Published January 1, 2001. Distribution Restricted to Sponsors Until April 1.

WHITE PAPER

The Electronic Product Code (EPC)
A Naming Scheme for Physical Objects

David L. Brock

MIT AUTO-ID CENTER MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVENUE, BUILDING 3-449G, CAMBRIDGE, MA 02139-4307

ABSTRACT

For over twenty-five years, the Universal Product Code (UPC or "bar code") has helped streamline retail checkout and inventory processes. As one of the most successful standards ever developed, UPC coding and labeling methods have grown to include numerous elements of the supply chain. The emergence of the Internet, the digitalization of information and the globalization of business offer new possibilities for product identification and tracking. To take advantage of this network infrastructure, we propose a new object identification scheme, the Electronic Product Code (EPC), which uniquely identifies objects and facilitates tracking throughout the product life cycle. The EPC is a short, simple and extensible code designed for efficient referencing to networked information.

WHITE PAPER

The Electronic Product Code (EPC)
A Naming Scheme for Physical Objects

Biography

by David L. Brock
Co-Director

Dr. David Brock received Bachelors degrees in theoretical mathematics and mechanical engineering from MIT, and his Masters and Ph.D. degrees from the Department of Mechanical Engineering at MIT with an affiliation to the Artificial Intelligence Lab. He is currently a Principal Research Scientist in the Laboratory for Manufacturing and Productivity and Co-Director of the MIT Auto-ID Center. Dr. Brock is also the Founder of Brock Rogers Surgical, a manufacturer of robotic medical devices. Dr. Brock has worked with a number of organizations including the Artificial Intelligence Laboratory, the Massachusetts Eye and Ear Infirmary, DARPA, Lockheed-Martin, Loral, BBN and Draper Laboratories.

Published January 1, 2001. Distribution Restricted to Sponsors Until April 1.

WHITE PAPER

The Electronic Product Code (EPC)

A Naming Scheme for Physical Objects

Contents

WHITE PAPER

The Electronic Product Code (EPC)

A Naming Scheme for Physical Objects

Contents

For over twenty-five years, the Universal Product Code (UPC or "bar code") has helped streamline retail checkout and inventory processes. As one of the most successful standards ever developed, UPC coding and labeling methods have grown to include numerous elements of the supply chain. The emergence of the Internet, the digitalization of information and the globalization of business offer new possibilities for product identification and tracking. To take advantage of this network infrastructure, we propose a new object identification scheme, the Electronic Product Code (EPC), which uniquely identifies objects and facilitates tracking throughout the product life cycle. The EPC is a short, simple and extensible code designed for efficient referencing to networked information.

1. INTRODUCTION

Our vision is to create a "Smart World," that is, an intelligent infrastructure linking objects, information and people through the computer network. This new infrastructure will allow universal coordination of physical resources through remote monitoring and control by humans and machines. Our objective is to create open standards, protocols and languages to facilitate worldwide adoption of this network – forming the basis for a new **"Internet of Things**."

This paper describes one of the first activities in the creation of this new network infrastructure; that is, the design of a naming scheme to enumerate and identify physical objects. Termed the Electronic Product Code (EPC), this standard will serve not only as a next generation Universal Product Code (UPC), but also as a method to identifying components, assemblies and systems.

In the formation of this standard, we must carefully consider its purpose and scope, as well as its practical application in commerce and industry. We must also evaluate current product and shipment numbering standards, as well as other schemes from other domains, which have been successfully adopted.

2. THE INTELLIGENT INFRASTRUCTURE

We envision an intelligent infrastructure which automatically and seamlessly links physical objects to the global Internet. This network will have a number of important characteristics. First, unlike conventional bar-codes, this system will network physical objects without human intervention or manipulation by automatic machines. Second, the network will be seamless. In other words, it will operate continuously throughout the environment rather than just at the checkout or exit, as with current scanning and Electronic Article Surveillance (EAS) systems.

Third, the network will be relatively inexpensive. Although it was technically possible to create such a network in the past, new technologies make its practical implementation a reality. Fourth, the network will be ubiquitous, operating in many diverse environments–from manufacturing plants and distribution centers to retail establishments and homes. Finally, the network will adopt standards in cooperation with governing bodies, such as the Uniform Code Council (UCC), the European Article Number (EAN) Association, the American National Standards Institute (ANSI) and the International Standards Organization (ISO), as well as commercial consortium and industry groups.

This intelligent infrastructure has four major components: (1) electronic tags, (2) Electronic Product Code (EPC), (3) Physical Markup Language (PML) and (4) Object Name Service (ONS). We will briefly present these components and describe how the Electronic Product Code is a key element in this vision.

2.1. Electronic Tagging

Electronic tags refer to a family of technologies which transfer data wirelessly between tagged objects and electronic readers. Traditional bar codes may be considered electronic tags under this definition, since the printed bars reflect light and communicate their data to the laser scanner. Electronic Article Surveillance (EAS) labels also communicate information wirelessly, but only a single bit of data – whether or not an item has been purchased [1].

Radio Frequency Identification (RFID) tags, often used in "smart cards," have small radio antennas, which transmit information over a short range [2]. RFID technology may use both powered and non-powered means to activate the electronic tags. Powered devices use batteries to actively transmit data to more distant readers. Electronic highway toll systems are good examples of active RFID tags. Passive RFID devices typically use inductive coupling from an active reader to both power the tag and transmit the data.

Electromagnetic Identification (EMID) technology is a more broad designation for wirelessly tagging systems. The Motorola BiStatix™ tags, for example, use capacitive coupling to transmit data [3]. Both the reader and tag use simple conductive surfaces for the antenna. Since conductive surfaces can be printed, fabrication costs are significantly reduced as compared to the inductive windings of the RFID tags. Future versions may cost a few cents, making widespread adoption practical.

Electronic tags, when coupled to a reader network, allow continuous tracking and identification of physical objects. Reader arrays have been fabricated and integrated in floor tiles, carpeting, shelf paper, cabinets and appliances. Similar to cellular phone grids, the reader network may provide seamless and continuous communication to tagged objects. In order to access and identify these objects, we need a means to uniquely name them.

2.2. Electronic Product Code

The Electronic Product Code (EPC) was conceived as a means to identify all physical objects. The EPC code must be sufficiently large to enumerate all objects, and to accommodate all current and future naming methods. It should provide for industry coding standards, such as those from the Uniform Code Council (UCC) and the European Article Numbering (EAN) International. These standards include original the Uniform Product Code (UPC), as well as other numbering schemes, such as the Shipping Container Code (SCC-14) and the Serial Shipping Container Code (SSCC-18). The EPC should, as much as possible, be universally and globally accepted. Since the EPC is used primarily to link physical objects to the network, it should serve as an efficient information reference. Finally, the code should be extensible, allowing future expansion in both size and design.

2.3. Product Markup Language

The Electronic Product Code serves as a reference to information on the computer network. There are, of course, many methods for storing information "on-line." These include properitary and commercial databases, and relational databases, such as Structured Query Language (SQL) file systems. Web pages written in the HyperText Markup Language (HTML) are now one of the most common means of storing digital information. New approaches, such the eXtensible Markup Language (XML), promise a universal means for structured information.

In order to describe physical objects, we propose a new language specific for that purpose – the Physical Markup Language (PML). Rather than a new syntax, the PML will be based on the eXtensible Markup Language and include a set of schema describing common aspects of physical objects. Industry specific representations could then be "plugged into" the common framework or derived from the shared data.

2.4. Object Name Service

The Object Name Service (ONS) is the "glue," which links the Electronic Product Code (EPC) with its associated Physical Markup Language (PML) data file. More specifically, the ONS is an automated networking service, which, when given an EPC number, returns a host address on which the corresponding PML file is located.

The ONS, currently under development, is based on the standard Domain Naming Service (DNS). When complete, the ONS will be efficient and scaleable, designed to handle the billions of transactions which are expected.

3. IDENTIFICATION CODES

To properly develop a new coding standard, it is important to examine current methods, not only in the same domain, but in other areas which have been successfully deployed. In this section, we will briefly consider various identification schemes including the Uniform Code Council (UCC) and European Article Numbering (EAN) International standards, as well as numbering systems from other areas, such as vehicle numbers, addresses and telephone numbers. There are common and successful methods in their design which we will want to adopt.

3.1. The Universal Product Code (U.P.C.)

The Universal Product Code (UPC.) is perhaps one of the most successful standards ever developed. It has for years sped transactions in retail stores, shopping malls and supermarkets. The Uniform Code Council (UCC) UPC code numbering system is a 12 digit numeric sequence. The standard symbol, which represents the UPC code is a series of light and dark lines, together with a human readable numeric equivalent, as shown in Figure 1.

Figure 1. The Universal Product Code (UPC) is a numeric sequence representing manufacturer and product.

THE UNIVERSAL PRODUCT CODE

The UPC numeric code (which is distinct from its “bar code symbol” representation) consists of four partitions [4]. The first partition consists of a single digit indicating the numbering system used to interpret the remaining characters. A ‘0’, for example, designates a regular UPC code, a ‘3’ a National Drug Code and a ‘5’ a coupon.

The next five digits designate the manufacture identification number, and the following five, the item number. Item numbering is maintained by the associated manufacture that must ensure unique numbers for each product type. Finally, a single digit is added as a check character used to validate the correct interpretation of the machine scan.

The UPC code can, therefore, provide up to 10 unique numbering systems, 100,000 manufacturer identifiers, and 100,000 product types for each manufacturer.

In addition to the standard symbol shown in Figure 1, there is a zero-suppression version for smaller bar codes used on bottles and cans [5]. There is also a European variant, the UCC/EAN-13, administered by the European Article Number Numbering Association (EAN).

3.2. Supply Chain Identification

In addition to the item numbering standards outlined above, the Uniform Code Council and the European Article Numbering Association have developed many other numbering standards for supply chain tracking and identification [6]. These include numbering standards for books, magazines, greeting cards, audio/video products and drugs. There are also numbering standards for packaging and containers.

These coding methods have various configurations, lengths, partitions and bar code symbols. In addition to product identification, many of these coding standards have descriptive information. These include information on currency, price, weight, volume, expiration date, etc. In fact a variation, the UCC/EAN-128 Application Identifiers (AI), was developed specifically to carry information about items and shipments [7].

The need to carry greater information about a product, and the desire to track a wider variety of item, has lead to a great disparity in numbering methods. Furthermore, the inherent limitations of linear bar codes have lead to other approaches, such as two-dimensional “bar-codes” and embedded electronic memory. In fact, it is the recognition of these needs, and an acknowledgement of pervasive networking, which lead to the approach presented here.

3.3. Other Numbering Schemes

3.3.1. License Plate Numbers

Automobile license plates are typically composed of a series of letters and numbers. In the United States license numbers are issued and administered by each state, and generally composed of six or fewer characters.

Perhaps more than any other scheme, license numbers must be easily read and remembered after only a brief encounter. For this reason, they are typically grouped into one or two partitions consisting of three or four alphanumeric characters.

A typical license number, composed of six characters, allows a maximum address space of 36^6 or approximately 2.2 billion numbers. Clearly, this exceeds the number of operating vehicles in any one state or province.

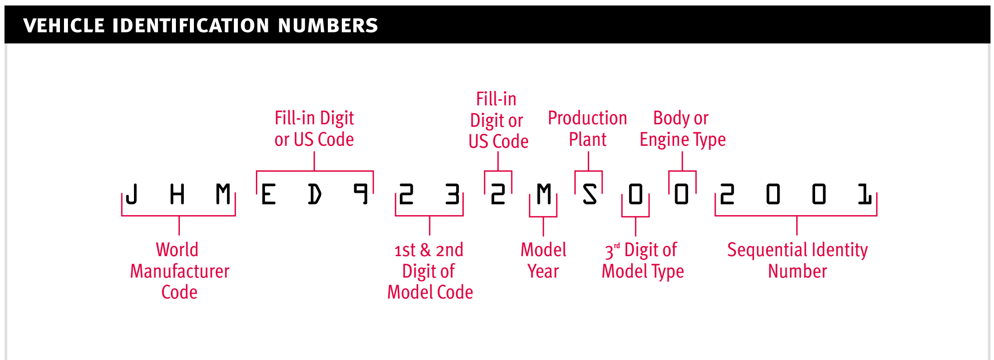
This calculation, however, over estimates the usable namespace, since more restrictive numbering partitions are currently adopted, such as three letters and three numbers, or four numbers and two letters, etc. We can assume, though, that these partitions could be changed should the need arise.

3.3.2. Vehicle Identification Numbers

Vehicle Identification Numbers (VIN) first appeared on American automobiles in 1954. The VIN was intended to identify the vehicle and provide a condensed description (see Figure 2). Early VIN numbers came in a variety of formats and configurations. The only way to interpret these numbers was through a VIN decoder.

In 1981, the U.S. National Highway Traffic Safety Administration required all over-the-road vehicles to contain a standard, 17-character VIN. In later years – in an effort to prevent theft – the Department of Transportation required VIN numbers to marked on major components [8].

Figure 2. The Vehicle Identification Numbers (VIN) identify vehicles and provide a condensed vehicle description.

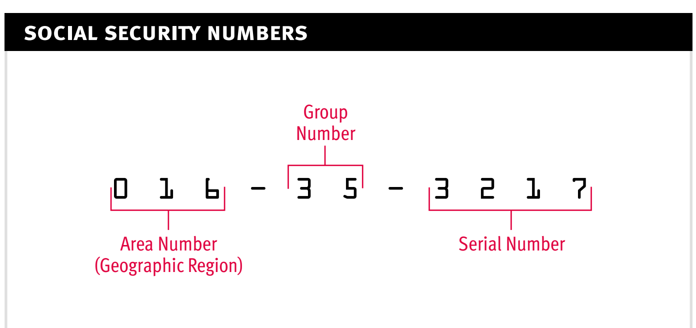


3.3.3. Social Security Numbers

In 1935, the U.S. established a permanent national pension system, financed through contributions by both the employee and employer. As a means of identifying the beneficiaries, unique Social Security Numbers were issued. In the United States, these Social Security Numbers (SSN) also serve as a common means of personal identification – from driver's licenses, to voter registration, to medical records.

The Social Security Number is a simple 9-digit number, divided into three sections of three, two and four digits, as shown in Figure 3 [9]. The SSN can therefore provide only a total of 1 billion unique identifiers. Certainly enough for the U.S. population, but insufficient if applied globally.

Figure 3. The Social Security Numbers are used in the United States for personal identification.



3.3.4. ISBN Numbers

The International Standard Book Number (ISBN) system was established in 1968 as a standard means for identifying books and publications. The ISBN numbering scheme has three data partitions, comprising the country of origin, publisher, and title and edition [10]. The ISBN number is an example of a variable partition scheme. The publisher and title share 8 digits, with 3 to 5 digits for the publisher and the remainder for the title, as illustrated in Figure 4.

In this way, the fewer publishers with many volumes have more title numbers available, while the many publishers with few titles have a smaller allotment.

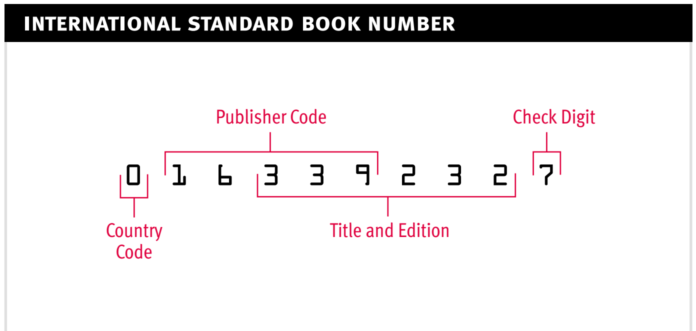


Figure 4. The International Standard Book Number (ISBN) uniquely identifies the publication.

3.3.5. Address Codes

While address codes do not correspond to any particular item, they do define a geographic region, and, in some sense, may be considered a type of object code.

In 1962, President Kennedy appointed an Advisory Board of the United States Post Office Department to study the mechanics of mail movement and to develop recommendations for its improvement. One of the prime objectives was the development of a coding system, which had been under consideration for many years. In 1963, the Department finalized the code, and on April 30, 1963 Postmaster General John A. Gronouski announced the ZIP (Zoning Improvement Plan) Code.

The five-digit code ZIP code was assigned to every address in the United States [11]. The first digit represented broad geographic areas, from zero in the Northeast to nine in the far West. The following two digits were related to population concentrations or regions accessible to common transportation. The final two digits designated post offices or postal zones in larger cities.

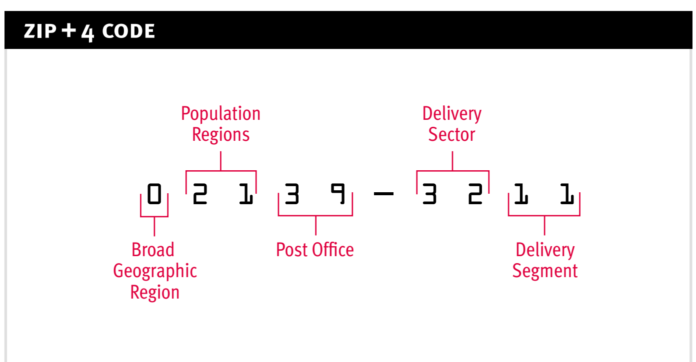


Figure 5. The ZIP + 4 code appends four digits to the traditional ZIP code.

In 1983, the ZIP+4 code was introduced, which added a four digit number following the traditional ZIP code, as shown in Figure 5. The sixth and seventh numbers of the new code denote a delivery sector, which may be several blocks, a group of streets, a group of post office boxes, several office buildings, a single high-rise office building, a large apartment building, or a small geographic area. The last two numbers denote a delivery segment, which might be one floor of an office building, one side of a street between intersecting streets, specific departments in a firm, or a group of post office boxes.

3.3.6. Telephone Numbers

As with the other numbering schemes, the telephone number format varies throughout the world. In the United States, the telephone number partitions correspond to network regions, subregions and individual extensions, as shown in Figure 6 [12]. International telephone numbers follow this pattern, though many have variable partition schemes allowing some regions to have a greater number of possible numbers. In addition, the country code varies from 2 to 3 digits.

Telephone numbers, in particular, require small partitions of few numbers to allow simple recall. In fact, the number of partitions, and the number of characters in each partition, are closely related to the capacity of human short-term memory. Any new number system requiring memory and recall must follow this same pattern.

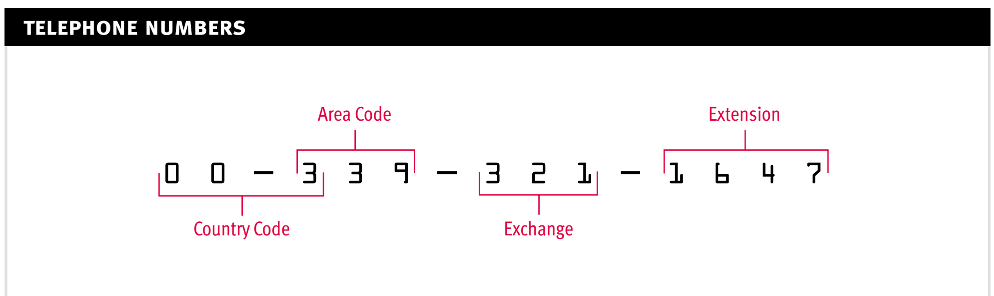


Figure 6. Telephone numbers in the United States have fixed partitions, which correspond to network regions, subregions and individual extensions.

3.3.7. Internet Address

The Internet address is, perhaps, one of the latest encoding schemes [13]. The **Internet Address** or **IP address** is a globally accepted method for identifying computers attached to the Internet. The address is a standardized, compact binary address that makes computations like routing decisions efficient. Specifically, an IP address encodes the identification of the network to which a host attaches as well as the identification of a unique host on that network.

Each host on a TCP/IP internet is assigned a unique 32-bit internet address that is used in all communication with that host.

Each address consists of a header and a pair (**netid, hostid**), where **netid** identifies the network, and **hostid** identifies a host on that network. The header varies in length from 1 to 5 bits. The header indicates the structure and partitions of the remaining bits. A Class A address, for example, is indicated by a 0 in a 1-bit header, and divides the remaining address into 7 bits for the **netid** and 24 bits for the **hostid**. This provides a total address space of 2^7 , or 128 networks, and 2^{24} , and 1,6777,216 hosts. A Class B address, shown in Figure 7, is used for intermediate sized networks allowing 2^{14}, or 16,384 networks, and 2^{16}, or 65,536 hosts. For readability, IP addresses are typically reduced to a four decimal integers, separated by dots as shown in Figure 7.

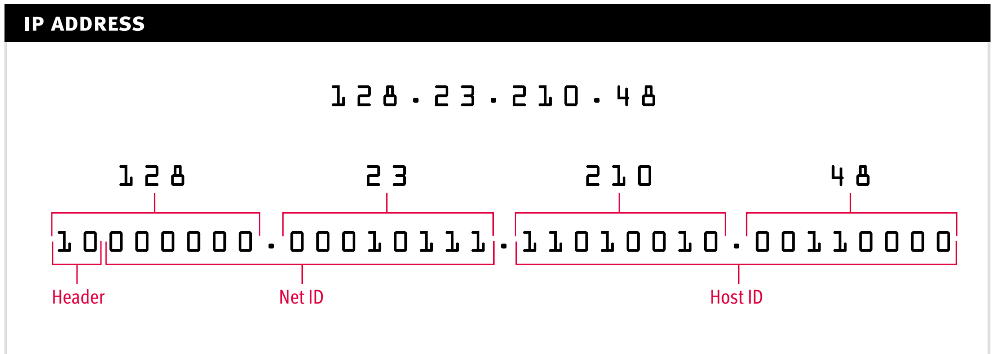


Figure 7. The **Internet address** or **IP address** is a 32-bit number which uniquely identifies a networked computer host. The first two bits for a Class B address, shown here, define the remaining partitions for the **netid** and **hostid**.

3.3.8. Summary

The examples given above illustrate the range and application of various coding methods. It is by no means an exhaustive list, nor greatly detailed. The discussion, however, shows common approaches and general directions which we will use in designing the new EPC standard. It may even be possible to accommodate some of these standards under the common framework of the EPC code.

4. DESIGN

The objective of the Electronic Product Code (EPC) is to provide unique identification of physical objects. The EPC will be used to address and access individual objects from the computer network, much as the Internet Protocol (IP) Address allows computers to identify, organize and communicate with one another. In the following sections, we will consider various aspects of an object numbering scheme, and describe strategies used to design the EPC.

4.1. Unique Identification

Unlike the current Uniform Product Code (UPC), the Electronic Product Code is intended to **uniquely** identify physical objects. In other words, a unique EPC will be assigned to one and only one physical item. There are some immediate consequences to this decision.

First, there must be a sufficient number EPC codes to accommodate past, current and future needs in object identification. Such an objective, leads to the consideration of the number of all physical objects. Consider items listed in Table 1. From the population of the world (six billion) to the total number of grains of rice (roughly ten million billion), we see the EPC must have a sufficiently large address space to reference all these objects.

Table 1: Unique physical objects number into the quadrillions [14].

BITS	UNIQUE NUMBER	OBJECTS
23	6.0×10^{6} per annum	Automobiles
29	5.6×10^{8} in use	Computers
33	6.0×10^{9} total	Humans
34	2.0×10^{10} per annum	Razor blades
54	1.3×10^{16} per annum	Grains of rice

Second, there must be some means to ensure uniqueness in EPC number assignment, as well as a way to resolve conflict should identical numbers be created. This also leads to the issue of who, or what organization, should be responsible for unique EPC assignment. As with the current UPC approach, the answer may lie in distributed responsibility. In other words, multiple managers administrate a portion of the EPC namespace assigned to them. In addition to organizational and legislative management, automated software tools may aid in the creation and resolution of the EPC namespace.

Finally, there is the issue of longevity and recycling of EPC numbers. Given that some organizations must track products indefinitely, coupled with the growing capability of the tagging electronics, there may be no need or desire to reuse EPC assignments. At least in the foreseeable future, we will consider unique identification permanently associated with particular items.

4.2. Manufacturers and Products

The Uniform Code Council has nearly one million member companies. Most of these companies are larger organizations with products that require UPC codes. When we consider smaller companies, service organizations and private enterprises, we have a much larger number. In fact, current estimates of the number of companies worldwide exceed 25 million. This number is expected to grow to over 39 million in the next 10 years [15]. Clearly we must construct a standard numbering system consistent with these projections.

Every company manages a set of products and services. One issue to consider is how many different **types** of products a company manages. Most companies have relatively few products, though some – especially in the apparel industry – have as many as 100,000 different products. Again this represents a product class or, as termed in the retail industry, a skew number or stock keeping unit (SKU).

The range of product numbers varies greatly, as shown in Table 2. It is interesting to note (in this survey of EAN member companies) there are no organizations with different product types numbering over 100,000. Furthermore, when we consider many smaller companies, who are not members of any standards bodies, this number will be even less.

Table 2: The range of product numbers varies greatly from a few dozen to over 100,000, as represented in an EAN survey.

DOMAIN	MEDIAN	RANGE
Emerging market economies	37	0 – 8,500
Emerging industrial economies	217	1 – 83,400
Advanced industrial countries	1,080	0 – 100,000

4.3. Containers

Traditionally items, containers and pallets received different numbering structures, such as the Shipping Container Codes and Serial Shipping Container Codes, as mentioned in the preceding section. In the architecture we propose, there is no reason these containment objects receive a different code than the EPC. The record of items within the container, as well as the shipping data commonly associated with the traditional codes may be stored on the computer network and automatically associated with the container.

Even further, the truck in which the containers are transported, as well as the boxcar, ship or warehouse, may also received the same EPC format as the individual product. An EPC hierarchy, as illustrated in Figure 8, will represent the state of an item shipment. This hierarchy will shift and change over time.

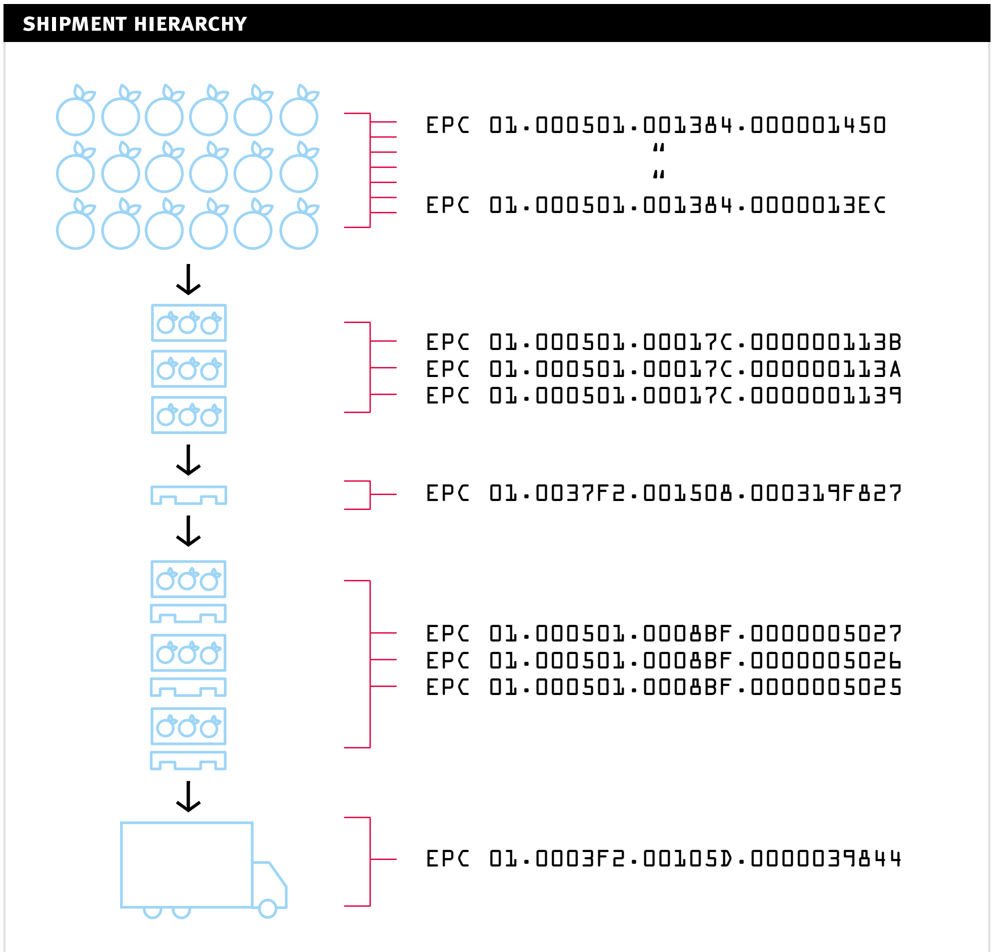


Figure 8. Elements in a shipment form a hierarchy composed of a transport vehicle, pallets, containers and items

Thus by recording the EPC containment structure, along with transition times, a partial history of product shipment can be logged.

The shipment records, however, are not complete without location information. It is clear shipping data will need to include the position of the item within a container, as well as its location along a route. Along these lines, we are developing an extension to the Electronic Product Code, which will describe item location within a container – whether that container is as pallet or a warehouse.

4.4. Assemblies, Aggregates and Collections

In addition to individual objects, the Electronic Product Code (EPC) may also identify object assemblies, aggregates and collections. An automobile, for example, is a discrete object, but is composed of numerous subassemblies and components.

We propose to use EPCs to designate assemblies and subassemblies, as well as individual items. In this case, we can describe an assembly in the same manner as the shipping data from the previous section. In fact, there is no substantive difference between a container and an assembly. Traditionally, assemblies are thought of as more complex, with linkages between components. In any case, the topological arrangement of containers and assemblies can be represented in the same hierarchical graph structure shown in Figure 8.

Beyond assemblies and containers, there are many aggregates, or collections of objects that, while having no physical connection, are associated with each other nonetheless. A dinner place setting, for example, is one unit composed of disparate items having no physical connection. These collections, however, may be assigned a unique Electronic Product Code. Furthermore, the same objects may be reorganized to form new collections, each with different EPC codes.

It is clear from the preceding discussion that the total number of Electronic Product Codes may, in fact, exceed the number of physical items. This discussion also emphasizes the need to provide a system to prevent redundant encoding.

4.5. Embedded Information

One of the major issues, and an area of continued debate, is whether to embed information in the Electronic Product Code. Current encoding standards, such as the UCC/EAN-128 Application Identifiers (AI), have data within their structure. This information, for example, may include item weight, size, expiration date, destination, etc.

We propose to eliminate, or minimize, the amount of information directly encoded within the EPC. The fundamental idea is to use the existing computer network and current information resources to store data. The EPC would then be an information reference. This desire to minimize information, however, must be balanced against pragmatic concerns, such as ease of use, simplicity and compatibility with legacy systems.

Whether or not information is stored in the EPC, we would certainly want to identify the object. Along these lines, we define **the Electronic Product Code to be that portion of an encoding scheme which uniquely identifies an item**. Therefore in the propose design, we will focus on those data elements needed to identify objects.

4.6. Categorization

Categorization; that is, the classification or grouping, of objects with common characteristics, is one of the fundamental abilities of intelligent systems and the primary means of reducing data complexity. Developing an effective taxonomy is a difficult task, since it depends intrinsically on the viewpoint of the observer.

For example, a manufacturer may classify a case of paint as an inventory asset, a shipper as a “stackable container” and a recycler as toxic waste. In every case, classification is the activity of grouping items with characteristics common to a particular viewpoint, and not an intrinsic feature of the object.

Therefore, we propose to eliminate, or minimize, classification information in the Electronic Product Code. Since classification is still an important activity, we propose to move this function to the network. More specifically, we propose higher-level software will operate on the fundamental data descriptions and “filter” objects into classification groups.

4.7. Information Reference

The primary function of the Electronic Product Code is to serve as a reference to networked information. In other words, the EPC is essentially a "pointer" to on-line data.

A common reference in use on the Internet is the Uniform Resource Identifier (URI), which includes the older Uniform Resource Locators (URL) and the Uniform Resource Names (URN). These identifiers are translated by the Domain Naming Service (DNS) into associated Internet Protocol (IP) addresses, at whose location networked information resides.

In the same way, we propose an Object Name Service (ONS) which will translate EPC codes directly into IP addresses. The hosts identified by the IP address then store relevant information about the product. The ONS is essentially the "glue" between the EPC code and networked information. The structure of the code should therefore facilitate the lookup of host addresses, and maximize the search efficiency through this object "Yellow pages."

4.8. Meta-Data

By meta-data, we mean information within the EPC, which encodes the type and structure of subsequent data. Since meta-data, does not add to the object identification process nor to embedded information, it should be minimized as much as possible. The trade-off, however, is ability to identify internal structure and to provide for future, unforeseen code requirements.

Essentially, meta-data is the version number of the EPC code embedded into its structure. We must, at a minimum, provide some ability to upgrade the code structure as requirements change.

4.9. Simplicity

There have been many standards and naming schemes devised in the past, but few see widespread adoption. One of the impediments to acceptance is complexity. More difficult schemes require longer learning times, and must be balanced by greater benefit to the user. It is our desire to create an Electronic Product Code, which is as simple as possible, yet still provides for object identification. Along these lines, it would be helpful, though not necessary, for the EPC to be immediately and intuitively obvious to the human user.

4.10. Human Interaction

Beyond simplicity, many coding systems are designed specifically for human interaction. Given our limited short-term memory (typically seven or fewer items), many codes, particularly license plate and telephone numbers, contain few partitions with few numbers. These codes are design specifically for instant recognition and simple recollection. Other codes, for example IP Addresses, are intended for machine use, but include representations for human readability. Though not intended for memory, the dotted representation of the IP Address can be written down and easily typed in manually.

In the design of the EPC code, direct human interaction is less important. Readability – as with the current UPC code and IP Address – will be necessary, but human recollection will not be required. Therefore, the EPC code should have a simple and consistent representation, which can easily be transcribed, dictated and keyed manually.

4.11. Extensibility

One of the difficulties in developing a global standard is anticipating all possible uses and future applications. Without perfect vision into the future, we must provide a simple method for expansion. Thus rather than a complete specification, we propose an initial design, leaving the majority of the EPC address space open for future definition.

4.12. Media

Physically, the EPC will be stored on some type of physical media, such as a bar-code, electronic memory or printed character. The data will be transmitted via temporal or spatially encoded electromagnetic radiation; in other words, a laser scanner or radio antenna.

For almost all media, the cost of storage and transmission raises with the amount of data. Since, we desire the EPC to be a pervasive code – with perhaps trillions of tagged items – the media must be as inexpensive as possible.

For this reason, the EPC should be of absolute minimum size, to limit cost and complexity.

4.13. Data Transmission Mechanisms

The Uniform Product Code has, as part of its definition, a checksum digit used to validate the transmission of digits from the printed label to the scanner. Check digits, start and stop bits and handshaking protocol are all devices used to aid in the reliable transmission of information. These mechanisms vary depending on the method of data transmission and the level of reliability.

Rather than embedding data transmission mechanisms into the Electronic Product Code, we propose to decouple the code from any communications protocol. All of these techniques will, of course, be used in the transmission of the EPC data, but will not be part its definition. In this way we separate object identification strategies from data transmission methods.

4.14. Privacy and Security

In the same way we separate content from transmission method, we propose to decouple the EPC definition from any security and cryptographic technique. Privacy and security are perhaps one of the primary issues in effective deployment of this networked infrastructure. Secure EPC transmission, storage and archive are vitally important for wide spread adoption.

However, rather than impose any particular cryptographic method, we propose to leave Electronic Product Code simply as a method for naming and identifying objects. All privacy and security systems would be layered on this basic identification system.

5. PROPOSAL

Given the design considerations discussed in the previous section, we must now finalize the structure of the Electronic Product Code. In the following sections, we will decide on (a) the meta-data; that is the portion of the code which describes the code itself, (b) the size of the code; that is the number of digits, characters or bits, the code will contain and (c) the number, nature and size of the code partitions.

5.1. Meta-Data

Although somewhat abstract, we must first decide on the size and nature of the meta-data. This is the portion of the code, which describes the structure of the code. The meta-data – or header – may define overall size of the code, number and size of the partitions, meaning of each partition and any other structural information of the schema.

Although meta-data is useful for providing extensibility, it limits information content for a given code size. Since the EPC will be a pervasive, lightweight code, we naturally want to limit the size of the meta-data, and rely on external standards to convey this information.

At the same time, given the ambition of the EPC to include not only traditional UCC.EAN codes, but other coding standards, we will want a high degree of extensibility.

5.2. Size

The number of bits in the EPC determines the theoretical upper bound on the number of identifiable objects. This upper bound is, of course, given by

$$N = 2^n,$$

where N is the total address space and n is the number of bits in the code. Relatively few bits should be necessary to provide unique object identity. A 96-bit code, for example, provides for 7.9×10^{28} addressable objects.

Partitions in the EPC reduce the practical address space, since underutilized segments produce gaps in subsequent sections. For example, assuming a uniform utilization rate within each partition, the total address space is

$$N = f_1 2^{n_1} x f_2 2^{n_2} x \ldots x f_m 2^{n_m},$$

where f_1 is the utilization rate, n_i is the number of bits for partition i and m is the number of partitions. Suppose a uniform 50% utilization for each of four identical partitions on a 96-bit code, the total address space is 4.9×10^{27}, a 16 times reduction in address space.

For uniform utilization over uniform partitions, the reduction in address space is

$$R = f^m.$$

Address space is thus reduced exponentially in the number of partitions, and therefore the partition count should be minimized.

It is likely partition utilization will be much less than the 50% used in this illustration. For example, the differences between items produced between manufactures vary by many orders of magnitude – from a single product to many billions.

Efficient utilization would argue strongly for a variable partition method. On the other hand, simplicity of design favors fixed partitions.

5.3. Partitions

Generally, partitions reduce address space, yet greatly increase the efficiency of searching through that address space. Suppose we have a catalog or cross reference of EPC and networked information references, such as Uniform Resource Locators (URLs), the size of any given catalog is given by

$$C = 2^{ni},$$

where C is the catalog size and ni is the number of bits for each partition i. From the previous illustration, for a 24-bit uniform partition, this catalog size is 1.7×10^7 entries – certainly within the capacity of conventional storage systems. Assuming a catalog search is proportional to its size, seek time reduces exponentially with partition count.

Partition design and catalog maintenance will also be determined by political influences. It is most likely catalogs will be managed by those who are oversee particular object sets, such as manufactures of products, facility managers and government organizations.

5.4. EPC Schema

Given the above considerations, we present a simple scheme for the Electronic Product Code. We propose a 96-bit code with a fixed, 8-bit header. This header defines the number, type and length of all subsequent data partitions. Thus, a single byte provides 256 ways to partition the remaining bits. Rather than defining all possible EPC configurations, we propose only one primary type and reserve all others for future use.

5.4.1. EPC Type I

The first EPC configuration, **EPC Type I**, is intended as a public object identifying number. It is to be used in much the same way the Uniform Product Code is used today.

The intention is to function not only as the UCC.EAN item codes, but also as the shipping identification methods, such as the Shipping Container Code (SCC-14), Serial Shipping Container Code (SSCC-18) and the Application Identifiers (AI). Together with the networked information infrastructure, the code can also consolidate many of the other standards in use today.

The EPC Type I has three data partitions, shown in Figure 9. The first data partition identifies the EPC manager; that is the manufacturer, or entity, responsible for maintaining the subsequent codes. In other words, it is the EPC manager's responsibility to maintain both object type codes and serial numbers in their domain. The EPC manager must also ensure reliable operation of the Object Name Service and for maintaining and publishing associate product documents.

Figure 9. The EPC Type 1 encoding is public object identifying number similar to the standard UPC used today.

ELECTRONIC PRODUCT CODE

01.0000A89.00016F.000169DC0

Header	EPC Manager	Object Class	Serial Number
0–7 bits	8–35 bits	36–59 bits	60–95 bits

The EPC manager partition spans a 28-bit section, encoding a maximum of 2^{28} = 268,435,456, or approximately 268 million, manufacturers. This far exceeds the current allocation provided by the UPC-12 and EAN-13.

It is likely, as with the current UPC codes, this allocation may eventually be insufficient. This is precisely why the EPC code is extensible. A repartitioning of the code may be necessary to accommodate, perhaps, billions of managers, who may oversee only hundreds of items. This may be necessary in the near future if many small entities or individuals use the EPC infrastructure to manage private, internal physical resources.

The next partition, object class, will occupy the next 24-bits. The object class may be considered the product skew or stock keeping unit (SKU). It may also be used for lot number, or any other object-grouping scheme developed by the EPC manager. Since each manufacturer is allowed more than 16 million object types, this partition could encode all the current UPC SKUs, as well as many other object classes. This is important as we expand beyond current retail applications into general supply chain and manufacturing.

The final partition encodes a unique object identification number. For all objects of a similar type, the EPC serial number provides 36-bits, or 2^{36} = 68,719,476,736, unique identifiers. Together with the product code, this provides each manufacturer with 1.1×10^{18} unique item numbers – currently beyond the range of all identified products.

6. CONCLUSION

The Universal Product Code has had phenomenal success, and demonstrates the power of standardization and industry consensus. Together with the more recent success of the Internet and the World Wide Web, we are on the verge of yet another revolution in networking and supply chain efficiency.

The Electronic Product Code, presented in this paper, is intended to provide a simple, efficient and uniform naming system for physical objects. The EPC code spans the breadth of application – from foods and drugs to assemblies and components. This naming scheme also accommodates depth throughout the supply chain, including vehicles, pallets, containers, packages and items.

The Electronic Product Code is only one piece of the new "intelligent" infrastructure. Together with the Object Name Service and the Physical Markup Language, these elements form the foundation of the "networked physical world."

7. REFERENCES

1. **Electronic Article Surveillance (EAS) summary from the AIM Global Network**
(http://www.aimglobal.org). http://www.aimglobal.org/technologies/eas/

2. **Radio Frequency Identification (RFID) summary from the AIM Global Network**
(http://www.aimglobal.org). http://www.aimglobal.org/technologies/rfid/

3. **Motorola BiStatix Technology**
http://www.motorola.com/GSS/SSTG/smartcard/3_0_bst_home.htm
http://www.motorola.com/GSS/SSTG/smartcard/white_papers/BiStatix_Whitepaper.pdf

4. **Uniform Code Council (UCC), Uniform Product Code (U.P.C.)**
Symbol Specification Manual, January 1986.
http://www.uc-council.org/reflib/01302/d36-t.htm.

5. **Uniform Code Council (UCC), Reduced Space Symbology.**
http://www.uc-council.org/rss14/.

6. **Uniform Code Council (UCC), Application Standard for Shipping Container Codes (ANSI/UCC6).**
http://www.uc-council.org/reflib/00602/d33-t.htm.

7 **Uniform Code Council (UCC), UCC/EAN-128 Application Identifier Standard (ANSI/UCC4).**
http://www.uc-council.org/reflib/00403/d30-t.htm

8. **United States Department of Transportation (DOT),
National Highway Traffic Safety Administration (NHTSA),
Vehicle Identification Number Requirements, 1995.**
http://198.17.75.65/fril/1995/19951025/95-26499.txt.

9. **Social Security Numbering Scheme, From the Social Security Administration.**
(http://www.ssa.gov). http://www.ssa.gov/history/geocard.html.

10. **International ISBN Agency,**
Berlin, Germany, The ISBN Users' Manual, 1999, ISBN 3-88053-075-0.
http://www.bowker.com/standards/home/isbn/international/isbnmanual.html.

11. **United States Postal Service**
(http://www.usps.gov), History of the U.S. Postal Service, ZIP Codes.
http://www.usps.gov/history/his2_75.htm#ZIP.

12. **Summary information on United States telephone numbering from the
North American Numbering Plan (NANP).**
http://www.nanpa.com/

13. **W. Richard Stevens, TCP/IP Illustrated, Volume 1:**
The Protocols, Addison-Wesley, 1993.

14. **Farighetti, Robert, ed., The World Almanac and Book of Facts 2000,**
World Almanac Books, 2000.

15. **Brady, Thomas, "EAN Survey,"**
Uniform Code Council Internal Document, 2000.