RFC 1287 未来的Internet体系结构

网络工作组
RFC: 1287

D.clark
MIT
L.Chapin
BBN
V. Cerf
CNRI
R. Braden
ISI
R. Hobby
UC Davis
1991年12月

提示

这是一个提供信息的RFC，它讨论Internet体系结构未来可能演变的重大方向，及走向期望目标的建议步骤。它是提供给Internet社区讨论和评议用的。本文是为Internet社区提供信息，而不是指定Internet标准，本文的分发不受限制。

目录

1. 绪论 …… [1]
2. 选路与寻址 …… [3]
3. 多协议体系结构 …… [5]
4. 安全性体系结构 …… [7]
5. 业务流控制与状态 …… [9]
6. 现代应用 …… [10]
7. 参考文献 …… [11]
附录A 建立步骤 …… [12]
附录B 组成员 …… [15]
安全性考虑 …… [16]
作者地址 …… [16]

1. 绪论

1.1 Internet体系结构

作为TCP/IP协议集之后的巨大计划， Internet体系结构在70年代后期由一个网络研究小组[1、2、3、4]开发并测试。80年代初期，体系结构中加入了若干重要的特性，如子网化、自治系统和域名系统[5、6]。最近又加入了IP组播[7]。

在本体系结构框架内， Internet工程任务组(IETF)一直以极大的活力和效率为Internet策划、

定义、推广、测试和标准化协议进行不懈的工作。已完成的三个特别重要的领域是选路协议、TCP性能与网络管理。同时，Internet基础设施继续以惊人的速度增长。自1983年1月ARPANET第一次从NCP转换成TCP/IP时，Internet的销售商、管理员、专家和研究人员一直以极大的努力坚强地工作，为他们的成功而奋斗不息。

定义Internet体系结构的一组研究人员形成了Internet活动董事会(IAB)的初始成员。IAB由1981年DARPA建立的一个技术顾问组发展成为Internet总技术和策略监督实体。IAB的成员年年有变化，以便更好地体现Internet社区中需求和议题的变化，及反映Internet的国际化，但仍旧保持协议体系结构制定的关系。

IAB创建了IETF，为Internet实施协议的开发和工程设计。为了管理不断发展的IETF活动，IETF主持在IETF内建立了Internet工程指导组(IESG)。IAB和IESG密切配合，共同批准IETF内开发的协议标准。

过去几年中，对基础体系结构有着不断增加的严峻考验的迹象，大部分由于Internet持续不断的增长引起。对这些问题的讨论经常反映在许多主要的发送文件清单中。

1.2 假设

对当前Internet体系结构中的问题，解决的优先次序取决于人们对TCP/IP与OSI协议集未来关系的观点。一种观点是让TCP/IP夭折在成功之中，然后转换到OSI协议。然而，许多在Internet协议产品和服务上花过大力气并获得成功的人们，急于要在已有的框架内尝试解决新的问题。而且，有些人相信OSI协议将会遇到许多同样类型的问题。

为了着手解决这些问题，IAB和IESG于1991年1月联合组织了一天有关Internet体系结构议题的讨论会。这次会议的框架是由Dave Clark(见附录A中的幻灯片)整理的。关于TCP/IP与OSI协议的关系和未来方向问题上的讨论生气勃勃、富有挑战，有时还有激烈的争论。会议的重要成果是在下一个5~10年涉及网络世界的下述四个基本假设上达成了共识。

(1) TCP/IP和OSI协议集将在一个长时期内共存。

OSI协议集引入的背后是强有力的政治和市场力量，以及以某些技术优势作后盾。然而，TCP/IP牢固确立的市场位置意味着在可预见的未来非常可能继续使用。

(2) Internet将继续包括各种各样的网络和服务，永远不会是由一个单网络技术构成的。

实际上接到Internet上的网络技术和特性的范围在下一个十年还将增加。

(3) 商用和专用网也将加入，但不能期望公共通信提供全部服务。将会形成公用网和专用网、公共通信与专用线路混用的局面。

(4) Internet体系结构要能达到10^9个网络的规模。

Internet的规模历史性地呈指数增长，在将来的某个时候估计可能会饱和，但预测到什么时候饱和，差不多和预测未来的经济一样容易。在任何情况下，负责工程设计的需要考虑一个有能力扩展到最坏情况规模的体系结构。指数9是比较模糊的数字，估计在7~10之间变化。

1.3 开始一个规划过程

IAB和IESG会议的另一个成果是在体系结构进化中的下列五个最重要的领域上形成了共识。

(1) 选路与寻址

这是一个最急需解决的体系结构的问题，因为它直接关系到Internet继续成功增长的能力。

(2) 多协议体系结构

Internet正在朝着广泛的既支持 TCP/IP又支持OSI协议集的方向迈进。对两个协议集的支持带来了技术难题，需要有一个计划，也就是一个体系结构来增加成功的机会。人们开玩笑地把这个领域看成是：“ 为了造福人类，问题变得更艰难 ”。

Clark观察到转发网关 (如邮件网关)在Internet运行中是非常有生命力的，但是它不属于体系结构或规划的一部分。该组成员讨论了围绕包含这样的网关的部分网络连接来建立体系结构的可能性。

(3) 安全性体系结构

在设计Internet体系结构时，虽然考虑到了军事上的安全性，可是现代安全性议题是非常广泛的，它也包括了商业上的需求。还有，经验表明除非一开始就把它建立到体系结构中去，否则是很难在协议集内再加入安全性。

(4) 数据流控制及状态

Internet将扩展以支持如语音和视频这样的 “ 实时 ” 应用。这就需要网关中有新的包排队机制(数据流控制)和附加的网关状态。

(5) 现代应用

随着基础的Internet通信机制的成熟，需要不断革新和标准化，以创建新形式的应用。

IAB和IESG于1991年6月再次在SDSC召开三天的会议，讨论这 5个课题。这次会议多少有点反常，被称为 “ 体系结构的再处理 ”，召集在一起开会，表明有坚强的决心朝着规划体系结构的进化迈出第一步。除了 IAB和IESG 以外，由32人组成的小组，包括了研究指导组 (IRSG)的成员及少量的特邀客人。会议的第二天，分成 5个组讨论，每个组讨论 1个领域的问题。附录B列出了成员名单。

本文件是从这些组的主席报告中收集得到的。该材料在亚特兰大召开的 IETF会议上介绍过，同时发表在会议记录中[8]。

2. 选路与寻址

为了应付Internet预期的增长和功能的演变， IP寻址和选路结构需要改变。人们预测：

- Internet将用完IP网络地址的某些地址类，如 B类地址。
- 尽管该地址空间当前已被子分和管理， Internet将用完全部32位IP地址空间。
- IP网络号的总数将增长到一定时刻，就连较好的选路算法也不再能完成基于网络号的选路。
- 为了允许适应不同的 TOS和策略，从一个源到一个目的地需要多个路由。这将需要新的应用和多种多样的转运服务来推动。源或源的代理，必须控制路由的选择。

2.1 建议的方法

处理这些事情所需方法有通用的约定。

(1) 必须改变寻址方案，使网络号集聚成较大的单位，以此作为选路的基础。自治系统或行政管理域(AD)就是一种集聚的例子。

集聚将完成若干目标：确定采用策略的范围，控制选路部件数，以及为网络管理提供部件。有些人认为如一个嵌套的 AD那样，可进一步组合一些集聚。

(2) 必须提供某些有效的方法来计算公共路由，以及某些通用的方法来计算 “ 特定 ” 的路由。

特定路由的通用方法将由“源路由”指定的形式来建立路由。

会上，对期望AD如何集聚或选路协议如何组织来处理集聚边界，尚未达成完全一致的意见。可能使用一个非常通用的方案 (参考文献Chiappa)，可是某些人倾向于一个更受限制的方案，并定义期望的网络模型。

为了处理地址空间耗尽的问题，必须要么扩展地址空间，要么在网的不同部分重用 32位地址字段。下一节将描述几种可能的地址格式。

或许更重要的问题是如何向新的方案迁移。所有迁移计划都需要某些路由器 (或者Internet内的其他部件)能重写包头，以适应只处理老格式或者只处理新格式的主机。除非格式变换能够进行算法上的推理，迁移本身需要在变换元素中建立某种状态。

我们并不计划对体系结构进行一系列“小”的改变。从现在起，我们将着手实施一个计划，以便能渡过地址空间耗尽的难关。比起 Internet社区近期承担的任务而言，这是一系列更长的规划行动，但迁移问题需要一个漫长的研制周期，同时很难发现有效的方法来处理某些更直接的问题。诸如B类地址的耗尽问题，从某种意义上讲，就其本身而论，不用长时间。因此，一旦我们着手进行一项变更的计划，就要求全部替代当前的 32位全球空间。(如果有非常巧妙、能很快应用、而又留有发展空间的想法浮现出来的话，本结论总是可以被修订的。这并不意味着我们不鼓励对于短期行动的创新性设想。但需要指出的是即使小的变更也要花长时间去推广应用)。

仅有地址空间变换是不够的。同时还需要提供一个规模可伸缩的选路体系结构以及能更完善地管理Internet的工具。建议的方法就是把 AD作为选路的集聚单位。我们已经有部分方法来实现。IDRP能实现这一功能。 BGP的OSI版本(IDRP)也能实现。 BGP改进后也可实现。另外需要的附加设施是要有一张网络号到 AD的映象表。

为了若干原因(特定的路由和地址变换以及计费和资源分配)，我们将从“无状态”网关模型做起，在该模型中仅把预先计算好的路由存放在网关中，然后发展成另一个模型，该模型至少在某些网关中每个连接有状态。

2.2 扩展的IP地址格式

扩展的IP地址格式有三个比较好的选择。

(1) 用同样位数不同含义的地址字段代替 32位地址字段。由此地址的唯一性只是在某个较小的区域(一个AD或者一个集聚的AD)内，而不在全球范围。当包穿越边界时，边界上的网关重写包地址。

问题：(a)必须找到并重写包内的地址；(b)主机软件需作修改；(c)必须用某些方法建立地址映象。

本方案是Van Jacobson的研究成果，也可参见Paul Tsuchiya为NAT所作的工作。

(2) 将32位地址字段扩展至64位(或其他值)，用以保持一个全球主机地址和主机所在的AD。

这样的选择方案提供一个从主机到作为选路根据的 AD的烦琐的映射。普通路由(是指基于目的地址而不需考虑源地址的选路)可直接从包地址中得到，正如目前进行的，不需要任何事先建立过程。

(3) 将32位地址字段扩展至64位(或其他值)，并用该字段作为“平面”主机标识符。需要时，用建立连接来为路由器提供从主机标识符到 AD的映射。

这64位地址如以太网地址一样，可用来简化主机标识符的分配问题。

所有以上选择方案作为迁移的一部分，都需要一个地址重写模块。第二和第三方案 IP头需要改变，所以主机软件也要随之改变。

2.3 建议的行动

建议采取下列行动：

(1) 时间表。

对于上面提出的各种问题，要编制出一个估计的详细时间表，又要对一个新的寻址 /选路体系结构编制出开发和推广应用的相应时间表。用这些时间表作为根据来评价用于变革的各种提案。这是IETF的任务。

(2) 新地址格式。

探索下一代地址格式的可选方案并提出一个迁移计划。特别是要构造一个作地址映射的网关样机。要理解这个任务的复杂性，以便指导我们思考有关迁移的方案选择。

(3) 基于AD的选路。

采取步骤做出作为选路基础的网络集聚 (AD)。特别是要为映射网络号到 AD的一张全球映射表探索若干可选方案。这是 IETF的事情。

(4) 基于策略的选路。

基于策略的选路要继续当前的工作。有下列明确的目标：

- 寻求方法以控制指定策略的复杂性 (这是一个人类的接口议题，而不是算法复杂性议题)。
- 充分了解在网关中保持连接状态的议题。
- 充分了解连接状态建立议题。

(5) 进一步集聚的研究。

作为研究活动，探索如何将 AD集聚到仍然较大的选路元素中。

- 考虑体系结构应定义 AD的“ 角色 ”，还是集聚的“ 角色 ”。
- 考虑用一个万能的选路方法，还是在 AD和集聚以内和以外用不同的选路方法。

现有的计划如 DARTnet 工程项目帮助解决这些议题中的几个：如网关内的状态、状态建立、地址映射和计费等。研究开发界的其他试验也承担本领域的研究。

3. 多协议体系结构

改变Internet以支持多协议集引起以下三个特殊的体系结构问题：

- 如何正确地定义Internet？
- 如何设计支持多个又不论何种协议集的 Internet？
- 是为部分还是过滤了的网络设计连通性？
- 如何在体系结构中加入能明显地支持应用的网关？

3.1 什么是“ Internet ”？

如果不首先确定我们认为的 Internet是什么或者应该是什么的话，要想建设性地处理 “ 多协议Internet ” 议题将是非常困难的。我们要把“ Internet ”和“ Internet社区 ”区别开来，前者是由通信系统组成的，而后者是一群人和组织。大部分人接受后者的松散定义，即“ 认为他

们自己是Internet社区的一部分”。Internet本身这种“社会学的”定义似乎是没有用的。

不久以前，Internet被定义为IP网络连通性(IP和ICMP过去是、现在仍然是唯一“需要”的Internet协议)。如果我能ping你，你能ping我，那么我们都在Internet上，同时，Internet令人满意的工作定义可构想为IP对话系统的接近可过渡的最后结果。这样的Internet模型是简单的、统一标准的，或许最重要的是可测试的。IP网络连通性模型可清楚地判别系统是否“在Internet上”。

随着Internet的增长及其使用的技术已广泛的被商界接受，对一个系统“在 Internet上”的含义已经有所改变，应当包括：

- 具有部分IP网络连通性，受限于策略过滤器的任何系统。
- 运行TCP/IP协议集，不管是否从 Internet 的其他部分实际上可接入的任何系统。
- 能交换RFC 822邮件，无需邮件网关的干预或邮件对象的转换的任何系统。
- 有e-mail连通到Internet，不论是否需要邮件网关或邮件对象转换的任何系统。

对Internet的这些定义仍是基于原始的网络连通性概念，只是“栈的向上移动”。

在此，基于有区别的统一概念，提出 Internet的新定义：

- “老的”Internet概念：以IP为基础,组织的原则是IP地址，也就是一个公共的网络地址空间。
- “新的”Internet概念：以应用为基础,组织的原则是域名系统和目录，也就是一个公共的(虽然必定是多形式的)应用名字空间。

这就告诉我们，“连接的状况”概念传统上是与 IP地址(通过网络号)紧密联系在一起的，应该代之以与存放在分布式 Internet目录中的名字和相关标识信息联系在一起的。 Internet基于名字的定义意味着一个大得多的 Internet社区，以及一个更为动态 (和不可预测的) 可运转的Internet。对Internet体系结构的争论，是基于在很宽的范围内对未来可能发展的适应性，而不局限于原来的设想。

3.2 基于过程的多协议Internet模型

与其制订一个特殊的“多协议 Internet”，接受一个预先确定特定协议数量的体系结构，倒不如建议采用一个面向过程的 Internet模型，它可以适应不同的协议体系结构，符合传统的“能工作”原则。

面向过程的Internet模型，作为一个基本前提，主张不包括稳定状态“多协议 Internet”的体系结构。最基本的驱动 Internet进化的力量不是推动它朝多协议多样化发展 (虽然事实上永远不可能达到)。要说明的是 Internet发展的趋势是向同质性进化，作为最“热动稳定”状态，下面描述一个新的基于过程的 Internet 体系结构的四个部分：

第1部分：核心Internet体系结构。

这是传统的基于 TCP/IP的体系结构。是 Internet进化的“磁铁中心”，公认(1)同质性仍然是处理互连网多样化的最好方法； (2)IP网络连通性仍然是 Internet的最佳基本模型 (在全球Internet中，不论IP无处不在的实际状态是否是一个现实)。

一开始，Internet体系结构只包括第 1部分。然而Internet的成功在于它超出了原来的设想。无处不在和高度统一，对极大地丰富 Internet“基因库”作出了贡献。

新Internet体系结构增加的两个部分扩大了 Internet的广度和深度。

第2部分：链路共享。

传输媒体、网络接口及低层链路协议等物理资源是由多个非交互的协议集所共享的。这部分体系结构被认为是必须且适合于共存的，但不涉及到互操作性；被称为 ships in the night(S.I.N.)。

当然，共存的协议集实际上不是纯粹孤立的；在真正的Internet系统中，S.I.N.会引发管理、无冲突、协调和公平性等议题。

第3部分：应用互操作性。

虽然缺乏互连普遍性(即“基础栈”的互操作性)，但只要在Internet系统的不相邻社区之间安排应用的基本语义能以传递信息，仍然可能获得普遍的应用功能。这可以通过应用转发站，或者由用户代理，对不同的由共同语义表示的应用服务提供一致的虚访问方法来完成。

体系结构的这一部分，强调了Internet的最终作用是作为应用间的通信基础，而不是它本身的结局。在一定程度上，使一个应用群体和他们的用户能够从一个基础协议集过渡到另外一个，而不会发生难以接受的功能丢失，这可被称为“过渡起动器”。

将第2和第3部分加入到原始的Internet体系结构中，充其量是一件好坏半掺的事情。虽然大大增加了Internet的广度和Internet社区的规模，但也会引入复杂性、价格和管理等重大问题，同时还会出现功能的丢失(特别对第3部分而言)。第2和第3部分不可避免地背离了第1部分所表示的同质性，但这是我们所不希望的。为了扩展Internet广度，某些功能丢失了，还要承受附加的系统复杂性和成本。而在一个完美的世界中，应该不需付出这些代价就可换得Internet的进化和扩展。

目前有一种趋势，Internet的进化倾向于第1部分表示的同质性体系结构，而不是第2和第3部分表示的折衷的体系结构。第4部分表达了这种趋势。

第4部分：混合/集成。

第4部分认识到可以从不同的Internet协议体系结构中集成类似的元素以形成混合体，以便减少Internet系统的多样化和复杂性。同时也认识到可以影响已存在的 Internet基础设施以便Internet吸收“新东西”，并把已建立的Internet的测试、评价和应用实践融入到“新东西”中去。

本部分表达了Internet的发展趋势，作为一个系统，试图回到原来由第1部分统一的体系结构所表示的“美好的状态”。虽然Internet将永远不会在未来的任意时刻回到统一的状态，但这是一个对Internet进化起作用的力量。

按照这个动态的进程模型，在TCP/IP栈上通过RFC 1006运行X.400邮件，集成IS-IS选路，传送网关以及对IP和CLNP协议的单个共同后继协议的开发，都是很好的例子。 在第1部分主张的“磁场”影响下，参照第4部分混合的动态，它们显示了背离第2和第3部分的非统一性，而走向更好的同质性。

4. 安全性体系结构

4.1 哲学准则

Internet安全性体系结构开发的主题是简单、可测试性、可信度、采用的技术和安全周界标识。

- 安全性比协议和密码保密措施要求更多。
- 安全性体系结构和策略应该简单且容易理解。复杂性会引起错误理解和不良的实现。
- 实现应该是可测试的，以便确定是否满足了策略。

- 我们认为使任何安全性体系结构运行的硬件、软件和人是可以信任的。假设安全性策略实施的技术设备至少和个人计算机及工作站具有同样的能力。我们不需要能力差的部件受到自保护(但可能会用诸如链路级密码编码设备进行外部补救)。
- 最后，认定安全性有效保护的周界是最根本的。

4.2 安全性周界

有4种可能的安全性周界：链路级、网络 /子网级、主机级和进程 /应用级。每种施加不同的需求，能够接纳不同的技术，并能对何种系统部件可以被认为是有效的做出不同的假设。

隐私强化邮件是一种进程级安全系统的例子，另一个例子是为SNMP提供身份验证和保密。主机级安全性一般是在主计算机的通信口上用一个外部安全机制。网络或子网安全性则应用从子网到“ 外部 ”的网关 /路由器上的外部安全性能力。 链路级安全性采用传统的点对点或媒体级(如以太网)密码编码机制。

关于网络/子网安全性保护存在许多开放的问题，不单是主机级 (端/端)安全性方法与网络/子网级安全性方法之间存在潜在的不匹配，而且网络级保护也不能处理安全性周界内出现的威胁。

在进程级采用保护，假定基础的程序和操作系统机制是可以信任的，不会由于使用了相应的安全机制而妨碍应用程序。当安全性周界在系统体系结构中向下移至链路级，就要做有关安全威胁的许多假设，以便得出这样的论点，就是在一个特定周界的实施是有效的。例如，如果只有链路级使用加密编码，我们可以假设来自外部的攻击，只通过通信线，那么主机、交换机和网关实际上是被保护的，同时人和所有部件中的软件都是可以信任的。

4.3 期望的安全性服务

如果在系统的应用级和较低级实现选定和非选定的接入控制，则需要可验证的正规的名字。除此之外，还需要实施完整性 (防修改、防欺骗、防重放)，保密性和防止否认服务。在某些情况下，可能还需要防止报文传输的否认或防止秘密信道。

已经有一些标准部件用以建立 Internet安全系统。可以采用密码算法 (如DES、RSA、El Gama1和其他可能的公共密钥和对称密钥算法)，也可以采用如MD2和MD5的散列函数。

根据OSI的意义需要可鉴别的名字，并且为了便于人们了解标识符和目录服务，非常需要一个指派标识符以及广泛使用目录服务的基础设施。把公共密钥与可鉴别的名字捆绑在一起，并把能力和许可与可鉴别的名字捆绑在一起的认证概念，具有很多优点。

在路由器/网关级，采用地址和协议过滤器及其他配置控制，能有助于形成一个安全性系统。把建议的OSI安全性协议3(SP3)和安全性协议4(SP4)作为Internet安全性体系结构的可能要素, 要给予认真的考虑。

最后，必须看到，在未实施安全存储的 PC或笔记本电脑系统上，安全地存储秘密信息 (诸如一个公共密钥对的秘密部分)， 还没有好的解决方案。

4.4 建议的行动

建议采取下列行动：

(1) 安全性参考模型。

需要建立一个Internet的安全性参考模型，并迅速地得到开发。该模型应该建立目标周界，并用文件形式建立安全性体系结构目标。

(2) 隐私强化邮件(PEM)。

对于隐私强化邮件，最关键的步骤看来是建立： (1)认证生成和管理基础设施； (2)X.500目录服务以提供通过可鉴别的名字访问公共密钥。在推广使用本系统时，还需要对专利方面的限制和出口限制给予认真关注。

(3) 分布式系统安全性。

对分布式系统的应用，不论是简单的客户机 /服务器系统还是复杂的分布式计算环境，都需要检查安全设施。例如，对授予与可鉴别的名字捆绑在一起的许可 /能力的认证的实用性应受到检查。

(4) 主机级安全性。

对面向主机的安全性，应当对SP4予以评估，SP3也在考虑之列。

(5) 应用级安全性。

不论是为了服务的直接实用性 (如PEM.SNMP身份验证)，还是为了获得能够形成 Internet安全性体系结构精华的有价值的实际经验，都应该实施应用级安全性服务。

5. 业务流控制与状态

目前的Internet平等地处理所有的IP数据报。每个数据报对同一连接、同一应用、同一应用类别、同一用户类，不论它和其他包有任何关系，都是独立地转发的。虽然在 IP头中定义了服务类型位和优先权位，但通常都没有实施，事实上还不清楚如何去实施它们。

众所周知，未来的Internet需要支持尽力而为所不能满足的大量应用，如电视会议的包图像和语音。为了处理实时业务流，要求在路由器中有以附加的状态来控制的业务流控制机制。

5.1 假设和原则

• 假设：Internet需要为业务流的特定子集支持性能保证。

遗憾的是对术语“ 性能 ”、“ 保证 ”或“ 子集 ”，远不能给出精确的定义。研究仍需要对这些问题做出回答。

• 默认的服务将继续是当前无服务保证的“ 尽力而为 ”数据报分发服务。

• 路由器机制可分割为两部分： (1)转发路径； (2)发生在后台的控制计算(如选路)。

转发路径必须高度优化，有时由硬件辅助完成，因此相对而言很昂贵，而且难于变更。运行在转发路径上的业务流控制机制，是由发生在后台的选路和资源控制计算创建的状态来控制的。在改变路由器的转发路径时，最多起动一次，所以最好一开始就使它正确。

• 新的扩展必须运行在一个高度异质的环境中。在该环境中，某些部分将永远不支持保证。对一个路径上的某些段 (如高速局域网)，即使当显式资源预留不用时，“ 超配给 ”(即超过容量)也会对实时业务给予满足要求的服务。

• 组播分发或许是最根本的。

5.2 技术议题

需要解决的技术议题，包括：

(1) 资源建立。

为了支持实时业务流，从源到目的地的路由上的路由器中需要预留资源。该新的路由器状态应该是“硬”的(如建立连接)还是“软”的(即缓冲的状态)？

(2) 资源捆绑与路由捆绑。

选择从源到目的地的一条路由传统上是由一个动态选路协议来完成的。资源捆绑和选路可以重叠在单个复合进程中，或者也可以基本上独立地完成。这就要求在复杂性和效率之间折衷考虑。

(3) 另一组播模型。

IP组播用一个逻辑寻址模型，在该模型中，目标地址本身与一个组联系。在 ST-2中，一个组播会话中的每个主机在它的建立包中包括一系列显式目标地址。每一种方法都有优点和缺点。当前还不十分清楚对n路电视会议而言，哪个会占优势。

(4) 资源建立与行政管理域间的选路。

不论倾向于哪种资源保证，必须保持穿越一条任意的端对端并包括多个 AD的路由。因此，任何资源建立机制必须与包含在 IDPR 中的路由建立机制平稳地配合。

(5) 计费。

资源保证子集(“类别”)可以是自然的计费单位。

5.3 建议的行动

此处所谓的行动是指对上面列出的技术议题的进一步研究，紧随其后的是相应协议的开发和标准化。DARTnet，DARPA研究测试床网络，在本研究中将起重要作用。

6. 现代应用

人们不禁要问“我们想要何种基于网络的应用，为什么现在还没有？”很容易列出一张潜在应用的大表，其中许多都将基于客户机 /服务器模型。然而问题中更有意思的是：“为什么还没有人来做呢？”回答是：方便应用程序编写的工具尚不存在。

首先，对于许多将用于穿越网络的数据术语，需要一套公共交换格式。定义了公共交换格式后，还需便于开发应用程序移动数据的工具。

6.1 公共交换格式

为使信息有意义， 应用程序必须知道它们要交换的信息的格式。考虑下面的格式类型：

(1) 文本 —— 文本是最标准的，但今天的国际性 Internet还需要有除了USASCII 以外的字符集。

(2) 图像 —— 当进入“多媒体时代”，图像变得越来越重要，但需要对如何在信息包中表示图像信息取得一致。

(3) 图形 —— 和图像一样，矢量图形信息需要一个共同的定义。有了定义的格式才能交换类似结构蓝图的细节。

(4) 视频 —— 先要知道从网络上来的视频信息的格式，才能在工作站上运行视频窗口。

(5) 模拟音频 —— 当然，人们需要的是伴有声音的视频，但这样的格式应该可以表示所有类型的模拟信号。

(6) 显示 —— 我们打开工作站上的窗口，并打开另一个人的工作站上的窗口，给它显示与研究项目有关的某些数据，所以需要一个通用的窗口显示格式。

(7) 数据对象 —— 对进程间的通信，类似整数、实数、串等数据的格式需要一致。

这些格式的相当一部分正在由几个标准组定义。我们需要为 Internet 的每一类取得一种一致的格式。

6.2 数据交换方法

应用程序将需要下列的数据交换方法：

(1) 存储转发。

不是每个人所有时间都在网上。需要一个标准手段向有时连在网上的主机提供信息流，也就是需要一个通用的存储转发服务。组播也应包括在这一服务中。

(2) 全球文件系统。

在网上，大部分数据访问可以被分解成单个文件访问。如果有一个真正的全球文件系统，那就能访问Internet上的任何文件(假定被许可的话)。你曾经需要用FTP吗？

(3) 进程间通信。

对一个真正的分布式计算环境，需要通过一些手段使进程在网络上能通过一个标准方法来交换数据。这样的需求包括 RPC、API等。

(4) 数据广播。

许多应用程序要求发送同样的信息到其他许多主机，因此需要一个标准且高效的方法来完成这功能。

(5) 数据库访问。

对于好的信息交换，需要为访问数据库指定一个标准方法。全球文件系统能使你获得数据，但数据库访问方法将告诉你有关它的结构和内容。

上述许多项正在由其他组织着手拟订，但对Internet的互操作性，还需要在方法上取得一致。

最后，现代应用需对本文中两个较早领域的问题寻找解决方案。从业务流控制与状态领域而言，应用需要发送实时数据的能力。这意味着数据能在一个确定的时间范围内分发。从安全性领域而言，应用也需要全球身份验证和访问控制系统。今天的 Internet由于缺乏可信度和安全，失去了许多有用的应用。这要求在明天的应用中得到解决。

7. 参考文献

[1] Cerf, V. and R. Kahn, "A Protocol for Packet Network Intercommunication," IEEE Transactions on Communication, May 1974.

[2] Postel, J., Sunshine, C., and D. Cohen, "The ARPA Internet Protocol," Computer Networks, Vol. 5, No. 4, July 1981.

[3] Leiner, B., Postel, J., Cole, R., and D. Mills, "The DARPA Internet Protocol Suite," Proceedings INFOCOM 85, IEEE, Washington DC, March 1985. Also in: IEEE Communications Magazine, March 1985.

[4] Clark, D., "The Design Philosophy of the DARPA Internet

Protocols", Proceedings ACM SIGCOMM '88, Stanford, California, August 1988.

[5] Mogul, J., and J. Postel, "Internet Standard Subnetting Procedure", RFC 950, USC/Information Sciences Institute, August 1985.

[6] Mockapetris, P., "Domain Names - Concepts and Facilities", RFC 1034, USC/Information Sciences Institute, November 1987.

[7] Deering, S., "Host Extensions for IP Multicasting", RFC 1112, Stanford University, August 1989.

[8] "Proceedings of the Twenty-First Internet Engineering Task Force", Bell-South, Atlanta, July 29 - August 2, 1991.

附录A 设定步骤

幻灯片1

Internet向何处去？

体系结构的选择

IAB/IESG -- 1990年 1 月

David D. Clark

幻灯片2

设定讨论的课题

目的：

- 为IAB、IESG及Internet社区建立一个理解的共同框架。
- 了解要解决的问题集
- 了解为我们敞开的解决方案的范围
- 得出某些结论或“ 总结论 ”。

幻灯片3

若干声明 —— 我的见解

两个不同的目标：

- 使建立Internet成为可能。
- 定义Internet的一套协议。

声明：这些目标有非常不同的含义。协议只不过是一种手段，然而是一种有力的手段。

声明：如果Internet获得成功及增长，就将需要专门的设计。这就需要至少另一个十年的继续努力。

声明：不加控制的增长将会导致混乱。

声明：从根本上解决问题看来是走向成功的唯一方法。从上向下命令是无力的。

幻灯片4

报告提纲

(1) 问题空间和解决方案空间。

(2) 一系列专门问题 —— 供讨论用。
(3) 回到顶层问题 —— 供讨论用。
(4) 行动计划 —— 供总体讨论用。
设法从技术研究中将功能需求分离出来。
了解我们是如何受到问题空间和解决方案空间的限制。
是否体系结构除了协议以外别无其他？

幻灯片5

问题空间是什么？

选路与寻址：
大到什么程度，采用何种拓扑结构及选路模型？
逐渐变大：
用户服务；主机和网络采用何种技术？
Internet的舍弃：
计费、控制的使用和修复故障。
新服务：
视频？事务处理？分布计算？
安全性：
终端节点还是网络？ 路由器还是转发器？

幻灯片6

限定解决方案的空间

从当前的状态能迁移到多远？
- 我们能改变IP头吗(除了OSI外)？
- 我们能以命令方式改变主机的需求吗？
- 我们能管理一个长期迁移目标吗？—— 始终如一的方向与多种多样的目标、资金来源。

我们能接受网络级的连通性吗？
- 转发将来会被抛弃吗？
- 安全性以及变换是一个关键议题。
- 需要一个基于转发的体系结构吗？

如何能够和必须管理Internet？
- 我们能管理或者限制网络的连通性吗？

研究开发什么协议？一个还是多个？

幻灯片7

多协议Internet

“ 把问题想得难一点对人类有好处。”
我们是迁移、互操作还是容忍多协议？
- 不是所有的协议集在同一时期都有同样的功能范围。
- Internet需要特定的功能。

声明：基本的矛盾(非宗教性的或恶意的)：

- 满足Internet积极进取的需求。
- 处理OSI迁移。

结论：一种协议必定为主导，其他协议必定为辅助。我们什么时候“切换”到 OSI？

请考察本文下面的每张幻灯片。

幻灯片8

选路和寻址

什么是Internet的目标规模？

- 如何将地址和路由联系起来？
- 拓扑模型是什么？
- 什么是可能的解决方案？

选路要求什么样的策略范围？

- BGP和IDRP是两个解答。问题在那儿？
- 固定类别或可变路径？
- 源控制的选路是最低要求。

如何无缝地支持移动主机 ？

- 新地址类，再捆绑到本地地址，用 DNS吗？

是否要推动Internet组播？

幻灯片9

逐渐变大 —— 一个老题目

(寻址与选路在前一张幻灯片上。)

在下一个十年中需要什么样的用户服务？

- 我们能否构筑一个计划？
- 需要体系结构方面的改变吗？

是否有更好的处理速度、包大小等范围的需求。

- 是否取消分段策略？

我们将支持什么主机范围(如UNIX 环境)？

幻灯片10

处理舍弃

Internet是由独立管理和控制的部分组成的。

为网络收费需要什么支持？

- 体系结构不隐含按容量收费、重记帐和为丢失包付费。
- 是否需要控制以提供记帐标识符或选路？

需求：必须支持有控制共享的链路。 (简单的形式是基于链路标识符的类别)。

- 如何一般化？

对故障隔离是否更加需要？ (我投赞成票！)

- 我们如何能找到可以交谈的经理们？
- 我们需要主机上的服务吗？

幻灯片11

新服务

要支持视频和音频吗？是实时吗？百分比多少？

- 需要计划从研究结果得到什么，什么样的质量？
- 向供货商交底的目标日期。

我们能“更好”地支持事务处理吗？

- TCP能做吗？VMTP呢？介绍呢，还是刹车？

哪些象样的应用即将出笼？

- 分布计算——它真的将发生吗？
- 信息网络技术吗？

幻灯片12

安全性

能坚持说终端节点是唯一防线吗？

- 在网络内部我们能做什么？
- 能要求主机做什么？

能容忍转发器或安排它们的结构吗？

能找到一个更好的方法来构筑安全性边界吗？

需要全球身份验证吗？

有新的主机需求吗？

- 登录。
- 身份验证。
- 管理接口。电话号码或访问点。

附录B 组成员

第1组：选路与寻址

```
Dave Clark, MIT [Chair]
Hans-Werner Braun, SDSC
Noel Chiappa, Consultant
Deborah Estrin, USC
Phill Gross, CNRI
Bob Hinden, BBN
Van Jacobson, LBL
Tony Lauck, DEC.
```

第2组：多协议体系结构

```
Lyman Chapin, BBN [Chair]
Ross Callon, DEC
Dave Crocker, DEC
Christian Huitema, INRIA
Barry Leiner,
Jon Postel, ISI
```

第3组：安全性体系结构

```
Vint Cerf, CNRI [Chair]
```

Steve Crocker, TIS
Steve Kent, BBN
Paul Mockapetris, DARPA

第4组：业务流控制与状态

Robert Braden, ISI [Chair]
Chuck Davin, MIT
Dave Mills, University of Delaware
Claudio Topolcic, CNRI

第5组：现代应用

Russ Hobby, UCDavis [Chair]
Dave Borman, Cray Research
Cliff Lynch, University of California
Joyce K. Reynolds, ISI
Bruce Schatz, University of Arizona
Mike Schwartz, University of Colorado
Greg Vaudreuil, CNRI.

安全性考虑

安全性议题在第4节讨论。

作者地址

David D. Clark
Massachusetts Institute of Technology
Laboratory for Computer Science
545 Main Street
Cambridge, MA 02139

Phone: (617) 253-6003
EMail: ddc@LCS.MIT.EDU

Vinton G. Cerf
Corporation for National Research Initiatives
1895 Preston White Drive, Suite 100
Reston, VA 22091

Phone: (703) 620-8990
EMail: vcerf@nri.reston.va.us

Lyman A. Chapin
Bolt, Beranek & Newman
Mail Stop 20/5b
150 Cambridge Park Drive
Cambridge, MA 02140

Phone: (617) 873-3133
EMail: lyman@BBN.COM

Robert Braden
USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90292

Phone: (310) 822-1511
EMail: braden@isi.edu

Russell Hobby
University of California
Computing Services
Davis, CA 95616

Phone: (916) 752-0236
EMail: rdhobby@ucdavis.edu

RFC 1454 下一版本IP提案的比较

网络工作组 T.Dixon
RFC：1454 RARE
1993年5月

提示

本文为Internet社区提供信息，不指定Internet标准，它的分发不受限制。

摘要

本文是经过少许编辑后重印的RARE技术报告(RTC(93)004)。

下面是当前IP的三个主要替代提案的特点的简短总结。本文并不打算作为详尽的或最终的文本(最后给出简要的参考文献目录以提供更多的信息源)，但可作为讨论这些提案时的参考，由RARE和RIPE来协调。应该承认这些提案本身是“推动目标”的，它反映了在华盛顿举行的第25届IETF会议的见解是完全正确的。 Ross Callon和Paul Tsuchiya对原始草案的评议也包括在内。有一个时期，术语 IPv7用来指IP的下一个版本，但该术语与一个特别提案有关，所以现在用术语IPng来标识下一代IP。

在个别讨论提案前，本文先对为解决问题和达到特定目标的机制作一般性的讨论。

1. 为何当前的IP能力不足？

该问题已经由ROAD小组研究并阐述过，此处简述如下：

- IP B类地址空间耗尽。
- IP 地址空间会全部耗尽。
- 地址分配的非分级结构导致平面的选路空间。

虽然IESG对于新的IP要求比简单选路和寻址议题更深入一步，但正是这些议题使扩展当前协议成为不实际的选择。因而，对提出的各种协议的大部分讨论和开发集中在这些专门问题上。

对这些问题的近期补救办法，包括使用 CIDR提案(CIDR允许以C类网络的集聚来选路)以及以发挥CIDR优势的方式分配 C类网络地址的分配策略。支持 CIDR的选路协议有 OSPF和BGP4。以上这些都不是新IP(IPng)必须具备的先决条件，但是必须延长当前 Internet的生存期，以满足长期解决方案工作的要求。 Ross Callon指出为延长IP生存期有其他选择，他的一些想法已被列在TUBA清单中。正在考虑可使 Internet进一步增长的长期提案。这些提案的时间进度如下：

- 12月15日提出作为 RFC 的议题选择准则。
- 2月12日两个可互操作的实施就绪。
- 2月26日第二个提案的草案文件就绪。

有雄心的目标是在1993年3月在哥伦比亚举行的第26届IETF会议上能作出提案贯彻应用的决定。

当前可选的候选对象有：

• PIP(P IP —— 一个全新的协议)。

• TUBA(具有大地址的TCP/UDP —— 用ISO CLNP)。

• SIP(简单IP —— 具有较大地址和较少选项的IP)。

Robert Ullman有一个更好的提案，不过我对它了解不多。与每个提案候选对象相联系的是过渡计划，但大都独立于提案本身且包含的元素可分开采用，即使对 IPv4，也还要延长当前的设备和系统的生存期。

2. 提案具有的共同点

2.1 较长的地址

所有的提案都为较长的地址字段做了准备，不仅增加了可寻址系统的数量，也方便了路由集聚的地址分级分配。

2.2 基本原理

提案也起源于世界性的“ 选路实现 ”观点 —— 也就是说集中在网络内的选路内部部件而不是集中在终端用户或应用看得见的网络服务上。这或许是不可避免的，尤其是给生产可互操作的设备的时间非常紧。然而在第 25届IETF会议上少数真正的用户代表显然不高兴，因为他们支持最终必须采用新的主机设备。

提案中有一个内置的假设，就是 IPng企图成为一个环球协议：也就是同一网络层协议将可用在同一局域网上的主机之间、主机和路由器之间、同一自治域的路由器之间和不同自治域的路由器之间。在定义分开的“ 接入 ”协议和“ 远程 ”协议上有某些优点，这在需求中没有排除。尽管这是Internet内少有的重要变革机会 ，但要求加速开发和低风险导致提案数不断递增，而不是从根本上变革到经过很好验证的已有的技术上。

一个未进一步陈述的假设是体系结构的目标定在单个连接的主机。目前，要设计允许主机有多个接口，并和单个连接的主机相比，可从增加的带宽和可靠性中获益 (是地址属于接口而不属于主机的缘故)的IPv4网络很困难。正如这些文件中提到的，倾向于拓扑是否存在限制。已经认定不一定是 PIP或TUBA提案的制约，但是相信这是一个议题，到现在为止还没有出现在相当的准则中。

2.3 源选路

已有的 IPv4 对源指定路由已有保证，然而很少用， (有人要反驳我吗？)部分原因是由于需要了解直至路由器级的网络的内部结构。源路由通常是需要使用的，当用户根据策略，要求源和目的地之间的业务倾向或强令通过特殊行政管理域时，源路由也可被行政管理域内的路由器用来指定通过特殊的逻辑拓扑。源指定的选路需要一些性质不同的部件：

(1) 根据技术规范中源的策略来选择路由。

(2) 路由的选择要与其策略相适应。

(3) 用已标识的路由对业务流做标记。

(4) 为已加标记的业务流相应地选路由。

这些步骤不是完全独立的。在这种方法中，第 (3)步标识的路由可能会约束前面步骤中能被选择的路由种类。目的地不可避免地、或者通过告知准备接受的策略，或者通过一个协商过程，加入到源路由的技术规范中去。

所有提案都是通过在每个包中加一串直接地址 (或许部分地指定)来标记源路由。没有规定一个主机取得指定这些直接地址所需信息的过程 (这个阶段不完全不合理，但期望有更多的信息)。这些决定的负面后果是：

- 根据必须指定的直接地址数，包头会变得很长 (虽然当前有指定的机构或想象该机构只指定直接地址的重要部分)。
- 如果个别的直接地址不再可达到，源路由可能必须周期性地重新指定。

正面影响是：

- 域间路由器不必了解策略，只是机械地跟随源路由。
- 路由器不必存储标识路由的上下文，因为信息被指定在每个包头中。
- 路由服务器可定位在网络的任何地方，只要主机知道如何找到它们。

2.4 封装

封装是将一个网络层包封装到另一个包中，以使有效的包能直接通过一条路径，否则就不能到达能移去最外面包的路由器，并指引结果包到它的目的地。封装需要：

(1) 在包中有一指示位，以指示它包含另一个包。

(2) 路由器具备这样的功能，它能在收到一个包后，移去封装并再启动包转发进程。

所有提案都支持封装。由源进行的合适的封装可能会获得源选路的效果。

2.5 组播

所有提案都能协调在地址规范许可的多种范围的组播。 Internet范围的组播尚待进一步研究。

2.6 分段

所有提案都支持中间路由器对包的分段，然而最近的一些讨论，主张从提案中取消该机制，而改为使用MTU发现过程以避免中间分段。这样的决定实际上排除在网络上使用报文计数序列编号的传输协议 (如OSI传输协议)，只有用字节计数并确认的协议 (如TCP)才能在一个连接激活期间处理 MTU还原。OSI传输协议可能不会特别地与 IP界有关联，但是它可能与提供多协议服务的供应商有关联，但是应该注意到对于 IPng支持的服务类型的影响。

2.7 包的生存期

IPv4中的“ 生存期 ” (TTL)字段在每种情况下，作为一个简单的段计数被重新计算，很大程度上以实施方便为基础。虽然老的 TTL很大程度上以这种方式实现，但它以服务于体系结构为目的，在网络中为一个包的生存期设置了一个上限。如果该字段作为一个跳计数而重新计算，那么必定对网络中包的最大生存期有其他的技术规范，所以源主机能保证网络层分段标识符和传输层序列号，当存在混淆危险时，从来不会有重用的危险。事实上，有三个分开

的议题：

(1) 防止选路形成回路(由跳计数解决)。

(2) 限制网络层包的生存期(需要，但目前为止未指定)，支持传输层的设想。

(3) 允许源对包设置更多限制 (例如在拥塞情况下丢弃老的实时业务流，让位给新的业务流，这是一个选项，到目前为止还没作规定)。

3. 提案略为提及的内容

3.1 资源预留

应用日益要求确定的带宽和传输延时，两者对实时视频和音频传送都是必须的。这样的应用需要过程向网络指出它们的需求以及必要的资源预留。这样的过程在某种程度上类似于源路由选择。

(1) 源提出需求的技术规范。

(2) 确认需求能被满足。

(3) 用需求来标记业务流。

(4) 为已加标记的业务流相应地选路由。

按照同样资源需求规定路线发送的业务流有时也称其为流。流的标识需要一个建立过程，人们可能设想与建立源路由使用同样的过程，但两者是有区别的，表现在：

- 在一条路径上的所有路由器必须同意并参予资源预留。
- 由此在每个路由器中相对直接地保持前后关系和短的流标识。
- 在失效时，网络可选择重定路由。

每个提案用各种方法来携带流标识，然而这是目前十分超前的研究。没有确定建立机制。实际上预留资源过程是一个高层次的问题。源选路和资源预留间的交互作用，还需进一步试验：虽然两者性质截然不同，实现的制约也不一样，但两个不同的机制将使得在选择路由时，既要满足策略，又要满足性能指标，变得困难。

3.2 地址分配策略

在IPv4中，地址与系统捆绑在一起是长期的。且在多数情况下，能与 DNS名字互相交换使用。默许地接受地址和一个特殊系统的联系，在 IPng 中可能更为短暂。提案之一的 PIP 是使系统的标识和它的地址之间有区别，并允许捆绑能暂时改变。没有提案规定地址生存期的限制，也未规定地址分配方式必须受特殊协议的约束。例如，由 IPng 提供的较大的地址空间中分配分级地址的高位部分，可以选择是根据与地理位置相关方式，还是参照服务供应商方式。基于地理位置的地址是不变的，也易于分配，但意味着在分配区域内有重新退化到“ 平面 ”地址的危险，除非采取确实的拓扑上的限制。基于供应商的地址分配会造成地址改变 (如果供应商改变)或多个地址(如果多个供应商)。移动主机(依赖于基础技术)不论是基于地理位置还是基于供应商方案都会出现问题。

对地址分配方案以及对地址生存期的影响没有严密的提案，假设捆绑名字到地址的已有的DNS 模型仍然有效是不可能的。

值得提出的是，在地址分配机制和可能采用的自动配置方法之间有交互作用。

3.3 自动配置

对当前IP服务用户最大的担忧是维护基本配置信息的管理工作，诸如为主机分配名字和地址，并要保证信息正确地反映在 DNS中。部分问题是由不良的实施造成的 (或者盲目相信vi和awk是网络管理工具)。不过许多问题通过使过程更自动化而得到减轻。这些可能性 (有些是互斥的)有：

- 分配主机地址使用相对恒定的值，如 LAN 地址。
- 在子网内定义一个动态地址分配协议。
- 定义“ 通用地址 ”，通过使用它，主机不需预先配置便可达本地服务器 (DNS、路由服务器等)。
- 通过检索 DNS，主机便能确定它们的名字。
- 当主机配置改变时，由主机更新它们的名字 /地址捆绑。

当很多提案提及某些以上的可能性时，选择合适的解决方案在一定程度上依赖于地址分配策略。同时，动态配置引起某些困难的理性和实际议题 (确切地说，地址的作用是什么？从什么意义上来讲，当一个主机地址改变时，还是同一个主机吗？如何处理 DNS映射的动态变化，又如何对它们进行身份验证)。

提案小组会发现大部分问题在他们讨论范围以外。当定义和选择 IPng的候选者时，像“ 系统 ”这样的议题没有很好的讨论，看来是一种疏忽。参加者意识到了这一点，看来即使做了决定，某些观念还会在更多的读者范围内重新研究。

然而IP不可能在非技术环境中对有专利权的连网系统 (如Netware、AppleTalk)产生影响，在体系结构中或供应商都没有严格地采用自动配置。我坚信在人们头脑中对如何解决这些议题有想法，只是没有写在纸上而已。

3.4 应用接口/应用协议改变

一些公共应用协议 (如FTP、RPC等)已经确定专门传送32位IP地址，无疑还有其他标准的和专用的协议。也有许多应用简单地把 IP地址当成32位整数来处理。甚至用 BSD套接字试图不透明地处理地址的一些应用，也不明白如何分析语法或打印长地址 (即使套接字结构大到足够容纳它们)。

因此，每个提案需要指定机制，以便当变换发生时，允许已存在的应用程序和接口运行在新的环境下。对于TCP和IPng(也能运行IPv4)，有一个程序设计接口参考技术规范是有用的，它允许开发者现在就开始改变应用程序。从指定过渡机制的所有提案中，就能推断出已存在的应用兼用性。现在还没有迹象表示一个新的接口技术规范独立于所选的协议。

3.5 DNS改变

显然，必须要有能支持新的、长地址的名字到地址的映射服务。所有提案都认为这种服务应该由具有合适定义的新资源记录的 DNS来提供。关于为响应某种查询，用返回“ A ”记录信息的合适性，以及什么信息该首先请求的讨论正在进行。在为建立正确地址所必须的询问次数和由于返回非期望信息破坏已存在的执行过程之间存在着折衷。

为自动配置使用DNS和寻址方案反向转换的规模的讨论很热烈，但没有实质性进展。

4. 提案中没有真正提到的

4.1 拥塞避免

IPv4中路由器用“源抑制”控制报文，向源指示拥塞并有可能不久会丢失包。 TUBA/PIP有一“遭受拥塞”位，它给目的地提供类似的信息。但这些技术规范都没有提供如何使用这些设施的详细说明。因而近年来有许多研究分析实体，他们建议这样的设施不仅可以用来报告拥塞(向传输协议提供信息)，也可减少通过网络层的延迟。每种提案都提供某种形式的拥塞信令，但没有一个指定它使用的机制(或分析该机制实际上是否可用)。

作为网络服务的用户，目前大约有30%的丢弃率，且仅在500英里内来回时间就多至2秒。我对某些提案感兴趣的是网络服务在额外负载的情况下性能下降得很少。

4.2 移动主机

移动主机的一个特征是相对快地移动它们的物理位置和到网络拓扑的连接点。显然对寻址和选路是重要的(不管是地理的还是拓扑的)。到目前为止，没有解决方案的详细技术规范，看来这是一个认识问题。

4.3 计费

IESG的选择准则只要求提案不会阻止为审计和收费目的而进行的信息收集，因此没有提案考虑潜在的计费机制。

4.4 安全性

“网络层安全性议题有待进一步研究”，最好每个候选者都能扩展以显示能提供一定程度的安全性，例如对抗地址欺骗。当资源分配特性允许某些主机为特殊应用要求大量可用带宽时，这一点特别重要。

值得提出的是，提供某种程度的安全性意味着在网络内人工配置安全信息，还必须考虑与自动配置目标的关系。

5. 提案的不同点

每个提案互不相同，正像不同于IPv4一样，原理差别虽小，但会产生重大影响(地址规模的扩充，原理上仅是一个小的差别)。主要的特性差别是：

• PIP

PIP有一个创新的头格式，从而简化了分级、策略和虚电路选路。头中也有一个“含糊”的字段，它的语义在不同的行政管理域可以有不同的定义，它的使用和解释在穿越边界时协商解决，还没有指定控制协议。

• SIP

SIP提供了“最小者”方法——从IPv4头中移去所有不常用的字段，并将地址长度扩展到64位。控制协议基于对ICMP的修改。该提案有处理效率高和易于熟悉的优点。

• TUBA

TUBA基于CLNP(ISO 8473)和ES-IS(ISO 9542)控制协议。TUBA是考虑为了TCP和UDP能在CLNP网络上运行。倾向于TUBA的主要论点是认为能处理网络层协议的路由器已经存在，可扩展的地址提供了宽范围的可供“未来验证”的余量，同时是一个标准和产品会聚的机会。

5.1 PIP

PIP包头包含了一个指令集，供路由器中的转发处理器完成对包的某些动作。在传统协议中，某些字段的内容隐含某些动作。PIP为源端编写指引包通过网络选路的小“程序”提供了灵活性。

PIP地址长度实际上不受限制：网络拓扑分级的每一级成为地址的一部分，同时地址随网络拓扑改变而改变。在完全分级的网络拓扑中，每级所需的选路信息数量可以非常小。因而在实际上，分级的级数将更多地由商界和实用因素来决定，而不是受任何特定的选路协议的制约。一个明显优点是地址的高位部分在本地交换时可以省略，低位部分在源路由中可以省略，减少了主机系统需要知道的拓扑信息数量。

这里有一个假设，就是PIP地址易于改变，所以为了标识，给系统指定另一个参数 PIP标识符。不清楚该参数有何用途，它不能同等地受到 DNS名字的服务(更加紧凑，但同样不需要携带在每个包中，但需要一个额外的检索)。因此，提出了这样的问题：两个潜在可通信的主机系统如何找到可使用的正确的地址。

PIP最复杂的部分莫过于某些头字段的意义是由特定域中相互之间的合同来确定的。专门处理设施的语义(如排队优先级)是全球登记的，但实际使用和在包头中为这些设施申请的编码在不同的域中可以是不同的。在两个域之间用不同编码的边界路由器必须从一种编码映射成另一种。因为路由器和其他域在物理上不一定是相邻的，而是通过“隧道”，因此一个路由器必须了解的潜在编码规则数十分大。相对于更熟悉的“选项”而言，虽然用这样的方案可以节省包头的空间，但是协商这些设施使用和编码的复杂性导致成本增加，以及在每个域边界上对包的再编码，这些才是关心的主题。虽然主机为它们的本地域有可能“预编译”编码规则，还是存在许多潜在的实施上的困难。

虽然PIP在三个提案中提供了最大的灵活性，但对于“希望可用”的情况还需进行更多工作，使其潜在的优点和缺点能暴露得更具体。

5.2 SIP

SIP是一个简单而具有较大地址和较少选项的IP。它的主要优点是甚至比IPv4更容易处理。它的主要缺点是：

- 如果32位地址不够的话，那么64位地址在可预见的未来是否就够了，还远远不清楚。
- 虽然在头字段中有少量“保留”位，但 SIP支持新特性的扩展不明显。真是没有其他什么可说的！

5.3 TUBA

ISO CLNS的特征相当有名，协议与IPv4有很强的文化上的相似性，然而有20字节供网络层寻址。除了谬误的(不是这儿发明的)偏见之外，反对TUBA的主要争论在于TUBA太像IPv4了。除了更大、更灵活的地址外，别无其他贡献。采样试验证明路由器能高效地处理非常长

的地址，但同时长的头很少不给网络带宽带来负面影响。

对建议的控制协议(ISO 9542)有下列异议：

- 根据我以前的经验，如果要合理地容纳大的局域网，路由器发现主机的过程将是低效的，而且会消耗路由器资源，同时在主机上需要十分精确的时间分辨力。 TUBA支持者建议，根据最近的经验， ARP不适用了，但是我想本议题还需要检验。
- 重定向机制实际上是基于 LAN 地址，而不是网络地址，意思是本地路由器将复合的选路决定交给同一LAN上其他路由器。同样，重定向方案 (如IPv4中的方案)重定向到网络地址会造成不必要的额外跳数。要分析哪个解决方案比较好，依赖于构筑的情况。客观地说，该协议的路由器发现部分提供了一个其他提案所没有的机制。通过该机制，主机能定位就近的网关，并能自动配置它们的地址。

6. 过渡计划

为使“老”主机能与“新”主机对话，显然需要一个过渡：

(1) IPng主机也能用IPv4，或

(2) 通过一个中间系统转换。

或者：

(1) 系统间的基础设施有能力承载IPng和IPv4，或

(2) 网络的某些部分用隧道或转换方法将一个协议附在另一个协议中。

各种提案拥护的过渡计划只是简单地将上述方案进行组合。经验表明，不管选择那个协议，事实上以上情况都会发生。

隧道/转换过程的一个问题是必须携带在穿过网络中的 IPv4隧道时的附加信息(外加地址部分)，这可以在数据封装在IPv4包前加一个附加的“头”来实现，或者通过将信息编码作为新IPv4 选项类型来实现。

在前一种情况下，可能要正确地映射出错报文会有困难，因为原始包在返回前被截取；后一种情况，包有被丢弃的危险 (因为IPv4选项不是自描述的，新的选项可能无法通过 IPv4路由器)。这就是为了支持IPng隧道方法而引入IPv4的“新”版本的理由。

另一个替代方案 (在该方案中， IPng主机有两个栈，基础设施可以支持、也可以不支持IPng或IPv4)当然需要一个机制来解决用哪个协议做试验。

7. 随意评议

这是Internet协议中发生的首次根本性改变。因为 Internet是一个可管理的实体，它的发展是与美国政府合同紧密联系的。或许 IETF/IESG/IAB组织结构不可避免地无法管理如此大幅度的改变，但希望提议的新结构在促进共识上获得更大成功。值得注意的是许多觉察到的OSI过程问题(如进步慢，在琐事上派别内争，聚焦在最低层共同特性解决方案上，缺乏对终端用户的考虑等)，用它们来处理IPng是危险的，同时关注着由网络设计的广泛参与所带来的困难会到什么程度。

三个主要提案在IPng上很少有实质性的差别，但选择 IPng的竞争过程如不成功就是失败。在这方面，选择过程的结果没有什么特别意义，但在过程本身中为了修复 Internet工程过程的社会和技术凝聚力，或许是必要的。

8. 更多的信息

提案的主要讨论清单如下：

```
TUBA:          tuba@lanl.gov
PIP:           pip@thumper.bellcore.com
SIP:           sip@caldera.usc.edu
General:       big-internet@munnari.oz.au

(Requests to: <list name>-request@<host>)
```

各种提案的Internet草案和RFC，仍能在惯常场合找到。

安全性考虑

安全性议题未在本文中讨论。

作者地址


```
Tim Dixon
RARE Secretariat
Singel 466-468
NL-1017AW Amsterdam
(Netherlands)

Phone: +31 20 639 1131 or + 44 91 232 0936
EMail: dixon@rare.nl or Tim.Dixon@newcastle.ac.uk
```

RFC 1671 向IPng过渡和其他考虑的白皮书

网络工作组 B.Carpenter
RFC：1671 CERN
类别：信息类 1994年8月

提示

本文为Internet社区提供信息，不指定任何种类的Internet标准。本文的分发不受限制。

摘要

本文是响应RFC 1550而向IETF IPng领域提交的文件。本文的发布并不意味着 IPng领域接受文中所表达的任何思想。评议请提交给 big-internet@munnari.oz.au邮件列表。

总结

本白皮书在所选领域勾画了IPng某些通用需求。下面表示的是逐级过渡的需求：

(1) 在网络的每级和每层实现互通。
(2) 包头转换被认为是有害的。
(3) 共存。
(4) IPv4到IPng地址映射。
(5) 双栈主机。
(6) 域名系统(DNS)。
(7) 智能双栈代码。
(8) 智能管理工具。

接受某些关于物理和逻辑组播的论点，并建议需要一个 IPng在ATM上运行的模型。最后，本文建议的策略选路、计费和安全防火墙等需求，需要所有 IPng包携带所涉及事务处理类型的踪迹，以及它们的源和目的地址。

过渡和发展

显然过渡需要几年的时间，同时网络中的每个站点不得不决定它自己的阶段过渡计划。只有那些最小的站点可能在 ISP的压力下，考虑一步到位 (“ 标志日 ”)的过渡。此外，一旦决定采用IPng，那么Internet和所有用Internet协议集的专用网在下一十年 (或更长)的活动，将受到IPng发展的强大影响。用户站点注视着决策，是否和他们过去所看到的在改变程序设计语言或操作系统时所用的同样方法来改变 IPv4。向IPng转变可能不是必然的结果。他们主要担心是，改变是否能使成本和影响生产的风险减到最低。

这样的担心立刻对IPng过渡和发展的模型产生了强大的约束。这些约束中的某些列在下面，并对每种约束赋予简短的解释。

术语“ IPv4主机 ”是一个和今天的主机运行同样内容，而没有维护版本及配置改变的主机。“ IPng主机 ”是一个运行IP新版本，经过重新配置的主机。它类似于路由器。

1. 网络的每级和每层实现互通

这是主要约束。计算机系统、路由器和应用软件厂商肯定不会协调他们产品的发布日期。用户将继续运行他们的老设备和软件。因此， IPv4和IPng主机和路由器的任何组合必须能互通(即加入到UDP和TCP会话中)。一个IPv4包必须能找到从任何 IPv4主机到其他任何 IPv4或IPng主机的路径，反之亦然，穿过 IPv4和IPng路由器的混合路径， IPv4主机无需修改。 IPv4路由器无需修改可与 IPng路由器互通。另外，一个“ 明白 ” IPv4但还“ 不明白 ” IPng的应用软件包必须能在运行 IPv4的计算机系统上运行，并和 IPng主机通信。例如，欧洲的一个老 PC机应该可访问美国的 NIC服务器，即使NIC服务器运行的是IPng，且北美的选路机制只是部分地变换过。或者某个公司某个部门的一个 C类网络应该保持对运行IPng的公司服务器完全的访问，尽管C类网络内部什么也没有改变。

(并不要求一个只能在 IPv4上运行的应用程序到一个 IPng主机上运行。因此，我们承认某些主机一直要等到所有它们的应用程序是 IPng兼容后才能升级。换句话说，我们承认某种程度上API要改变。然而，即使这样的放松，还是有争议的，甚至有些厂商要求在 IPng主机上严格保持IPv4 API。)

2. 包头转换被认为是有害的

该作者相信在任何过渡情况下，要求 IPv4和IPng间动态包头转换将会造成几乎是不可克服的实际困难：

(1) 可以认为IPng功能将是IPv4功能的一个超集。然而，协议间的成功转换要求被转换的两个协议的功能事实上应该相同。为此，应用需要知道它们什么时候通过 IPng API和设在网络中某处的转换器与IPv4主机互通，以便只用IPv4功能。这是不现实的约束。

(2) 转换器的管理对大的站点而言是完全行不通的，除非转换机制是完全隐蔽和自动的。特别是任何转换机制要求为每个主机中的表格 (如DNS表或路由器表)人工地保持专门标志以指示需要转换，这样做是完全不可能进行管理的。在一个有几千台运行多种操作系统的主机的站点上，主机在不同软件版本上前进或后退，使得继续用这种方法来不断跟踪所需要的这些标志的状态是不可能的。整个 Internet的多样化，将会导致混乱、复合的失效模式和困难的诊断。特别是不可能遵守(1)的约束。

实践中为了避免混乱，对转换所需要的知识、所涉及的站点将决不泄漏，并且如果还没有这样的知识的话，当需要时，应用程序不能将其本身限制在 IPv4功能上。

为了避免混淆，此处所讨论的包头转换和地址转换 (NAT)不是同一件事。本文不讨论地址转换。

本文不详细处理性能议题，但转换的另一个明显缺点是带来必然的开销。

3. 共存

Internet基础设施(不论是公用的还是专用的)必须允许IPv4和IPng在同一路由器和同一物理路径上共存。

为了在不要求主机步调一致地更新，及不使用转换器的情况下，网络基础设施能更新至IPng，共存是必须的。

值得注意的是，这种需求并不强制使用有关公共的或是分离的方法来进行选路。作为共存机制，也不排斥使用封装。

4. IPv4到IPng地址映射

人们必须明白过渡期间会遇到什么问题。虽然IPng地址的自动配置可能是所期望的目的，如

果在给定的站点上，IPv4和IPng地址之间有一个可选的简单映射，那么过渡的管理就会大大简化。

因此，IPng地址空间应包含IPv4地址的映射，这样(如果一个站点或服务供应商愿意做的话)一个系统的IPv4地址能机械地被转换成IPng地址，大多数倾向于加一个前缀。对每个站点而言，前缀不一定是相同的，可能至少是服务供应商指定的。

这并不意味着这种地址映射会用作动态转换(虽然有可能是)，或将IPv4选路嵌入IPng选路内(虽然有可能是)。主要目的是简化网络运行者的过渡规划。

顺便指出，这样的需求实际上没有假设IPv4地址是全球唯一的。在建立IPv4和IPng选路域与分级之间的关系上也没有太多帮助。没有理由设想它们之间是1：1对应关系。

5. 双栈主机

无转换的逐步过渡是很难想象的，除非大部分主机同时能运行 IPng和IPv4。如果A想和B(IPng主机)以及和C(IPv4主机)交谈，于是A或者B必须能运行IPv4和IPng两者。换句话说，所有运行IPng主机必须仍能运行IPv4。只能运行IPng的主机在过渡期间是不允许的。

这样的需求并不意味着IPng主机真的有两种完全分离的IP实现(双栈和双API)，但是表现出好像是分离的。封装是兼容的(即两个栈中的一个可为另一个封装包)。

显然，对双栈主机的管理，由于上面提到的地址映射而简化了。除了 IPv4地址以外，只有站点前缀必须配置(人工或动态地)。

在双栈主机中，即使IPng API和IPv4 API是作为一个单个实体实现的，但在逻辑上是可区别的。应用程序将从API得知它们在用IPng还是IPv4。

6. DNS

双栈要求隐含了DNS必须给IPng主机回答IPv4和IPng两者的地址，或将两者编码在一起的单个回答。

如果在DNS中，一个主机附属于一个IPng地址，但该主机实际上还没有运行 IPng，尤如在IPng空间中出现一个黑洞 —— 见下一点。

7. 智能双栈代码

双栈代码可从DNS得到两个地址，用哪一个呢？多年的过渡期间 Internet将包含黑洞。例如，从IPng主机A到IPng主机B路上某处有时(不可预测)会遇到只运行IPv4的路由器，它将丢弃IPng包。同样，DNS的状态也不一定与现实是一致的。 DNS声称知道IPng地址的主机可能在一特别的时刻并没有运行 IPng，因此到那个主机的IPng包在传递时将被丢弃。知道一个主机具有IPv4和IPng两种地址并没有给出有关黑洞的信息。对此必须有一种解决方案，这方案不能依赖于人工保持的信息。(如果这个不解决，双栈方法是不会好于转换方法的。)

8. 智能管理工具

过渡期间需要一整套管理工具。为什么 IPng路由不同于IPv4路由？如果要转换的话，该发生在何处？何处有黑洞？(宇宙学家喜欢同样的工具。)今天的主机是否真正有IPng能力？

组播

众所周知，IPng必须支持组播应用，体系结构上一个明显的规则是：不论是 LAN还是WAN线路，组播包不应在同一线路上经过两次。如果做不到这点，则意味着同时组播的事务处理最大数量会减半。

LAN上的IPv4的一个负面特征是：轻率地使用物理广播包，诸如 ARP(各种非IETF盲目模

仿者)协议。在大的LAN上，这将导致一系列不希望有的后果 (经常是由于差的产品或差的用户，而不是协议设计本身造成的。)如有可能的话，体系结构上明显的规则是改用单播 (或最坏情况，用组播)来代替物理广播。

ATM

网络工业界正在ATM上大量投资。没有IPng提案似乎是可取的(从获得管理部门批准的意义上)，除非它是“ ATM兼容 ”的，也就是要有一个如何运行在 ATM网络上清晰的模型。虽然不马上需要一个像RFC 1577那样十分详细的文件，但必须显示该基本模型是能工作的。

类似的论点同样可用于X.25、帧中继、 SMDS等，但ATM是当前呼声最高的。

策略选路与计费

遗憾的是，这不能被忽略，许多人对此感兴趣。基金代理希望业务流流经提供基金的线路，且在以后他们要知道有多少业务流。计费信息也可用作网络规划和反向付费的根据。

所以IPng及它的选路过程允许根据详细的源和目的地址来指定业务流的途径。 (作为一个例子，从 MIT物理系输出的业务流和任何其他系输出的业务流可能通过不同的路由到CERN。)

满足该需求的一个简单途径是坚持IPng必须支持基于供应商的寻址和选路方案。

业务流的计费要求同样的详细程度 (甚至更详细，例如ftp的业务流是多少， www的业务流又是多少)。

两者都需要花费时间和金钱，并且不仅影响IP层，所以IPng不该回避它们。

安全性考虑

公司网络运行者和校园网络运行者曾受到过好几次安全问题的困扰，他们对待此事比许多协议专家更认真。实际上，许多公司网络运行者希望在向 IPng过渡中，作为一个比其他任何议题更为紧迫的议题，安全性能得以改善。

因为IPng估计是一个数据报协议，限制了它为端到端的安全性所能做的工作， IPng必须允许路由器中有比IPv4更有效的防火墙。特别需要基于源和目的地址以及事务处理类型的高效的业务流阻挡。

看来需要同样的特征允许策略选路和详细计费来改善防火墙的安全性。讨论这些特征的细节超出了本文范围，但是看来不大可能在边界路由器中限制实现细节。为了检查有害的业务流，允许基于策略的源选路和/或允许详细计费，包必须携带某些三重验证的踪迹 (源、目的地、事务处理)。可能所有IPng可以在每个包中以某种格式携带源和目的地的标识符，但是标识事务处理的类型或甚至于个别的事务处理，是一个额外需求。

声明和致谢

以下是个人观点，未必代表我老板的观点。

近几年来， CERN已经通过三个网络的过渡 (由John Gamble解决的IPv4重编号，由Mike Gerard解决的AppleTalk从阶段I向阶段II的过渡，以及由Denise Heagerty解决的DECnet阶段IV向DECnet/OSI选路的过渡)。如果没有从他们那儿获取知识，我是不能写出这个文件的。我也

从许多人，特别是从IPng董事会的各个成员的讨论或作品中，获益非浅。多位董事会成员及波音公司的Bruce L Hutfless提出了意义帮助阐明本文。不过意见是我本人的，并非董事会全体成员的共同意见。

作者地址

Brian E. Carpenter
Group Leader, Communications Systems
Computing and Networks Division
CERN
European Laboratory for Particle Physics
1211 Geneva 23, Switzerland

Phone: +41 22 767-4967
Fax: +41 22 767-7155
Telex: 419000 cer ch
EMail: brian@dxcoms.cern.ch

RFC 1715 地址分配效率比例系数H

网络工作组 C.Huitema
RFC：1715 INRIA
类别：信息类 1994年11月

提示

本文是为Internet社区提供信息，并不指定任何类的Internet标准。本文的分发不受限制。

摘要

本文是响应RFC 1550而向IETF IPng领域提交的文件。本文的发布并不意味着IPng领域接受文中所表达的任何思想。评议可提交给作者或发邮件至sipp@sunroof.eng.sun.com邮件列表。

目录

1. 地址分配的效率 …… [1]
2. 估计合理的系数H值 …… [1]
3. 评估提出的地址计划 …… [2]
4. 安全性考虑 …… [2]
5. 作者地址 …… [2]

1. 地址分配的效率

IPng争论的实质性部分集中于地址长度的选择。一个重复的概念是“分配效率”，参加讨论的大部分人表示，效率是网络中有效的系统数对最大理论值之比。例如， 32位的IP 寻址计划理论上超过70亿个系统，而目前DNS中记录有大约3500万个地址，说明效率为 0.05%。

但是这种经典评估是误导，因为它没有考虑分级的级数。例如 IP地址至少分为三级：网络、子网和主机。为了排除这些相关性，建议对效率系数用一个对数尺度：

H=log(目标数)/可用位数

系数H不至于太依赖于分级的数量。例如：设想在两级之间选择，每级用 8位编码。若单级则用16位编码。如果在每个8位级平均分配100个元素，或在单个16位级平均分配10 000个元素，我们将获得同样的效率。

为便于心算，以下用的是以10为底的对数。当数变大时，人们习惯于用 10的指数来表示。这样可以说“ IPng能编号1 E+15个系统”。如果遵循这样的单位选择， H就在0与理论最大值0.30103(log 2)之间变化。

2. 估算合理的系数H值

我们并不指望在实际中获得系数值为 0.3。关心的问题是推断该值为合理的期望值。我们可以试着从已有的编号计划来评估它。特别感兴趣的是考虑计划打破时，即当人们被迫对电话号码增加数字时，或者计算机地址增加位数时。我手头有若干这样的数字：

- 当电话号码数达到一个门限 1.0E+7时，所有法国的电话号码加 1位数字，由8位加到9位数。对数值为7，位数大约为27(一个十进制数大约为3.3位)。则系数就是0.26。

- 扩展美国电话系统的区域号，使其为 10位数，可以有1.0E+8个用户。对数值为8，位数为33位，系数大约为0.24。
- 扩展Internet地址长度，从32位至某一值。当前32位网上大约有300万个主机。3.0E+6的对数大约为6.5，这样得出的系数为0.2。我们相信32位还足够用好几年。如果主机数乘10，系数升至0.23。
- 扩展SITA 7号码地址的长度。按照他们的文件，在他们的网络中，大约有 64 000个可编址的点，分散在1200个城市，180个国家中。一个上限情况提供5位码供寻址用，造成效率为0.14。这是一种极端情况，因为SITA在它的分级中用固定长度的令牌。
- 全球连通的物理/空间科学DECnet网(阶段IV)到15 000个节点时停止增长(即新节点隐藏)，在16位空间中给出系数为0.26。
- 在46位空间中，大约有2亿个IEEE 802节点，给出系数为0.18。然而，这个号码空间没有饱和。

从以上例子，可以推测出效率系数通常在0.14到0.26之间。

3. 评估提出的地址计划

用反向计算，可得到网络中寻址的设备总数：

	悲观的估计(0.14)	乐观的估计(0.26)
32位	3 E+4(!)	2 E+8
64位	9 E+84	E+16
80位	1.6E+11	2.6 E+27
128位	8 E+17	2 E+33

数字对于为什么有些人认为64位“不够”，而另一些人则认为“有足够余量”解释得很好。根据分配效率，或者远低于目标，或者远高于目标。我的观点是 128位足足有余。甚至我们假设效率最低，仍有超过1.E+15台Internet主机的冗余估计。

同时值得提出的是，如果我们给网络贡献80位，并为“缺少自动配置的服务器”提供48位，在悲观的情况下，仍能编号多于E+11个网络；要达到E+12个网络，只要取效率系数为0.15。

这就是为什么我认为128位在下一30年中是完全安全的解释。必须包括在地址分配内的制约程度，显示出与今天我们所知道如何做的是非常一致的。

4. 安全性考虑

安全性议题不在本文中讨论。

5. 作者地址

Christian Huitema
INRIA, Sophia-Antipolis
2004 Route des Lucioles
BP 109
F-06561 Valbonne Cedex
France

Phone: +33 93 65 77 15
EMail: Christian.Huitema@MIRSA.INRIA.FR

RFC 2373 IPv6寻址体系结构

网络工作组 R.Hinden
RFC：2373 诺基亚公司
撤销：1884 S.Deering
分类：标准跟踪 Cisco公司

提示

本文为Internet社区指定一个Internet标准跟踪协议，并请求为改进进行讨论和提出建议。对标准化状态和本协议的状况，请参考“ Internet正式协议标准 ”(STD.1)最新版本。本文的分发不受限制。

版权声明

摘要

本技术规范定义IPv6[IPv6]的寻址体系结构。本文件包括IPv6寻址模型、IPv6地址的文字表示、IPv6单播地址、任意点播地址和组播地址的定义以及IPv6节点需要的地址。

目录

1. 概述 …… [2]
2. IPv6寻址 …… [2]
2.1寻址模型 …… [2]
2.2地址的文本表示 …… [3]
2.3地址前缀的文本表示 …… [3]
2.4地址类型表示 …… [4]
2.5单播地址 …… [5]
2.5.1接口标识符 …… [5]
2.5.2未指定的地址 …… [6]
2.5.3回返地址 …… [6]
2.5.4嵌有IPv4地址的IPv6地址 …… [6]
2.5.5NSAP地址 …… [7]
2.5.6IPX地址 …… [7]
2.5.7可集聚全球单播地址 …… [7]
2.5.8本地使用的IPv6单播地址 …… [7]
2.6任意点播地址 …… [8]
2.7组播地址 …… [9]
2.7.1预定义的组播地址…… [10]
2.7.2新IPv6组播地址的分配…… [10]

2.8节点需要的地址 …… [11]
3. 安全性考虑 …… [11]
附录A 创建基于EUI 的接口标识符 …… [11]
附录B 文本表示的ABNF描述 …… [13]
附录C RFC 1884的变化…… [13]
参考文献 …… [14]
作者地址 …… [15]
版权声明 …… [15]

1. 概述

本技术规范定义了IPv6的寻址体系结构。包括当前定义的IPv6[IPv6]地址格式的详细描述。

作者衷心感谢Paul Francis, Scott Bradner, Jim Bound, Brian Carpenter, Matt Crawford, Deborah Estrin, Roger Fajman, Bob Fink, Peter Ford, Bob Gilligan, Dimitry Haskin, Tom Harsch, Christian Huitema, Tony Li, Greg Minshall, Thomas Narten, Erik Nordmark, Yakov Rekhter, Bill Simpson和Sue Thomson所做的努力。

2. IPv6寻址

IPv6地址为接口和接口组指定了128位的标识符。有三种地址类型：

- 单播。一个单接口有一个标识符。发送给一个单播地址的包传递到由该地址标识的接口上。
- 任意点播。一般属于不同节点的一组接口有一个标识符。发送给一个任意点播地址的包传送到该地址标识的、根据选路协议距离度量最近的一个接口上。
- 组播。一般属于不同节点的一组接口有一个标识符。发送给一个组播地址的包传递到该地址所标识的所有接口上。

在IPv6中没有广播地址，它的功能正在被组播地址所代替。在本文中，地址内的字段给予一个规定的名字，例如“用户”。当名字后加上标识符一起使用(如“用户ID”)时，则用来表示名字字段的内容。当名字和前缀一起使用时(如“用户前缀”)则表示一直到包括本字段在内的全部地址。

在IPv6中，任何全“0”和全“1”的字段都是合法值，除非特殊地排除在外的。特别是前缀可以包含“0”值字段或以“0”为终结。

2.1 寻址模型

所有类型的IPv6地址都被分配到接口，而不是节点。一个IPv6单播地址属于单个接口。因为每个接口属于单个节点，多个接口的节点，其单播地址中的任何一个可以用作该节点的标识符。所有接口至少需要有一个链路本地单播地址(见2.8节额外需要的地址)。一个单接口可以指定任何类型的多个IPv6地址(单播、任意点播、组播)或范围。具有大于链路范围的单播地址，对这样的接口是不需要的，也就是从非邻居或者到非邻居的这些接口，不是任何IPv6包的起源或目的地。这有时适用于点到点接口。对这样的寻址模型有一个例外：

如果处理多个物理接口的实现呈现在Internet层好像一个接口的话，一个单播地址或一组单播地址可以分配给多个物理接口。这对于在多个物理接口上负载共享很有用。

目前的IPv6延伸了IPv4模型，一个子集前缀与一条链路相关联。多个子集前缀可以指定给同一链路。

2.2 地址的文本表示

用文本串表示的IPv6地址有三种规范形式：

(1) 优先选用的形式为x:x:x:x:x:x:x:x:，其中x是8个16位地址段的十六进制值。

例如：

FEDC : BA98 : 7654 : 3210 : FEDC : BA98 : 7654 : 3210

1080 : 0 : 0 : 0 : 8 : 800 : 200C : 417A

个别字段中前面的0可以不写，但是每段必须至少有一位数字 ((2)中描述的情形除外)。

(2) 在分配某种形式的IPv6地址时，会发生包含长串0位的地址。为了简化包含0位地址的书写，指定了一个特殊的语法来压缩 0。使用“ ::”符号指示有多个0值的16 位组。“ ::”符号在一个地址中只能出现一次。该符号也能用来压缩地址中前部和尾部的 0。

用下面的例子来说明：

1080:0:0:0:8:800:200C:417A	单播地址
FF01:0:0:0:0:0:0:101	组播地址
0:0:0:0:0:0:0:1	回返地址
0:0:0:0:0:0:0:0	未指定地址

可用下面的压缩格式表示：

1080::8:800:200C:417A	单播地址
FF01::101	组播地址
::1	回返地址
::	未指定地址

(3) 当谈到IPv4和IPv6节点这样一个混合环境时，有时更适合于采用另一种表示形式：

x:x:x:x:x:x:d.d.d.d,其中x是地址中6个高阶16位段的十六进制值， d是地址中4个低价8位段的十进制值(标准IPv4表示)。举例说明：

0:0:0:0:0:0:13.1.68.3

0:0:0:0:0:FFFF:129.144.52.38

写成压缩形式为：

::13.1.68.3

::FFFF.129.144.52.38

2.3 地址前缀的文本表示

IPv6地址前缀的表示方式和 IPv4地址前缀在CIDR中的表示方式很相似。一个 IPv6地址前缀可以表示为如下的形式：

IPv6地址/前缀长度

其中，IPv6地址是2.2节中表示的任何形式的 IPv6地址。而前缀长度是组成前缀的十进制值，说明地址最左边的连续的地址位的长度。

例如，60位长的前缀12AB00000000CD3(十六进制)可用下面的合法格式来表示：

12AB:0000:0000:CD30:0000:0000:0000:0000/60

12AB::CD30:0:0:0:0/60

12AB:0:0:CD30::/60

但是，下面的表示方式是不合法的。

12AB:0:0:CD3/60 在任何一个16位段的地址块中，可以省略前部的0。但不能省略尾部的0。

12AB::CD30/60 /左边的地址会展开成12AB:0000:0000:0000:0000:000:0000:CD30

12AB::CD3/60 /左边的地址会展开成12AB:0000:0000:0000:0000:000:0000:0CD3

当书写节点地址和它的子网前缀两者时，可以组合成如下表示：

节点地址：

12AB:0:0:CD30:123:4567:89AB:CDEF

和它的子网号：

12AB:0:0:CD30::/60

可以缩写成为：

12AB:0:0:CD30:123:4567:89AB:CDEF/60

2.4 地址类型表示

一个IPv6地址的具体类型是由地址的前面几位来指定的。包含这前面几位的可变长度字段称为格式前缀(FP)。这些前缀的初始分配如下：

分 配	前缀(二进制)	占地址空间的百分率
保留	0000 0000	1/256
未分配	0000 000	11/256
为NSAP地址保留	0000 001	1/128
为IPX地址保留	0000 010	1/128
未分配	0000 011	1/128
未分配	0000	11/32
未分配	0001	1/16
可集聚全球单播地址	001	1/8
未分配	010	1/8
未分配	011	1/8
未分配	100	1/8
未分配	101	1/8
未分配	110	1/8
未分配	1110	1/16
未分配	1111 0	1/32
未分配	1111 10	1/64
未分配	1111 110	1/128
未分配	1111 1110 0	1/512
链路本地单播地址	1111 1110 10	1/1024
站点本地单播地址	1111 1110 11	1/1024
组播地址	1111 1111	1/256

注：1. 未指定地址(见2.5.2节)、回返地址(见2.5.3节)，和嵌入IPv4地址的IPv6地址(见2.5.4节)的分配在格式前缀空间0000-0000以外。

2. 除了组播地址(1111 1111)外，格式前缀空间001到111，在EUI-64格式中都要求必须有64 位接口标识符。参见2.5.1节中的定义。

这样的分配方案支持可集聚地址、本地用地址和组播地址的直接分配，并有保留给 NSAP地址和IPX地址的空间。其余的地址空间留给将来用。可用于已有使用的扩展 (如附加可集聚地址等)或者新的用途(如将定位符和标识符分开)。地址空间的15%是初始分配的，其余85%的地址空间留作将来使用。

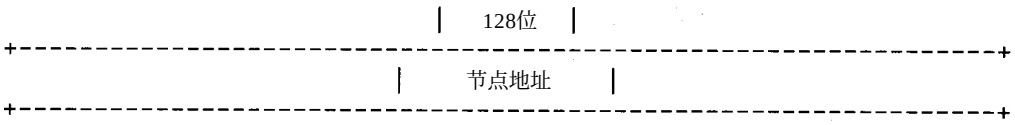
单播地址和组播地址是由地址的高阶字节值来区分的：值为 FF(1111 1111)标识一个地址为组播地址，其他值则标识一个地址为单播地址。任意点播地址取自单播地址空间，和单播地址在语法上是无法区分的。

2.5 单播地址

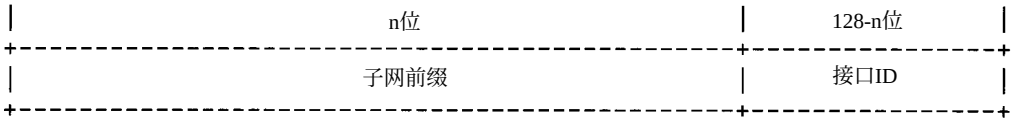
IPv6单播地址是用连续的位掩码集聚的地址，类似于 CIDR的IPv4地址。

IPv6中的单播地址分配有多种形式，包括全部可集聚全球单播地址、 NSAP地址、IPX分级地址、站点本地地址、链路本地地址以及运行 IPv4的主机地址。将来还可以定义另外的地址类型。

IPv6节点对IPv6地址的内部结构可能知之甚多或知之甚少，这是由节点的作用决定的 (例如，主机还是路由器)。在最简单的情况下，节点把单播地址 (包括它本身)看成是无内部结构的、如下图所表示的128位地址。



一个稍完善但仍很简单的主机可能还知道它所连接的链路的子网前缀，在这种场合下，不同地址可能有不同值。



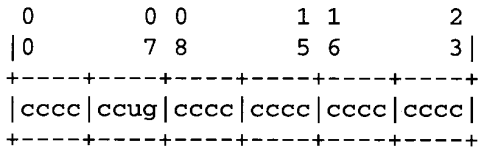
更完善的主机可能知道单播地址中其他分级边界。虽然一个非常简单的路由器可能对 IPv6单播地址的内部结构一无所知，但为了运行选路协议，路由器对一个或多个分级边界要有更为普遍的知识。知道边界随路由器不同而不同，是由路由器在选路分级中所处的位置决定的。

2.5.1 接口标识符

在IPv6单播地址中接口标识符用来标识链路的接口。标识符在该链路上应是唯一的。也可能在较宽范围内是唯一的。在许多情况下，一个接口标识符与该接口的链路层地址相同。在一个单节点上，同一个接口标识符可以用在多个接口上。

在一个单节点的多个接口上，用同样的接口标识符不会影响接口标识符的全球唯一性，或由接口标识符创建的每个IPv6地址的全球唯一性。

在许多格式前缀中 (见2.4节)，接口标识符要求64位长，并构成IEEE EUI-64格式。基于EUI-64的接口标识符，当全球令牌可用时 (如IEEE 48位MAC)，具有全球范围的意义。当全球令牌不可用时 (如串行链路、隧道终点等)，则只具有本地范围的意义。当由 EUI-64形成接口标识符时，若u位(IEEE EUI-64术语中称全球/本地位)置1，则表示全球范围；若u位置0，则表示本地范围。一个EUI-64标识符的头三个字节的二进制表示如下所示。



按Internet标准中的位序，其中u是全球/本地位，g是个体/团体位，c是公司标识符。“ 附录A 创建基于EUI-64接口标识符 ” 为不同的基于EUI-64接口标识符的创建提供了实例。

当形成接口标识符时，使用 u位的动机是当硬件令牌不可用，即在串行链路、隧道终点等情况下，便于系统管理员人工配置本地范围标识符。另一种方法是用 0200:0:0:1、0200:0:0:2等形式代替十分简单的::1、::2等形式。

在IEEE EUI-64标识符中使用全球/本地位的目的是为了将来技术的发展能利用具有全球范围的接口标识符所带来的好处。

形成接口标识符的细节定义在 IP over<link>技术规范中，诸如IP over Ethernet[ETHER]、IP over FDDI[FDDI]等。

2.5.2 未指定地址

地址0:0:0:0:0:0:0:0称为未指定地址。它不能分配给任何节点。意思是没有这个地址。它的一个应用示例是初始化主机时，在主机未取得自己的地址以前，可在它发送的任何 IPv6包的源地址字段放上未指定地址。

未指定地址不能在IPv6包中用作目的地址，也不能用在IPv6选路头中。

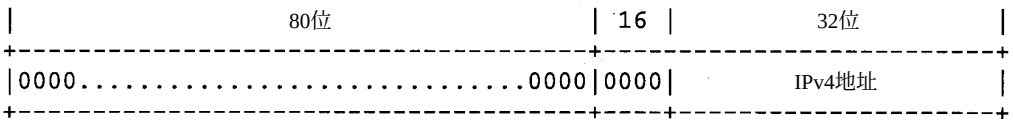
2.5.3 回返地址

单播地址0:0:0:0:0:0:0:1称为回返地址。节点用它来向自身发送 IPv6包。它不能分配给任何物理接口。可以设想它正在与一个虚拟接口相关联 (如回返接口)。

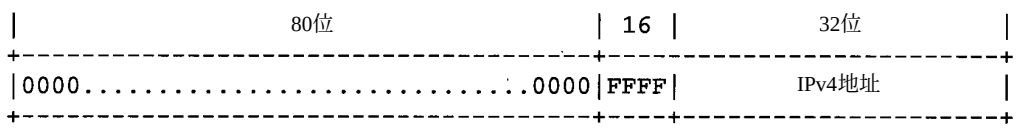
发送到单节点外的IPv6包回返地址必须用作源地址。具有一个目的地址为回返地址的包不应发送出单节点之外，IPv6路由器也不会转发这样的包。

2.5.4 嵌有IPv4地址的IPv6地址

IPv6过渡机制[TRAN]包括一种技术，使主机和路由器能在 IPv4选路基础设施上动态地以隧道方法传送IPv6包。使用该技术的IPv6节点要指定特殊的IPv6单播地址，它在低阶 32位上携带IPv4地址。这种地址类型称其为 “ 与IPv4兼容的IPv6地址 ”，并具有下面的格式：



第二种类型的IPv6地址嵌有IPv4地址。该地址用来表示只支持IPv4，而不支持IPv6的节点的IPv6地址。这种地址类型称为 “ 与IPv4映射的IPv6地址 ”，并具有下面的格式：

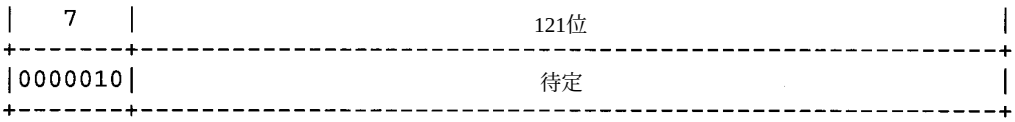


2.5.5 NSAP地址

NSAP地址到IPv6地址的映射定义在[NSAP]中。对于已经规划或应用OSI NSAP寻址计划，并希望应用IPv6或向IPv6过渡的网络实现者，该文件应该重新设计成IPv6寻址计划来满足他们的需要。另外还定义了一套机制，用来在IPv6网络中支持OSI NSAP寻址。如果需要这种支持的话，则必须要有这样的机制。该文件还定义了OSI地址格式内IPv6地址的映射，这应该是必需的。

2.5.6 IPX地址

IPX地址到IPv6地址的映射表示如下：

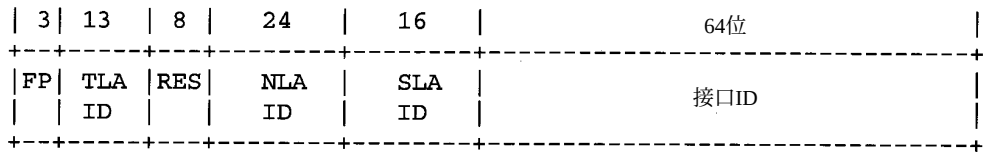


本草案的定义、动机和使用正在研究中。

2.5.7 可集聚全球单播地址

全部可集聚全球单播地址定义在[AGGR]中。设计这样的地址格式为了既支持基于当前供应商的集聚，又支持被称为交换局的新的集聚类型。其组合使高效的选路集聚可用于直接连接到供应商和连接到交换局两者的站点上。站点可以选择连接到两种类型中的任何一种集聚点。

IPv6可集聚全球单播地址格式如下所示：

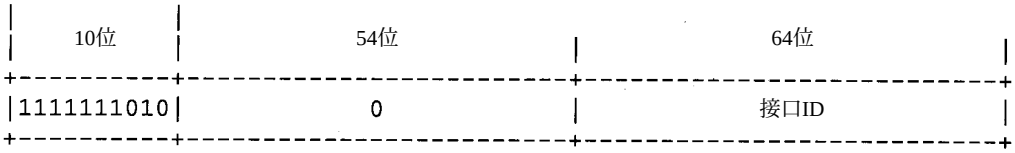


其中，001(FP)用于可集聚全球单播地址的格式前缀(3位)；TLA ID为顶级集聚标识符；RES保留将来用；NLA ID为下一级集聚标识符；SLA ID为站点级集聚标识符；INTERFACE ID为接口标识符。

在[AGGR]中，还规定了内容、字段长度和分配规则。

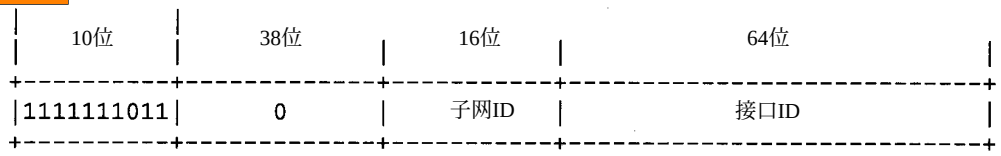
2.5.8 本地用IPv6单播地址

规定了链路本地和站点本地两种类型的本地使用单播地址。链路本地地址用在单链路上，而站点本地地址用在单站点上。链路本地地址格式表示如下：



设计链路本地地址的目的是为了用于诸如自动地址配置、邻居发现或无路由器存在的单链路的寻址。路由器不能将带有链路本地源地址或目的地址的任何包转发到其他链路上去。

站点本地地址具有下面的地址格式：



站点本地地址的设计目的是为了用于无需全球前缀的站点内部寻址。

路由器不应转发站点外具有站点本地源或目的地址的任何包。

2.6 任意点播地址

IPv6任意点播地址是分配给一般属于不同节点的多个接口。根据这个特性，发送给任意点播地址的包，总是发送到具有该地址并按照选路协议测得距离为最近的接口。

任意点播地址从单播地址空间分配而来，可用任何一种规定的单播地址格式。这样，任意点播地址和单播地址在语法上是无法区别的。当一个单播地址分配给多个接口时，如果把它转为任意点播地址，那么被分配该地址的节点，必须显式地配置，以便知道这是一个任意点播地址。

对于任何已分配的单播地址，有一个最长的地址前缀 P用于标识拓扑地区。在该地区中，所有接口均属于该任意点播地址。在由 P标识的区域内，任意点播组的每个成员，被告知在选路系统中作为一个独立实体 (通常称之为“ 主机路由 ”)。在P标识的区域以外，任意点播地址可以集合在前缀P的选路通告中。

在最坏情况下，一个任意点播组的前缀 P可以是0前缀，那组成员可能没有拓扑位置。在这种情况下，任意点播地址在整个 Internet中，必须被告知作为一个分离的选路实体，这就为可以支持多少这样的全球任意点播组，带来严格的规模限制。因此，期望支持全球任意点播组似乎是不可能的或者说是非常受限制的。

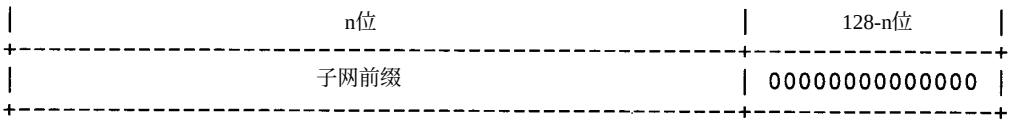
任意点播地址的用途之一是标识一组路由器，该组路由器是属于提供 Internet服务的一个组织的。这样的地址在 IPv6选路头中可用作直接地址，造成包的传递通过一个特定的集聚或集聚系列。其他可能的用途是标识连到一个特定子网的一组路由器，或者标识提供入口到一个特定选路域的一组路由器。

Internet任意点播地址在广泛传播及随意使用方面经验不多，然而已知使用它们所带来的复杂性和麻烦却很普遍 [ANYCST]。在获得更多的经验，并对一些问题有一致的解决方案之前，IPv6任意点播地址的下列限制始终存在。

- 任意点播地址不能用作IPv6包的源地址。
- 任意点播地址不能指定给 IPv6主机，只能指定给 IPv6路由器。

要求的任意点播地址

预定的子网路由器任意点播地址，其格式如下：



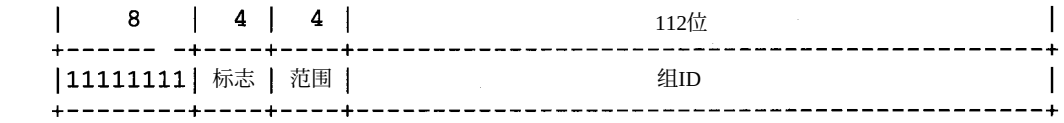
在任意点播地址中，子网前缀用来标识一条特定链路。对于接口标识符置 0的链路上的一个接口，其任意点播地址和单播地址语法上是相同的。

发送给子网路由器任意点播地址的包会传递到子网上的一个路由器。与子网有接口的所有路由器需要支持子网路由器任意点播地址。

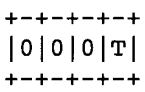
子网路由器任意点播地址企图用在某些应用场合，即一个节点需要和远程子网上一组路由器中的一个进行通信的场合。例如当移动主机要和一个位于本子网的移动代理通信的场合。

2.7 组播地址

IPv6组播地址是一组节点的标识符。一个节点可以归属于任意数量的组播组。组播地址具有下面的格式：



地址开始的1111 1111标识该地址为组播地址。标志由4位组成：



前面3位为保留位，初始设置为0。

T=0指示一个永久分配的(熟知的)组播地址，由全球Internet编号机构进行分配。

T=1指示一个非永久分配(临时)的组播地址。

4位的组播范围值用来限制组播组的范围。该字段的可能值如下表。

值	描　述	值	描　述
0	保留	8	组织本地范围
1	节点本地范围	9	(未分配)
2	链路本地范围	A	(未分配)
3	(未分配)	B	(未分配)
4	(未分配)	C	(未分配)
5	站点本地范围	D	(未分配)
6	(未分配)	E	全球范围
7	(未分配)	F	保留

组标识符字段标识给定范围内的组播组，可以是永久的，也可以是临时的。

永久分配的组播地址，意思是独立于范围值。例如，如果为 NTP服务器组指定一个组标识符为101(十六进制)的永久组播地址，于是：

FF01:0:0:0:0:0:0:101意指在同一节点上的所有NTP服务器。

FF02:0:0:0:0:0:0:101意指在同一链路上的所有NTP服务器。

FF05:0:0:0:0:0:0:101意指在同一站点上的所有NTP服务器。

FF0E:0:0:0:0:0:0:101意指Internet上的所有NTP服务器。

非永久分配的组播地址仅在给定范围内才有意义。例如，在某个站点由非永久的站点本地组播地址FF15:0:0:0:0:0:0:101标识的组与一个不同站点中使用同一个组标识符的组没有关系，与不同范围内使用同一个组标识符分配非永久地址的组也没有关系，与具有同一个组标识符的永久组也没有关系。

组播地址在IPv6包中不能用作源地址或出现在任何选路头中。

2.7.1 预定义的组播地址

下面为熟知的预定义的组播地址：

保留的组播地址：

```
FF00:0:0:0:0:0:0:0
FF01:0:0:0:0:0:0:0
FF02:0:0:0:0:0:0:0
FF03:0:0:0:0:0:0:0
FF04:0:0:0:0:0:0:0
FF05:0:0:0:0:0:0:0
FF06:0:0:0:0:0:0:0
FF07:0:0:0:0:0:0:0
FF08:0:0:0:0:0:0:0
FF09:0:0:0:0:0:0:0
```

上面列出的是保留的组播地址，且永远不能分配给任何组播组。

所有节点地址：

```
FF0A:0:0:0:0:0:0:0
FF0B:0:0:0:0:0:0:0
FF0C:0:0:0:0:0:0:0
FF0D:0:0:0:0:0:0:0
FF0E:0:0:0:0:0:0:0
FF0F:0:0:0:0:0:0:0
```

上面列出的组播地址标识了范围1(节点本地)或范围2(链路本地)内的所有IPv6节点的组。

所有路由器地址：

```
FF01:0:0:0:0:0:0:1
FF02:0:0:0:0:0:0:1
```

以上的组播地址标识了范围1(节点本地)、范围2(链路本地)或范围5(站点本地)内的所有IPv6路由器的组。

```
FF01:0:0:0:0:0:0:2
FF02:0:0:0:0:0:0:2
FF05:0:0:0:0:0:0:2
```

请求节点地址：FF02:0:0:0:0:1:FFXX:XXXX

上面的组播地址是从节点的单播和任意点播地址计算而得的。取单播或任意点播地址的低24位，并将其附加到前缀FF02:0:0:0:0:1:FF00::/104上形成一个请求节点组播地址，其范围在FF02:0:0:0:0:1:FF00:0000至FF02:0:0:0:0:1:FFFF:FFFF之间。

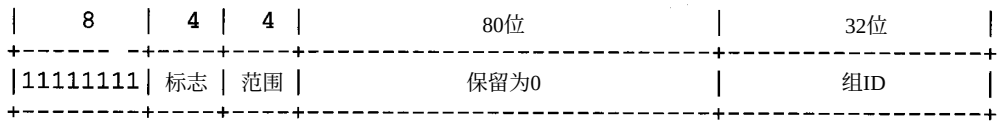
例如，对应IPv6地址4037::01:800:200E:8C6C的请求节点组播地址是FF02::1:FF0E:8C6C 。IPv6地址差别仅在高位，譬如，由于与不同的集聚相关联的多个高位前缀，将映射到同一个请求节点地址，因此减少了一个节点必须加入的组播地址数。

对每个指定的单播和任意点播地址，一个节点需要计算并加入相关的请求节点组播地址。

2.7.2 新IPv6组播地址的分配

目前将IPv6组播地址映射到IEEE 802 MAC地址的方法是用IPv6组播地址的低阶32位来创建

MAC地址。值得提出的是令牌网有不同的处理方法，定义见[TOKEN]。32位组标识符将生成唯一的MAC地址。由于新IPv6组播地址应当分配，所以组标识符总是在低阶32位上，如下图所示：



尽管将永久IPv6组播组数限制在2^{32}，但在将来不可能成为极限。如果将来必须要超过这个限度，组播仍然能工作，只是处理稍慢而已。

其他IPv6组播地址的定义和注册由IANA[MASGN]完成。

2.8 节点要求的地址

主机需要识别下面的地址以辨识它自身：

- 它的每个接口的链路本地地址。
- 分配的单播地址。
- 回返地址。
- 所有节点的组播地址。
- 每一个分配的单播和任意点播地址的请求节点组播地址。
- 主机所属的所有其他组的组播地址。

主机需要识别的所有地址，要求路由器都能识别，路由器还要能识别用来识别其本身的下列地址：

- 配置路由器工作的接口所用的子网路由器任意点播地址。
- 完成路由器配置要用的所有其他任意点播地址。
- 所有路由器组播地址。
- 路由器归属于所有其他组的组播地址。

在实现中应该预定义的地址前缀包括：

- 未指定地址。
- 回返地址。
- 组播前缀(FF)。
- 本地用前缀(链路本地和站点本地)。
- 预定义的组播地址。
- IPv4兼容的前缀。

实现时除非专门配置(如任意点播地址)，应假设所有其他地址均为单播地址。

3. 安全性考虑

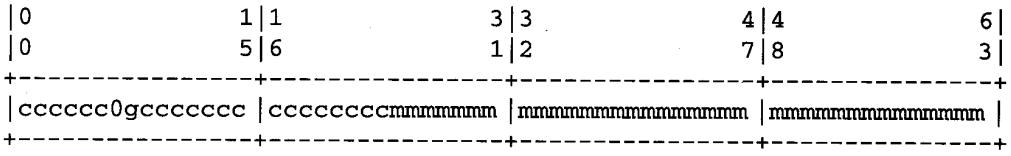
IPv6寻址文件对Internet基础设施的安全性没有任何直接影响。 IPv6包身份验证的定义见[AUTH]。

附录A 创建EUI-64接口标识符

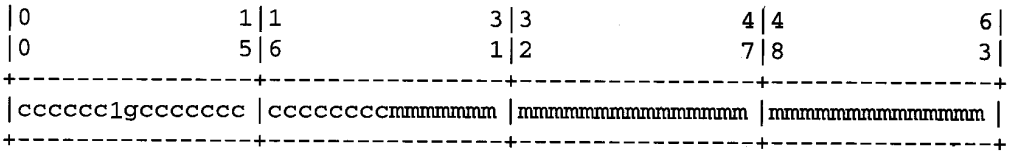
根据特定链路或节点的特性，有不少方法可以创建 EUI-64接口标识符。本附录介绍了其中的某些方法。

A.1具有EUI-64标识符的链路或节点

将一个EUI-64标识符转换成一个接口标识符，只需改变 u位的值。例如，一个全球唯一的EUI-64标识符具有下面的形式。



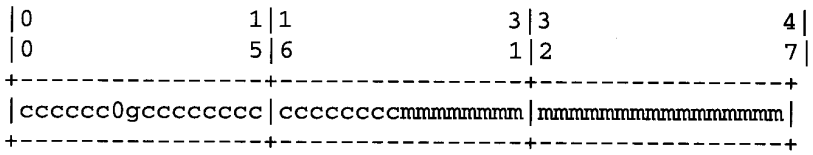
其中，c位是分配给公司的标识符； 0是全球/本地位的值，此处指本地范围； m是生产商选择的扩展标识符。 IPv6接口标识符的形式如下：



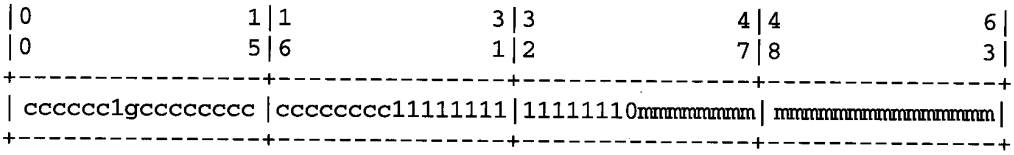
唯一改变的是转变全球/本地位的值。

A.2具有IEEE 802 48位MAC地址的链路或节点

[EUI64]规定了从一个IEEE 48位MAC标识符创建一个EUI-64标识符的方法。就是将以十六进制表示的两个字节 OxFF和OxFE插入到48位MAC地址中间(公司标识符与厂商配给的标识符之间)。下面的例子是一个具有全球范围的 48位MAC地址。



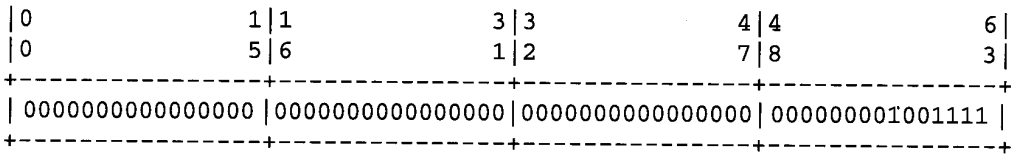
其中，c位是分配给公司的标识符； 0是指示全球范围的全球 /本地位值； g是个体/团体位；m是生产厂选择的扩展标识符。这样，接口标识符便具有下面的形式。



当接口或节点上IEEE 892 48位MAC地址可用时，由于它们具备的可用性和唯一性特性，就可以用它来实现创建接口标识符。

A.3具有非全球标识符的链路

有许多链路类型，当多个接入时，例如包括 LocalTalk和Arcnet,就无全球唯一的链路标识符。创建EUI-64格式化标识符的方法是取链路标识符 (如LocalTalk 8位节点标识符)，并在其前面填充0。下面就是一个具有十六进制值 Ox4F的LocalTalk 8位节点标识符生成的接口标识符的例子。



注意其中的全球/本位置为0，以指示本地范围。

A.4无标识符的链路

有一些链路无任何类型内置标识符。最普遍的就是一些串行链路和配置的隧道。为链路选择的接口标识符必须是唯一的。

当一条链路上无内置标识符可用时，最好是用从另一个接口的，或分配给节点本身的，一个全球接口标识符。使用这种方法就不会有连接同一链路的同一节点的其他的接口会用同样的标识符。

如果在链路上无全球接口标识符可使用时，就需要创建一个本地范围接口标识符。唯一的要求就是在该链路上是唯一的。有许多可能的方法用来选择一条链路唯一的接口标识符，包括如下方法：

- 人工配置。
- 生成随机数。
- 节点串行号(或其他节点特殊令牌)。

链路唯一接口标识符的生成方法应该使一个节点启动后或者接口从节点中删除或加入时都不应该有变化。

合适算法的选择，取决于链路和实现。形成接口标识符的细节规定在相应的 IPv6 over <link>技术规范中。强烈建议在任何自动算法中要实现冲突检测算法。

附录B 文本表示的ABNF描述

本附录定义了ABNF[ABNF]中的IPv6地址及前缀的文本表示，仅供参考用。

```
IPv6address = hexpart [ ":" IPv4address ]
IPv4address = 1*3DIGIT "." 1*3DIGIT "." 1*3DIGIT "." 1*3DIGIT

IPv6prefix = hexpart "/" 1*2DIGIT

hexpart = hexseq | hexseq "::" [ hexseq ] | "::" [ hexseq ]
hexseq = hex4 *( ":" hex4)
hex4= 1*4HEXDIG
```

附录C 对RFC 1884的修改

对RFC 1884(IPv6寻址体系结构)作了如下的修改：

- 增加了一个描述文本表示的ABNF的附录。
- 澄清了链路唯一标识符在自举或其他接口重新配置后不会改变。
- 阐述了评议后的地址模型。
- 改变了集聚格式术语，以便和集聚草案一致。
- 增加了在同一节点上，接口标识符可用于多个接口的文字说明。
- 增加了定义新组播地址的规则。

- 增加了创建基于EUI-64接口标识符的描述过程。
- 增加了定义IPv5前缀的标记方法。
- 用一个长的前缀改变请求节点组播的定义。
- 增加了站点范围所有路由器组播地址。
- 规定可集聚全球单播地址用001格式前缀。
- 将010(基于供应商的单播)和100(保留为地理上的)格式前缀改成未指定的格式前缀。
- 增加了对单播地址的接口标识符的定义部分；对单播地址增加了接口标识符定义的选择。要求在格式前缀范围内使用EUI-64以及在EUI-64中置全球/本地范围位的规则。
- 更新了NSAP文本部分以反映RFC 1888的工作。
- 删去协议特定的IPv6组播地址(如DHCP)，并参考了IANA中的定义。
- 删去了“单播地址例子”部分，变成OBE。
- 增加了新参考文献，并更新了参考文献。
- 对少量文字说明进行了澄清和改进。

参考文献

[ABNF] Crocker, D., and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, November 1997.

[AGGR] Hinden, R., O'Dell, M., and S. Deering, "An Aggregatable Global Unicast Address Format", RFC 2374, July 1998.

[AUTH] Atkinson, R., "IP Authentication Header", RFC 1826, August 1995.

[ANYCST] Partridge, C., Mendez, T., and W. Milliken, "Host Anycasting Service", RFC 1546, November 1993.

[CIDR] Fuller, V., Li, T., Yu, J., and K. Varadhan, "Classless Inter-Domain Routing (CIDR): An Address Assignment and Aggregation Strategy", RFC 1519, September 1993.

[ETHER] Crawford, M., "Transmission of IPv6 Pacekts over Ethernet Networks", Work in Progress.

[EUI64] IEEE, "Guidelines for 64-bit Global Identifier (EUI-64) Registration Authority", http://standards.ieee.org/db/oui/tutorials/EUI64.html, March 1997.

[FDDI] Crawford, M., "Transmission of IPv6 Packets over FDDI Networks", Work in Progress.

[IPV6] Deering, S., and R. Hinden, Editors, "Internet Protocol, Version 6 (IPv6) Specification", RFC 1883, December 1995.

[MASGN] Hinden, R., and S. Deering, "IPv6 Multicast Address Assignments", RFC 2375, July 1998.

[NSAP] Bound, J., Carpenter, B., Harrington, D., Houldsworth, J., and A. Lloyd, "OSI NSAPs and IPv6", RFC 1888, August 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[TOKEN] Thomas, S., "Transmission of IPv6 Packets over Token Ring Networks", Work in Progress.

[TRAN] Gilligan, R., and E. Nordmark, "Transition Mechanisms for IPv6 Hosts and Routers", RFC 1993, April 1996.

作者地址

Robert M. Hinden
Nokia
232 Java Drive
Sunnyvale, CA 94089
USA

Phone: +1 408 990-2004
Fax:+1 408 743-5677
EMail: hinden@iprg.nokia.com

Stephen E. Deering
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

Phone: +1 408 527-8213
Fax:+1 408 527-8254
EMail: deering@cisco.com

版权声明

RFC 2374 IPv6可集聚全球单播地址格式

网络工作组
RFC 2374
撤销：2073
类别：标准跟踪

R. Hinden
Nokia
M. O'Dell
UUNET
S. Deering
Cisco
1998年7月

提示

本文为Internet社区指定一个Internet标准跟踪协议，并请求为改进进行讨论和提出建议。对标准化状态和本协议的状况请参考“ Internet正式协议标准 ”(STD-1)最新版本。本文的分发不受限制。

版权声明

1. 引论

本文定义了可用于Internet上的IPv6可集聚全球单播地址格式。本文定义的地址格式与IPv6协议[IPv6]以及“ IPv6寻址体系结构 ”[ARCH]是一致的。它的设计是为了推进规模可伸缩的Internet选路。

本文件取代了RFC 2073(基于供应商的IPv6单播地址格式)。RFC 2073成为了历史文件。可集聚全球单播地址格式是对RFC 2073某些方面的改进。主要的改变包括删去了对路由集聚、EUI-64接口标识符的支持，对供应商和交换局集聚的支持，公共和站点拓扑的分割以及新集聚术语等所不需要的注册位。

2. IPv6地址概述

IPv6地址是为接口和接口组指定的128位标识符。有三类地址：单播、任意点播和组播。本文专门定义单播地址类。

在本文中，地址内的字段，赋予如“ 子网 ”这样的专门名字。当名字与其后的名词“ 标识符 ”一起使用时(如“ 子网标识符 ”)，被称为名字字段的内容。当名字与名词“ 前缀 ”一起用时(如“ 子网前缀 ”)，则表示包括本字段在内的所有左边的寻址位。

IPv6单播地址的设计使Internet选路系统在不需要了解IPv6地址内部结构的情况下，在任意位边界上，使用一个最长的前缀匹配“ 算法 ”，就可作出包的转发决定。IPv6地址的结构是为指派和分配用的。唯一的例外就是要在单播和组播地址之间加以区别。

IPv6地址的特定类型由地址的前几位指出。包含这前几位的可变长度字段叫做格式前缀(FP)。

本文为可集聚全球单播地址定义地址格式，其格式前缀为001(二进制)。其他格式前缀也

可以采用同样的地址格式，只要这些格式前缀是标识 IPv6单播地址的。只是本文只定义了这一种格式前缀而已。

3. IPv6可集聚全球单播地址格式

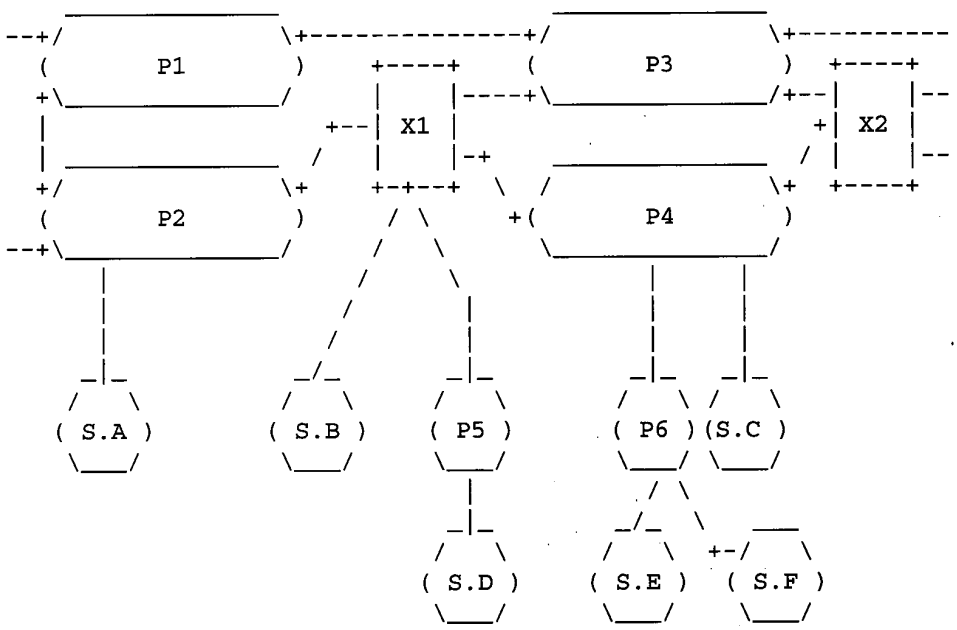
本文为IPv6可集聚全球地址格式的分配定义一种地址格式。作者相信这地址格式会广泛用于连到Internet的IPv6节点。设计该地址格式时，考虑到既支持当前基于供应商的集聚，也支持新的基于交换局的集聚。其组合既允许直接连接到供应商的站点能高效率地选路集聚，也允许连接到交换局的站点能高效率地选路集聚。站点可以选择连接到两者中的任一个集聚实体。

当该地址格式的目的是支持基于交换局的集聚 (除了当前基于提供商的集聚外)时，它的总体路由集聚特性与交换局无关。只有用基于供应商的集聚，才能提供效率高的路由集聚。

可集聚地址安排成一个三层次的分级结构：

- 公用拓扑。
- 站点拓扑。
- 接口标识符。

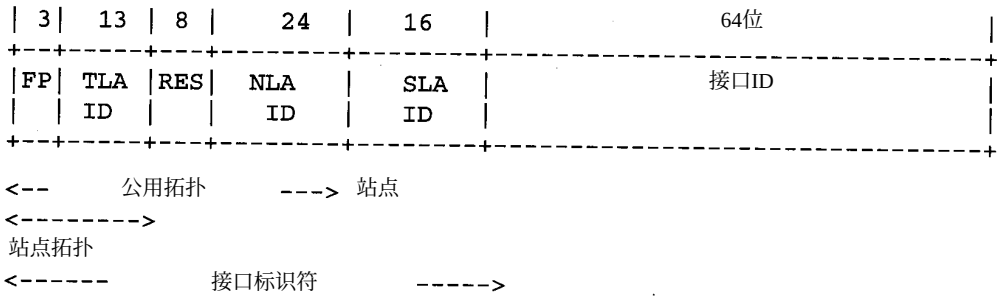
公用拓扑是提供公用 Internet传送服务的供应商和交换局群体。站点拓扑是本地的特定站点或组织，它不提供到本站点以外节点的公用传送服务。接口标识符是标识链路上的接口。



正如上面图所表示的可集聚地址格式，其目的是支持长途供应商 (如图中P1、P2、P3、P4)，交换局(如图中X1和X2)，多级供应商(如图中P5和P6)和用户(如图中S.x)。交换局(不像目前的NAPs、FIXes等)将分配IPv6地址。连接到这些交换局的组织，也要从一个或多个长途供应商那里预订(直接、间接地通过交换局等)长途服务。这样做可使寻址与长途转运供应商无关。这使得在改换长途供应商时，无需给它们的组织重新编号。组织也能成为多家的，也就是通过交换局连到一个以上的长途供应商，而不需要从每个长途供应商处获得地址前缀。用于此类供应商的选择及移植性的机制不在本文中讨论。

3.1 可集聚全球单播地址结构

可集聚全球单播地址格式表示如下：



其中，FP为格式前缀(001)；TLA ID为顶级集聚标识符；RES保留为将来用；NLA ID为下一级集聚标识符；SLA ID为站点级集聚标识符；INTERFACE ID为接口标识符；

下面分别给出IPv6可集聚全球单播地址格式的每一部分的说明。

3.2 顶级集聚标识符

顶级集聚标识符(TLA ID)是选路分级结构中的最高级。非默认路由器必须为每个激活的TLA ID保留一个路由表项，同时也许还有为TLA ID提供选路信息的附加项。附加项的目的是为它们的特定拓扑优先选路，但是所有级的选路拓扑，必须使提供给非默认路由表的附加项数量最少。

这样的寻址格式支持8192(2^{13})个TLA ID。要增加TLA ID的数量可以向右扩展TLA字段到保留字段，或者在另外的格式前缀上使用此格式。

关系分配TLA ID的议题，超出了本文范围，将在正在进行准备的文件中说明。

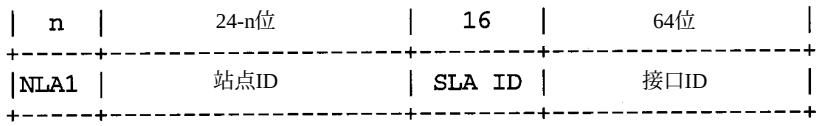
3.3 保留字段

保留字段留作将来用，当前必须置成0。

保留字段可留作TLA和NLA字段扩展时用。见第4节的讨论。

3.4 下一级集聚标识符

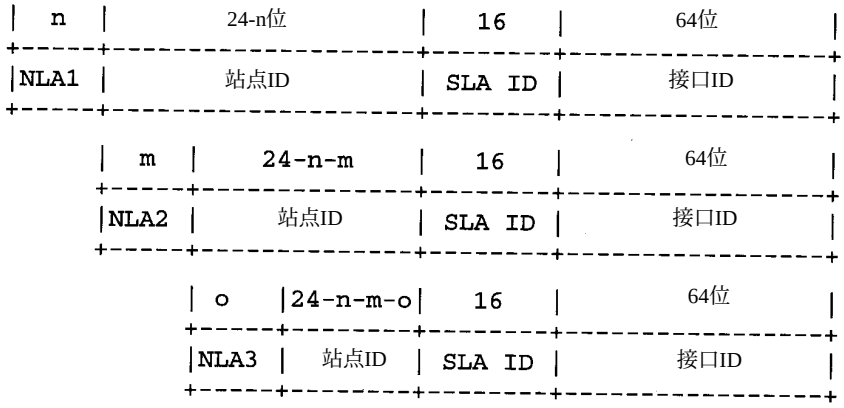
下一级集聚标识符被得到一个TLA ID的机构用来创建寻址分级结构和标识站点。该机构可以指定NLA ID字段的前n位，用来创建适合于它的网络的寻址分级结构。该字段的其余部分用来标识它愿为之服务的站点。表示如下：



每个得到一个TLA ID的机构可以有24位NLA ID空间。NLA ID空间使每个机构能够为相当于目前IPv4 Internet能够支持的总网络数几乎一样多的组织提供服务。

得到TLA ID的机构，也支持他们自己站点ID空间中的NLA ID。这就允许得到TLA ID的机构，能给提供公用传送服务的机构提供服务，也能给不提供公用传送服务的机构提供服务。

得到NLA ID的机构，也可以选择用他们的站点 ID空间去支持其他的NLA ID。这种情况表示如下：



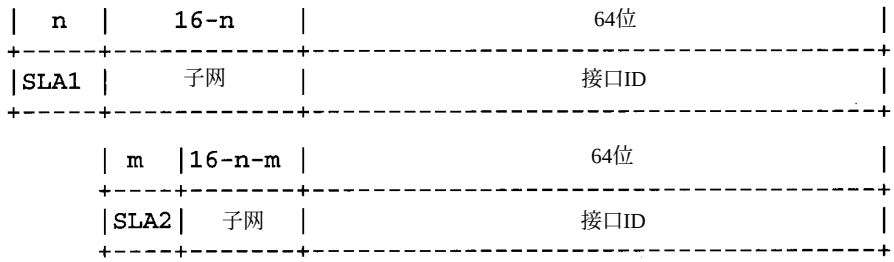
对一个特定的TLA ID，设计NLA ID位的安排，留给负责该 TLA ID的机构去做。同样，设计下一级NLA ID位的安排，由前面一级 NLA ID负责。在此建议分配NLA地址空间的机构用类似于[RFC2050]中的“慢启动”分配过程。

设计NLA ID分配计划，要在选路集聚效率和灵活性之间进行权衡。创建分级结构允许较大集聚数，从而使得路由表较小。平面 NLA ID的分配能使分配容易和连接灵活，但使得路由表较大。

3.5 站点级集聚标识符

SLA ID字段被单个机构用来创建他自己的本地寻址分级结构与标识子网。除了每个机构有一个数量很大的子网以外，类似于IPv4中的子网。16位的SLA ID字段支持65 535个单个子网。

机构可以选择他们的SLA ID为平面路由(如在SLA标识符之间不创建任何逻辑关系，这会使得路由表较大)，或者在SLA ID字段中，创建一个两级或多级分级结构 (使路由表较小)。后一种情况表示如下：



构成SLA ID字段所选择的方法，由个别机构负责。

在这种地址格式下支持的子网数，除了最大的机构之外，对其他所有机构应该是足够的。对于需要更多子网的组织，可以和它获得 Internet服务的机构商量，以获得附加的站点标识符，从而用来创建更多的子网。

3.6 接口ID

接口标识符用来标识一条链路上的接口。对链路来说，应该是唯一的。也可以在一个更

宽的范围内是唯一的。许多情况下，一个接口标识符与接口的链路层地址相同，或者根据接口的链路层地址而得的。用在可集聚全球单播地址格式中的接口标识符要求 64位长，并能构成IEEE EUI-64格式[EUI-64]。这些标识符，当全球令牌(如IEEE 48位MAC)可用时，具有全球范围意义；当全球令牌不可用时 (如串行链路、隧道终点等)，则只具有本地范围意义。 u位(在IEEE EUI-64术语中称为全球/本地位)在EUI-64标识符中必须根据[ARCH]的规定，正确地置位以指示是全球还是本地范围。

创建基于EUI-64接口标识符的过程定义见[ARCH]。形成接口标识符的细节，规定在相应的IPv6 over<link>技术规范中，诸如IPv6 over Ethernet[ETHER]，IPv6 over FDDI[FDDI]等。

4. 技术动机

在可集聚的地址格式中，字段长度的设计选择需要满足许多技术需求。这些将在下面段落中介绍。

顶级集聚标识符的长度是13位。可有8192个TLA ID。选择这样的长度，可使Internet上顶级路由器的非默认路由表，能在当前的选路技术且合理地留有余量的情况下，保持有限的范围。

因为非默认路由器为优化 TLA内部路径和TLA之间的路径，还要含有大量的长的前缀，所以保留余量是重要的。

重要的议题不仅是非默认选路表的长度问题，还有拓扑的复杂性决定了当计算一个转发表时，路由器必须考察非默认路由的拷贝数。当前 IPv4的实践是通常一个前缀要通过不同的路径通告15次。

Internet拓扑的复杂性将来还可能增加。重要的是 IPv6非默认选路应支持更大的复杂性以及巨大的Internet。

应该提出的是，在写作本文时(1998年春)，作者作了一个比较，IPv4非默认路由表包含大约50 000个前缀，表示可能支持大于8192个的路由。现在争论的问题是在当前的选路技术下，是否IPv4目前支持的前缀数已经足够多了。一些需要认真考虑的议题是路由稳定性以及供应商不支持所有顶级前缀的情况。技术上要求挑选 TLA ID的长度，在考虑合理余量的情况下，低于IPv4所具有的。

选择TLA ID字段为13位是出于工程的综合考虑。位数太少将不足以支持足够的顶级组织，位数太多将会超过合理协调的能力。为了处理前面所提到的议题，用当前的选路技术考虑一个合理的余量是合适的。

如果将来选路技术改进到在非默认路由表中能支持大量的顶级路由，那么如何加大 TLA标识符，就有两种选择：第一种是扩大TLA ID字段占用保留字数，这将使TLA ID数大约增加二百万个；第二种途径是为这样的地址格式分配另一个格式前缀 (FP)。或者将这两种途径组合，使TLA ID数量大大地增加。

保留字段的长度为8位，是为了使TLA ID字段和NLA ID字段有大的增长余地。

下一级集聚标识符的长度为 24位。如果用平面结构的话，可容纳大约 1600万个NLA ID。如果分级使用的话，合成起来大致等效于 IPv4的地址空间(假定平均网络规模为 254个接口)。如果NLA ID将来需要更多的空间，那么可以将 NLA ID扩展到保留字段来协调。

站点级集聚标识符字段的长度是 16位。每个站点可支持 65 535个子网。本字段长度的设计目标，对除了最大组织以外的所有组织是足够的。对于需要更多子网的组织，可以和它获

得Internet服务的机构商量，以得到附加的站点标识符，从而用来创建更多的子网。

站点集聚标识符字段是固定长度，这是为了强制标识特定站点的所有前缀，具有同样的长度(即48位)。这样会方便站点在拓扑中的移动(如变更ISP以及接到多个ISP的多家站点)。

接口标识符字段为64位。选择这个长度是为了满足[ARCH]中指定的需求，以支持基于EUI-64接口标识符。

致谢

作者对Thomas Narten, Bob Fink, Matt Crawford, Allison Mankin, Jim Bound, Christian Huitema, Scott Bradner, Brian Carpenter, John Stewart和Daniel Karrenberg的评论和建设性意见表示衷心的谢意。

参考文献

[ALLOC] IAB and IESG, "IPv6 Address Allocation Management", RFC 1881, December 1995.

[ARCH] Hinden, R., "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[AUTH] Atkinson, R., "IP Authentication Header", RFC 1826, August 1995.

[AUTO] Thompson, S., and T. Narten., "IPv6 Stateless Address Autoconfiguration", RFC 1971, August 1996.

[ETHER] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", Work in Progress.

[EUI64] IEEE, "Guidelines for 64-bit Global Identifier (EUI-64) Registration Authority", http://standards.ieee.org/db/oui/tutorials/EUI64.html, March 1997.

[FDDI] Crawford, M., "Transmission of IPv6 Packets over FDDI Networks", Work in Progress.

[IPV6] Deering, S., and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 1883, December 1995.

[RFC2050] Hubbard, K., Kosters, M., Conrad, D., Karrenberg, D., and J. Postel, "Internet Registry IP Allocation Guidelines", BCP 12, RFC 1466, November 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

安全性考虑

IPv6寻址文件对Internet基础设施安全性无任何直接影响。IPv6包的身份验证定义见[AUTH]。

作者地址

Robert M. Hinden
Nokia

232 Java Drive
Sunnyvale, CA 94089
USA

Phone: 1 408 990-2004
EMail: hinden@iprg.nokia.com

Mike O'Dell
UUNET Technologies, Inc.
3060 Williams Drive
Fairfax, VA 22030
USA

Phone: 1 703 206-5890
EMail: mo@uunet.uu.net

Stephen E. Deering
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

Phone: 1 408 527-8213
EMail: deering@cisco.com

版权声明