Published June 1, 2003. Early Release July, 2003

TECHNICAL REPORT

On the Future of RFID Tags and Protocols

Sanjay Sarma, Daniel W. Engels

AUTO-ID CENTER MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 400 TECHNOLOGY SQ, BUILDING NE46, 6TH FLOOR, CAMBRIDGE, MA 02139-4307, USA

ABSTRACT

We present a roadmap for the future of RFID protocols. We cover two dimensions: added functionality and new versions. We show that functionality can be added to RFID tags without compromising modularity. We also show that new versions of tags can evolve and co-exist with older versions without interference. A software defined radio reader can make the investments of end-users future-proof and insulate end-users from the gradual evolution of technology. This will enable the Auto-ID community to keeps its system and standards abreast of the state-of-the-art.

Published June 1, 2003. Distribution restricted to Sponsors until September 1, 2003.

TECHNICAL REPORT

On the Future of RFID Tags and Protocols

Biography

Sanjay E. Sarma
Research Director

Sanjay Sarma received his Bachelors from the Indian Institute of Technology, his Masters from Carnegie Mellon University and his PhD from the University of California at Berkeley. In between degrees he worked at Schlumberger Oilfield Services in Aberdeen, UK, and at the Lawrence Berkeley Laboratories in Berkeley, California. Prof. Sarma's Masters thesis was in the area of operations research and his PhD was in the area of manufacturing automation. From 1995 to 1999, Dr. Sarma was an Assistant Professor in the Department of Mechanical Engineering at the Massachusetts Institute of Technology. He is now an associate professor.

Daniel W. Engels
Associate Director

Daniel W. Engels received his B.S. from the University of Buffalo, his M.S. from the University of California, Berkeley, and his Ph.D. from the Massachusetts Institute of Technology all in Electrical Engineering and Computer Science. His master's thesis is in the area of computer-aided design for electronic systems, and his doctoral thesis is in the field of theoretical computer science. Dr. Engels joined the Auto-ID Center after obtaining his doctoral degree where he leads the day-to-day research activities of the Center. Dr. Engels' research interests include scheduling theory and applications, real-time system design, distributed and mobile computing, and computer-aided design for embedded systems.

Published June 1, 2003. Distribution restricted to Sponsors until September 1, 2003.

TECHNICAL REPORT

On the Future of RFID Tags and Protocols

Contents

1. Introduction ... 3
2. The Wish-list for RFID ... 3
 2.1. The Constraints ... 3
 2.2. Exploring the Options: Modes, Functions and Resources ... 4
3. RFID Classes ... 5
 3.1. The Class Structure ... 5
 3.2. Principles of Platform Modularity and Default States ... 6
 3.3. On Frequencies ... 7
4. Ongoing Generations of Standards ... 8
 4.1. On the “Generation Gap” ... 8
 4.2. A Roadmap ... 9
5. Conclusions ... 10
6. References ... 10

1. INTRODUCTION

Recent history indicates that success in communication systems comes from diligent adherence to three principles:

– **Modularity**, i.e., the abstraction of the larger system into smaller, self-contained sub-systems, each with clearly defined behaviors and sacred boundaries;
– **Standardization**, i.e., the establishment of a broad agreement on architectures, module definitions, interface definitions, languages and protocols; and
– **Renewal**, i.e., the ability and willingness to adopt improvements in technology and architecture as they become available (a facility that modularity sustains).

Together, these principles yield number of outcomes that are desirable to end-users, such as: better designed products, ease of maintenance, product choice, competition, and eventually, lower costs. For vendors too, these principles yield benefits such as: easier targets to design to, more component reusability, and most importantly, adoption on a large scale. These principles have been the foundation of our work at the Auto-ID Center. In previous papers, we have described the larger Auto-ID System, also referred to as the EPC™ system, which includes tags, readers, software, architecture, and applications. In this paper we turn the magnifying glass towards the reader-tag system and describe a coherent vision for the future progression of RFID hardware modules and standards as technology, functionality and applications evolve.

2. THE WISH-LIST FOR RFID

The wish-list of performance attributes that end-users associate with Auto-ID reader-tag systems is long, and will doubtless continue to grow in coming years. The characteristic that makes RFID system design challenging is that the resources available to the designer are very limited. The most basic requirements that come to mind are the obvious ones: range, speed and reliability. Additional functions that end-users have viewed with increasing interest include writable memory, security/authentication/anonymity functions, and sensors on tags. At the same time, the growth of other wireless technologies in the general "neighborhood" of the RFID space such as wireless peer-to-peer networking and wireless actuators must also be taken into in any roadmap. We will present the view in this paper that all these concepts fit into a rather seamless continuum in the evolution of RFID systems. While the Center has always been, and continues to be, a proponent of minimalism – for all its benefits, ranging from cost to performance – we have also always been "pro-choice." End users are always welcome to make informed choices based on their application needs and their cost constraints. The modular underpinnings of the Auto-ID System enable almost endless opportunities for RFID.

2.1. The Constraints

RFID system designers are constrained by several factors, principal among which are power consumption of the circuit, the cost of the silicon chip, and the regulatory limits imposed by the government. At a secondary level, power and cost constraints together limit the amount of silicon area available for a chip designer, and power and regulatory limits in turn impact the speed of communication of data from the reader to the tag. There are also tertiary effects related to reliability. For example, in some regulatory jurisdictions, like the 902-928 MHz ISM band in the US, communication between the reader and the tag is complicated by a frequency-hopping requirement. For a variety of reasons, frequency hopping can cause certain tags to drop in and out of the area viewed by the reader. This places an additional requirement on the reader-tag communication: it is best that transactions be either very brief, or stateless, or that the tag have a battery so that when it drops out of the reader's "zone of illumination," it continues to remember where it was in the negotiation with the reader. (For all these reasons, simplistic measures of reader-tag performance like "read rate" don't actually tell the whole story of the practical speed of reading a population.) This impacts the ability of passive tags to reliably carry out long transactions like encryption.

2.2. Exploring the Options: Modes, Functions and Resources

There are a number of ways in which the designer can mix and match possibilities to meet the user's wishes in the face of the constraints described above. Obviously it may not be possible to meet all the wishes at once, and the user must pick what he or she is willing to forgo on one front to achieve gratification on another. Cost is one constraint that the end-user must always be willing to relax for any additional functionality beyond the basic Class 0 or Class I tags. Here are some commonly discussed functions:

Memory: At the very basic level, using today's technology, additional memory will occupy more silicon, and will cost more money. Writeable memory in particular usually requires higher power, and therefore will result in a loss of performance over and above increased cost. Furthermore, the more the data stored on the chip, the longer it will take to read it and the more susceptible the transaction is to noise.

Sensors: Sensors on the tag are obviously an attractive proposition, but they come with a number of requirements beyond basic silicon area and power considerations. First, note that a sensor tag is most likely to be used in situations where another device, like a reader, is not present – otherwise the reader could itself monitor the condition that the sensor is trying to capture. Second, note that most sensors need power to operate. Putting these two points together, it is easy to see that passive RFID tags will rarely be used as "sensor-tags;" they will most likely be used with batteries which permit the sensor-tag to operate in the absence of a reader.

Another point to remember is that the data collected by sensors will in many cases be extensive. Transmitting this data reliably will require higher bandwidth than typical passive RFID tags are designed to operate at. For this reason, sensor tags will also often use broad-band protocols.

We address these two points next: batteries and bandwidth.

Batteries: The Auto-ID Center's Class 0 and Class I protocols are designed for passive, or battery-less operation. It is possible to place batteries on tags, and there are two flavors of battery-operation in tags: semi-passive and active. In order to understand these terms, it is necessary to understand that passive tags use scavenged power from the reader to run the digital logic on the chip, but not for communication – passive tags use a form of reflection to convey signals from the reader to the tag. Semi-passive tags use battery power to power the digital logic, but still use reflection, rather than active transmission to send signal from the tag to the reader. Semi-passive tags have ranges which are an order of magnitude greater than for passive tags, and have significantly higher reliability. However, semi-passive tags have shorter lives than passive tags, and larger, more delicate packaging.

Active tags use the battery not only to power the logic, but also to transmit the signal. This gives active tags another order of magnitude increase in range over semi-passive tags, and also improved reliability. However, active tags have substantially shorter life-times in terms of the number of transmissions that can be sustained by a single battery. Active tags have a "bonus" benefit: because the tags are stand-alone units which don't depend on an external power source, they can also communicate tag-to-tag. This is also referred to as "peer-to-peer networking" or "wireless mesh networking." If RFID tags could support this functionality, there is no reason why it should not be built on similar standards to today's Class 0 and Class I RFID protocols.

Bandwidth: If RFID tags are used for sensors or for mesh networking, they will need higher bandwidth than they can currently support. There are two channels to consider: the forward channel, from reader to tag, and the reverse channel, from tag to reader. The bandwidths that passive RFID protocols can support today in the forward channel are limited because the reader has two functions, power supply to the tag and signal transmission to the tag. These two functions compete. By and large, in most regulatory jurisdictions, the higher the power or field strength the reader must put out, the lower the bandwidth it is permitted to occupy. With semi-passive or active tags, however, the reader is relieved of the burden of actually powering the tag. This enables the speed of communication from reader to tag to be increased considerably. Therefore, it is possible to achieve much higher bandwidths in the forward channel for semi-passive and active tags. Of course, this additional functionality comes at a cost: the tag will be more expensive and bulky, and will have a shorter life.

Since passive RFID tags are not active transmitters, and because they merely modulate a reflected signal in the reverse channel, they are not subject to the same regulatory constraints as readers are in the forward channel. In principle, the return channel can be broadband, and much higher-speed communication is possible. The problem with a broadband return channel is that there is an increased likelihood of reader-collisions if the return signal is out-of-band. However, in sparse deployments, this may not manifest itself as a serious problem. In fact the Class 0 and Class I protocols are different in this way: the Class 0 protocol uses out-of-band signaling in the reverse channel, and the Class I protocol uses in-band signaling in the reverse channel.

Encryption: Finally, there may be a number of encryption-related functions necessary in various circumstances in RFID systems. A summary is available in [Engels 03]. At the very basic level, it may be necessary for a tag to be able to authenticate itself to a reader. This would help, for example, in detecting counterfeit tags. It may also be necessary for a reader to be able to authenticate itself to the tag. This would be necessary to protect the contents of the tag from unauthorized access. For example, consider a sensor-tag which has collected some sensitive data about Company X's process. Company Y should not be able to read that tag and acquire the sensitive information. Security or encryption primitives of various types may also be necessary to preserve anonymity. Any encryption, however, comes with costs: the tag becomes more expensive and power-hungry. Furthermore, encryption algorithms will need many cycles to complete computation. This will make encryption-based transactions difficult to complete within the length of time for which a passive RFID tag can reliably be assumed to be powered.

Summary: Several mutations of tags are possible: read-only vs. read/write memory; passive vs. semi-passive vs. active; low forward-bandwidth vs. high forward-bandwidth; low return-bandwidth vs. high return-bandwidth; and several types of encryption. It would be tragic of each of these options, or mutation thereof, demanded a different protocol and a different reader. Fortunately, we can avoid fragmentation, a problem that has plagued the industry in the past, with a modular approach as we explain below.

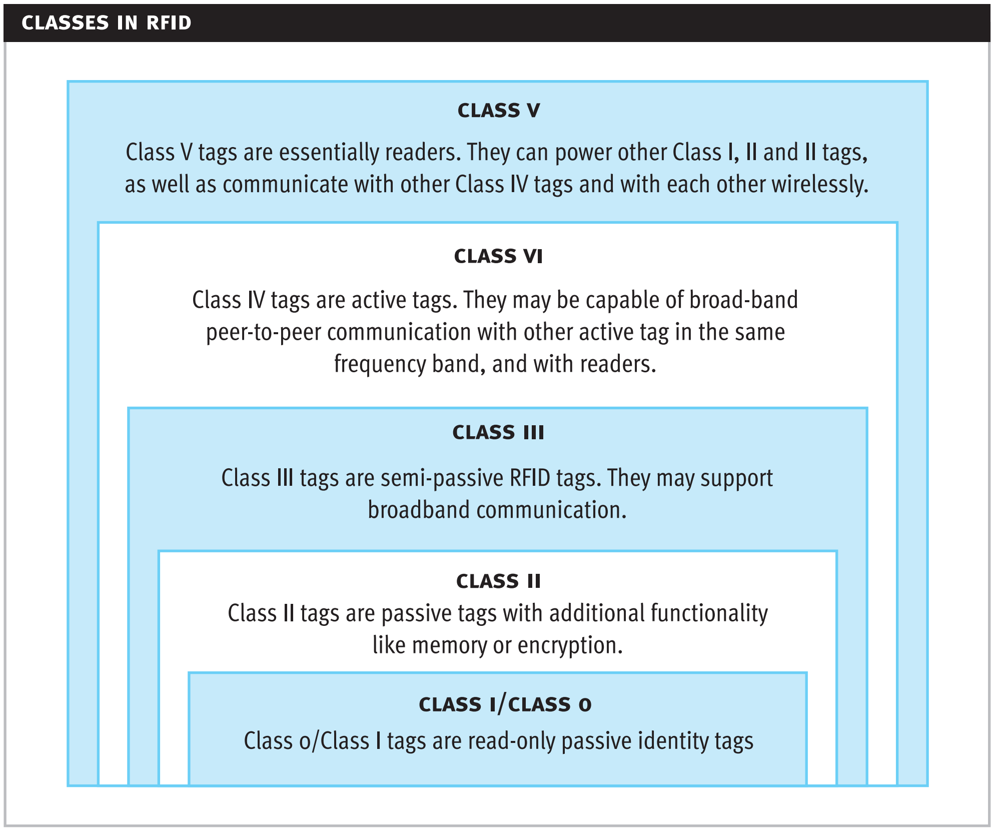
3. RFID CLASSES

3.1. The Class Structure

The Auto-ID Class structure emerged in a number of discussions between the Auto-ID Center and sponsor companies, especially Alien Technology. It was first articulated in a working draft by Roger Stewart, then CTO of Alien [Stewart 01]. The Class Structure has since evolved, and while it does not completely capture all mutations of tags, it does provide a handy vocabulary for informed discussions of RFID tags. The updated Class Structure is shown in Figure 1.

Figure 1

CLASSES IN RFID



CLASS V

Class V tags are essentially readers. They can power other Class I, II and II tags, as well as communicate with other Class IV tags and with each other wirelessly.

CLASS VI

Class IV tags are active tags. They may be capable of broad-band peer-to-peer communication with other active tag in the same frequency band, and with readers.

CLASS III

Class III tags are semi-passive RFID tags. They may support broadband communication.

CLASS II

Class II tags are passive tags with additional functionality like memory or encryption.

CLASS I/CLASS 0

Class 0/Class I tags are read-only passive identity tags

Class 0 and Class I tags represent basic capability. Class II represents the addition of functionality like encryption or memory to passive tags. Classes III, IV and V represent the incremental addition of power, to various degrees, as a resource. Class III tags use batteries only to power the logic portion of the circuit. Class IV tags are active tags. Class V devices have enough power to activate other tags. Note the implicit statement here: readers themselves will have EPC's™.

Why are the layers drawn this way and not in some other order? This is because they incorporate a practical guess – for example, that the incremental cost of adding a few thousand gates to an active tag (which costs more than US$1.00) is so low that it is unlikely that active tags will ever be made for read-only applications. Class III, IV and V tags will also likely be used in applications where there is a great deal of environmentally generated data; therefore, these protocols will also most likely need to support the broadband modes.

3.2. Principles of Platform Modularity and Default States

We are firmly committed to the view that the protocols for all the species of tags will be modular, and built off the same platform. Three tenets are summarized below:

– Additional **functionalities and modes** will be accessed through extra commands on top of the existing RFID protocol. For example, say a tag has additional memory. The commands to read and write to this memory will be additions to the existing Class 0 and Class I protocols.[1] Similarly, if a tag can operate in a broadband mode, it can be commanded to go to that mode from the Class 0 or Class I mode, one of which the tag must support.

[1] Obviously, the number of functions will proliferate, and we would rather not waste precious command space. In a separate paper, we will discuss an embedded command approach which will enable an almost limitless set of commands.

All commands which access higher functionalities or modes must be acknowledged for success (`an <ack>`) or an error code (`a <nack>`). This will help with the discovery process, which we describe later.

– All tags that are Auto-ID Center compliant will operate in the basic Class 0 or Class I mode by **default**. In other words, the tags will operate in narrow band, with the simplest protocol in the default mode. There will be a **global default call method** which returns all tags to the default mode. This may be as simple as a pause in power or modulation (or both). For example, a broadband Class IV tag will need to operate as a Class 0 or Class I tag by default. It may take on broadband capabilities after it has been discovered and commanded to do so.

– Upon failure of an enhanced function or mode, the tag must be designed to fail safely to a lower class. This makes tags **failsafe**. For example, when a semi-passive tag loses battery power, it must act as a passive tag. When an active tag loses power, it must become a semi-passive tag.[2] This makes failure graceful, and is an advantage of the modular approach.

This strategy establishes a common baseline default mode of operation at start of the transaction, and all tags can be returned to this mode at will. Our approach permits efficient **discovery**. A reader which is unaware of the tags in its neighborhood is incapable of harnessing their functionality. Discovering the tags is a very important step in transacting any business with the tags, and the common default baseline ensures quick discovery of the EPC's™ of the tags. Looking up a database which captures the additional capabilities of the tags in the population then permits the reader to transact business with the tag.

For example, consider a temperature logging semi-passive sensor-tag attached to a package of beef. When the tag appears on a shelf with a number of other passive RFID tags, the reader on the shelf reads the EPC's of its population, and sends the information to the host software, which we call the Savant. The Savant has a process which looks for temperature-sensor tags. When the process detects the temperature-sensor tag, it requests the reader to execute certain commands on that particular EPC tag: first, it asks that particular tag to go into broadband mode, then it commands the tag to download the temperature log. The reader then forwards the log to the software authority which can process the temperature log and look for anomalies. If in fact an anomaly had occurred, the Savant in charge of the shelf sounds an alarm or alerts the clerk responsible for that aisle. This approach demonstrates the importance of a common baseline mode, and the power of a unified discovery mechanism in a heterogeneous population of tags.

If we didn't use a default mode, the reader would have to continuously look for other modes, wasting bandwidth and exacerbating the reader collision problem. With the default mode, the search is unified and efficient.

The concepts of this section can be summarized with a simple statement: a technical document describing the protocol of a higher functionality tag can be reduced to a document describing the Cass 0 or Class I tag by simply deleting text.

3.3. On Frequencies

Our discussion is independent of frequencies. Today, the Auto-ID Center has a Class I protocol for the HF band (primarily 13.56 MHz) and Class 0 and Class I protocols for the UHF band (868-870 MHz in Europe, 902-928 MHz ISM in the US, and 2.45 GHz worldwide). In future, the Center may develop protocols for LF. The hierarchy of classes we have described applies across all frequencies. The protocols themselves are designed to provide the same functionality regardless of frequency. For example, all the protocols permit the reader to pre-select a portion of the EPC™ and search only for tags which match that section. The logical layer should be able to access the physical layer implementations in different frequencies in the same way. Developing a constant abstraction is an important step in the quest for modularity.

[2] Semi-passive tags will need a low-power monitoring circuit to turn the battery on. Active tags must support this circuit even though they may not use it.

4. ONGOING GENERATIONS OF STANDARDS

Successful standards are generated by processes which embrace change. The inexorable progress of technology and human knowledge renders static or slow thinking obsolete. A model for success is the Internet, which has gracefully melded itself to adopt new waves of technology and applications. RFID, which is today in its infancy, will need to be similarly dynamic. Changes in regulation, improvements in technology and the emergence of new applications will all challenge the RFID community to develop newer, better, more efficient solutions, and these solutions will not yield commercial fruit if the standard does not reflect them.

4.1. On the "Generation Gap"

The protocols under discussion today are all in the first version (we will refer to a version in this document as a "generation" to avoid confusion with versioning of the EPC™) of the Auto-ID System. However, the Center is committed to developing Generation 2 at the very earliest. New generations may seem onerous, but adherence to the following principles will ensure that these changes are transparent to the user.

– The Mixed Population Guarantee. A mixed population of tags from different generations of the Auto-ID family will not confuse the reader, or any of the tags.

  This can be achieved in two ways: **backward compatibility** or **complete backward incoherence.** Backward compatibility is the more difficult of the two approaches. A new protocol will be seen as backward compatible with an old protocol if the signaling schemes and the commands are similar, but the changes in the new generation are ignored in all implementations of chips for the old generation of the tag. In other words, the new generation of the protocol is a subtle, non-interfering modification of the older protocol – though rarely will a subtle change yield the benefits to justify a new protocol at all.

  More likely is the possibility of an entirely different protocol in the new version – so different in fact, that there is utter incompatibility to the point that there is no possibility of confusing tags or readers. In other words, when a reader communicates with tags of the new version, tags of the old version hear gibberish. The possibility of interference is completely eliminated by complete incoherence.

  Note that all Version 1 protocols are non-interfering with each other, and the Guarantee of Mixed Population is satisfied implicitly.

– Future Proofing with Agile Readers. The largest capital investment in the adoption of the Auto-ID System will likely be in readers. This investment will need to be protected. We feel that there is no better way to future-proof readers than to make sure that they are a flexible platform with network access and software defined functionality.

  Software Defined Radio is not a new concept. The military has considered and used Software Radio for many years, and it has recently generated excitement in the cellular phone industry, and in civilian and military telecommunications in general, as a guardian against obsolescence [Mitola 95, Bose 96]. Neil Gershenfeld of the Media Lab at MIT first envisioned the use of an RFID reader based on software defined radio, which he called the "agile reader." At the Auto-ID Center, Kevin Ashton then contracted ThingMagic Corporation, now a Center sponsor, to develop a reference design to show that an agile reader, in addition to guarding the capital investment, could also be made to be very inexpensive. That reference design is described in [Reynolds 02].

> Readers like the reference agile reader provide three classes of benefits. First, the guts of a software defined radio reader are **frequency agile**. In other words, they can be used to read multiple frequencies from the same platform. Second, they are **protocol agile**. In other words, the protocol that can be read by the reader can easily and quickly be updated in the reader. Third, they are **functionality agile**. This is an important attribute; the fact that the reference reader has the Linux operating system and a standard network interface, for example, makes it possible to run a small, simple web server on it. Other software implementations are also possible, and can be updated on demand. This turns the reader into a network appliance, and lets end-users morph its functionality over time, as needs, applications and standards evolve. In principle, the agile reader is also easier to configure, monitor and maintain remotely.

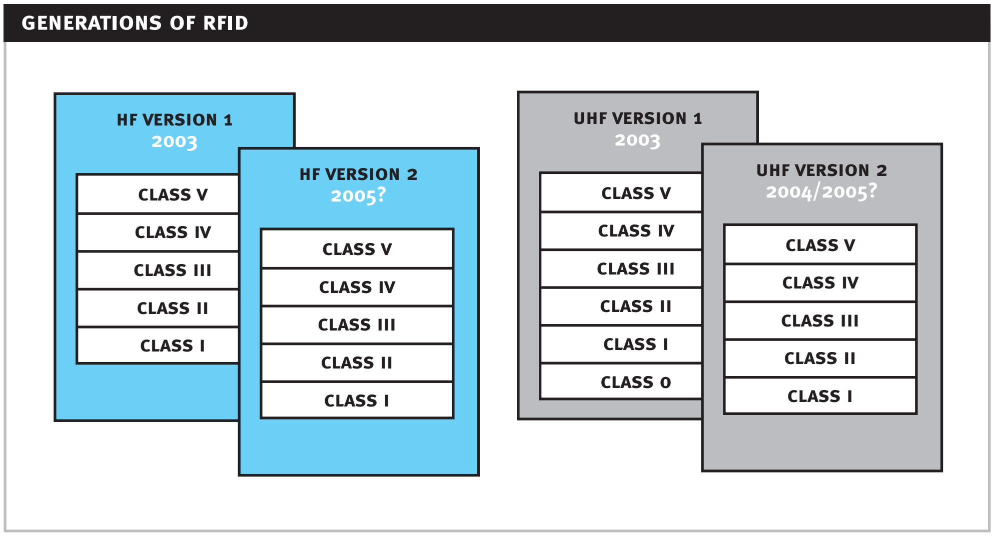
Modularity endows the agile-reader approach with especial benefits. The fact that tags of all classes will be built on top of the Class I protocol will ensure that the proliferation of functionality is not burdened with an accompanying explosion of standards or of software. Instead, the software (or firmware) that is necessary to run a specific sensor-tag, say, will be only a small incremental module in addition to the software (or firmware) necessary to communicate with the Class I tag of the same frequency. There will be only two significant pieces of software (or firmware) on the reader for today's Auto-ID System: the HF software and the UHF software. These pieces will be reused for all mutations of tags in each band.

4.2. A Roadmap

As the RFID industry graduates from generation to generation, the basic Class structure built on top of the Class I or Class 0 tags will also inherit the basic protocols. In other words, a Class III tag in Generation 2 will be modular to the Class I protocol of Generation 2. The evolution of RFID systems from generation to generation will not lead to a proliferation of protocols. Instead, for the transition period, as one generation is being phased out and as a new generation is being phased in, the reader will simply need to carry both protocols. In all cases, based on the Mixed Population Guarantee, the user is assured that no complications will occur.

The roadmap for HF and UHF protocols is shown in Figure 2 below, and is fairly self-evident. In UHF, we envision the disappearance of separate Class 0 and Class I protocols, and the emergence of a single Class I Generation 2 protocol. We believe this will be a natural outcome of our experience with the current Class 0 and Class I protocols, and of the emerging lessons from field trials and pilots.

Figure 2



5. CONCLUSIONS

The modularity of the Auto-ID System admits a layered standard in which many functions can co-exist. This avoids the one-size-fits-all approach that has plagued the industry in the past and permits a range of options from the low-cost tag to sensor-tags. The Auto-ID Center believes in embracing advances in technology and thinking, and renewing the standard to keep it vital and cutting-edge. Together, these principles will keep the Auto-ID Center and its system state-of-the art. The system, modules and roadmap we propose can be transparent to users if they invest in agile readers which can absorb new protocols seamlessly.

6. REFERENCES

[Engels 03]
1. **D.W. Engels, R.L. Rivest, S.E. Sarma & S.A. Weis, "Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems".**
Accepted for publication to the First International Conference on Security in Pervasive Computing (SPC 2003), March 12–14, 2003.

[Stewart 01]
2. **R. Stewart, "Auto-ID Working Paper".**
2001.

[Mitola 95]
3. **J. Mitola, "The Software Radio Architecture".**
IEEE Communications Magazine, 33(5):26-38, May 1995.

[Bose 99]
4. **V. Bose, "Design and Implementation of Software Radios Using a General Purpose Processor".**
Ph.D. Thesis, Department of Electrical Engineering and Computer Science, June 1999.

[Reynolds 02]
5. **M. Reynolds, J. Richards, S. Pathare, H. Tsai, Y. Maguire, R. Post, R. Pappu & B. Schoner, "Multi-Band, Low-Cost EPC™ Tag Reader".**
Auto-ID WH-012, June, 2002.

Designed by Foxner. www.foxner.com