Distributed Systems

COMP90015 2021 Semester 1
Tutorial 05

Today’s agenda

- Discussion / Q & A about Assignment 1
- Inter Process Communication (IPC) data formats
- Code Demonstration : JSON format
- Glance at a research paper

Assignment 1 Q & A

External Data Representation and Marshalling

- Data structures in programs are flattened to a sequence of bytes before transmission
- Different computers have different data representations- e.g., number of bytes for an integer, floating point representation, ASCII vs Unicode. Two ways to enable computers to interpret data in different formats:
  - Data is converted to an agreed external format before transmission and converted to the local form on receipt
  - Values transmitted in the senders format, with an indication of the format used
- **Marshalling:** Process of converting the data to the form suitable for transmission
- **Unmarshalling:** Process of disassembling the data at the receiver
- **External data representation:** Agreed standard for representing data structures and primitive data

Extensible Markup Language (XML)

A **markup** language is a textual encoding representing data and the details of the structure (or appearance)

XML is:

- a markup language defined by World Wide Web Consortium (W3C)
- tags describe the logical structure of the data
- is extensible - additional tags can be defined
- tags are generic - unlike HTML where tags give display instructions
- self describing - tags describe the data
- tags together with namespaces allow the tags to be meaningful
- since data is textual, it can be read by humans and platform independent
- since data is textual the messages are large causing longer processing and transmission times and more space to store

XML Elements and Attributes

Element

- consists of data surrounded by tags - e.g. <name>Smith</name>
- elements can be enclosed within elements - e.g. elements with the tag “name” is enclosed within the elements with tag “person”. This allows hierarchical representation.

Attributes

- a start tag may optionally contain attributes (names and values) - e.g. id="12345678"

```
<Books>
    <Book ISBN="0553212419">
        <title>Sherlock Holmes: Complete Novels...
        <author>Sir Arthur Conan Doyle</author>
    </Book>
    <Book ISBN="0743273567">
        <title>The Great Gatsby</title>
        <author>F. Scott Fitzgerald</author>
    </Book>
    <Book ISBN="0684826976">
        <title>Undaunted Courage</title>
        <author>Stephen E. Ambrose</author>
    </Book>
    <Book ISBN="0743203178">
        <title>Nothing Like It In the World</title>
        <author>Stephen E. Ambrose</author>
    </Book>
</Books>
```

JavaScript Object Notation (JSON)

JSON is a lightweight data-interchange format.

JSON is a syntax for storing and exchanging data.

JSON is an easier-to-use alternative for XML

It is based on the subset of JavaScript Programming Language

It is text based and completely language independent

IPC Data formats– JSON vs XML

- JSON is lightweight thus simple to read and write.
- JSON supports array data structure.
- JSON files are more human readable.
- JSON has no display capabilities .
- Provides scalar data types and the ability to express structured data through arrays and objects.
- Native object support. Similarities between JSON

- XML is less simple than JSON.
- XML doesn't support array data structure.
- XML files are less human readable.
- XML provides the capability to display data because it is a markup language.
- Does not provide any notion of data types. One must rely on XML Schema for adding type information.
- Objects have to be expressed by conventions, often through a mixed use of attributes and elements.

IPC Data formats– JSON vs XML

JSON serialization produces shorter strings than XML, reducing the amount of data transmission and improved performance

JSON:

```
{"employees":[
    { "firstName":"John", "lastName":"Doe" },
    { "firstName":"Anna", "lastName":"Smith" },
    { "firstName":"Peter", "lastName":"Jones" }
]}
```

XML:

```
<employees>
    <employee>
        <firstName>John</firstName> <lastName>Doe</lastName>
    </employee>
    <employee>
        <firstName>Anna</firstName> <lastName>Smith</lastName>
    </employee>
    <employee>
        <firstName>Peter</firstName> <lastName>Jones</lastName>
    </employee>
</employees>
```

Demo time!

- IPC with JSON

Glance at a research paper

- Understanding what it takes to build a good Ditsributed System
  - https://queue.acm.org/detail.cfm?id=2482856