ZERO CONFIGURATION AND THE FUTURE HOME NETWORK

William Garrison
Motorola, Inc., Broadband Communications Sector

Abstract

This paper will review the state-of-the-art in Zero Configuration initiatives as it applies to an in-home network. The IETF Zeroconf, UPnP (Universal Plug-and-Play), Apple Rendezvous and IPv6 Stateless Address Autoconfiguration initiatives will be covered. In addition, from a MSO's perspective, does Zero Configuration represent a lost revenue opportunity or a lost headache opportunity? How will the Zero Configuration home network connect to the internet?

INTRODUCTION

Consumers do not want to own a home network. Instead, they want applications and services to simply work, as if by magic. If this magic requires a network, then it should be easily invoked with, at most, a simple incantation. This is how the consumer sees it. This is how we need to see it to get consumers to deploy applications that rely on home networks.

The purpose of a home network is to enable new services by combining the capabilities of both new and existing elements. Why can't I view a program that happens to be stored on my PVR downstairs on my TV upstairs? Why can't I listen to the music stored on my iPod on my home theater without hooking anything up? Well, with a Zero Configuration home network, you will be able to do all these things and more!

Zero Configuration is not a new idea. In the past, AppleTalk handled Zero Configuration for Macs and NETBIOS provided similar features for small networks of Windows PCs. However, these protocols were completely separate from any WAN (Wide Area Network) Protocol and served a limited range of devices. They did not allow a wide variety of appliances, PCs and Macs to interoperate.

An additional Zero Configuration networking benefit will be the creation of new kinds of networked products. These products will become commercially viable only when the inconvenience and support costs of traditional networking technologies are removed.

This paper will cover the Zero Configuration of IP (Internet Protocol) based networks. IP was selected because it is the native protocol for the ubiquitous Internet.

COVERAGE

A whole home network should be comprised of all of the data paths available between devices. These data paths include wired and wireless technologies. Wired technologies include power line, phone line, coax and dedicated CAT5 wiring. Wireless includes 802.11a/b/g, Zigbee and UWB (Ultra Wide Band). These technologies have significantly different data rate, delay and jitter characteristics. However, we need them to interoperate seamlessly and with zero user intervention.

REQUIREMENTS

As described in the internet draft entitled "Requirements for Automatic Configuration of IP Hosts" by Aidan Williams [1], Zero Configuration requires that we:

- Distribute IP addresses without a DHCP (Dynamic Host Configuration Protocol) server
- Provide name resolution without a DNS (Domain Name System) server
- Find and list services
- Distribute multicast addresses

In addition, while the system must operate in the absence of DHCP and DNS, it also must operate properly in their presence. Zero Configuration must not thwart their normal function.

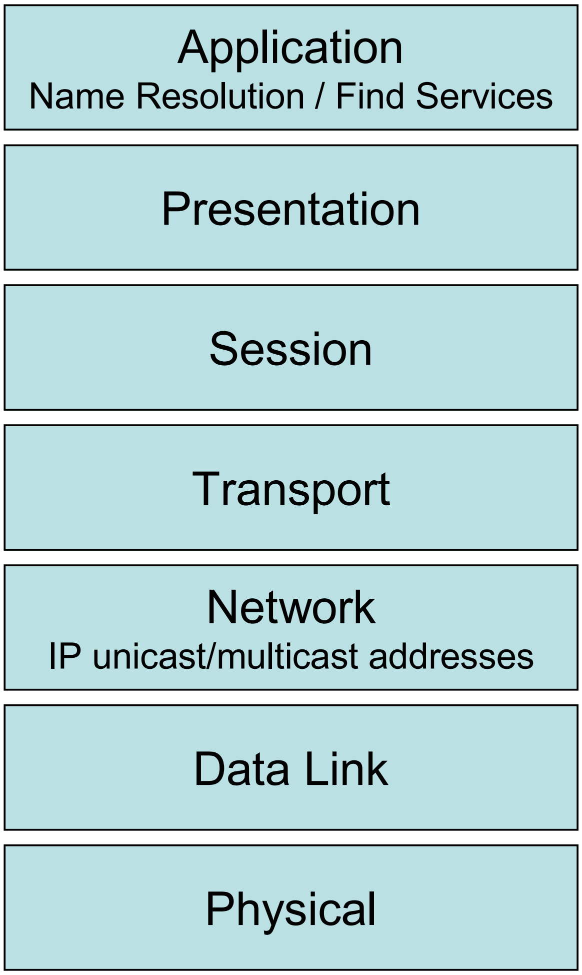


Figure 1 – The OSI Reference Model

The OSI Basic Reference Model divides a networking system into seven layers. This layering system enables an entity in one host to interact with a corresponding entity at the same layer in a remote host. Zero Configuration applies to the Application and Network layers, as outlined in Figure 1.

SECURITY

Security is also a Zero Configuration requirement. In the wired world, there is some physical security in a local network. In the wireless world, it is a challenge to tell the difference between your wireless device and your neighbor's. Zero Configuration may bring tighter security requirements. You cannot depend on the user noticing an intrusion because the user is further removed from what is actually going on. At the very least, the protocols used on a Zero Configuration network must be as secure as a non-Zero Configuration network.

Does Zero Configuration degrade security? Not really, a cracker can find out the same network information as offered by Zero Configuration using standard tools. There are plenty of tools floating around that let the unsophisticated "script kiddies" get into lightly protected networks. Zero Configuration's most likely security effect is to increase the number of networks that someone might try to intercept and increase the number of devices connected to those networks.

ZERO CONFIGURATION APPROACHES

Many people are currently trying to solve the Zero Configuration problem. Current initiatives include:

- Apple Rendezvous
- IETF Zeroconf
- UPnP™ (Universal Plug-and-Play)
- IPv6 Stateless Address Autoconfiguration

Each of these will be covered in its own section.

IETF ZEROCONF

The Zeroconf Working Group of the IETF (Internet Engineering Task Force) was chartered in September, 1999. Its goal is to “enable networking in the absence of configuration and administration.” Their goal is so inclusive that it goes as far as to include allowing “impromptu networks as between the devices of strangers on a train.” [2]

Zeroconf is a link-local technology. This means that the link-local addressing and naming are only meaningful to devices directly connected to the local network. Because these addresses and names are not unique globally, Zeroconf only applies to small wired or wireless networks. Zeroconf is appropriate for:

- Home and small office networks
- Ad hoc networks at meetings and conferences
- Two devices needed to share information

Inappropriate applications of Zeroconf would potentially result in serious networking problems.

Security

Zeroconf security is primarily based on the requirement that all included devices must be connected to a single link. Therefore, a Zeroconf connection can only be hacked by a device that is close by and easier to detect. However, if you are indeed going to wirelessly network with the stranger on the train, you will want some way to prevent networking with the stranger in the next train compartment.

A Zeroconf network is relatively vulnerable to some fairly standard attacks. Vulnerability to denial of service attacks is probably unavoidable. Even a simple ploy, such as a rogue device responding to every ARP (Address Resolution Protocol) so as to claim all available IP addresses, could shut down a Zeroconf network. However, is this any worse than being unable to speak to the person next to you on the train due to an unruly child? Given local nature of a Zeroconf network, the universe of people who could interfere with your network is small. Given the local nature of the network, in the case of a network interruption you could simply search for the scoundrel.

Industry Support

A standard needs industry support to have real-world relevance. Support for Zeroconf has been announced by:

- Apple
- Canon
- Epson
- HP
- Lexmark
- Philips
- Sybase
- Xerox
- World-Book

Working Group Status

The Zeroconf Working Group could not reach a consensus on security and service discovery issues. Therefore, it is not going to produce a specification on those issues. The Working Group is producing a protocol specification, describing automatic generation and assignment of link-local IPv4 addresses in environments lacking host configuration (static or using DHCP). This document is in draft form and will be submitted in the spring of 2004 for consideration as a Standards Track RFC.

Further information on the IETF Zeroconf Working group can be found at http://www.zeroconf.org/.

APPLE RENDEZVOUS

Rendezvous is Apple's name for IETF's Zeroconf. You might want to think of it like Rendezvous is to Zeroconf as Firewire is to 1394. Rendezvous is an open protocol, which Apple has submitted back to the IETF as a part of the ongoing standards creation process. Apple is using Rendezvous to transition from AppleTalk to an all-IP network.

Rendezvous matches the Apple customer's expectations of a friendly, easy to use system. Apple added Rendezvous services in the Jaguar release of the Mac OS X Operating System and is using those services in its own applications.

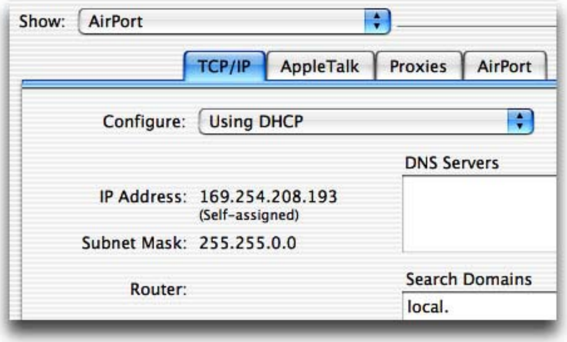
Rendezvous uses three technologies

- Automatic IP addresses (IPv4)
- Name to Address Translation (DNS queries via IP Multicast)
- Service Discovery

A Rendezvous device first tries to obtain an IP address by a standard DHCP request. If there is no DHCP response, it:

- picks an address at random in the 169.254/16 range
- Checks to see if this IP address is used via an ARP (tries another IP if it is)
- Periodically checks for DHCP server

The device periodically checks for a DHCP server because it wants to participate in the network with the widest possible reach.



Show: AirPort

Configure: Using DHCP

IP Address: 169.254.208.193
(Self-assigned)

Subnet Mask: 255.255.0.0

Router:

DNS Servers

Search Domains
local.

Figure 2 – A Rendezvous Assigned IP Address

In the above figure, you can tell the IP address came from Rendezvous both by the 164.256/16 address and the annotation "Self-assigned".

Apple then uses mDNS (multicast DNS) to handle DNS requests. In mDNS, each device runs their own mDNS responder. The mDNS responder provides traditional domain name services by having every device respond to the name queries that they know how to answer. Traditionally, devices contact a single known DNS server for name lookups. In mDNS, when a host needs to look up a name, it sends the query out to a local multicast group that includes all of devices that have locally registered Rendezvous services.

Rendezvous also uses a DNS-based Service Discovery called mDNS-SD. Essentially, services are resolved to devices similar to the way host names are resolved using mDNS.

Rendezvous Applications

Apple uses Rendezvous in its iChat instant messaging application. Besides working with AOL Instant Messenger, it also works with Rendezvous-enabled Macs. So, if the stranger on the train has a Mac, you can easily chat.

And if you change your status from "Available" to "Away", all Rendezvous clients are notified of the change automatically as a part of the Rendezvous mDNS Service Discovery.

Further information is available at http://developer.apple.com/macosx/rendezvous.

UPnP™

UPnP is being developed by the UPnP Forum. The Forum was formed in 1999 and now consists of over 650 member companies. The primary purpose of this Forum is to produce DCPs (Device Control Protocols) that describe standard methods for device interaction. UPnP is based entirely on open standards such as IP, TCP, UDP, HTTP and HTTPMU (a variant of HTTP that works on top of UDP multicast).

What is universal about UPnP? UPnP uses common protocols rather than vendor-specific device drivers. UPnP is independent of the physical media and can be implemented in any programming language and on any operating system. The basic foundation of UPnP is a client-server architecture, where the client is called a "Control Point" and the server is called a "Device."

Operation

UPnP covers the following device operations:

- Obtaining an IP address
- Discovering other devices
- Controlling other devices
- Receiving state change notifications (Eventing)
- Presenting User Interface for other devices

A UPnP device obtains an IP address the same way as described for a Rendezvous device in the previous section. Once it has an IP address, a device will use the IETF's SSDP (Simple Service Discovery Protocol) to find an interesting device with an interesting service on offer. This is accomplished using a multicast search message (HTTP over UDP over IP). Replies are unicast to the requestor. The multicast address, as well as the mechanism for advertising, searching, and revoking, are defined by the SSDP.

If a Control Point wants to know more about the services offered by a device, it requests an XML format description document. The XML document describes the device and all its embedded devices. The description includes services supported by the device, manufacturer information, version of the device, device web site, serial numbers and other relevant information.

The Control Point can now access the advertised services on the destination device via the SOAP (Simple Object Access Protocol). For a control point to invoke an action on device, it must get the device address, discover the device, retrieve descriptor, get URL for control and then send actions. UPnP is somewhat unique in that control is included as a part of its Zero Configuration standard.

UPnP also supports Eventing, where a Control Point will be notified of device state changes. In order for a Control Point to register for Eventing, it must get the IP address of the device, discover the device, retrieve the device description, get the URL for Eventing and then subscribe to the events from device. The subscription must be for all events on the device. There is no way to subscribe to just a single type of event.

While Rendezvous does not explicitly support Eventing, similar features are provided through its mDNS-SD service.

Security

UPnP does not directly specify any security measures in the basic protocol. The basic protocol relies on the security features in the standards-based protocols on which it is based. In addition, UPnP is seen as relatively secure because it sends only data and keeps the implementation private. Because no executables are exchanged, there are fewer security concerns.

UPnP has recently (Nov 2003) added a Device Security standard. UPnP security adds Access Control Lists and a Security Console which runs on Control Points that lets you edit Access Control Lists. Security is controlled down to the Service level. So, a Control Point might be able to set a clock's alarm but not its time. While this adds security, it requires significant manual intervention and therefore does not qualify as part of Zero Configuration.

Compatibility

UPnP and Apple Rendezvous use essentially the same link-local address specification. In both protocols, the IP address 192.164/16 is understood to be a link-local address. The Rendezvous version is based on a slightly newer version of the RFC than the one used by UPnP. Therefore, UPnP and Rendezvous devices can exist on the same network. The differences are Rendezvous can communicate with devices with routable addresses and Rendezvous uses a packet TTL (Time to Live) of 255 while UPnP (as implemented by Microsoft Windows) uses 128. The routable address feature is just an added benefit and the TTL can be handled by changing the default TTL value.

Windows XP and Windows ME provide various levels of UPnP support. More information on UPnP is available at http://www.upnp.org.

IPv6 Stateless Address Autoconfiguration

IPv6 is a redesign of the original Internet protocols. Most often you hear about how IPv6 supports a larger (128 bit addresses vs. 32 bit) address space. This removes the need for NATs (Network Address Translations) and private addresses, so you have end-to-end transparency. However, it also includes a number of Zero Configuration features that are well suited for the home environment. In addition, IPv6/IPv4 translation mechanisms let us take advantage of these features before the whole world transitions to IPv6.

IPv6 Autoconfiguration requires a multicast-capable link and begins when a multicast-capable interface is enabled. When a device interface is enabled, the host will generate a link-local address. The link-local address is sufficient for communication among nodes attached to the same link. The Link-local address is constructed by appending the well known local prefix fe80:0000:0000:0000:0000: to the device's 64 bit interface ID. For Ethernet, the interface ID is based on the 48 bit MAC address and generated according to IEEE EUI-64 (Extended Universal Identifier).

Before any address can be assigned to an interface and used, however, a node must attempt to verify that this "tentative" address is not already in use by another node on the link. To do this, it sends a Neighbor Solicitation message containing the tentative address as the target. If another node is already using that address, it will reply with a Neighbor Advertisement.

One unfortunate part of this specification is that if a duplicate address is found, the device must be configured manually. However, given that the interface ID should be unique, you should never have a duplicate address.

Once an interface has a link-local address, it can use this address to obtain site-local and/or Global-scope addresses, if desired. To get these addresses, the link first tries a DHCP request. If it does not get a response, the stateless mechanism allows the host to generate its own addresses using a combination of its interface ID and the subnet prefix advertised by a router. An address created this way will be a proper Site-local or Global-scope address, depending on the router configuration. IPv6 is designed for interfaces to have multiple IP addresses.

IPv6 also supports the easy renumbering of an entire site. This means that if a home is suddenly connected to the Internet, there is a simple way to change all the device addresses from local addresses to global addresses.

Security

So, a device can use a Link-local address that limits its traffic to inside the home or a link level address to limit communication to devices to which they directly attach. A clock radio might select a Site-local address to make sure only people within the radio's household can set the alarm. Or, it may select a global address so you can set it from the office.

All IPv6 nodes support the IP Security protocols (IPSec) standard for cryptographic authentication and encryption. So, all devices can send and receive packets with some confidence that the packets are from the expected source and their contents have not been modified.

.

LOST REVENUE OR LOST HEADACHE?

This is an easy answer. Zero Configuration is a lost headache opportunity. A service provider can bill for services, but it is hard to bill for support of the backbone network. You could charge a per-hour or a per-incident fee, but in the end the call center will not be a profit center. Zero Configuration is a revenue generator because the "lost headache" of configuration will allow the roll-out of new and profitable services.

CONNECTING TO THE INTERNET

Zero Configuration is designed for link-local connections. This means that you connect to devices on the same wire (if a wired network) or same channel (if a wireless network). In order to connect a Zeroconf/Rendezvous device to the Internet, you will need a DHCP server or a bridge device to go from the link-local domain to the Internet. UPnP explicitly allows a single device interface to have multiple IP addresses. So if a DHCP server became available, a UPnP device could seamlessly connect to the Internet. IPv6 allows a device to get a Global address from DHCP or create one based on a router advertisement.

CONCLUSION

The biggest remaining question is "When will Zero Configuration become commonplace?" Zero Configuration is more difficult to roll out than many other services because it requires wide adoption before it has significant value. Metcalfe's Law says that the value of a network grows in proportion to the square of the number of users. By analogy, the value of Zero Configuration will grow rapidly as the number of devices which offer it grow. Connecting one device that supports Zero Configuration and one device that does not will not leave you with two devices that are half configured! It unfortunately leaves you with two devices to configure, one of which may not have been intended to be user configured.

When will we get to the tipping point? When will people refuse to buy a device unless it has Zero Configuration? Well, first one of the Zero Configuration approaches must become the clear winner. Or some combination of approaches will become the clear winner. Why not use Zeroconf for the link layer and then above that use an evolution of UPnP which has evolved to support IPv6? This certainly sounds like a practical and powerful combination. We will have to wait for the market to decide this one.

Because of its support in Microsoft Windows, and the widespread deployment of Windows, UPnP looks like it could be the winner. However, IPv6 is certainly a very vendor neutral option and it too is supported in the latest versions of Windows. The race is far from over. When the market decides the winner (perhaps in a matter of 2-3 years), the winner will rapidly become ubiquitous.

REFERENCES

[1] Aidan Williams, “Requirements for Automatic Configuration of IP Hosts”, draft-ietf-zeroconf-reqts-12.txt.

[2] Erik Guttman, “Zeroconf Host Profile Applicability Statement”, draft-ietf-zeroconf-host-prof-01.txt.