

Jul – Dec 2019
IT302 WTA



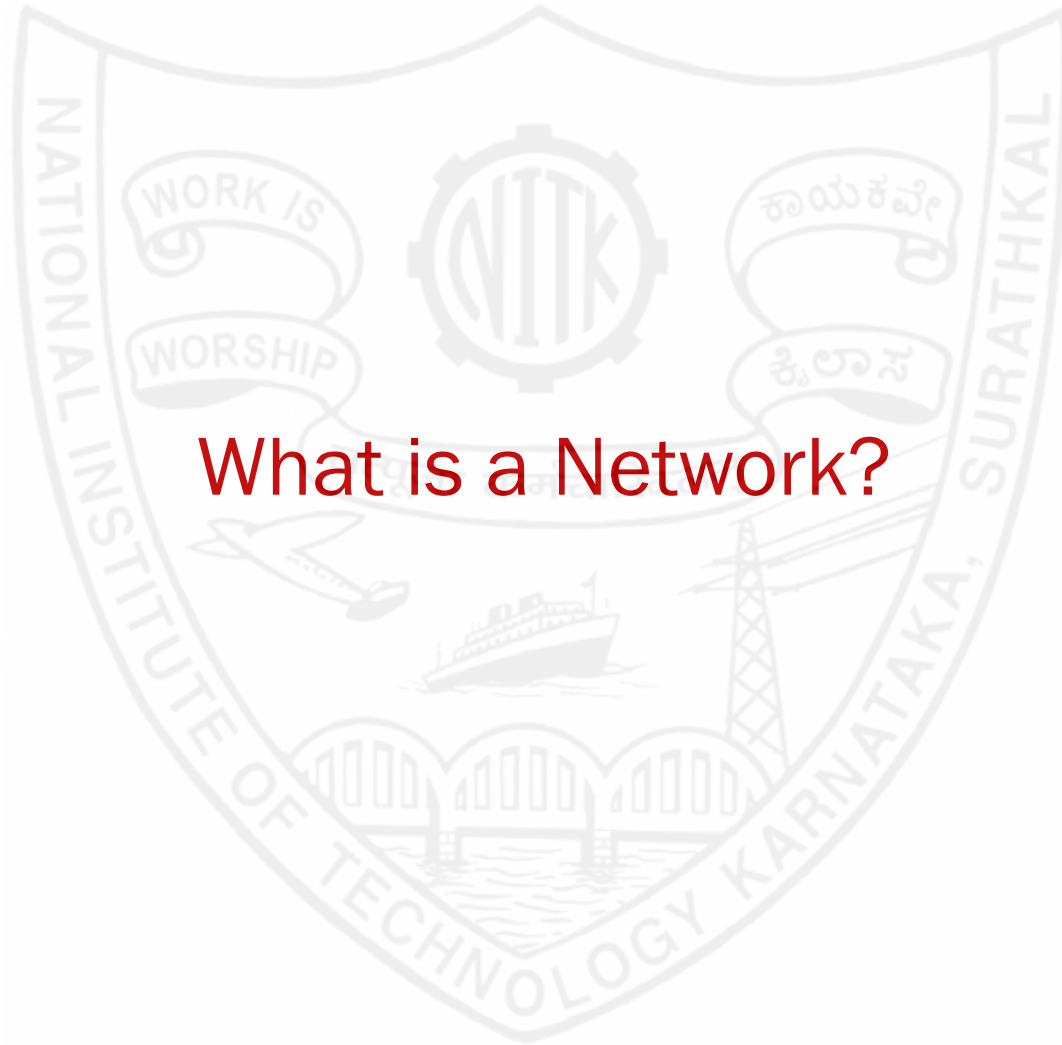
The World Wide Web

Basics, Concepts, Protocols

Topics of Coverage



- ▶ *Introduction to the WorldWideWeb:*
 - ▶ History of development
 - ▶ Internet and the WWW
 - ▶ Web Concepts
 - ▶ Web Architecture and Components
 - ▶ Web Protocols.



What is a Network?

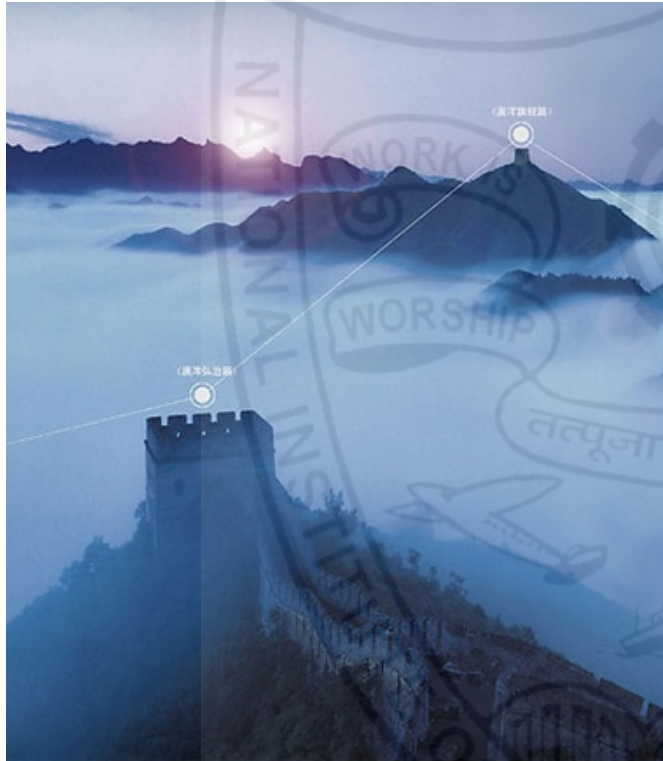
What is a Network?

- ❖ **topology:** The physical arrangement of entities in a network.
- ❖ **protocol:** The protocol defines a common set of rules/signals that entities on the network use to interact/communicate.
- ❖ **architecture :** e.g. Direct interaction (*peer-to-peer*) or Indirect interaction (*hierarchy, client/server architecture etc*)

Some early networks ...



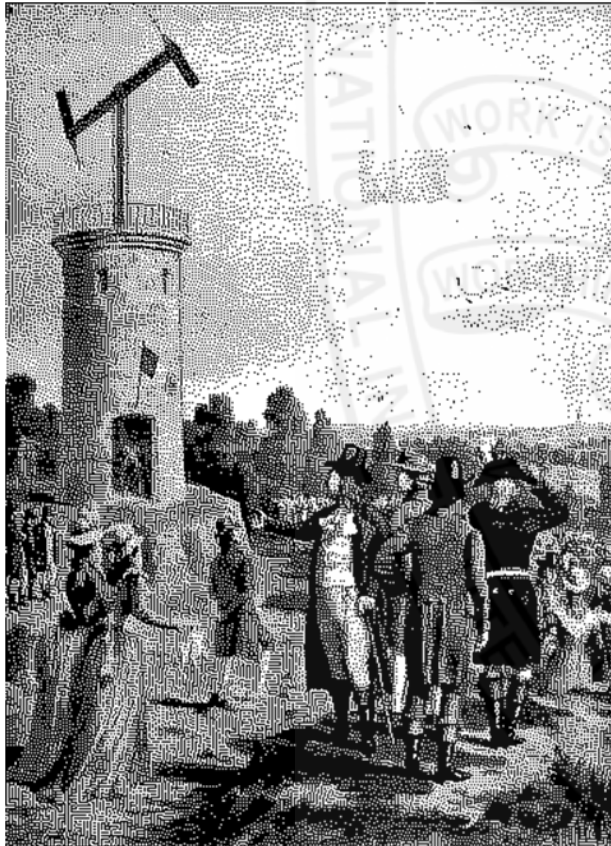
Beacon Chain Networking (400 BC)



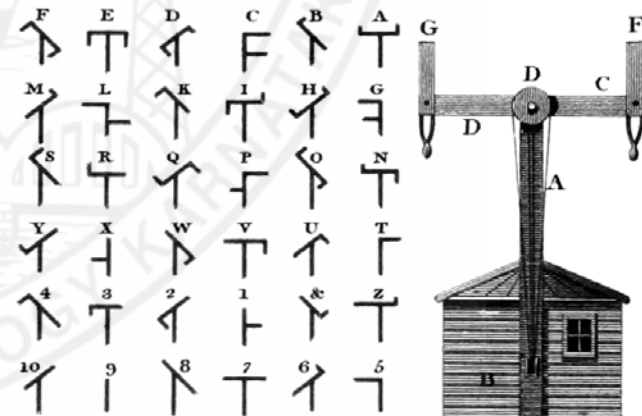
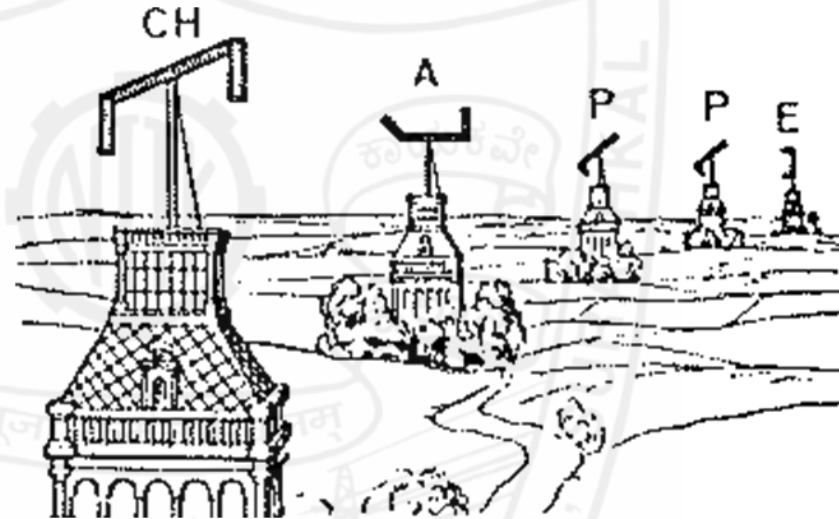
The Watch towers on the Great Wall of China

○ Signals used - smoke, fire, drums, coloured flags, gunshots

Chappe's Semaphore Network (1794)



First Line (Paris to Lille), 1794



The 'Victorian' Internet (1840s)

- ▶ The Telegraph (1839), Transatlantic Telegraph (1858)
 - ▶ Signals sent over wires that were established over vast distances.
 - ▶ Used Morse Code consisting of dots and dashes (short / long signals)
 - ▶ Electronic signal standard of +/- 15V.



A	• —
B	— • • •
C	— • — •
D	— • •
E	•
F	• • — •
G	— — •
H	• • • •
I	• •
J	• — — —
K	— • — —
L	• — • •
M	— —
N	— •
O	— — —
P	• — — •
Q	— — • —
R	• — •
S	• • •
T	—

U	• • —
V	• • • —
W	— — —
X	— • • —
Y	— — — •
Z	— — • •

1	• — — — —
2	• • — — —
3	• • • — —
4	• • • • —
5	• • • • •
6	— • • • •
7	— — • • •
8	— — — • •
9	— — — — •
0	— — — — —

Mondothèque (1934)

- ▶ a massive “search engine” envisioned by Paul Oltet.
 - ▶ An imaginary device that could be at the same time, an archival, instrument, workstation, catalogue and broadcasting machine.
- ▶ Goal: to collect, organize, and share all the world’s knowledge.

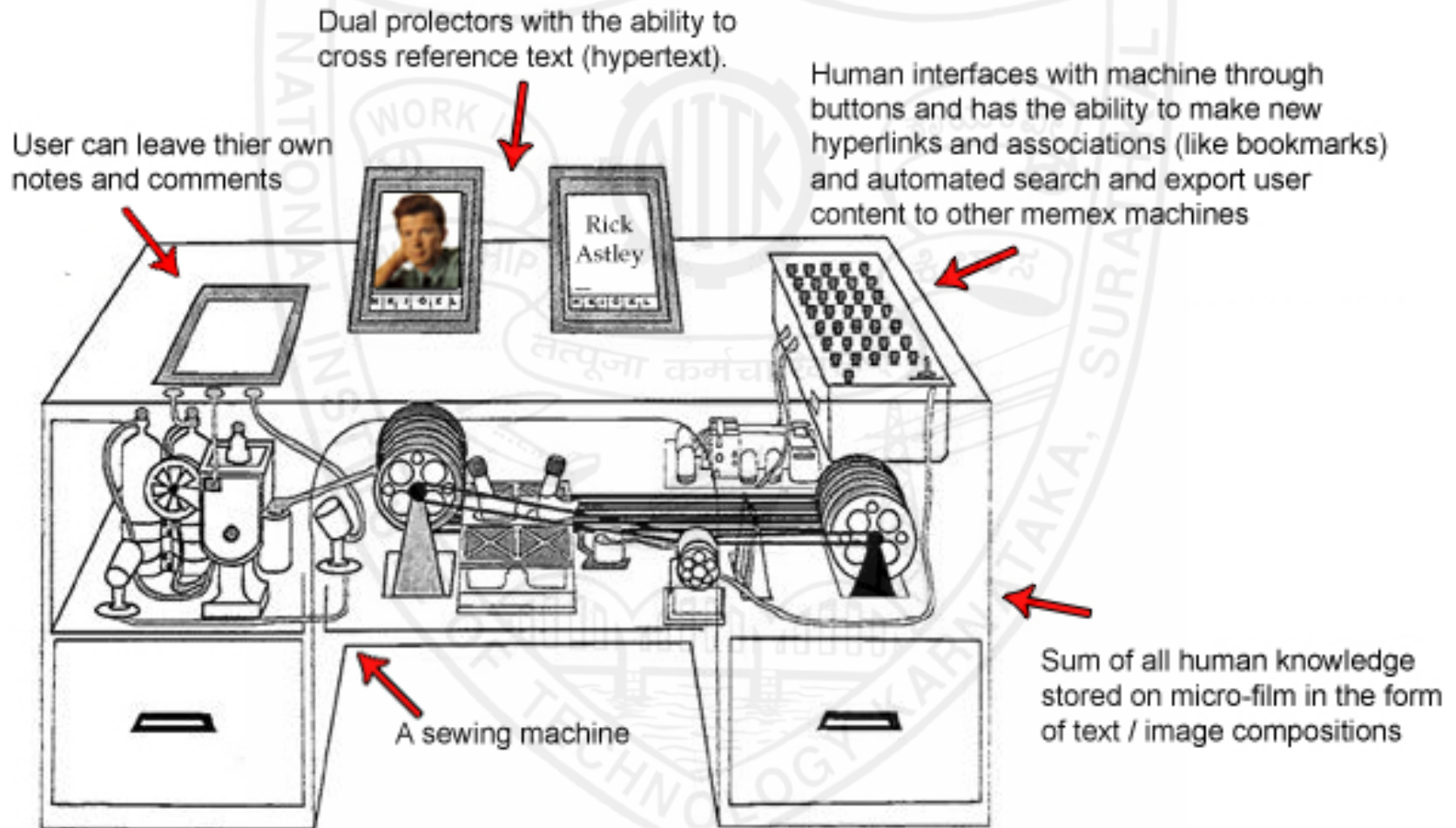




The Memex (1945)

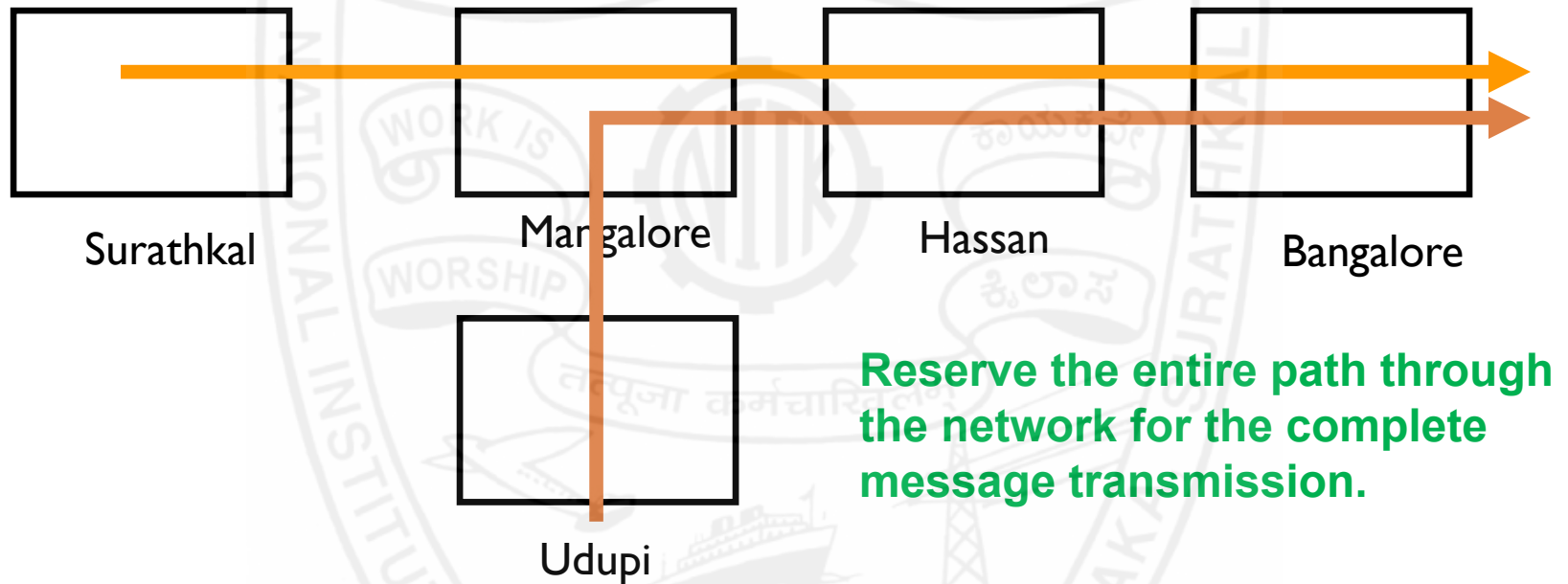
- ▶ hypothetical system that Vannevar Bush described in his 1945 *The Atlantic Monthly* article "As We May Think".
- ▶ Envisioned as a device in which individuals would compress and store all of their books, records, and communications
- ▶ influenced the development of early hypertext systems
 - ▶ eventually leading to the creation of the World Wide Web
 - ▶ personal knowledge base software

The Memex (1945)



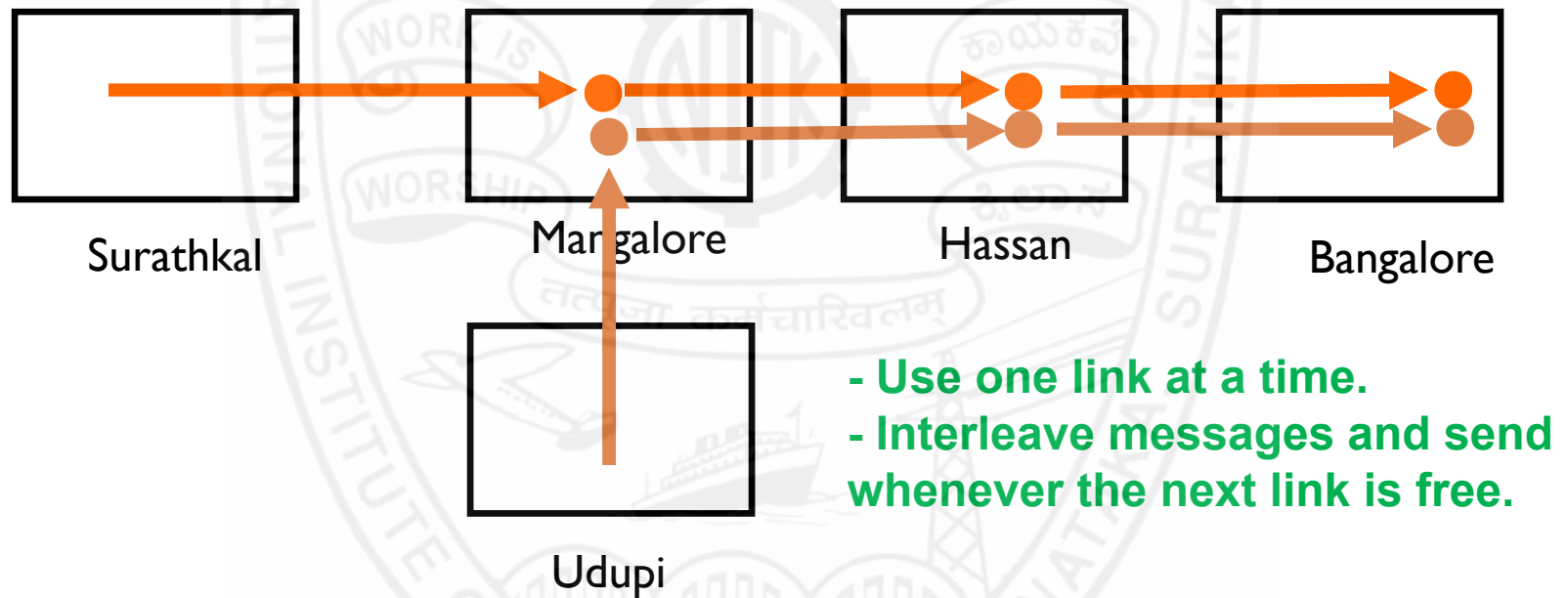


Circuit Switching (1878)



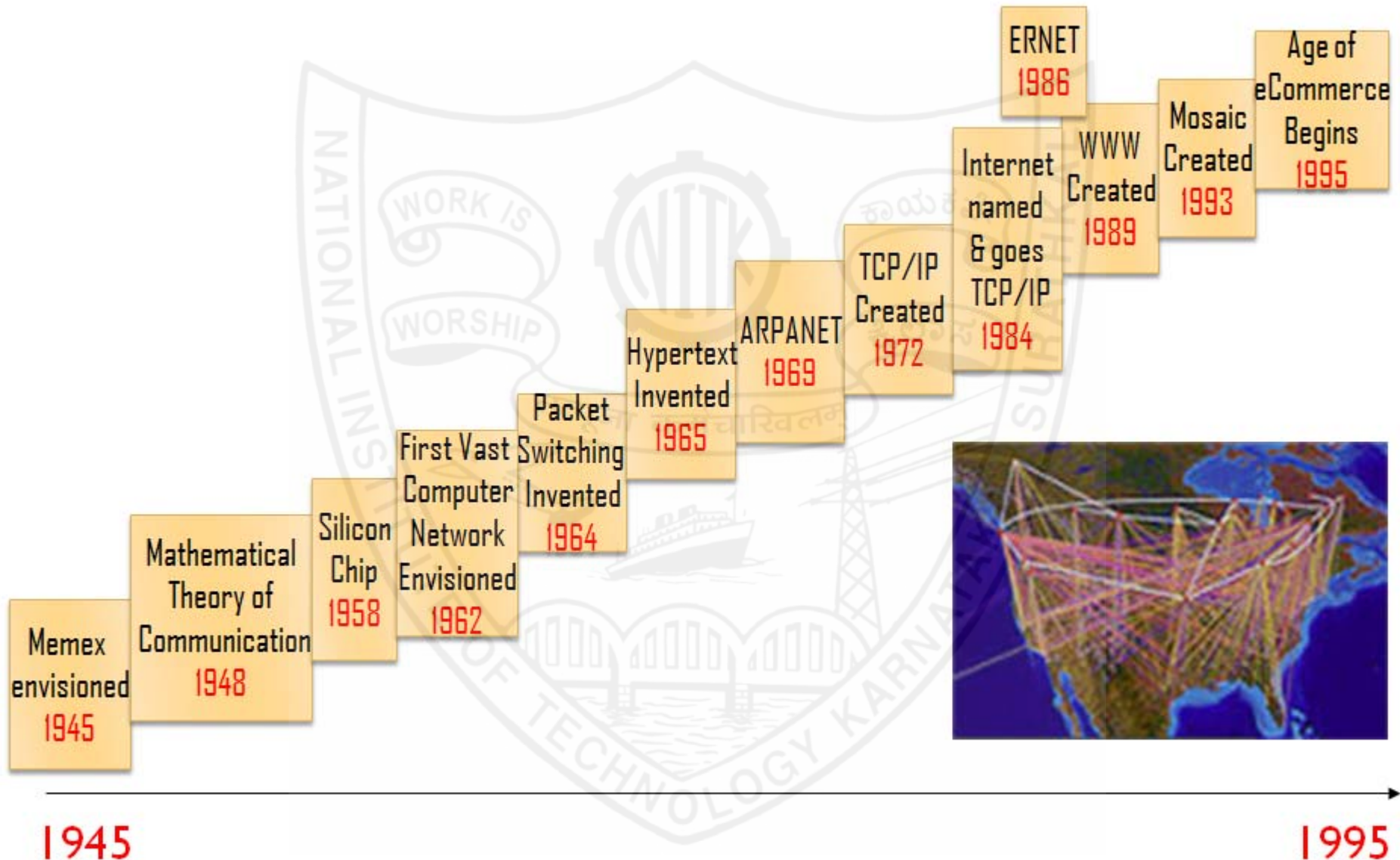
- ▶ Example: Landlines (Telephone Network)

Packet Switching (1964)



- ▶ Example: Modern networks, The Internet

A brief summary of Internet/Web evolution...



Internet and the WWW



▶ Difference ?

Internet is to **WWW** as **NITK** is to **IT**

- ▶ **Internet** - an infrastructure of millions of cables & computers, or several smaller sub-networks that share these cables and computers.
- ▶ **The WWW (or simply the Web)** - the largest and most popular sub-network on the internet.



Internet and WWW Development Forums

- ▶ Standards and specifications for the design of the Internet - ***Internet Engineering Task Force (IETF)***.
- ▶ Web standards like HTML, CSS, XML, RDF were introduced and standardized - ***World Wide Web Consortium (W3C)***
- * *IETF and W3C are open forums, allowing anyone interested in contribution to participate in the policy making/standardization process.*

Design Principles of the Internet and WWW



► Interoperability/Universality:

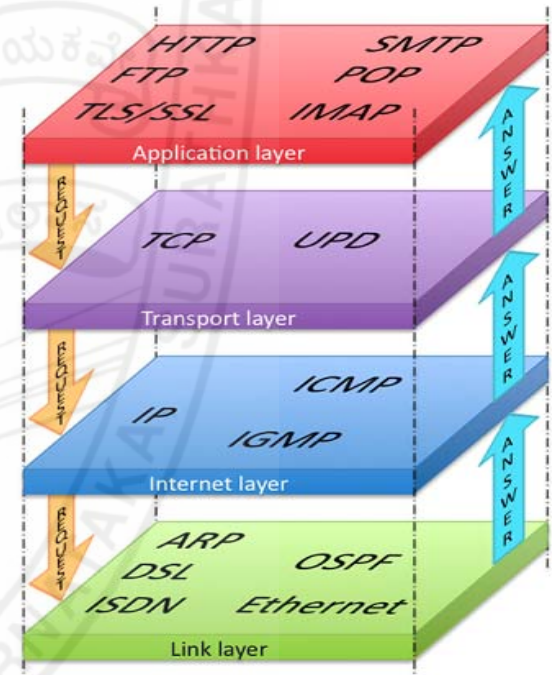
- Different implementations of Internet Protocols actually work together.
 - Adoption of open standards to facilitate interoperability.
- Systems can be assembled using client/server computers and software from different vendors.
- * For applications like e-commerce, buyers and sellers do not have to change/buy/upgrade software or systems to do business with each other.

Design Principles of the Internet and WWW



▶ Layering

- ▶ Internet protocols are designed to work in layers, higher layers building on the facilities provided by the lower layers.
- ▶ E.g.
 - ▶ TCP builds on IP to create reliable byte streams
 - ▶ Application layer protocols such as email build on TCP capabilities.

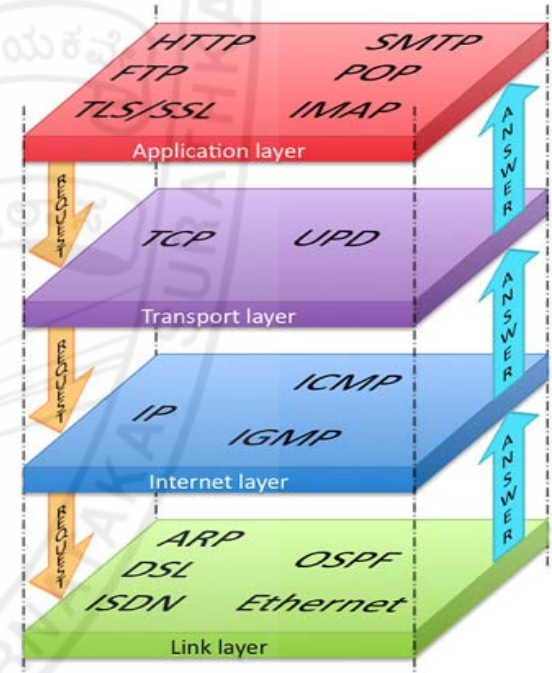


Design Principles of the Internet and WWW



► **Simplicity:**

- Layering of the Internet simplifies application development.
- IP hides the complexities of the layers below it.



Design Principles of the Internet and WWW



- ▶ **Uniform Naming and Addressing:**
 - ▶ **IP addressing:** Use of the dotted decimal form, assigning a 32/64 bit address to each computer connected to the network.
 - ▶ **DNS** - a standard way to translate human readable names for computers.
 - ▶ **URI** - a standard way to link and locate resources on the Web

Design Principles of the Internet and WWW



► End to End:

- The internet/Web is concerned only with the transmission of data, not its interpretation.
- Interpretation of data happens on the sending & receiving systems (computer/browser), not on the network.

* Analogy: Mailing a letter.

Design Principles of the Internet and WWW



► Decentralisation:

- No permission is needed from a central authority to post anything on the web, there is no central controlling node, and so no single point of failure ... and no “kill switch”!
- freedom from indiscriminate censorship and surveillance.*

* No longer holds good.

Design Principles of the Internet and WWW



► **Non-discrimination:**

- If I pay to connect to the internet with a certain quality of service, and you pay to connect with that or a greater quality of service, then we can both communicate at the same level.
- Also known as **Net Neutrality**.

Web System Architecture



Web System Architecture

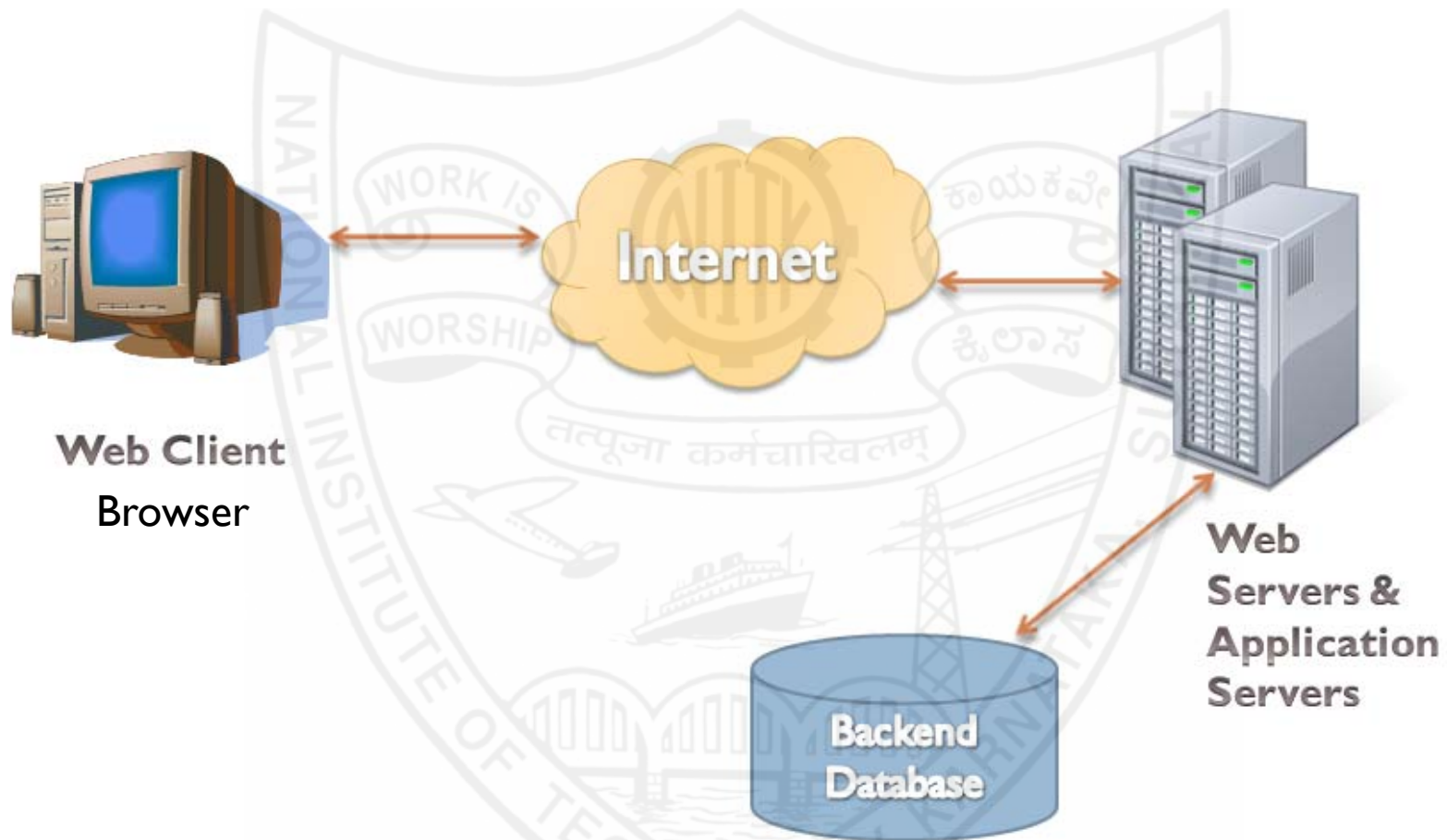


Fig: Basic Components of a Web based system



Web Clients

- ▶ Types of web clients may vary depending on the application.

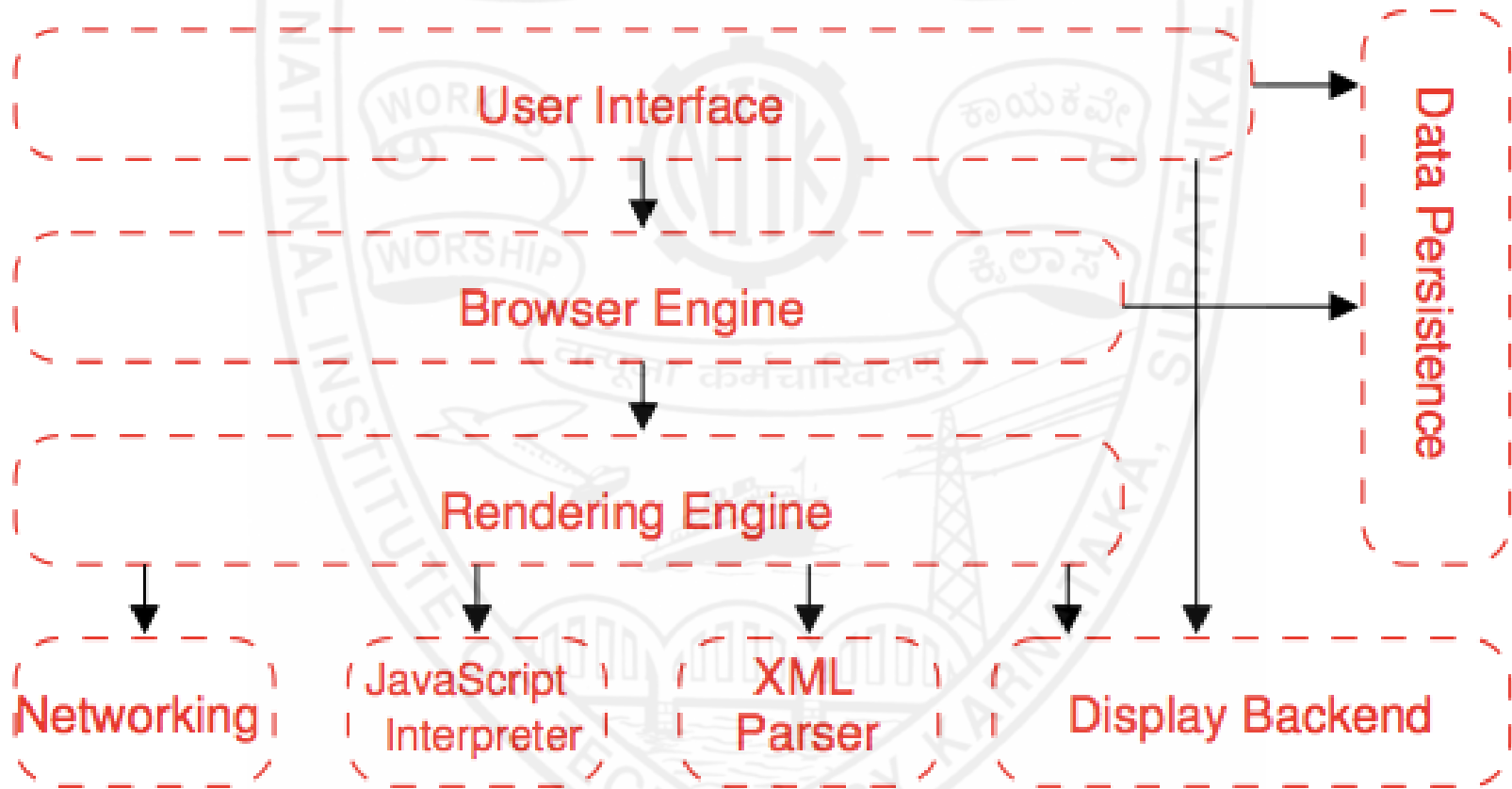
For e.g. —

- ▶ Web browsers (PC, mobile, text-only browsers, voice browsers)
- ▶ Chat browsers/interfaces
- ▶ **Software robots** (*no direct user contact*)
- ▶ **Software agents on the Web** (*initiated with user action*)

Web System Components (contd.)



Web Browser





Web Clients

- ▶ Basic tasks to be handled by a browser
 - ▶ Reformat the URL entered as a valid HTTP request message.
 - ▶ Use DNS to convert the host name to the appropriate IP address.
 - ▶ Establish a TCP connection using the IP address of the specified web server.
 - ▶ Send the HTTP request over TCP connection and wait for the server's response.
 - ▶ Display document contained in the response. (e.g. direct display of plain text, rendering HTML pages etc.)



- ▶ Some important additional functionalities provided by modern browsers
 - ▶ Automatic URL completion (...data persistence)
 - ▶ Script Management
 - ▶ Event Handling
 - ▶ Management of form GUI
 - ▶ Secure Communication
 - ▶ Session/Cookie Management
 - ▶ Handling extension mechanisms



Extension Mechanisms for the Web Client

- ▶ Mechanisms that add additional capabilities to the browser, either automatically or by user intervention.
- ▶ Types -
 - a. MIME Types or Internet Media Types
 - b. Plug-ins
 - c. Add-ons
 - d. Scripts
 - e. Applets
 - f. Controls

Extension Mechanisms for the Web Client



a. MIME Types or IM Types

- ▶ **Multipurpose Internet Mail Extensions/Internet media type**
 - ▶ standard identifier used on the Web/Internet to indicate the type of data that a file contains.
 - ▶ Common uses include :
 - ▶ In web browsers - how to display or output files that are not in HTML format
 - ▶ In search engines - to classify data files on the web.
 - ▶ In email clients - to identify attachment files.

Extension Mechanisms for the Web Client



a. MIME Types or IM Types (contd.)

- ▶ Each document is tagged with a *type* to identify what kind of resource it is.
- ▶ Format — *class/subclass*
- ▶ E.g. *text/html*, *image/gif*, *application/pdf*, *audio/mp3* etc.

Extension Mechanisms for the Web Client



a. MIME Types or IM Types (contd.)

file type	MIME type
avi	video/x-msvideo
bmp	image/bmp
css	text/css
doc	application/msword
dtd	application/xml-dtd
dvi	application/x-dvi
gif	image/gif
html	text/html
ico	image/x-icon
midi	audio/midi
mov	video/quicktime
mp3	audio/mpeg
mpeg	video/mpeg
pdf	application/pdf

Complete List maintained by **Internet Assigned Numbers Authority (IANA)**

<http://www.iana.org/assignments/media-types/media-types.xhtml>



b. Plug-ins

- ▶ Allow adding new capabilities for handling third party software in the browser itself rather than launching a separate application.
- ▶ Applications provide plug-ins to –
 - ▶ support easy adding of new features to browsers.
 - ▶ enable third-party developers to provide abilities to handle their formats in the native browser.
 - ▶ separate source code from an application because of incompatible software licenses.



b. Plug-ins (contd.)

- ▶ Features -
 - ▶ Must be manually installed by user before new data type can be used.
 - ▶ Browser plug-ins can modify the behavior of the browser. (e.g. adding new toolbar commands, menu items etc.)
- ▶ For e.g. Quicktime player, Adobe Reader, Macromedia Flash etc.

Extension Mechanisms for the Web Client

b. Plug-ins (contd.)



The screenshot displays the Firefox Add-ons Manager interface. On the left, a sidebar contains navigation links: 'Get Add-ons', 'Extensions', 'Appearance', 'Plugins' (selected), and 'Services'. The main panel shows a list of installed plugins, each with a blue brick icon, name, version, description, a 'More' link, and an activation status dropdown.

Plugin Name	Version	Description	More Info	Status
Adobe Acrobat	11.0.6.70	Adobe PDF Plug-In For Firefox and Netscape	More	Always Activate
Google Talk Plugin	5.1.4.17398	Version 5.1.4.17398	More	Always Activate
Google Talk Plugin Video Accelerator	0.1.44.29	Google Talk Plugin Video Accelerator version:0.1.44.29	More	Always Activate
Google Talk Plugin Video Renderer	5.1.4.17398	Version 5.1.4.17398	More	Always Activate
Google Update	1.3.22.3	Google Update	More	Always Activate
Warning: Java Deployment Toolkit 7.0.250.17 is known to be vulnerable. Use with caution. More Information				
Java Deployment Toolkit	7.0.250.17 10.25.2.17	NPRuntime Script Plug-in Library for Java(TM) Deploy	More	Ask to Activate
McAfee Security Scanner +	3.8.130.0	McAfee MSS+ NPAPI Plugin	More	Always Activate
Microsoft Office 2010	14.0.4761.1000			

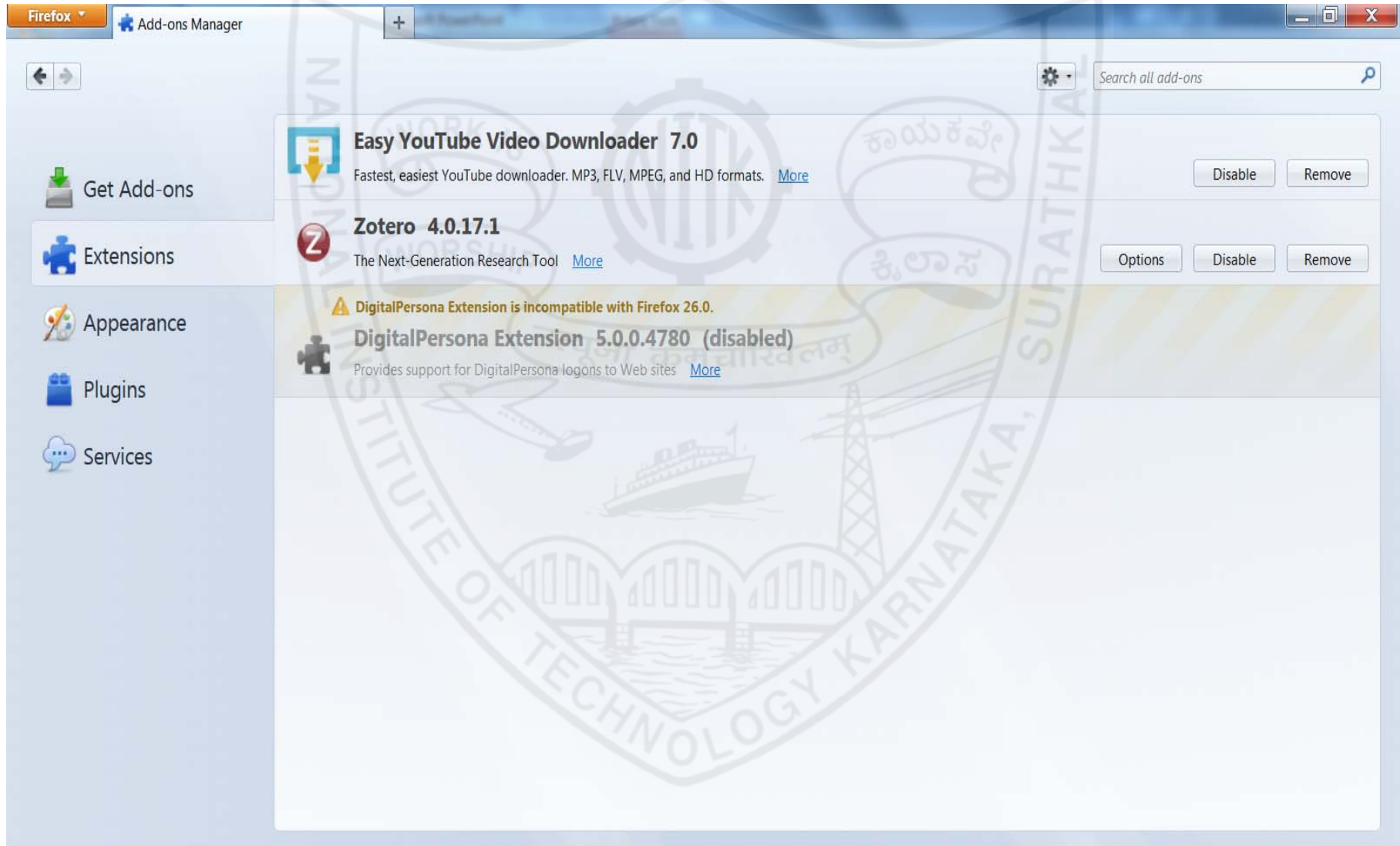


c. Add-ons

- ▶ used to refer to features that enhance an application.
 - ▶ Types – *extensions*, *themes* and *skins*.
- ▶ *An extension add-on* tailors the *core* features of an application by adding an optional module.
- ▶ *Theme or skin add-on* tailors the *outer* layers of an application to personalize functionality.

Extension Mechanisms for the Web Client

c. Add-ons (contd.)





d. Scripts

- ▶ Executable scripts can be embedded in web pages.
 - ▶ run when encountered on a page or when specified events occur.
 - ▶ written in languages like JavaScript, VBScript, ActionScript etc, and are executed by an interpreter in the browser when page is displayed.
- ▶ Scripts can modify page display and increase interactivity of the page, but have limited power. (for security reasons)

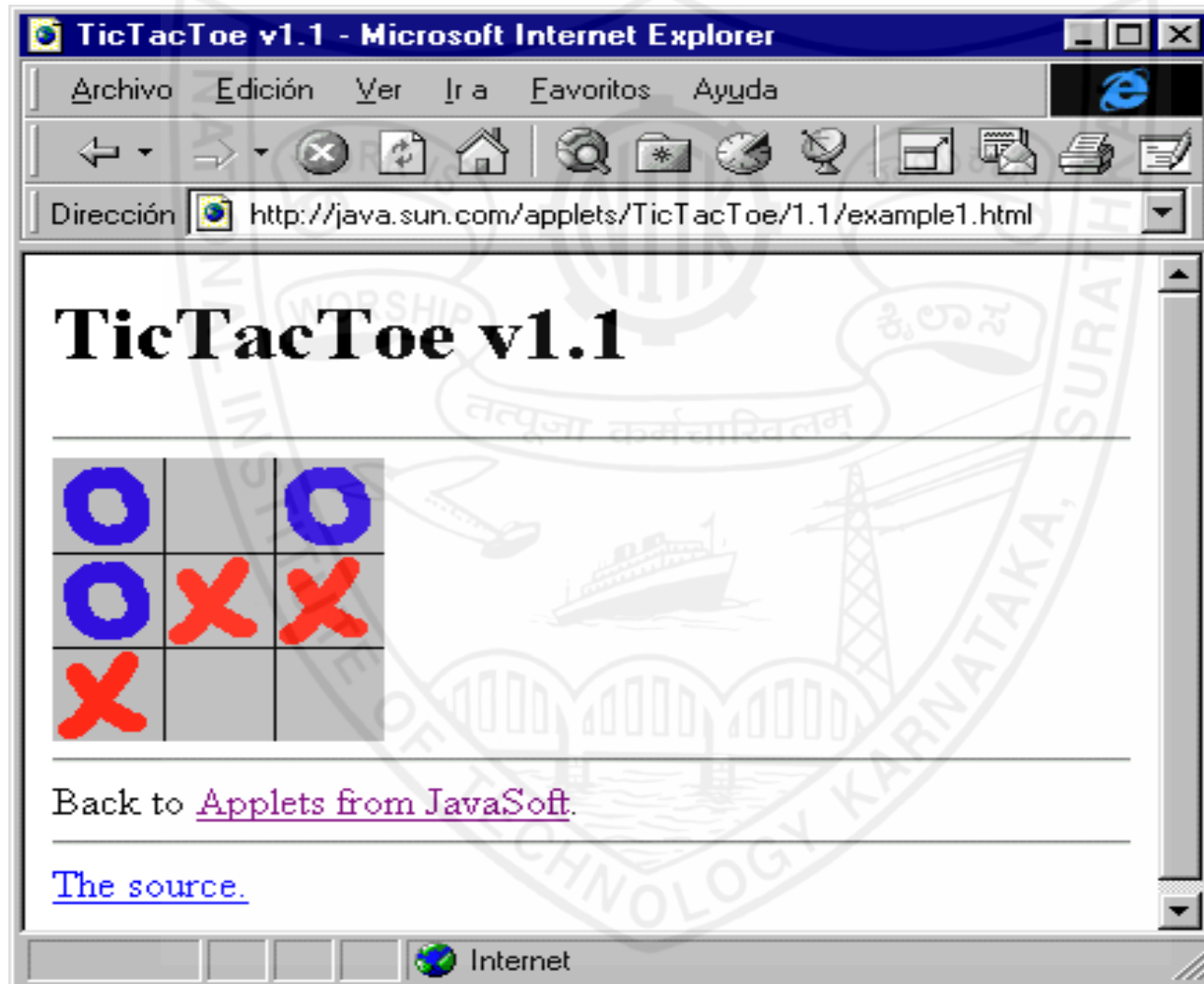


e. Applets

- ▶ Java applets are downloaded on demand from a server.
 - ▶ Used to create animation effects and other interactive behavior in the browser.
 - ▶ Are executed in the Java Virtual Machine supplied by the browser, thus limiting its effect on the system.
- ▶ User experience may be affected as download time can be significantly higher than that of scripts.

Extension Mechanisms for the Web Clients

e. Applets (contd.)





f. Controls

- ▶ Are software modules that are automatically downloaded and installed when a webpage containing them is encountered.
- ▶ On future references, it is automatically activated without having to be downloaded again.
 - ▶ Contain compiled code that can make changes to your machine.

Extension Mechanisms for the Web Client (contd.)



f. Controls

- ▶ Features -
 - ▶ Controls have full system control, hence user needs to allow only trusted providers.
 - ▶ Each control is digitally signed by its authoring organization. (Code Signing)
- ▶ Disadvantage:
 - ▶ If user decides not to install the control, then user experience with website may be disrupted.
 - ▶ E.g. Flickr Photo Uploader, software download controls, IE, MS-Office.

Web System Components (contd.)



Web Servers

- ▶ Basic functionality of a web server:
 - ▶ Accept HTTP requests from web clients and return an appropriate resource (if available) in the HTTP response.
 - ▶ Functionalities provided –
 - ▶ Communicating with TCP.
 - ▶ Handling multiple incoming requests and their corresponding responses.
 - ▶ Identifying resource location based on request URL.
 - ▶ Session Management.

Web System Components (contd.)

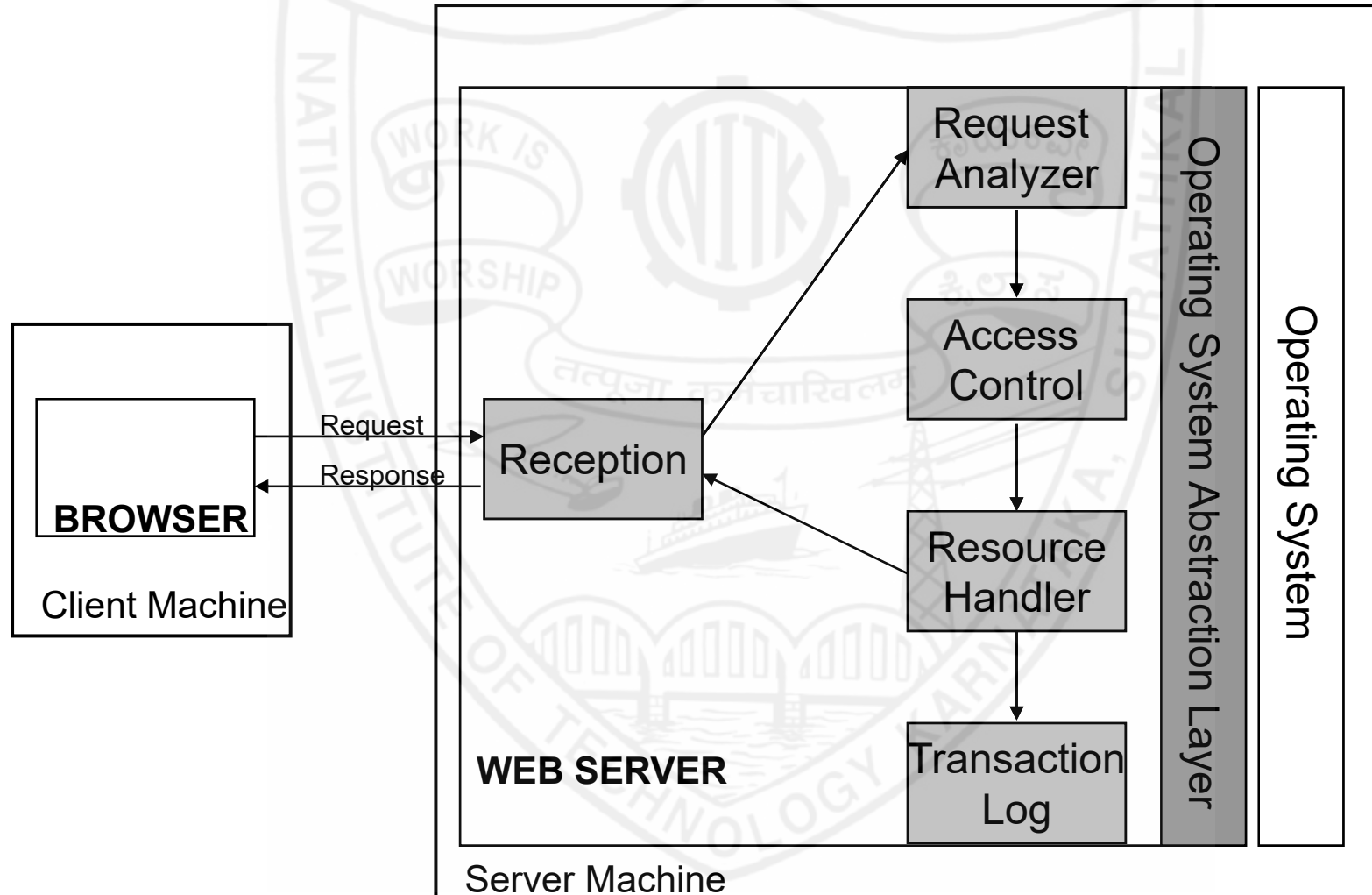


Web Servers

- ▶ *HTTPd* web server
 - ▶ supports Windows, Unix and Mac Systems.
- ▶ Microsoft's Internet Information Server (IIS)
 - ▶ runs only on Windows systems, while Apache
- ▶ Others – nginx, GWS



Web Server Architecture



Web System Components (contd.)



Web Servers

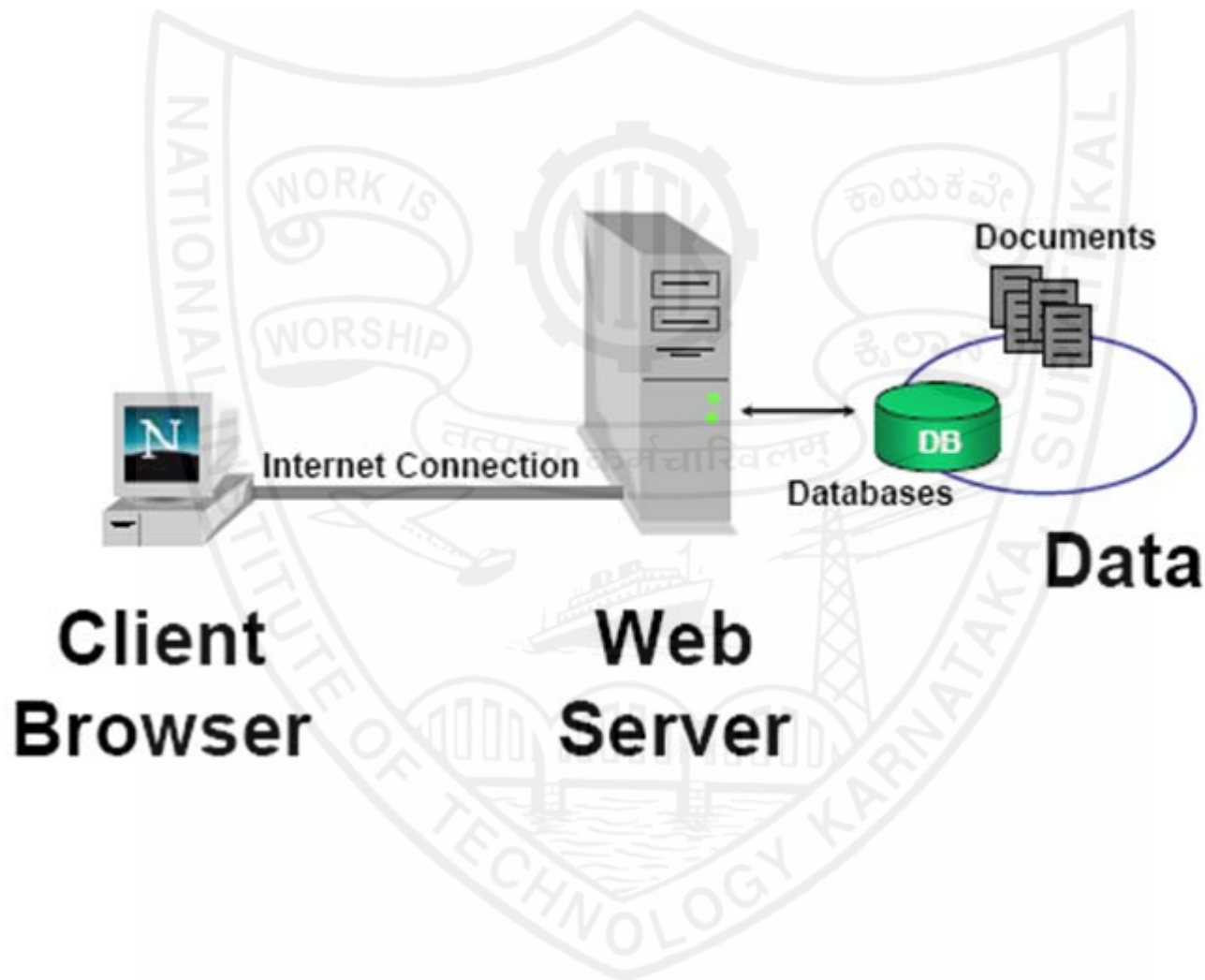
- ▶ *HTTPd* web server was the very first web server implementation.
(developed by NCSA)
 - ▶ *HTTPd* became the starting point for the free, open source Apache Server (April 1995)
- ▶ Microsoft's Internet Information Server (IIS) offers all the features of Apache.
 - ▶ IIS runs only on Windows systems, while Apache supports Windows, Unix and Mac Systems.
- ▶ Others – nginx, GWS

Application Servers



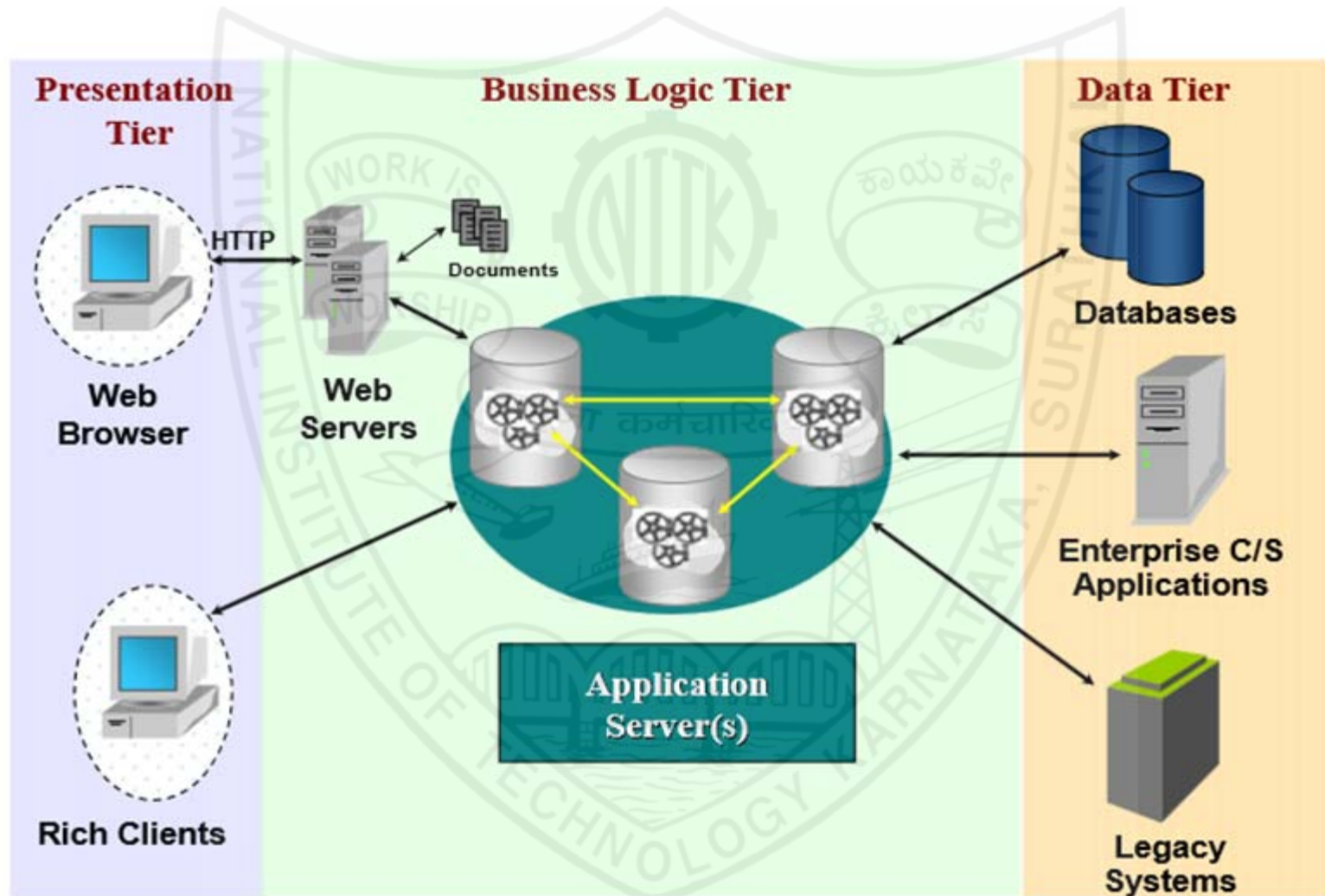
- ▶ Application Servers are middleware for Web Applications.
 - ▶ Used for connecting remote clients with applications over Internet and effectively integrating applications.
- ▶ provides middleware logic - for e.g. for transactions, security, data persistence, heterogeneous clients for complex Web Systems.
 - ▶ The goal is to provide an environment for hosting all kinds of application logic:
 - ▶ Can be used for EAI as well as Web-based integration.

Basic Web Applications



Enterprise Web Applications

Application Server(s) as Web Middleware



Application Server support for Presentation Layer



- ▶ provides extensive support for client-side interaction. A typical app server can support -
 - ▶ Web browsers
 - ▶ Applications and Devices
 - ▶ Chat clients
 - ▶ Mobile clients
 - ▶ E-mail programs
 - ▶ Web services clients
- ▶ Presentation logic support includes
 - ▶ Multi-device content delivery
 - ▶ Servlets, JSPs, XML support, etc.
 - ▶ Personalization logic



Application Server Functions

- ▶ Flexibility and Scalability
 - ▶ Provides a flexible, secure, highly scalable, and fault-tolerant infrastructure for all types of e-business activities.
- ▶ Universal Business Server
 - ▶ Provides a dynamic, Web-enabled environment - scales applications, balances loads, manages transactions.
- ▶ XML Server
 - ▶ Provides the ability to dynamically exchange/modify XML documents externally, or internally as per user request.



Application Server Functions (contd.)

- ▶ Universal Listener Framework
 - ▶ Monitors server ports to identify the presence and protocol of an incoming message.
- ▶ Application Manager
 - ▶ Agent-based management component providing real-time performance and status information.
- ▶ Security Console
 - ▶ User, group and role-based access control to every system level.



Application Server Functions (contd.)

- ▶ Fault-Tolerance
 - ▶ Customer-Facing Fault Tolerance
 - ▶ Ensuring that software/hardware system failures or upgrades don't adversely affect users.
- ▶ Fast Fail-over
 - ▶ Speed-up application recovery
- ▶ State Management
 - ▶ Storing State information (session, user activities)



Web System Components (contd.)

Backend System

- ▶ Supports the service system by fulfilling the user's request.
- ▶ In many cases, this is a Database Management System.

Internet

- ▶ The communication platform for web server and web client.
- ▶ Web client and web server are not connected directly, hence use a protocol (HTTP) to communicate with each other.



More reading...

- ▶ History of the Internet: by Gregory Gromov —
 - ▶ http://www.netvalley.com/cgi-bin/intval/net_history.pl
- ▶ Architecture of the World Wide Web (W3C)
<https://www.w3.org/TR/webarch/>