# Chapter 6: Establishing (secure) communication using TCP/IP Sockets

## Objective

At the end of Chapter 6 will will understand how to use the WICED-SDK to send and receive data using TCP/IP sockets. You will also understand the fundamentals of symmetric and asymmetric encryption and how that is used to provide security to your IoT device.

## Time: 4 Hours

## Fundamentals

### Sockets – Fundamentals of TCP Communication

For Applications, i.e. a web browser, to communicate via the TCP transport layer they need to open a **Socket**. A Socket, or more properly a TCP Socket, is simply a reliable, ordered pipe between two devices on the internet. To open a socket you need to specify the IP Address and [Port](https://en.wikipedia.org/wiki/Port_(computer_networking)) Number (just an unsigned 16-bit integer) on the Server that you are trying to talk to. On the Server there is a program running that listens on that Port for bytes to come through. Sockets are uniquely identified by two tuples (source IP/ source port) an (destination IP/ destination port) e.g. 192.168.15.8/3287 + 184.27.235.114/80. This one reason why there can be multiple open connections to a webserver running on port 80. The local (or ephemeral port) is allocated by the TCP and new ports are allocated on the initiator (client) for each connection to the receiver (server).

There are a bunch of [standard ports](https://en.wikipedia.org/wiki/List_of_TCP_and_UDP_port_numbers) (which you might recognize) for Applications including:

* HTTP 80
* SMTP 25
* DNS 53
* POP 110
* MQTT 1883

These are typically referred to as “Well Known Ports” (WKP) and are managed and maintained by the IETF Internet Assigned Numbers Authority (IANA); IANA ensures that no two applications designed for the Internet use the same port (whether for UDP or TCP)

WICED easily supports TCP sockets (wiced\_tcp\_create\_socket) and you could create your own protocol to talk between your IoT device and a server. For instance, we could define the WACEP (aka Wiced Academy Example Protocol) as a method to transfer the value of one of 256 registers (on our IoT device) in the form of a 32-bit integer. Morever, lets assume that we have assigned “device ids” to all of the IoT devices in our system as a 2 byte unsigned integer (aka 0-🡪65535). Then, we define the protocol to run on port 40507. Lastly we agree that the Server will respond with “1” if the message is accepted and “0” if the message fails. So, the communication would look like this:

* Client looks up server IP address
* Client opens a TCP connection on the server to port 40507
* The Client send 2-byte little endian format of its “device id”
* The Client send 1 byte indicating which register it is reporting on
* The Client sends 4-byte little endian format of the Integer value
* The Server responds with 1 byte indicating success or failure
* The Client closes the connection

There are a number of problems with using “raw” sockets including, most notably security.

Sockets are available in WICED and will enable you to build your own custom protocol. However, in general developers are mostly using one of the standard Application Protocols (HTTP, MQTT etc) which are discussed later.

### Symmetric and Asymmetric Encryption: A Foundation

When you see “HTTPS” in your browser window, the “S” stands for Secure. The reason it is called Secure is that it uses an encrypted channel for all communication. But how can that be? How do you get a secure channel going? And what does it mean to have a secure channel? What is secure? This could be a very complicated topic as establishing a fundamental mathematical understanding of encryption requires competence in advanced mathematics that is far beyond most everybody on the face of this planet. It is also beyond what there is room to type in this manual. It is also far beyond what I have the ability to explain. But, don’t despair. The practical aspects of getting this going are actually pretty simple.

All encryption does the same thing. It takes un-encrypted data, combines it with a key, and runs it through an encryption algorithm to produce encrypted data. The original data is called plain or clear text and the encrypted data is known as “cipher-text”. You then transmit the cipher-text over the network. When the other side receives the data it decrypts the cipher-text by combining it with a key, and running the decrypt algorithm to produce clear-text a.k.a the original data.

There are two types of encryption schemes, symmetric and asymmetric.

[Symmetric](https://en.wikipedia.org/wiki/Symmetric-key_algorithm) means that both sides use the same key. That is, the key that you encrypt with is the same as the key you decrypt with. Examples of this type of encryption include [AES](https://en.wikipedia.org/wiki/Advanced_Encryption_Standard) and [DES](https://en.wikipedia.org/wiki/Data_Encryption_Standard). Symmetric encryption is preferred because it is very fast and secure. Unfortunately, both sides need to know the key before you can use it (i.e. the encrypt key is exactly the same as the decrypt key). The problem is, if you have never talked before how do you get both sides to know the key? The other problem with symmetric key cryptography is that once the key is lost or compromised, the system will be compromised as well. Changing the key at regular intervals is one solution to the problem. Another technique to help protect privacy is to use one way algorithms to create the keys.

[Asymmetric](https://en.wikipedia.org/wiki/Public-key_cryptography), often called Public Key, encryption techniques use two keys that are mathematically related. The keys are often referred to as the “public” and the “private” keys. The private key can be used to decrypt data that the public key encrypted and vice versa. This is super cool because you can give out your public key to everyone, they can encrypt data, then only your private key can be used to decrypt it. What is amazing about Asymmetric encryption is that even knowing the Public key you cant figure out the private key ( one-way function). The problem with this encryption technique is that it is slow and requires large key storage on the device (usually in the OTP or FLASH) to store the public key (192 bytes for PGP).

What now? The most common technique to communicate is:

* You open an unencrypted connection to a server
* You give out your public key to the server
* The server then creates a random symmetric key
* The server then encrypts its newly created random symmetric key using your public key and sends it back to you
* You use your private key to decrypt the symmetric key
* You open a new channel using symmetric key encryption



This scheme is completely effective against eavesdropping. What happens if someone eavesdrops the original public key? That is OK because they won’t have the “client private key” required to decrypt the symmetric key. So, what’s the hitch? What this scheme doesn’t work against is called man-in-the-middle (MIM). An MIM attack works by:

* You open an unencrypted connection to a server [but it really turns out that it is a MIM]
* You send your public key to the MIM
* The MIM opens a channel to the server
* The MIM sends its public key to the server
* The Server encrypts a symmetric key using the MIMs public key and send it back to the MIM
* The MIM decrypts the symmetric key using its private key
* The MIM sends you the symmetric key encrypted with your public key
* You unencrypt the MIM symmetric key using your private key
* Then you open new channel to the MIM using the symmetric key
* The MIM opens up a channel to the server using the symmetric key

Once the MIM is in the middle it can read all of the traffic. You are only vulnerable to this attack if the MIM gets in the middle on the first transaction. After that, things are secure.

However, the MIM can easily happen if someone gets control of an intermediate connection point in the network e.g. WiFi Access Point. There are only two ways to protect against MIM attacks

* Pre Share the key
* Use a [Certificate Authority](https://en.wikipedia.org/wiki/Certificate_authority) (CA)

A CA is a server on the internet that has a huge dictionary of keys. To use a CA, you embed the CAs verified public key in your system (so you can make a secure connection to the CA). Then when you get a key from someone you don’t know, you open a secure connection to the CA and it verifies the key that you have matches the key you were sent.

If the MIM sends you its public key then you check with the CA and find out that the MIM public key does not belong to the server that you are trying to connect to, then you know that you are being subjected to an MIM attack. How do you prevent an MIM when talking to a CA? This is done by building in known valid certificates into your program. This morning when I looked at the certificates on my Mac there were 179 built in, valid certificates.

### [Secure Sockets Layer (SSL) / Transport Layer Security (TLS)](https://en.wikipedia.org/wiki/Transport_Layer_Security)

SSL and TLS are two Application Layer Protocols that handle the key exchange described in the previous section and present an encrypted data pipe to the layer above it - i.e the Web Browser or the WICED device running MQTT. SSL is a fairly heavy (memory and CPU) protocol and has largely been displaced by the lighter weight and newer TLS.

Both of these protocols are generally ascribed to the Application layer but to me it has always felt like it really belongs between the Application and the Transport Layer. TLS is built into WICED and if you give it the keys (from the DCT) when you initialize a connection its operation appears transparent to the layer above it. Several of the application layers protocols that are discussed in this chapter rest on a TLS connection i.e. HTTP🡪TLS🡪TCP🡪IP🡪WiFi Datalink 🡪 WiFI 🡪 Router 🡪 Router🡪Server Ethernet🡪Server Datalink🡪Server IP🡪Server TCP🡪TLS🡪HTTP Server

## Further Reading

[1] RFC1700 – “Assigned Numbers”; Internet Engineering Task Force (IETF) - https://www.ietf.org/rfc/rfc1700.txt

[2] RFC2045 – “Multipurpose Internet Mail Extensions”; Internet Engineering Task Force (IETF) - https://tools.ietf.org/html/rfc2045

[3] IANA Service Name and Port Registry - <http://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml>

[4] RFC2616 – “Hypertext Transfer Protocol (HTTP) “ ; Internet Engineering Task Force (IETF) - <https://tools.ietf.org/html/rfc2616>

[5] RFC7159 – “The Javascript Object Notation (JSON) Data Interchange Format”; Internet Engineering Task Force (IETF) - <https://tools.ietf.org/html/rfc7159>

[6] MQTT - <http://mqtt.org/>

[7] RFC7959 – “The Constrained Application Protocol (CoAP)” ; Internet Engineering Task Force (IETF) - <https://tools.ietf.org/html/rfc7252>

[8] AMQP - <http://www.amqp.org/>