测距的节点加入能够更加准确地实现系统时延保障的约束，节点加入过程中灵活的节点间关系判定方法也使系统结构具有更好的均衡性和对用户更高的覆盖率，引入虚拟节点提升了系统结构的完整性和连通性。

附图说明

图1是本发明的分布式分级时延提供的名字解析系统构建方法流程图；

图2是本发明的选择根节点流程图；

图3是本发明的找祖先节点流程图；

图4是本发明的找子节点流程图。

具体实施方式

本发明的方法是采用去中心化的管理，节点分布式自组织地进行局部结构生成，支持从第一个点开始的、通用的增量部署场景，基于节点间时延测量进行划分，保障生成的系统结构满足提供服务不同等级的时延保障，更好地满足时延敏感应用对使用服务的需求。

分级时延提供的名字解析系统由一个管理系统和多个服务提供节点组成，所有服务提供节点构成一个森林结构，森林中的每一棵树代表一个服务区域，树中的每一个节点对应一个服务时延等级下为用户提供解析服务的覆盖范围；在同一颗树中，同一层的节点具有相同的服务时延等级，上一级服务节点所服务的用户集合包含与其具有嵌套关系的下一级服务节点所服务用户的全集；分布式分级时延提供的名字解析系统具有支持逐步部署和构建的特点，对于某一个服务时延等级的服务提供节点尚未实际部署，而其覆盖范围内包含的较低服务时延等级的服务提供节点已部署的情况，采用虚拟节点代替该未部署节点从而构成完整的树结构；在构建过程中，以服务提供节点间实际往返时延测量获

得的节点距离作为依据，通过嵌套关系判定条件来确定每个节点在森林中的位置及与其他节点的连接关系。具体地，构建过程包括节点加入的找祖先节点和找子节点两个过程。其中，找祖先节点指确定新加入节点从根到该节点加入位置父节点之间的所有分支节点的过程；找子节点指将系统中可能的较低服务时延等级的节点加入到新加入节点下并随之调整树结构的过程。

下面结合附图和实施例对本发明的技术方案进行详细的说明。

实施例 1

在本实施例中，以在网络中部署现场名字解析系统作为主要场景进行描述，现场名字解析系统是一种分布式分级时延提供的名字解析系统，能够在多个时延等级上给用户服务的分布式查询服务系统，系统由多个分布式部署的解析服务节点组成，每个解析服务节点为其覆盖区域内的用户所提供解析服务的通信时延小于预定的时延上限。

覆盖区域是由该服务节点所服务的所有用户组成，一个服务节点覆盖某个用户，表示该服务节点可以为这个用户提供其服务时延等级所对应的通信时延上限的查询服务；反之，一个用户在某个服务时延等级上的可服务节点，表示可以为该用户在该服务时延等级提供查询服务的所有节点。

系统中的分布式解析服务节点构成一个森林结构，解析服务节点在森林中树的层级与该节点服务时延等级相对应，森林结构中的节点分为实节点和虚拟节点，实节点执行的功能包括结构管理和用户服务的提供，而虚拟节点只执行结构管理功能；每个节点的结构信息包括但不限于父节点指针和子节点列表。

位于森林中相同层级的节点具有相同的服务时延等级，不同层级的节点具有不同的服务时延等级，且具有较高服务时延上限的服务时延等级对应森林结构中较靠近根节点的层级。

在现场名字解析系统中，一个用户在相同服务时延等级上可以有一个或多个可解析服务节点，但其中有且只有一个节点为主节点，主节点即为节点加入过程中每个层级所选择的祖先节点；覆盖某个用户的各个服务时延等级的所有主节点需要满足嵌套关系，即一个解析服务节点所覆盖用户的集合必须是其所有祖先节点所覆盖用户的子集。

如图 1 所示，基于该场景，本发明提供了一种分等级时延保障服务系统的结构生成方法，系统的结构是由每个服务节点通过自组织的方式协同生成的，方法包括：首先待加入节点配置有服务时延等级，通过入口查询节点获取当前

森林结构的根节点列表，对其中的虚拟根节点进行替代；然后从根节点列表开始，待加入节点通过由根到叶的方向在每个层级选择一个祖先节点，逐层迭代地，在当前层级祖先节点的子节点中进行下一层级祖先节点的选择，直到找到父节点后结束迭代；最后，待加入节点从父节点处获取其虚拟子节点列表，分别将该列表中节点的子节点尝试进行过继，至此完成新节点的加入；其中，祖先节点的选择和子节点进行过继通过时延测量技术和节点关系判定约束进行确定。

在上述技术方案中，所述的虚拟节点是一种在结构生成中出于完整性和连通性需求提供占位功能的逻辑节点，虚拟节点在实节点加入过程中生成，与该实节点位于同一台服务器。

所述的结构管理功能，指位于系统中每个分布式节点上的功能模块，在分布式节点的加入过程中进行系统结构信息的获取、存储和更新，以及分布式节点在构建系统结构时的节点间通信和协同决策。其中，系统结构信息包括但不限于：父节点指针、子节点列表；

在上述技术方案中，所述的根节点管理系统，指位于分等级时延保障服务系统外、可以为待加入节点提供现有系统结构根节点入口查询功能的节点，该系统仅提供根节点的查询服务，而不执行复杂的拓扑结构计算和节点加入位置指示；提供的根节点信息包括但不限于节点标识、网络地址、地理位置和节点虚实类型；根节点信息通过部署后作为根的服务节点通过通知报文进行注册。

在现场名字解析系统的结构生成过程中采用时延测量技术，节点在加入过程中对待测节点进行两个节点真实往返通信时延的测量。节点间时延使用如下定义：如果两个节点均为实节点，则节点间时延为这两个节点间的通信时延；如果其中一个节点为虚拟节点，则节点间时延为实节点到该虚拟节点下等级最高的所有实节点通信时延的平均值。

祖先节点，指位于当前服务节点与其根节点连接路径上，所有的比当前节点服务等级更高的所有服务节点。所述的找祖先节点过程中，如果在一个层级有多个节点满足节点间关系判定约束，则优先选择确定性嵌套的节点，其次在可能性嵌套的节点中选择节点间时延最小的节点作为祖先节点；如果没有任何节点满足节点间关系判定约束，则新加入节点在该节点在该层级生成一个虚拟节点并将这个虚拟节点作为该层级的祖先节点。

如图 2 所示，对于根节点的选择过程，包括下述步骤：

步骤 101) 比较新加入节点 s 预设的时延等级 L_s 与系统最高时延等级 L_r , s 的时延等级对应时延值为 T_s : 如果 $L_s = L_r$, 转步骤 102); 如果 $L_s < L_r$, 转步骤 103);

步骤 102) s 从根节点管理系统处查询所有虚拟根组成的集合 VR , s 对 VR 中每个节点进行确定性可替代判定, 将满足判定的节点加入确定性可替代集合 GR : 如果 GR 不为空, 则用 s 替代 GR 中的所有节点, 节点 s 完成加入; 如果 GR 为空, 则 s 作为一个新的根节点加入;

步骤 103) s 下属用户集合为 U , 查询 U 中所有用户当前所属的根节点 UR , 选取 UR 中距离 s 时延最短的 ur , 如果该 ur 与 s 间时延小于 $T_{ur} + T_s$, 则选定 ur 为根节点, 将不以 ur 为根节点的用户过滤, 完成根节点选择过程。如果无法选出满足的 ur , 转步骤 104);

步骤 104) s 从根节点管理系统处查询所有实根节点 组成的集合 R , 对于 R 中的每个节点 r , 根据其到 s 间的时延从小到大排序, 依序遍历 R , 如果存在一个 r 与 s 间的时延小于 $T_r - T_s$, 则选定 r 为根, 完成根节点选择过程。如果无法选出满足的 r , 则依序遍历 R , 如果存在一个 r 与 s 间的时延小于 $T_r + T_s$, 则选定 r 为根, 完成根节点选择过程。如果无法选出满足的 r , 则转步骤 105);

步骤 105) s 从根节点管理系统处查询所有虚拟根组成的集合 VR , 对于 VR 中的每个节点 vr , 根据其到 s 间的时延从小到大排序, 依序遍历 VR , 如果存在一个 vr 与 s 间的时延小于 $T_{vr} - T_s$, 则选定 vr 为根, 完成根节点选择过程。如果无法选出满足的 vr , 则依序遍历 VR , 对于 vr 的子树中的任一节点 q , 都满足 q 与 s 间的距离小于 $2 * L_r + L_s - L_q$, 则选定 vr 为根, 完成根节点选择过程。如果无法选出满足的 vr , 则 s 生成一个新的虚拟根。

如图 3 所示, 对于找祖先节点过程, 包括下述步骤:

步骤 201) s 将选定的根节点 r 作为当前搜索分支的入口;

步骤 202) 如果 s 的时延等级比当前搜索分支时延等级低一级, 转步骤 205), 否则转步骤 203);

步骤 203) s 向搜索分支入口节点查询其下所属的所有实子节点加入待搜索集合 P , 对于 P 中的每个节点 p , 根据其到 s 的时延从小到大排序, 依序遍历 P , 如果存在一个 p 与 s 间的时延小于 $T_p - T_s$, 则选定 p 作为后续搜索分支的入口。如果无法选出满足的 p , 则依序遍历 P , 如果存在一个 p 与 s 间的时延小于 $T_p + T_s$, 则选定 p 作为后续搜索分支的入口, 转步骤 202)。如果无法选出满足的 p , 则

转步骤 204)；

步骤 204) s 向搜索分支入口节点查询其下所属的所有虚拟子节点加入待搜索集合 PV，对于 PV 中的每个节点 pv，根据其到 s 的时延从小到大排序，依序遍历 PV，如果存在一个 pv 与 s 间的时延小于 $T_{pv}-T_s$ ，则选定 pv 作为后续搜索分支的入口。如果无法选出满足的 pv，则依序遍历 PV，如果存在一个 pv 与 s 间的时延小于 $T_{pv}+T_s$ ，则选定 pv 作为后续搜索分支的入口，转步骤 202)。如果无法选出满足的 pv，则 s 生成一个新的虚拟节点并加入当前搜索分支入口下。

步骤 205) 完成找祖先节点过程。

如图 4 所示，对于找子节点过程，包括下述步骤：

步骤 301) s 向选定的父节点 rf 查询 rf 的所有虚拟子节点加入虚拟兄弟节点集合 SV，如果 SV 为空，转 304)。否则转 302)；

步骤 302) s 向 SV 中每个节点 sv 查询 sv 的子节点，并加入待选子节点集合 CS 中，对于 CS 中的每个节点 cs，根据其到 s 的时延从小到大排序，依序遍历 CS，找到所有满足到 s 时延小于 T_s-T_{cs} 的 cs，将其加入确定性子节点集合 GCS 中。如果 GCS 不为空，则 s 将 GCS 中的所有节点由其原父节点处过继；如果 GCS 为空，则转 304)；

步骤 303) s 检测每个被过继的子节点的原父节点 pp，如果 pp 因子节点被过继走而成为虚拟叶子节点，则向上循环删除虚拟叶子节点，直至不存在虚拟叶子节点。

步骤 304) s 完成子节点选择过程，完成部署。

最后所应说明的是，以上实施例仅用以说明本发明的技术方案而非限制。尽管参照实施例对本发明进行了详细说明，本领域的普通技术人员应当理解，对本发明的技术方案进行修改或者等同替换，都不脱离本发明技术方案的精神和范围，其均应涵盖在本发明的权利要求范围当中。

权 利 要 求

1、一种分布式分级时延提供的名字解析系统构建方法，所述分布式分级时延提供的名字解析系统包括管理系统和多个服务提供节点；所述方法包括：

以服务提供节点间实际往返时延测量获得的节点距离作为依据，通过嵌套关系判定条件确定每个服务提供节点在森林结构中的位置及与其他服务提供节点的连接关系；通过找祖先节点过程确定待加入服务提供节点从根节点到加入位置的父节点之间的所有分支节点；通过找子节点过程将较低服务时延等级的节点加入到待加入服务提供节点下，并随之调整树结构；其中，

森林结构具体包括：森林中的每一棵树代表一个服务区域，一棵树包括的节点层数与具有的服务时延等级对应，在同一棵树中，根节点为具有最大服务时延等级的节点，位于森林中相同层级的节点具有相同的服务时延等级，较高层节点比较低层节点的服务时延等级高；所述节点包括实节点和虚拟节点，在树的构建过程中，采用虚拟节点代替未部署的实节点从而构成完整的树结构。

2、根据权利要求1所述的分布式分级时延提供的名字解析系统构建方法，其特征在于，

所述实节点对应实际部署的服务提供节点，为用户提供解析服务和系统结构管理功能；

所述虚拟节点对缺失部署的服务时延等级起占位作用，提供系统结构管理功能；在系统结构中不存在父节点的虚拟节点为虚拟根节点；虚拟节点由一个实节点在加入过程中生成，并由生成该虚拟节点的实节点所部署的服务器进行承载；

虚拟节点与实节点之间的节点距离，用实节点与以该虚拟节点为根的子树中最高层的所有实节点间节点距离的平均值表示。

3、根据权利要求1或2所述的分布式分级时延提供的名字解析系统构建方法，其特征在于，所述服务时延等级对应一个具体的服务时延，用于表示位于该服务提供节点服务覆盖范围内的用户获取该服务所需时延的上限。

4、根据权利要求3所述的分布式分级时延提供的名字解析系统构建方法，其特征在于，所述服务提供节点服务覆盖范围为与该服务提供节点的距离符合服务时延要求的所有用户节点、该节点的子树节点以及每个子树节点的所有用户节点，位于不同层级树节点的覆盖范围为嵌套关系，较高层节点的覆盖范围

包含较低层节点的覆盖范围。

5、根据权利要求4所述的分布式分级时延提供的名字解析系统构建方法，其特征在于，所述嵌套关系判定条件包括：

对于两个节点 p 和 r ，各自的服务时延等级分别为 L_p 和 L_r ，两个节点间距离的测量时延为 $D(p,r)$ ，节点 q 为以 r 为根的子树中存在的最高层的实节点；

当节点 p 和 r 均为实节点时，如果 $D(p,r) < L_p + L_r$ ，则满足实节点等级判定条件；

当节点 p 和 r 均为实节点时，如果 $L_p > L_r$ ，且 $D(p,r) < L_p - L_r$ ，则满足实节点同树判定条件；

当节点 p 为实节点， r 为虚拟节点，如果 $L_p = L_r$ ，且对于所有节点 q ，都有 $D(p,q) < L_p - L_q$ ，则满足虚拟节点确定性替代判定条件；

当节点 p 为实节点， r 为虚拟节点，如果 $L_p < L_r$ ，且对于所有节点 q ，都有 $D(p,q) < 2 * L_r + L_p - L_q$ ，则满足虚拟节点同树判定条件；或者如果 $L_p > L_r$ ，且对于所有节点 q ，都有 $D(p,q) < L_p + L_q$ ，则满足虚拟节点等级判定条件；

当节点 p 为实节点， r 为虚拟节点，如果 $L_p < L_r$ ，且存在一个节点 q ，使得 $D(p,q) < |L_q - L_p|$ ，则满足虚拟节点同树判定条件；或者如果 $L_p > L_r$ ，且对于所有节点 q ，都有 $D(p,q) < L_p - L_q$ ，则满足虚拟节点同树判定条件。

6、根据权利要求5所述的分布式分级时延提供的名字解析系统构建方法，其特征在于，所述通过找祖先节点过程确定待加入服务提供节点从根节点到加入位置的父节点之间的所有分支节点；具体包括：

步骤601) 比较待加入服务提供节点 s 的服务时延等级 L_s 与系统最高服务时延等级 L_{max} ：如果 $L_s = L_{max}$ ，转步骤602)；如果 $L_s < L_{max}$ ，转步骤603)；

步骤602) 根据待加入服务提供节点 s ，从管理系统查询所有虚拟根，组成集合 VR ，对待加入服务提供节点 s 和集合 VR 中的每个虚拟节点分别进行虚拟节点确定性替代判定，将满足的虚拟节点加入确定性替代集合 GR ，如果 GR 不为空，则用待加入服务提供节点 s 替代确定性替代集合 GR 中的所有节点，转步骤609)；

步骤603) 根据待加入服务提供节点 s 的已知可服务用户集合 U ，查询已知可服务用户集合 U 中所有用户当前所属的根节点，组成集合 UR ，将集合 UR 中与待加入服务提供节点 s 不满足实节点等级判定条件的节点删除，选定其中对应用户的优先级最高的根节点 ur ，删除根节点不是 ur 的用户，以 ur 为待加

入服务提供节点 s 的根, 转步骤 606); 如果无法选出满足的根节点 ur , 转步骤 604);

步骤 604) 根据待加入服务提供节点 s , 从管理系统查询所有实根, 组成集合 R , 对待加入服务提供节点 s 和集合 R 中的每个实节点进行实节点同树判定, 如果存在满足判定条件的实节点 r , 则待加入服务提供节点 s 以实节点 r 作为根, 转步骤 606); 如果无满足判定条件的实节点 r , 则对待加入服务提供节点 s 与 R 中的每个节点分别进行实节点等级判定, 将不满足判定条件的节点从集合 R 中删除, 并按与待加入服务提供节点 s 间时延从小到大的顺序对集合 R 进行排序, 依序遍历集合 R 取出实节点 r , 如果实节点 r 的负载未超过预设的负载阈值, 则以该实节点 r 为待加入服务提供节点 s 的根, 记当前节点 V_{cur} 为 r , 转步骤 606); 如果无法选出满足判断条件的实节点 r , 转步骤 605);

步骤 605) 根据待加入服务提供节点 s , 从管理系统查询所有虚拟根, 组成集合 VR , 对待加入服务提供节点 s 和集合 VR 中的每个虚拟节点分别进行虚拟节点同树判定, 如果存在满足判定条件的虚拟节点 vr , 则以虚拟节点 vr 为待加入服务提供节点 s 的根, 转步骤 606); 如果无满足判定条件的虚拟节点 vr , 则对待加入服务提供节点 s 与集合 VR 中的每个虚拟节点分别进行虚拟节点等级判定, 删除不满足的虚拟节点, 从集合 VR 中选取与对待加入服务提供节点 s 间时延最小的虚拟节点 vr 作为待加入服务提供节点 s 的根, 记当前节点 V_{cur} 为 vr , 转步骤 606); 如果无满足判定条件的虚拟节点 vr , 则待加入服务提供节点 s 生成一个新的虚拟根节点, 并生成中间服务时延等级对应的所有虚拟节点, 并令这些虚拟节点按等级建立对应的父子关系, 转步骤 609);

步骤 606) 待加入服务提供节点 s 记录当前节点 V_{cur} 为该层的祖先节点, 待加入服务提供节点 s 从 V_{cur} 查询其所有实子节点, 组成集合 RC , 对待加入服务提供节点 s 和 RC 中的每个实节点进行实节点同树判定, 如果存在满足判定条件的实节点 rc , 则令 V_{cur} 等于 rc , 转步骤 608); 如果无满足判定条件的实节点, 则对待加入服务提供节点 s 与 RC 中的每个节点分别进行实节点等级判定, 将不满足判定条件的节点从集合 RC 中删除, 并按与待加入服务提供节点 s 间时延从小到大的顺序对集合 RC 进行排序, 依序遍历集合 RC 取出实节点 rc , 如果实节点 rc 的负载未超过预设的负载阈值, 则令 V_{cur} 等于 rc , 转步骤 606); 如果无法选出满足判断条件的实节点 rc , 转步骤 607);

步骤 607) 待加入服务提供节点 s 从 V_{cur} 查询其所有虚拟子节点, 组成集

合 VC, 对待加入服务提供节点 s 和集合 VC 中的每个虚拟节点分别进行虚拟节点同树判定, 如果存在满足判定条件的虚拟节点 vc, 则令 Vcur 等于 vc, 转步骤 608); 如果无满足判定条件的虚拟节点, 则对待加入服务提供节点 s 与集合 VC 中的每个虚拟节点分别进行虚拟节点等级判定, 删除不满足的虚拟节点, 从集合 VC 中选取与待加入服务提供节点 s 间时延最小的 vc, 令 Vcur 等于 vc, 转步骤 608); 如果无满足判定条件的虚拟节点, 则待加入服务提供节点 s 生成一个新的虚拟节点, 并生成中间服务时延等级对应的所有虚拟节点, 并令这些虚拟节点按等级建立对应的父子关系, 转步骤 609);

步骤 608) 如果待加入服务提供节点 s 的服务时延等级比 Vcur 的服务时延等级低一级, 则转步骤 609); 否则转步骤 606);

步骤 609) 待加入服务提供节点 s 完成找祖先节点的过程。

7、根据权利要求 6 所述的分布式分级时延提供的名字解析系统构建方法, 其特征在于, 所述通过找子节点过程将较低服务时延等级的节点加入到待加入服务提供节点下, 并随之调整树结构; 具体包括:

步骤 701) 如果父节点 f 为实节点, 且父节点 f 无子节点, 则待加入服务提供节点 s 的子节点为空, s 为叶子节点, 转步骤 706); 否则转步骤 702);

步骤 702) 如果父节点 f 为虚拟节点, 且父节点 f 无子节点, 则待加入服务提供节点 s 的子节点为空, s 为叶子节点, 转步骤 706); 否则转步骤 703);

步骤 703) 根据待加入服务提供节点 s, 查询得到父节点 f 的所有虚拟子节点, 组成虚拟兄弟节点集合 SV, 如果 SV 为空, 则待加入服务提供节点 s 的子节点为空, 转步骤 706); 否则, 转步骤 704);

步骤 704) 根据待加入服务提供节点 s, 查询虚拟兄弟节点集合 SV 中每个虚拟兄弟节点的子节点, 组成待选子节点集合 CS, 对待加入服务提供节点 s 与待选子节点集合 CS 中的每个节点进行虚拟节点同树判定, 将满足判定条件的虚拟节点加入确定性子节点集合 GCS, 待加入服务提供节点 s 向确定性子节点集合 GSS 中的所有节点发起过继邀请请求, 对于回复拒绝的节点, 从确定性子节点集合 GSS 中删除, 如果 GSS 不为空, 则待加入服务提供节点 s 将确定性子节点集合 GSS 的所有节点过继到自身下, 转步骤 705);

步骤 705) 待加入服务提供节点 s 检测每个被过继子节点的原父节点, 如果其因子节点被过继而成为一个虚拟叶子节点, 则其向上循环删除虚拟叶子节点, 直至该分支不存在虚拟叶子节点, 转步骤 706);

步骤 706) 找子节点过程结束。

8、根据权利要求 7 所述的分布式分级时延提供的名字解析系统构建方法，其特征在于，所述待加入服务提供节点 s 将确定性子节点集合 GSS 的所有节点过继到自身下；具体包括：

将确定性子节点集合 GSS 中选定的子节点的父节点改为待加入服务提供节点 s，并通知选定子节点的原父节点，从其子节点列表中删除该子节点。

9、根据权利要求 8 所述的分布式分级时延提供的名字解析系统构建方法，其特征在于，所述管理系统采用弱中心化的部署，接受有限节点上报信息，其中，作为根节点的服务提供节点上报的信息，包括节点的标识、网络地址、地理位置、节点虚实类型和节点状态；提供根节点查询服务，但不执行复杂的拓扑结构的计算和分析。