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

## Objective

At the end of Chapter 6 you 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 it is used to provide security to your IoT device.

## Time: 2 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) and (destination IP/ destination port) e.g. 192.168.15.8/3287 + 184.27.235.114/80. This is 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 stack 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” and are managed 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).

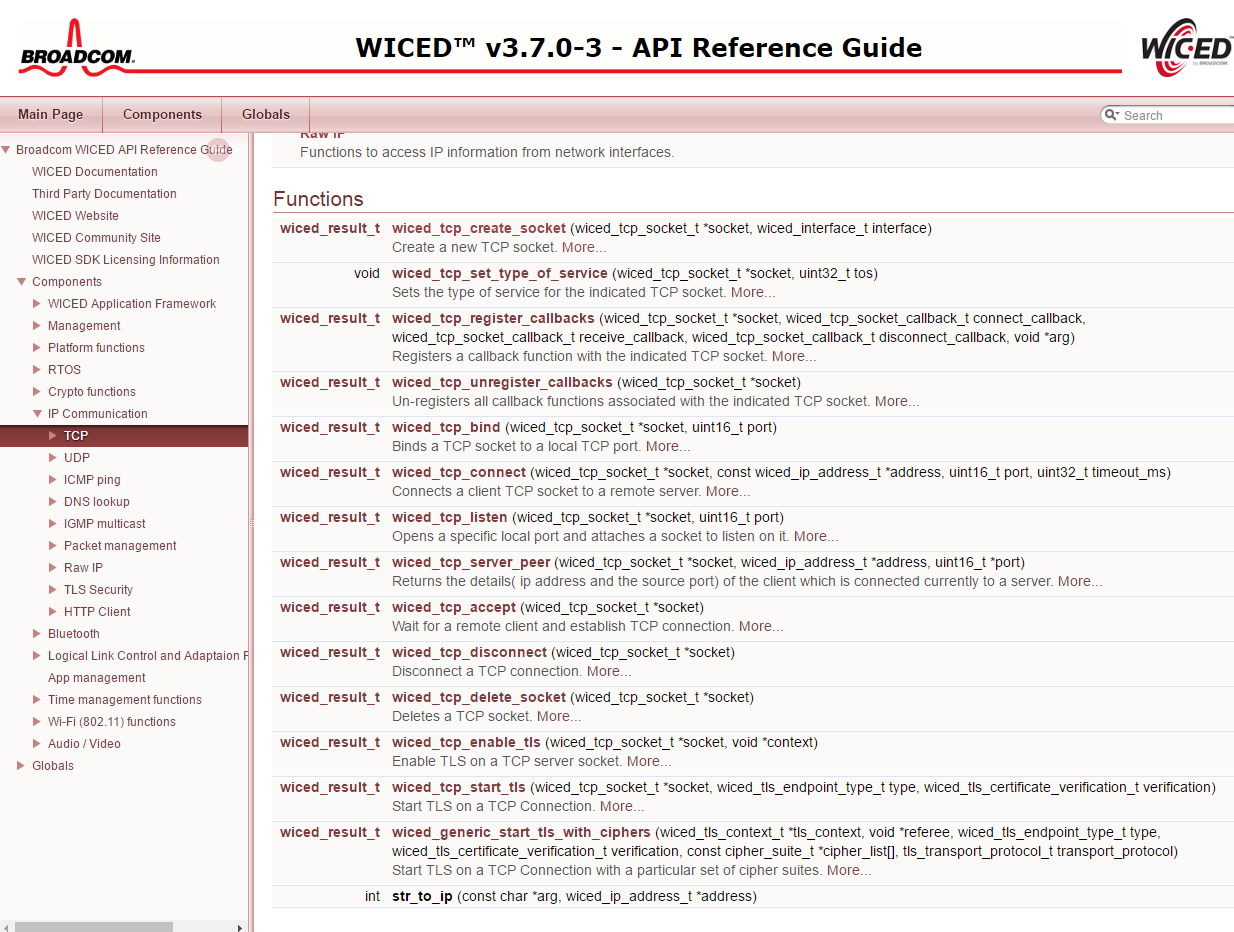
WICED easily supports TCP sockets (wiced\_tcp\_create\_socket) and you can create your own protocol to talk between your IoT device and a server or you can implement a custom protocol as defined by someone else.

To build a custom protocol, for instance, we can define the Wiced Academy Example Protocol (WAEP) as a method to transfer the value of one of 256 registers (on our IoT device) in the form of a 32-bit integer. Moreover, let’s 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 27708 (for the insecure version) and 40508 (for the TLS version). Lastly, we agree that the Server will respond with one byte, a 0x01, if the message is accepted and 0x00 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 27708
* The Client sends 2-bytes little endian format of its “device ID”
* The Client sends 1 byte indicating which register it is reporting on
* The Client sends 4-bytes 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 - most notably security - as the data going over the internet in a TCP socket is not encrypted (which we will talk about in the next section).

Sockets are available in the WICED-SDK 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 in Chapter 7. The WICED documentation on sockets resides in Components🡪IP Communication 🡪 TCP.



### Symmetric and Asymmetric Encryption: A Foundation

Given that we have the problem that TCP/IP sockets are not encrypted, now what? 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 is be a very complicated topic, as establishing a fundamental mathematical understanding of encryption requires competence in advanced mathematics that is far beyond almost everyone. 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, remember, 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 entire system is compromised as well.

[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 using your public key, then only your private key can be used to decrypt it. What is amazing about Asymmetric encryption is that even when you know the Public key you can’t 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 FLASH) to store the public key (e.g. 192 bytes for PGP).

What now? The most common technique to communicate is to use public key encryption to pass a private symmetric key which will then be used for the rest of the communication:

* 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. But, 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 sends 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. a WiFi Access Point. There are only two ways to protect against MIM attacks:

* Pre Share the public key (so you are sure you have the right 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 CA’s 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.

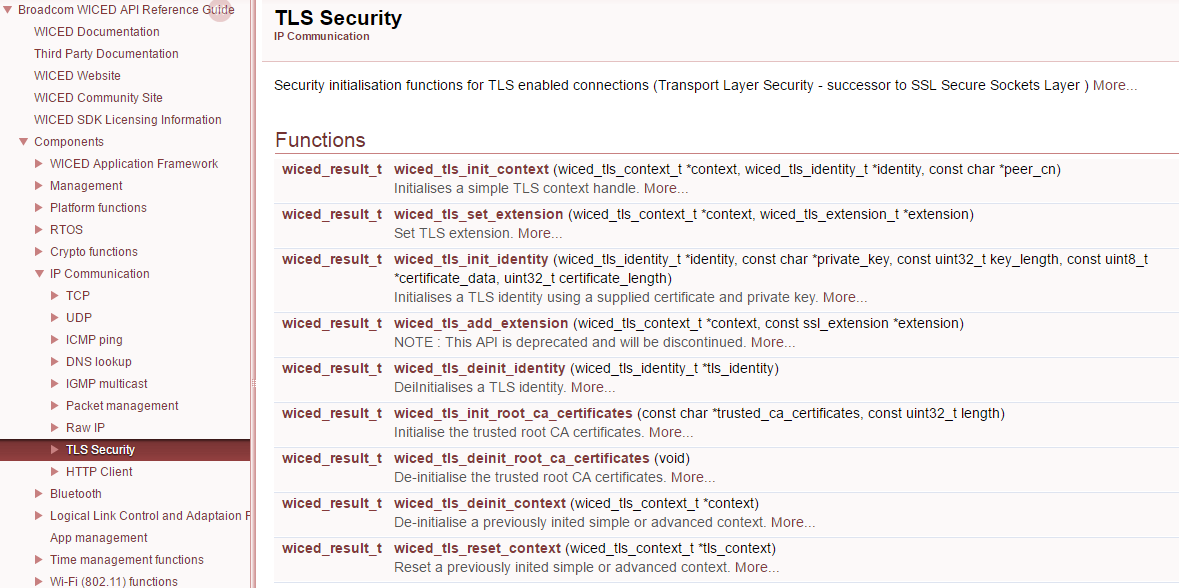
If the MIM sends you its public key then you check with the CA and find out that the MIM public key does not belong to the server that you are trying to connect to, then you know that you are being subjected to an MIM attack. How do you prevent an MIM when talking to a CA? This is done by building in 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)

For the key sharing to work, everyone must agree on a standard way to implement the key exchanges and resulting encryption. That method is SSL and and its successor TLS which 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 layer protocols that are discussed in the next 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

The documentation for TLS resides in Components🡪IP Communication🡪TLS Security.



## Exercise(s)

### 01 Create an IoT Device to write data to a server running WAEP when a button is pressed on your board

We have implemented a server using the WICED-SDK running the insecure version of the WAEP protocol as described above with the following:

* DNS name: waep.wa101.cypress.com
* IP Address: xyz
* Port: 27708.

You application will monitor button presses on the board and will toggle an LED in response to each button press. In addition, your application will connect to the WAEP server and will send the state of the LED each time the button is pressed. For the application:

* The LED characteristic number is 5. That is, the LED is the 6th register of the 256 byte register space.
* The “value” of the LED is 0 for OFF and 1 for ON
* Your device number is the 16-bit checksum of your device’s MAC

The steps the application must perform are:

1. Connect to WiFi
2. Figure out your device number by adding the MAC bytes together in a uint16\_t (effectively a checksum)
3. Open a socket to xyz
   1. Hint: you need to create the socket, bind it to a local port, and connect it to the remote port.
4. Initialize the LED to OFF
5. Setup the GPIO to monitor the button
6. If the button is pressed
   1. Flip the LED state
   2. Create a packet with 7 bytes:
      1. 2-bytes for your device ID
      2. ‘0x05’ (the number of the LED characteristic)
      3. 0x01000000 (for LED On) or 0x00000000 for LED Off
   3. Write your packet to the Socket
7. Go look at the console of the class WAEP server and make sure that your transactions happened.

### 02 Modify your first project to check the return code

### 03 Modify your program to use DNS to get the IP address of the server

### 04 Implement (1-3) using TLS

1. Don’t forget it is a different port #.
2. Get the required certificates from ?

### 05 Implement the server side of the insecure WAEP protocol

### 06 Implement the server side of the secure WAEP protocol

## Further Reading

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

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