life.augmented

eSIM: hassle-free, secure global connectivity for mobile-enabled products

Michele SCARLATELLA

System Technology & Systems Architecture

Agostino VANORE

SIM/eSIM Product Marketing Manager

Cellular connectivity is a key enabler of connected devices. Leading to a greater diversity of smart objects, it paves the way for new market opportunities. In addition to being required to be connected at all times and everywhere with worldwide coverage and interoperability, mobile-connected devices must also ensure asset security whether for consumer IoT applications, industrial machine-to-machine (M2M) communications or automotive-grade solutions for smart mobility.

A company wishing to sell and distribute mobile-enabled products for a worldwide market must make sure that once the device is turned on it can connect to one of the available local mobile networks. The process should not be a hassle for the end user. These products can be static, like a smart meter, or nomadic, like a container that is loaded and unloaded from a truck or train. For the latter, the product must be able to switch to the different networks along its journey. Before the embedded SIM, adding mobile connectivity to a product implied many operational and commercial hurdles that prevented widespread adoption.

WHAT REMOTE SIM PROVISIONING MEANS

The embedded SIM with Remote SIM Provisioning (RSP eSIM) specification was defined by the GSMA to simplify the process of adding and replacing a mobile subscription to an IoT device. Mobile is expected to be the most common long range technology for the IoT, with over 40% market share by 2026 [Ref.1].

The eSIM technology will have a significant impact for a faster adoption. Mobile connectivity has many advantages compared to other types. It is virtually available everywhere in world, no need to build an infrastructure network, as it is already in place, and Mobile Network Operators (MNO) are quickly building specialized commercial and support offers for this type of service with NB-IoT and LTE CAT-M networks.

Another non-negligible advantage of mobile connectivity is that it is very easy to establish a connection. There is no need to pair devices as with Bluetooth or configure Wi-Fi credentials; all of this is managed securely and transparently by the eSIM.

The eSIM eliminates these barriers and makes global deployment easy, manageable, and transparent for the end user.

While specifications and products are now available, the eSIM remains still an intricate subject for many. To use it effectively, engineers must grasp certain concepts before designing flexible mobile connectivity into their product.

In this paper we try to explain what you need to know to develop a robust mobile-enabled product and get it to market as soon as possible.

THE STARTING POINT: WHAT IS A SIM

In order to understand how an eSIM works and what it can do, here is an overview of what a basic SIM is and some of its key features.

Invented at the end of the last century based on a set of standards issued by ETSI (European Telecommunications Standards Institute), the Subscriber Identity Module, or SIM, lets users independently select their handset and Mobile Network Operator (MNO). Users were no longer forced to buy the handset offered by the network operators. This boosted the handset market, spurring innovation and creativity by the OEMs. The introduction of the SIM simplified the logistics of handset distribution and allowed companies to sell them on the open market. MNOs and newly created mobile virtual network operators (MVNO) started to buy SIMs from companies specialized in smart card security, ST being one of them.

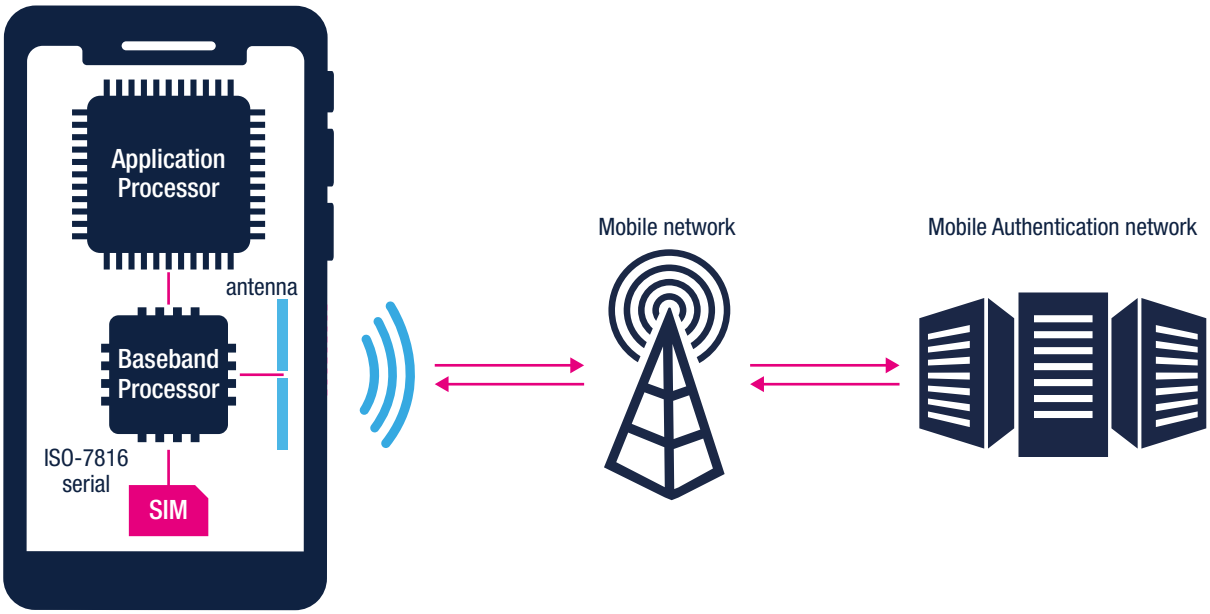


Figure 1: Traditional handset network architecture

More than 20 years later, the handset-network architecture (Figure 1) is still basically true. A removable SIM is connected to the modem baseband processor via a half-duplex serial link. In modern smartphones, the modem is integrated inside the mobile system-on-chip. The SIM connects to networks via the modem, without having to use the main processor, using a challenge-response authentication system with pre-shared secret keys that are known to the network.

In addition to authentication, the SIM also performs functions related to network operations and call control.

- Temporary Session keys are set up to encrypt the current call, preserving privacy.
- Parameters of the operating network are hosted inside the SIM for the handset to use: a list of preferred and barred networks (depending on roaming agreements) and the subscriber identity known as IMSI (International Mobile Subscriber Identity).
- Call control function: allowed and prohibited numbers, dial code modification like adding an international prefix or number substitution, for example to reach the closest customer center based on location.

The SIM can also manage a contact list, but this has become of little use as handset-based contact lists are much more sophisticated and the preferred choice of consumers.

INTEROPERABILITY FOR MOBILE DEVICES

Based on ISO 7816 specifications, and further refined by ETSI, the data exchange protocol between the SIM and modem uses a poll-response protocol that sends “Application Data Packet Units” (APDU). The protocol can multiplex up to four logical channels, with channel 0 reserved for the mobile part, while the other channels can carry additional “services”.

HOW COMMUNICATIONS ARE SECURED USING SIMs

Original 2G cellular networks lacked network authentication and made it possible for a bogus device to impersonate a base-station, relaying the data and spoofing the conversation.

Starting with 3G, an improved internal SIM architecture was introduced based on the Universal Integrated Circuit Card [15] (UICC). The network application running on top of the UICC is the Universal SIM (USIM) which now authenticates the mobile network using a set of secret encryption keys. We will continue to use the word SIM in this paper as this is commonly used despite being not formally correct. Let’s be aware of the difference.

The secret keys held in the SIM are preloaded during manufacturing at a certified secure location. The keys and unblocking PINs are provided by the MNO during the procurement process, and it is not uncommon for MNOs to simply instruct their suppliers to generate keys and PIN codes and let them do the job. Modern SIMs also offer cryptographic accelerators for more efficient security functions.

Typically, SIMs have DES/AES engines for symmetric cryptography, RSA and ECC for asymmetric, and a Hash engine to authenticate message integrity. SIMs also offer other protection functions with sets of sensors and features to defeat “side-channel” attacks, improve memory protection, and for real-time memory bus encryption/decryption.

WHY IS 2G TECHNOLOGY STILL IS USE?

With almost worldwide coverage guaranteed, 2G technology lets networks provide basic services such as text messages, picture messages, and MMS (multimedia messages). 2G is sometimes preferred for M2M applications, as the lower operating frequency allows for better connectivity from inside buildings. This is acceptable as long as you are aware of the security risks implied by the simple 2G authentication procedure. You have been warned.

HOW CAN I BE SURE THAT MY SUPPLIER'S PERSONALIZATION SERVICES ARE TRULY SECURE?

The GSMA has defined a Security Accreditation Scheme (GSMA-SAS) that defines what facilities should comply with, and the procedure to implement, in order to build a secure personalization service facility. Most SIM suppliers have undergone such accreditation processes, mandatory by most MNOs in order to become a qualified supplier. GSMA keeps an up-to-date publicly available list of GSMA-SAS authorized sites on their website.

Interoperability and inherent security: the benefits of the JavaCard SIM

Eventually MNOs wanted to deploy services exploiting the SIM functions. This clashed with fragmentation of the different SIM brands: different MCU core, memory size and layout, and incompatible software APIs made the task of developing a common service across all SIMs nearly impossible. This led to the development of a simplified Java version optimized to run on a smart card, the JavaCard, to make software deployable to all JavaCard-based SIMs, regardless of their make.

Both a language and a runtime environment with a rich set of security APIs, JavaCard lets MNOs deploy multiple applications on a single SIM. In addition to being able to update applications post-issuance.

The SIM Security Domains

A SIM can host several applications maintained by different entities not necessarily trusting each other, pursuing different business scopes, for example mobile payments and network localization.

The concept of Security Domain (SD) was introduced and standardized by the Global Platform. An SD defines how a set of files and code can be managed securely. Designed for the secure storage of access keys, it obeys a set of protocols that ensure a secure connection (using the secret keys) and can fully manage an application under the SD itself.

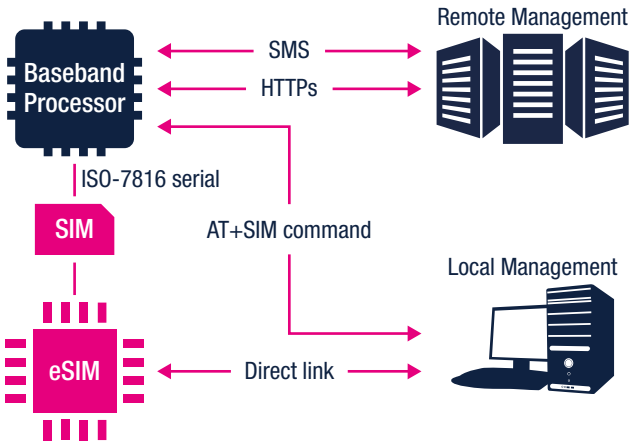
An SD manager, can enable/disable an SD, create sub-Security Domains, and can even entrust some of its functions to another SD manager (delegation).

The specifications related to Security Domains are issued by the GlobalPlatform [4] who extended other security products including banking cards, mobile Secure Elements, and electronic documents (passports, ID, health cards) as well as consumer and IoT applications.

SIM Remote management

Along with the deployment of JavaCard-based applications and GlobalPlatform SDs came the need to manage remotely the SIM. A set of standards to perform over-the-air (OTA) remote updates was defined by ETSI. Now, one could remotely manage files, download applets, and install/activate new features. OTA messages go directly from air to SIM, passing thru the baseband, not easily accessible from the application processor; all data is encrypted and authenticated, making SIM-based OTA a fairly secure process.

The set of SD parameters, keys, files, and applications for network connectivity is known as an Operational (MNO) profile. A typical Operational profile can occupy from 10 to 100 Kbytes of memory depending on network technology and services present in the profile itself.



OPERATIONAL (MNO) PROFILE

A Profile containing one or more Network Access Applications and associated Network Access Credentials and Operator's (e.g. SIM Tool Kit) applications and 3rd party applications.

As defined by the GSM Association

THE NEED FOR AN EMBEDDED SIM

There were two main drivers that led to the deployment of the embedded SIM, a solderable SIM in a standard IC package. In the consumer market, the miniaturization of components in smartphones pushed the SIM to smaller form factors. But even the nanoSIM, known as the 3rd Form Factor in the standards (3FF) failed to meet design requirements. In addition to occupying a significant area of the PCB (see for example Figure 2 that shows an iPhone 6S motherboard tear down), the SIM slot also introduces a design complexity for the antenna and waterproofing.

For M2M devices where reliability is the primary concern, the SIM connector is a point of failure due to vibrations, moisture, and corrosion. The automotive industry was the first vertical application to push the most for a soldered solution.

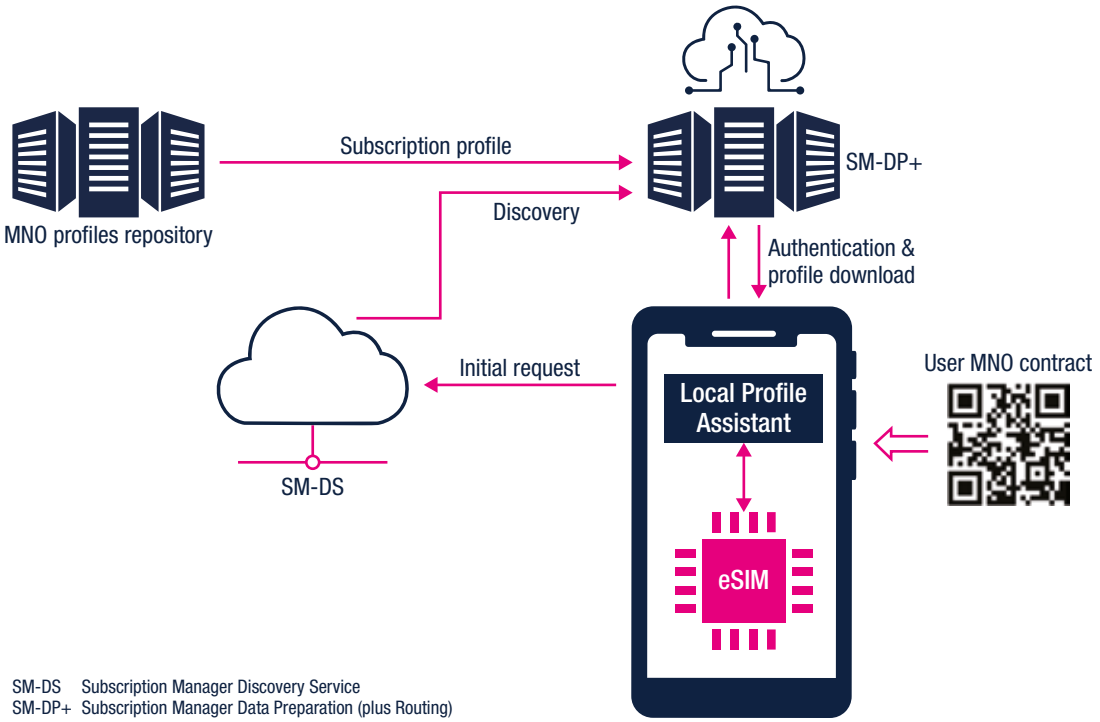
In 2015, GSMA officially launched the embedded SIM initiative to allow a soldered SIM to be capable of changing the connectivity provider just with an OTA update.

The complexities behind this feature are many. MNOs were used to carefully selecting their SIM products and suppliers. With embedded SIMs, they had to rely on OEMs for the selection. The loading of the initial subscription was also a very sensitive topic as devices (like smartphones or cars) are manufactured in only a few manufacturing facilities, then distributed worldwide with a very limited idea in which country the product will be sold.

After several years of work, GSMA issued the Remote SIM Provisioning specifications, first for the M2M then for the consumer eSIM market. One of the main differences between these two schemes is who manages the connectivity.

In Consumer applications, the user selects and manages the carrier. For M2M, this is done by the service entity running the M2M system, which in a vast majority of cases is not the end user.

Figure 2. iPhone 6S motherboard with SIM slot



SM-DS Subscription Manager Discovery Service
SM-DP+ Subscription Manager Data Preparation (plus Routing)

Figure 3: Consumer eSIM architecture

THE CONSUMER eSIM

A user that wants to activate a new subscription will sign a commercial contract with the MNO. The MNO will provide a link where to activate the new subscription. The link can be in a variety of forms such as a QR code, a URL link in an SMS, or even by manually entering data in a smartphone app. The Local Profile Assistant (LPA), a program acting as an interface between the user app and the eSIM, coordinates the process with the goal to download authenticated and encrypted profiles into the eSIM. The LPA has an external part in the smartphone, and internal part in the eSIM. The LPA acts as a local user agent for the management of the downloaded profiles.

After having gathered the necessary data from the eSIM, the LPA will contact the Subscription Manager – Discovery Service (SM-DS). The SM-DS functions as a go-between the eSIM and Subscription Manager Data Preparation Plus (SM-DP+). It allows the SM-DP+ to locate the eSIM without needing to ascertain which network the eSIM is connected to. Simply put, it acts as a secure notice board where the SM-DP+ can send a notification. The SM-DP+ then tells the LPA that a profile can be downloaded. Once the internal LPA receives this notification, it verifies and downloads the profile.

The Operational profile is retrieved from an MNO store, encrypted, signed and downloaded. Once in the eSIM, the profile is authenticated, instantiated and activated. This simplified flow does not cover many operational details and cases that are not reported here. The GSMA specifications listed at the end of the paper have all the gory details.

An OEM will have to choose the eSIM during the product design and provide all the needed instructions to the eSIM supplier, including security credentials and digital certificates. Depending on the business arrangement, one or more subscriptions can be preloaded at the manufacturing stage, for example in pre-branded smartphones.

At the end of manufacturing and delivery, the eSIM manufacturer will send the list of eSIMs sold to the SM-DP+ server indicated by the OEM.

THE M2M eSIM

The connectivity model for M2M applications is different than for the consumer market, and the business model is also different. In M2M applications, the end user simply benefits from the network. In some cases, there is no end-user at all, like for an asset tracking device on containers. Here, the shipping company manages the relation with the MNO, while the user is just personnel loading/unloading the containers.

In M2M applications, a different model for the subscription provision process is employed where the provision process works in Push mode.

To explain better, let's take a mobile-connected security camera as an example. The key players (see Figure 4) in the provisioning process are:

- eSIM manufacturer
- Security camera manufacturer buying an eSIM on the open market
- Security service provider, managing a fleet of security cameras
- MNO, with whom the Security Provider establishes a commercial connectivity agreement

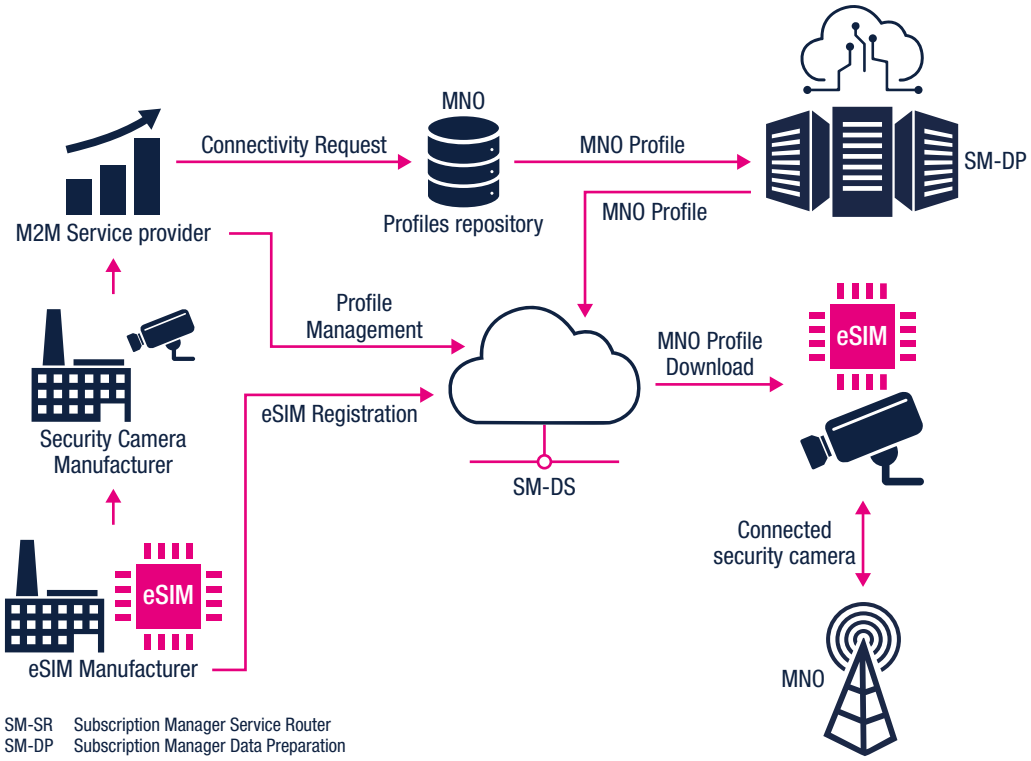


Figure 4: M2M-eSIM system architecture

The security camera manufacturer designs mobile-connected products distributed to security service companies worldwide. It does not know which MNO the security service company will select. Other points in question: how does the SIM manufacturer indicate where the eSIM should connect when the device first powers up? And how does the eSIM connect initially without a subscription?

- When the eSIM is manufactured, it is pre-loaded with a bootstrap Provisioning Subscription that lets it connect to the mobile network only to activate a regular, fully functional subscription. Normally in "bootstrap mode", the device cannot send or receive operational calls or data, but it is not uncommon for a bootstrap operator to also provide full working connectivity. This depends on the specific commercial arrangements.
- Based on the bootstrap network provider, the eSIM manufacturer loads into the eSIM the identity of the SM-SR (Subscription Management-Service Routing) it should connect to in order to get the operational subscription information (Operational profile).
- Once the device is activated, the eSIM will connect to the SM-SR informing the service operator of its location.

It is up to the security service company to buy connectivity from an MNO of its choice and provide it with the identities of the eSIMs in its devices. Once the commercial relation is in place, the SM-DP can find out where the device is located by querying the SM-SR, thus the subscription can be delivered and activated. Only then, the M2M device becomes fully operational.

PROVISIONING SUBSCRIPTION

A special purpose contract, with its associated Provisioning Profile, that enables a machine-to-machine device to access a mobile network only for the purpose of management of Operational Profiles on the eUICC.

As defined by the GSM Association

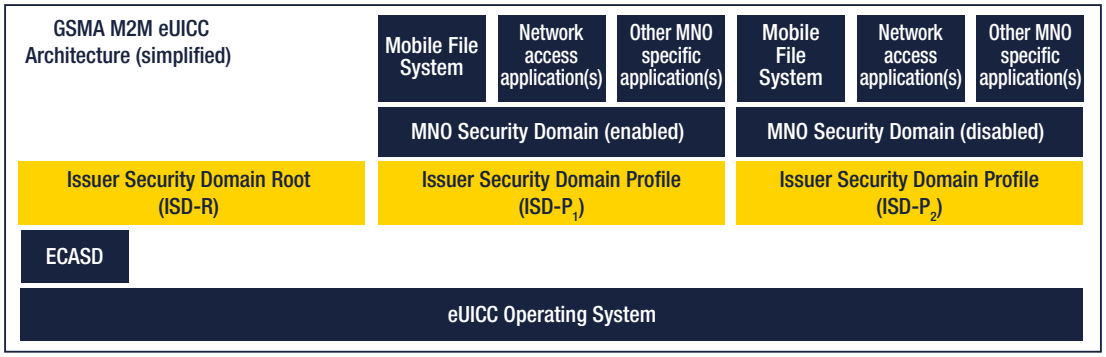


Figure 5: M2M eSIM architecture

To fulfill the requirements, a M2M-eSIM internal architecture was defined, as shown in Figure 5. On bottom sits the underlaying hardware and Operation System software of the eUICC. A number of security domains are present. Starting from the left:

- ECASD: (eUICC Certificate Authority SD)
  - This a root for all subsequent SDs. It contains several immutable security data, such as the root private key and associated certificate, the Certificate Issuer (CI) certificate, and eSIM manufacturer keyset for key/certificate renewals.
  - The ECASD is created at manufacturing time, and cannot be disabled or removed.
- ISD-R (Issuer Security Domain – Root):
  - This is also created at manufacturing time, and like the above, cannot be disabled or removed. The ISD-R contains to which SM-SR the eSIM should connect to. It also offers services that allow the security keys for wrapping/un-wrapping of the profile data at download.
  - It can create the additional SDs needed to host the MNO profiles (ISD-P).
- ISD-P is the umbrella SD to host and manage all MNO-SD:
  - It can create/destroy/enable/disable the MNO-SD,
  - receive and install the profile,
  - has all the OTA security to be remotely managed.

While more than one Operational profile can be hosted in an eSIM, only one can be active at any given time.

One may ask how you can handle an application that is not related to a specific MNO? Today, this is not possible as applications can only be hosted inside an operator profile, and it is lost after an MNO swap, and must be re-loaded with the new profile. To circumvent this problem, GSMA is working on a “Secure Application for Mobile” (SAM) framework that will let applications survive MNO swaps. Agnostic of specific contracts, the specification should be available sometime in 2021. Stay tuned.

Will this eSIM work in my region?

All bootstrap MVNOs claim they have “worldwide” coverage, which they usually do. However, they may not have commercial relations (and consequent SP-SR/SP-DP links) with all the MNOs in the world. Before committing to a specific eSIM bootstrap network, it is worth checking if it can provide the connectivity you are likely to need.

ACCESSING THE SIM FROM AN EMBEDDED APPLICATION

The SIM can also be accessed from the application processor, using the AT+SIM or similar “tunneling” commands to the modem.

The data inside the SIM is organized in tree-like directories, and files. The EFname = Elementary File and DFname = Directory File are identified by hex digits. The set of a directory, files and sub-directories is referred to as an “Application”.

An initial 2-byte AID was extended to a longer format, from 6 to 16 bytes. The first 5 bytes indicate the entity providing the application (Registered application provider ID), the 6th byte and onwards are the unique number for the application itself. The Application concept was also extended besides files and directories, and it includes a program (applet) to handle specific commands, plus read/writes to files.

Starting with 3G, a more formal definition of the internal SIM architecture was introduced. The base of the hardware and software platform is the Universal Integrated Circuit Card (UICC). The network application running on top of the UICC is the Universal SIM (USIM).

With 3G, SIM configuration became much more complex. The DF containing the 3G network configuration consists of hundreds of parameters contained in tens of files.

In 3G, the (U)SIM also authenticates the network. The lack of network authentication in 2G by the SIM allows a bogus device to impersonate a base-station, relaying the data and spoofing the conversation. 2G is sometimes preferred for M2M applications, as the lower operating frequency allows for better connectivity from inside buildings. This is acceptable as long as you are aware of the security risks implied by the simple 2G authentication procedure.

HOW TO GET STARTED WITH YOUR DESIGN AND REDUCE TIME-TO-MARKET

While there are many products available on the market, ST offers a tailored, diversified connectivity portfolio with a wide range of SIMs and embedded SIMs (eSIM) compatible with commercial-, industrial- and automotive-grade applications.

A good place to start is with our **ST4SIM SIM and eSIM cellular connectivity solutions**.

Part of a complete ecosystem with trusted partners specialized in connectivity and subscription management platforms, the ST4SIM portfolio is a scalable offer that makes sure that your product is always connected and that you are always in control.

Compliant with Remote SIM Provisioning (RSP), our ST4SIM devices have an ISO 7816 serial interface, with an SPI link (WLCSP package only). All SIM functions are sent via the serial port to the modem.

The SPI is an interesting option, where a dedicated JavaCard application can interact directly with the host microcontroller or microprocessor. Thanks to this feature, the ST4SIM can double-up as an embedded Secure Element for functions including secure bootstrap, authentication token, digital signature and validation of data, or just as a very secure cryptographic engine. In addition to saving the cost of adding a dedicated Secure Element, it can also be managed remotely, a feature not normally found on most Secure Elements.

ST4SIM devices are available in a range of packaging options, and can also be bought in a removable, hardened plastic package (traditional SIM form) if required.

Parameter	SIM / eSIM	Application	Generic Description	Hardware	Certification / Qualification	Packages
ST4SIM-100S	Basic SIM	IoT	SoC Card OS	ST32H480	-	Card plug-in
ST4SIM-110S	Crypto SIM/eSIM		SoC Card OS with advanced crypto services	ST33G1M2	CC EAL5+	Card plug-in
ST4SIM-200S	GSMA eSIM		SoC Card OS compliant with GSMA SGP.02	ST33G1M2	CC EAL5+	Card plug-in
ST4SIM-100M	Basic SIM/eSIM	M2M Industrial	SoC Card OS	ST32F512M	Industrial Grade (JEDEC 47)	Card plug-in, MFF2
ST4SIM-110M	Crypto SIM/eSIM		SoC Card OS with advanced crypto services	ST33G1M2M	CC EAL5+, Industrial Grade (JEDEC 47)	Card plug-in, MFF2, WLCSP
ST4SIM-200M	GSMA eSIM		SoC Card OS compliant with GSMA SGP.02	ST33G1M2M	CC EAL5+, Industrial Grade (JEDEC 47)	Card plug-in, MFF2, WLCSP
ST4SIM-110A	Crypto SIM/eSIM	Automotive	SoC Card OS with advanced crypto services			MFF2, TSSOP
ST4SIM-200A	GSMA eSIM		SoC Card OS with advanced crypto services	ST33G1M2A0	CC EAL5+, AEC-Q100 Grade 2	

Table 1. eSIM selection guide

From removable SIMs to GSMA-certified eSIMs, ST4SIM is a flexible and scalable offer which can be integrated in various environments. High-quality and reliable, ST4SIM devices are ready-to-use solutions available in commercial-, industrial-, and automotive-grade versions to satisfy the different market requirements.

Worldwide coverage and interoperability are enabled thanks to connectivity solutions from approved ST partners that ensure the provisioning process. ST has four partners to select from for the bootstrap connectivity: Truphone, Arkessa, Soracom, and Pelion. Pick the most convenient for your needs.

If your business model already has a clear view of the MNO that will provide connectivity for final use, it is also possible to pre-load the final profile during production. ST has production facilities that are GSMA-SAS certified for this kind of work.

ST PARTNER PROGRAM

The ST Partner Program helps companies easily identify trusted partners able to supply expertise for their critical design projects; reducing their development efforts and accelerating time to market.

life.augmented Partner Program

CONCLUSION

Thanks to the Remote SIM Provisioning (RSP) feature, the M2M-eSIM is a significant improvement to enabling cellular connectivity on IoT devices, avoiding a life-long commitment to a specific MNO, and facilitating the distribution logistics in different part of the world.

The sourcing of the eSIM can follow the typical electronic components sourcing process, from a distributor or direct from ST.

While the RSP eSIM is considered a standard electronic component, OEMs must pay attention to certain peculiar facets when selecting the final product. Bootstrap connectivity, MNO profile, and SM-SR are important to consider, as they will affect the device connectivity, once deployed in the field.

We hope that with this paper we have given you better insight into eSIM technology and guidance for selecting the best product for your application, as well as showing the benefits of using a product from the ST4SIM family.

RESOURCES

Internet of Everything Market Tracker, ABIresearch, Nov. 2020

Introduction to Smart Cards by ETSI

Smart Card Handbook - Wolfgang Rankl, Wolfgang Effing, Kenneth Cox

ISO/IEC 7816-4: "Integrated circuit cards, Part 4: Organization, security and commands for Interchange"

ETSI TS 131 102 UMTS; LTE; Characteristics of the Universal Subscriber Identity Module (USIM) application

ETSI TS 102 221 Smart Cards; UICC-Terminal interface; Physical and logical characteristics

ETSI TS 102 226 Smart Cards; Remote APDU structure for UICC based applications

GSMA eSIM Consumer Specification

GSMA eSIM M2M Specification

GSMA Security Accreditation Scheme

GlobalPlatform Card Specification

A personal History of the Java Card

Order code: **BRWPESIM0621**

For more information on ST products and solutions, visit www.st.com

life.augmented