DPC QUESTIONS FOR COMPUTER DISCIPLINE

1. Why you call it as portal?

   A **website** is a collection of interlinked web pages typically hosted from a single domain. A website is accessible over the internet or a private network such as Local Area Network (LAN) through an address known as Uniform Resource Locator (URL).

   The URLs organize the web pages into a hierarchical form which help a user to navigate through the pages. A home page is included that usually is the starting point of all navigation artifacts, including links to other pages, an “About Us” page, or a “Contact Us” page.

   Static vs Dynamic Website

   - A static website is composed of web pages the content of which remains constant until a developer wishes to alter.
   - The information on a static website remains the same throughout.
   - A static website may consist of plain text or rich media. However, on visiting a static site, you will see the same content at all times irrespective of the time of visit.
   - On the other hand, a dynamic website updates itself frequently depending on a set of parameters.
   - In other words, a dynamic website’s content is renewed every time a user visits the website.
   - A dynamic web page is created using a wide range of software and languages such as, Java Server Pages (JSP), Active Server Pages (ASP), PHP, Python, Perl, etc.

   A **web portal** is a customized website that immerses information from a wide array of sources in a consistent and uniformed manner. For example, web portals are served in the form of dashboards for company executives and managers. How the content on a portal should be organized and presented depends largely on the requirements of the end users?

   A web portal may be customized based on the restrictions of domain searches. An enterprise portal usually has a consistent design and has the capability to interact with applications and databases.

   Website:

Definition:	A Location on the internet, publicly accessible with a unique URL (web address)

Features:	• No Login required • Anyone can see content • Content does not change for different individuals • Can have interactive features, but does not reference personalized • database
Property Management AppLication:	"Front-facing", corporate site for marketing and web presence Educate site visitors about products, services, and industry information Location for generic content and Login access to private portal sites

Portal:

Definition:	A private location on the internet, accessible with a unique URL (web address) and unique username and password
Features:	• Personal Login is required only portal members can see content • Content is unique to user based on linked account information and group member settings/permissions • Secure access point for personalized information • Communication features with other portal members or groups Dynamic content changes more frequently than typical websites Interactive functionality for portal site members
Property Management Application:	"Back-end" site for designated set of users Educate portal site members about association information and provide association content, such as governing documents in a self-serve environment Provide association residents with access to account balances, transaction histories, and online payments Direct homeowners to one access point to answer common questions, disseminate information, and take advantage of convenient services

2. What happens when you do more normalization?

3. What is difference between authentication and authorization?

   Authentication is the process of verifying who you are. When you log on to a PC with a user name and password you are authenticating.

   Authorization is the process of verifying that you have access to something. Gaining access to a resource (e.g. directory on a hard disk) because the permissions configured on it allow you access is authorization.

4. What is RDBMS?

   A relational database management system (RDBMS) is a database management system (DBMS) that is based on the relational model as invented by E. F. Codd, of IBM's San Jose Research Laboratory. Many popular databases currently in use are based on the relational database model.

   RDBMSs are a common choice for the storage of information in new databases used for financial records, manufacturing and logistical information, personnel data, and other applications since the 1980s. Relational databases have often replaced legacy hierarchical databases and network databases because they are easier to understand and use. However, relational databases have been challenged by object databases, which were introduced in an attempt to address the object-relational impedance mismatch in relational databases, and XML databases.

No.	DBMS	RDBMS
1)	DBMS applications store **data as file**.	RDBMS applications store **data in a tabular form**.
2)	In DBMS, data is generally stored in either a hierarchical form or a navigational form.	In RDBMS, the tables have an identifier called primary key and the data values are stored in the form of tables.
3)	**Normalization is not** present in DBMS.	**Normalization is** present in RDBMS.
4)	DBMS does **not apply any security** with regards to data manipulation.	RDBMS **defines the integrity constraint** for the purpose of ACID (Atomocity, Consistency, Isolation and Durability) property.
5)	DBMS uses file system to store data, so there will be **no relation between the**	in RDBMS, data values are stored in the form of tables, so a**relationship** between these data values

	tables.	will be stored in the form of a table as well.
6)	DBMS has to provide some uniform methods to access the stored information.	RDBMS system supports a tabular structure of the data and a relationship between them to access the stored information.
7)	DBMS **does not support distributed database**.	RDBMS **supports distributed database**.
8)	DBMS is meant to be for small organization and **deal with small data**. it supports **single user**.	RDBMS is designed to **handle large amount of data**. it supports**multiple users**.
9)	Examples of DBMS are file systems, **xml** etc.	Example of RDBMS are **mysql**, **postgre**, **sql server**, **oracle** etc.

5. What is the value stored in a null field?

   Each row has a null bitmap for columns that allow nulls. If the row in that column is null then a bit in the bitmap is 1 else it's 0.

6. How do you take care of remote server security?

7. What is a regression testing and what types of tests are carried out?

   Regression testing is a type of software testing that verifies that software previously developed and tested still performs correctly after it was changed or interfaced with other software. Changes may include software enhancements, patches, configuration changes, etc. During regression testing, new software bugs or regressions may be uncovered. Sometimes a software change impact analysis is performed to determine what areas could be affected by the proposed changes. These areas may include functional and non-functional areas of the system.

   The purpose of regression testing is to ensure that changes such as those mentioned above have not introduced new faults. One of the main reasons for regression testing is to determine whether a change in one part of the software affects other parts of the software.

   Common methods of regression testing include rerunning previously completed tests and checking whether program behavior has changed and whether previously fixed faults have re-

emerged. Regression testing can be performed to test a system efficiently by systematically selecting the appropriate minimum set of tests needed to adequately cover a particular change.

8. What is the customization of software?

   Customisation: bringing existing software features closer to the specification a buyer is interested in.

   Extensibility: creating new software features to match the specification a buyer is interested in.

   Configuration: changing no software features but defining existing content items (such as options, variables, dashboards, etc) to match the specification a buyer is interested in.

   To give an example in terms of whether VAT applies to a product price, extensibility would mean introducing VAT to the software, customization would be to allow a buyer to include two different VAT tiers, and configuration would mean presetting these two VAT values once for easy inclusion each time the product is processed.

9. How do you integrate new modules?

   The life of a software project usually involves the integration of new modules. Methods used to integrate these new modules vary greatly but usually encounter many of the same problems. Typical problems involve management of data, exchange of information and program flow control. To reduce the impact of these problems, software must be designed cognizant of the fact that it may some time be integrated with other modules. A new method for integration, presented here, called dynamic integration, integrates not only module data but module functionality as well. Associativity of data elements persists after data are transferred to new modules. In addition, this integration does not require that modules explicitly call each other. As a direct result, applications are developed by easily assembling several modules.

10. How do you take care of web application security?

    The list of key security requirements from OWASP ASVS (Application Security verification Standards) that an online application need to comply with,

    - Authentication & Authorization
    - Session Management
    - Input Validation
    - Output Encoding
    - Cryptography – provided the application uses any sensitive data (credit cards, SSN etc)
    - Error Handling and Logging
    - Data Protection
    - Communication Security

- HTTP Security and
- Security configuration

The following descriptions refer to the sections in the OWASP Web Guide Project document. These references provide general guidance to the technologies addressed in these sections and the specific recommendations contained therein.

1. Building Secure Web Services and AJAX Topics

   Web Services

   This section deals with the common issues facing web developers as they work to build secure web apps, whether that includes Java, pHp, AJAX or other web languages and/or technologies.

2. Secure Web Application and Secure Coding Topics

   Authentication

   This section deals with authentication issues associated with secure web apps, such as basic/digest authentication, form-based authentication, integrated (SSO) authentication, etc.

   Authorization

   This section addresses authentication issues, ensuring a user has the appropriate privileges to view a resource. Topics such as principle of least privilege, client-side authorization tokens, etc. are addressed here.

   Session Management

   This section addresses topics such as authenticated users having a robust and cryptographically secure association with their session, applications enforcing authorization checks and applications avoiding or preventing common web attacks, such as replay, request forging and man-in-the-middle.

   Data Validation

   This section deals with applications being robust against all forms of input data, whether obtained from the user, infrastructure, external entities or databases.

   Interpreter Injection

   This section addresses application issues so they are secure from well-known parameter manipulation attacks against common interpreters.

   Canoncalization, Locale and Unicode

This section addresses issues that help to ensure the application is robust when subjected to encoded, internationalized and Unicode input.

Error Handling, Auditing and Logging

This section deals with designing well-written applications that have dual-purpose logs and activity traces for audit and monitoring. This makes it easy to track a transaction without excessive effort or access to the system. They should possess the ability to easily track or identify potential fraud or anomalies end-to-end.

Distributed Computing

This section deals with synchronization and remote services to web applications, by hardening applications against: time of check, time of use race conditions distributed synchronization issues common multi-programming, multi-threaded and distributed security issues

Buffer Overflow

This section addresses issues such as: Applications do not expose themselves to faulty components Applications create as few buffer overflows as possible Developers are encouraged to use languages and frameworks that are relatively immune to buffer overflows.

Administrative Interfaces

This section addresses issues such that: Administrator level functions are appropriately segregated from user activity Users cannot access or utilize administrator functionality Provide necessary audit and traceability of administrative functionality

Cryptography

This section helps to ensures that cryptography is safely used to protect the confidentiality and integrity of sensitive user data

Configuration

This section is focused on creating secure web applications which are as well-built and secure out-of-the-box as possible.

Software Quality Assurance (QA)

According to the OWASP 3.0 guide, “The software quality assurance goal is to confirm the confidentiality and integrity of private user data is protected as the data is handled, stored, and transmitted. The QA testing should also confirm the application cannot be hacked, broken, commandeered, overloaded, or blocked by denial of service attacks, within acceptable risk levels. This implies that the acceptable risk levels and threat modeling scenarios are

established up front, so the developers and QA engineers know what to expect and what to work towards."

Deployment

This section deals with the issues surrounding secure deployment of web applications.

Maintenance

This section addresses issues such as: Products are properly maintained post deployment Minimize the attack surface area throughout the production lifecycle Security defects are fixed properly and in a timely fashion.

List of vulnerabilities (top 10) stated by OWASP that needs effective handling during coding phase.

1. Injection (SQL Injection):Attacker can access and modify databases
2. Broken Authentication and Session Management: attackers can assume users' identity
3. Cross-Site Scripting (XSS): Allows attackers to hijack user sessions.
4. Insecure Direct Object References: Attackers can access data.
5. Security Misconfiguration: Attacker can us gaps in configuration to attack.
6. Sensitive Data Exposure: Attackers may steal or modify Sensitive data.
7. Missing Function Level Access Control: Attackers will be able access functionality.
8. Cross-Site Request Forgery (CSRF): Allows the attacker to control victim's browser.
9. Using Components with Known Vulnerabilities: Attacker exploits components that run with full privileges.
10. Invalidated Redirects and Forwards: Attackers redirect victims to phishing or malware sites.

Security testing: Termed into different areas of Static and Dynamic testing of the application code base and of the application at run time respectively.

SAST - Static Application Security testing

This phase tests the code base as and when it is ready for release into UAT/ production environments. There are various tools both open source from OWASP and licensed software for SAST that scans through the code and generates report detailing the vulnerabilities at code level. Although, the automated report provides a few false positives, security team needs to work with the application team and ensure only the appropriate SAST defects are taken forward for fixing.

DAST - Dynamic Application Security testing

This is the final phase of testing wherein the application is tested at runtime after it is functionally and nonfunctionally ready in terms of security requirements, design, code and SAST fixes

incorporated with agreed upon SAST defects. The DAST test is executed and report provides details of the vulnerabilities and recommendations for fixing them.

Web application security is a branch of Information Security that deals specifically with security of websites, web applications and web services.

Cross-site Scripting (or XSS) is one of the most common application-layer web attacks. XSS commonly targets scripts embedded in a page which are executed on the client-side (in the user's web browser) rather than on the server-side. XSS in itself is a threat which is brought about by the internet security weaknesses of client-side scripting languages, with HTML and JavaScript (others being VBScript, ActiveX, HTML, or Flash) as the prime culprits for this exploit. The concept of XSS is to manipulate client-side scripts of a web application to execute in the manner desired by the malicious user. Such a manipulation can embed a script in a page which can be executed every time the page is loaded, or whenever an associated event is performed.

In a typical XSS attack the hacker infects a legitimate web page with his malicious client-side script. When a user visits this web page the script is downloaded to his browser and executed.

Cross-site scripting (XSS) is a type of computer security vulnerability typically found in Web applications. XSS enables attackers to inject client-side script into Web pages viewed by other users. A cross-site scripting vulnerability may be used by attackers to bypass access controls such as the same origin policy. Cross-site scripting carried out on websites accounted for roughly 84% of all security vulnerabilities documented by Symantec as of 2007. Their effect may range from a petty nuisance to a significant security risk, depending on the sensitivity of the data handled by the vulnerable site and the nature of any security mitigation implemented by the site's owner.

SQL injection: The ability to inject SQL commands into the database engine through an existing application.

SQL injection is a code injection technique, used to attack data-driven applications, in which malicious SQL statements are inserted into an entry field for execution (e.g. to dump the database contents to the attacker). SQL injection must exploit a security vulnerability in an application's software, for example, when user input is either incorrectly filtered for string literal escape characters embedded in SQL statements or user input is not strongly typed and unexpectedly executed. SQL injection is mostly known as an attack vector for websites but can be used to attack any type of SQL database.

11. Why not get off the shelf product instead of developing in-house?

- They may not meet all our organisational requirements
- Maintenance and customization needs extra money
- Off shelf product may not be exactly be compatible with our organisational softwares.

12. Business Process Re-engineering

    Business process reengineering (BPR) is the analysis and redesign of workflows within and between enterprises in order to optimize end-to-end processes and automate non-value-added tasks.

13. What are the different approaches to s/w development (structured programming, oo programming etc.)

    A software development methodology or system development methodology in software engineering is a framework that is used to structure, plan, and control the process of developing an information system.

    There are the following methodologies:

    - Agile Software Development
    - Crystal Methods
    - Dynamic Systems Development Model (DSDM)
    - Extreme Programming (XP)
    - Feature Driven Development (FDD)
    - Joint Application Development (JAD)
    - Lean Development (LD)
    - Rapid Application Development (RAD)
    - Rational Unified Process (RUP)
    - Scrum
    - Spiral
    - Systems Development Life Cycle (SDLC)
    - Waterfall (a.k.a. Traditional)

    Several software development approaches have been used since the origin of information technology, in two main categories. Typically an approach or a combination of approaches is chosen by management or a development team.

    "Traditional" methodologies such as waterfall that have distinct phases are sometimes known as software development life cycle (SDLC) methodologies, though this term could also be used more generally to refer to any methodology. A "life cycle" approach with distinct phases is in contrast to Agile approaches which define a process of iteration, but where design, construction, and deployment of different pieces can occur simultaneously.

    Waterfall development

The activities of the software development process represented in the waterfall model. There are several other models to represent this process. The waterfall model is a sequential development approach, in which development is seen as flowing steadily downwards (like a waterfall) through several phases, typically:

- Requirements analysis resulting in a software requirements specification
- Software design
- Implementation
- Testing
- Integration, if there are multiple subsystems
- Deployment (or Installation)
- Maintenance

Project is divided into sequential phases, with some overlap and splashback acceptable between phases. Emphasis is on planning, time schedules, target dates, budgets and implementation of an entire system at one time. Tight control is maintained over the life of the project via extensive written documentation, formal reviews, and approval/signoff by the user and information technology management occurring at the end of most phases before beginning the next phase.

The waterfall model is a traditional engineering approach applied to software engineering. A strict waterfall approach discourages revisiting and revising any prior phase once it is complete. This "inflexibility" in a pure waterfall model has been a source of criticism by supporters of other more "flexible" models. It has been widely blamed for several large-scale government projects running over budget, over time and sometimes failing to deliver on requirements due to the Big Design Up Front approach. Except when contractually required, the waterfall model has been largely superseded by more flexible and versatile methodologies developed specifically for software development. See Criticism of Waterfall model.

Prototyping

Software prototyping, is the development approach of activities during software development, the creation of prototypes, i.e., incomplete versions of the software program being developed.

The basic principles are:

Not a standalone, complete development methodology, but rather an approach to handle selected parts of a larger, more traditional development methodology (i.e. incremental, spiral, or rapid application development (RAD)).

Attempts to reduce inherent project risk by breaking a project into smaller segments and providing more ease-of-change during the development process.

User is involved throughout the development process, which increases the likelihood of user acceptance of the final implementation.

Small-scale mock-ups of the system are developed following an iterative modification process until the prototype evolves to meet the users' requirements.

While most prototypes are developed with the expectation that they will be discarded, it is possible in some cases to evolve from prototype to working system.

A basic understanding of the fundamental business problem is necessary to avoid solving the wrong problems.

Incremental development

Various methods are acceptable for combining linear and iterative systems development methodologies, with the primary objective of each being to reduce inherent project risk by breaking a project into smaller segments and providing more ease-of-change during the development process.

The basic principles are:

A series of mini-Waterfalls are performed, where all phases of the Waterfall are completed for a small part of a system, before proceeding to the next increment, or

Overall requirements are defined before proceeding to evolutionary, mini-Waterfall development of individual increments of a system, or

The initial software concept, requirements analysis, and design of architecture and system core are defined via Waterfall, followed by iterative Prototyping, which culminates in installing the final prototype, a working system.

Iterative and incremental development

Iterative development prescribes the construction of initially small but ever-larger portions of a software project to help all those involved to uncover important issues early before problems or faulty assumptions can lead to disaster.

Spiral development

In 1988, Barry Boehm published a formal software system development "spiral model," which combines some key aspect of the waterfall model and rapid prototyping methodologies, in an effort to combine advantages of top-down and bottom-up concepts. It provided emphasis in a key area many felt had been neglected by other methodologies: deliberate iterative risk analysis, particularly suited to large-scale complex systems.

The basic principles are:

Focus is on risk assessment and on minimizing project risk by breaking a project into smaller segments and providing more ease-of-change during the development process, as well as providing the opportunity to evaluate risks and weigh consideration of project continuation throughout the life cycle.

"Each cycle involves a progression through the same sequence of steps, for each part of the product and for each of its levels of elaboration, from an overall concept-of-operation document down to the coding of each individual program."

Each trip around the spiral traverses four basic quadrants: (1) determine objectives, alternatives, and constraints of the iteration; (2) evaluate alternatives; Identify and resolve risks; (3) develop and verify deliverables from the iteration; and (4) plan the next iteration.

Begin each cycle with an identification of stakeholders and their "win conditions", and end each cycle with review and commitment.

Rapid application development

Rapid application development (RAD) is a software development methodology, which favors iterative development and the rapid construction of prototypes instead of large amounts of up-front planning. The "planning" of software developed using RAD is interleaved with writing the software itself. The lack of extensive pre-planning generally allows software to be written much faster, and makes it easier to change requirements.

The rapid development process starts with the development of preliminary data models and business process models using structured techniques. In the next stage, requirements are verified using prototyping, eventually to refine the data and process models. These stages are repeated iteratively; further development results in "a combined business requirements and technical design statement to be used for constructing new systems".

The term was first used to describe a software development process introduced by James Martin in 1991. According to Whitten (2003), it is a merger of various structured techniques, especially data-driven Information Engineering, with prototyping techniques to accelerate software systems development.

The basic principles of rapid application development are:

- Key objective is for fast development and delivery of a high quality system at a relatively low investment cost.
- Attempts to reduce inherent project risk by breaking a project into smaller segments and providing more ease-of-change during the development process.

- Aims to produce high quality systems quickly, primarily via iterative Prototyping (at any stage of development), active user involvement, and computerized development tools. These tools may include Graphical User Interface (GUI) builders, Computer Aided Software Engineering (CASE) tools, Database Management Systems (DBMS), fourth-generation programming languages, code generators, and object-oriented techniques.
- Key emphasis is on fulfilling the business need, while technological or engineering excellence is of lesser importance.
- Project control involves prioritizing development and defining delivery deadlines or “timeboxes”. If the project starts to slip, emphasis is on reducing requirements to fit the timebox, not in increasing the deadline.
- Generally includes joint application design (JAD), where users are intensely involved in system design, via consensus building in either structured workshops, or electronically facilitated interaction.
- Active user involvement is imperative.
- Iteratively produces production software, as opposed to a throwaway prototype.
- Produces documentation necessary to facilitate future development and maintenance.
- Standard systems analysis and design methods can be fitted into this framework.

Agile development

"Agile software development" refers to a group of software development methodologies based on iterative development, where requirements and solutions evolve via collaboration between self-organizing cross-functional teams. The term was coined in the year 2001 when the Agile Manifesto was formulated.

Agile software development uses iterative development as a basis but advocates a lighter and more people-centric viewpoint than traditional approaches. Agile processes fundamentally incorporate iteration and the continuous feedback that it provides to successively refine and deliver a software system.

There are many variations of agile processes:
Dynamic systems development method (DSDM)
Kanban
Scrum

14. What are quality parameters for software?

Software quality is the characteristic of the software that defines how well the software meets the customer requirements, business requirements, coding standards etc. It can be divided into two categories:

Software Functional Quality: characteristics that define how well the software meets functional requirements, and how well it satisfies the end-users.

Software Non-Functional Quality: characteristics that define how well structural requirements are met that support the delivery of functional requirements. It is usually related to software code and internal structure.

The different software qualities can be measured through various software testing techniques and tools. Following are the different attributes (parameters) that are used to measure the software quality:

Testability – How easy it is to test the software and to what extent it can be tested.

Usability – It defines how user friendly the software is.

Understandability – How easily the software can be made understood to a layman about its functions/purpose

Consistency – How far the software is consistent / uniform in terms of GUI, terminologies, conventions.

Efficiency – It defines the amount of work that the software can do in defined timeframe with minimum resources used.

Effectiveness – The extent to which the software satisfies the user and meets user requirements

Accuracy – How accurately the software works with gives the correct results.

Maintainability – How easily the software can be maintained in terms of enhancements/new features/bug fixes.

Reliability – How reliable the software is in performing required functions under different scenarios/conditions.

Portability – How easily the software can be transported and run in different environments e.g. platforms like operating systems (Linux, Mac, Windows), machines it can run on.

Security – How secured the software is in terms of access authorization and personal data like passwords.

Robustness – How robust the software is under unexpected events like software crash, power-off etc and saves its data.

15. From Data you get Information? What do you get from Information? (Ans: Knowledge)
16. From Knowledge what do you get? (Ans: Wisdom)
17. What is the protocol used by Ping Command?

ICMP means Internet Control Message Protocol and is always coupled with the IP protocol (There's 2 ICMP variants one for IPv4 and one for IPv6.) echo request and echo response are the two operation codes of ICMP used to implement ping.

18. What are the Advantages of Open source softwares?

    Open source software is generally free software than you can use in your business. Open source developers choose to make the source code of their software publicly available for the good of the community and publish their software with an open source license, meaning that other developers can see how it works and add to it. Examples of open source products include Open Office, the internet browser Mozilla Firefox, Wikipedia, the GNU/Linux operating system and its derivative Android, an operating system for mobile devices.

    - It's generally free - it has been estimated that open source software collectively saves businesses $60 billion a year. These days for virtually every paid for proprietary software system you will find an open source version.
    - It continually evolving in real time as developers add to it and modify it, which means it can be better quality and more secure and less prone to bugs than proprietary systems, because it has so many users poring over it and weeding out problems.
    - Using open source software also means you are not locked in to using a particular vendor's system that only work with their other systems.
    - You can modify and adapt open source software for your own business requirements, something that is not possible with proprietary systems.

19. What is the purpose of the button "I am Feeling lucky" in Google?

    Google's homepage includes a button labeled "I'm Feeling Lucky". Prior to a change in 2012 when a user typed in a search and clicked on the button the user would be taken directly to the first search result, bypassing the search engine results page. The idea was that if a user is "feeling lucky", the search engine would return the perfect match the first time without having to page through the search results. According to a study by Tom Chavez of "Rapt", this feature cost Google $110 million a year as 1% of all searches use this feature and bypass all advertising

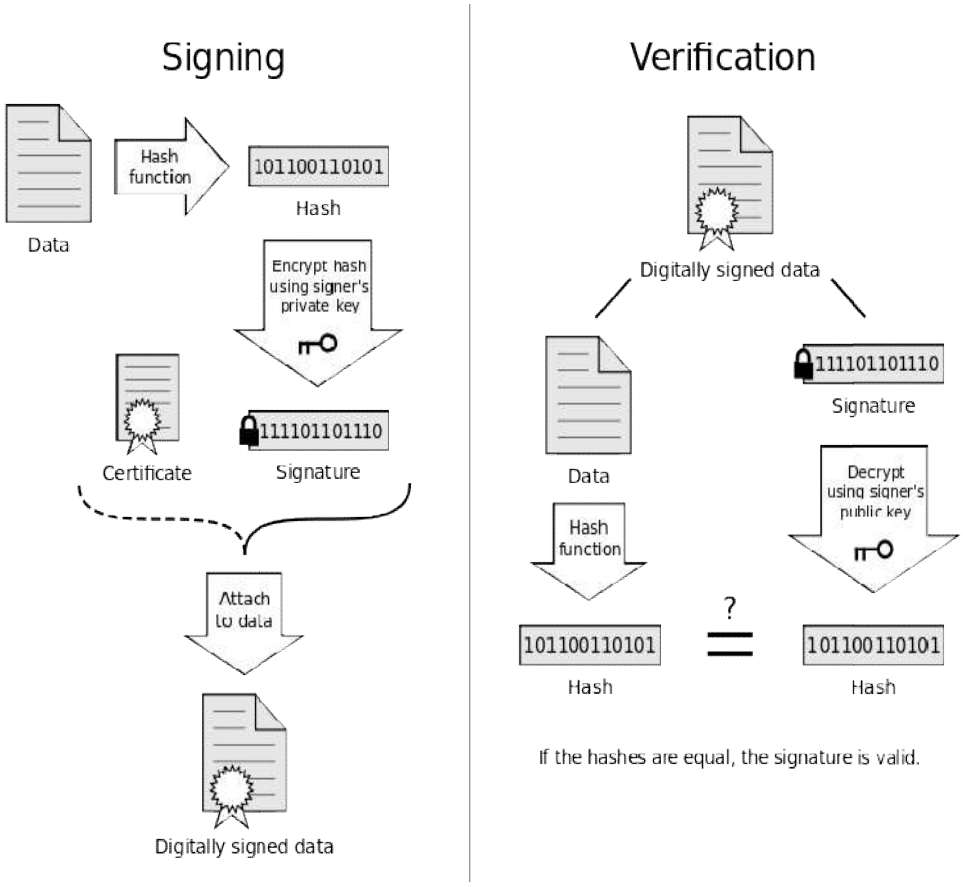
20. What is Digital Signature?

    A digital signature is a mathematical scheme for demonstrating the authenticity of a digital message or document. A valid digital signature gives a recipient reason to believe that the message was created by a known sender, such that the sender cannot deny having sent the message (authentication and non-repudiation) and that the message was not altered in transit (integrity). Digital signatures are commonly used for software distribution, financial transactions, and in other cases where it is important to detect forgery or tampering.

A digital signature scheme typically consists of three algorithms;

- A key generation algorithm that selects a private key uniformly at random from a set of possible private keys. The algorithm outputs the private key and a corresponding public key.
- A signing algorithm that, given a message and a private key, produces a signature.
- A signature verifying algorithm that, given a message, public key and a signature, either accepts or rejects the message's claim to authenticity.

Two main properties are required. First, the authenticity of a signature generated from a fixed message and fixed private key can be verified by using the corresponding public key. Secondly, it should be computationally infeasible to generate a valid signature for a party without knowing that party's private key. A digital signature is an authentication mechanism that enables the creator of message to attach a code that act as a signature. It is formed by taking the hash of message and encrypting the message with creator's private key.



In cryptography, a public key certificate (also known as a digital certificate or identity certificate) is an electronic document used to prove ownership of a public key. The certificate includes information about the key, information about its owner's identity, and the digital signature of an entity that has verified the certificate's contents are correct. If the signature is valid, and the person examining the certificate trusts the signer, then they know they can use that key to communicate with its owner.

In a typical public-key infrastructure (PKI) scheme, the signer is a certificate authority (CA), usually a company which charges customers to issue certificates for them. In a web of trust scheme, the signer is either the key's owner (a self-signed certificate) or other users ("endorsements") whom the person examining the certificate might know and trust.

Certificates are an important component of Transport Layer Security (TLS, sometimes called by its older name SSL), where they prevent an attacker from impersonating a secure website or other server. They are also used in other important applications, such as email encryption and code signing.

21. Difference between HTTP and HTTPS.

Hypertext Transfer Protocol (HTTP) is a protocol used in networking. When you type any web address in your web browser, your browser acts as a client, and the computer having the requested information acts as a server. When client requests for any information from the server, it uses HTTP protocol to do so. The server responds back to the client after the request completes. The response comes in the form of web page which you see just after typing the web address and press “Enter”.

Hypertext Transfer Protocol Secure (HTTPS) is a combination of two different protocols. It is more secure way to access the web. It is combination of Hypertext Transfer Protocol (HTTPS) and SSL/TLS protocol. It is more secure way to sending request to server from a client, also the communication is purely encrypted which means no one can know what you are looking for. This kind of communication is used for accessing those websites where security is required. Banking websites, payment gateway, emails (Gmail offers HTTPS by default in Chrome browser), and corporate sector websites are some great examples where HTTPS protocols are used.

HTTP functions as a request-response protocol in the client-server computing model. A web browser, for example, may be the client and an application running on a computer hosting a web site may be the server. The client submits an HTTP request message to the server. The server, which provides resources such as HTML files and other content, or performs other functions on behalf of the client, returns a response message to the client. The response contains completion status information about the request and may also contain requested content in its message body.

A web browser is an example of a user agent (UA). Other types of user agent include the indexing software used by search providers (web crawlers), voice browsers, mobile apps, and other software that accesses, consumes, or displays web content.

HTTP is designed to permit intermediate network elements to improve or enable communications between clients and servers. High-traffic websites often benefit from web cache servers that deliver content on behalf of upstream servers to improve response time. Web browsers cache previously accessed web resources and reuse them when possible to reduce network traffic.

HTTP proxy servers at private network boundaries can facilitate communication for clients without a globally routable address, by relaying messages with external servers.

HTTP is an application layer protocol designed within the framework of the Internet Protocol Suite. Its definition presumes an underlying and reliable transport layer protocol,[3] and Transmission Control Protocol (TCP) is commonly used. However HTTP can use unreliable protocols such as the User Datagram Protocol (UDP), for example in Simple Service Discovery Protocol (SSDP).

For HTTPS connection, public key trusted and signed certificate is required for the server. These certificate comes either free or it costs few dollars depends on the signing authority. There is one other method for distributing certificates. Site admin creates certificates and loads in the browser of users. Now when user requests information to the web server, his identity can be verified easily.

Here are some major differences between HTTP and HTTPS:

HTTP	HTTPS
URL begins with “http://”	URL begins with “https://”
It uses port 80 for communication	It uses port 443 for communication
Unsecured	Secured
Operates at Application Layer	Operates at Transport Layer
No encryption	Encryption is present
No certificates required	Certificates required

22. How many layer does SSL has? What are its functions?

SSL (Secure Sockets Layer) is a standard security technology for establishing an encrypted links between a server and a client - typically a web server (website) and a browser; or a mail server and a mail client (e.g., Outlook).

SSL allows sensitive information such as credit card numbers, social security numbers, and login credentials to be transmitted securely. Normally, data sent between browsers and web servers is sent in plain text - leaving you vulnerable to eavesdropping. If an attacker is able to intercept all data being sent between a browser and a web server they can see and use that information.

- Handshaking Protocol
  Establish communication variables
- ChangeCipherSpec Protocol
  Alert to a change in communication variables
- Alert Protocol
  Messages important to SSL connections
- Application Encryption Protocol
  Encrypt/Decrypt application data

23. Explain what is Struts 2? Advantages and disadvantages of Struts 2.

Apache Struts 2 is an open-source web application framework for developing Java EE web applications. It uses and extends the Java Servlet API to encourage developers to adopt a model–view–controller (MVC) architecture. The WebWork framework spun off from Apache Struts aiming to offer enhancements and refinements while retaining the same general architecture of the original Struts framework.

Advantages of Struts2

- Simplified Design: Code is not tightly coupled to Struts framework or Servlet API.
- Easy plug-in: Developers can use other technologies plug-in easily. It includes SiteMesh, Spring, Tiles, etc.
- Simplified ActionForm: ActionForms are POJOs, we do not need to implement any interface or extend from any class.
- Annotations introduced: Use of annotation results in reduction in length and complexity of code. It is also used in configuration file for simplicity.
- Better tag features: It includes theme based tags and Ajax enabled tags.
- Simplified Testability: Unit testing of Struts 2 Action class is very easy because it doesn't need complex HttpServletRequest and HttpServletResponse objects.
- Simplified Action: Similar to ActionForms, Actions are also simple POJOs and they do not need to implement any interface or extend any class.
- OGNL integration: It uses OGNL to fetch data from ValueStack and type conversion which reduces code.
- Ajax support: Struts2 tags are Ajax enabled.
- Multiple View options: View is not restricted to JSP. Freemarker, velocity templates can also be used as a view.

Disadvantages of Struts2

- Compatibility: Struts 2 is completely different from the Struts 1. So it's difficult to perform migration of applications from Struts 1 to Struts 2.
- Limited Documentation: Limited documentation is available for Struts 2. In addition to this, new users find it difficultly to understand its concepts due to poorly managed documentation by Apache

24. How to develop a Search Engine?

A web search engine is a software system that is designed to search for information on the World Wide Web. The search results are generally presented in a line of results often referred to as search engine results pages (SERPs). The information may be a mix of web pages, images, and other types of files. Some search engines also mine data available in databases or open directories. Unlike web directories, which are maintained only by human editors, search engines also maintain real-time information by running an algorithm on a web crawler.

25. Why do you use Java for development of Web application

Web application development is at an all time high as more and more people are gaining access to the web, as well as the rise of many web-based businesses. Web applications are apps that run in a web browser and also are popular because of the ubiquity of web browsers. It's also advantageous due to the fact that you don't have to install the software or application on dozens, hundreds, or even thousands of machines, which in turn makes the application easier to maintain. Java is a popular web application programming language and has several advantages.

One of the main advantages of Java in software and application development is "that it is a cross-platform tool. Thanks to the JVM (Java Virtual Machine), Java's run-time environment is able to translate code into machine code compatible with the native operating system, whether that's Windows, iOS or Linux. This versatility, and specifically the cross-platform functionality, immediately makes it a powerful tool for larger organizations doing software development work", as described by Inside Technology.

Java is a specialist language, but is also a C based programming language which makes it easy for developers to learn or "pick-up" if they've worked with other C-based languages. This makes finding developers much easier. Java is also not a "committee-driven" language and is owned by Sun, which means that code is often cleaner and has a vast library of classes that work well for rapid development.

26. How do you develop Software?

Planning

The decision that is to be made on whether to create your own software should be preceded by a thorough analysis, e.g. for identifying the costs and stages of the process. The first and most crucial requirement is to develop very detailed documentation with lists of all the required solutions, functionalities, scoping for the system, etc. This exhaustive specification is the basis for the subsequent work that is to be performed by the programming team. It is also the reference for technology choices, user interface layout design, and workflow processes. Importantly, the specification should not be an outline of just a few pages, with vaguely formulated requirements. It should total more than 100 pages of detailed descriptions of all the components and methods for carrying out the subsequent workflow stages. Roughly speaking, the specification development work takes a few to several weeks.

The next thing to consider is who will run the project and be the interface between the client and the project team. The role of the designated person, who is to be the project manager, is to translate the initial vision into a structured system, and then into a set of components that are gradually carried out by the team. It should also be remembered that IT professionals use a highly specialised lingo and, therefore, the project manager also needs to translate requirements and

functionalities that are agreed upon with the system concept owner into IT speak, which can be understood by those who are involved in writing the software code.

Technology

The next step in the project is the choice of technology. We should now start looking for an expert, usually charging high rates, who is experienced in the solutions available in the market. This person should, in line with the formulated needs and expectations, propose a technology that is not only good for the specific point in time, but one that also enables free development and software modifications in the future. Here, we should consider how the company's expectations might evolve over time, whether the environment and market requirements are likely to change, and whether modifications or extensions are likely to be needed in the future. Mistakes in the technology selection stage may thwart the company later on in the project, and involve additional costs.

Programming team

Now we can move on to building a team of programmers. We need experts who have hands-on experience in the selected technology and can reliably carry out the subsequent stages of the project. It may not appear so, but this stage is quite difficult and risky. Just as in any other domain, the risk always exists that we may recruit people who do not meet our expectations. We need to be aware that although many may declare that they can do the job, they may in fact lack in the necessary experience and skills. The consequences of poor recruitment are dire for the company in the first place. Therefore, we need to come up with an effective method of candidate review and accept the costs involved along with the time needed to build a good and creative team.

Development process

When we have the necessary human resources in place, we can then proceed to the project delivery, which are the practical details of "code writing." The time needed for this stage depends on the level of complexity of the system that we are designing and the size of the programming team. We can expect, for the software that is used to manage translation departments and agencies, the period to be from a few to several months. This is also the period for the hours of consultations, fine-tuning of individual matters addressed by the initial specification, and searching for optimum solutions to problems as they emerge. Clearly, the more extensive and accurate the initial specification, the less time that should need to be spent on interaction with the project manager.

Testing

After several months of work, when the system is becoming fully formed, we can embark on the testing task. The new product is uploaded on test servers and its operation is simulated so that all

the functionalities are within the tested scope. These tests are usually run at both the programming and user levels, so that they require the full commitment of the company's staff. The testing is usually a painstaking and time-consuming process, and it may even interfere with the usual operations of the company. Still, we need to remember that the more accurate and meticulous we are at this stage, the easier and smoother the final go live will be.

The output of the testing work is a detailed list of components that require improvements. The product comes back to the programmers, where all the bugs are removed, and the expected changes implemented. This is how the next version of the software is developed, and the process is again followed by the same testing procedure. After a few such cycles have been repeated, we finally arrive at the system that makes the client happy and moves us to another stage, namely implementation and staff training.

27. How many ports do the protocols FTP & TELNET use?

The File Transfer Protocol (FTP) is a standard network protocol used to transfer computer files from one host to another host over a TCP-based network, such as the Internet.
FTP is built on a client-server architecture and uses separate control and data connections between the client and the server.[1] FTP users may authenticate themselves using a clear-text sign-in protocol, normally in the form of a username and password, but can connect anonymously if the server is configured to allow it. For secure transmission that protects the username and password, and encrypts the content, FTP is often secured with SSL/TLS (FTPS). SSH File Transfer Protocol (SFTP) is sometimes also used instead, but is technologically different.

28. What are the types of backup? Differentiate Hot backup and cold backup

There are quite a number of backup types and terms used when it comes to backups of your digital content. This is a compilation of the most common types of backup with a brief explanation of their meaning, common examples, advantages and disadvantages of each backup type.

Full Backup

Full backup is a method of backup where all the files and folders selected for the backup will be backed up. When subsequent backups are run, the entire list of files and will be backed up again. The advantage of this backup is restores are fast and easy as the complete list of files are stored each time. The disadvantage is that each backup run is time consuming as the entire list of files is copied again. Also, full backups take up a lot more storage space when compared to incremental or differential backups. Read more…

Incremental backup

Incremental backup is a backup of all changes made since the last backup. With incremental backups, one full backup is done first and subsequent backup runs are just the changes made since the last backup. The result is a much faster backup then a full backup for each backup run.

Storage space used is much less than a full backup and less then with differential backups. Restores are slower than with a full backup and a differential backup. Read more…

Differential backup

Differential backup is a backup of all changes made since the last full backup. With differential backups, one full backup is done first and subsequent backup runs are the changes made since the last full backup. The result is a much faster backup then a full backup for each backup run. Storage space used is much less than a full backup but more then with Incremental backups. Restores are slower than with a full backup but usually faster then with Incremental backups. Read more...

Mirror Backup

Mirror backups are as the name suggests a mirror of the source being backed up. With mirror backups, when a file in the source is deleted, that file is eventually also deleted in the mirror backup. Because of this, mirror backups should be used with caution as a file that is deleted by accident or through a virus may also cause the mirror backups to be deleted as well. Read more…

Full PC Backup or Full Computer Backup

In this backup, it is not the individual files that are backed up but entire images of the hard drives of the computer that is backed up. With the full PC backup, you can restore the computer hard drives to its exact state when the backup was done. With the Full PC backup, not only can the work documents, picture, videos and audio files be restored but the operating system, hard ware drivers, system files, registry, programs, emails etc can also be restored. Read more…

Local Backup

Local backups are any kind of backup where the storage medium is kept close at hand or in the same building as the source. It could be a backup done on a second internal hard drive, an attached external hard drive, CD/ DVD –ROM or Network Attached Storage (NAS). Local backups protect digital content from hard drive failures and virus attacks. They also provide protection from accidental mistakes or deletes. Since the backups are always close at hand they are fast and convenient to restore. Read more…

Offsite Backup

When the backup storage media is kept at a different geographic location from the source, this is known as an offsite backup. The backup may be done locally at first but once the storage medium is brought to another location, it becomes an offsite backup. Examples of offsite backup include taking the backup media or hard drive home, to another office building or to a bank safe deposit box.

Beside the same protection offered by local backups, offsite backups provide additional protection from theft, fire, floods and other natural disasters. Putting the backup media in the next room as

the source would not be considered an offsite backup as the backup does not offer protection from theft, fire, floods and other natural disasters. Read more…

Online Backup

These are backups that are ongoing or done continuously or frequently to a storage medium that is always connected to the source being backed up. Typically the storage medium is located offsite and connected to the backup source by a network or Internet connection. It does not involve human intervention to plug in drives and storage media for backups to run. Many commercial data centres now offer this as a subscription service to consumers. The storage data centres are located away from the source being backed up and the data is sent from the source to the storage data centre securely over the Internet. Read more…

Remote Backup

Remote backups are a form of offsite backup with a difference being that you can access, restore or administer the backups while located at your source location or other location. You do not need to be physically present at the backup storage facility to access the backups. For example, putting your backup hard drive at your bank safe deposit box would not be considered a remote backup. You cannot administer it without making a trip to the bank. Online backups are usually considered remote backups as well. Read more…

Cloud Backup

This term is often used interchangeably with Online Backup and Remote Backup. It is where data is backed up to a service or storage facility connected over the Internet. With the proper login credentials, that backup can then be accessed or restored from any other computer with Internet Access. Read more…

FTP Backup

This is a kind of backup where the backup is done via FTP (File Transfer Protocol) over the Internet to an FTP Server. Typically the FTP Server is located in a commercial data centre away from the source data being backed up. When the FTP server is located at a different location, this is another form of offsite backup. Read more…

Cold database backup: copy of the database which can be restored exactly.

During a cold backup, the database is closed or locked and not available to users. The data files do not change during the backup process so the database is in a consistent state when it is returned to normal operation.

Normal system backups, referred to as either Hot or Cold backups, are used to protect from media failure. A Cold backup, that is, one done with the database in a shutdown state, provides a complete **Hot database backup**

Some database management systems offer a means to generate a backup image of the database while it is online and usable ("hot"). This usually includes an inconsistent image of the data files

plus a log of changes made while the procedure is running. Upon a restore, the changes in the log files are reapplied to bring the copy of the database up-to-date (the point in time at which the initial hot backup ended)

A Hot backup, or one taken while the database is active, can only give a read consistent copy, but doesn't handle active transactions. All data in the Oracle or system buffers and all non-committed changes may be lost unless a redo log switch is forced, the resulting archive log and a control file copy taken along with the hot file backup. In order to use the hot backup methodology, the database must be in archivelog mode.

29. What is a Port?

In computer networking, a port is a software construct serving as a communications endpoint in a computer's host operating system. A port is always associated with an IP address of a host and the protocol type of the communication. It completes the destination or origination address of a communications session. A port is identified for each address and protocol by a 16-bit number, commonly known as the port number.

1) On computer and telecommunication devices, a port (noun) is generally a specific place for being physically connected to some other device, usually with a socket and plug of some kind. Typically, a personal computer is provided with one or more serial ports and usually one parallel port. The serial port supports sequential, one bit-at-a-time transmission to peripheral devices such as scanners and the parallel port supports multiple-bit-at-a-time transmission to devices such as printers.

2) In programming, a port (noun) is a "logical connection place" and specifically, using the Internet's protocol, TCP/IP, the way a client program specifies a particular server program on a computer in a network. Higher-level applications that use TCP/IP such as the Web protocol, Hypertext Transfer Protocol, have ports with preassigned numbers. These are known as "well-known ports" that have been assigned by the Internet Assigned Numbers Authority (IANA). Other application processes are given port numbers dynamically for each connection. When a service (server program) initially is started, it is said to bind to its designated port number. As any client program wants to use that server, it also must request to bind to the designated port number.

Port numbers are from 0 to 65535. Ports 0 to 1024 are reserved for use by certain privileged services. For the HTTP service, port 80 is defined as a default and it does not have to be specified in the Uniform Resource Locator (URL).

30. What is Software Evaluation, How do you evaluate software?

A software evaluation is a type of assessment that seeks to determine if software or a combination of software programs is the best possible fit for the needs of a given client. The idea is to look closely at the resources and tools provided by the software that is either currently in use or is

being examined as a possible addition to programs already in use by that client. Based on a prepared list of criteria along with some practical experimentation, a software evaluation makes it possible to determine if the products would be helpful to the client or if some other combination of software products would serve to better advantage.

There are several factors to consider with any software evaluation. One has to do with compatibility of the software with the hardware resources already in place on the client's network or computer equipment. Here, the focus is on the type of operating system the software requires in order to function, as well as the amount of memory and capacity that the hardware currently provides. This is particularly important if there is no budget for hardware and memory upgrades that would accommodate the software under consideration.

Another key factor in software evaluation is how well the proposed software package will interact with other applications already in place. For example, if a proposed word processing program were found to be unable to easily import and export data from the sales database currently in use by the company, this would mean additional time spent preparing mailing pieces to customers and prospects. In like manner, if software used to process customer orders will not download to the accounting software, this can add more manual steps to the preparation of invoices. The right combination of software programs can streamline essential functions, allowing employees more time to devote to other activities that help to enhance the process of revenue generation.

Software evaluation is necessary to make sure that all software used by an individual or business is actually increasing the efficiency of the operation rather than creating additional work loads. While individuals and companies can conduct this type of evaluation on their own, there are also consultants who can engage in product and software assessment for a client, making suggestions for any changes or additions that would be in the best interests of the client. This approach can often uncover issues that would be overlooked otherwise, ultimately saving the company a great deal of money in terms of labor and other types of operational costs.

31. What happens when a LINUX system is interrupted from electricity?

    If the server used ext2 file system you could be very susceptible to file corruption in the event of a power failure. Compare that to a system that used something like Btrfs which has more data integrity features than NTFS does.

32. What is Firewall?

    In computing, a firewall is a network security system that controls the incoming and outgoing network traffic based on an applied rule set. A firewall establishes a barrier between a trusted, secure internal network and another network (e.g., the Internet) that is assumed not to be secure and trusted. Firewalls exist both as software to run on general purpose hardware and as a

hardware appliance. Many hardware-based firewalls also offer other functionality to the internal network they protect, such as acting as a DHCP server for that network.

33. What is Script? Differentiate between Psuedocode and Scripting.

    A high-level programming language that is interpreted by another program at runtime rather than compiled by the computer's processor as other programming languages (such as C and C++) are. Scripting languages, which can be embedded within HTML, commonly are used to add functionality to a Web page, such as different menu styles or graphic displays or to serve dynamic advertisements. These types of languages are client-side scripting languages, affecting the data that the end user sees in a browser window. Other scripting languages are server-side scripting languages that manipulate the data, usually in a database, on the server.

    Scripting languages came about largely because of the development of the Internet as a communications tool. JavaScript, ASP, JSP, PHP, Perl, Tcl and Python are examples of scripting languages.

    JavaScript is a dynamic computer programming language.[5] It is most commonly used as part of Web browsers, whose implementations allow client-side scripts to interact with the user, control the browser, communicate asynchronously, and alter the document content that is displayed

    Pseudocode is an informal high-level description of the operating principle of a computer program or other algorithm. or An outline of a program, written in a form that can easily be converted into real programming statements.

    Pseudocode cannot be compiled nor executed, and there are no real formatting or syntax rules. It is simply one step - an important one - in producing the final code. The benefit of pseudocode is that it enables the programmer to concentrate on the algorithms without worrying about all the syntactic details of a particular programming language. In fact, you can write pseudocode without even knowing what programming language you will use for the final implementation.

34. What are tools used for web development?

    - In the case of ASP.NET, a developer can use Microsoft Visual Studio to write code. But, as with most other programming languages, he/she can also use a text editor. Notepad++ is an example. WebORB Integration Server for .NET can be used to integrate .NET services, data and media with any web client. It includes developer productivity tools and APIs for remoting, messaging and data management.
    - For ColdFusion and the related open source CFML engines, there are several tools available for writing code. These include Adobe Dreamweaver CS4, theCFEclipse plugin for Eclipse (software) and Adobe CF Builder. You can also use any text editor such as Notepad++ or TextEdit.

- For PHP, the Zend Development Environment provides numerous debugging tools and provides a rich feature set to make a PHP developer's life easier. WebORB Integration Server for PHP can be used to integrate PHP classes and data with any web client. It includes developer productivity tools and APIs for remoting, messaging and data management. Tools such as Hammerkit abstract PHP into a visual programming environment and utilise component-based software methods to accelerate development.
- For Java (programming language), there are many tools. The most popular is Apache Tomcat, but there are many others. One very specific one is WebORB Integration Server which can be used to integrate Java services, data and media with any web client. It includes developer productivity tools and APIs for remoting, messaging and data management.
- Several code generation tools such as nuBuilder, dbQwikSite or M-Power are available to automate the development of code. Using such tools, non-technical users can produce working code, and experienced coders can accelerate the development cycle.
- Other tools include various browsers, FTP clients, etc.

35. What is the difference between Servlet and Applet?

Applets

- Applets are applications designed to be transmitted over the network and executed by Java compatible web browsers.
- An Applet is a client side java program that runs within a Web browser on the client machine.
- An applet can use the user interface classes like AWT or Swing.
- Applet Life Cycle Methods: init(), stop(), paint(), start(), destroy()

Servlets

- Servlets are Java based analog to CGI programs, implemented by means of servlet container associated with an HTTP server.
- Servlet is a server side component which runs on the web server.
- The servlet does not have a user interface.
- Servlet Methods: doGet(), doPost()

36. How do you measure the performance of software?

In software engineering, performance of a software is in general testing performed to determine how a system performs in terms of responsiveness and stability under a particular workload.

37. What is CAPTCHA? What are it advantages and Disadvantages?

A CAPTCHA (an acronym for "Completely Automated Public Turing test to tell Computers and Humans Apart") is a type of challenge-response test used in computing to determine whether or not the user is human.

Advantages:

- Distinguishes between a human and a machine
- Makes online polls more legitimate
- Reduces spam and viruses
- Makes online shopping safer
- Diminishes abuse of free email account services

Disadvantages:

- Sometimes very difficult to read
- Are not compatible with users with disabilities
- Time-consuming to decipher
- Technical difficulties with certain internet browsers
- May greatly enhance Artificial Intelligence

Don't you think, OCR's can read the CAPTCHA values?

It depends on captcha. Standard OCR aren't meant for CAPTCHA breaking. Anyway simple captcha can be preprocessed (e.g. Captcha OCR Tutorial) and then feeded to an OCR engine, sometimes it works... In general CAPTCHA breaking is much more complex than downloading the Tesseract binaries. If it were that easy, all of the paid services would be out of business over night.

38. What is CERT?

Computer emergency response teams (CERT) are expert groups that handle computer security incidents. The history of CERTs is linked to the existence of malware, especially computer worms and viruses. Whenever a new technology arrives, its misuse is not long in following. The first worm in the IBM VNET was covered up. Shortly after, a worm hit the Internet on 3 November 1988, when the so-called Morris Worm paralysed a good percentage of it. This led to the formation of the first computer emergency response team at Carnegie Mellon University under U.S. Government contract.

39. What is Proxy Server? How it has been set up in ISAC?

A proxy server is a server (a computer system or an application) that acts as an intermediary for requests from clients seeking resources from other servers. A client connects to the proxy server, requesting some service, such as a file, connection, web page, or other resource available from a different server and the proxy server evaluates the request as a way to simplify and control its complexity. Proxies were invented to add structure and encapsulation to distributed systems

40. Differentiate between DNS and NAT

Domain Name System

Domain Name System (or Service or Server), an Internet service that translates domain names into IP addresses. Because domain names are alphabetic, they're easier to remember. The Internet however, is really based on IP addresses. Every time you use a domain name, therefore, a DNS service must translate the name into the corresponding IP address. For example, the domain name www.example.com might translate to 198.105.232.4.

The DNS system is, in fact, its own network. If one DNS server doesn't know how to translate a particular domain name, it asks another one, and so on, until the correct IP address is returned.

In computer networks, a proxy server is a server (a computer system or an application) that acts as an intermediary for requests from clients seeking resources from other servers. A client connects to the proxy server, requesting some service, such as a file, connection, web page, or other resource available from a different server and the proxy server evaluates the request as a way to simplify and control its complexity. Proxies were invented to add structure and encapsulation to distributed systems.

NAT - Network Address Translation

Short for Network Address Translation, an Internet standard that enables a local-area network (LAN) to use one set of IP addresses for internal traffic and a second set of addresses for external traffic. A NAT box located where the LAN meets the Internet makes all necessary IP address translations.

NAT serves three main purposes:

Provides a type of firewall by hiding internal IP addresses

Enables a company to use more internal IP addresses. Since they're used internally only, there's no possibility of conflict with IP addresses used by other companies and organizations.

Allows a company to combine multiple ISDN connections into a single Internet connection.

41. What is Firewall? Types of Firewall

In computing, a firewall is a network security system that controls the incoming and outgoing network traffic based on an applied rule set. A firewall establishes a barrier between a trusted, secure internal network and another network (e.g., the Internet) that is assumed not to be secure and trusted. Firewalls exist both as software to run on general purpose hardware and as a hardware appliance. Many hardware-based firewalls also offer other functionality to the internal network they protect, such as acting as a DHCP server for that network.

Conceptually, there are two types of firewalls:

- Network layer
- Application layer

They are not as different as you might think, and latest technologies are blurring the distinction to the point where it's no longer clear if either one is ``better'' or ``worse.'' As always, you need to be careful to pick the type that meets your needs.

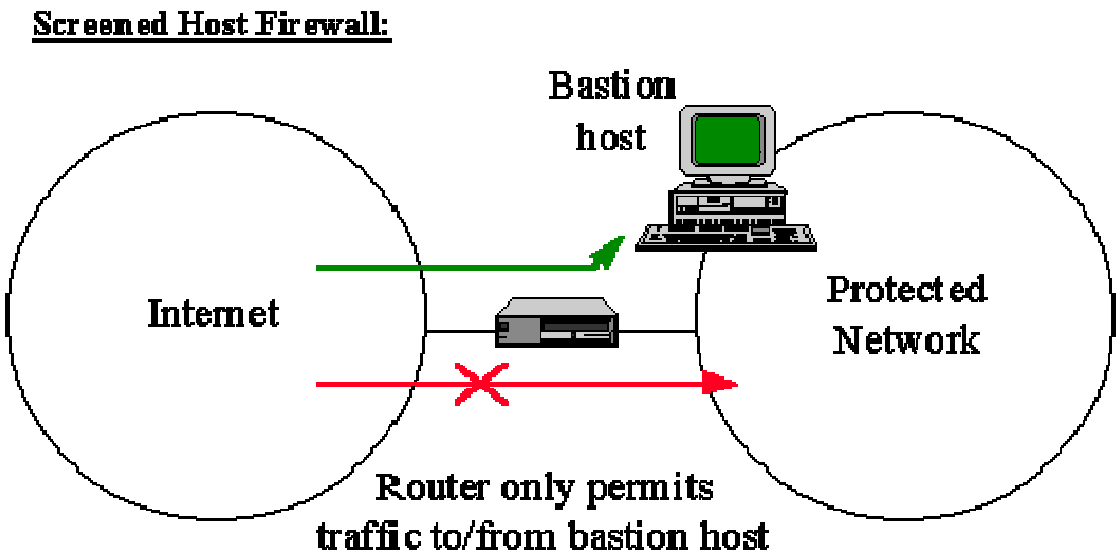
Which is which depends on what mechanisms the firewall uses to pass traffic from one security zone to another. The International Standards Organization (ISO) Open Systems Interconnect (OSI) model for networking defines seven layers, where each layer provides services that ``higher-level'' layers depend on. In order from the bottom, these layers are physical, data link, network, transport, session, presentation, application.

The important thing to recognize is that the lower-level the forwarding mechanism, the less examination the firewall can perform. Generally speaking, lower-level firewalls are faster, but are easier to fool into doing the wrong thing.

Network layer firewalls

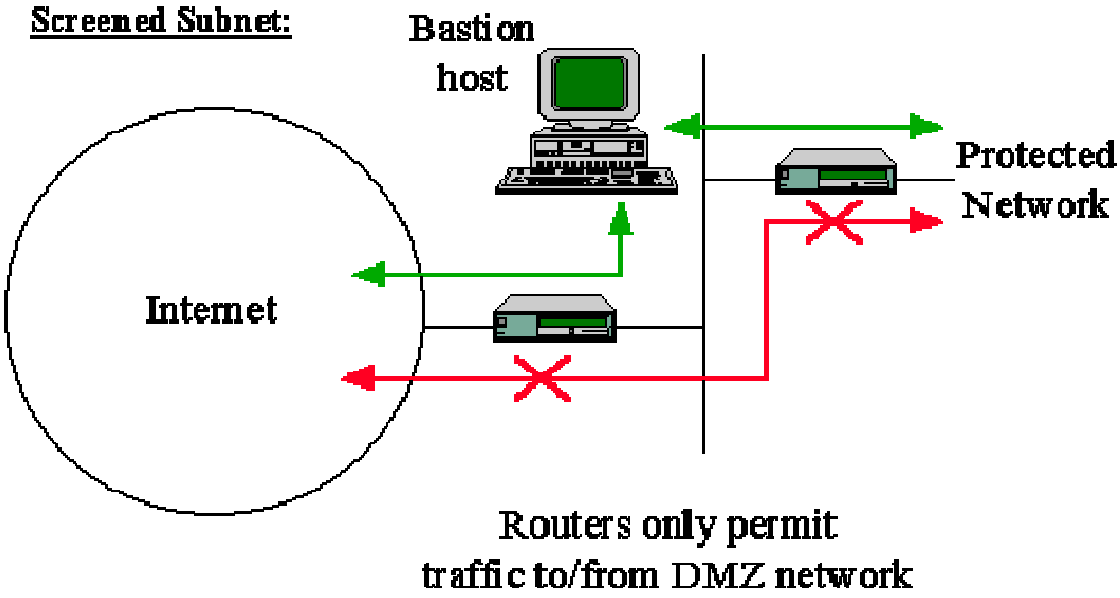
These generally make their decisions based on the source, destination addresses and ports (see Appendix C for a more detailed discussion of ports) in individual IP packets. A simple router is the ``traditional'' network layer firewall, since it is not able to make particularly sophisticated decisions about what a packet is actually talking to or where it actually came from. Modern network layer firewalls have become increasingly sophisticated, and now maintain internal information about the state of connections passing through them, the contents of some of the data streams, and so on. One thing that's an important distinction about many network layer firewalls is that they route traffic directly though them, so to use one you either need to have a validly assigned IP address block or to use a ``private internet'' address block [3]. Network layer firewalls tend to be very fast and tend to be very transparent to users.

Figure 1: Screened Host Firewall



In Figure 1, a network layer firewall called a ``screened host firewall'' is represented. In a screened host firewall, access to and from a single host is controlled by means of a router operating at a network layer. The single host is a bastion host; a highly-defended and secured strong-point that (hopefully) can resist attack.

Figure 2: Screened Subnet Firewall

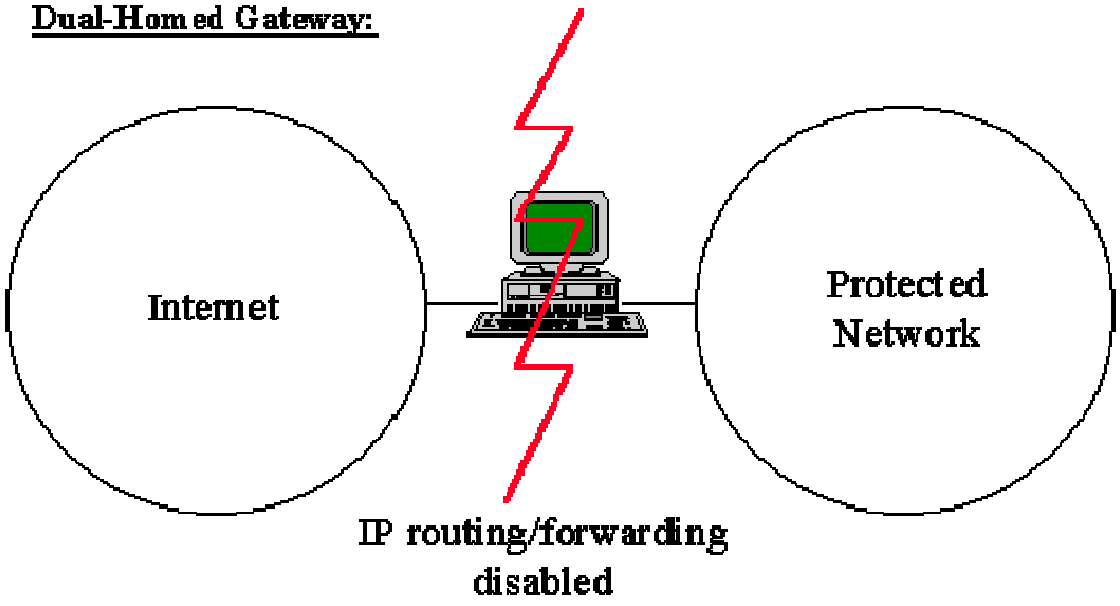


Example Network layer firewall : In figure 2, a network layer firewall called a ``screened subnet firewall" is represented. In a screened subnet firewall, access to and from a whole network is controlled by means of a router operating at a network layer. It is similar to a screened host, except that it is, effectively, a network of screened hosts.

Application layer firewalls

These generally are hosts running proxy servers, which permit no traffic directly between networks, and which perform elaborate logging and auditing of traffic passing through them. Since the proxy applications are software components running on the firewall, it is a good place to do lots of logging and access control. Application layer firewalls can be used as network address translators, since traffic goes in one ``side" and out the other, after having passed through an application that effectively masks the origin of the initiating connection. Having an application in the way in some cases may impact performance and may make the firewall less transparent. Early application layer firewalls such as those built using the TIS firewall toolkit, are not particularly transparent to end users and may require some training. Modern application layer firewalls are often fully transparent. Application layer firewalls tend to provide more detailed audit reports and tend to enforce more conservative security models than network layer firewalls.

Figure 3: Dual Homed Gateway



Example Application layer firewall : In figure 3, an application layer firewall called a ``dual homed gateway'' is represented. A dual homed gateway is a highly secured host that runs proxy software. It has two network interfaces, one on each network, and blocks all traffic passing through it.

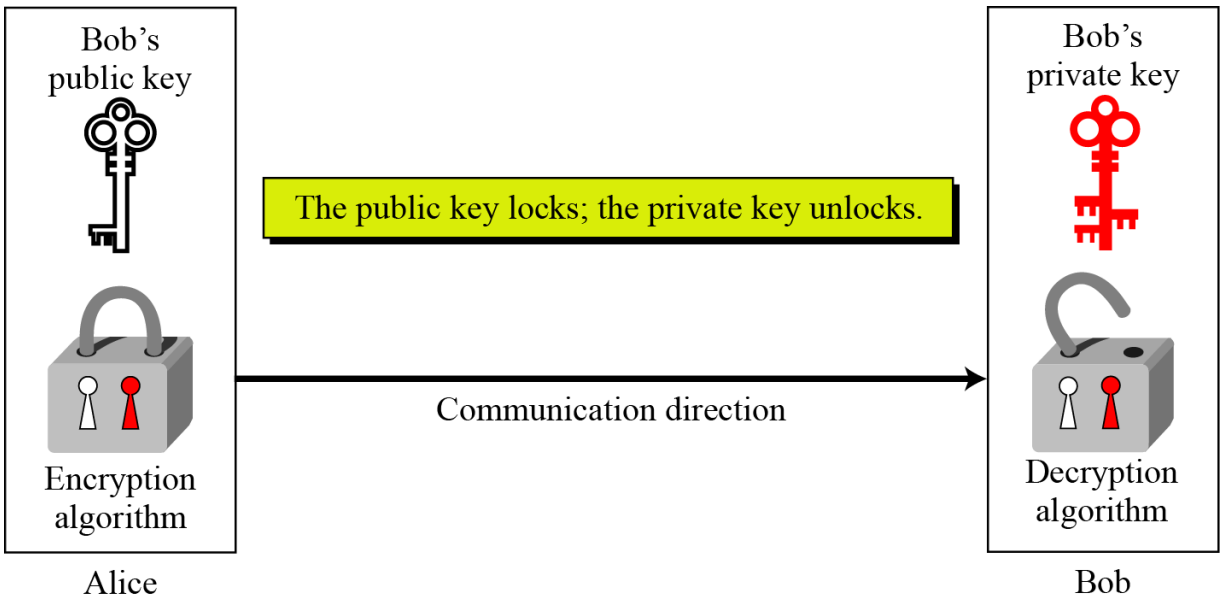
The Future of firewalls lies someplace between network layer firewalls and application layer firewalls. It is likely that network layer firewalls will become increasingly ``aware'' of the information going through them, and application layer firewalls will become increasingly ``low level'' and transparent. The end result will be a fast packet-screening system that logs and audits data as it passes through. Increasingly, firewalls (network and application layer) incorporate encryption so that they may protect traffic passing between them over the Internet. Firewalls with end-to-end encryption can be used by organizations with multiple points of Internet connectivity to use the Internet as a ``private backbone'' without worrying about their data or passwords being sniffed.

42. Explain about Asymmetric key-pair.

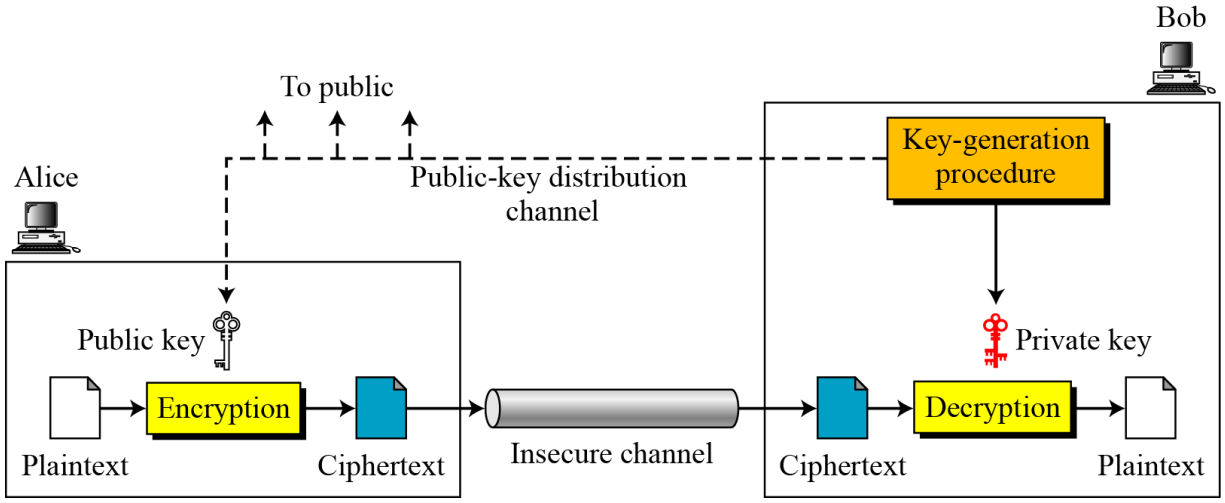
Asymmetric algorithms (public key algorithms) use different keys for encryption and decryption, and the decryption key cannot (practically) be derived from the encryption key. Asymmetric algorithms are important because they can be used for transmitting encryption keys or other data securely even when the parties have no opportunity to agree on a secret key in private.

Asymmetric key cryptography uses two separate keys: one private and one public.

Figure shows Locking and unlocking in asymmetric-key cryptosystem



General idea of asymmetric-key cryptosystem



General idea of asymmetric-key cryptosystem

Types of Asymmetric algorithms (public key algorithms):

RSA

Diffie-Hellman

Digital Signature Algorithm

ElGamal

ECDSA

XTR

Asymmetric algorithms examples:

RSA Asymmetric algorithm

Rivest-Shamir-Adleman is the most commonly used asymmetric algorithm (public key algorithm). It can be used both for encryption and for digital signatures. The security of RSA is generally considered equivalent to factoring, although this has not been proved.

RSA computation occurs with integers modulo n = p * q, for two large secret primes p, q. To encrypt a message m, it is exponentiated with a small public exponent e. For decryption, the recipient of the ciphertext c = me (mod n)computes the multiplicative reverse d = e-1 (mod (p-1)*(q-1)) (we require that e is selected suitably for it to exist) and obtains cd = m e * d = m (mod n). The private key consists of n, p, q, e, d (where p and q can be omitted); the public key contains only n and e. The problem for the attacker is that computing the reverse d of e is assumed to be no easier than factorizing n.

The key size should be greater than 1024 bits for a reasonable level of security. Keys of size, say, 2048 bits should allow security for decades. There are actually multiple incarnations of this algorithm; RC5 is one of the most common in use, and RC6 was a finalist algorithm for AES.

Diffie-Hellman

Diffie-Hellman is the first asymmetric encryption algorithm, invented in 1976, using discrete logarithms in a finite field. Allows two users to exchange a secret key over an insecure medium without any prior secrets.

Diffie-Hellman (DH) is a widely used key exchange algorithm. In many cryptographical protocols, two parties wish to begin communicating. However, let's assume they do not initially possess any common secret and thus cannot use secret key cryptosystems. The key exchange by Diffie-Hellman protocol remedies this situation by allowing the construction of a common secret key over an insecure communication channel. It is based on a problem related to discrete logarithms, namely the Diffie-Hellman problem. This problem is considered hard, and it is in some instances as hard as the discrete logarithm problem.

The Diffie-Hellman protocol is generally considered to be secure when an appropriate mathematical group is used. In particular, the generator element used in the exponentiations should have a large period (i.e. order). Usually, Diffie-Hellman is not implemented on hardware.

Digital Signature Algorithm

Digital Signature Algorithm (DSA) is a United States Federal Government standard or FIPS for digital signatures. It was proposed by the National Institute of Standards and Technology (NIST) in August 1991 for use in their Digital Signature Algorithm (DSA), specified in FIPS 186 [1], adopted in 1993. A minor revision was issued in 1996 as FIPS 186-1 [2], and the standard was expanded further in 2000 as FIPS 186-2 [3]. Digital Signature Algorithm (DSA) is similar to the one used by ElGamal signature algorithm. It is fairly efficient though not as efficient as RSA for signature verification. The standard defines DSS to use the SHA-1 hash function exclusively to compute message digests.

The main problem with DSA is the fixed subgroup size (the order of the generator element), which limits the security to around only 80 bits. Hardware attacks can be menacing to some implementations of DSS. However, it is widely used and accepted as a good algorithm.

ElGamal

The ElGamal is a public key cipher - an asymmetric key encryption algorithm for public-key cryptography which is based on the Diffie-Hellman key agreement. ElGamal is the predecessor of DSA.

ECDSA

Elliptic Curve DSA (ECDSA) is a variant of the Digital Signature Algorithm (DSA) which operates on elliptic curve groups. As with Elliptic Curve Cryptography in general, the bit size of the public key believed to be needed for ECDSA is about twice the size of the security level, in bits.

XTR

XTR is an algorithm for asymmetric encryption (public-key encryption). XTR is a novel method that makes use of traces to represent and calculate powers of elements of a subgroup of a finite field. It is based on the primitive underlying the very first public key cryptosystem, the Diffie-Hellman key agreement protocol.

From a security point of view, XTR security relies on the difficulty of solving discrete logarithm related problems in the multiplicative group of a finite field. Some advantages of XTR are its fast key generation (much faster than RSA), small key sizes (much smaller than RSA, comparable with ECC for current security settings), and speed (overall comparable with ECC for current security settings).

43. How do u design a test set?

A test set is a set of data used in various areas of information science to assess the strength and utility of a predictive relationship. Test sets are used in artificial intelligence, machine learning, genetic programming and statistics. In all these fields, a test set has much the same role.

44. Difference between C & Java

1. JAVA is Object-Oriented while C is procedural. Different Paradigms, that is.

Most differences between the features of the two languages arise due to the use of different programming paradigms. C breaks down to functions while JAVA breaks down to Objects. C is more procedure-oriented while JAVA is data-oriented.

2. Java is an Interpreted language while C is a compiled language.

We all know what a compiler does. It takes your code & translates it into something the machine can understand-that is to say-0's & 1's-the machine-level code. That's exactly what happens with our C code-it gets 'compiled'. While with JAVA, the code is first transformed to what is called the bytecode. This bytecode is then executed by the JVM(Java Virtual Machine). For the same reason, JAVA code is more portable.

3. C is a low-level language while JAVA is a high-level language.

C is a low-level language(difficult interpretation for the user, closer significance to the machine-level code) while JAVA is a high-level lagunage(abstracted from the machine-level details, closer significance to the program itself).

4. C uses the top-down {sharp & smooth} approach while JAVA uses the bottom-up {on the rocks} approach.

In C, formulating the program begins by defining the whole and then splitting them into smaller elements. JAVA(and C++ and other OOP languages) follows the bottom-up approach where the smaller elements combine together to form the whole.

5. Pointer go backstage in JAVA while C requires explicit handling of pointers.

When it comes to JAVA, we don't need the *'s & &'s to deal with pointers & their addressing. More formally, there is no pointer syntax required in JAVA. It does what it needs to do. While in JAVA, we do create references for objects.

6. The Behind-the-scenes Memory Management with JAVA & The User-Based Memory Management in C.

Remember 'malloc' & 'free'? Those are the library calls used in C to allocate & free chunks of memory for specific data(specified using the keyword 'sizeof'). Hence in C, the memory is managed by the user while JAVA uses a garbage collector that deletes the objects that no longer have any references to them.

7. JAVA supports Method Overloading while C does not support overloading at all.

JAVA supports function or method overloading-that is we can have two or more functions with the same name(with certain varying parameters like return types to allow the machine to differentiate between them). That it to say, we can overload methods with the same name having different method signatures. JAVA(unlike C++), does not support Operator Overloading while C does not allow overloading at all.

8. Unlike C, JAVA does not support Preprocessors, & does not really them.

The preprocessor directives like #include & #define, etc are considered one of the most essential elements of C programming. However, there are no preprocessors in JAVA. JAVA uses other alternatives for the preprocessors. For instance, public static final is used instead of the #define preprocessor. Java maps class names to a directory and file structure instead of the #include used to include files in C.

9. The standard Input & Output Functions.

Although this difference might not hold any conceptual(intuitive) significance, but it's maybe just the tradition. C uses the printf & scanf functions as its standard input & output while JAVA uses the System.out.print & System.in.read functions.

10. Exception Handling in JAVA And the errors & crashes in C.

When an error occurs in a Java program it results in an exception being thrown. It can then be handled using various exception handling techniques. While in C, if there's an error, there IS an error

C++	Java
Extension of C with object-oriented programming, still able to run C code.	Strongly influenced by C++/C syntax.
Compatible with C source code, except for a few corner cases.	Provides the Java Native Interface and recently Java Native Access as a way to directly call C/C++ code.
Write once, compile anywhere (WOCA).	Write once, run anywhere / everywhere (WORA / WORE).
Allows procedural programming, functional programming, object-oriented programming, generic programming, and template metaprogramming. Favors a mix of paradigms.	Allows procedural programming, functional programming (since Java 8) and generic programming (since Java 5), but strongly encourages the object-orientedprogramming paradigm. Includes support for the creation of scripting languages.
Runs as native executable machine code for the target instruction set(s).	Runs in a virtual machine.
Provides object types and type names. Allows reflection through RTTI.	Is reflective, allowing metaprogramming and dynamic code generation at runtime.
Has multiple binary compatibility standards (commonly Microsoft (for MSVC compiler) and Itanium/GNU (for virtually all other compilers)).	Has a single, OS- and compiler-independent binary compatibility standard.
Optional automated bounds checking (e.g., the at() method in vector and stringcontainers).	All operations are required to be Bound-checked by all compliant distributions of Java.HotSpot can remove bounds checking.

Supports native unsigned arithmetic.	No native support for unsigned arithmetic.
Standardized minimum limits for all numerical types, but the actual sizes are implementation-defined. Standardized types are available through the standard library `<cstdint>`.	Standardized limits and sizes of all primitive types on all platforms.
Pointers, references, and pass-by-value are supported for all types (primitive or user-defined). Memory management can be done manually through `new` / `delete`, automatically by scope, or by smart pointers. Supports deterministic destruction of objects. Garbage collection ABI standardized in C++11, though compilers are not required to implementgarbage collection.	All types (primitive types and reference types) are always passed by value.[1] Automatic garbage collection. Supports a non-deterministic finalize() method whose use is not recommended.[2]
Resource management can be done manually or by automatic lifetime-based resource management (RAII).	Resource management must be done manually, or automatically via finalizers, though this is generally discouraged. Has try-with-resources for automatic scope-based resource management (version 7 onwards).
Supports classes, structs (POD-types), and unions, and can allocate them on the heapor the stack.	Classes are allocated on the heap. Java SE 6 optimizes with escape analysis to allocate some objects on the stack.
Allows explicitly overriding types as well as some implicit narrowing conversions (for compatibility with C).	Rigid type safety except for widening conversions.
The C++ Standard Library was designed to have a limited scope and functionality but	The standard library has grown with each release. By version 1.6, the library included

includes language support, diagnostics, general utilities, strings, locales, containers, algorithms, iterators, numerics, input/output, random number generators, regular expression parsing, threading facilities, type traits (for static type introspection) and Standard C Library. The Boost library offers more functionality including network I/O. A rich amount of third-party libraries exist for GUI and other functionalities like: ACE,Crypto++, various XMPP Instant Messaging (IM) libraries,[3] OpenLDAP, Qt, gtkmm.	support for locales, logging, containers and iterators, algorithms, GUI programming (but not using the system GUI), graphics, multi-threading, networking, platform security, introspection, dynamic class loading, blocking and non-blocking I/O. It provided interfaces or support classes for XML, XSLT, MIDI, database connectivity, naming services (e.g. LDAP), cryptography, security services (e.g. Kerberos), print services, and web services. SWT offers an abstraction for platform-specific GUIs.
Operator overloading for most operators. Preserving meaning (semantics) is highly recommended.	Operators are not overridable. The language overrides + and += for the String class.
Single and Multiple inheritance of classes, including virtual inheritance.	Single inheritance of classes. Supports multiple inheritance via the Interfacesconstruct, which is equivalent to a C++ class composed of abstract methods.
Compile-time templates. Allows for Turing complete meta-programming.	Generics are used to achieve basic type-parametrization, but they do not translate from source code to byte code due to the use of type erasure by the compiler.
Function pointers, function objects, lambdas (in C++11), and interfaces.	References to functions achieved via the reflection API. OOP idioms using Interfaces, such as Adapter, Observer, and Listener are generally preferred over direct references to methods.
No standard inline documentation mechanism.	Extensive Javadoc documentation standard on all

Third-party software (e.g. Doxygen) exists.	system classes and methods.
const keyword for defining immutable variables and member functions that do not change the object. Const-ness is propagated as a means to enforce, at compile-time, correctness of the code with respect to mutability of objects (see const-correctness).	final provides a version of const, equivalent to type* const pointers for objects and const for primitive types. Immutability of object members achieved through read-only interfaces and object encapsulation.
Supports the goto statement.	Supports labels with loops and statement blocks.
Source code can be written to be platform-independent (can be compiled for Windows,BSD, Linux, Mac OS X, Solaris, etc., without modification) and written to take advantage of platform-specific features. Typically compiled into native machine code, must be re-compiled for each target platform.	Compiled into byte code for the JVM. Byte code is dependent on the Java platform, but is typically independent of operating system specific features.

45. What is a simulator?

    Computer program (such as a game or animated flowchart) or a dedicated device that models (simulates) some aspects of a real life situation (such as flying an aircraft) and can be manipulated to observe the outcomes of different assumptions or actions, without exposing the experimenter to any danger or risk.
    Simulation is the imitation of the operation of a real-world process or system over time.[1] The act of simulating something first requires that a model be developed; this model represents the key characteristics or behaviors/functions of the selected physical or abstract system or process. The model represents the system itself, whereas the simulation represents the operation of the system over time.

46. What do u mean by statistical quality?

    Statistical process control (SPC) is a method of quality control which uses statistical methods. SPC is applied in order to monitor and control a process. Monitoring and controlling the process ensures that it operates at its full potential. At its full potential, the process can make as much conforming product as possible with a minimum (if not an elimination) of waste (rework or scrap). SPC can be applied to any process where the "conforming product" (product meeting

specifications) output can be measured. Key tools used in SPC include control charts; a focus on continuous improvement; and the design of experiments.

47. What is Ethernet? What are its specifications (length, data rate)? How are the nodes connected in ethernet?

    A local-area network (LAN) architecture developed by Xerox Corporation in cooperation with DEC and Intel in 1976. Ethernet uses a bus or star topology and supports data transfer rates of 10 Mbps. The Ethernet specification served as the basis for the IEEE 802.3 standard, which specifies the physical and lower software layers. Ethernet uses the CSMA/CD access method to handle simultaneous demands. It is one of the most widely implemented LAN standards.

    A newer version of Ethernet, called 100Base-T (or Fast Ethernet), supports data transfer rates of 100 Mbps. And the newest version, Gigabit Ethernet supports data rates of 1 gigabit (1,000 megabits) per second.

    Ethernet is a family of computer networking technologies for local area networks (LANs) and metropolitan area networks (MANs). It was commercially introduced in 1980 and first standardized in 1983 as IEEE 802.3, and has since been refined to support higher bit rates and longer link distances. Over time, Ethernet has largely replaced competing wired LAN technologies such as token ring, FDDI, and ARCNET. The primary alternative for contemporary LANs is not a wired standard, but instead a wireless LAN standardized as IEEE 802.11 and also known as Wi-Fi.

    Over the course of its history, Ethernet data transfer rates have been increased from the original 3 megabits per second (Mbit/s) to the latest 100 gigabits per second (Gbit/s), with 400 Gbit/s expected by early 2017.

48. What is TCP/IP?

    TCP/IP (Transmission Control Protocol/Internet Protocol) is the basic communication language or protocol of the Internet. It can also be used as a communications protocol in a private network (either an intranet or an extranet). When you are set up with direct access to the Internet, your computer is provided with a copy of the TCP/IP program just as every other computer that you may send messages to or get information from also has a copy of TCP/IP.

49. What are the characteristics of real-time systems?

    Large and complex - vary from a few hundred lines of assembler or C to 20 million lines of Ada estimated for the Space Station Freedom

    Concurrent control of separate system components - devices operate in parallel in the real-world; better to model this parallelism by concurrent entities in the program

    Facilities to interact with special purpose hardware - need to be able to program devices in a reliable and abstract way

Mixture of Hardware/Software - some modules implemented in hardware, even whole systems, SoC

Extreme reliability and safety - embedded systems typically control the environment in which they operate; failure to control can result in loss of life, damage to environment or economic loss

Guaranteed response times - we need to be able to predict with confidence the worst case response times for systems; efficiency is important but predictability is essential

Cloud Computing

Cloud computing is a general term for anything that involves delivering hosted services over the Internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). The name cloud was inspired by the symbol that's often used to represent the Internet in flowcharts and diagrams.

Cloud computing enables companies to consume compute resources as a utility -- just like electricity -- rather than having to build and maintain computing infrastructures in-house.

Cloud computing promises several attractive benefits for businesses and end users. Three of the main benefits of cloud computing include:

- Self-service provisioning: End users can spin up computing resources for almost any type of workload on-demand.
- Elasticity: Companies can scale up as computing needs increase and then scale down again as demands decrease.
- Pay per use: Computing resources are measured at a granular level, allowing users to pay only for the resources and workloads they use.

Cloud computing services can be private, public, or hybrid.

Private cloud services are delivered from a business' data center to internal users. This model offers versatility and convenience, while preserving management, control and security. Internal customers may or may not be billed for services through IT chargeback.

In the public cloud model, a third-party provider delivers the cloud service over the Internet. Public cloud services are sold on-demand, typically by the minute or the hour. Customers only pay for the CPU cycles, storage or bandwidth they consume. Leading public cloud providers include Amazon Web Services (AWS), Microsoft Azure, IBM/SoftLayer and Google Compute Engine.

Hybrid cloud is a combination of public cloud services and on-premises private cloud – with orchestration and automation between the two. Companies can run mission-critical workloads or sensitive applications on the private cloud while using the public cloud for bursty workloads that must scale on-demand. The goal of hybrid cloud is to create a unified, automated, scalable environment which takes advantage of all that a public cloud infrastructure can provide, while still maintaining control over mission-critical data.

Although cloud computing has changed over time, it has always been divided into three broad service categories: infrastructure as a service (IaaS), platform as a service (PaaS) and software as service (SaaS).

IaaS providers such as AWS supply a virtual server instance and storage, as well as application program interfaces (APIs) that let users migrate workloads to a virtual machine (VM). Users have an allocated storage capacity and start, stop, access and configure the VM and storage as desired. IaaS providers offer small, medium, large, extra-large, and memory- or compute-optimized instances, in addition to customized instances, for various workload needs.

In the PaaS model, providers host development tools on their infrastructures. Users access those tools over the Internet using APIs, Web portals or gateway software. PaaS is used for general software development and many PaaS providers will host the software after it's developed. Common PaaS providers include Salesforce.com's Force.com, Amazon Elastic Beanstalk and Google App Engine.

SaaS is a distribution model that delivers software applications over the Internet; these are often called Web services. Microsoft Office 365 is a SaaS offering for productivity software and email services. Users can access SaaS applications and services from any location using a computer or mobile device that has Internet access..

Grid Computing

Grid computing is a distributed architecture of large numbers of computers connected to solve a complex problem. In the grid computing model, servers or personal computers run independent tasks and are loosely linked by the Internet or low-speed networks. Computers may connect directly or via scheduling systems.

Most applications for grid computing projects have no time dependency, and large projects typically deploy across many countries and continents. Search programs and others use the idle power of computers, also known as cycle-scavenging, running in the background for many weeks.

Utility Computing

Conventional Internet hosting services have the capability to quickly arrange for the rental of individual servers, for example to provision a bank of web servers to accommodate a sudden surge in traffic to a web site.

“Utility computing” usually envisions some form of virtualization so that the amount of storage or computing power available is considerably larger than that of a single time-sharing computer. Multiple servers are used on the “back end” to make this possible. These might be a dedicated computer cluster specifically built for the purpose of being rented out, or even an under-utilized supercomputer. The technique of running a single calculation on multiple computers is known as distributed computing.

Distributed Computing

A method of computer processing in which different parts of a program are run simultaneously on two or more computers that are communicating with each other over a network. Distributed computing is a type of segmented or parallel computing, but the latter term is most commonly used to refer to processing in which different parts of a program run simultaneously on two or more processors that are part of the same computer. While both types of processing require that a program be segmented divided into sections that can run simultaneously, distributed computing also requires that the division of the program take into account the different environments on which the different sections of the program will be running. For example, two computers are likely to have different file systems and different hardware components.

Cluster Computing

A computer cluster is a group of linked computers, working together closely so that in many respects they form a single computer. The components of a cluster are commonly, but not always, connected to each other through fast local area networks. Clusters are usually deployed to improve performance and/or availability over that provided by a single computer, while typically being much more cost-effective than single computers of comparable speed or availability. Clustering is used for parallel processing, load balancing and fault tolerance.

50. Name any requirements other than functional requirements? Is there any non-functional requirement?

Domain requirements

- Describe system characteristics and features that reflect the domain
- May be new functional requirements, constraints on existing requirements or may define specific computations
- If domain requirements are not satisfied, the system may be unworkble
- Example: Library system

Non-functional requirements

- Product requirements – Requirements which specify that the delivered product must behave in a particular way, e.g. execution speed, reliability etc.
- Organisational requirements – Requirements which are a consequence of organisational policies and procedures, e.g. process standards used, implementation requirements etc.
- External requirements – Requirements which arise from factors which are external to the system and its development process, e.g. interoperability requirements, legislative requirements etc.

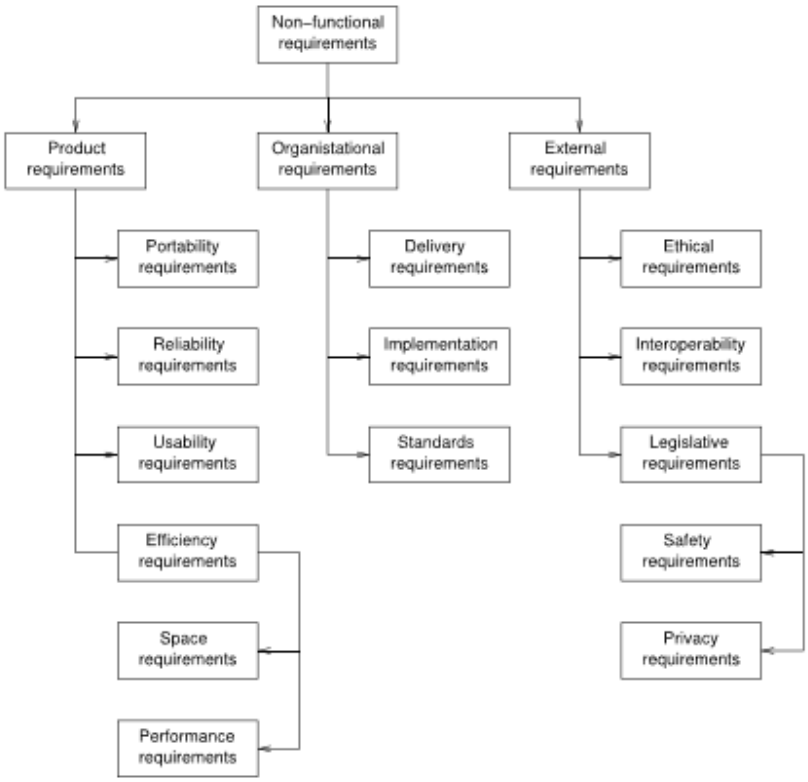


Figure: Non-functional requirements

51. Why do we need to have network address? Is the MAC address not sufficient? explain

    IP address (network address) is just used for transferring information from one network to another network. Travelling of information among networks uses IP addresses. It is used to identity the network and host.

    A MAC address (physical addresses) is used for distribution of information with in the network segment. The Ethernet uses MAC address to transfer data between hosts. When its used with IP network, the IP address is resolved using ARP protocol to find the MAC address of the end device and the data is transmitted.

    Port numbers are used by the TCP/UDP protocol to isolate the traffic which is multiplexed and sent by the user application. For example, the user device, can open multiple applications at the same time like, multiple web browsers, email and FTP. To identify the data individually the port number are used.

52. What is the file system used by Linux? What are the features of ext3 file system?

    The ext2 or second extended filesystem is a file system for the Linux kernel. It was initially designed by Rémy Card as a replacement for the extended file system (ext). Its metadata structure was inspired by the earlier Unix File System (UFS).

ext2 was the default filesystem in several Linux distributions, including Debian and Red Hat Linux, until supplanted more recently by ext3, which is almost completely compatible with ext2 and is a journaling file system. ext2 is still the filesystem of choice for flash-based storage media (such as SD cards, and USB flash drives), since its lack of a journal increases performance and minimizes the number of writes, and flash devices have a limited number of write cycles. Recent kernels, however, support a journal-less mode of ext4, which would offer the same benefit, along with a number of ext4-specific benefits.

ext3, or third extended filesystem, is a journaled file system that is commonly used by the Linux kernel. It is the default file system for many popular Linux distributions. Stephen Tweedie first revealed that he was working on extending ext2 in Journaling the Linux ext2fs Filesystem in a 1998 paper, and later in a February 1999 kernel mailing list posting. The filesystem was merged with the mainline Linux kernel in November 2001 from 2.4.15 onward.[2][3][4] Its main advantage over ext2 is journaling, which improves reliability and eliminates the need to check the file system after an unclean shutdown. Its successor is ext4.

53. How do you do testing? What is White, Black and gray box testing?

Black-Box Testing

The technique of testing without having any knowledge of the interior workings of the application is called black-box testing. The tester is oblivious to the system architecture and does not have access to the source code. Typically, while performing a black-box test, a tester will interact with the system's user interface by providing inputs and examining outputs without knowing how and where the inputs are worked upon.

White-Box Testing

White-box testing is the detailed investigation of internal logic and structure of the code. White-box testing is also called glass testing or open-box testing. In order to performwhite-box testing on an application, a tester needs to know the internal workings of the code.

Grey-Box Testing

Grey-box testing is a technique to test the application with having a limited knowledge of the internal workings of an application. In software testing, the phrase the more you know, the better carries a lot of weight while testing an application.

54. Why do you do testing?

Software testing is very important because of the following reasons:

- Software testing is really required to point out the defects and errors that were made during the development phases.
- It’s essential since it makes sure of the Customer’s reliability and their satisfaction in the application.

- It is very important to ensure the Quality of the product. Quality product delivered to the customers helps in gaining their confidence.
- Testing is necessary in order to provide the facilities to the customers like the delivery of high quality product or software application which requires lower maintenance cost and hence results into more accurate, consistent and reliable results.
- Testing is required for an effective performance of software application or product.
- It's important to ensure that the application should not result into any failures because it can be very expensive in the future or in the later stages of the development.
- It's required to stay in the business.

55. What is PERT?

    PERT is a method to analyze the involved tasks in completing a given project, especially the time needed to complete each task, and to identify the minimum time needed to complete the total project. It was developed primarily to simplify the planning and scheduling of large and complex projects.

56. Mysql, Oracle Sybase follows under which database?

    Relational DBMS

57. When a database does really become relational? (when it follows 9 or more codds rule)

58. How to generate test plan? (along with SRS and SDD)
59. Exact definition of reliability
    Reliability the ability of a system or component to perform its required functions under stated conditions for a specified period of time.
60. What is URL? What is the difference between website and URL?

    A URL is one type of Uniform Resource Identifier (URI); the generic term for all types of names and addresses that refer to objects on the World Wide Web. The term "Web address" is a synonym for a URL that uses the HTTP / HTTPS protocol.

    Difference between website and URL

    - URL stands for Uniform Resource Locator, it is the "technical term" for the address of a web page.
    - Website refers to a collection of web pages, which of course also means it refers to a collection of URLs.
    - So when someone says "Yahoo" they usually mean web site and if someone says "www.rainy-day-games.com" then usually mean URL.

61. Why you have used jsp for programming? What is the difference between jsp and Servlet?

There are plenty advantages of using JSP. In general, JSP allows developers to easily distribute application functionality to a wide range of page authors. These authors do not have to know the Java programming language or know anything about writing servlet code, so they can concentrate on writing their HTML code while you concentrate on creating your objects and application logic.

- JSP pages easily combine static templates, including HTML or XML fragments, with code that generates dynamic content.
- JSP pages are compiled dynamically into servlets when requested, so page authors can easily make updates to presentation code. JSP pages can also be precompiled if desired.
- JSP tags for invoking JavaBeans components manage these components completely, shielding the page author from the complexity of application logic.
- Developers can offer customized JSP tag libraries that page authors access using an XML-like syntax.
- Web authors can change and edit the fixed template portions of pages without affecting the application logic. Similarly, developers can make logic changes at the component level without editing the individual pages that use the logic.

Difference between jsp and Servlet

JSP	Servlets
JSP is a webpage scripting language that can generate dynamic content.	Servlets are Java programs that are already compiled which also creates dynamic web content.
JSP run slower compared to Servlet as it takes compilation time to convert into Java Servlets.	Servlets run faster compared to JSP.
It's easier to code in JSP than in Java Servlets.	Its little much code to write here.
In MVC, jsp act as a view.	In MVC, servlet act as a controller.

JSP are generally preferred when there is not much processing of data required.	servlets are best for use when there is more processing and manipulation involved.
The advantage of JSP programming over servlets is that we can build custom tags which can directly call Java beans.	There is no such custom tag facility in servlets.
We can achieve functionality of JSP at client side by running JavaScript at client side.	There are no such methods for servlets.

Jsp or Servlets ? which is best…??

While JSP may be great for serving up dynamic Web content and separating content from presentation, some may still wonder why servlets should be cast aside for JSP. The utility of servlets is not in question. They are excellent for server-side processing, and, with their significant installed base, are here to stay. In fact, architecturally speaking, you can view JSP as a high-level abstraction of servlets that is implemented as an extension of the Servlet 2.1 API. Still, you shouldn't use servlets indiscriminately; they may not be appropriate for everyone. For instance, while page designers can easily write a JSP page using conventional HTML or XML tools, servlets are more suited for back-end developers because they are often written using an IDE -- a process that generally requires a higher level of programming expertise.

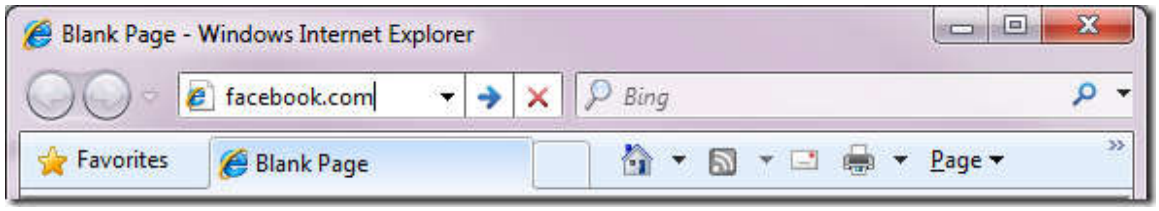
When deploying servlets, even developers have to be careful and ensure that there is no tight coupling between presentation and content. You can usually do this by adding a third-party HTML wrapper package like htmlKona to the mix. But even this approach, though providing some flexibility with simple screen changes, still does not shield you from a change in the presentation format itself. For example, if your presentation changed from HTML to DHTML, you would still need to ensure that wrapper packages were compliant with the new format. In a worst-case scenario, if a wrapper package is not available, you may end up hardcoding the presentation within the dynamic content. So, what is the solution? One approach would be to use both JSP and servlet technologies for building application systems.

62. When you type URL, how you get the webpage on your browser? (Whole process, DNS…..)

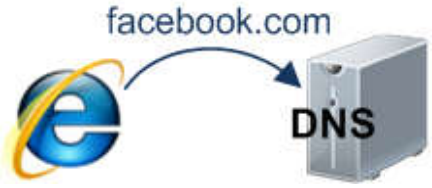
In this article, we will take a deeper look at the sequence of events that take place when you visit a URL.

1. You enter a URL into the browser

It all starts here:



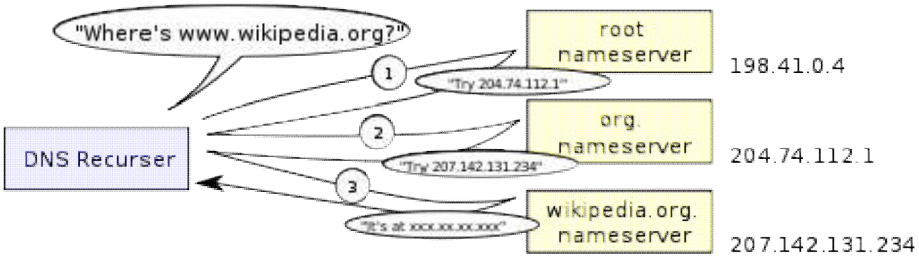
2. The browser looks up the IP address for the domain name



The first step in the navigation is to figure out the IP address for the visited domain. The DNS lookup proceeds as follows:

- **Browser cache –** The browser caches DNS records for some time. Interestingly, the OS does not tell the browser the time-to-live for each DNS record, and so the browser caches them for a fixed duration (varies between browsers, 2 – 30 minutes).
- **OS cache** – If the browser cache does not contain the desired record, the browser makes a system call (gethostbyname in Windows). The OS has its own cache.
- **Router cache** – The request continues on to your router, which typically has its own DNS cache.
- **ISP DNS cache** – The next place checked is the cache ISP's DNS server. With a cache, naturally.
- **Recursive search** – Your ISP's DNS server begins a recursive search, from the root nameserver, through the .com top-level nameserver, to Facebook's nameserver. Normally, the DNS server will have names of the .com nameservers in cache, and so a hit to the root nameserver will not be necessary.

Here is a diagram of what a recursive DNS search looks like:



One worrying thing about DNS is that the entire domain like wikipedia.org or facebook.com seems to map to a single IP address. Fortunately, there are ways of mitigating the bottleneck:

- **Round-robin DNS** is a solution where the DNS lookup returns multiple IP addresses, rather than just one. For example, facebook.com actually maps to four IP addresses.
- **Load-balancer** is the piece of hardware that listens on a particular IP address and forwards the requests to other servers. Major sites will typically use expensive high-performance load balancers.
- **Geographic DNS** improves scalability by mapping a domain name to different IP addresses, depending on the client's geographic location. This is great for hosting static content so that different servers don't have to update shared state.
- **Anycast** is a routing technique where a single IP address maps to multiple physical servers. Unfortunately, anycast does not fit well with TCP and is rarely used in that scenario.

Most of the DNS servers themselves use anycast to achieve high availability and low latency **of the DNS lookups**.

3. The browser sends a HTTP request to the web server

You can be pretty sure that Facebook's homepage will not be served from the browser cache because dynamic pages expire either very quickly or immediately (expiry date set to past).

The GET request names the **URL** to fetch**:** "http://facebook.com/". The browser identifies itself (**User-Agent** header), and states what types of responses it will accept (**Accept** and **Accept-Encoding**headers). The **Connection** header asks the server to keep the TCP connection open for further requests.

The request also contains the **cookies** that the browser has for this domain. As you probably already know, cookies are key-value pairs that track the state of a web site in between different page requests. And so the cookies store the name of the logged-in user, a secret number that was assigned to the user by the server, some of user's settings, etc. The cookies will be stored in a text file on the client, and sent to the server with every request.

There is a variety of tools that let you view the raw HTTP requests and corresponding responses. My favorite tool for viewing the raw HTTP traffic is fiddler, but there are many other tools (e.g., FireBug) These tools are a great help when optimizing a site.

In addition to GET requests, another type of requests that you may be familiar with is a POST request, typically used to submit forms. A GET request sends its parameters via the URL (e.g.: http://robozzle.com/puzzle.aspx**?id=85**). A POST request sends its parameters in the request body, just under the headers.

The trailing slash in the URL "http://facebook.com/" is important. In this case, the browser can safely add the slash. For URLs of the form http://example.com/folderOrFile, the browser cannot automatically add a slash, because it is not clear whether folderOrFile is a folder or a file. In such cases, the browser will visit the URL without the slash, and the server will respond with a redirect, resulting in an unnecessary roundtrip.

4. The facebook server responds with a permanent redirect

This is the response that the Facebook server sent back to the browser request:

The server responded with a 301 Moved Permanently response to tell the browser to go to "http://www.facebook.com/" instead of "http://facebook.com/".

There are interesting reasons why the server insists on the redirect instead of immediately responding with the web page that the user wants to see.

One reason has to do with **search engine rankings**. See, if there are two URLs for the same page, say http://www.igoro.com/ and http://igoro.com/, search engine may consider them to be two different sites, each with fewer incoming links and thus a lower ranking. Search engines understand permanent redirects (301), and will combine the incoming links from both sources into a single ranking.

Also, multiple URLs for the same content are not **cache-friendly**. When a piece of content has multiple names, it will potentially appear multiple times in caches.

5. The browser follows the redirect

The browser now knows that “http://www.facebook.com/” is the correct URL to go to, and so it sends out another GET request:

The meaning of the headers is the same as for the first request.

6. The server ‘handles’ the request

The server will receive the GET request, process it, and send back a response.

This may seem like a straightforward task, but in fact there is a lot of interesting stuff that happens here – even on a simple site like my blog, let alone on a massively scalable site like facebook.

- **Web server software:** The web server software (e.g., IIS or Apache) receives the HTTP request and decides which request handler should be executed to handle this request. A request handler is a program (in ASP.NET, PHP, Ruby, …) that reads the request and generates the HTML for the response.

  In the simplest case, the request handlers can be stored in a file hierarchy whose structure mirrors the URL structure, and so for example http://example.com/folder1/page1.aspx URL will map to file /httpdocs/folder1/page1.aspx. The web server software can also be configured so that URLs are manually mapped to request handlers, and so the public URL of page1.aspx could behttp://example.com/folder1/page1.

- **Request handler**: The request handler reads the request, its parameters, and cookies. It will read and possibly update some data stored on the server. Then, the request handler will generate a HTML response.

One interesting difficulty that every dynamic website faces is how to store data. Smaller sites will often have a single SQL database to store their data, but sites that store a large amount of data and/or have many visitors have to find a way to split the database across multiple machines. Solutions include sharding (splitting up a table across multiple databases based on the primary key), replication, and usage of simplified databases with weakened consistency semantics.

One technique to keep data updates cheap is to defer some of the work to a batch job. For example, Facebook has to update the newsfeed in a timely fashion, but the data backing the “People you may know” feature may only need to be updated nightly (my guess, I don’t actually know how they implement this feature). Batch job updates result in staleness of some less important data, but can make data updates much faster and simpler.

7. The server sends back a HTML response

Here is the response that the server generated and sent back:

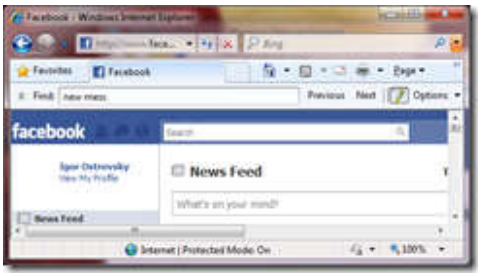
HTTP/1.1 200 OK

In addition to compression, headers specify whether and how to cache the page, any cookies to set (none in this response), privacy information, etc.

Notice the header that sets **Content-Type** to **text/html**. The header instructs the browser to render the response content as HTML, instead of say downloading it as a file. The browser will use the header to decide how to interpret the response, but will consider other factors as well, such as the extension of the URL.

8. The browser begins rendering the HTML

Even before the browser has received the entire HTML document, it begins rendering the website:



9. The browser sends requests for objects embedded in HTML

As the browser renders the HTML, it will notice tags that require fetching of other URLs. The browser will send a GET request to retrieve each of these files.

Here are a few URLs that my visit to facebook.com retrieved:

- **Images**
  http://static.ak.fbcdn.net/rsrc.php/z12E0/hash/8q2anwu7.gif
  http://static.ak.fbcdn.net/rsrc.php/zBS5C/hash/7hwy7at6.gif
  …
- **CSS style sheets**
- http://static.ak.fbcdn.net/rsrc.php/z448Z/hash/2plh8s4n.css
  http://static.ak.fbcdn.net/rsrc.php/zANE1/hash/cvtutcee.css
  …
- **JavaScript files**
- http://static.ak.fbcdn.net/rsrc.php/zEMOA/hash/c8yzb6ub.js
  http://static.ak.fbcdn.net/rsrc.php/z6R9L/hash/cq2lgbs8.js
  …

Each of these URLs will go through process a similar to what the HTML page went through. So, the browser will look up the domain name in DNS, send a request to the URL, follow redirects, etc.

However, static files – unlike dynamic pages – allow the browser to cache them. Some of the files may be served up from cache, without contacting the server at all. The browser knows how long to cache a particular file because the response that returned the file contained an Expires header. Additionally, each response may also contain an ETag header that works like a version number – if the browser sees an ETag for a version of the file it already has, it can stop the transfer immediately.

Can you guess what **"fbcdn.net"** in the URLs stands for? A safe bet is that it means "Facebook content delivery network". Facebook uses a content delivery network (CDN) to distribute static content – images, style sheets, and JavaScript files. So, the files will be copied to many machines across the globe.

Static content often represents the bulk of the bandwidth of a site, and can be easily replicated across a CDN. Often, sites will use a third-party CDN provider, instead of operating a CND themselves. For example, Facebook's static files are hosted by Akamai, the largest CDN provider.

As a demonstration, when you try to ping static.ak.fbcdn.net, you will get a response from an akamai.net server. Also, interestingly, if you ping the URL a couple of times, may get responses from different servers, which demonstrates the load-balancing that happens behind the scenes.

10. The browser sends further asynchronous (AJAX) requests

In the spirit of Web 2.0, the client continues to communicate with the server even after the page is rendered.

For example, Facebook chat will continue to update the list of your logged in friends as they come and go. To update the list of your logged-in friends, the JavaScript executing in your browser has to send an asynchronous request to the server. The asynchronous request is a programmatically constructed GET or POST request that goes to a special URL. In the Facebook example, the client sends a POST request to http://www.facebook.com/ajax/chat/buddy_list.php to fetch the list of your friends who are online.

This pattern is sometimes referred to as “AJAX”, which stands for “Asynchronous JavaScript And XML”, even though there is no particular reason why the server has to format the response as XML. For example, Facebook returns snippets of JavaScript code in response to asynchronous requests.

Among other things, the fiddler tool lets you view the asynchronous requests sent by your browser. In fact, not only you can observe the requests passively, but you can also modify and resend them. The fact that it is this easy to “spoof” AJAX requests causes a lot of grief to developers of online games with scoreboards. (Obviously, please don’t cheat that way.)

Facebook chat provides an example of an interesting problem with AJAX: pushing data from server to client. Since HTTP is a request-response protocol, the chat server cannot push new messages to the client. Instead, the client has to poll the server every few seconds to see if any new messages arrived.

Long polling is an interesting technique to decrease the load on the server in these types of scenarios. If the server does not have any new messages when polled, it simply does not send a response back. And, if a message for this client is received within the timeout period, the server will find the outstanding request and return the message with the response.

63. What is search engine?

Search engine is a program that searches for and identifies items in a database that correspond to keywords or characters specified by the user, used especially for finding particular sites on the World Wide Web.

64. What is deep web?

Deep Web (also called the Deepnet), [1] Invisible Web,[2] or Hidden Web[3] is the portion of World Wide Web content that is not indexed by standard search engines.

(Or)

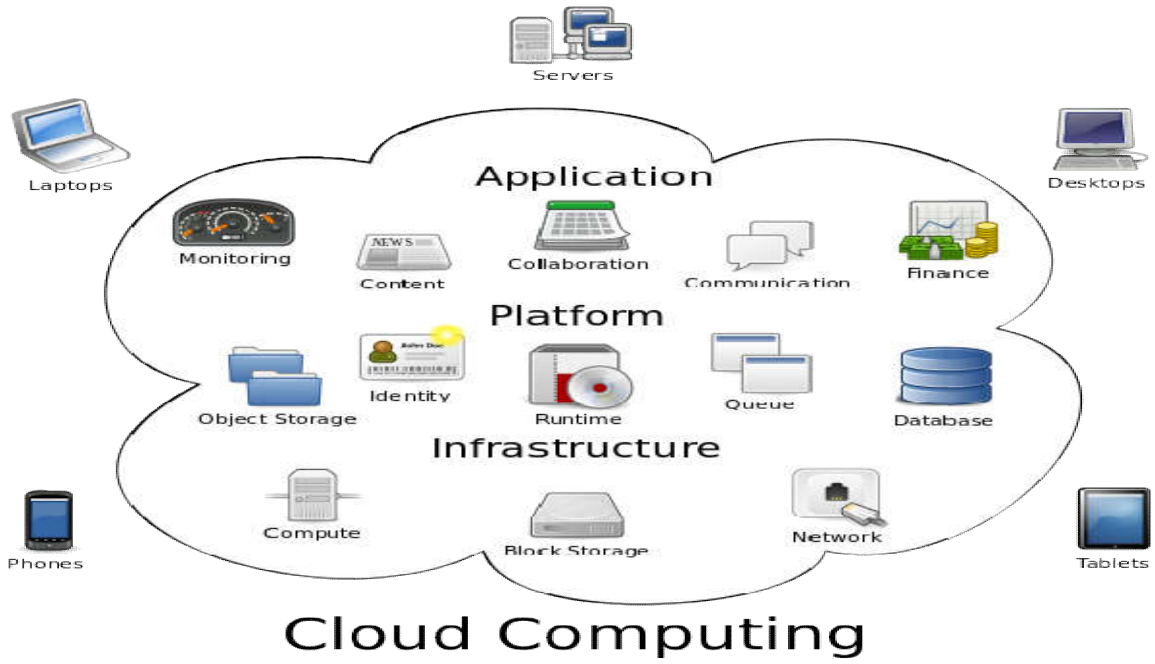
Deep Web is the part of the World Wide Web that is not discoverable by means of standard search engines, including password-protected or dynamic pages and encrypted networks.

65. Explain Grid computing & Cloud computing

Grid computing is the collection of computer resources from multiple locations to reach a common goal. The grid can be thought of as a distributed system with non-interactive workloads that involve a large number of files. Grid computing is distinguished from conventional high performance computing systems such ascluster computing in that grid computers have each node set to perform a different task/application.[1] Grid computers also tend to be more heterogeneous and geographically dispersed (thus not physically coupled) than cluster computers.[2] Although a single grid can be dedicated to a particular application, commonly a grid is used for a variety of purposes. Grids are often constructed with general-purpose grid middleware software libraries.

Grid size varies a considerable amount. Grids are a form of distributed computing whereby a “super virtual computer” is composed of many networked loosely coupled computers acting together to perform large tasks. For certain applications, “distributed” or “grid” computing, can be seen as a special type of parallel computing that relies on complete computers (with onboard CPUs, storage, power supplies, network interfaces, etc.) connected to a network (private or public) by a conventional network interface, such as Ethernet. This is in contrast to the traditional notion of a supercomputer, which has many processors connected by a local high-speed computer bus.

Cloud computing is a computing term or metaphor that evolved in the late 2000s, based onutility and consumption of computer resources. Cloud computing involves deploying groups of remote servers and software networks that allow different kinds of data sources be uploaded for real time processing to generate computing results without the need to store processed data on the cloud. Clouds can be classified as public, private or hybrid.



66. If you have one system, how you will connect it to internet?

Choosing an Internet Access Method

The specific steps required to connect a computer to the Internet depend on the type of Internet access involved. Most Internet access methods used in homes involve a small hardware unit called a modemthat connects to a physical medium supporting one of these fixed location services:

- phone line (for DSL)
- cable Internet (CATV) line
- fiber optic cable
- wireless antenna (for satellite and wireless broadband services)

Portable computers, like tablets, can be connected to fixed location networks inside a home, but they additionally support mobile broadband Internet access via cellular networks that can be used at home and while traveling. Finally, outside the home, portable computers can also reach the Internet via Wi-Fi hotspots, hardware access points installed in fixed locations that are in turn networked to Internet service through one of the other above methods.

Configuring an Internet Gateway (if applicable)

A network gateway is the hardware device that joins a local network to the Internet. On fixed location networks, the modem connects to the gateway device. Home networks most commonly use a broadband router as their gateway device, although technically any modern home computer can be set up as the gateway instead.

When using mobile broadband networks or Wi-Fi hotspots, the gateway hardware that directly connects a computer to the Internet is set up and maintained by service providers. However, some end users prefer to add a portable network router (typically advertised as a travel router) into their configuration. Travel routers serve as an additional layer of Internet gateway, helping to more conveniently connect a group of devices to the same Internet service and share data between them. Administrators configure travel routers similarly to other types of consumer routers.

Configuring the Internet Client Device

Configuration parameters must be set on a computer to match the type of network gateway and Internet service being used. Typical required settings for client computers include:

- user name and password – required for log in to Internet services based on PPPoE
- choice of network by name (SSID) – for Wi-Fi home networks and hotspots
- wireless security key (or passphrase) that matches the gateway – for Wi-Fi networks
- Wi-Fi turned off – for connecting via mobile broadband (cellular) networks
- Domain Name System (DNS), MTU and other service-specific settings - as required by the provider

Troubleshooting Internet Connection Problems

Mistakes in configuring network equipment often lead to failure connecting to the Internet. In wireless networking, entering incorrect security keys is one of the most common errors. Loose cables or cables plugged into the wrong locations cause similar errors on wired networks. Broadband modems must be connected to a home router uplink port and not any other of the router's ports, for example.

It may also be necessary to contact the Internet service provider to resolve connection problems. When connecting to a provider's network for the first time, the customer subscription must be activated and any special settings the provider requires (such as login information) set via the gateway. Once a computer has successfully connected to the provider's network the first time, subsequent problems tend to be unexpected outages due to weather or technical issues the provider is having with their own equipment (assuming the home network itself is functioning normally).

Advanced Internet Connection Topics

In some cases you can set up two (or more) Internet services on one device or on one home network. Smartphones, for example, can be connected via a Wi-Fi to a home wireless router but can communicate over the cell network instead when Wi-Fi isn't available. These so-called multi-homed configurations help keep you connected the Internet with fewer interruptions, as one of the network paths can still work even if the other one fails.

An Internet connection can be established, but computers may still not be able to reach Web sites normally, if the local network has an incorrect DNS configuration (or the DNS provider experiences a service outage).

67. What is virus and spam? Whats the difference between them? What is firewall?

    Virus

    A virus is a program which can self-replicate and insert itself into other applications on your computer. They vary in the amount of damage they can do, from simply slowing your computer down so much it becomes almost unusable to the worst scenario, which is destroying data, disabling software and deleting files.

    Spam

    Spam is unsolicited or junk email that clogs up your email inbox. The best way to deal with spam is not to open it or reply to it. Use your spam filter and if you are using a email site such as Gmail, be sure to report the spam, as every piece reported is added to a database. Be aware of which websites you enter your email address on, make sure the site states they will not sell your details onto a third party. When you sign up to some legitimate sites they will try to opt you in to as many advertising emails as possible, simply un-tick all the opt-in boxes.

    (FireWall definition is given above already)

68. What is CGI and what are its disadvantages? Why we should go for JAVA? (multi processing and multi threading concept)

    The common gateway interface (CGI) is a standard way for a Web server to pass a Web user's request to an application program and to receive data back to forward to the user.

    When you are dealing with Web pages, you will often hear people talk about CGI or CGI scripts without ever explaining exactly what that is. Essentially, CGI is the connection (or interface) between a form on a Web page and the Web server.

    Web pages cannot interact directly with the reader. In fact, until JavaScript came along, Web pages had no way of interpreting reader reaction except through interaction with the server they were running on. This interaction is done through scripts and programs that use common gateway interface to create interactive programs on your Web pages.

    Disadvantages of using CGI

    The biggest drawback to CGI Scripts is that they can put a lot of load on a Web server. Poorly written programs can fall into endless loops tying up server processor time. TheWeb browser will time out (usually after around 5 minutes), but often the server will continue to run the program until a system administrator comes in and shuts off the faulty script. The browser based scripting tools

mentioned above have the advantage of running off the reader's computer. They use the processor locally rather than on the Web server itself, and so are less intense on the Web server.

The Advantages of Servlets Over “Traditional” CGI.

Java servlets are more efficient, easier to use, more powerful, more portable, safer, and cheaper than traditional CGI and many alternative CGI-like technologies.

EFFICIENT

With traditional CGI, a new process is started for each HTTP request. If the CGI program itself is relatively short, the overhead of starting the process can dominate the execution time. With servlets, the Java Virtual Machine stays running and handles each request using a lightweight Java thread, not a heavyweight operating system process. Similarly, in traditional CGI, if there are Nsimultaneous requests to the same CGI program, the code for the CGI program is loaded into memory N times. With servlets, however, there would be N threads but only a single copy of the servlet class. Finally, when a CGI program finishes handling a request, the program terminates. This makes it difficult to cache computations, keep database connections open, and perform other optimizations that rely on persistent data. Servlets, however, remain in memory even after they complete a response, so it is straightforward to store arbitrarily complex data between requests.

Convenient

Servlets have an extensive infrastructure for automatically parsing and decoding HTML form data, reading and setting HTTP headers, handling cookies, tracking sessions, and many other such high-level utilities. Besides, you already know the Java programming language. Why learn Perl too? You’re already convinced that Java technology makes for more reliable and reusable code than does C++. Why go back to C++ for server-side programming?

Powerful

Servlets support several capabilities that are difficult or impossible to accomplish with regular CGI. Servlets can talk directly to the Web server, whereas regular CGI programs cannot, at least not without using a server-specific API. Communicating with the Web server makes it easier to translate relative URLs into concrete path names, for instance. Multiple servlets can also share data, making it easy to implement database connection pooling and similar resource-sharing optimizations. Servlets can also maintain information from request to request, simplifying techniques like session tracking and caching of previous computations.

Portable

Servlets are written in the Java programming language and follow a standard API. Consequently, servlets written for, say, I-Planet Enterprise Server can run virtually unchanged on Apache, Microsoft Internet Information Server (IIS), IBM Web Sphere, or Star Nine Web Star. For example, virtually all of the servlets and JSP pages in this book were executed on Sun's Java Web

Server, Apache Tomcat and Sun's Java Server Web Development Kit (JSWDK) with no changes whatsoever in the code. Many were tested on BEA Web Logic and IBM Web Sphere as well. In fact, servlets are supported directly or by a plug-in on virtually every major Web server. They are now part of the Java 2 Platform, Enterprise Edition (), so industry support for servlets is becoming even more pervasive.

Secure

One of the main sources of vulnerabilities in traditional CGI programs stems from the fact that general-purpose operating system shells often execute them. So the CGI programmer has to be very careful to filter out characters such as back quotes and semicolons that are treated specially by the shell. This is harder than one might think, and weaknesses stemming from this problem are constantly being uncovered in widely used CGI libraries. A second source of problems is the fact that languages that do not automatically check array or string bounds process some CGI programs. For example, in C and C++ it is perfectly legal to allocate a 100-element array then write into the 999th "element," which is really some random part of program memory. So programmers who forget to do this check themselves open their system up to deliberate or accidental buffer overflow attacks. Servlets suffer from neither of these problems. Even if a servlet executes a remote system call to invoke a program on the local operating system, it does not use a shell to do so. And of course array bounds checking and other memory protection features are a central part of the Java programming language.

Inexpensive

There are a number of free or very inexpensive Web servers available that are good for "personal" use or low-volume Web sites. However, with the major exception of Apache, which is free, most commercial-quality Web servers are relatively expensive. Nevertheless, once you have a Web server, no matter its cost, adding servlet support to it (if it doesn't come pre configured to support servlets) costs very little extra. This is in contrast to many of the other CGI alternatives, which require a significant initial investment to purchase a proprietary package.

69. What is a process?

In computing, a process is an instance of a computer program that is being executed. It contains the program code and its current activity. Depending on the operating system (OS), a process may be made up of multiple threads of execution that execute instructions concurrently.

70. What is a test vector? Why is u not calling it as a test sample?
    In computer science and engineering, a test vector is a set of inputs provided to a system in order to test that system.

71. What is the life cycle model? What are different types? What is the difference between waterfall and iterative waterfall model? Why have u not chosen spiral? Have u analyzed and then chosen the model?
72. If you have estimated 12 man hours and schedule results in 24 man hours, what does that mean?
73. What is testing? What is a test case? How do you generate it? What are the methods of test case generation?
74. For square root of a number find the equivalence partitioning and boundary value.
75. If I give u software, will u give me a certificate that it is 100% bug free?
76. After clearing software for prom fuse if ur boss asks u to find the faults in the software, how will u do it?
77. What is mutation theory? What is error seeding? Link it with 9^{th} question.
78. What is the difference between object oriented and object based language? give example.
79. What is the difference between space based and ground system testing?
80. How are you carrying out database verification? What are requirement verification tools? Which tool are you using?
81. Which tool are u using to carry out testing?
82. For DBMS projects what is the front end and backend? Why did you choose them?
83. Why did you validate the compiler?
84. How have you checked for test case adequacy?
85. You have generated so many non-conformances? Does it mean that u have not followed the process?
86. How did you carry out the IEEE compliance audit?
87. What are design patterns?
88. What is remote programming? how do u do it? what is the limit?
89. How did you carry out the IEEE compliance audit for OOD software?
90. How did you build the fault repository?
91. If a non-conformance is generated, then has it happened that it goes without correction?
92. What is BMU? What are its functions?
93. What is Teamcentre? Challenging task while installing it in ISAC.
94. Explain about NetApp Mass Storage
95. PKI and Digital Signature.

A public key infrastructure (PKI) supports the distribution and identification of public encryption keys, enabling users and computers to both securely exchange data over networks such as the Internet and verify the identity of the other party.

A digital certificate is an electronic "passport" allowing people, computers or organizations to exchange secure information over the Internet using the public key infrastructure (PKI).

96. What is Cloud Computing?

    Cloud computing is is typically defined as a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications.

97. Why you have used CGI and Perl for developing Software

    Advantages of CGI

    1. One advantage of CGI programs is that they are language independent

    2. CGI programs can be written in any language that allows one to write normal programs since they are executed in the same way as the normal programs. Another advantage of CGI is that it's a very simple interface

    3. It's not necessary to have any special library to create a CGI program, or write programs using a particular API. Instead, CGI programs rely on the standard concepts of standard input, standard output, and environment variables to communicate with the Web server

98. Need for Remote Job Submission Portal?

    "Remote job submissions" means that jobs are submitted by scripts or programs running on hosts.

    - Save time: Simplify job submission and management thanks to a powerful GUI with smart, simplified interfaces.
    - Be more productive: Spend more time focused on work and not IT tasks - for example, monitor jobs graphically without having to download huge job files.
    - Increase ROI: Consolidate access to applications and optimize license availability.
    - Reduce errors and improve consistency: Embed your company's best-practice "know how" directly into Application Definitions used for job submission

99. What is File system? How do you decide which file system to be mounted on a server

    A file system is the method an operating system uses to name files and assign them locations for efficient storage and retrieval.

1. Difference between SAN and NAS

    A storage area network (SAN) is storage connected in a fabric (usually through a switch) so that there can be easy access to storage from many different servers. From the server application and operating system standpoint, there is no visible difference in the access of data for storage in a

SAN or storage that is directly connected. A SAN supports block access to data just like directly attached storage.

Network-attached storage (NAS) is really remote file serving. Rather than using the software on your own file system, the file access is redirected using a remote protocol such as CIFS or NFS to another device (which is operating as a server of some type with its own file system) to do the file I/O on your behalf. This enables file sharing and centralization of management for data.

The difference between SAN and NAS is that SAN is for block I/O and NAS is for file I/O. One additional thing to remember when comparing SAN vs. NAS is that NAS does eventually turn the file I/O request into a block access for the storage devices attached to it.

2. What are clustered systems and Distributed Systems?

   Clustered system: systems having many computers with shared storage and linked by a lan or network.

   Distributed system:systems having many computers connected by a network and there is no shared storage.

3. What problems you face when using Open Source Applications?

   Problems faced when using Open Source Applications

   - Because there is no requirement to create a commercial product that will sell and generate money, open source software can tend to evolve more in line with developers' wishes than the needs of the end user.
   - For the same reason, they can be less "user-friendly" and easy to use because less attention is paid to developing the user interface.
   - There may also be less support available for when things go wrong – open source software tends to rely on its community of users to respond to and fix problems.
   - Although the open source software itself mostly free, there may still be some indirect costs involved, such as paying for external support.
   - Although having an open system means that there are many people identifying bugs and fixing them, it also means that malicious users can potentially view it and exploit any vulnerabilities.

4. Why was the consolidation required? (isacmail1 and isacmail2 into isacmail.)
5. What is Environment Monitoring System? What sensors are installed?
6. Do you get instantaneous readings for the sensors? Or will there be any delay?
7. Is it OK, even if there is any delay?
8. Which are the two protocols used for Mail?

   IMAP (Internet Message Access Protocol) , POP3 (Post Office Protocol 3), SMTP (Simple Mail Transfer Protocol)

9. Difference between IMAP and POP?

   IMAP (Internet Message Access Protocol) – Is a standard protocol for accessing e-mail from your local server. IMAP is a client/server protocol in which e-mail is received and held for you by your Internet server. As this requires only a small data transfer this works well even over a slow connection such as a modem. Only if you request to read a specific email message will it be downloaded from the server. You can also create and manipulate folders or mailboxes on the server, delete messages etc.

   The POP (Post Office Protocol 3) protocol provides a simple, standardized way for users to access mailboxes and download messages to their computers.

   When using the POP protocol all your eMail messages will be downloaded from the mail server to your local computer. You can choose to leave copies of your eMails on the server as well. The advantage is that once your messages are downloaded you can cut the internet connection and read your eMail at your leisure without incuring further communication costs. On the other hand you might have transferred a lot of message (including spam or viruses) in which you are not at all interested at this point.
10. What is OTP? Which type of OTP method you have deployed?

    A one-time password (OTP) is a password that is valid for only one login session or transaction, on a computer system or other digital device.
11. What is the other security setups implemented on your Internet network?
12. Will it be possible to have a single ISRO mail server for all ISRO staff?
13. Does the conference management software has content management built in?
14. If no, how do you adapt to the same software to different Conferences?
15. If I want CMS for COSPAR, how much time u will require to set it up?
16. What is SPAMAssassin and how it works?

    SpamAssassin is a computer program released under the Apache License 2.0 used for e-mail spam filtering based on content-matching rules. It is now part of the Apache Foundation.SpamAssassin uses a variety of spam-detection techniques, that includes DNS-based and fuzzy-checksum-based spam detection, Bayesian filtering, external programs, blacklists and online databases.

    How it works

    SpamAssassin comes with a large set of rules which are applied to determine whether an email is spam or not. Most rules are based on regular expressions that are matched against the body or header fields of the message, but SpamAssassin also employs a number of other spam-fighting techniques. The rules are called "tests" in the SpamAssassin documentation.

Each test has a score value that will be assigned to a message if it matches the test's criteria. The scores can be positive or negative, with positive values indicating "spam" and negative "ham" (non-spam messages). A message is matched against all tests and SpamAssassin combines the results into a global score which is assigned to the message. The higher the score, the higher the probability that the message is spam.

SpamAssassin has an internal (configurable) score threshold to classify a message as spam. Usually a message will only be considered as spam if it matches multiple criteria; matching just a single test will not usually be enough to reach the threshold.

If SpamAssassin considers a message to be spam, it can be further rewritten. In the default configuration, the content of the mail is appended as aMIME attachment, with a brief excerpt in the message body, and a description of the tests which resulted in the mail being classified as spam. If the score is lower than the defined settings, by default the information about the tests passed and total score is still added to the email headers and can be used in post-processing for less severe actions, such as tagging the mail as suspicious.

SpamAssassin allows for a per-user configuration of its behaviour, even if installed as system-wide service; the configuration can be read from a file or a database. In their configuration users can specify individuals whose emails are never considered spam, or change the scores for certain rules. The user can also define a list of languages which they want to receive mail in, and SpamAssassin then assigns a higher score to all mails that appear to be written in another language.

SpamAssassin is based on heuristics (pattern recognition), and such software exhibits some false positives, blocking email that may be entirely innocent, hence the need for the software to go through a "learning" exercise. This is similar to heuristic software utilized by credit card issuing banks, that will block a credit card number based upon "suspicious" usage patterns, such as a large number of purchases made within a short time period. As there is no way to tell the "bad guys" from the "good guys" with one-hundred percent accuracy, there are going to be mistakes made determining the appropriate category for some email

17. How many attributes / tables are used for DARS and how it is implemented?
18. Have you done load testing for any server? Did you do load testing for ISACMail?
19. Is CAPTCHA with audio has any threats?
20. What are the features implemented in your Proxy Server.
21. How the caching works in your proxy server?
22. In your LAN there are about 1700 PCs. They remain idle after office hours? How can you make use of all such PCs for some benefit?
23. What are thin clients? Are you aware?

24. What are the fault tolerant features?

    The ability of a system to respond gracefully to an unexpected hardware or software failure. There are many levels of fault tolerance, the lowest being the ability to continue operation in the event of a power failure. Many fault-tolerant computer systems*mirror* all operations -- that is, every operation is performed on two or more duplicate systems, so if one fails the other can take over.

25. What is fault injection?

    In software testing, fault injection is a technique for improving the coverage of a test by introducing faults to test code paths, in particular error handling code paths, that might otherwise rarely be followed. It is often used with stress testing and is widely considered to be an important part of developing robust software.[1] Robustness testing[2] (also known as Syntax Testing,Fuzzing or Fuzz testing) is a type of fault injection commonly used to test for vulnerabilities in communication interfaces such as protocols, command line parameters, or APIs.

    The propagation of a fault through to an observable failure follows a well defined cycle. When executed, a fault may cause an error, which is an invalid state within a system boundary. An error may cause further errors within the system boundary, therefore each new error acts as a fault, or it may propagate to the system boundary and be observable. When error states are observed at the system boundary they are termed failures. This mechanism is termed the fault-error-failure cycle [3] and is a key mechanism in dependability.

26. Diff betwn V-model and iterative water fall model

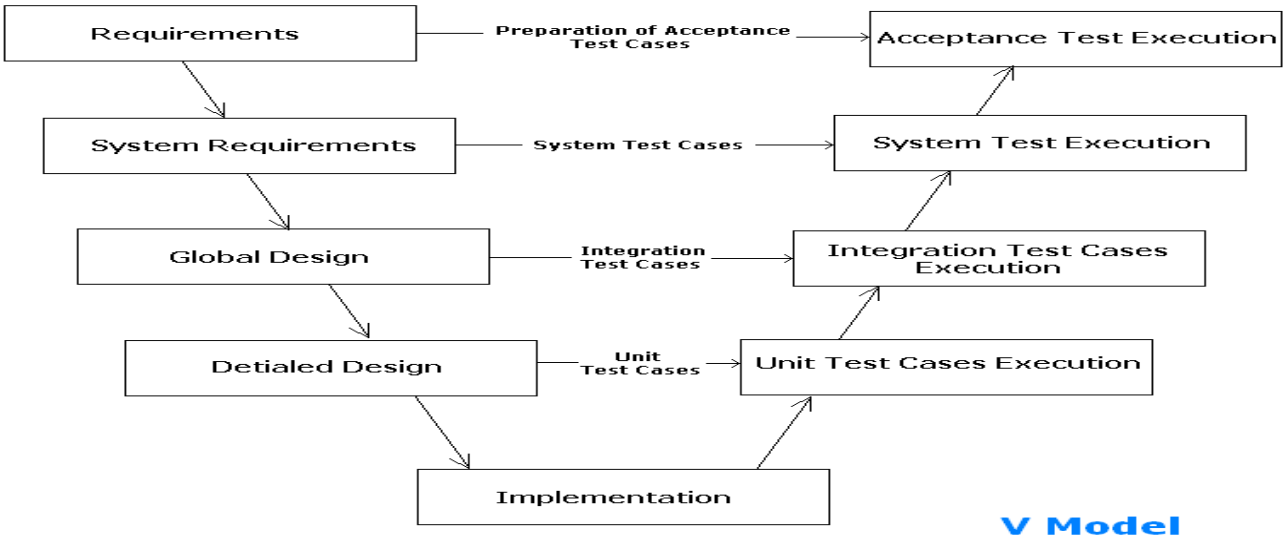
    V Model

    In the basic Waterfall model process seen some disadvantages or limitations in the model which started a new SDLC model. As we seen in the Waterfall model the issues found in the end of the SDLC, this is due to the testing is occurred in the end phases of the you SDLC. To overcome this problem the **V-Model** is comes into the picture. It is always better to introduce testing in the early phase of SDLC, as in this model the testing activity gets started from the early phase of the SDLC.

    Before starting the actual testing, testing team has to work on various activities like preparation of Test Strategy, Test Planning, Creation of Test cases & Test Scripts which is work parallel with the development activity which help to get the test deliverable on time.

    V Model – Software Development Life Cycle

V Model - Software Development Life Cycle V Model - Software Development Life Cycle V Model - Software Development Life Cycle



V Model - Software Development Life Cycle

In the V Model Software Development Life Cycle, based on same information(requirement specification document) the development & testing activity is started. Based on the requirement document developer team started working on the design & after completion on design start actual implementation and testing team starts working on test planning, test case writing, test scripting. Both activities are working parallel to each other. In Waterfall model & V-model they are quite similar to each other. As it is most popular Software Testing Life Cycle model so most of the organization is following this model.

The V-model is also called as Verification and Validation model. The testing activity is perform in the each phase of Software Testing Life Cycle phase. In the first half of the model Verification testing activity is integrated in each phase like review user requirements, System Design document & in the next half the Validations testing activity is come in picture.

Typical V-model shows Software Development activities on the Left hand side of model and the Right hand side of the model actual Testing Phases can be performed.

In this process “Do-Procedure” would be followed by the developer team and the “Check-Procedure” would be followed by the testing team to meets the mentioned requirements.

In the V-Model software development life cycle different steps are followed however here we will taking a most common type of V-model example. The V-model typically consist of the following phases:

1. Unit Testing: Preparation of Unit Test Cases

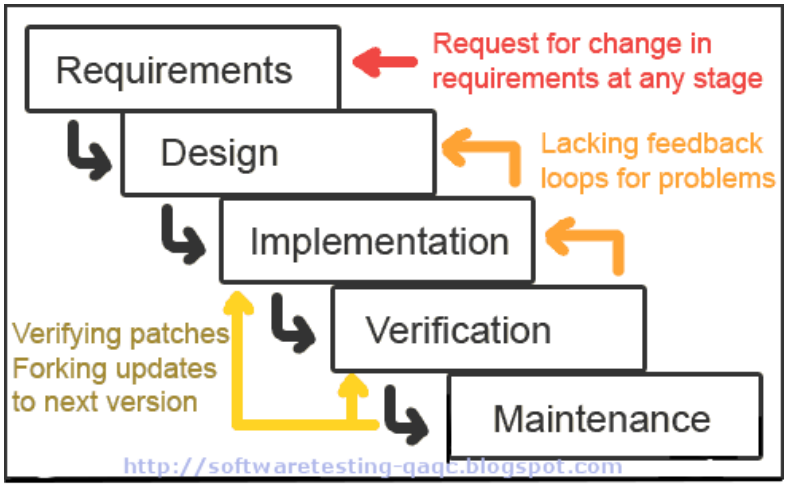
2. Integration Testing: Preparation of Integration Test Cases

3. System Testing: Preparation of System test cases

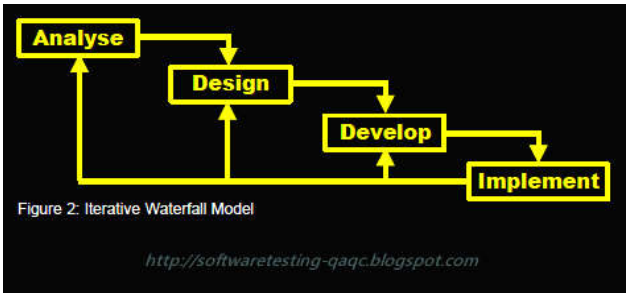
4. Acceptance Testing: Preparation of Acceptance Test Cases

Iterative Waterfall Model

- Water Fall model has been revised later to be Iterative.
- Everything remains the same as Waterfall Model.
- It is easy to use.



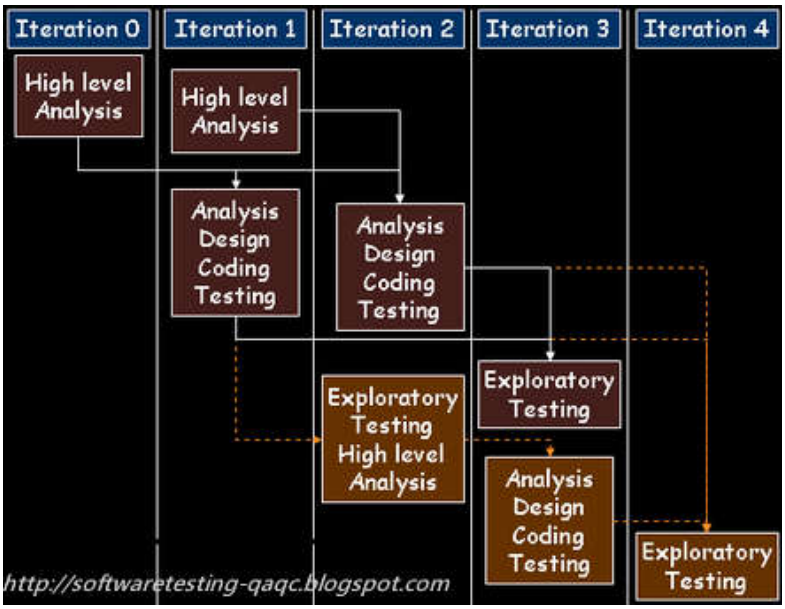
Iterative Waterfall Model



Iterative Waterfall Model.

- Each iteration involves Design Analysis and Implementation as well as verification of the current build/version of the system.
- If the application lacks any requirements or has any problems then the phase will be looped back to previous iteration.

- It supports redesign, acceptance and review of any new requirement.
- It involves analysis of usability, achievement of goals, reliability, efficiency, and structure.
- It has many draw backs - each process is iterative and has many problems to meet the project deadlines.



Waterfall Model Iteration.

1. Difference between SQLITE and MySQL database

MySQL	**SQLite**
MySQL is a database system from Oracle corporation.	SQLite is a database developed by D. Richard Hipp.
MySQL ranks 2nd in DB engines ranking.	SQLite ranks 8th in DB engines ranking.
Type system of MySQL is static.	SQLite type system is dynamic.
MySQL is available under GPL and proprietary license.	SQLite is a public domain work.
MySQL runs on (i) z/OS, (ii) BSD, (iii) Symbian, (iv) AmigaOS, (v) Windows, (vi) Mac OS X, (vii) Linux and (viii) UNIX operating system.	SQLite also is supported on (i) z/OS, (ii) BSD, (iii) Symbian, (iv) AmigaOS, (v) Windows, (vi) Mac OS X, (vii) Linux and (viii) UNIX operating system. But z/OS does not support SQLite.
SQL interface is provided by MySQL. MySQL	SQLite provides SQL interface but not GUI.

does not provide GUI. MySQL Workbench provides an integrated GUI environment for MySQL.	SQLite Database browser can be used as a GUI editor for SQLite databases.
MySQL does not provide audit control. By placing the plugin, audit_log in the MySQL plugin directory, audit control can be achieved.	SQLite provides audit control
MySQL supports native network encryption.	SQLite does not support native network encryption.SQLite Encryption Extension (SEE) can be used to support encryption in SQLite.
MySQL does not have resource limit. By setting max_user_connections system variable to a nonzero value, server resources can be limited.	SQLite has resource limit
MySQL can create indexes with hash, R/R+ tree, and fulltext. MySQL does not support reverse index.	SQLite creates indexes with reverse, R/R+ tree, and fulltext. SQLite does not support hash index.
MySQL supports composite (range + hash), hash, list and range partitioning methods.	SQLite does not support composite (range + hash), hash, list and range partitioning methods. By placing each table in a separate database file and then attaching (ATTACH) the databases, table partitioning can be supported.
MySQL does not support except and intersect operators.	SQLite supports except and intersect operators
MySQL provides merge joins and outer joins	SQLite does not provide merge joins. Only LEFT OUTER JOIN is implemented. RIGHT OUTER JOIN and FULL OUTER JOIN are not implemented in SQLite
MySQL provides Java support.	SQLite does not support Java.
MySQL supports (i) 32bit INTEGER, (ii) 24bit MEDIUMINT, (iii) 8bit TINYINT, (iv) 64bit BIGINT and (v) SMALLINT data types.	SQLite does not support (i) 32bit INTEGER, (ii) 24bit MEDIUMINT, (iii) 8bit TINYINT, (iv) 64bit BIGINT and (v) SMALLINT. SQLite supports 64bit INTGER8 data type.

MySQL supports FLOAT and 64bit DOUBLE data type. MYSQL does not support REAL data types.	SQLite does not support FLOAT and 64bit DOUBLE, instead supports REAL data type.
MySQL supports DECIMAL data type.	SQLite does not support DECIMAL data type. Decimal data type is set to UNKNOWN in SQLite.

2. Which Network monitoring tool has been used in ISAC, How it is beneficial to ISAC
3. What is testing? Simple example for Unit test(testing definition given below)

   Example of Unit testing is explain below

   For example you are testing a function; whether loop or statement in a program is working properly or not than this is called as unit testing. A beneficial example of a framework that allows automated unit testing is JUNIT (a unit testing framework for java). XUnit [20] is a more general framework which supports other languages like C#, ASP, C++, Delphi and Python to name a few.

   Tests that are performed during the unit testing are explained below:

   - Module Interface test: In module interface test, it is checked whether the information is properly flowing in to the program unit (or module) and properly happen out of it or not.
   - Local data structures: These are tested to inquiry if the local data within the module is stored properly or not.
   - Boundary conditions: It is observed that much software often fails at boundary related conditions. That's why boundary related conditions are always tested to make safe that the program is properly working at its boundary condition's.
   - Independent paths: All independent paths are tested to see that they are properly executing their task and terminating at the end of the program.
   - Error handling paths: These are tested to review if errors are handled properly by them or not.

1. Chairman Question: give difference between COWAA and other banking systems (like SBI, ICICI)
2. What is evaluation in 'Testing and Evaluation Process'
3. Which are different testing methodologies? What is functional testing and performance testing

   Software Testing

   Software testing is the process of evaluation a software item to detect differences between given input and expected output. Also to assess the feature of A software item. Testing assesses the quality of the product. Software testing is a process that should be done during the development process. In other words software testing is a verification and validation process.

Verification

Verification is the process to make sure the product satisfies the conditions imposed at the start of the development phase. In other words, to make sure the product behaves the way we want it to.

Validation

Validation is the process to make sure the product satisfies the specified requirements at the end of the development phase. In other words, to make sure the product is built as per customer requirements.

Basics of software testing

There are two basics of software testing: blackbox testing and whitebox testing.

Blackbox Testing

Black box testing is a testing technique that ignores the internal mechanism of the system and focuses on the output generated against any input and execution of the system. It is also called functional testing.

Whitebox Testing

White box testing is a testing technique that takes into account the internal mechanism of a system. It is also called structural testing and glass box testing.

Black box testing is often used for validation and white box testing is often used for verification.

Types of testing

There are many types of testing like

- Unit Testing
- Integration Testing
- Functional Testing
- System Testing
- Stress Testing
- Performance Testing
- Usability Testing
- Acceptance Testing
- Regression Testing
- Beta Testing

Unit Testing

Unit testing is the testing of an individual unit or group of related units. It falls under the class of white box testing. It is often done by the programmer to test that the unit he/she has implemented is producing expected output against given input.

Integration Testing

Integration testing is testing in which a group of components are combined to produce output. Also, the interaction between software and hardware is tested in integration testing if software and hardware components have any relation. It may fall under both white box testing and black box testing.

Functional Testing

Functional testing is the testing to ensure that the specified functionality required in the system requirements works. It falls under the class of black box testing.

System Testing

System testing is the testing to ensure that by putting the software in different environments (e.g., Operating Systems) it still works. System testing is done with full system implementation and environment. It falls under the class of black box testing.

Stress Testing

Stress testing is the testing to evaluate how system behaves under unfavorable conditions. Testing is conducted at beyond limits of the specifications. It falls under the class of black box testing.

Performance Testing

Performance testing is the testing to assess the speed and effectiveness of the system and to make sure it is generating results within a specified time as in performance requirements. It falls under the class of black box testing.

Usability Testing

Usability testing is performed to the perspective of the client, to evaluate how the GUI is user-friendly? How easily can the client learn? After learning how to use, how proficiently can the client perform? How pleasing is it to use its design? This falls under the class of black box testing.

Acceptance Testing

Acceptance testing is often done by the customer to ensure that the delivered product meets the requirements and works as the customer expected. It falls under the class of black box testing.

Regression Testing

Regression testing is the testing after modification of a system, component, or a group of related units to ensure that the modification is working correctly and is not damaging or imposing other modules to produce unexpected results. It falls under the class of black box testing.

Beta Testing

Beta testing is the testing which is done by end users, a team outside development, or publicly releasing full pre-version of the product which is known as beta version. The aim of beta testing is to cover unexpected errors. It falls under the class of black box testing.

4. What are testing metrics
   Test metrics accomplish in analyzing the current level of maturity in testing and give a projection on how to go about testing activities by allowing us to set goals and predict future trends.
5. How do you design test cases
6. What are stubs and drivers

   Driver :A software component or test tool that replaces a component that takes care of the control and/or the calling of a component or system.

   Stub: A skeletal or special-purpose implementation of a software component, used to develop or test a component that calls or is otherwise dependent on it. It replaces a called component.

7. If zero downtime of the server is mandatory how you ensure it, if operational server is down.
8. What is software quality? And what are the software quality factors? What is reliability?

   Software quality refers to two related but distinct notions that exist wherever quality is defined in a business context:

   - Software functional quality reflects how well it complies with or conforms to a given design, based on functional requirements or specifications. That attribute can also be described as the fitness for purpose of a piece of software or how it compares to competitors in the marketplace as a worthwhile product;[1]
   - Software structural quality refers to how it meets non-functional requirements that support the delivery of the functional requirements, such as robustness or maintainability, the degree to which the software was produced correctly.

Structural quality is evaluated through the analysis of the software inner structure, its source code, at the unit level, the technology level and the system level, which is in effect how its architecture adheres to sound principles of software architecture outlined in a paper on the topic by OMG.[2] In contrast, functional quality is typically enforced and measured throughsoftware testing.

Some **software quality factors** are listed here:

- Understandability is possessed by a software product if the purpose of the product is clear. This goes further than just a statement of purpose - all of the design and user documentation must be clearly written so that it is easily understandable. This is obviously subjective in that the user context must be taken into account, i.e. if the software product is to be used by software engineers it is not required to be understandable to the layman.
- A software product possesses the characteristic completeness to the extent that all of its parts are present and each of its parts is fully developed. This means that if the code calls a sub-routine from an external library, the software package must provide reference to that library and all required parameters must be passed. All required input data must be available.
- A software product possesses the characteristic conciseness to the extent that no excessive information is present. This is important where memory capacity is limited, and it is important to reduce lines of code to a minimum. It can be improved by replacing repeated functionality by one sub-routine or function which achieves that functionality. It also applies to documents.
- A software product possesses the characteristic portability to the extent that it can be operated easily and well on computer configurations other than its current one. This is particularly important with PC applications where, for example, a product is expected to work on all 80486 processors.
- A software product possesses the characteristic maintainability to the extent that it facilitates updating to satisfy new requirements. Thus the software product which is maintainable should be well-documented, not complex, and should have spare capacity for memory usage and processor speed.
- A software product possesses the characteristic testability to the extent that it facilitates the establishment of acceptance criteria and supports evaluation of its performance. Such a characteristic must be built-in during the design phase if the product is to be easily testable - a complex design leads to poor testability.
- A software product possesses the characteristic usability to the extent that it isconvenient and practicable to use. This is affected by such things as the human-computer interface. The component of the software which has most impact on this is the graphical user interface (GUI).
- A software product possesses the characteristic reliability to the extent that it can be expected to perform its intended functions satisfactorily. This implies a time factor in that a reliable product is expected to perform correctly over a period of time. It also encompasses

environmental considerations in that the product is required to perform correctly in whichever conditions it finds itself - this is sometimes termed robustness.

Software Reliability: Ability of a computer program to perform its intended functions and operations in a system's environment, without experiencing failure (system crash).

1. Can you name ONE Specific drawback of Sun Solaris OS (Ans: Support to 3rd Party tools?)
   - Uncompatible : It is not recommended to run Solaris on other architectures such as Intel, AMD. It is possible to install Solaris on Intel however, the performance would degrade considerably since Solaris cannot make use of Intel that efficiently.
   - Lack of good GUI: Solaris does have GUI support - Common Desktop Environment, OpenWindows etc. but they are far way from the other GUI environments seen in Windows or Mac.
   - Costlier: With other cheaper alternatvies such as Linux available, it proves to be costlier to acquire a license of Solaris. Since it is intended to be used on SPARC so the end user often ends up in buying the hardware as well.
2. What is a patch? What are different types of Patches? What is Difference between a Patch and a Code revision?

   A patch is a piece of software designed to update a computer program or its supporting data, to fix or improve it. This includes fixing security vulnerabilities and other bugs, and improving the usability or performance.

   Different types of Patches

   Patches for proprietary software are typically distributed as executable files instead of source code. This type of patch modifies the program executable—the program the user actually runs—either by modifying the binary file to include the fixes or by completely replacing it.

   Patches can also circulate in the form of source code modifications. In this case, the patches usually consist of textual differences between two source code files, called "diffs". These types of patches commonly come out of open source projects. In these cases, developers expect users to compile the new or changed files themselves.

   Because the word "patch" carries the connotation of a small fix, large fixes may use different nomenclature. Bulky patches or patches that significantly change a program may circulate as "service packs" or as "software updates". Microsoft Windows NT and its successors (including Windows 2000, Windows XP, and later versions) use the "service pack" terminology.

3. How will you ensure the systems performance from degrading, when you add a new application to the Server?

4. What are the different types of Maintenance? Preventive , Corrective and which is the other
5. Corrective, preventive, risk-based and condition-based maintenance.
6. How do you monitor your Systems/ Servers performance? When you have a large number of systems or servers to be monitored which tools you will use?

   Maintaining Server Performance is a widely used service business. Systems administrators need to ensure that their computer systems' performance is efficient. Server performance is optimized by collecting data supplied when server performance is monitored. This includes checking the servers' activities and maintaining its tasks. To monitor the performance of your system configuration, collect different types of data over a small period of time.

   Step 1

   Understanding the workload of the system and the other effects on the system's resources is required to monitor a server performance. Providing the data is given by the resources used by some parts of the Operating System and by other programs. The data provided by the system shows the alerts of the system.

   Step 2

   Using a Task Manager, you can monitor the computer system. Open a Task Manager by pressing right click on your task bar and clicking the 'Task Manager' button, or pressing the 'CTRL' + 'ALT' + 'DEL' at the same time. Check the system's performance by looking at the 'CPU Usage' meter and the 'Memory Usage' meter. We could check if the system is running low on memory.

   Step 3

   To check the processes of your system, open the Task Manager then click on the 'Processes' tab. It shows the memory usage of each program running on the system. To stop an active program that is not required on your system, choose the program and click the 'End Process' button on the lower right of the window. Terminating a program that is not required speeds up the system.

   Step 4

   One way of checking out the Server Performance is by opening up the 'Performance Tool' on the system. To open up the 'Performance Tool', click 'Start' then 'All Programs'. Go to 'Administrative Tools' and click 'Performance'. Checking out the system by looking out at the graph and monitoring the graph are helpful for your system. It scales the graph of the Memory, the Physical Disk, and the Processor of your Server.

   Step 5

   It is important to make a Counter Log on your computer system. A Counter Log may be created by simply opening Performance, double-clicking 'Performance Logs and Alerts', and then clicking 'Counter Logs'. Click on the blank area of the program, and choose 'New Log Settings'. Enter the name of the counter log, and click 'OK'.

   Step 6

Using the Counter Log helps generate server's activities and data. Using the data collected checks and optimizes the system to locate the cause of delays and tune the server for optimal performance.

7. What contingency plans are built by you to minimize the down time? What is Cluster? Have you configured? What is RAID, Explain RAID 5, what is RAID 6?

Cluster Definition:

Connecting two or more computers together in such a way that they behave like a single computer. Clustering is used for parallel processing, load balancing and fault tolerance.

Clustering is a popular strategy for implementing parallel processing applications because it enables companies to leverage the investment already made in PCs and workstations. In addition, it's relatively easy to add new CPUs simply by adding a new PC to the network.

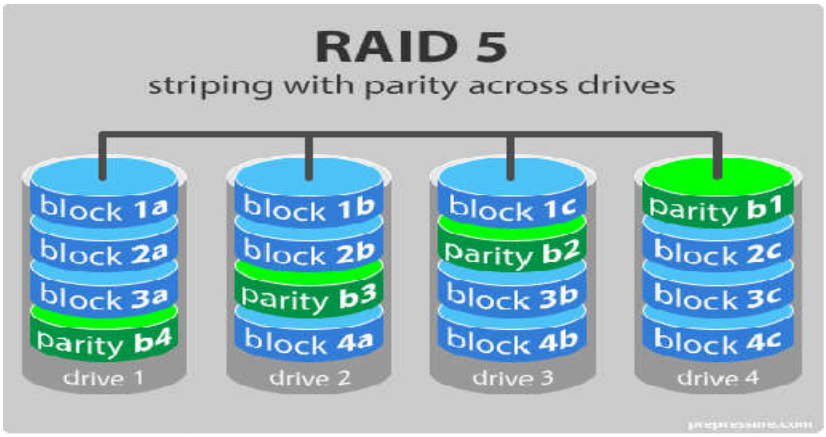
Raid Definition:

RAID (originally redundant array of inexpensive disks; now commonly redundant array of independent disks) is a data storage virtualization technology that combines multiple disk drive components into a logical unit for the purposes of data redundancy or performance improvement.

Data is distributed across the drives in one of several ways, referred to as RAID levels, depending on the specific level of redundancy and performance required. The different schemes or architectures are named by the word RAID followed by a number (e.g. RAID 0, RAID 1). Each scheme provides a different balance between the key goals: reliability, availability,performance, and capacity. RAID levels greater than RAID 0 provide protection against unrecoverable (sector) read errors, as well as whole disk failure.

RAID level 5

RAID 5 is the most common secure RAID level. It requires at least 3 drives but can work with up to 16. Data blocks are striped across the drives and on one drive a parity check sum of all the block data is written. The parity data are not written to a fixed drive, they are spread across all drives, as the drawing below shows. Using the parity data, the computer can recalculate the data of one of the other data blocks, should those data no longer be available. That means a RAID 5 array can withstand a single drive failure without losing data or access to data. Although RAID 5 can be achieved in software, a hardware controller is recommended. Often extra cache memory is used on these controllers to improve the write performance.



Advantages

- Read data transactions are very fast while write data transactions are somewhat slower (due to the parity that has to be calculated).
- If a drive fails, you still have access to all data, even while the failed drive is being replaced and the storage controller rebuilds the data on the new drive.

Disadvantages

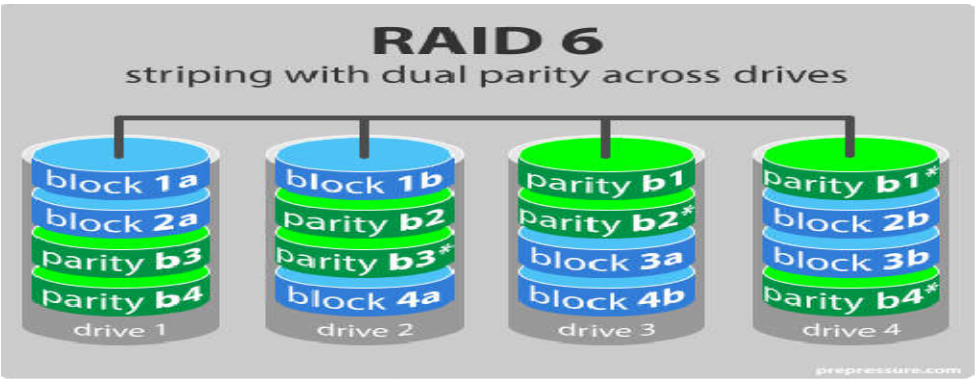
- Drive failures have an effect on throughput, although this is still acceptable.
- This is complex technology. If one of the disks in an array using 4TB disks fails and is replaced, restoring the data (the rebuild time) may take a day or longer, depending on the load on the array and the speed of the controller. If another disk goes bad during that time, data are lost forever.

Ideal use

RAID 5 is a good all-round system that combines efficient storage with excellent security and decent performance. It is ideal for file and application servers that have a limited number of data drives.

RAID level 6 – Striping with double parity

RAID 6 is like RAID 5 but it writes the parity data to two drives. That means it requires at least 4 drives and can withstand 2 drives dying simultaneously. The chances that two drives break down at exactly the same moment are of course very small. However, if a drive in a RAID 5 systems dies and is replaced by a new drive, it takes hours to rebuild the swapped drive. If another drive dies during that time, you still lose all of your data. With RAID 6, the RAID array will even survive that second failure.



Advantages

- Like with RAID 5, read data transactions are very fast.
- If two drives fail, you still have access to all data, even while the failed drives are being replaced. So RAID 6 is more secure than RAID 5.

Disadvantages

- Write data transactions are slowed down due to the parity that has to be calculated.
- Drive failures have an effect on throughput, although this is still acceptable.
- This is complex technology. Rebuilding an array in which one drive failed can take a long time.

Ideal use

RAID 6 is a good all-round system that combines efficient storage with excellent security and decent performance. It is preferable over RAID 5 in file and application servers that use many large drives for data storage.

8. Can you pl. tell me the different OS, you have worked on, and the advantages and disadvantages of each?
9. What is the software development platform you have used? , Can you name some Open source Web Development tools?

   Open source Web Development tools: Eclipse, Apache, Jquery.

10. What is time synchronization? What is its importance? How is time synchronization done in ACS, e-Procurement? , What is NTP? Does any group in ISAC maintain the timing services?

    Time synchronization is a problem from computer science and engineering which deals with the idea that internal clocks of several computers may differ. Even when initially set accurately, real clocks will differ after some amount of time due to clock drift, caused by clocks counting time at slightly different rates. There are several problems that occur as a repercussion of clock rate differences and several solutions, some being more appropriate than others in certain contexts.

The Importance of Time Synchronization for Your Network

In modern computer networks time synchronization is critical because every aspect of managing, securing, planning, and debugging a network involves determining when events happen. Time also provides the only frame of reference between all devices on the network. Without synchronized time, accurately correlating log files between these devices is difficult, even impossible. Following are just a few specific reasons:

- Tracking security breaches, network usage, or problems affecting a large number of components can be nearly impossible if timestamps in logs are inaccurate. Time is often the critical factor that allows an event on one network node to be mapped to a corresponding event on another.
- To reduce confusion in shared file systems, it is important for the modification times to be consistent, regardless of what machine the file systems are on.
- Billing services and similar applications must know the time accurately.
- Some financial services require highly accurate timekeeping by law.
- Sarbanes-Oxley and HIPAA Security Rules both require accurate timestamping.

NTP Definition: NTP stands for Network Time Protocol, and it is an Internet protocol used to synchronize the clocks of computers to sometime reference. NTP is an Internet standard protocol originally developed by Professor David L. Mills at the University of Delaware.

11. What are 3 tier / n-tier Architecture? Can you map your Sandesh and COWAA to the Architecture models?(Answered)
12. In which tier the business logic stored in COWAA? What are the advantages of it?
13. What is e-procurement? How the Tamper proofing is achieved in e-procurement Bids?
14. What is Hashing? , What are Encryption / Decryption? , What is PKI?

    Hashing is the transformation of a string of characters into a usually shorter fixed-length value or key that represents the original string.

    Encryption is the conversion of electronic data into another form, called ciphertext, which cannot be easily understood by anyone except authorized parties.

    Decryption is the process of transforming data that has been rendered unreadable through encryption back to its unencrypted form.

    (PKI Explained already)
15. Do you do any backend correction to COWAA Database? , How can you prevent doing this?

1. How to protect from malicious data/code placed in options field of TCP header?
2. What are the processes for Business Process Management and how to use workflow for them?

3. Which is the parameter to be tuned to improve performance of a website when it is not responding at all due to heavy load?
4. In Bhuvan website, how to monitor who has accessed how many times a particular location within the image data of GIS?
5. What would you like to do in future from professional career point of view?
6. What is RSS, Web 2.0, Proxy Server, Reverse Proxy Server, and Workflow?

RSS

RSS is the acronym used to describe the de facto standard for the syndication *of Web content*. RSS is an XML-based format and while it can be used in different ways for content distribution, its most widespread usage is in distributing news headlines on the Web. A Web site that wants to allow other sites to publish some of its content creates an RSS document and registers the document with an RSS publisher. A user that can read RSS-distributed content can use the content on a different site. Syndicated content can include data such as news feeds, events listings, news stories, headlines, project updates, excerpts from discussion forums or even corporate information.

Web 2.0

Web 2.0 is term that was introduced in 2004 and refers to the second generation of the World Wide Web. The term "2.0" comes from the software industry, where new versions of software programs are labeled with an incremental version number. Like software, the new generation of the Web includes new features and functionality that was not available in the past. However, Web 2.0 does not refer to a specific version of the Web, but rather a series of technological improvements.

Some examples of features considered to be part of Web 2.0 are listed below:

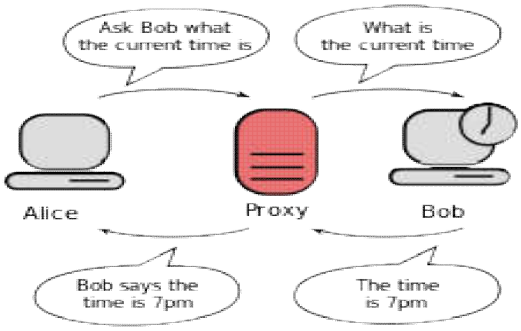
- Blogs - also known as Web logs, these allow users to post thoughts and updates about their life on the Web.
- Wikis - sites like Wikipedia and others enable users from around the world to add and update online content.
- Social networking - sites like Facebook and MySpace allow users to build and customize their own profile sand communicate with friends.
- Web applications - a broad range of new applications make it possible for users to run programs directly in a Web browser.

Web 2.0 technologies provide a level user interaction that was not available before. Websites have become much more dynamic and interconnected, producing "online communities" and making it even easier to share information on the Web. Because most Web 2.0

features are offered as free services, sites like Wikipedia and Facebook have grown at amazingly fast rates. As the sites continue to grow, more features are added, building off the technologies in place. So, while Web 2.0 may be a static label given to the new era of the Web, the actual technology continues to evolve and change.

Proxy Server

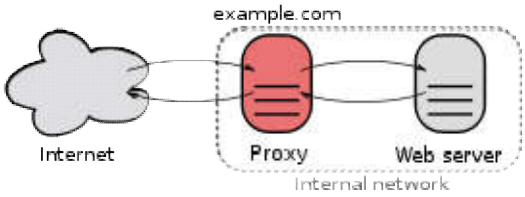
In computer networks, a proxy server is a server (a computer system or an application) that acts as an intermediary for requests from clients seeking resources from other servers. A client connects to the proxy server, requesting some service, such as a file, connection, web page, or other resource available from a different server and the proxy server evaluates the request as a way to simplify and control its complexity. Proxies were invented to add structure and encapsulation to distributed systems. Today, most proxies are web proxies, facilitating access to content on the World Wide Web and providing anonymity.



Communication between two computers (shown in grey) connected through a third computer (shown in red) acting as a proxy. Bob does not know whom the information is going to, which is why proxies can be used to protect privacy.

Reverse Proxy Server

In computer networks, a reverse proxy is a type of proxy server that retrieves resources on behalf of a client from one or more servers. These resources are then returned to the client as though they originated from the proxy server itself.[1] While a forward proxy acts as an intermediary for its associated clients to contact any server, a reverse proxy acts as an intermediary for its associated servers to be contacted by any client.



A reverse proxy taking requests from the Internet and forwarding them to servers in an internal network. Those making requests to the proxy may not be aware of the internal network.

Workflow?

7. What is the difference between IPv4 and IPv6, HTTP and HTTPS?

IPv4	IPv6
IPv4 addresses are 32 bit length.	**IPv6 addresses** are 128 bit length.
IPv4 addresses are **binary numbers** represented in decimals.	**IPv6 addresses** are **binary numbers** represented in**hexadecimals**.
IPSec support is only optional.	Inbuilt **IPSec** support.
Fragmentation is done by sender and forwarding routers.	**Fragmentation** is done only by sender.
No packet flow identification.	Packet flow identification is available within the **IPv6 header**using the **Flow Label** field.
Checksum field is available in **IPv4 header**	No checksum field in **IPv6 header**.
Options fields are available in **IPv4 header**.	No option fields, but **IPv6 Extension headers** are available.
Address Resolution Protocol (ARP) is available to map IPv4 addresses to MAC addresses.	**Address Resolution Protocol (ARP)** is replaced with a function of **Neighbor Discovery Protocol (NDP)**.
Internet Group Management Protocol (IGMP) is used to manage multicast group membership.	IGMP is replaced with Multicast Listener Discovery (MLD) messages.
Broadcast messages are available.	**Broadcast messages** are not available. Instead a link-local scope "All nodes" **multicast IPv6**

	address (FF02::1) is used for broadcast similar functionality.
Manual configuration (Static) of IPv4 addresses or DHCP (Dynamic configuration) is required to configure IPv4 addresses.	Auto-configuration of addresses is available.

8. Even though you are securing your network through Router, why do you use firewall?

   Routers Function as Hardware Firewalls but still we use software firewall because of the following reasons:

   - A hardware firewall sits between your computer and the Internet, while a software firewall sits between your computer and the network. If other computers on your network become infected, the software firewall can protect your computer from them.
   - Software firewalls allow you to easily control network access on a per-application basis. In addition to controlling incoming traffic, a software firewall can prompt you when an application on your computer wants to connect to the Internet and allow you to prevent the application from connecting to the network. This feature is easy to use with a third-party firewall, but you can also prevent applications from connecting to the Internet with the Windows firewall.

9. How do you say that your security protection is better than a general Internet security setup?
10. Why not are you using Reverse Proxy?

    Of course on big disadvantage is that if your **proxy server crashes nothing will work** as it is all dependent upon the proxy server. Another disadvantage is that if your **reverse proxy is compromised you will be providing them with a comprehensive view of the internal network.** The final disadvantage is that you **may see some issues with speed**. Speed must be weighed against the value of a cache, and how much time it takes the request to traverse several firewalls. This issue may be trivial or may not be an issue at all. Much of it depends on the resources and number of requests that your server receives.

11. What is OS hardening?

    Hardening of the OS is the act of configuring an OS securely, updating it, creating rules and policies to help govern the system in a secure manner, and removing unnecessary applications and services. This is done to minimize a computer OS's exposure to threats and to mitigate possible risk. Although it is impossible to reduce risk to zero.

    (Or)

    Os Hardening is nothing but making an operating system more secure. It often requires numerous actions such as configuring system and network components properly, deleting unused files and

applying the latest patches. There are hardening checklists available for popular operating systems that administrators can follow.

1. Why did you choose AES Encryption for capturing final grade? Latest encryption algorithm that you know?

   Benefits of AES Encryption over other encryption algorithms are :

   - High efficiency
   - not complex
   - high secure
2. How did you know about Lamda-probe, Apache Jmeter, Jhat and Jconsole profiling tools?
3. How do you improve performance of an application which is already developed?
4. What is knowledgebase system? Is it similar to Wikipedia?

   A knowledge-based system (KBS) is a computer program that reasons and uses a knowledge base to solve complex problems.

5. What could be the alternative method if the opensource FreeTTS was not available?

   FreeTTS is an open source speech synthesis system written entirely in the Java programming language. It is based upon Flite. FreeTTS is an implementation of Sun's Java Speech API.

   If the opensource FreeTTS was not available :

   A number of markup languages have been established for the rendition of text as speech in an XML-compliant format. The most recent is Speech Synthesis Markup Language (SSML), which became a W3C recommendation in 2004. Older speech synthesis markup languages include Java Speech Markup Language (JSML) andSABLE. Although each of these was proposed as a standard, none of them have been widely adopted.

   Speech synthesis markup languages are distinguished from dialogue markup languages. VoiceXML, for example, includes tags related to speech recognition, dialogue management and touchtone dialing, in addition to text-to-speech markup.

6. Real time example of n-tier architecture system

   In software engineering, multi-tier architecture (often referred to as n-tier architecture) is a client-server architecture in which, the presentation, the application processing and the data management are logically separate processes. For example, an application that uses middleware to service data requests between a user and a database employs multi-tier architecture. The most widespread use of "multi-tier architecture" refers to three-tier architecture.

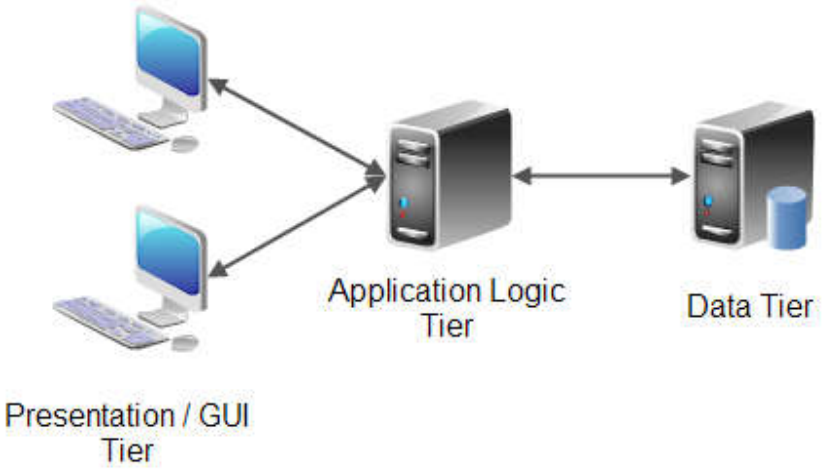
   Three tier architecture having three layers. They are

Presentation tier: This is the topmost level of the application. The presentation tier displays information related to such services as browsing merchandise, purchasing and shopping cart contents. It communicates with other tiers by which it puts out the results to the browser/client tier and all other tiers in the network. (In simple terms it is a layer which users can access directly such as a web page, or an operating systems GUI)

Application tier (business logic, logic tier, or middle tier): The logical tier is pulled out from the presentation tier and, as its own layer, it controls an application's functionality by performing detailed processing.

Data tier

The data tier includes the data persistence mechanisms (database servers, file shares, etc.) and the data access layer that encapsulates the persistence mechanisms and exposes the data. The data access layer should provide an Application Programming Interface (API) to the application tier that exposes methods of managing the stored data without exposing or creating dependencies on the data storage mechanisms. Avoiding dependencies on the storage mechanisms allows for updates or changes without the application tier clients being affected by or even aware of the change. As with the separation of any tier, there are costs for implementation and often costs to performance in exchange for improved scalability and maintainability.



In the web development field, three-tier is often used to refer to websites, commonly electronic commerce websites, which are built using three tiers:

- A front-end web server serving static content, and potentially some cached dynamic content. In web based application, Front End is the content rendered by the browser. The content may be static or generated dynamically.
- A middle dynamic content processing and generation level application server, for example Ruby on Rails, Java EE, ASP.NET, PHP, ColdFusion, Perl, Pythonplatform.
- A back-end database or data store, comprising both data sets and the database management system software that manages and provides access to the data.

7. What are Unicode fonts? Can Unicode be used for any language? Is it necessary for the developer to know the language for which he is using Unicode?
8. If the approver for approving the notice is not available. And I need to post & display a notice immediately, how do I achieve this?
9. How frequently you synchronize SAMWAD database with COWAA database? What happens if one the database is down?
10. What are the security measures undertaken for development of APAR scheduler?
11. Difference between SNMPv1, SNMPv2 and SNMPv3.
    - SNMP version 1: the oldest flavor. Easy to set up – only requires a plaintext community. The biggest downsides are that it does not support 64 bit counters, only 32 bit counters, and that it has little security. A community string sent in plaintext, possibly from a restricted range of allowed IP addresses, is as good as the security gets. In other words, no security from someone with access to the network – such a person will be able to see the community string in plaintext, and spoofing a UDP packet's source IP is trivial. (On the other hand, if your device is set up to only allow SNMP read only access – the risk is fairly small, and confined to evil people with access to your network. If you have evil people with this access, SNMP is probably not what you need to be worrying about.)
    - SNMP version 2c: in practical terms, v2c is identical to version 1, except it adds support for 64 bit counters. This matters, especially for interfaces. Even a 1Gbps interface can wrap a 32 bit counter in 34 seconds. Which means that a 32 bit counter being polled at one minute intervals is useless, as it cannot tell the difference between successive values of 30, 40 due to the fact that only 10 octets were sent in that minute, or 30, 40 due to the fact that 4294967306 (2^32 +10) octets were sent in that minute. Most devices support snmp V2c nowadays, and generally do so automatically. There are some devices that require you to explicitly enable v2c – in which case, you should always do so. There is no downside.

- SNMP version 3: adds security to the 64 bit counters. SNMP version 3 adds both encryption and authentication, which can be used together or separately. Setup is more complex than just defining a community string – but then, what security is not? But if you require security, this is the way to do it.

Hacking:

A hacker is someone who seeks and exploits weaknesses in a computer system or computer network. Hackers may be motivated by a multitude of reasons, such as profit, protest, challenge.

White hat

A white hat hacker breaks security for non-malicious reasons, perhaps to test their own security system or while working for a security company which makes security software. The term "white hat" in Internet slang refers to an ethical hacker. This classification also includes individuals who perform penetration tests and vulnerability assessments within a contractual agreement. The EC-Council, also known as the International Council of Electronic Commerce Consultants, is one of those organizations that have developed certifications, courseware, classes, and online training covering the diverse arena of ethical hacking.

Black hat

A "black hat" hacker is a hacker who "violates computer security for little reason beyond maliciousness or for personal gain" (Moore, 2005). Black hat hackers form the stereotypical, illegal hacking groups often portrayed in popular culture, and are "the epitome of all that the public fears in a computer criminal". Black hat hackers break into secure networks to destroy, modify, or steal data; or to make the network unusable for those who are authorized to use the network. Black hat hackers are also referred to as the "crackers" within the security industry and by modern programmers. Crackers keep the awareness of the vulnerabilities to themselves and do not notify the general public or the manufacturer for patches to be applied. Individual freedom and accessibility is promoted over privacy and security. Once they have gained control over a system, they may apply patches or fixes to the system only to keep their reigning control. Richard Stallman invented the definition to express the maliciousness of a criminal hacker versus a white hat hacker who performs hacking duties to identify places to repair.

Grey hat

A grey hat hacker lies between a black hat and a white hat hacker. A grey hat hacker may surf the Internet and hack into a computer system for the sole purpose of notifying the administrator that their system has a security defect, for example. They may then offer to correct the defect for a fee. Grey hat hackers sometimes find the defect of a system and publish the facts to the world instead of a group of people. Even though grey hat hackers may not necessarily perform hacking for their personal gain, unauthorized access to a system can be considered illegal and unethical.

Attacks

A typical approach in an attack on Internet-connected system is:

1. Network enumeration: Discovering information about the intended target.

2. Vulnerability analysis: Identifying potential ways of attack.

3. Exploitation: Attempting to compromise the system by employing the vulnerabilities found through the vulnerability analysis.

Security exploits

A security exploit is a prepared application that takes advantage of a known weakness. Common examples of security exploits are SQL injection, cross-site scripting and cross-site request forgery which abuse security holes that may result from substandard programming practice. Other exploits would be able to be used through File Transfer Protocol (FTP), Hypertext Transfer Protocol (HTTP), PHP, SSH, Telnet and some Web pages. These are very common in Web site and Web domain hacking.

Vulnerability

a vulnerability is a weakness which allows an attacker to reduce a system's information assurance. Vulnerability is the intersection of three elements: a system susceptibility or flaw, attacker access to the flaw, and attacker capability to exploit the flaw. To exploit a vulnerability, an attacker must have at least one applicable tool or technique that can connect to a system weakness.

Computer virus

A virus is a self-replicating program that spreads by inserting copies of itself into other executable code or documents. By doing this, it behaves similarly to a biological virus, which spreads by inserting itself into living cells. While some viruses are harmless or mere hoaxes, most are considered malicious.

Computer worm

Like a virus, a worm is also a self-replicating program. It differs from a virus in that (a.) it propagates through computer networks without user intervention; and (b.) does not need to attach itself to an existing program. Nonetheless, many people use the terms "virus" and "worm" interchangeably to describe any self-propagating program.

Malware

Malware, short for malicious software, is any software used to disrupt computer operation, gather sensitive information, or gain access to private computer systems.[1] Malware is defined by its malicious intent, acting against the requirements of the computer user, and does not include software that causes unintentional harm due to some deficiency. The term badware is sometimes used, and applied to both true (malicious) malware and unintentionally harmful software.

Viruses

A computer program usually hidden within another seemingly innocuous program that produces copies of itself and inserts them into other programs or files, and that usually performs a malicious action (such as destroying data).

Trojan horses

For a malicious program to accomplish its goals, it must be able to run without being detected, shut down, or deleted. When a malicious program is disguised as something normal or desirable, users may unwittingly install it. This is the technique of the Trojan horse or trojan. In broad terms, a Trojan horse is any program that invites the user to run it, concealing harmful or malicious executable code of any description. The code may take effect immediately and can lead to many undesirable effects, such as encrypting the user's files or downloading and implementing further malicious functionality.

In the case of some spyware, adware, etc. the supplier may require the user to acknowledge or accept its installation, describing its behavior in loose terms that may easily be misunderstood or ignored, with the intention of deceiving the user into installing it without the supplier technically in breach of the law.

Rootkits

Once a malicious program is installed on a system, it is essential that it stays concealed, to avoid detection. Software packages known as rootkits allow this concealment, by modifying the host's operating system so that the malware is hidden from the user. Rootkits can prevent a malicious process from being visible in the system's list of processes, or keep its files from being read.

Some malicious programs contain routines to defend against removal, not merely to hide themselves. An early example of this behavior is recorded in the Jargon File tale of a pair of programs infesting a Xerox CP-V time sharing system:

Each ghost-job would detect the fact that the other had been killed, and would start a new copy of the recently-stopped program within a few milliseconds. The only way to kill both ghosts was to kill them simultaneously (very difficult) or to deliberately crash the system.

Backdoors

A backdoor is a method of bypassing normal authentication procedures, usually over a connection to a network such as the Internet. Once a system has been compromised, one or more backdoors may be installed in order to allow access in the future, invisibly to the user.

The idea has often been suggested that computer manufacturers preinstall backdoors on their systems to provide technical support for customers, but this has never been reliably verified. It was reported in 2014 that US government agencies had been diverting computers purchased by those

considered "targets" to secret workshops where software or hardware permitting remote access by the agency was installed, considered to be among the most productive operations to obtain access to networks around the world. Backdoors may be installed by Trojan horses, worms, implants, or other methods.

SPAM

Email spam, also known as as junk email or unsolicited bulk email (UBE), is a subset of electronic spam involving nearly identical messages sent to numerous recipients by email.

Content management system:

A content management system (CMS) is a computer application that allows publishing, editing and modifying content, organizing, deleting as well as maintenance from a central interface.[4] Such systems of content management provide procedures to manage workflow in a collaborative environment.[5] These procedures can be manual steps or an automated cascade. CMSs have been available since the late 1990s.

CMSs are often used to run websites containing blogs, news, and shopping. Many corporate and marketing websites use CMSs. CMSs typically aim to avoid the need for hand coding, but may support it for specific elements or entire pages.

OpenCms

OpenCms is an open source content management system[1] written in Java.[2][3] It is distributed by Alkacon Software under the LGPL license.[4] OpenCms requires a JSP Servlet container such as Apache Tomcat.[4]

It is a CMS application with a browser-based work environment, asset management, user management, workflow management, a WYSIWYG editor, internationalization support, content versioning, and many more features including proxying of requests to another endpoint.[2]

<u>General OpenCms features:</u>

- The page editor allows WYSIWYG inline editing of web pages and arrangement of content by drag & drop.
- The form based editor allows editing of structured content in a well defined form mask.
- The sitemap editor allows to create new pages and re-arrange the navigation tree by drag & drop.
- Responsive demo template based on Bootstrap 3.
- Content creation for mobile devices with preview and device specific content control.
- Structured contents can be defined using a simple XML schema.
- Easy to use "Online / Offline" workflow, changes must be approved before they become visible.
- Link management for all internal resources with broken link detection.

- Integrated image scaling and cropping.
- SEO features with automatic sitemap.xml generation and page alias support.
- Full featured user management that supports the concept of "Organizational Units" (OUs).
- Allows management of multiple websites within one installation.
- Contents can be served dynamically or exported to static HTML files.
- Direct access to the OpenCms content repository over a shared network drive.
- CMIS and WebDAV access to the OpenCms content repository.
- Integrates Apache SOLR for powerful content searching and noSQL like queries.
- Full text search for web pages as well as in office documents like PDF, MS Office and Open Office.
- Extensions can be added through a flexible module system.
- The "time warp" feature allows to view resources which are expired or not yet released.
- JSP integration for dynamic functionality in templates, dynamic forms etc.

Improved in OpenCms 9.5.1:

- Added "DependentSelectWidget" that obtains values depending on other content fields.
- Added alternative login handler to authorize users with a cookie.
- Added option to search for expired / unreleased resources to all gallery dialogs.
- Added option "galleryselect" to select a gallery to file selector widget.
- Added further configuration options for OpenCms auto-setup.
- Added option to use permanent instead of temporary redirects from a non-secure to a secure server.
- Added Solr "rows" parameter in workplace's source search.
- Added support for configuring a JDBC connection in the OpenCms db properties.
- Demo template updated to Bootstrap 3.3.2.

Drupal

Drupal is a free and open-source content-management framework written in PHP and distributed under the GNU General Public License.[3][5][6] It is used as a back-end framework for at least 2.1% of all Web sites worldwide[7][8] ranging from personal blogs to corporate, political, and government sites including WhiteHouse.gov and data.gov.uk.[9] It is also used for knowledge management and business collaboration.

Joomla

Joomla is a free and open-source content management system (CMS) for publishing web content. It is built on a model–view–controller web application framework that can be used independently of the CMS.

Joomla is written in PHP, uses object-oriented programming (OOP) techniques (since version 1.5[2]) and software design patterns,[3] stores data in a MySQL, MS SQL (since version 2.5), or PostgreSQL (since version 3.0) database,[4][5] and includes features such as page caching, RSS feeds, printable versions of pages, news flashes, blogs, polls, search, and support for language internationalization.

Triple DES (3DES)

In cryptography, Triple DES (3DES) is the common name for the Triple Data Encryption Algorithm (TDEA or Triple DEA) symmetric-key block cipher, which applies the Data Encryption Standard (DES) cipher algorithm three times to each data block.

The original DES cipher's key size of 56 bits was generally sufficient when that algorithm was designed, but the availability of increasing computational power made brute-force attacks feasible. Triple DES provides a relatively simple method of increasing the key size of DES to protect against such attacks, without the need to design a completely new block cipher algorithm.

Trigger:

A database trigger is procedural code that is automatically executed in response to certain events on a particular table or view in a database. The trigger is mostly used for maintaining the integrity of the information on the database. For example, when a new record (representing a new worker) is added to the employees table, new records should also be created in the tables of the taxes, vacations and salaries.

Advantages of trigger:

1) Triggers can be used as an alternative method for implementing referential integrity constraints.
2) By using triggers, business rules and transactions are easy to store in database and can be used consistently even if there are future updates to the database.
3) It controls on which updates are allowed in a database.
4) When a change happens in a database a trigger can adjust the change to the entire database.
5) Triggers are used for calling stored procedures.

Disadvantages of trigger:

1) Programmers don’t have full control: Since business rules are hidden, programmers don’t have full control over the database. BY this, a program can generate several actions.
2) Increase in complexity: Triggers increase the complexity of a database as they have hard coded rules already implemented.
3) Decrease in performance of the database: By the complex nature of the database programs take more time to execute and there can be hidden performance downtimes.

Stored procedures:

Stored procedures have been viewed as the de facto standard for applications to access and manipulate database information through the use of codified methods, or "procedures." This is largely due to what they offer developers: the opportunity to couple the set-based power of SQL with the iterative and conditional processing control of code development. Developers couldn't be happier about this; finally, instead of writing inline SQL and then attempting to manipulate the data from within the code, developers could take advantage of:

1) Familiar Coding Principles
    - Iterative Loops
    - Conditionals
    - Method Calls (the stored procedure itself is built and similarly called like a method)
2) One-time, One-place Processing
    - Instead of having inline SQL code spread throughout the application, now sections of SQL code can be encapsulated into chunks of named methods that are easily identifiable and accessible all within one location – the "Stored Procedure" folder of the database.
    - All complex data processing can now be performed on the server, allowing the client processing to focus more on presentation rather than manipulation of data.

Advantages of Using Stored Procedures

Stored procedures are so popular and have become so widely used and therefore expected of Relational Database Management Systems (RDBMS) that even MySQL finally caved to developer peer pressure and added the ability to utilize stored procedures to their very popular open source database. The list below details why stored procedures have gained such a stalwart following among application developers (and even Database Administrators for that matter):

Maintainability:

Because scripts are in one location, updates and tracking of dependencies based on schema changes becomes easier

Testing:

Can be tested independent of the application

Isolation of Business Rules:

Having Stored Procedures in one location means that there's no confusion of having business rules spread over potentially disparate code files in the application

Speed / Optimization:

Stored procedures are cached on the server

Execution plans for the process are easily reviewable without having to run the application

Utilization of Set-based Processing:

The power of SQL is its ability to quickly and efficiently perform set-based processing on large amounts of data; the coding equivalent is usually iterative looping, which is generally much slower

Security:

Limit direct access to tables via defined roles in the database

Provide an "interface" to the underlying data structure so that all implementation and even the data itself is shielded.

Securing just the data and the code that accesses it is easier than applying that security within the application code itself

Drawbacks to Using Stored Procedures:

There are certainly drawbacks to Stored Procedures that preclude them from being the one-stop shop solution to application database access. The list below contains some reasons why Stored Procedures might not be right for your application solution. Interestingly, you'll probably recognize some headings that also appear in the "Advantages" section above; this is because what one developer views as affirmative evidence for their use might cause another to see the same evidence to disprove their viability as a solution.

Limited Coding Functionality:

Stored procedure code is not as robust as app code, particularly in the area of looping (not to mention that iterative constructs, like cursors, are slow and processor intensive)

Portability:

Complex Stored Procedures that utilize complex, core functionality of the RDBMS used for their creation will not always port to upgraded versions of the same database. This is especially true if moving from one database type (Oracle) to another (MS SQL Server).

Testing:

Any data errors in handling Stored Procedures are not generated until runtime

Location of Business Rules:

Since SP's are not as easily grouped/encapsulated together in single files, this also means that business rules are spread throughout different Stored Procedures. App code architecture helps to ensure that business rules are encapsulated in single objects.

There is a general opinion that business rules / logic should not be housed in the data tier

Utilization of Set-based Processing:

Too much overhead is incurred from maintaining Stored Procedures that are not complex enough. As a result, the general consensus is that simple SELECT statements should not be bound to Stored Procedures and instead implemented as inline SQL.

Cost:

Depending on your corporate structure and separation of concern for development, there is the potential that Stored Procedure development could potentially require a dedicated database developer. Some businesses will not allow developers access to the database at all, requiring instead a separate DBA. This will automatically incur added cost.

Some companies believe (and sometimes it's true, but not always) that a DBA is more of a SQL expert than an application developer, and therefore will write better Stored Procedures. In that case, an extra developer in the form of a DBA is required.

SQL Query Processing - It involves 4 steps

- Parsing - converts to internal parse tree
- Query modification - parse tree is modified to incorporate any relevant view defn. and/security & integrity constraints
- Optimization - optimizer chooses an optimal access strategy for implementing that each query it process
- Execution

Functional Dependency (FD): is relationship b/w the attributes. An attribute ‘y’ is said to be fully functionally dependent on another attribute ‘x’, if the value of ‘x’ determines the value of ‘y’. Determinant: is an attribute on which other values depend.

Update anomaly: Results in data inconsistency because of possible partial update instead of the proper complete update.

Delete anomaly: Results in unintended loss of data because of possible deletion of data other than what must be deleted.

Addition anomaly: Results in inability to add the data to the database because of the absence of some data presently unavailable.

ACID

Atomicity - All actions in the Xact happens, or none happens - by commit tran or rollback transaction.

Consistency - If a Xact is consistent, and the DB starts consistent, it ends up consistent Isolation - Execution of one Xact is isolated from the other Transactions

Durability - A xact commits, its effects persists (WAL -write ahead log)

Locking - 3 - types

1. **Shared locks** - for read - all process can read - but no process can write- eg: select
2. **Exclusive Locks** - when a transaction gets an exclusive lock, other tran cannot obtain any other type of lock until the exclusive lock is released.
3. **Update locks** - allows many process to read, but no other process can get an exclusive or update lock- this lock is applied during the initial portion of an update or delete operation.- when the page is ready to be modified, the lock is promoted to an exclusive lock.

What is Digital Signature?

Digital signature is a mathematical scheme for demonstrating the authenticity of a digital message or document. This is compromising of the key generation, signing algorithm and signature verifying algorithms.

Key encryption is done based two types:

Symmetric - same key for encryption and de-encryption.

Asymmetric - two pairs of keys will be generated (private and public); private key used for encryption and public key for de-encryption.

Zero Based Budgeting (ZBB) System:

ZBB is budgeting system, where the budgeting process starts from a clean slate without any references to the past experience.

Ex: 1. Augmentation of Replacement Equipments

2. Salary and Admin are traditional budgeting

Considering the objectives of the organisation, priorities/alternatives and national needs.

What is cash flow?

The actual expenditure incurred during a specific period is called cash flow.

	SAN	**NAS**
Protocol	Fibre Channel	•TCP/IP

	Fibre Channel-to-SCSI	
Applications	• Mission-critical transaction-based database application processing • Centralized data backup • Disaster recovery operations • Storage consolidation	• File sharing in NFS and CIFS • Small-block data transfer over long distances • Limited read-only database access
Advantages	•High availability •Data transfer reliability •Reduced traffic on the primary network •Configuration flexibility •High performance •High scalability •Centralized management •Multiple vendor offerings	•Few distance limitations •Simplified addition of file sharing capacity •Easy deploymentand maintenance
	Data centric	Network centric
	Dedicated storage network	LAN or WAN
	Defines an architecture	Defines a product/appliance
	Low latency—thin protocol	High latency—overhead of TCP/IP
	Direct connection (switched)	Indirect connection (routed)
	Server owns a volume	Shared files, data

ISO's OSI Model

1. Application layer
2. Presentation Layer

3. Session Layer
4. Transport Layer
5. Network
6. Data link
7. Physical Layer

Object Oriented Programming

1. Encapsulation (Data and methods)
2. Reusability
3. Inheritance
4. Polymorphism

CMM

Capability Maturity Model (CMM) broadly refers to a process improvement approach that is based on a process model. A process model is a structured collection of practices that describe the characteristics of effective processes; the practices included are those proven by experience to be effective.

Level 1 - Initial (process is informal and ad-hoc; performance is unpredictable)

Level 2 - Repeatable (Project management system in place)

Level 3 - Defined (Software engineering and management processes defined and integrated)

Level 4 - Managed (product and process are quantified controlled)

Level 5 - Optimizing (process improvement is institutionalised)

Comparison between Oracle and Sybase

Developer perspective

- Oracle has no if / then syntax in SQL*Plus
- The arguments to stored procedures in Oracle take the form of function arguments. In other words, you use parenthesis. In Sybase, you list them out without the parenthesis.
- In Sybase, you can list the procedure arguments after the comment. In Oracle, you must list the arguments in juxtaposition with the procedure name. Oracle blows up if you don't.
- Oracle stored procedures cannot exit with a return code. They just return. Sybase procedures, by nature, always return with a status code. The Sybase language provides the mechanism to bubble

up errors from below. The Sybase programmer checks the status to see the procedure worked. Oracle's behavior is "let it blow up and clean up afterwards".

- There are differences in the way which Sybase and Oracle name and set variables. For example, Sybase begins all variable names with a '@' sign. This is pretty useful because you can always discern between column names and variables in Sybase.
- Oracle sets a variable using the ':=' while Sybase uses a select = statement.

What DWFS

DWFS system is a new work environment in which work moves from computer to computer digitally based on the events generated by the system.

Presently lot of paper movement and DWF system very thing is done online with proper authentication using digital signature.

1. Objectives of DWFS
    - Processes re-engineering for bring dramatic changes in the system
    - Paperless system
    - Seamless flow of information
    - Transparent - accurate
2. Advantages of DWFS
    - Dramatic improvement in the system - leave 2-3 days from 10-15 days; LTC advance - 3 from 10 days and LTC settlement - 7 from 15 days
    - Speedy administration
    - Automatic application of business logic - error free operations
    - Redeployment of existing admin personnel
    - Near paper less office
    - Transparent - accurate
3. Why DWFS when COWAA system is available
    - Some advantages of COWAA
        - Uniform software across ISRO centres
        - Integrated database for all the administration areas
        - Automation of administration areas
    - However there are certain limitation of COWAA
        - Installation of Client software
        - Automation existing functions without optimizing
        - Inter centre connectivity

    - Department centric not employee centric
  - To overcome these limitations, we are planning a new system by using new innovations in IT. Also it reduces paper which will provide eco friendly office and a common platform for all future applications needs.

1. What are all draw backs for COWAA system
2. What is AJAX
3. How do you secure your web applications
4. What is security threats explain SQL injection
5. What is 0 client
6. What is website and web portal
7. What is software
8. How do you test application and web based software
9. What is Web Embedded systems
10.