

Expert Group on the Internet of Things (IoT-EG)

Sub-Group on Identification

Identification of the Issue / Challenge

The Internet of Things envisions billions of devices of our daily lives interconnected in such a way that applications that were not possible in isolation emerge from the combination of capabilities and the cooperation of such ‘smart objects’. In such a vast network of interconnected objects, the issue of identification of a particular object and its addressing mechanism play a crucial role that affects all other aspects of the system, including its overall architecture, privacy characteristics, governance, etc., whose study is part of the work performed in other sub-groups

The Group identified three areas whose technological developments will be relevant to reflections on public policy in Europe: object identifiers, network addresses and resolution and discovery functions.

Multiple vs. Unique Identifiers:

Work to define identifiers for various classes of objects (sensors, actuators, tags etc.) is currently ongoing in industry and is also part of studies in standardisation bodies both public and private (e. g. ITU, IETF). It seems unclear at this stage whether usage will drive developments towards a globally unique scheme or several distinct ID spaces with varying degree of interoperability. These alternative scenarios have very different public policy implications. If a globally unique scheme were to emerge, like the Internet IP addresses, questions similar to the critical Internet resources debate would likely need to be addressed. A higher degree of flexibility in allowing different governance models would probably be suited to the alternative scenario of multiple identifier spaces. In the latter case, the degree of interoperability would be part of the debate.

With the standardization of protocols such as IPv6 and, more concretely for embedded devices, 6LoWPAN, it has been shown that it is feasible in practice to provide a unique identifier (ID) to arbitrarily small devices in an efficient way. However the issues of providing non-colliding unique addresses in a global scheme requires an infrastructure in place that supports highly dynamic devices that appear and disappear from the network at any time, move between different local and/or private networks and have the flexibility to either identify their user uniquely or hide his/her identity, thus preserving privacy as needed. Additionally, this infrastructure has to be able to retrieve information about an object as required for the interoperability and cooperation with other objects and networks and allow for the interchange of meta-data and data without compromising security and/or privacy.

Identifiers vs. Network Addresses:

Another issue that needs to be taken into account is the conceptual difference between the ID of an object and its network address (or addresses). In the most general case, the ID of an object and its address(es) are distinct and serve different purposes. The former provides a unique handle to the object itself whereas the latter might change depending on the physical location of the object, its logical membership in one or several networks, or the current role of the object. In cases where the ID of an object and its address are different, the ID is normally structured by different identification schemes. The Electronic Product Code (EPC) is one of the well known object identification schemes which could uniquely identify objects associated

with an RFID tag. Another identification scheme called ubiquitous Code (uCode) introduced by the uID center in Japan represents a different coding system that has support within Japan and Asia. In addition to these identification schemes a range of other, industry or application specific coding standards might be identified.

Similarly the addressing schemes could be different. Objects currently connected to the Internet use the global IP addressing scheme (IPv4 or IPv6). Some other objects may not use global IP addressing, using private addressing instead. Even in the case of objects that use a private network, these might be still connected to the Internet, which will often be used to bridge one private network to another. In this case a border gateway which uses global IP addressing is needed to transport the data from one private network to another. For this particular case, the addressing scheme could be heterogeneous and a device could potentially have the capability to “speak” different addressing schemes to operate as part of different networks.

Challenges:

1. It will be nearly impossible to have one global identification scheme for all the objects in the world, since most industries have been using their own proprietary coding standards (ie. identification schemes) for a long time. For this reason, it is highly unlikely that they will move to a different object identification system. Another difficulty is that it will require consideration of a wide variety of object identification schemes to achieve a global object identification schema.

Resolution and Discovery:

Finally, the issue of object discovery and resolution is a very important one that affects the choice of identification and addressing scheme. This is particularly true if the system is global and the issues of scalability, interoperability, etc. are crucial.

Domain Name System (DNS) [RFC1034], the name resolution service on the Internet was basically conceived for translating "human-friendly" computer host names on a TCP/IP network into their corresponding "machine-friendly" IP addresses. Besides translating host names to IP addresses, at present DNS is used for instance by Mail transfer agents to find out where to deliver mail for a particular address, a general mechanism for locating services in a domain using SRV records, resolution of identifiers that do not have traditional host components through DNS using NAPTR resource records etc.

There are overlay resolution mechanisms services such as Object Naming Service (ONS) and Object Directory Service (ODS) which use the DNS to resolve the object identifiers (their respective identification schemes) to its related digital information.

Case Study: Resolution Using ONS:

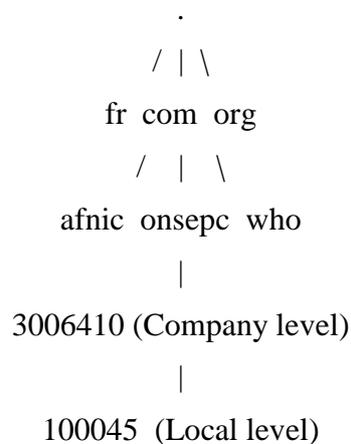
This case study briefly explains the case of an RFID associated object using the Internet to obtain digital information related to it. Let us suppose the object is a "pepper mint". A passive RFID tag is attached to it. The tag constitutes a microchip and an antenna which usually contains a unique code. There are different coding schemas for structuring this unique code and in this example we use the EPC coding standard. The unique code is just a reference for the object and the information about the object is stored in different servers across the Internet.

To resolve the object (pepper mint) information in the Internet, ONS is used. Since ONS uses DNS, the EPC has to be converted into a Fully Qualified Domain Name (FQDN) to query the ONS infrastructure.

To read the code from the tag associated with the pepper mint, an RFID reader is needed. The RFID reader connected to a computer reads the code in Binary form. A conversion tool is used to convert the binary data into Hexadecimal format as follows "3074B77F2861B34000000001". The obtained data is then converted into an URI of the following format "urn:epc:id:sgtin:3006410.100045.1".

Using the "narrowest to widest" structure for host name, the URI is rewritten as 100045.3006410.sgtin.id.onsepc.com. The serial number ("1") is ignored since the ONS resolution stops at the object reference level. The serial number is used to distinguish between different objects of the same class (for example to distinguish one pepper mint from another). The string "onsepc.com" is added at the end of the converted string to point to the ONS root. As per the current ONS standards there is only one root "onsepc.com".

The delegation at the DNS could be as follows:



By using the normal DNS resolving procedures the information about the pepper mint could be obtained from the server at the local level.

Challenges:

1. Object discovery is, for example, a trivial task in small networks of several hundreds or thousands of devices, where a broadcasting mechanism can be used easily to look for a particular device. However, using the same scheme in a network of millions of devices would impose significant performance problem on any network design.
2. Regarding object resolution, the assumption is the existence of a naming scheme that allows the user or another device in the system to find the object (or its meta-data/data) it is looking for.
3. For the most part, the Internet uses a hierarchical naming scheme, the URL, that is not suitable for a highly mobile environment such as the one envisioned as part of the Internet of Things. URLs are not network location transparent, that is, they disclose information about the network where a resource (a computer) is located. This makes it extremely expensive to move between networks, where the "name of the computer" has to change when moving from one network to the next.
4. In a dynamic environment such as the IoT, where new objects and services keep on evolving and network topologies keep on changing, automated discovery mechanisms are needed for overall communication management. The discovery mechanism should enable interaction between objects or identify suitable services for objects which are not pre-configured or hard coded as far as the objects addresses or service end points are concerned.

To provide a context for these issues and highlight some essential requirements of IoT Identification, consider the following example:

A societal medical system provides monitoring data on common human body parameters (for example temperature, blood pressure, pulse rate etc). This data is fed from non-intrusive body mounted sensors to IoT data repositories as part of a population-wide medical support service. Data mining applications process the data searching for signs of potential medical situations that can benefit from early or preventive treatment. An outbreak of influenza in North East England is signalled by a localized increase in body temperature within parameter ranges associated with influenza. The data mining application alerts the regional medical services of a potential influenza epidemic and early actions to screen and immunize the exposed population are put in place.

In this example, location is essential information for the necessary medical application described. This in turn requires that geographical location be determined from the Identification associated with the data. Other demographic information may need to be similarly derivable from IoT Identifiers for other applications, especially those applications not foreseen at the time the data was collected.

This also highlights that access needs to be equitable not only to the IoT infrastructure but also to IoT data since the true value of the IoT will be in its data not in its connectivity.

Objectives

The main objective is, therefore, to propose requirements and to identify policies that promote the desired identification, addressing and naming scheme. This scheme should be able to deal with the issues and challenges described in the previous section, while still fulfilling the following non-functional requirements:

Transparency / Network independence:

The requirement of transparency is an important one that will allow a device to own an identifier that is independent of the network it is currently connected to. If, as in the current identification scheme used for computer networks, the name contains information that might be used to infer the location of the device, it is violating this requirement.

From a technical point of view, an identifier is just a tag or a name assigned to an object under specific conditions. Given that we are dealing with Internet of Things and with objects that people use in their everyday lives, it is required to have the possibility to assign a tag to an object that does not change if the device is put into another context, another network or is used by another user. In technical terms, this is called network transparency and has already been identified as one of the key issues needed by a distributed system to increase the level of availability and scalability of its components.

Scalability up to billions of devices:

Related to the previous requirement, it is crucial to have an identification, addressing and naming scheme that supports billions of devices. This can be achieved in different ways but one of the best understood methods is the use of naming hierarchies to limit the applicability (the scope) of a name to a specific domain, location, or context. The desired identification, addressing and naming scheme has to deal with potential naming conflicts by treating additional information, such as the context of the device, to disambiguate the mapping between a name and a particular device.

Given the nature of the Internet of Things, it is possible that many of the devices that will be connected to the Internet will not be mobile and might be constrained to a very limited space.

This is a characteristic that will help with the scalability issue since they will not have to deal with this potential problem. On the other hand, objects that are capable of moving, changing context or can be carried by people or robots will need to interact with their changing environment. Therefore, complexity should be implemented and supported only on the devices that really require it.

Efficient for very simple devices:

Following the reasoning behind the previous requirement, the proposed schemes should be efficient enough that they can be implemented in devices with very limited capabilities, such as sensor nodes, RFID tags, etc. These devices have computational capabilities that are either non-existent (as is the case of a passive RFID tag), or are as limited as a small embedded computer with several bytes of RAM and a few kilobytes of ROM. If the addressing scheme requires a computational complexity that surpasses what can be done reasonably in small devices, the Internet of Things will remain a dream and will never penetrate in our daily lives as it has been predicted by visionaries like the late Mark Weiser, as he described his notion of the disappearing computer back in the early 90s.

Taking the notion of having small devices as providers and consumers of data, many sources and targets of IoT data will be low power constrained devices that either impose very long service lifetimes (minimum maintenance cost) or will be carried on individual people, animals or vehicles. Implications of the identification schemes should, therefore, promote efficient implementations.

Preservation of privacy:

The collection and use of data from embedded devices that people control and interact with on a daily basis has many implications with respect to privacy. It is clear that all of the privacy requirements cannot be taken into consideration only at the identification schemes, but support for privacy at different levels is a crucial feature. The main issue with a privacy-preserving identification scheme has to do with the guarantee that the scheme is able to protect the possibility to link information from a number of devices with a specific person or group of people.

As an example, consider data about the location, uptime and usage of a particular device. This, in itself does not constitute a problem with respect to privacy, unless it is possible to link this device to a particular person. It is this mapping that makes the data valuable and worthy of being protected. While in most cases the preservation of privacy should be granted, there are particular scenarios where the same device in the previous example, should provide information about the location and well-being of the person it is connected to. This is particularly important for life-or-death situations, where an “emergency call” from the device to the local authorities could make all the difference. Although this might seem an unlikely case, the notion of ambient assisted living and many of its applications are based on the voluntary sacrifice of certain private information for longer, more comfortable stays at home as opposed to a care centre.

Therefore, the identification schemes should support and encourage the preservation of privacy of the users linked to their devices, but it should also allow for the implementation of applications and scenarios where the users opt for a voluntary reduction of their level of privacy to improve their well-being.

Flexible device authentication:

The concept of authentication with respect to the IoT has quite different implications from the well-known, user-based authentication common in the existing Internet. IoT-devices may

communicate with each other without any human interaction or even (real time) human control. For the integrity and safety of the overall system it is essential that only authorized devices take part in this process. “Authorized,” in this context, also includes “genuineness.” i.e., the device is authentic rather than counterfeited, and it has not been modified or replaced in an illegal or unauthorized way. Unique identifiers alone cannot deliver this capability. “Uniqueness” in this case has to be accompanied by copy protection and protection against forgery of the identifier. This protection may be based on cryptographic mechanisms or intrinsic physical properties such as so-called “physical unclonable functions” implemented at the hardware level.

If authentication is used to link a particular device or devices to a particular person or group of individuals, then “accountability” becomes an important issue: Devices may incur service charges or raise liability issues. Devices linked to the same individual may have a special relationship to each other not shared with other devices, or, on the other hand, individuals may wish to keep certain such devices unaware of this relationship (unlinkability).

Additionally, daily use of devices also implies their trade (selling, buying, etc.) which means that the association of a particular device to an individual has to be flexible, transferable and revocable. Taking into account the sheer numbers of items traded every day, it is important that authentication mechanisms can support this flexibility.

Interoperability:

In order to allow for the seamless integration of devices from different manufacturers, it is crucial that devices use schemes that allow for interoperability. This can be achieved by the use of standards or by making sure that gateways are available to translate the identification of one device into another. This is especially important for the case where a global identification scheme is proposed that should operate at all levels. While the definition of standards is a long and sometimes tedious task, the use of technologies currently available on the Internet should be encouraged so as to allow for a smooth transition when the standards are ready. Organizations like the IETF or W3C need to embrace the Internet of Things requirements and extend their current documents to work with them.

Reliability:

Together with interoperability, reliability is another crucial requirement for the adoption of the Internet of Things. If the proposed identification scheme is not reliable, or does not always provide the right information (for example regarding authentication), people will move away from the technology and still rely on more trusted ways of interacting with other users. This is also related to security. Devices that are easily compromised to either not work as they are supposed to, or to break prematurely, will have a negative effect on the adoption of the IoT. Schemes that hinder such compromises and that make it very hard, if not computationally impossible, to tamper with devices in order to forge another identity will be necessary.

Flexibility and Extensibility:

Given the sheer numbers of devices envisioned for the Internet of Things, it is crucial to have an identification scheme that is flexible and extensible. Flexibility is necessary in order to allow for changes in the mapping of devices to location, to individuals, etc., whereas extensibility is required for the addition of devices to the “pool” owned by someone. These two characteristics are tied very strongly with reliability and interoperability since the Internet of Things will have to make it possible for anybody to go to a shop, buy an IoT device and plug it in his living room without having to deal with the technical problems associated with this activation.

This requirement goes well beyond the capabilities needed by the identification scheme and ties with the architecture proposed for IoT, but authentication and identification are the first steps a user needs to perform before he/she can start using a new device. On the other hand, the association of a particular device to a user, location or context, for example, needs to be performed in a flexible way. Otherwise, mobility capabilities will be reduced greatly.

Support for mobility:

Mobility plays a crucial role in our everyday lives, with people moving constantly between different spaces, such as home and work, gym, etc. In this regard, IoT devices will have to support the notion of mobility by allowing the association with the user without having to change their ID, even if the underlying network addressing scheme changes. Even in the case where mobility is normally performed locally, the notion of locality is superseded by the fact that Internet-connected objects can communicate with devices located thousands of kilometres away. Therefore, the main requirement for the identification scheme is to be able to support both static and mobile devices as well as devices that virtually move from one network to the other by communicating and associating themselves with other networks.

As an example, think of the fact that, independently of location, current smart phone users still connect to their offices, read e-mail and answer questions from colleagues virtually from any location. From the point of view of a device, if the user is connected to the office network, the underlying infrastructure views the device as a device local to the network.

Given the breadth and importance of the aforementioned requirements, it seems clear that an identification scheme that is able to cope with the needs of users in all settings of life, as it is the purpose of IoT, will have to deal with a number of very complex issues that, not only impose new challenges to the devices themselves, but also to the infrastructure that supports their interoperation. A Reference Architecture for identification, addressing and naming will have to interoperate with the rest of the system and provide the level of support required.

Policy Options

Like the Internet, the IoT is a complex phenomenon whose development will result from the interaction between two complementary areas: technological/market developments and political discourse. Each of these in turn, will have to account for a wide variety of stakeholders both from the private and public sector.

Considering the very early stage of technological developments related to the IoT, it will be important to promote a wide consultation of stakeholders to identify if/where an EU policy initiative could enhance the social benefits of these technologies while promoting innovation and European competitiveness. Any such initiative will also have to take into account the inherent global nature of many of the IoT applications, when choosing the appropriate level of intervention, global vs. regional.

Challenge / Issue description:

As discussed in the preceding sections, IoT identification needs to satisfy requirements of object identification, object resolution and network addressing. The structure of identifiers will most likely develop differently in different markets with IoT infrastructure standards providing the platform for unification of different identifier formats and standard means of resolving specific objects.

At the same time, object identifiers need to satisfy the requirements of independence of location, scalability to large populations of devices, support of mobility and provision of necessary characteristics to support Data Privacy laws.

European policy options to promote convergence towards efficient IoT platforms where required to support national, regional and European applications and services must be balanced against allowing market innovation and supporting global initiatives.

Policy Options:

Do nothing / Self regulation:

Identification structures and protocols suitable for the IoT will evolve from existing technologies and markets. Market driven collaborative standards are already developed where needed to support large markets and this will continue as those markets spill over into the IoT.

Soft law:

In general, it would be inadvisable to allow divergence at a member state level of key underlying technologies required for large or global scale markets. Although certain national applications such as medical care may differ in different member states, long term policy should encourage standardisation at the largest scale.

Co-regulation:

EU policy providing guidance towards long term converged standards for IoT identification and object resolution may prove all that is required to allow existing actors to provide the necessary market driving agreements. The preferential adoption of key IoT protocols will have significant effect on directing supporting technologies.

Binding law:

It would seem appropriate to consider binding law for aspects of IoT identification that are related to Data Privacy.

Potential Impact

In general, the impact on society, economy and the environment depends on the type of policy applied in the first place. For example, the issue of a non-binding recommendation could have a very limited impact whereas the implementation of a reference architecture for identification, especially if this is open source and freely available, could potentially shape the way identification in the Internet of Things is performed in the future.

On Policies and Impact:

1. Branch specific regulatory requirements are very much distinct; e.g. imagine industries like Power/Energy, Health Care Products, Public Traffic,... . They all have impacts on the use case specific business case, which cannot be easily harmonized.
2. Policies have to follow a branch specific “per-use-case-approach”
3. There will be no “one-fits-for-all” solution
4. Policy may be:
 - a. There will exist branch specific solutions also in the future, governed by already existing regulatory processes;
 - b. Opportunities and requirements for interoperability and convergence on technical and non technical level have to be identified and solutions have to be supported
 - c. Clarify the application issues of binding laws like Data Privacy Laws and work for international harmonization where necessary and appropriate (=interoperability issue)

d.

Impact - Promotion of IPv6 Addressing for IoT Identifier and Object Resolution:

If Policy options are adopted which lead to uniform application of preferred Identifier technology, it is important that the selected technology, or application of that technology, promotes all aspects of IoT Policy. For example, it is highly likely that IPv6 addressing will play a significant role in IoT Identification. Current practice in IPv6 allows the construction of unique identifiers by concatenating a 'network prefix' and a 'MAC Address', each 64-bits in length, to provide a 128-bit IPv6 Address. Clearly the use of an identifiable MAC Address in a Globally Unique Identifier would imply traceability of the device and, potentially, linkage to individual owners thus violating the requirements of Privacy.

Do nothing / Self regulation:

Large markets, especially those which are cost sensitive, may promote IoT Identification via IPv6 addressing without concern for potential impact on other important IoT policy matters.

Soft law:

Member states with strong national policies, especially in Privacy and Security, may demand technical differences in IoT deployments thus fragmenting European markets and limiting benefits of scale in IoT devices.

Co-regulation:

EU guidance on policy will ensure that important cross-policy issues are taken into account in IoT deployment

Binding law:

Certain specific cross policy issues would be avoided but potentially at a cost of limiting innovation if Binding Law is enacted too early. When IoT deployments are widely integrated in European Citizen daily life, Binding Law may be the correct option, but not before.