SERIAL BUS ARCHITECTURES

The peripheral devices and external buffer that operate at relatively low frequencies communicate with the processor using serial bus. There are two popular serial buses: Serial Peripheral Interface (SPI) and Inter-Integrated Circuit (I2C).

Serial Peripheral Interface (SPI)

Serial Peripheral Interface (SPI) is an interface bus designed by Motorola to send data between microcontrollers and small peripherals such as shift registers, sensors, and SD cards. It uses separate clock and data lines, along with a select line to choose the device.

The Serial Peripheral Interface (SPI) is a synchronous serial communication interface specification used for short-distance communication, primarily in embedded systems. The interface was developed by Motorola in the mid-1980s and has become a de facto standard. Typical applications include Secure Digital cards and liquid crystal displays.

SPI devices communicate in full duplex mode using a master-slave architecture with a single master. The master device originates the frame for reading and writing. Multiple slave-devices are supported through selection with individual slave select (SS), sometimes called chip select (CS), lines.

Sometimes SPI is called a four-wire serial bus, contrasting with three-, two-, and one-wire serial buses. The SPI may be accurately described as a synchronous serial interface,[1] but it is different from the Synchronous Serial Interface (SSI) protocol, which is also a four-wire synchronous serial communication protocol. The SSI protocol employs differential signaling and provides only a single simplex communication channel. SPI is one master and multi slave communication.

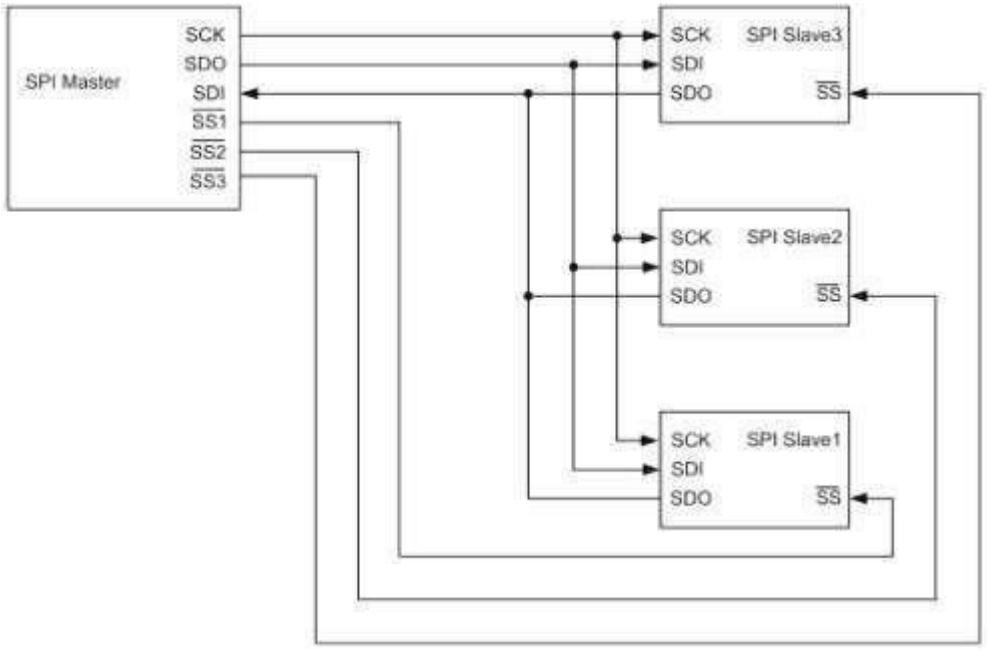

Fig 1: SPI master with three slaves

Source: Miles J. Murdocca and Vincent P. Heuring, —"Computer Architecture and Organization: An Integrated approach"

Modes in SPI

Mode	CPOL	CPHA
0	0	0
1	0	1
2	1	0
3	1	1

- In addition to the standard 4-wire configuration, the SPI interface has been extended to include a variety of IO standards including 3-wire for reduced pin count and dual or quad I/O for higher throughput.
- In 3-wire mode, MOSI and MISO lines are combined to a single bidirectional data line.
- Transactions are half-duplex to allow for bidirectional communication.

Reducing the number of data lines and operating in half-duplex mode also decreases maximum possible throughput; many 3-wire devices have low performance requirements and are instead designed with low pin count in mind.

- Multi I/O variants such as dual I/O and quad I/O add additional data lines to the standard for increased throughput.
- Components that utilize multi I/O modes can rival the read speed of parallel devices while still offering reduced pin counts. This performance increase enables random access and direct program execution from flash memory (execute-in-place).

Inter-Integrated Circuit (I2C)

An inter-integrated circuit (Inter-IC or I2C) is a multi-master serial bus that connects low- speed peripherals to a motherboard, mobile phone, embedded system or other electronic devices.

- Philips Semiconductor created I2C with an intention of communication between chips reside on the same Printed Circuit Board (PCB).
- It is a multi-master, multi-slave protocol.
- It is designed to lessen costs by streamlining massive wiring systems with an easier interface for connecting a central processing unit (CPU) to peripheral chips in a television.
- It had a battery-controlled interface but later utilized an internal bus system.
- It is built on two lines
- SDA (Serial Data) – The line for the master and slave to send and receive data
- SCL (Serial Clock) – The line that carries the clock signal.
- Devices on an I2C bus are always a master or a slave. Master is the device which always initiates a communication and drives the clock line (SCL). Usually a microcontroller or microprocessor acts a master which needs to read

data from or write data to slave peripherals.

- Slave devices are always responds to master and won't initiate any communication by itself. Devices like EEPROM, LCDs, RTCs acts as a slave device. Each slave device will have a unique address such that master can request data from or write data to it.
- The master device uses either a 7-bit or 10-bit address to specify the slave device as its partner of data communication and it supports bi-directional data transfer.

Working of I2C

- The I2C, data is transferred in messages, which are broken up into frames of data. Each message has an address frame that contains the binary address of the slave, and one or more data frames that contain the data being transmitted.
- The message also includes start and stop conditions, read/write bits, and ACK/NACK bits between each data frame.
- The following are the bits in data frames:

1. Start Condition: The SDA line switches from a high voltage level to a low voltage level before the SCL line switches from high to low.
2. Stop Condition: The SDA line switches from a low voltage level to a high voltage level after the SCL line switches from low to high.
3. Address Frame: A 7 or 10 bit sequence unique to each slave that identifies the slave when the master wants to talk to it.
4. Read/Write Bit: A single bit specifying whether the master is sending data to the slave (low voltage level) or requesting data from it (high voltage level).
5. ACK/NACK Bit: Each frame in a message is followed by an acknowledge/no-acknowledge bit. If an address frame or data frame was successfully received, an ACK bit is returned to the sender from the receiving device.

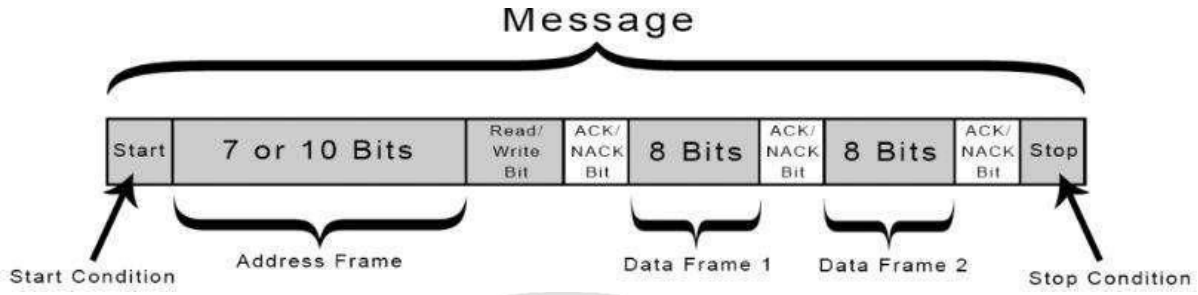

Fig 2: I_2C Message Format

Source: Miles J. Murdocca and Vincent P. Heuring, —"Computer Architecture and Organization: An Integrated approach"

Addressing:

I_2C doesn't have slave select lines like SP), so it needs another way to let the slave know that data is being sent to it, and not another slave. It does this by addressing. The address frame is always the first frame after the start bit in a new message

- The master sends the address of the slave it wants to communicate with to every slave connected to it. Each slave then compares the address sent from the master to its own address.
- If the address matches, it sends a low voltage ACK bit back to the master. If the address

Doesn't match, the slave does nothing and the SDA line remains high.

Read/Write Bit

- The address frame includes a single bit at the end that informs the slave whether the master wants to write data to it or receive data from it. If the master wants to send data to the slave, the read/write bit is a low voltage level. If the master is requesting data from the slave, the bit is a high voltage level.

Data Frame

- After the master detects the ACK bit from the slave, the first data frame is ready to be sent.

- The data frame is always 8 bits long, and sent with the most significant bit first.
- Each data frame is immediately followed by an ACK/NACK bit to verify that the frame has been received successfully.
- The ACK bit must be received by either the master or the slave (depending on who is sending the data) before the next data frame can be sent.
- After all of the data frames have been sent, the master can send a stop condition to the slave to halt the transmission.
- The stop condition is a voltage transition from low to high on the SDA line after a low to high transition on the SCL line, with the SCL line remaining high.

Steps in Data transmission

1. The master sends the start condition to every connected slave by switching the SDA line from a high voltage level to a low voltage level before switching the SCL line from high to low.

   The master sends each slave the 7 or 10 bit address of the slave it wants to communicate with, along with the read/write bit.
2. Each slave compares the address sent from the master to its own address. If the address matches, the slave returns an ACK bit by pulling the SDA line low for one bit. If the address from the master does not match the slave's own address, the slave leaves the SDA line high.
3. The master sends or receives the data frame.
4. After each data frame has been transferred, the receiving device returns another ACK bit to the sender to acknowledge successful receipt of the frame.
5. To stop the data transmission, the master sends a stop condition to the slave by switching SCL high before switching SDA high.

Advantages

- It uses two wires.

- This supports multiple masters and multiple slaves.
- ACK/NACK bit gives confirmation that each frame is transferred successfully.
- Well known and widely used protocol

Disadvantages

- Slower data transfer rate than SPI.
- The size of the data frame is limited to 8 bits
- More complicated hardware needed to implement than SPI.

Mass storage refers to various techniques and devices for storing large amounts of data. Mass storage is distinct from memory, which refers to temporary storage areas within the computer. Unlike main memory, mass storage devices retain data even when the computer is turned off.

MASS STORAGE

The mass storage medium includes:

- solid-state drives (SSD)
- hard drives
- external hard drives
- optical drives
- tape drives
- RAID storage
- USB storage
- flash memory cards

Solid State Devices

- Solid-state devices are electronic devices in which electricity flows through solid semiconductor crystals like silicon, gallium arsenide, and germanium rather than through
  vacuum tubes.
- It do not involve any moving parts or magnetic materials.
- RAM is a solid state device that consists of microchips that store data on non-

moving components, providing for fast retrieval of that data.

- Transistors are the most important solid state devices. The transistors contain two p– n junctions, have three contacts or terminals.
- They require the action of perpendicular electrical fields, their behavior is more difficult to understand than that of diodes.
- The different types of transistors are: bipolar junction transistor (BJT) where the current is amplified, while in the field effect transistor (FET) a voltage controls a current.
- In a solid-state component, the current is confined to solid elements and compounds engineered specifically to switch and amplify it.
- Current flows in two forms: as negatively charged electrons, and as positively charged electron deficiencies called holes.
- In some semiconductors, the current consists mostly of electrons; in other semiconductors, it consists mostly of holes. Both the electron and the hole are called charge carriers.

Hard Drives

- A hard disk drive is a non-volatile memory hardware device that permanently stores and retrieves data on a computer.
- A hard drive is a secondary storage device that consists of one or more platters to which data is written using a magnetic head, all inside of an air-sealed casing.
- Internal hard disks reside in a drive bay, connect to the motherboard using an ATA, SCSI, or SATA cable, and are powered by a connection to the power supply unit.

External Hard Drives

- An external hard drive is a portable storage device that can be attached to a computer through a USB or FireWire connection, or wirelessly.
- External hard drives typically have high storage capacities and are often used

to back up computers or serve as a network drive.

Optical Drives

- An Optical Drive refers to a computer system that allows users to use DVDs, CDs and Blu- ray optical drives.
- The drive contains some lenses that project electromagnetic waves that are responsible for reading and writing data on optical discs.
- An optical disk drive uses a laser to read and write data. A laser in this context means an electromagnetic wave with a very specific wavelength within or near the visible light spectrum.
- An optical drive that works with all types of discs will have two separate lenses: one for CD/DVD and one for Blu-ray.
- An optical drive has a rotational mechanism to spin the disc. Optical drives were originally designed to work at a constant linear velocity (CLV) (i.e.) the disc spins at varying speeds depending on where the laser beam is reading, so the spiral groove of the disc passes by the laser at a constant speed.
- An optical drive also needs a loading mechanism: A tray-loading mechanism, where the disc is placed onto a motorized tray, which moves in and out of the computer case and slot- loading mechanism, where the disc is slid into a slot and motorized rollers are used to move the disc in and out.

Tape disks

- A tape drive is a device that stores computer data on magnetic tape, especially for backup and archiving purposes.
- Tape drives work either by using a traditional helical scan where the recording and playback heads touch the tape, or linear tape technology, where the heads never actually touch the tape.
- Drives can be rewinding, where the device issues a rewind command at the end of a session, or non-rewinding.

- Rewinding devices are most commonly used when a tape is to be unmounted at the end of a session after batch processing of large amounts of data.
- Non-rewinding devices are useful for incremental backups and other applications where new files are added to the end of the previous session's files.
- The different types of tapes are audio, video and data storage tape.

Redundant Array of Inexpensive Disks (RAID) Storage

- RAID is a way of storing the same data in different places on multiple hard disks to protect data in the case of a drive failure.
- RAID works by placing data on multiple disks and allowing input/output (I/O) operations to overlap in a balanced way, improving performance. Because the use of multiple disks increases the mean time between failures (MTBF), storing data redundantly also increases fault tolerance.
- A RAID controller can be used as a level of abstraction between the OS and the physical disks, presenting groups of disks as logical units. Using a RAID controller can improve performance and help protect data in case of a crash.
- Levels in RAID:

1. RAID 0 (Disk striping):

   RAID 0 splits data across any number of disks allowing higher data throughput. An individual file is read from multiple disks giving it access to the speed and capacity of all of them. This RAID level is often referred to as striping and has the benefit of increased performance.

2. RAID 1 (Disk Mirroring):

   RAID 1 writes and reads identical data to pairs of drives. This process is often called data mirroring and it┼ s a primary function is to provide redundancy. If any of the disks in the array fails, the system can still access data from the remaining disk(s).

3. RAID 5 (Striping with parity):

RAID 5 stripes data blocks across multiple disks like RAID 0, however, it also stores parity information (Small amount of data that can accurately describe larger amounts of data) which is used to recover the data in case of disk failure. This level offers both speed (data is accessed from multiple disks) and redundancy as parity data is stored across all of the disks.

4. RAID 6 (Striping with double parity):

Raid 6 is similar to RAID 5, however, it provides increased reliability as it stores an extra parity block. That effectively means that it is possible for two drives to fail at once without breaking the array.

5. RAID 10 (Striping + Mirroring):

RAID 10 combines the mirroring of RAID 1 with the striping of RAID 0. Or in other words, it combines the redundancy of RAID 1 with the increased performance of RAID 0. It is best suitable for environments where both high performance and security is required.

Universal Serial Bus (USB) Devices

- USB is a system for connecting a wide range of peripherals to a computer, including pointing devices, displays, and data storage and communications products.
- The Universal Serial Bus is a network of attachments connected to the host computer.
- These attachments come in two types known as Functions and Hubs.
- Functions are the peripherals such as mice, printers, etc.
- Hubs basically act like a double adapter does on a power-point, converting one socket, called a port, into multiple ports.
- Hubs and functions are collectively called devices.
- When a device is attached to the USB system, it gets assigned a number called its address.

The address is uniquely used by that device while it is connected.

- Each device also contains a number of endpoints, which are a collection of sources and destinations for communications between the host and the device. The combination of the address, endpoint number and direction are what is used by the host and software to determine along which pipe data is travelling.

Flash Drives

- A flash drive stores data using flash memory. Flash memory uses an electrically erasable programmable read-only (EEPROM) format to store and retrieve data.
- Flash drives are non-volatile, which means they do not need a battery backup.
- Most computers come equipped with USB ports, which detect inserted flash drives and install the necessary drivers to make the data retrievable.
- Computer users can store and retrieve data once the operating system has detected a connection to the USB port.
- Flash drives have a USB mass storage device classification, which means they do not require additional drivers.
- The computer┼ s operating system recognizes a block-structured logical unit, which means it can use any file system or block addressing system to read the information on the flash drive.
- A flash drive enters emulation mode, or acts a hard drive, once it has connected to the USB port. This makes it easier to transfer data between the flash drive and the computer.
- Flash memory is known as a solid state storage device, meaning there are no moving parts — everything is electronic instead of mechanical.