

# SERVICE ORIENTED ARCHITECTURE

## UNIT - I

# eXtensible Markup Language (XML)

- Need for XML – xchange info' in open format
- SGML (Standardized Markup Language) build on generalized markup format
- System compiled with SGML spec can communicate with each other
- Use DTD for standardized communication
- Overweight and complex

# Contd...

- HTML – developed for web by Tim Berners Lee – has its own format
- Lightweight, eliminated DTD
- Create visually appeal documents in linked manner
- Not meant for storing data / metadata or xchange data in structured manner

# Benefits of XML

- Structured format
- Data and metadata exchanged
- XML needs well-formed (follows basic XML syntax)
- Valid XML (set validity using DTD or Schemas) —design time errors fixed
- Validity checking s/w low cost or open source

# Benefits of XML

- XML editor – free with simple UI
- Use existing IP
- Supports DOM tree, so easily parsed
- Machine readable format
- Loosely coupled / future proofed

# Advantages of XML over Other technologies

- XML over SGML
- Xml allows to be parsed without DTD, but with well-formedness
- Simple syntax
- Xml spec. small
- SGML internet implementation was difficult

# Contd...

- XML over HTML
- No capability of – metadata representation, validation, extensibility, even basic needs of e-business
- Only consumption by human

# Contd...

- Xml over EDI
- Internet for data exchange
- Cost savings and efficiency enhancements provision - i.e. Validity checking, low cost parsers, XSL (EDI, domain expertise with cost)
- Internationalization and localization



# Contd...

- Xml over DB and flat files
- Text files predefined locations
- Complex files (eg. excel) structured understandable by only machine

## Contd...

- Relational DB and OO DB use proprietary binary format
- Binary formats prevent access over internet (gateway protocol and other mechanism needed)
- Custom, proprietary, expensive

# Drawbacks of XML

- Xml is huge (3 to 20 times than binary and other text format)
- Xml based protocol are heavy weight (eg. SOAP) to manipulate on fly
- In memory processing difficult for parsers

# Contd...

- Code complexity more (than comma / delimited text files)
- Complexity of Xml is not often needed (simple text file ok)
- Validation techniques not sophisticated as EDI
- Has security issues associated with internet

# Where XML Heading?

- E-business, E-commerce – standardise message exchange for business transactions
- Content Management – Legacy systems (ERP, CRM, HR) able to communicate each other
  - New open technologies for content management
- Exchange programming functionality – RPC, COM, CORBA, Web Services

# Contd...

- Used in P2P network – instant file sharing technology
- Xml spec and protocols used in instant messaging systems in open manner
- Xml enabled initiatives used in semantic web