



Can We Make IPv4 Great Again?

Presentation to Internet Society - NY Chapter

2017 March 07
AYChen@Avinta.com

Avinta Communications, Inc.
142 N. Milpitas Blvd., #148, Milpitas, CA 95035-4401 U.S.A.
Tel: +1 (408) 942-1485 Web: www.Avinta.com

- ▶ **The IPv4 based Internet went through phenomenal growth during the last couple decades.**
- ▶
- ▶ **However, the future is dampened primarily due to the exhaustion of the assignable public address pool.**
- ▶
- ▶ **The challenge is to resolve this issue with minimum perturbation to the deployed equipment and the established operation.**
- ▶
- ▶ **After extensive study, the only solution seemed to be moving on to IPv6. However, IPv6 has not been delivering what was expected.**
- ▶
- ▶ **For example, the Dual-Stack approach has to be adopted apparently due to IPv6 not being a superset of IPv4, nor capable of encapsulating the latter. This handicap of not being able to achieve a smooth transition is a major surprise for a protocol that was set out to replace the one being heavily depended upon by everyone in daily use.**
- ▶
- ▶ **Consequently, the disadvantaged regions are concerned about handling the technical challenges and dealing with the financial implications associated with IPv6.**
- ▶
- ▶ **A recent review of the subject suggests a fresh possibility for an alternative.**



Outline

- A. Why this Proposal**
- B. What is the Solution**
- C. Optimum Approach**
- D. Technical Considerations**
- E. Nitty-Gritty**
- F. Goals / Takeaways**

- ▶ **Since this is a rather involved subject, to expedite this presentation, we will start with addressing two logistic items. First, it would be prudent to briefly describe my experience to shed some light about where this seemly unorthodox solution may come from. My bio is in the ISOC Member Profile area which was copied in from the LinkedIn. Secondly, there are two handouts utilized by this presentation. One is a Terminology, Abbreviation and Acronym list, and the other is URLs to technical references. Please utilize these to go along with the presentation. Hopefully, these supporting material may harmonize the diversity of the participants to allow us focusing on the main topics.**
- ▶
- ▶ *********
- ▶
- ▶ **This presentation will introduce system level concepts of the proposed solution that manifests to a range of possibilities. We will describe the idea by using the simplest form that inspired our effort, and then present the proposed approach.**
- ▶
- ▶ **Related technical considerations will be briefly reviewed to demonstrate the realizability. Supporting public references will be identified along the way to minimize the need of digging into the technical details.**
- ▶
- ▶ **Although the implementation of the proposal is rather rudimentary for networking professionals, examples will follow to provide a glimpse of the extend of efforts that may be involved.**
- ▶
- ▶ **At the conclusion, we would like to hear thoughts, comments as to whether this approach is technically realizable and economically responsible. If so, would it be in the interest of the general mass to consider deploying it.**



A. Why this proposal

- **IPv4 address pool almost exhausted**

- **IPv6 deployment sluggish**
<https://ams-ix.net/technical/statistics/sflow-stats/ether-type>
<https://stats.labs.apnic.net/ipv6>

- **Disadvantaged regions concerned with**
 - **being left behind due to technical complexity**
 - **financial burden of going to IPv6**

- **IPv4 and IPv6 coexist on DualStack for quite sometime to come**

- ▶ **IPv4 address pool exhaustion is a known fact.**
- ▶
- ▶ **Despite great promises and a lot of work, IPv6 deployment seems to be still rather sluggish, perhaps because it is lacking the technical base for a smooth transition from IPv4. Based on up-to-date statistics, both IPv6 traffic and equipment readiness percentages are surprising low.**
- ▶
- ▶ **These seem to correlate with concerns expressed by developing regions. They are worried about not only to keep up with the technology, but also to bear the cost of the transition.**
- ▶
- ▶ **Currently, the Internet community has settled with the Dual-Stack approach. Under such an environment, even though IPv6 has been chosen as the primary solution, one should not negate the other. We should not brush off valid technology that may improve IPv4 during the interim.**
- ▶
- ▶ **IPv6 could continue its way where it makes sense. In parallel, any viable IPv4 enhancement should be considered as a temporary relief for those who are unable to adopt IPv6 right away. This philosophy would go with the common goal for an open Internet and the mission of the Internet Society.**



Boundary Conditions

- **Demand - By Year 2020:**
 - **Worldwide Population to reach: 7.6B (Billion)**
 - **Number of IoT Devices: 50B**

<https://nishithsblog.files.wordpress.com/2014/04/internet-of-things-market-forecast.jpg>

 - **Supply - IPv4 Address Pool Size: 4B (256 x 256 x 256 x256 from 4 Octets, 8 Bits each)**

 - **Address Demand estimated to be 13 times over the supply**

 - **Actual usable IPv4 pool much smaller than 4B due to historical allocation practices**
- <http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml>

- ▶ **To start this analysis, we need to know the quantitative conditions that we are facing with.**
- ▶
- ▶ **Basically, the IPv4 address capacity is at best only 13th of the projected number of worldwide IoTs by Year 2020.**
- ▶
- ▶ **In reality, the ratio is even worse, because many of the IPv4 addresses have not been properly utilized.**
- ▶
- ▶ **As may be seen in the IPv4 Address Space Registry, many addresses are not been effectively utilized. In particular, those begin from 224/8 are "RESERVED". Furthermore, the block starting from 240/8 is even RESERVED for "Future use". The implication of this caught our attention.**
- ▶
- ▶ *********
- ▶
- ▶ **In the past, various techniques have been attempted to resolve the address shortage issue without success.**
- ▶
- ▶ **The scheme described here approaches the task from a slightly different angle, yet utilizes ingredient techniques that are either in use or have been studied and tested previously. So, no development effort is anticipated in implementing this proposal.**

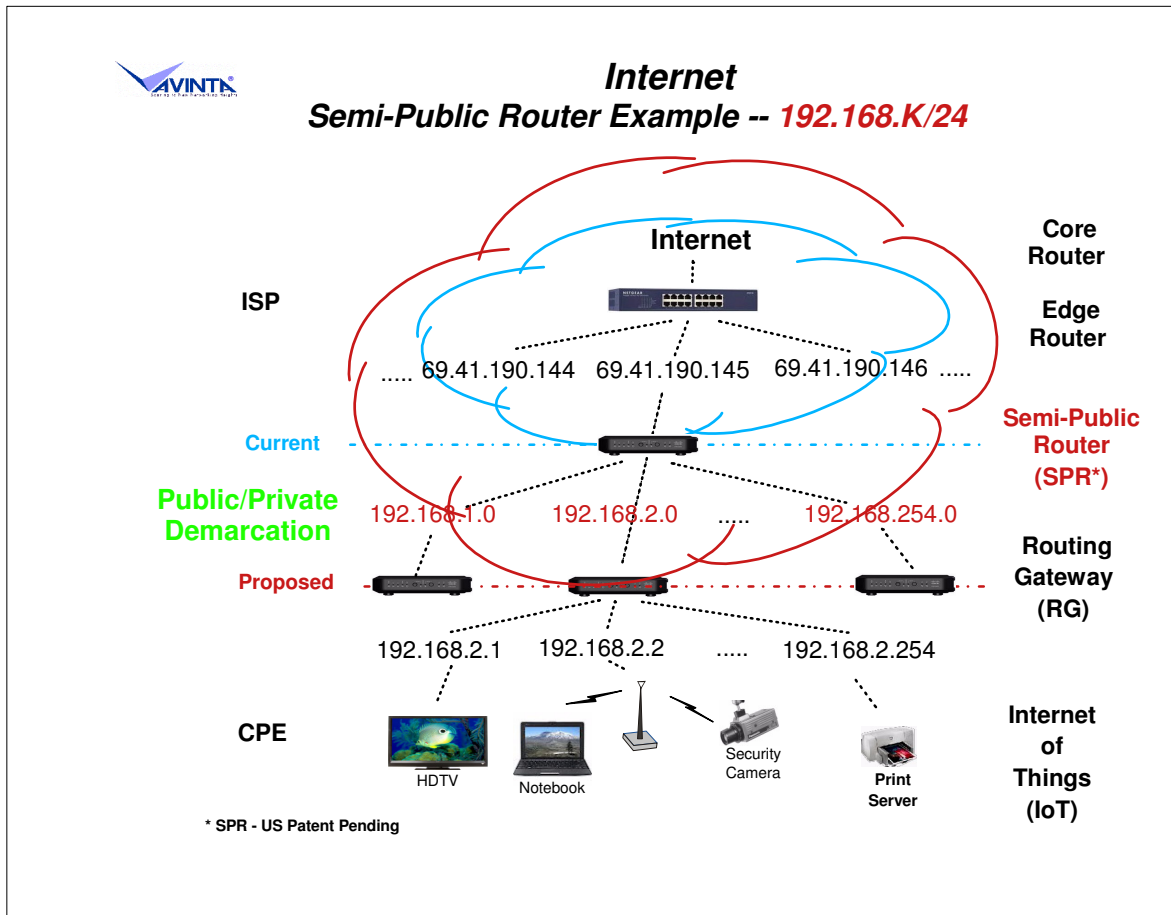


B. What is the solution

- **Expand assignable IPv4 public address pool:**
 - **Introduce Semi-Public Router (SPR) - A simple IPv4 compatible router**
 - **Insert a SPR between an Internet Edge Router (ER) and the private premises Routing Gateway (RG) it serves**
 - **Reclaim a block of currently reserved addresses within the IPv4 pool for SPR to utilize**

(Refer to example - next slide)

- ▶ **The basic idea is to deploy a new class of simple routers (SPR), each is inserted inline on an Internet public access channel between a public ER and the private RG that it serves.**
- ▶
- ▶ **Since SPRs provide public routing functions, the effective assignable IPv4 public address pool is expanded.**
- ▶
- ▶ **To maintain the address format compatibility, a subset of IPv4 addresses need be reserved from the main pool to be used by the SPR operation.**
- ▶
- ▶ *********
- ▶
- ▶ **The last step may sound contradictory, since the IPv4 pool is practically depleted. However, arithmetically removing a finite number of elements from a relatively large set has only minuscule effect to the overall set, because of the simple subtraction. On the other hand, if such a subset is geometrically applied to the main set, the effective overall number of elements is multiplied accordingly. This particular characteristics is the critical idea behind this proposal.**
- ▶
- ▶ **For example, reserving 10 addresses from the 4B IPv4 pool is a trivial reduction to the pool. However, if these 10 addresses are reused on each and every address in the remaining pool, the total assignable IPv4 addresses become 40B (minus 100 to be more accurate)!**
- ▶
- ▶ **This mechanism actually has been in operation under IPv4 protocol for a long time. Perhaps no one has examined it from the specific angle for relieving the address pool depletion issue.**



- ▶ Let's look at the everyday Residential Gateways (RGs) that route the 192.168/16 address block for private networks.
- ▶
- ▶ Most of such RGs are actually operating with 192.168.K/24, where K is an arbitrary parameter between 0 and 255 preset in the factory. The most common values for the K have been, 0, 1, 2, 10, etc. Under this condition, each private network is restricted to 256 addresses, available from the fourth octet. This practice is probably adopted by manufacturers to reduce the processing burden on the RGs.
- ▶
- ▶ A group of 256 RGs like the above, each with a different K value, may be operated from the same public IPv4 address without conflict.
- ▶
- ▶ The router that serves these 256 RGs is one of the basic SPRs.
- ▶
- ▶ Since the RGs have not changed, the conventional demarcation line remains where it has been from subscriber point of view.
- ▶
- ▶ Applying this configuration to every IPv4 address, we can get 256 x 4B or about 1000B assignable addresses that are IPv4 compatible. This is 20 times of the 50B IoTs by Year 2020. The Internet has thus expanded by 20 times.
- ▶
- ▶ *****
- ▶
- ▶ This finding encouraged us to look further and harder for similar configurations. For the convenience of referring to these possibilities, we created a name for this approach as the EzIP (phonetic for Easy IPv4).

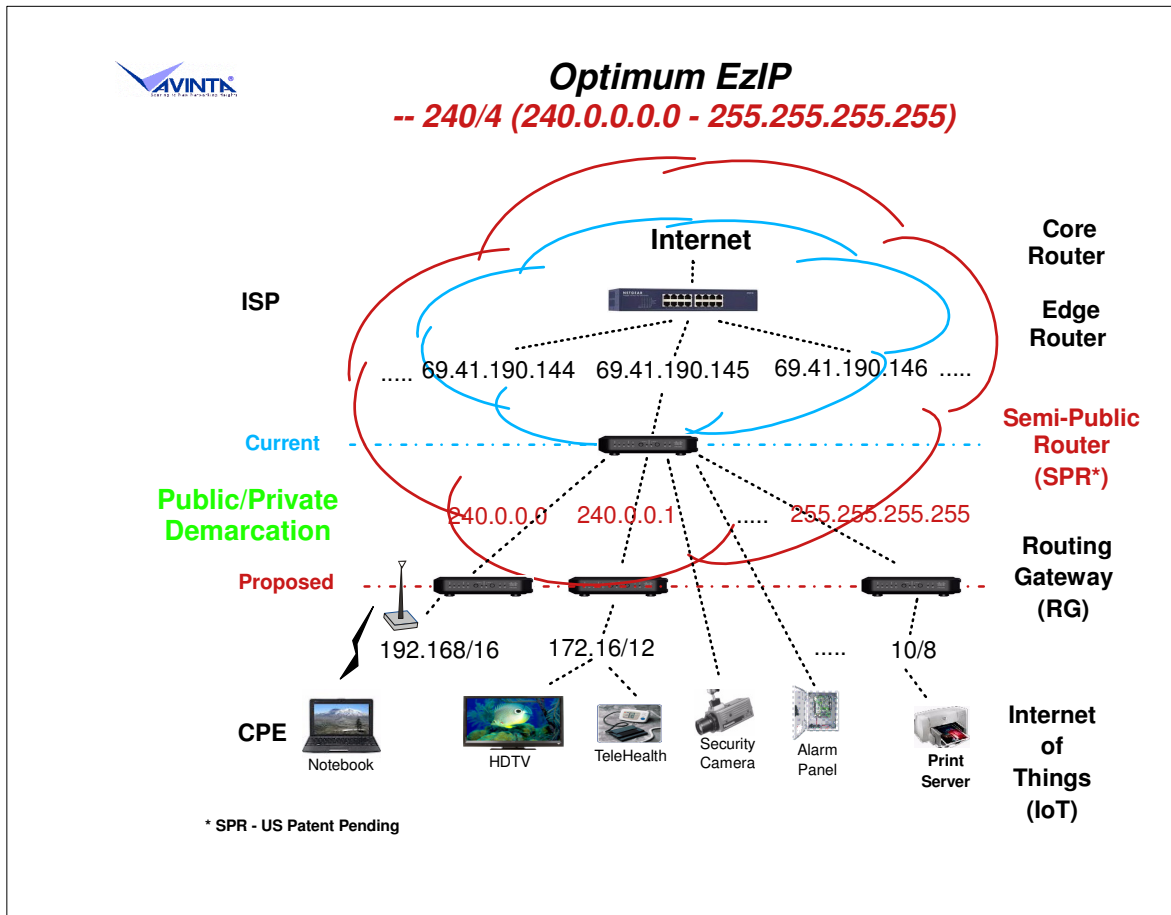


C. Optimum Approach

- **The 240/4 block:**
 - **From 240.0.0.0 through 255.255.255.255 whose first four bits are all "1", totaling 256M addresses**
 - **Reserved for "Future use": Not routable - neither publicly nor privately**
 - **Offers the potential of multiplying each current IPv4 public address by 256M times, yet**
 - **Does not impact existing private networks, nor directly connected IoTs**

(Refer to Optimum EzIP example - next slide)

- ▶ **The other two private network address blocks, 172.16/12 and 10/8, each can similarly support the SPR to multiply the IPv4 pool by 256 times for a total of 768 times. Together, the IPv4 public pool could become 60 times of the projected 50B IoTs by Year 2020.**
- ▶
- ▶ **However, there are drawbacks with this basic scheme:**
 - ▶ **As one octet is allocated for extending the public address, the corresponding private network is reduced accordingly to one 256th. This will not go well for private networks that are already heavily populated.**
 - ▶
 - ▶ **At the other end of the spectrum, we could utilize all available bits in these three private network blocks to extend the Internet to be fully end-to-end. This approach can multiply the assignable address pool by a factor of over 17M. However, no assignable address is left for any private network.**
 - ▶
 - ▶ **Furthermore, these address blocks have been used privately without coordination for many years. It will be difficult to reclaim any portion for a new and uniform application.**
- ▶
- ▶ *********
- ▶
- ▶ **One possibility is to utilize the 240/4 block as the SPR address. This block has been "RESERVED" for "Future Use" since 1981, yet very much under utilized at this juncture.**
- ▶
- ▶ **Since 240/4 is acceptable for neither public nor private routing, it is ideal to be redesignated as a new class of "Semi-Public" address for SPR. 240/4 will interfere neither of the current two types of networks, because both are designed to ignore or reject this address block.**



- ▶ Let's start with a basic Internet diagram similar as the previous one, except showing a 172.16/12 private network.
- ▶
- ▶ Making use of the 240/4 block, each SPR may expand an IPv4 public address to a publicly assignable pool of 256M addresses.
- ▶
- ▶ The Demarcation line will similarly stay where it is currently.
- ▶
- ▶ Applying to all 4B IPv4 addresses, the assignable address pool becomes 1000MB, or 1BB which is 20M times of the projected IoTs by Year 2020.
- ▶
- ▶ Since 256M is 0.25B that is reserved from the original IPv4 pool, the more accurate calculation of the above would be about 0.94BB (3.75B x 0.25B).
- ▶
- ▶ Note that with 256M assignable addresses, each SPR will be able to serve a measurable size community. Even with a good portion of the communications within a community being handled as the Intra-SPR traffic, the data stream flowing through each of the original IPv4 public address access channel will increase significantly, requiring a dedicated optical facility.
- ▶
- ▶ To make the EzIP scheme work, the additional address information between SPRs needs be transported through the Internet via a well-established payload mechanism in the IP header. Let's review the available technology.



D. Technical Considerations

- **RFC 791: Define Option mechanism in the IP Header (Figure 9).**
<https://tools.ietf.org/html/rfc791>
- **RFC1385: Utilize Option Number = 17 to carry "variable" EIP Extension address in the IP header (Figure 1)**
<https://tools.ietf.org/html/rfc1385>
- **APNIC: Request to redesignate 240/4 for "Private Use" (2. Caveats of Use)**
<https://tools.ietf.org/html/draft-wilson-class-e-02>
- **EnIP: Utilize the full private network addresses for end-to-end connectivity (Temporary Option Number = 26) (Figure 1)**
<https://tools.ietf.org/html/draft-chimiak-enhanced-ipv4-03>
- **Over a dozen of Option numbers available**
<http://www.iana.org/assignments/ip-parameters/ip-parameters.xhtml>
<https://tools.ietf.org/html/rfc6814>
- **EzIP: Transport the Semi-Public address by Option words in EzIP Header (Figures 12, 13)**
<https://tools.ietf.org/html/draft-chen-ati-adaptive-ipv4-address-space-00>

- ▶ **Before going through the technology considerations that make the EzIP concept realizable, we should first highlight two basic terminology.**
 - ▶ **RFC (Request For Comment) is an IETF technical specification document that solicits comments and eventually may, after all feedback has been properly addressed, become a recommendation, convention or standard for the Internet operations.**
 - ▶ **The Option mechanism is an IP header capability that carries numerically encoded information as payload.**
- ▶ *****
- ▶ **RFC791 (1981 - Page 38, Figure 9) defined the Option mechanism together with EOOL (End of Operation List) and NOP (No-Operation) Option codes for carrying flexible length information as payload**
- ▶ **RFC1385 (1992 - Page 4, Figure 1) proposed new routers linking existing CRs with variable length IP address. Although this opened the door for very long IP address (could be longer than IPv6), it led to issues, such as address format compatibility and router configuration transitions.**
- ▶ **A Draft by APNIC (2008 - Page 2, Section 2) proposed to use 240/4 as the fourth private network address block, but identified that many deployed private network routers would reject the 240/4 addresses.**
- ▶ **EnIP (Enhanced IPv4, 2016 - Page 6, Figure 1) is an ongoing parallel effort for expanding the address pool. It trades private network for end-to-end connectivity. It may be viewed as a special case of the EzIP scheme.**
- ▶ **An Option number is needed to identify the type of information conveyed by the Option word(s) that follows. RFC6814 (2012 - Page 4) deprecated nine Option numbers that were assigned to earlier experiments.**
- ▶ **The EzIP (2016 - Pages 15, 16, Figures 12, 13) transports the Semi-Public address across the Internet according to RFC791. The optimum address pool is the 240/4 block.**



E. Nitty-Gritty

1. IP Header variances

- IP Header
- TCP/IP Header
- EzIP Header and
- TCP/EzIP Header

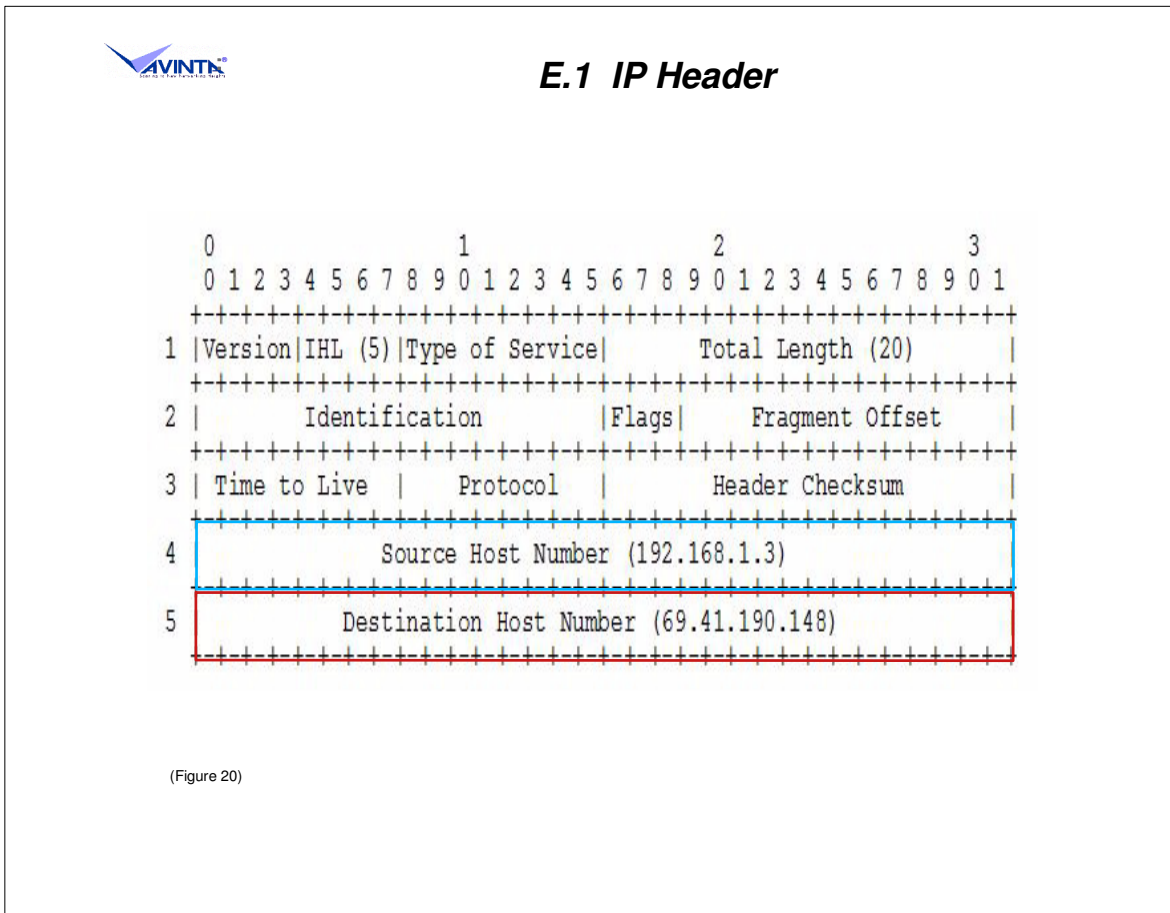
2. TCP/EzIP Header transitions

- Network Architecture Example
- T1z
- RG1
- SPR1
- ER1
- CR
- ER4
- SPR4

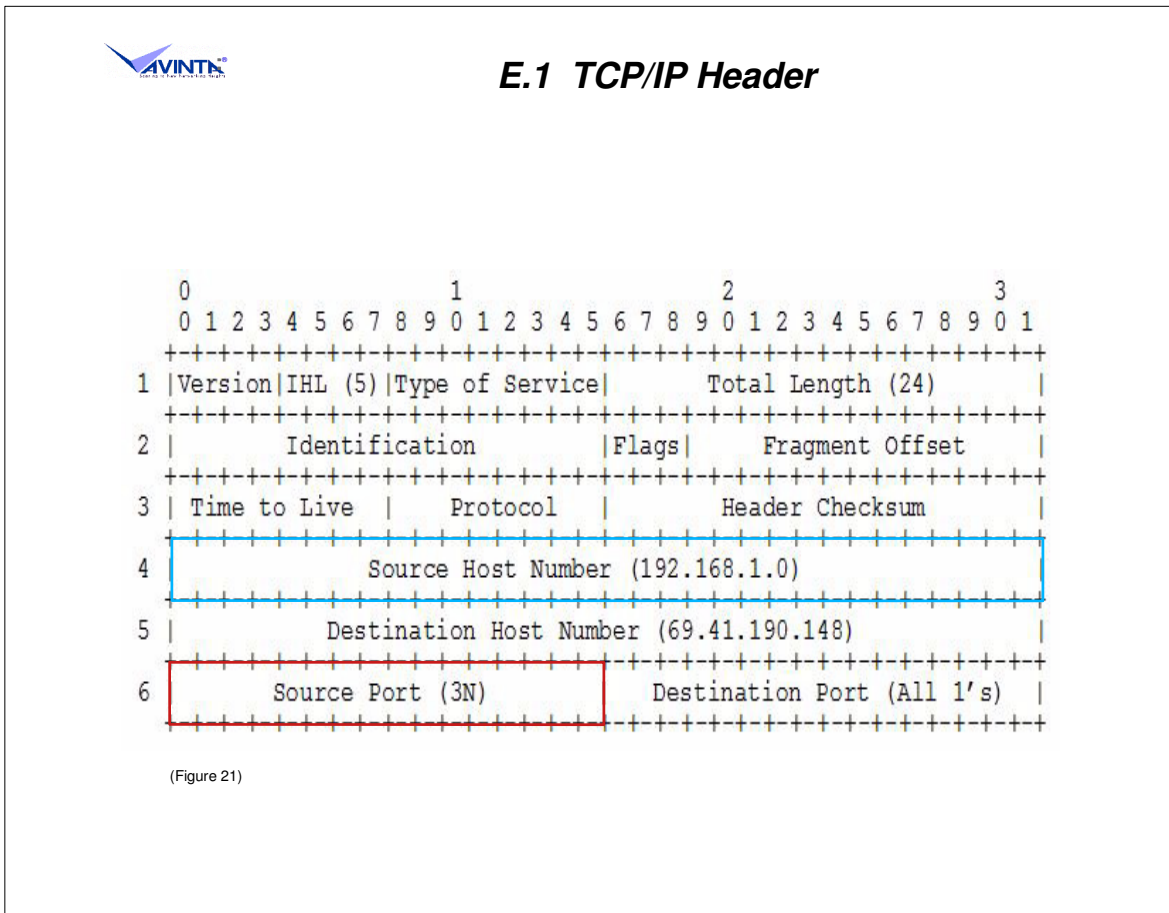
3. Generic full (4 octet) EzIP Header

(Refer to following slides for examples)

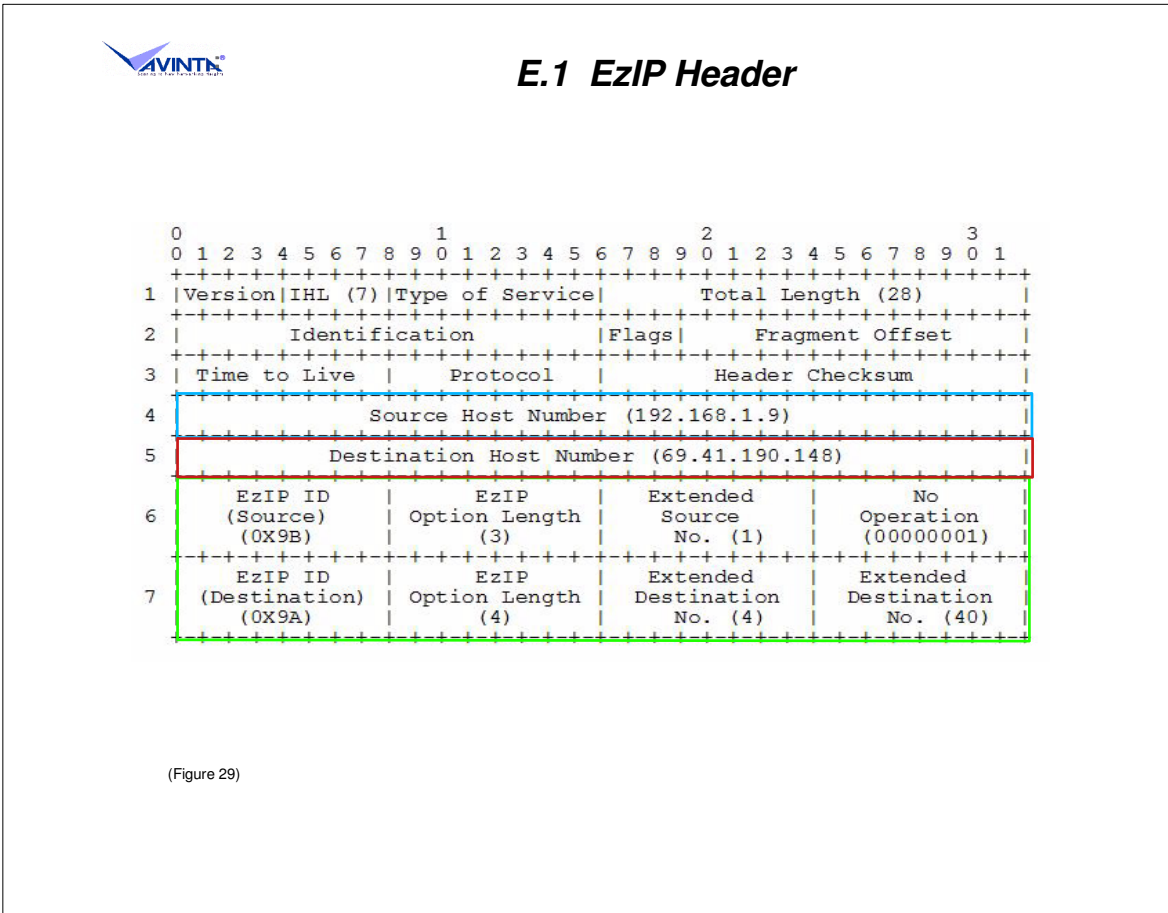
- ▶ **Briefly, the EzIP utilizes the Option mechanism defined by RFC791 to transport the extended address information (240/4 block) across the Internet as the payload in an IP header. A pair of SPRs at either end of a link (beyond ERs) encodes and decodes this information, respectively to effect the additional routing.**
- ▶
- ▶ **The TCP port number is not affected.**
- ▶
- ▶ **These are probably more than enough description for TCP/IP experts in the audience.**
- ▶
- ▶ *********
- ▶
- ▶ **For completeness, however, this presentation includes how these headers may look like under a couple situations.**
- ▶
- ▶ **The first set of figures are quick comparisons between the conventional IP and TCP/IP headers versus those with EzIP information.**
- ▶
- ▶ **The second set describes how the content of an EzIP header may change as an IP packet traverses through the Internet represented by a sample network architecture diagram.**
- ▶
- ▶ **Lastly, we will present a generic EzIP header for handling the 240/4 case as well as other EzIP manifestations.**



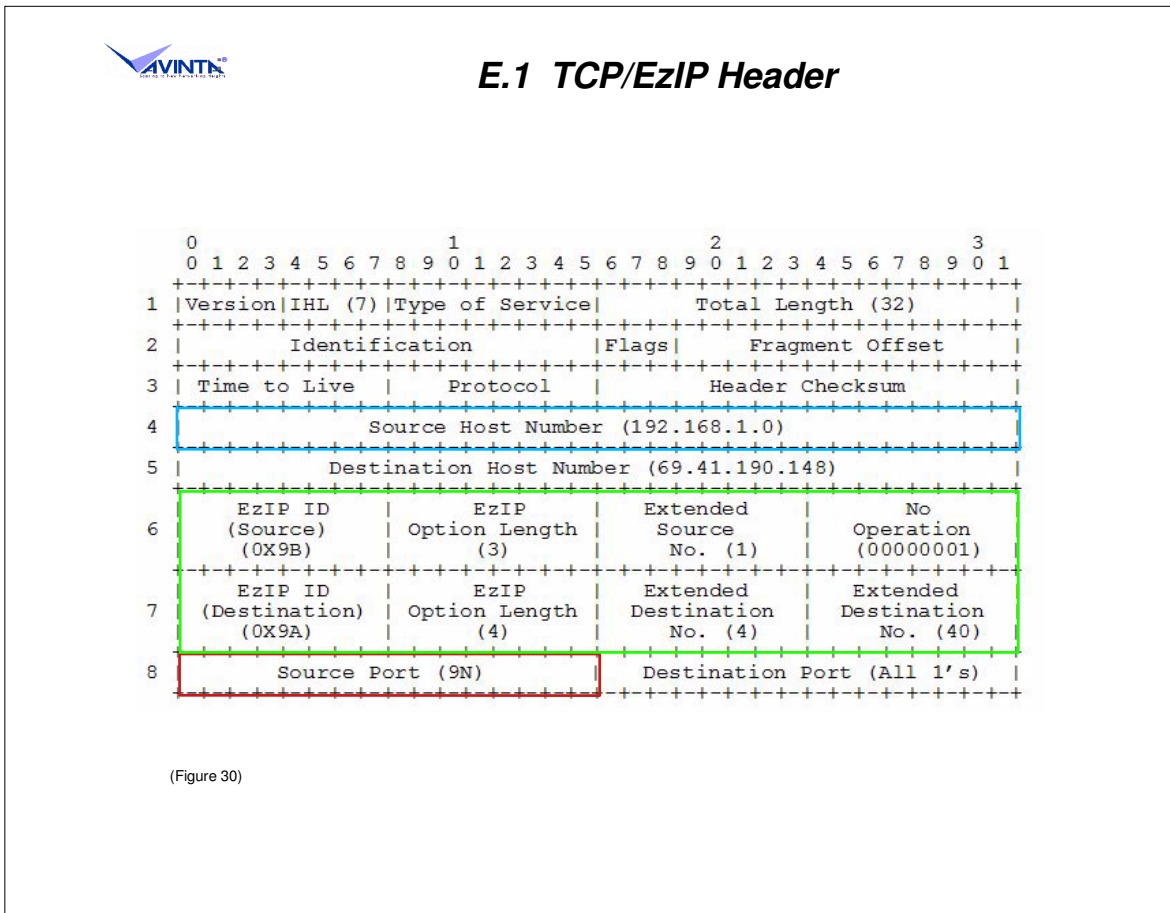
- ▶ **This is a basic 5 word IP header.**
- ▶
- ▶ **The Source and Destination Host Numbers as highlighted in blue and red colored rectangles, respectively, are used to direct the associated packet to be sent from the Source address to the Destination address.**
- ▶
- ▶ **Note that the Source Host Number in word 4 is a private network address, indicating that the initiating IoT device is on a private network.**



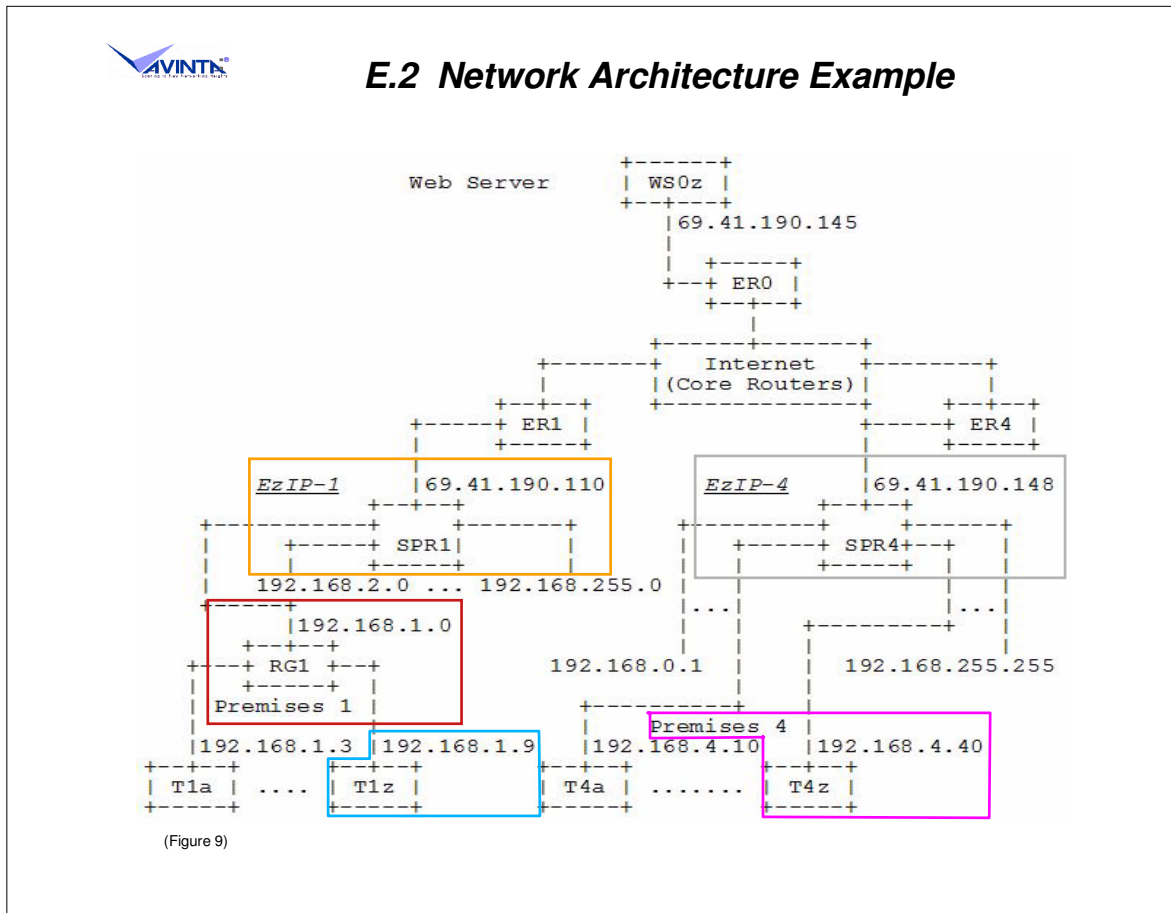
- ▶ When the preceding IP packet went through its RG, the NAT in the RG appends a word 6 to the IP header and assigns a Port number (red box) to the packet. Since the private network configuration is not known yet, the second half of word 6 is left blank (filled with all 1's).
- ▶
- ▶ The packet moves on with its TCP/IP header masquerading the RG's IP address (blue box).
- ▶
- ▶ Note that the Port Number "3N" is constructed by suffixing an "N" to the value of the fourth octet of the IoT's private network address, which is normally assigned by the DHCP server in the RG. This notation is to indicate that this NAT assigned number is associated with the IoT whose last octet value is 3



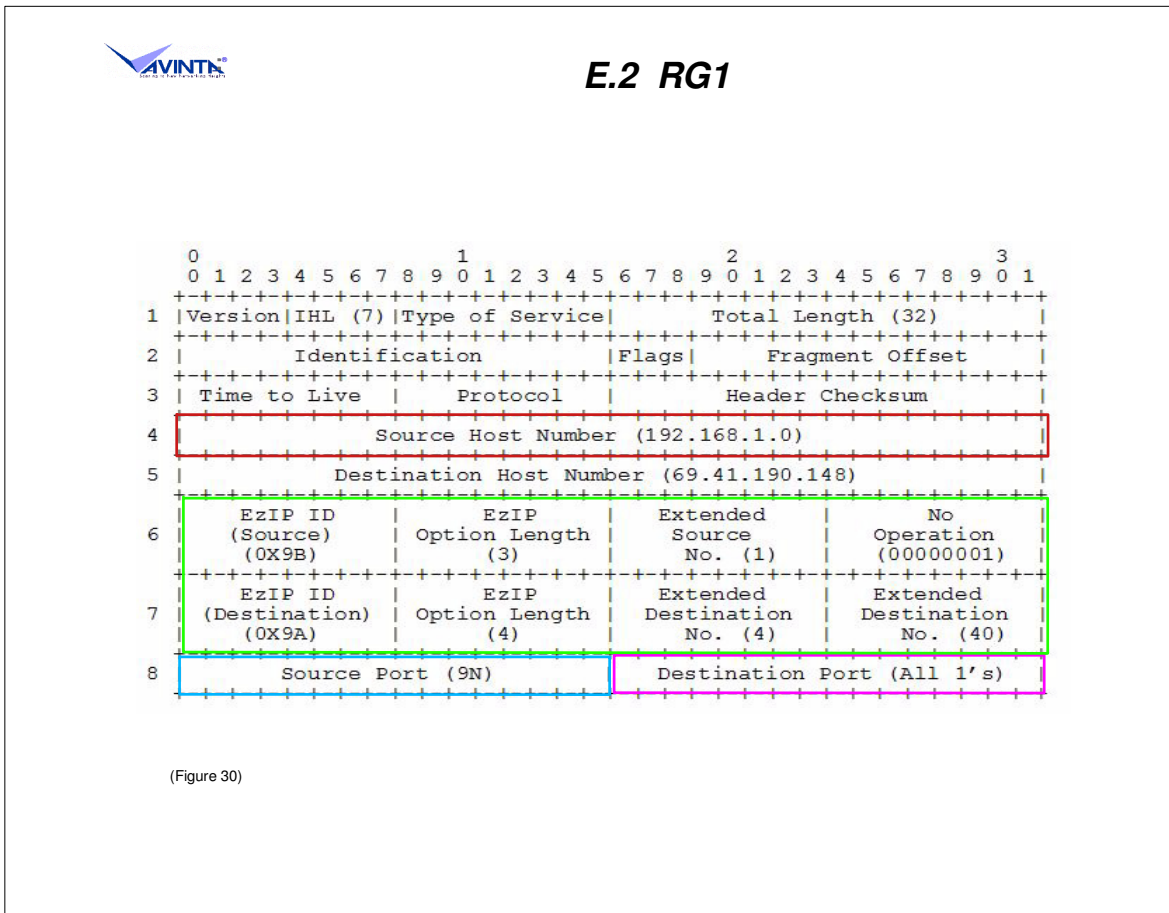
- ▶ An EzIP capable IoT, knowing the EzIP addresses of both ends, starts with an IP header by filling the Source Host Number (blue box) with its own private network address, and the Destination Host number with the destination's public address (red box), plus
- ▶
- ▶ Includes in the header with two words (#6 and #7) that carries the Semi-Public address extensions of either end of the intended link (green box).
- ▶
- ▶ A brief description of the Option word:
- ▶
- ▶ The first octet of an Option word is the Option type ID that defines the meaning of the remaining word. The second octet is the length of the Option word. So, the actually number of EzIP address octets of an Option word is "the Length minus two".
- ▶
- ▶ Word 6 is filled with three octets because the Extended Source Number has only one octet. The last octet is filled up with a NOP Option ID code.
- ▶
- ▶ For word 7, the EzIP address consists of two octets. Thus, a full word is used.



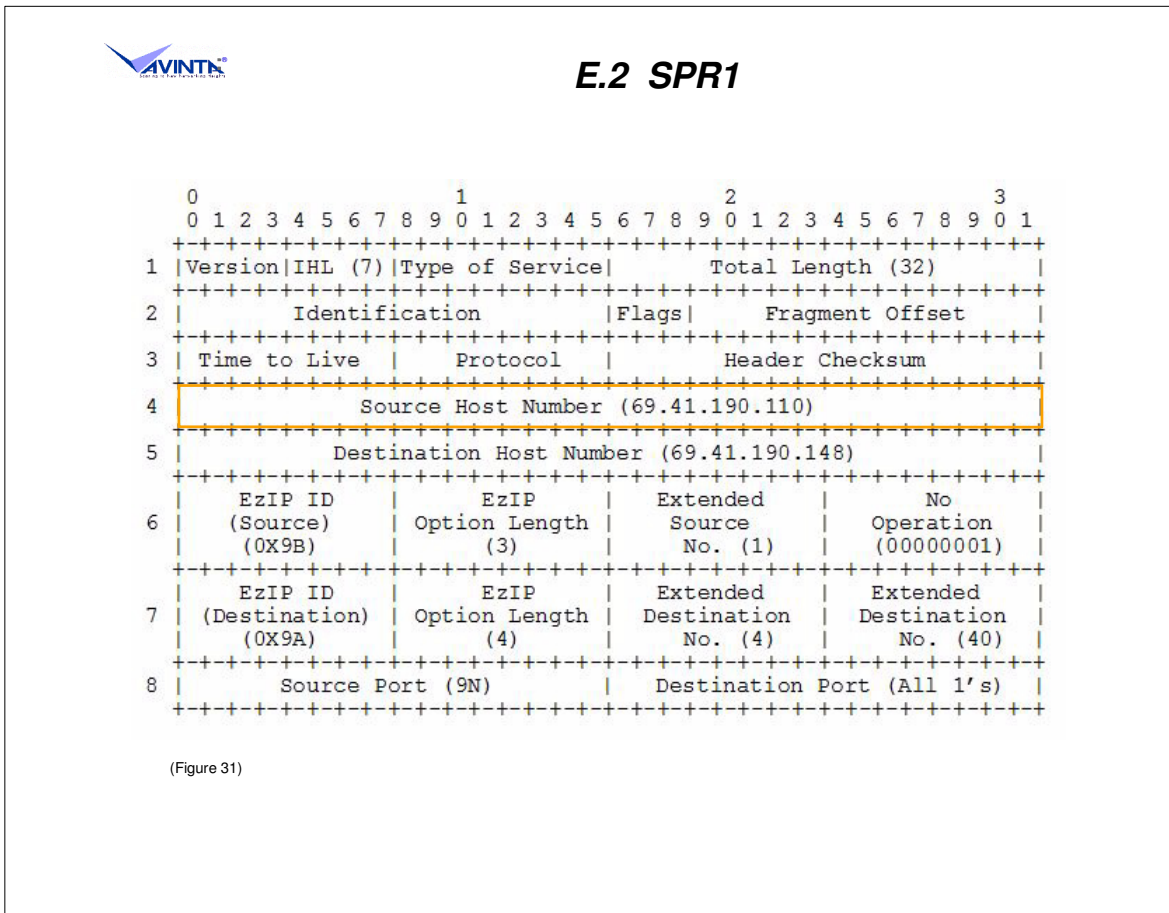
- ▶ **When the preceding EzIP packet went through its RG, the NAT in the RG appends a word 8 to the EzIP header and assigns a Port number (9N - red box) to the packet. The packet moves on with its TCP/EzIP header masquerading the RG's IP address (192.168.1.0 - blue box).**
- ▶
- ▶ **The EzIP portion of the address information (green box) is copied over but not altered.**



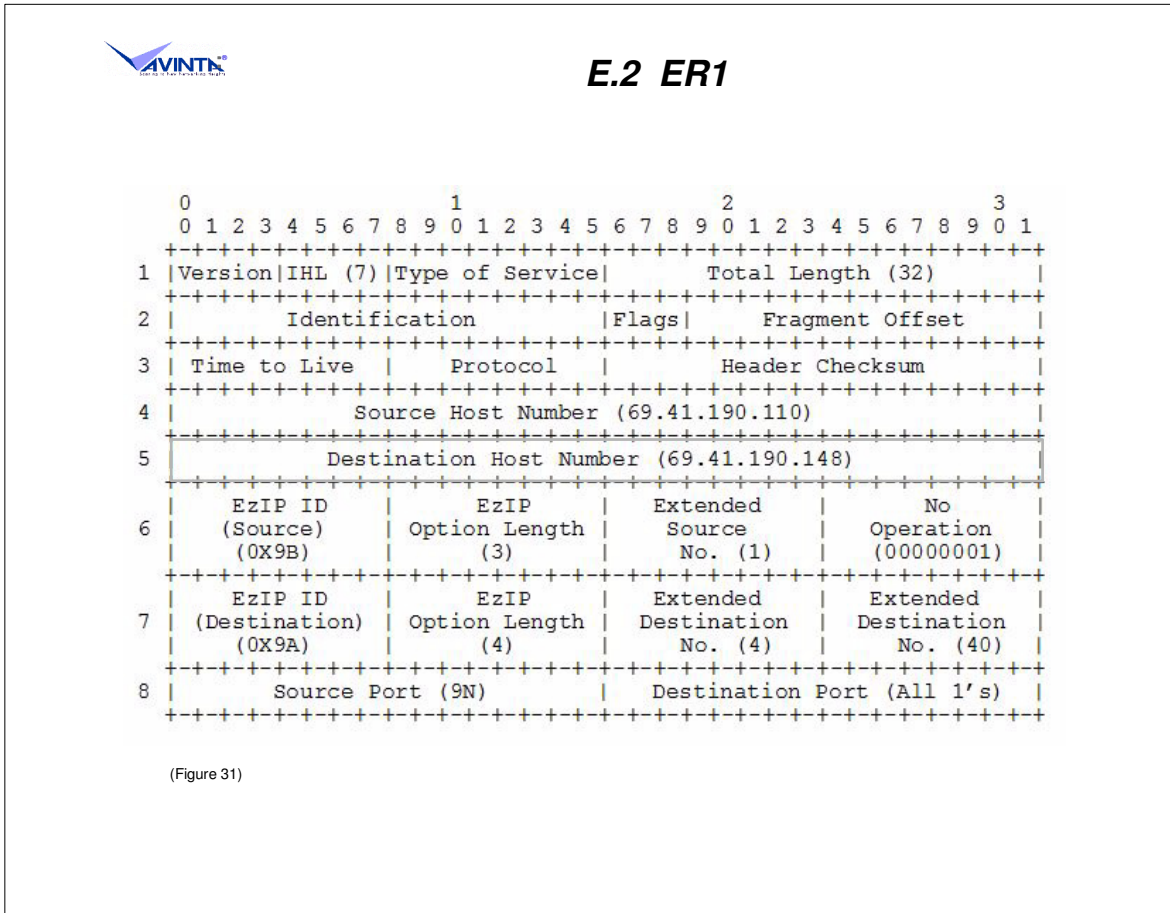
- ▶ To describe how the IP header may evolve as a packet traverses through the Internet, we need to follow a path between the two ends. Among quite a few possible EzIP configurations, we will use the most common private network block, 192.168/16 as an example.
- ▶
- ▶ This partial architecture diagram consists of two types of EzIP manifestations:
 - ▶ The EzIP-1 with SPR1 (69.41.190.110 - brown box) on the left makes use of only the third octet for Semi-Public routing in the form of 192.168.K/24, resulting in 256 premises, each may have an RG, like RG1 (192.168.1.0 - red box) on Premises 1, to distribute the remaining 256 addresses to IoTs, such as T1z (192.168.1.9 - blue box). This part of the diagram is the equivalent of the graphics shown earlier.
 - ▶
 - ▶ The EzIP-4 with SPR4 (69.41.190.148 - gray box) on the right makes use of both the third and the fourth octets of 192.168.K.L/32 for Semi-Public routing, resulting in 64K directly connected IoTs, such as T4z (192.168.4.40 - pink box) on Premises 4.
 - ▶
- ▶ On the second page of the Reference handout, there is a copy of this diagram. It would be handy to refer to it as we go through the following slides.



- ▶ **As the EzIP packet goes through RG1, the NAT function in the RG1 appends the header with word 8, assigns a TCP Port number ("9N" - blue box) to the packet and then lets the packet masquerade with the RG1's own IP address as the Source Host Number (192.168.1.0 - red box) .**
- ▶
- ▶ **Since there is no Destination Port number, the second half of word 8 is filled with 1's (pink box).**
- ▶
- ▶ **The EzIP Option words (green box) are copied to the new header unaltered.**

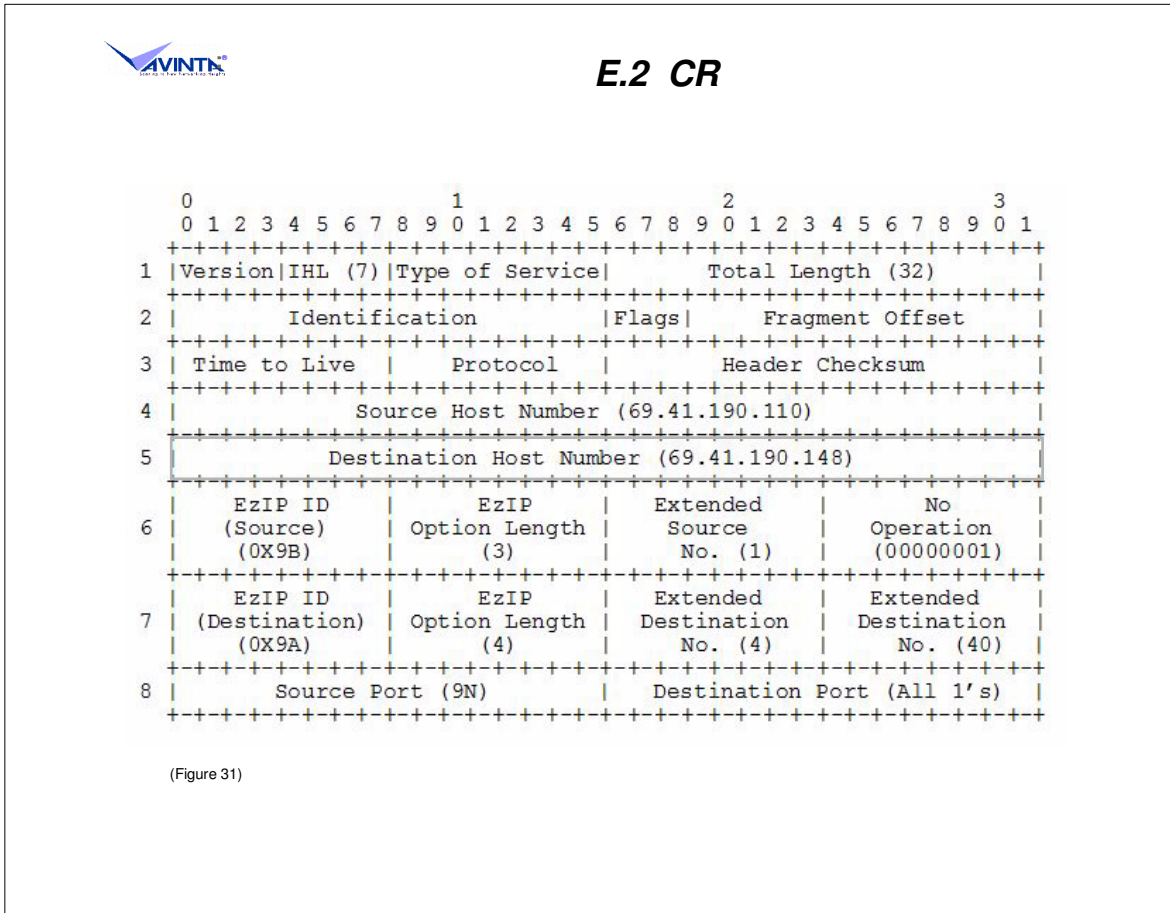


- ▶ **When the EzIP packet goes through the SPR1, the SPR1 replaces the Source Host Number by its own public address (69.41.190.110 - brown box),**
- ▶
- ▶ **while leaving everything else in the EzIP header unchanged.**
- ▶
- ▶ **The IP header, with the first five words having the standard IPv4 public addresses in words 4 and 5, is now ready to go through the Internet.**

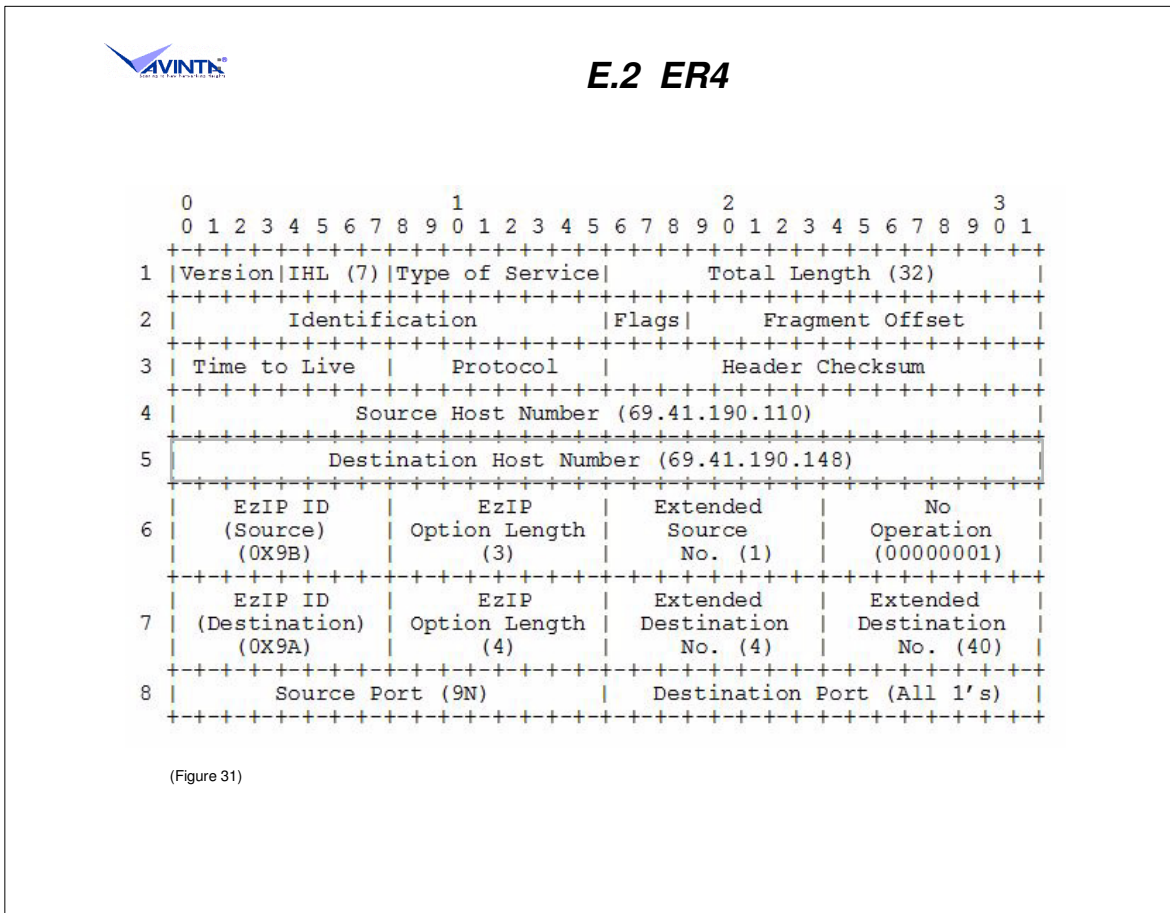


- ▶ **Without changing anything in the EzIP header, the IP packet is forwarded by ER1 into the Internet according to its Destination Host Number (69.41.190.148 - gray box).**

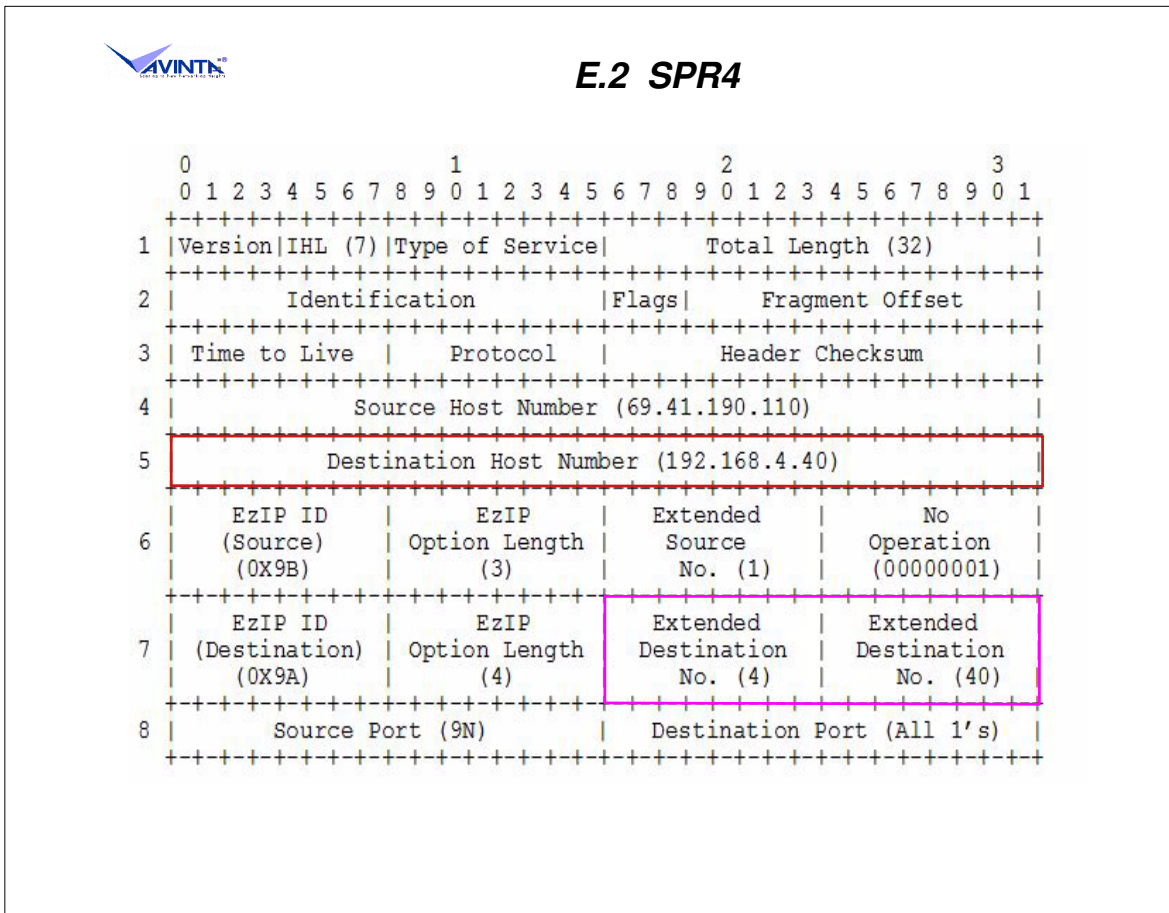
▶



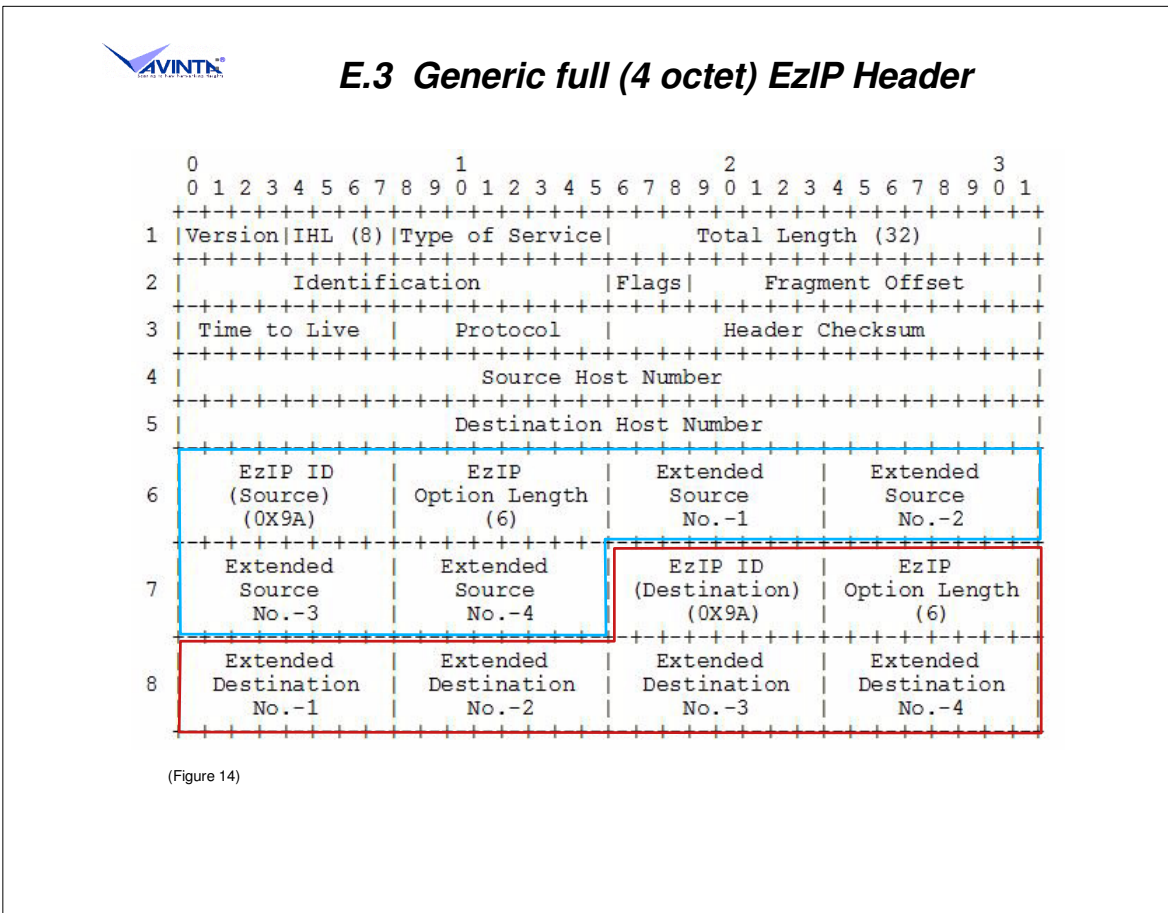
- ▶ **Without changing anything in the EzIP header, the packet is forwarded by CR according to the value in its Destination Host Number (69.41.190.148 - gray box) towards ER4.**



- ▶ **While all EzIP header information remain unchanged, the packet is forwarded by ER4 according to the value in its Destination Host Number (69.41.190.148 - gray box) towards SPR4.**



- ▶ **According to EzIP ID (Destination) information in the first octet of word 7 (0X9A - meaning the extended address prefix is 192.168/16, and the extension part consists of two octets), SPR4 reconstructs IoT, T4z address (192.168.4.40 - red box) from the Extended Destination Number (4.40 - pink box) as the Destination Host Number and then delivers the IP packet to T4z.**
- ▶
- ▶ **This completes the full sequence of header transitions through the Internet.**



- ▶ **This EzIP header format allocates two 6-octet Option words for transporting the Extended address information on either end of a link.**
- ▶
- ▶ **Each is capable of transporting a full IPv4 address (4 octets). This is necessary for making use of the proposed 240/4 block, because it requires at least 28 bits to carry its routable addresses.**
- ▶
- ▶ **This full length format is a superset of the EzIP header examples shown earlier. With EOOL and NOP Option codes to fill the gaps, this format is also applicable to Semi-Public address extensions of shorter and different lengths on either end of a link.**



F. Goals / Takeaways

- Technical Flaws?
- Merits / Benefits for the mass?
- Next step?

- ▶ **EzIP is a novel approach that utilizes the originally defined IPv4 protocol (RFC791) and the reserved addresses (240/4 block) to multiply the public assignable pool by 256M times. There is no new technology to develop, just system configuration efforts.**
- ▶
- ▶ **This scheme may be implemented as software or firmware enhancement to existing ERs, or RGs, or new stand-alone inline SPRs, depending on the deployment considerations. Either way, the engineering effort required is expected to be minimal.**
- ▶
- ▶ **Because the SPR is to be realized outside of the current Internet perimeter of ERs, the SPR deployment may be carried out by individual geographical areas or at the national level, on as-needed basis.**
- ▶
- ▶ **The goal is to deploy this scheme universally for benefiting the whole Internet. However, developing regions with limited access to technical and financial resources may derive the immediate relief by being able to continue utilizing their IPv4 investments through finite local reprogramming efforts.**