

# 基于 SkipNet 的 P2P-SIP 系统

朴冬临

北京邮电大学计算机科学与技术学院, 北京 (100876)

Email: pudonglin@gmail.com

**摘要:** 本文首次提出了一种基于跳图 (SkipNet) 路由算法的 P2P-SIP 网络系统, 充分利用 SkipNet 的局部特性, 保证本地 P2P-SIP 的网络结点仅在本地路由。首先介绍了 P2P 网络、SIP 协议的基本概念, 其次介绍 SkipNet 算法的基本理论, 最后介绍 P2P SkipNet 网络进行的 SIP 协议的扩展定义以及系统的主要工作流程。

**关键词:** P2P, SIP, P2P-SIP, SkipNet

## 1. 引言

在过去的几年, 曾经飞速发展的电信业正在逐渐失去往日的辉煌, 虽然电信业务的收入仍在增长, 但增速已经开始下降, ARPU (Average Revenue Per User) 值也呈下降趋势。在这种局面下, 人们需要能刺激电信业发展的新业务与新应用产生, 这种新业务应该是建立在开放平台上的由众多内容提供商支持的个性化多媒体服务, 而 SIP 协议的出现使这种新兴业务平台的出现成为可能。

但是, 尽管 SIP 协议具有可扩展性强, 独立于底层网络, 并支持终端移动性等优点, 它仍然存在着一些不足, 如单点失效问题和网络性能瓶颈问题, 而这些问题产生的主要原因都是 SIP 协议对于服务器过于依赖, P2P-SIP 系统的提出正是为了解决这个问题。P2P-SIP 系统利用 P2P 技术代替 SIP 系统中的网络服务器, 将现有的 C/S 架构的 SIP 应用系统改造成基于 P2P 模式的对等网络, 并取消现有的包括注册、代理和重定向在内的各种 SIP 服务器。利用 P2P 实现用户的注册、查找定位和信令路由功能。每个 SIP 终端都是一个 Peer, 既作为服务器, 也作为客户端, 有效改善了原有 SIP 系统的性能。本文主要讨论如何对 SIP 协议进行扩展, 以实现基于 SkipNet 路由算法的 P2P-SIP 系统。

## 2. P2P技术

### 2.1 P2P 技术简介

P2P 是一种分布式网络, 网络的参与者共享他们所拥有的一部分硬件资源 (处理能力、存储能力、网络连接能力、打印机等), 这些共享资源需要由网络提供服务和内容, 能被其它对等节点 (Peer) 直接访问而无需经过中间实体。在此网络中的参与者既是资源 (服务和内容) 提供者 (Server), 又是资源 (服务和内容) 获取者 (Client)。因此, P2P 打破了传统的 Client/Server (C/S) 模式, 在网络中的每个结点的地位都是对等的。每个结点既充当服务器, 为其他结点提供服务, 同时也享用其他结点提供的服务。

### 2.2 P2P 网络的主要特点

P2P网络的特点体现在以下几个方面<sup>[1]</sup>:

- 非中心化: 网络中的资源和服务分散在所有结点上, 信息的传输和服务的实现都直接在结点之间进行, 可以无需中间环节和服务器的介入, 避免了可能的瓶颈。
- 可扩展性: 在 P2P 网络中, 随着用户的加入, 不仅服务的需求增加了, 系统整体的资源和服务能力也在同步地扩充, 始终能较容易地满足用户的需要。整个体系是全分布的, 不存在瓶颈。理论上其可扩展性几乎可以认为是无限的。

- 健壮性：由于服务是分散在各个结点之间进行的，部分结点或网络遭到破坏对其它部分的影响很小，因此，P2P 架构天生具有耐攻击、高容错的优点，并能在部分结点失效时能够自动调整整体拓扑，保持其它结点的连通性。
- 负载均衡：P2P 网络环境下由于每个节点既是服务器又是客户机，减少了对传统 C/S 结构服务器计算能力、存储能力的要求，同时因为资源分布在多个节点，更好的实现了整个网络的负载均衡。

### 3. SIP协议

#### 3.1 SIP 协议简介

SIP (Session Initial Protocol, 会话初始协议)<sup>[2]</sup>是一种在IP网络中建立、修改和终止多媒体会话的应用层协议。它是因特网工程任务组(IETF)在不断进行标准化的多媒体协议体系的一部分，它与其它协议一起（如：完成会话描述的SDP协议，完成媒体实时传输的RTP协议，完成多媒体资源预留的RSVP协议等），提供基于IP网络的电话和多媒体会话业务。

#### 3.2 SIP 的组网结构

SIP 网络中主要包括以下功能实体：

- User agent (用户代理)：构成对话的端点，发送或者接收 SIP 请求和响应，包括 UAC 和 UAS。
- SIP Proxy Server (代理服务器)：将请求消息路由到 UAS 以及将响应消息路由到 UAC，即完成 SIP 消息的接收和转发。
- Registrar Server (注册服务器)：为用户提供注册服务，接受 REGISTER 请求，登记用户的当前位置，以便其他用户可以找到该用户。提供用户授权服务，与代理服务器一起提供用户定位服务。
- Location Server (地址服务器)：存储用户的位置信息，提供位置查询服务。
- Redirect Server (重定向服务器)：为终端提供 SIP 消息重定向服务。

图 1 给出了 SIP 协议的组网结构：

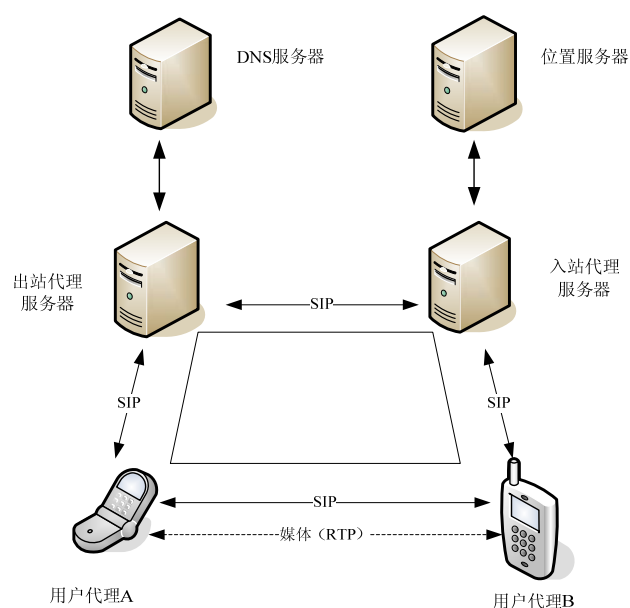


图 1 SIP 协议组网结构

## 4. 基于SkipNet的P2P网络

### 4.1 SkipNet 简介

SkipNet 是一个通过跳表 (SkipList) 数据结构实现局部特性的结构化 P2P 模型, 它的局部特性表现在两个方面: 1) 内容局部性, 即语义上可控的对象放置, 数据总是显式地被放到某个特定结点或者被分配到特定的网络组织; 2) 路由局部性, 即保证同一组织内两个结点的消息路由的路径一定位于该组织内部。

这两种特性使得 P2P 网络获取数据的可靠性和有效性大大提高。对于, 采用经典 P2P 路由算法 Chord 这样的网络来说, 网络一旦被分割, 性能就会急剧下降, 甚至无法工作陷入瘫痪, 而由于 SkipNet 能保证统一组织中结点间的消息路由路径一定位于该组织内部, 组织内部的结点仍然可以正常工作, 不影响组织内部的通信。这也正是本文选择 SkipNet 作为 P2P-SIP 网络路由算法的主要原因。

SkipNet 采用了两个独立但又相关的地址空间: 名标识 (NameID) 空间和数值标识 (NumericID) 空间。结点名和内容标识被直接映射到 NameID 空间, 而 NumericID 则是通过通过对结点名和内容标识进行哈希运算得到。

### 4.2 SkipNet 的网络结构

#### 4.2.1 跳表的概念

跳表 (SkipList) <sup>[3]</sup> 是一个有序链表, 其中一些指针会跳过很多链表中的元素指向跳表中距离很远的结点。每个指针有它的层, 层越高指的越远。在一个跳表中, 第  $h$  层指针指向距离  $2^h$  的结点, 如图 2 所示。每隔  $2^i$  个结点才会有一个  $i$  层结点, 其中最底层为第 0 层。

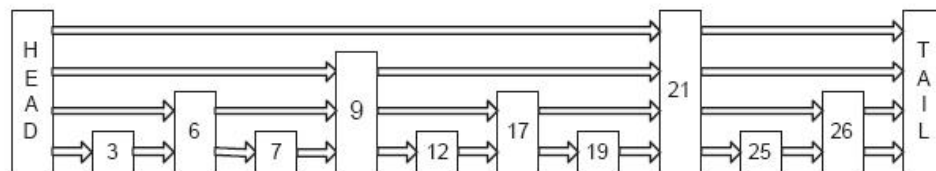


图 2 跳表结构

#### 4.2.2 SkipNet 的组织结构

借用跳表的思想, 用网络结点代替跳表结点, 将跳表头尾相连组成一个环, 并且每个指针变为双向的以保证路由局部性, 这就是 SkipNet 的基本结构。

SkipNet 的结点按照结点名或其数据对象的关键字来排列, 结点名或数据对象的关键字称为结点的 NameID。每个结点存储  $2\log N$  个指针,  $N$  为网络中结点数。结点的指针集也就是结点的路由表 (Routing Table, R-Table), 所以路由表的项数为  $2\log N$ , 第  $h$  层指针指向距离当前结点  $2^h$  跳的两个结点, 一个沿顺时针方向, 一个沿逆时针方向。

图 3 给出了一个包含 8 个结点的 SkipNet 结构。

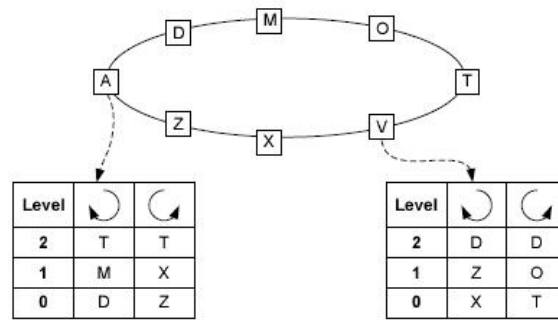


图3 SkipNet 网络结构和路由表

图4给出了图3中结点在各层的连接情况。所有结点通过0层指针相连,形成根环(Root Ring)。1层指针将根环分成两个环,因为1层指针是隔一个结点指的。类似的,通过2层指针可形成四个环。以此类推,第 $h$ 层指针将第 $h-1$ 层环分成两个 $h$ 层环,即将根环分成 $2^h$ 个 $h$ 层环。

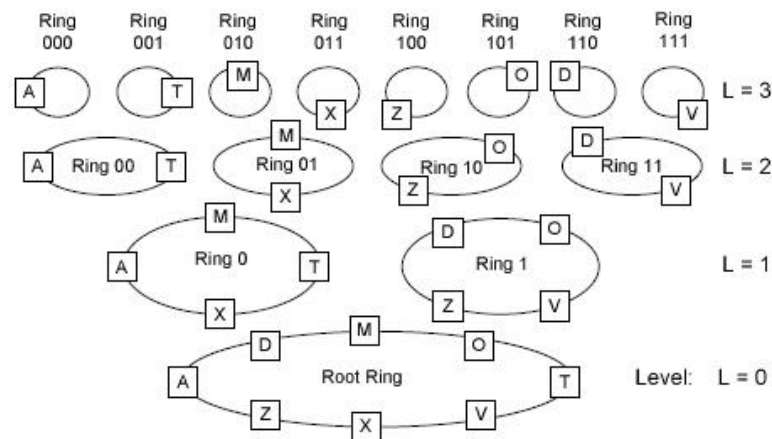


图4 SkipNet 分层结构

每个结点对环的随机选择可以由它唯一的二进制编码来决定,这个二进制编码就是结点的 NumericID。如图4,结点的 NumericID 的前 $h$ 位决定了该结点在第 $h$ 层所在的环。例如,结点 X 的 NumericID 为 011,它的前两位 01 决定了 X 在第2层属于环 01。SkipNet 并不一定要求使用哈希函数来生成结点的 NumericID,只要求 NumericID 是随机生成并且唯一即可。

### 4.3 SkipNet 的路由机制

#### 4.3.1 NameID 路由

基于 NameID 的路由遵循以下规则:在每一个结点,路由消息总是发送至路由表中不超过目的结点 NameID 的最高层结点,直到消息到达 NameID 与目的结点 NameID 最接近(或相等)的结点为止。NameID 路由跳数为  $O(\log N)$ 。

#### 4.3.2 NumericID 路由

基于 NumericID 的路由遵循以下规则:首先查找根环,直到发现一个结点与目的结点的 NumericID 第一位相匹配,此时通过这个结点跳升到它的1层环,再在1层环上查找与目的结点 NumericID 匹配前两位的结点。以此类推,每次上升一层就多匹配一位。当到达第 $h$ 层找不到能与目的结点 NumericID 匹配 $h+1$ 位的结点时,选择一个与目的结点的 NumericID 在数值上最接近的结点即为目的结点。NumericID 路由跳数为  $O(\log N)$ 。

## 5. 基于SkipNet结构的P2P-SIP系统设计

### 5.1 系统概述

当前分层结构在各种网络中被普遍采用,被证明是一种行之有效的路由方法。将网络分成不同的子网(即不同的层次),使每个子网内的用户通信只在本子网内部进行,与其它网络不发生任何关系;只有在属于不同子网的用户间的通信才会涉及到高一层的子网的参与。这样,既减轻了网络的负担,又可以最大限度的保证通信质量。因为本系统采用的是基于SkipNet的P2P模型,而SkipNet网络结构设计本身就具有局部特性,将P2P网络从逻辑上分为多个层次,更有利于实现P2P网络与SIP的完美结合。

根据SkipNet的组网思想,结点按照结点名(NameID)来排列,那么,如果将结点的URI地址映射到NameID地址空间,即使URI相近的结点在SkipNet网络中直接相邻或属于同一个组织内部,形成一个逻辑子网。SIP用户在P2P网络上进行注册,也就是在逻辑根环上进行注册,在查找用户时可直接利用SkipNet的路由算法进行查找。

### 5.2 针对P2P SkipNet网络结构的SIP扩展方案

在基于SkipNet结构的P2P-SIP系统中引入两个独立而关联的地址空间:名标识空间(NameID space)和数值标识空间(NumericID space)。

#### 5.2.1 结点URI

要在结点之间传送SIP消息,每个结点必须要有一个SIP URI。对结点的SIP URI定义如下:

NodeURI = "sip:"NodeID"@hostport  
";user-param";ring-param uri-parameters

其中,

- 1) NodeID 为结点域名通过哈希运算得到的哈希值映射到 NumericID 空间原则上, SkipNet 网络并不一定要求使用哈希函数来生成结点的 NumericID, 只要求 NumericID 是随机生成并且唯一即可。
- 2) hostport=反向 DNS 域名:端口(如果是结点的 IP 地址,需要转换为 DNS 域名)。hostport 可以看做是结点的 NameID, 因为网络需要保证局部特性, 通过使用反向域名更容易实现 NameID 的结构统一。例如, 将 dolly.bupt.edu.cn 写成 cn.edu.bupt.dolly, 从而在 bupt.edu.cn 域中的结点拥有共同的前缀, 也就能使它们在 SkipNet 网络中处于相邻近的位置, 所以, 这样的命名方案很好的服务于 SkipNet 网络的内容局部性和路由局部性。
- 3) user-param 规定结点为 node, 即 user=node。
- 4) ring-param 为结点所属的环, 通常默认为 0, 即结点位于根环, 当进行结点查找等过程时, ring-param 在 SIP 消息中用来指示当前环节结点所在的逻辑环。

例如, 域名为 dolly.bupt.edu.cn 的结点, 通过哈希该地址得到的结点标识 NodeID 为 0101001, 则该结点的 SIP URI 为:

sip:0101001@cn.edu.bupt.dolly;user=node;ring=0

#### 5.2.2 资源 URI

在P2P-SIP系统中,用户的SIP URI是作为资源存在于P2P网络中的。SIP用户的SIP URI与非P2P-SIP系统中的SIP URI相同,只不过增加了“resource-ID”参数来作为在数值标识空



间中对应的 ID，以减少运算，定义如下：

resourceID-param="resource-ID="ResourceID

其中，ResourceID 为用户的 SIP URI 通过哈希运算得到的哈希值，所以，ResourceID 在 SkipNet 网络中也是作为用户或资源的 NumericID，以方便查找。

例如，如若用户的 SIP URI 为 sip:dd@p2psip.com，哈希该 SIP URI 的值为 1010110，则该用户的资源 URI 表示为：

sip:dd@ p2psip.com;resource-ID=1010110

### 5.2.3 扩展标签

本系统使用“SkipNetP2P”这一个可选的标签来表示在基于 SkipNet 的 P2P-SIP 系统中所使用的为与该系统相应的扩展 SIP 协议。在这个系统中的结点应在消息的“Support”字段中列出该扩展，而使用该扩展协议发出请求的结点应在消息的“Require”字段使用该标签。

## 5.3 系统工作流程

### 5.3.1 结点加入

新结点加入操作从过 REGISTER 请求来完成。

新结点开始发起请求前应通过其他机制至少获知网络中的一个已存的结点，称为自举结点（bootstrap）。结点加入的 REGISTER 请求的 request URI 应该是自举结点的 IP 地址，无需带有任何扩展的 P2P-SIP 参数。例如，假设自举结点的 IP 地址为 59.64.159.10，则结点加入的 REGISTER 请求的 Request-URI 为：sip:59.64.159.10

REGISTER 请求的 To 和 From 为待加入结点的 NodeURI，ring 参数设置为 0。同时，在该请求的 Contact 字段中也应包含该结点的 NodeURI，以表明该请求是注册请求，而不是查询请求。在上例中，假设新结点是域名为 dolly.bupt.edu.cn 的结点，通过哈希该地址得到的结点标识 NodeID 为 0101001，则 REGISTER 请求消息应为：

```
REGISTER sip:59.64.159.10 SIP/2.0
To: sip:0101001@cn.edu.bupt.dolly;user=node;ring=0
From: sip:0101001@cn.edu.bupt.dolly;user=node;ring=0
Contact: sip:0101001@cn.edu.bupt.dolly;user=node;ring=0
Max-Forwards: 64
Call-ID: abcd1234
CSeq: 21 REGISTER
Require: SkipNetP2P //对 SIP 协议栈进行 P2P 扩展的标签
Supported: SkipNetP2P
Content-Length: 0
```

新结点将 REGISTER 请求发送给自举结点后，自举结点会将 REGISTER 请求继续转发，首先按照在数值标识空间路由（NumericID 路由）的方式自顶向下，在各层环（ring）上路由到与新结点 NumericID（即 0101001）具有相同前缀的结点，同时在每一层找到与其 NameID（即 cn.edu.bupt.dolly）拥有最长相同前缀的结点即为其邻居结点，并且据此更新每一层的邻居信息，将新结点逐层加入各层环。在各层转发注册信息时，用 ring 参数来表明层号以进行区分。与传统的 SIP 网络不通，在这个 P2P-SIP 网络中，不存在专门的注册服务器，所以结点的注册信息由在各层与之相连的邻居结点共同维护。每一层邻居结点将发送响应消息 200OK 给新结点表示确认注册，当新结点收到 ring=0 的 200OK 响应消息时，即得知已成功加入根环，至此，加入过程完成。

### 5.3.2 资源注册

资源的注册通过 REGISTER 消息来完成。

首先构建 REGISTER 请求。此请求与结点加入的 REGISTER 请求不同的是它在请求的 To、From、Contact 字段使用的是资源 URI 而不是 NodeURI。例如, SIP URI 为 dd@p2psip.com 的 SIP 用户的注册 REGISTER 请求 To 字段如下:

To: sip:dd@p2psip.com;resource-ID=1010110

其它字段与结点加入的 REGISTER 类似, 假设自举结点的 IP 地址为 59.64.133.20, 则 REGISTER 请求消息应为:

REGISTER sip: 59.64.133.20 SIP/2.0

To: sip:dd@p2psip.com;resource-ID=1010110

From: sip:dd@p2psip.com;resource-ID=1010110

Contact: sip:dd@p2psip.com;resource-ID=1010110

Max-Forwards: 64

Call-ID: efgh5678

CSeq: 45 REGISTER

Require: SkipNetP2P //对 SIP 协议栈进行 P2P 扩展的标签

Supported: SkipNetP2P

Content-Length: 0

对于资源的注册请求, 可以用下面的这个数据结构来表示:

域	ID
---	----

域对应资源 URI 的反向地址, ID 即为资源的 resource-ID 的值。在上面的例子里, 资源发送 REGISTER 请求后在其他结点中生成记录:

com.p2psip@dong	1010110
-----------------	---------

这里的结点记录采用这样的数据结构将资源的相关信息也做了类似于在 SkipNet 名标识空间和数值标识空间的映射。域可以作为 NameID, ID 作为 NumericID。

这样的记录方式有利于对资源的逐层注册。有了资源的 NameID 和 NumericID 就可以运用 SkipNet 的数值标识空间路由算法自顶向下在各层环上资源进行注册, 过程与上一节的结点注册过程类似。

此外, 这种记录结点注册信息的数据结构还有一个优点, 即它可以充分利用基于 SkipNet 模型的 P2P 网络的局部特性。由于在注册过程中会利用 NameID 查找各层的邻居结点, 所以资源会和各层中与它具有相同 NameID 前缀 (即相同的资源 URI 后缀部分) 的结点相连, 根据 SkipNet 的内容局部性和路由局部性, 这样就保证了同一域内的资源在 P2P 网络中仍然相连或位于同一组织内部, 并且查找资源的路径也一定会封闭在组织内部, 符合最初的设计目标。

### 5.3.3 资源查找

资源的查找也是通过 REGISTER 请求来完成的。构建 REGISTER 请求时与注册用的 REGISTER 的唯一区别是在进行资源查找时请求中不包含 Contact 字段。

设接收到查询请求的结点 SourceNode, 路由过程分以下两种情况:

(1) 如果 SourceNode 与目标资源具有相同的 NameID 前缀, 说明二者存在于同一域内。

根据请求查找的资源 URI 采用 NameID 路由算法进行查找。首先查找最高层的链表, 如果没有找到一个比自己更接近但不超过目标资源 NameID 的结点, 则开始向低一层链表搜索; 如果存在这样的结点, 则向其转发 REGISTER 请求, 重复这个过程直到找到该层最后一个最接近但不超过目标资源的结点。之后再从低一层结点里面用类似的方法继续转发 REGISTER 消息并找到前缀进一步接近目标资源的结点。重复上述动作, 直到找到目标资源所在结点 DestNode, DestNode 从自己的资源列表中找到该资源的信息, 将其置于 200OK 响应中的 Contact 字段发回给查询该资源的结点。

(2) 如果 SourceNode 与目标资源没有相同的 NameID 前缀, 可以在最高层链表中随机选择一个方向转发查找信息, 这样有利于减少路由某些中间结点的负载。在找到与目标资源具有相同的 NameID 前缀的某个结点后, 则进行与上述相同的查找过程。

## 6. 结论

将 P2P 技术与 SIP 协议相结合可以充分利用 P2P 的网络结构弥补 SIP 系统的不足。P2P-SIP 系统利用 P2P 技术代替 SIP 系统中的网络服务器, 将现有的 C/S 架构的 SIP 应用系统改造成基于 P2P 模式的对等网络, 使得整个网络范围内达到负载均衡, 网络性能也不再依赖于传统的 SIP 服务器。本文通过采用基于分层思想的 SkipNet 网络模型, 可利用 SkipNet 的局部特性使得在一定区域内部结点之间的通信可以通过很快捷的路由到达, 这在一定程度上也改善了 P2P-SIP 网络的性能。

## 参考文献

- [1] 陈贵海, 李振华. 对等网络: 结构、应用与设计. 清华大学出版社, 2007 年 9 月
- [2] 中华人民共和国通信行业标准会话初始协议技术要求. 中华人民共和国信息产业部, 2003 年 7 月
- [3] Nicholas J.A. Harvey, Michael B. Jones, Stefan Saroiu, Marvin Theimer, Alec Wolman. SkipNet: A Scalable Overlay Network with Practical Locality Properties. Microsoft Research Technical Report MSR-TR-2002-92

# SkipNet-based P2P-SIP System

Pu Donglin

Beijing University of Post and Telecommunications, Beijing (100876)

## Abstract

A solution using SkipNet network routing algorithm in P2P-SIP network system is proposed for the first time, especially taking advantage of the SkipNet's locality properties to guarantee that message traffic between two overlay nodes within the same organization is routed within that organization only. This paper introduce the basic concept of P2P network and SIP protocol first, then introduce SkipNet algorithm and focus on the design and working process of SkipNet-based P2P-SIP System.

**Keywords:** P2P, SIP, P2P-SIP, SkipNet