The Exponent Group of Journals For

Information Technology

In Association with

Shree Aniruddha Upasana Foundation

Volume: 4- Number 3 Jun 2016 - Aug 2016

The Exponent Group of Journals for Electronics is an E-journal published Quarterly by **Booktionary Publishing House.**

Contents

Editors Desk, Exponent Group of Journals
Email: editor.exponentjr@gmail.com, exponent.jr@gmail.com

Editorial

- Amit Kesari

E-mail: amitsinh.kesari@gmail.com

As we all know, Information Technology is such a rapidly changing field that newer and better technologies spring up every few months and in turn render existing technologies and tools obsolete. Hence, businesses need to be very agile in adopting right IT tools, software and frameworks to ensure that they don't fade away not only due to the stiff competition, but also because of lack of regulatory compliance. In this volume of Exponent Group of Journals – Information Technology, we address the above point by presenting those articles that primarily depict the roles, responsibilities and specializations in IT to take businesses to the next level of growth.

As has been the trend so far, we start our volume through the IT – News section by highlighting few of the key events and incidents that occurred in the past quarter around the globe. In the next article, we explore how data storage tools and mechanisms play a critical role in an enterprise. In this article "Storage Area Network - SAN", we study a specific type of storage mechanism designed to serve specific system architecture. This article covers the different components of SAN, its architecture and some of the key features and benefits of SAN. The article also briefly describes Network Attached Storage which is abbreviated as NAS and the key differences between SAN and NAS.

In the next article, we analyze a key element that forms the basis of information i.e. "data". In this article, we study an emerging field in information technology world called "Data Science". In simple terms, data science is the process of drawing meaningful insights from data. As analyzed by Ben Walker, Marketing Executive of vouchercloud, nearly 2.5 quintillion bytes of data are created every single day and 90% of world's total data has been created in the last 2 years. Unfortunately, there is not enough expertise in the market to analyze this humungous volume of data and draw relevant inferences. This article "Data Science – Data – Information – Insight" introduces the emerging field of data science to the readers by unfolding its meaning; exploring its need to the enterprise world and highlighting the different steps in the data science process. The article also talks about a case study of data science, skill sets required by a data scientist and some of the courses available for learning data science.

In the next two articles, we study some of the critical roles and responsibilities in the Information Technology domain. The first article among these two articles explores two executive roles, Chief information Officer (CIO) and Chief Technology Officer (CTO). The article "Role of CIO & CTO in IT Organization" not only explains how the CIO and CTO roles have evolved over the past several decades with changing technology, but also highlights the actual case studies and use cases as implemented by several organizations. This article also provides a glimpse into the career path primarily for the readers working in an IT organization so that they can take informed decisions on how they would like their careers to shape up. The second article among these two articles explores the role of solution architect in an IT organization. The article "Role of IT Solution Architect" describes the solution architecture practice, the activities involved in solution architecture and its relationship with enterprise architecture. The article also highlights the key aspects of the solution architect role including the pros and cons, thereby helping the readers to take a conscious decision whether to pursue this career path.

Then we move along to our final article of this volume wherein we revisit some of the basics of computer networking. This article “Computer Networks – Part 2”, a continuation of the series, takes a deep dive into each of the seven layers of the OSI (Open Systems Interconnection) model. Last, but definitely not the least, we wrap up the volume with some interesting facts and figures through our “IT –Trivia” section.

We hope you enjoy reading these articles!!

I.T. in News

How Unilever is leveraging mobile to change consumer behavior

Whatscook - The first live recipe service via WhatsApp

Unilever, one of the world's largest FMCG manufacturers, is placing mobile messaging at the heart of everything it does, most importantly how it communicates with consumers, and is now mulling a global roll out of a WhatsApp-based service for its Hellmann's brand initially launched in Brazil. Speaking at this year's Mobile World Congress, Unilever VP of global media in Europe and the Americas, Sarah Mansfield presented the FMCG giant's attitudes towards employing mobile into its media mix modelling citing different activities from around the world. During the presentation she discussed an engagement programme for its Hellmann's brand in Brazil called WhatsCook where it would send users recipes via the Facebook-owned messaging service once they send pictures of their ingredients (see video).

Reference

http://www.thedrum.com/news/2016/02/25/how-unilever-leveraging-mobile-change-consumer-behaviour

The complete beginner's guide to chatbots

What are chatbots? Why are they such a big opportunity? How do they work? How can I build one? How can I meet other people interested in chatbots? These are the questions we're going to answer for you right now. Ready? Let's do this.

"~90% of our time on mobile is spent on email and messaging platforms. I would love to back teams that build stuff for places where the consumers hang out!" — Niko Bonatsos, Managing Director at General Catalyst. What is a chatbot? A chatbot is a service, powered by artificial intelligence, which you interact with via a chat interface. The service could be any number of things, ranging from functional to fun, and it could live in any major chat product (Facebook Messenger, Slack, Telegram, Text Messages, etc.).

"Many businesses already have phone trees and they do work, though most users get grumpy using them. Text based response trees are much easier and faster and that is what I expect a lot of early bot interactions to be. Sometimes with ability to chat with a live person." — Josh Elman, Partner at Greylock

Reference

https://www.techinasia.com/talk/complete-beginners-guide-chatbots

Oracle India hosts its Third Annual Global CloudWorld in Mumbai

Oracle India announced that it is witnessing unprecedented demand for its cloud solutions in the country. Oracle's Software as a Service (SaaS) business has seen the highest and fastest adoption rate till date, with Platform as a Service (Paas) close behind. This can be attributed to the company's best-of-breed SaaS applications, which are integrated with social, mobile, and analytic capabilities to help companies deliver the experiences they expect, the talent to succeed, and the performance the market demands.

Over the last few years, Oracle has acquired several hundred cloud customers in India, all using different Oracle Cloud solutions. Nearly 50 percent of its SaaS customers are net new. The Oracle Customer Experience Cloud (Oracle CX Cloud), which includes Oracle Sales Cloud, Oracle Service Cloud, Oracle Marketing Cloud, and Oracle Social Cloud, had the highest uptake in customer adoption and accounts for more than 50 percent of the company's SaaS customer base.

References

1. http://www.deccanchronicle.com/150409/business-companies/article/oracle-india-hosts-its-third-annual-global-cloudworld-mumbai
2. https://www.oracle.com/in/cloudworld/index.html

YouTube live-streams in virtual reality and adds 3D sound

YouTube is introducing live-streamed 360-degree videos on its service. The Google-owned platform said select concerts from California's Coachella festival would be the first to use the virtual reality facility.

In addition, it announced videos on its service could now be enhanced with "spatial audio", which simulates the effect of sounds coming from different directions and distances. One expert said the innovation could greatly enhance VR experiences. The announcements were timed to coincide with the National Association of Broadcasters (NAB) conference in Las Vegas, where several new virtual reality products are being launched. GoPro, Adobe and Sky Italia are among those demoing new tech at the event. YouTube made it possible to play interactive 360-degree videos on its site just over a year ago.

If viewed on a laptop, viewers can use a mouse to adjust their perspective. But if watched on a smartphone or VR headset, the experience becomes more immersive, as changes in viewpoint correspond to the movements of the device being used.

Reference

http://www.bbc.com/news/technology-36073009

Ultra-Thin E-Skin Turns Your Arm Into An LED Display

This flexible, transparent and easily attachable e-skin display could measure blood pressure and blood oxygen levels in the future.

In this age of ultra-connectedness, we are constantly surrounded by smart displays—on mobile phones, laptops, tablets and medical devices. So why not make it a part of our bodies? Researchers in Japan have developed an ultra-thin, flexible e-skin that has a organic light-emitting diode display on it—bringing us that much closer to becoming one with our gadgets. Their technology, described in Science Advances, could potentially be used for medical applications such as displays for blood oxygen concentration or pulse rate—giving whole new meaning to wearing your heart on your sleeve. The team from the University of Tokyo developed the protective film by alternating layers of inorganic and organic material—silicon oxynitrite and parylene respectively. The result is a high-quality protective film, which can be incorporated into ultra-thin, ultra-flexible, high performance wearable electronic displays and other devices.

References

1. http://www.asianscientist.com/2016/04/tech/ultra-thin-e-skin-turns-arm-led-display-tokyou/
2. http://www.livescience.com/54440-ultrathin-e-skin-turns-hand-into-electronic-display.html

Twitter rolling out big update for Periscope that includes live broadcasting from drones

Twitter has revealed that a major update is on the cards for its Periscope live video app. Some of the new features that will be added include the ability to save broadcasts longer than 24 hours and new search

options that will make it easier for users to find live streams and replays. In the past, users could only search for livestreams from people that they followed, but soon they will be able to perform searches using titles, topics, or even hashtags. In addition, new categories will be added to make it easier for people find videos of interest. These will be pre-populated and pertain to topics such as Food, Travel and Music. Video creators will be able to add the relevant topic hashtag to the title of their broadcast or select a topic after starting the broadcast. Twitter is hoping this will help people stay on top of live events around the world.

Reference

http://www.fourthsource.com/news/twitter-rolling-big-update-periscope-includes-live-broadcasting-drones-20940

Build an App in 15 Minutes With Zero Coding Using Appian Quick Apps

Appian takes PCMag inside the new Appian 16.2 and wizard-based low-code application development platform for enterprise apps.

Many application development platforms market themselves as true "low-code" or "no-code" options for business, but few can actually pull it off. Appian Quick Apps—a new capability in the latest release of the Appian business process management (BPM) platform—lets users create a fully functioning web or mobile app in a completely visual interface without typing a single line of code.

Malcolm Ross, Vice President of Product Marketing at Appian, announced Appian 16.2 at the company's Appian World user conference in Washington, DC. Ross gave PCMag a demo of the new no-code Appian Quick Apps features in advance of the event, building a working app within the Appian platform in about 15 minutes, with pre-defined objects before adding deeper customization on top in the full Appian BPM app designer.

References

1. http://in.pcmag.com/salesforcecom-professional-edition/103286/feature/build-an-app-in-15-minutes-with-zero-coding-using-appian-qui
2. http://www.appian.com/

Top 10: Best Apps For Android – May 2016

New applications come and go and every single month it seems like there are thousands of more apps that are all looking to compete for your time and attention. While many of these new apps will fall on deaf ears, there are some which could be considered a diamond in the rough or an app that absolutely can't be missed. We've put together a list of ten apps this month that are worth giving a shot.

r for Reddit: Kicking off this list is another attendee at the Reddit party. You can never have too many Reddit apps, or maybe you can, but it's nice to have a little variety so people have a choice on which they use. The app features material design and most of your standard Reddit features like creating a new post, upvoting, commenting and more.

Reference

http://www.androidheadlines.com/2016/05/top-10-best-apps-for-android-may-2016.html

Storage Area Network - SAN

- Vishal Parte
Email : vishalparte@gmail.com

Storage Overview

Information technology is all about world of information which we manage and use for our day to day work. Information is nothing but the processed Data which has a meaning, which makes sense to us. The very next question is where this information store. The simple answer is on RAM as a volatile location and may be on Hard Disk, CD, and DVD or on Pen Drive. Now days SD and Micro SD Cards are more popular among smart phone users. What are these gadgets? These are nothing but the various Storages which we use to store our data in our day to day life. This means storage is playing a vital role in IT to supports its infrastructure. Every bit of information is getting stored and archive on storages. During the course of time the IT has discovered the variety of Storages which satisfies various business needs. It increases reliability and availability of information to any of the information system.

Types of Storages

- Hard Disk Drives (HDDs): A known and widely used device in IT. There are varieties of HDDs viz. SCSI, iSCSI, SATA etc. available in market. They are chosen and user as per need.
- Pen Drive: A portable device which can carry any data or information easily. Although it has limited capacity, it is very handy and popular now days.
- Compact Disks (CDs), Digital Versatile Disks (DVDs) and Blue Rays (BRs): Again a portable device which required a player to read content on it. This again a useful storage, more popular in multimedia industry.
- Solid State Drives (SSDs): Solid State Drives or SSDs are nonvolatile storage device that stores persistent data. SSDs are nothing but integrated circuit device which works much faster than traditional HDDs. SSDs are getting more popular in desktops and server versions.
- Storage Area Network (SAN): A special type of storage which is designed to serve specific system architecture. Unlike NAS it is available and being use by the information system which is attached to it or as a part of architecture where SAN exists. NAS provides a direct access to user date like we access our files over Cloud. The information stored on SAN is accessible to users only through the servers and these servers are host with information system which is ultimately using this information. For example, the bank accounts information which is stored on SAN but accessible only through banking information system. SAN cannot be accessed independently as we use NAS.

Image 1 - SAN Front side and back side

SAN Components

SAN comprises of components like Hard Disks, RAID Controllers, Storage Controllers, Host Base Adapters and Power Banks etc. Let's see the core components in terms of HDD and Storage Controllers in details.

Hard Disk Drives

We all are familiar with this hardware component. SAN is nothing but the cluster of HDDs size as per need. There may be a variety and number of HDDs use as per usage and performance demand by an application. Every type of disk has its own specialty which can be use as per need. The greatness of SAN is all these types of HDDs can be in one single SAN based on data access is required by an application. For examples, an array of SSDs can serve as cache for frequently accessed data. An intermediate data can reside on SAS* (Serial-attached SCSI**) drives and data which is less accessible and in archive mode can be kept on SATA# drives.

Following is the brief description of the terms SCSI, SAS, and SATA.

1. SCSI (pronounced as SKUH-zee and sometimes colloquially known as "scuzzy"), the Small Computer System Interface, is a set of American National Standards Institute (ANSI) standard electronic interfaces that allow personal computers (PCs) to communicate with peripheral hardware such as disk drives, tape drives, CD-ROM drives, printers and scanners faster and more flexibly than previous parallel data transfer interfaces.

2. Serial-attached SCSI (SAS) is a method used in accessing computer peripheral devices that employs a serial (one bit at a time) means of digital data transfer over thin cables. The method is specified in the American National Standard Institute (ANSI) standard called Serial-attached SCSI (Small Computer System Interface), also known as ANSI/INCITS 376-2003. In the business enterprise, serial-attached SCSI is especially of interest for access to mass storage devices, particularly external hard drives and magnetic tape drives.

3. Serial ATA (Serial Advanced Technology Attachment or SATA) is a new standard for connecting hard drives into computer systems. As its name implies, SATA is based on serial signaling technology, unlike current IDE (Integrated Drive Electronics) hard drives that use parallel signaling.

Another feature of SAN in terms of HDDs is that SAN gives the flexibility of Scalability. Based on Storage Architecture, arrays of HDDs can be added in SAN to meet the capacity requirements. Yes, this also has dependency on number of controllers per array etc. This we will see in subsequent sections. The SANs can be scalable upto its optimum capacity.

Storage Controllers

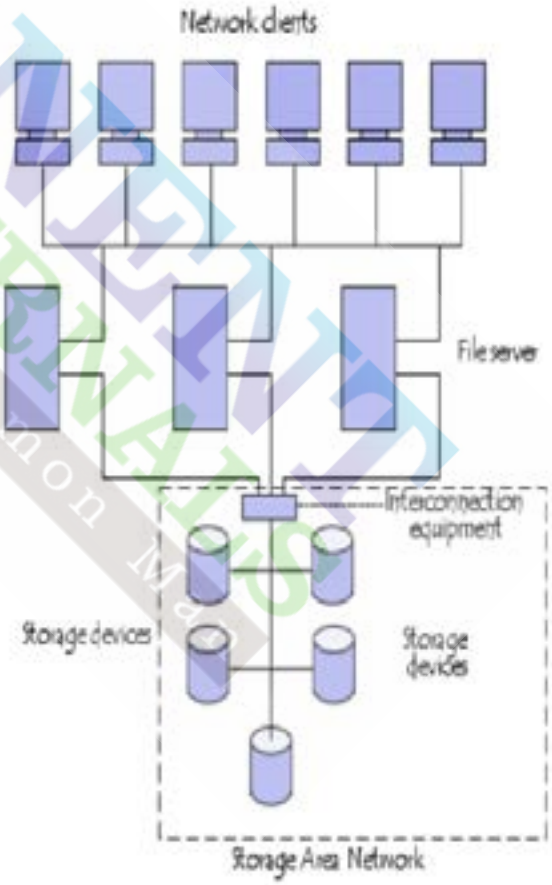


Image 2 – A schematic presentation of SAN base architecture

Image 2 - A schematic presentation of SAN base architecture

In simple terms, as name suggests, it controls the operations of SAN. There are always two controllers attached to SAN and are redundant with each other. But these controllers has a limited ports available to connect servers. This imitation can be come over by using SAN Switch between Server and SAN.

The storage controller is a kind of a server which has a processor embedded within itself. Storage Controller is responsible for performing wide range of functions for the storage system. Controller is available as a single controller or as a dual controller for redundancy.

Each Controller has I/O (Input-Output) path where it is connected through SAN Switch or may be directly to servers. It performs operations like data handling as well as functions like volume management and RAID management.

In modern day's data centers, the performance of storage system can directly be impacted by Speed and Capability of Storage Controller. The controllers processing capability are highly important. The reason for this is the availability of high speed network. It can go upto 10GBPS over UTP and FC can support scaling from 4 GBPS to 32GBPS. This helps upto 256 GBPS bidirectional throughput. This is allowing Storage to support high volume of IOP (Input-Output Performance) which improves performance and functionality.

Network Attached Storage (NAS)

Network-attached storage is a type of dedicated file storage device that provides local-area network local area network (LAN) nodes with file-based shared storage through a standard Ethernet connection.

In simple term, it is storage on a network. Best examples are the storage available over Cloud like Google Drive, One Drive and Dropbox. You simply have to have an account to have your storage space to store your information. These are nothing but NAS attached with these services over Cloud.

Features of SAN

1. Dynamic Storage Provision. Adding storage capacity to a server is a unique feature of SAN. Dynamic Storage Provisioning allows storage administrators to present extra storage to the server systems on a short notice. The SAN administrator can present the requested additional disk space and the server administrator can immediately extend the server's disk to accept the newly presented space.

2. Dynamic Snapshot. Dynamic snapshot provisioning is another advanced feature SAN. It allows administrators to back up data from one storage tier to another without leaving the SAN. This feature creates very fast and very reliable backups. SAN snapshots allow quick point-in-time backups of critical data for databases. For databases, these snapshots provide a rollback point for data. The snapshot, plus transaction logs since the previous snapshot backup, create a rapid restore point for database administrators.

3. WAN Replication. WAN Replication is an advance Disaster Recovery (DR) feature provided by SAN. This feature also supported with features like encrypted data transfer, deduplication and ability to automate the replication process. In case of Cloud base WAN replication, it can be used to replicate from Primary Site Data Center to DR Site Data Center.

4. SAN-to-disk backup. Backing up from SAN to disk is different than the dynamic snapshot concept. SAN-to-Disk (S2D) is an advanced feature that enables an administrator to create impromptu backups onto USB or eSATA disks. This feature is not intended for regular snapshots or backups. It is meant to be used only in unusual cases or as an emergency backup. The reason that you don't want to use this feature as a regular solution has to do with speed and reliability. USB and eSATA disks are not fast in comparison to RAIDed arrays; nor are they reliable. The occasional snapshot is harmless and a nice feature for your SAN system to accommodate, but don't

rely on it long term or with regularity.

5. Multi RAID Configuration. SAN must have multiple RAID configurations available along with the ability to mix them within the same environment. Different workloads require different levels of redundancy. Therefore, SAN needs flexibility to accommodate any type of workload. It's unwise to purchase multiple types of SAN technologies or levels to accommodate those different needs. Higher-end SAN solutions can provide that one-size-fits-all scenario with ease. Database systems, file servers, application servers and web servers all have different RAID requirements, but there's no need to purchase three or four different technologies to handle them and their backups.

6. Multi Disk Configuration. Similar to RAID, SAN may have array with variety of disk as per role and functionality of SAN. SAN may contain SSDs for highly use data like database files. This may followed by SAS or iSCSI or FC disk for application software. Then if infrastructure demands the backups and archive on to be on same SAN then this type of data can store on SATA drives. There is no need to populate Server Systems with range of disks. The variety of disks can be installed in a single SAN supported by Controller System.

Advantages of SAN over NAS

Following are the advantages of SAN over NAS.

- SAN is allowing a block level data access to server where NAS is allowing file level access to server.
- FC is a primary media used in SAN. Ethernet is a primary media used in NAS.
- SCSI is main I/O protocol in SAN. NAS use NFS/ CIFS as main I/O protocol.
- SAN appears to be an owned storage to server. NAS appears as a share drive or folder on server.
- SAN can have excellent speeds and performance when used with fiber channel media. NAS can sometimes worsen the performance, if the network is being used for other things as well (which normally is the case).
- SAN is used primarily for higher performance block level data storage. NAS is used for long distance small read and write operations.

Disadvantages of SAN

- SAN is an architecture based device. Hence the SAN oriented setups are expensive and maintenance cost is also higher.
- SAN based architecture needs professional management. The handling of SAN for configuration and maintenance demands professional services.
- SAN cannot be used as standalone device. Access of SAN is limited to the servers which are exposed to the storage either directly or through SAN switches.

Using SAN over NAS

- Performance is basic criteria of various IT infrastructures. SAN helps and plays crucial part in overall systems performance in terms of data processing and response to various activities and process over data.
- SAN is use for block level data access. Mainly database applications are storing data over SAN.
- The data over SAN is more secure. Unlikely file base access, the data store over SAN is accessible through application. Hence data security is handle and managed through application.
- SAN can be use a combination of drives depend on need. A single SAN box can comprises of SSDs, SAS and SATA drives together. For example, the frequently access

or persistent data can reside on SSDs. On SAS drives, application software and database information can be store. SATA drives can be used to store files, data which is less access and can use for data archive purpose. All these variety of drives can be handled by single controller.

Future of SAN is future of Storages

Storage technology develops in two dimensions: how it works and how it's used. The current decade will see major movements in both directions, although these will be evolutionary rather than revolutionary. Storage demands are doubling every two years and users still require easy access. As data management and storage become big business, the storage area network looks set to come out on top.

The major competition will be seen in Solid State Devices and Rotating Media. The development over SAN since its introduction as an architecture base system is significant. SAN became more popular as a storage system as it delivers optimum performance, return on investment to users. Most of the current IT infrastructure are SAN oriented which helps the storage management and availability compare to traditional server base storages. SAN is playing important role in any IT infrastructure in terms of data security and confidentiality. Next generation SANs are coming with built in operating environment for better management and control.

SAN also can offer Storage as a Service over Cloud based environment where a single box SAN can provide services to variety of applications over Cloud and manageability is still simple and can be scaled up to the limit. Hence SAN will continue to dominate the Storage Industry and will remain backbone of any IT infrastructure.

References:

1. Google Search Engine
2. Wikipedia
3. Forums and Blogs
4. SNIA (Storage Networking Industry Association) web site.
5. http://searchstorage.techtarget.com/

Data Science - Data, Information, Insight

- Poorva Sawant

E-mail: puranik.poorva@gmail.com

Definition of Data Science

Quite recently a new term "Data Science" is becoming popular. Now what does that mean? - Its simple it's Science about Data. Science can be defined as Knowledge or Study based on facts learned through observations and experiments. This is exactly what we do with Data while practicing Data Science. So in simple terms, Data Science can be defined as the extraction of knowledge from large volumes of data.

Data Science is an interdisciplinary field about processes and systems to extract knowledge or insights from data in various forms, either structured or unstructured, which is a continuation of some of the data analysis fields such as statistics, data mining, and predictive analytics, similar to Knowledge Discovery in Databases (KDD). The fact that we now have huge amounts of data, in itself justifies the need for a new term.

There's always a database behind every web front end, and middleware that talks to a number of other databases and data services (e.g. -credit card processing companies, banks, and so on). But merely using data isn't really what we mean by "data science." A data application acquires its value from the data itself, and creates more data as a result. It's not just an application with data; it's a data product. Data science enables the creation of such data products.

Need of Data Science - Why do we suddenly care about statistics and about data?

A famous quote by 20th-century American statistician W. Edwards Demming – "In God we trust, everyone else please bring data"- has come to characterize the new orientation, from intuition-based decision making to fact-based decision making.

Over the last decade there's been a massive explosion in both the data generated and retained by companies, as well as you and me. Sometimes we call this "big data," and like a pile of lumber we'd like to build something with it. Data scientists are the people who make sense out of all this data and figure out just what can be done with it.

Data science utilizes data preparation, statistics, predictive modelling and machine learning to investigate problems in various domains such as agriculture, marketing optimization, fraud detection, risk management, marketing analytics, public policy, etc. It emphasizes the use of general methods such as machine learning that apply without changes to multiple domains. This approach differs from traditional statistics with its emphasis on domain-specific knowledge and solutions. The development of machine learning has enhanced the growth and importance of data science.

Data scientists use their data and analytical ability to find and interpret rich data sources; manage large amounts of data despite hardware, software, and bandwidth constraints; merge data sources; ensure consistency of datasets; create visualizations to aid in understanding data; build mathematical models using the data; and present and communicate the data insights/findings. They are often expected to produce answers in days rather than months, work by exploratory analysis and rapid iteration, and to get/present results with dashboards rather than papers/reports, as statisticians normally do.

Data Science – Task Breakdown

At a high level following are the tasks a Data scientist should perform.

1. Data Cleansing
2. Data analysis
3. Modelling /statistics – Building Statistical Models Based on the analysis
4. Engineering/prototyping – Complete Solution for business users to enable and influence decision making.

The order of these tasks roughly reflects the life cycle of a data science project. Let's look at these tasks in more detail.

Data cleaning

There's a lot's of data available, but much of it is not in an easy to use format. This part of data scientist's job involves making sure that data is nicely formatted and conforms to some set of rules.

As an example, consider a sheet where each row describes the finances of a fast food franchise. There might be columns for city, state, and number of sandwiches sold in the last year. But, rather than having all this data in one document, it probably comes spread across many different files, which need to be joined together. Doing this is in some sense the easy part. The hard part is making sure the resulting combination makes sense. Typically there will be some formatting inconsistencies, and floating somewhere in the data set is a row where the number of sandwiches sold is 'India' and the state is 25,000. Data cleaning is all about finding these inconsistencies, fixing them, and making sure they'll be fixed automatically in the future.

Data analysis

This is the sort of work most people think of using Excel for. A data scientist will typically work with data sets that are too large to open in a spreadsheet program, and may even be too large to manipulate on a single computer.

Data analysis is the realm of visualization. This is where you make lots of plots of the data in an attempt to understand it. Through this process, a data scientist is trying to craft a story, explaining the data in a way that will be easy to communicate and easy to act on. Sometimes this can be something simple, like figuring out what event signals when new users convert into long-term users, or something more complex, like figuring out when someone is slowly scamming you for lots of money ala Office Space. For example, data scientists at Facebook figured out that having at least ten friends helps guarantee that a user will stay active on the site, which is why there is so much machinery on the site devoted to finding new friends.

Modelling/statistics

This is where deep theoretical knowledge creeps in to data science. Once you've got clean data and an understanding of that data, you generally want to make predictions either from that data or similar looking data that you'll get in the future. One example can be, predicting how many visitors a website gets. To do this there are models built based on what is known about traffic to individual websites as well as how people interact with the web. There's a lot to elaborate on this topic, and it's really the subject of a separate write-up. However, this step is often very complicated. We live in a golden age of machine learning, where very powerful algorithms are available as black boxes that produce good results but still will often find problem that no model is going to work well on. So a data scientist spends a lot of time evaluating and tweaking models, as well as going back to the data to bring out new features that can help make better models.

Engineering/prototyping

Having clean data and a good model is only the tip of the iceberg. Let's consider the visitor model again; even if we've got a good model for predicting how many people visit a site, it doesn't do anyone much good if we can't give those predictions and meaningful story to our customers and do it consistently. This needs building some sort of data product that can

be used by people who aren't data scientists. This data product can take many forms: a visualization (or chart), a metric on a dashboard, or an application. Whether a data scientist is building a full application or just a proof of concept often depends on the answers to following questions.

1. How much data is involved?
2. What story you want to narrate?
3. Who the final consumers are going to be?

We're still in the early days of engineering with a slant towards projects that utilize large amount of data. You work on a problem build a Tool; at times you realise the techniques don't work quite well in their new context. You go back and redo your analysis because you had a great insight while discussing with your friend, a new source of data comes in and you have to incorporate it, or your prototype gets far more use than you expected.

This is the best thing about data science: you do a lot of things and you do them together, and it's a nice challenge.

At times, it's all about the difference between explaining and predicting. Data analysis has been generally used as a way of explaining some phenomenon by extracting interesting patterns from individual data sets with well-formulated queries. Data science, on the other hand, aims to discover and extract actionable knowledge from the data, that is, knowledge that can be used to make decisions and predictions, not just to explain what's going on.

The raw materials of data science are not independent data sets, no matter how large they are, but heterogeneous, unstructured data set of all kinds – text, images, and video. The data scientist will not simply analyze the data, but will look at it from many angles, with the hope of discovering new insights.

One of the problems with conducting such an in-depth, exploratory analysis is that the multiple data sets that are typically required to do so are often found within organizational silos; be they different lines of business in a company, different companies in an industry or different institutions across society at large. Data science platforms and tools aim to address this problem by working with, linking together and analyzing data sets previously locked away in disparate silos.

There is explosion of online data, in an increasingly interconnected world. Following are some statistics. As of 2014, every minute

1. Youtube Users share 72 hours of video
2. Google receives over 4 million search queries
3. Tinder Users swipe more than four hundred and ten thousand times
4. Skype Users connect for 23, 300 hours
5. Whatsapp users share more than three hundred and forty thousand photos
6. Apple user download forty eight thousand apps
7. Facebook users share two million and four hundred and sixty thousand pieces of content

With increasing popularity, different sectors have gained tremendously by implementing data science projects and programs for their needs.

Data science is a capability that must be carefully developed over time and invested in before it begins to produce returns.

Data Science – Case Study

One example of Data science can be – "Map Based Real Estate Portal- Housing.com"

Housing.com, the Mumbai-based start-up, is one among the Indian online businesses that explored data science and machine learning algorithms. The realty portal has come up with many tools such as Traffic Flux, Heat Maps, Listing Decay, and more in their efforts to present information to users in a visually appealing way that is more interactive than traditional plain listings.

What drove housing.com to consider data science?

Once Housing.com made it a practice to collect

data for each listing, a few early engineers in the company realized that more could be done with the data than just presenting it to users in its raw form. So the Data Science Lab was formed and that set the foundation for rigorous and ambitious data science at Housing.com ever since. The data science team of the realty portal follow a set of practices to help home buyers and investors manage risk and make wise investments. The potential of machine learning algorithms in realty business. There is still tremendous potential in using machine learning to develop accurate and useful indicators of demand in the real estate market. Housing.com has number of unique users searching their inventory every day. They do a lot of data science work to track and structure that click stream behaviour so that they can develop models of consumer demand and provide decision advantage for property brokers and developers and other business customers like banks making home loans and providing other financial services around real estate.

Forecasting demand helps enable prices to rise and fall efficiently and helps home buyers and investors manage risk and make wise investments.

If you look at where data science has made the greatest impact, it's almost always on top of well-running technology products that have automated data collection as an integrated part of the system. Then machine learning is applied to the product's data streams, insights are generated and product is improved and entirely new features and products are developed.

What Technology & Tool Skills Do Data Scientists Require?

As per a recent survey of - what programming and scripting languages were required in these Data Scientists, the number one listed programming language was Python because of its inherent friendliness to data analysis & supporting libraries such as NumPy, SciPy & Pandas. The second most popular language was Java, followed by C++, Perl, and Ruby & C #.

We then performed a similar analysis on statistical tools required; the tool most required in these job postings was R, followed by SAS, Matlab, SPSS, Stata & Minitab.

But only acquiring skills on the programming languages or statistical tools doesn't make you a Data Scientist. You need to have An Insatiable Curiosity: A data scientist's job is to answer deep questions using data and to gather insights to improve the business or the product. This is facilitated by a natural curiosity about the product, when you can form your own questions about the data that you have the power to answer.

There are going to be a lot of metrics that will be wrong and a lot of metrics that will drop. A strong product intuition based on prior experience can identify what's wrong and generate good hypotheses on why things dropped. Domain knowledge will give any data scientist an advantage on diagnosing problems and figuring out where to start.

As it's rightly said "Machines do analytics, Humans do analysis". Humans have to find the patterns, ask the right questions and make the connections in the data. Computers are good at detail and examining the past, but real data science requires imagination and cognitive ability.

Additional Reading

1. The Open Source Data Science Masters - http://datasciencemasters.org/
2. Introduction to Data Science – Udacity - https://www.udacity.com/wiki/ud359
3. http://analyticsindiamag.com/what-should-one-learn-to-be-a-data-scientist/
4. Why Do We Need Data Science When We've Had Statistics for Centuries? - http://blogs.wsj.com/cio/2014/05/02/why-do-we-need-data-science-when-weve-had-statistics-for-centuries/
5. You can access the public details of innovation @TCS from this link: www.tcs.com/research
6. Data Science and Prediction – by Vasant Dhar - http://cacm.acm.org/magazines/2013/12/169933-data-science-

and-prediction/abstract#F3

Online courses

Following are some Data Science Courses available on Coursera.

1. https://www.coursera.org/course/datasci
2. https://www.coursera.org/specializations/statistics
3. https://www.coursera.org/specializations/data-analysis
4. https://www.coursera.org/learn/python-data
5. https://www.coursera.org/learn/python-databases
6. https://www.coursera.org/learn/wharton-customer-analytics
7. https://www.coursera.org/learn/practical-machine-learning
8. https://www.coursera.org/learn/exploratory-data-analysis
9. https://www.coursera.org/learn/data-science-course

Role of CIO & CTO in IT Organization

- Rucha Khale

E-mail: ruchakhale@gmail.com

Definitions of CIO & CTO

Who is Chief Information Officer?

Chief Information Officer (CIO) is an executive job title that is given to a person who heads the information technology **WITHIN** an organization. He or she plans the critical strategic, technical as well as management decisions needed for effective IT implementation. CIO analyses the benefit of a technology for an organization and works on improving existing business process. Being at the executive board level, he or she is responsible to derive a substantial demonstrable business value from IT spend. CIO has to work within the budget set by the organization and needs to report the same to the CFO (Chief Financial Officer).

Who is Chief Technology Officer?

Chief Technology Officer (CTO) is an executive job title that is given to a person who manages the Research and Development needed for developing new technologies that would give the organization a competitive edge in the industry. He or she is responsible to overview the short and long- term needs of an organization. CTO is responsible to ensure the organization reaches its goals by planning the investments in technology as per the capital allocated. Generally the CTO reports directly to the CEO (Chief Executive Officer).

Evolution of the role of CIO

Mainframe Era

The role of CIO started back in the late 1950s and 1960s when organizations started to incorporate computing into their business operations. These first-generation IT systems were used in the Electronic Data Processing department. Later, this department was called Information Systems (IS) department. Spanning from 1960s to early 1980s, this era – called Mainframe Era – saw the use of Mainframe Computers that were used by businesses in order to automate their back office processes as well as some large computing tasks. The data processing or IS managers, in this era, were not involved in IT strategy or business strategy. They were only responsible to deliver the new IT systems on time, within budget, and with a high level of reliability. Eventually, the mainframe technology was applied to the business. The role of CIO emerged from the need of aligning IT with business.

Distributed, Digital Era

By early 1980s, IT systems were not focused merely in the data processing departments that employed only IT professionals. Now, business units outside the central IS department felt the need to procure their own systems and began implementing independent systems. Systems were distributed across the entire enterprise in silos. This, not only proved to be expensive, but also was inefficient to cater to the goals of the enterprise. Also, employees too began using powerful desktops to interact with the systems implemented by their respective department. This resulted in the need for implementing an end-to-end Enterprise Resource Planning (ERP) Systems that would be networked throughout the enterprise and thus remove the autonomous pockets of IT systems. However, this became the most challenging job for the CIO, since it sometimes led to a massive failure of the implementation.

Figure 1. Evolution of role of CIO

Web-based Era

With the advent of the Internet in early 1990s, the role of CIO was further extended to step up and chalk out a business strategy for this Web-based Era. Along with the responsibilities of managing the IT Systems and Services for a company, CIO now was also responsible to visualize the commercial use of the Internet and how it would affect the enterprise. This was a role that triggered the need for leadership skills and vision in order to re-engineer business processes.

The rapid growth of technology and the tremendous adoption of digital technology by end consumers expected the businesses to work in both physical and virtual channels. CIOs now had to evaluate the market forces that affect the organization, innovations, vendor offerings, and emerging technologies. This led to rethinking of IT in every aspect of the businesses in order to sustain the business goals and ensure its survival.

Digital-Natives, Digital-Disruptors Era

Early 2000s saw the birth of digital-native companies such as Google, Facebook, Amazon, etc. Cloud-based IT, Big Data Analytics, and Social Collaboration Platforms had to be analyzed by the CIOs. Advance technologies like Artificial Intelligence and Internet of Things also raised concerns for CIOs. Digital Disruptors posed a big challenge on CIO, since the competitors were not limited only within their own industries. It was a big responsibility on the CIO to ensure that their enterprise IT systems as well as data was safeguarded. Cyber-attacks that could result in high-profile data breach would cost CIOs their jobs!

Evolution of the role of CTO

Distributed, Digital Era

The role of CTO started in 1980s when the organizations felt a need of investing in Research and Development activities in order to come up with unique products in the industry. CTOs were mainly leading the R&D activities. The focus of the R&D department was predominantly long-term. They were shielded from short-term business pressures.

Web-based Era

Eventually the Web-based era in the 1990s saw the development of global markets and technologies. CTOs

became extremely important to the companies for providing high-quality IT products and services, in order to maximize their shareholder value. The focus began to shift to short-term results. The R&D team was expected to come up with IT products and services that would fetch immediate gains for the businesses.

Digital-Natives, Digital-Disruptors Era

With the birth of digital-native companies in early 2000s, CTO's role evolved further to become the architects of future products and offerings of their companies. Their contribution was directly linked in the company's growth and profit. Monitoring R&D projects and their profitability for long-term results as well as short-term results became a challenge for the CTOs.

The focus area for CTO changed from managing their department to leading a transformation. Their role was now more strategic in order to drive innovation and business direction.

CTOs were then entrusted with chalking out critical competencies of the company. Working with the CFO, the CTO had to recommend as well as justify the funding for future investments. Competencies were investigated, and then either developed and grown, or pruned, or even acquired from other companies.

Roles & Responsibilities of CIO & CTO

The roles and responsibilities of a Chief Information Officer are as follows.

- Runs the IT organization by focusing on internal customers i.e. business units and users
- Bridges the gap between IT and Business
- Streamlines internal IT operations
- Aligns business process with technology
- Manages technology infrastructure of the organization
- Evaluates new technologies as per budget defined by CFO to optimize business processes
- Monitors IT purchases, interacts and negotiates with vendors, drives implementation of projects
- Manages internal IT projects like ERP, CRM, SCM, etc.
- Defines information strategy by describing the roles and uses of information as per desired business goals
- Enables interdepartmental information sharing
- Establishes organizations IT Security policies
- Sets up IT Governance and Compliance policies with Business Units
- Defines metrics to evaluate performance of internal IT systems.

The roles and responsibilities of a Chief Technology Officer are as follows.

- Runs the Research & Development and Engineering groups
- Holds the post of company's "Chief Navigator"
- Defines strategies to enhance product portfolio
- Aligns product development with business goals
- Defines competencies of the business
- Focuses on external customers
- Identifies risks and opportunities for the business
- Promotes innovation and product development by shaping up company culture
- Defines product acquisition strategy to increase product portfolio
- Communicates company's technology strategy with investors, management, employees, and to some extent, partners
- Collaborates with vendors who work on the solutions to enhance company's products
- Maintains technology standards of the product portfolio
- Monitors emerging technology trends that might impact the company.

Case study – How Internal IT Systems are enhanced

by CIO

How Cisco eliminated its disparate systems, multiple ERPs and non-standardized processes into one global ERP instance.

Background

- Cisco supply chain was highly diverse, extensive, and global.
- It had more than 300 product families.
- Large number of customers had extremely different expectations and fulfillment requirements.
- Acquisitions by Cisco formed a major part in Cisco's growth. These acquisitions brought along their own supply chain requirements and processes, that had to be integrated into core operations.
- Most of Cisco's products used Configure-to-Order Production Model, while some of the acquisitions used Build-to-Stock production Model.
- Cisco had different supply chain processes.

Cisco supply chain problems

- More than 1000 suppliers, manufacturing partners, etc.
- 16 Configure-to-Order manufacturing sites
- 4 Build-to-Stock manufacturing sites
- 8 strategic logistics centers
- More than 25000 orderable product IDs
- Millions of shipments annually.

Challenge faced by Cisco

- Separate Configure-to-Order and Build-to-Stock processes
- Multiple ERP instances
- Redundant and non-standardized processes
- Heavily customized Supply Chain Management System with more than 2500 customizations, and nearly half of them unused
- 250 custom applications, 19 separate databases, 30000 custom data objects
- Adding even one data field to an existing report could take efforts of at least one person month
- Data regarding parts were stored in 7 different systems.

Pain points of old IT Systems

- Very difficult to scale, improve productivity and customer experience
- Not able to adapt IT to changing supply chain business requirements
- Difficult to tap market transitions and opportunities quickly.

Simplified Large-Scale Services (LSS)

In early 2012, Chief Information Officer (CIO) Rebecca Jacoby made significant efforts to standardize the enterprise-wide processes. Large-Scale Services (LSS) had to be simplified. Large-Scale Services are the foundational services that are required by any company to operate – Supply Chain Management, Human Resource Management, and Customer Care.

The directive issued about LSS

- Use standardized processes and common practice, wherever possible
- Customize only when necessary for flexibility.

The first LSS to receive funding was Supply Chain, which consisted of SCM, Order Management, Product Lifecycle Management, Revenue Recognition, and Country Enablement.

Business-IT Alignment

Business and IT worked side-by-side to achieve a massive Business-IT alignment. They used the approach of "lead by architecture" for the planning cycles. To get maximum business value out of the Business-IT alignment, they used enterprise architecture framework **BOST** from Proact Business Transformation Inc. It helped business and IT stakeholders to have inter-linked planning models that were based on 4 architecture views: Business, Operations, Systems, and Technology (BOST).

Evaluation of Oracle R12

Oracle R12 Out-Of-The-Box (OOTB) functionalities were assessed against existing processes. OOTB are the functionalities directly provided by Oracle itself, without any customization. CIO's focus was clear – use standardized processes wherever possible, and customize only when necessary.

The team worked for 3 months to analyze the business processes at higher level, and the transformation and changes that had to be made to run the business if they used at least 80 percent of Oracle R12 OOTB functionalities.

Results

Cisco achieved its end result by using 95 percent Oracle R12 OOTB functionality. They were successful in consolidating the Supply Chain Business Process and the ERP System. Standardized end-to-end workflow was automated.

- New factory can be added in 6 months, as compared to 18 months
- Time to market reduced by 30 to 50 percent
- Order Cycle Time reduced by more than 70 percent
- Support Costs reduced by 30 to 50 percent

This successful transformation of Cisco has enabled

- Ease to expand into new markets
- Faster launch of new products and offers
- Rapid Acquisition Integration
- Business-to-Business model with the manufacturing partners
- Reduced freight costs
- Intelligent scheduling
- Better Customer and Partner Experience

Lessons Learned

The best practices for achieving a similar supply chain transformation

- Set up focus of the executives and commitment beforehand
- Ensure Business-IT alignment across all the programs at all time
- Engage stakeholders at functional level
- Communicate decisions and resolutions, and establish the support of senior management, internal project teams, partners, and vendors.

Case Study – How Product Portfolio is enhanced by CTO

A CTO ensures business growth of the organization by enhancing its technology portfolio. They are also recognized as "Chief Navigators". It is hence the responsibility of the CTO to channelize the allocated budget. They lead the innovation in R&D, and help model the innovation culture in the organization. CTOs ensure that the organization has the best technology on their portfolio, irrespective of the source. At times, it is even cheaper to license or acquire any new technology, rather than finding ways to develop it in-house.

Companies like Dow Chemical Co. and Royal Philips Electronics laid emphasis on promoting **entrepreneurship** and new business creation. In pharmaceutical industry, collaboration between internal research centers as well as collaboration with external technology partners was the prime focus. It is the CTO who has to make this transformation in the mindset of the R&D team, which is an on-going task.

Cisco Systems Inc.

The company's vision was to position itself as the preferred internet-enabling company and also as a total solution provider based on the superiority of its products. This was converted into action by chalking out a **serial technology acquisition strategy** - selecting a technology target, immediately acquiring them, and rapidly integrating them into the company's system.

P&G

P&G came up with its '**connect & develop**' strategy – P&G would post technical challenges, and vendors or other companies would send them bids to fulfill their technology needs. This was not done to eliminate all the scientists and R&D team, but to support them. From 15 percent of external innovations, this strategy helped to increase up to 45 percent of P&G's Product

Development Portfolio. This gave a significant boost to R&D productivity as well as increase in the success rate of innovation.

3M

3M Corporation projects itself as a global science company that creates inventions to improve daily life of hundreds of millions of people across the globe.

Technical managers realized that the technologies that are deployed don't belong to individual business units, but to the company. Hence a systematic sharing of technology across the entire organization – business units and labs – was created.

> "The culture in R&D in 3M is one of collaboration and sharing among our global technical community. We routinely build on each other's ideas in a boundary-less way to solve the problems of our customers and to bring relevant and breakthrough innovation to our global markets."
>
> - says Dr. Ashish Khandpur, 3M's Chief Technology Officer and Senior Vice President of Research and Development.

3M has an integrated team of scientists who are provided rich pool of practical ideas that can be nurtured. Customers too interact with the scientists at the **Innovation Centers** to discuss their pain areas. Scientists also have **Technical Council** to discuss progress on technology projects, and **Technical Forum** to present papers in an internal professional society.

The **three-tiered research structure** creates the right tension to balance present and future concerns.

1. **Business Unit Laboratories** – that focus on short-term products for specific industries or markets
2. **Sector Laboratories** – that work on products in a 3 to 10 year term
3. **Corporate Laboratories** – that work on a longer term, say 20 years.

3M encourages **experimental doodling**, where scientists can spend 15 percent of their time on any project of their choice. They are encouraged to look at any breakthrough innovations and opportunities that they find interesting.

3M's **Huge Opportunity Teams (HOT)** study customer needs very closely. They lay stress on what the customer does, than what they say.

In addition to this, 3M also encourages resourcefulness in the scientists through a rich set of systems.

1. **Seed Capital** – Inventors can seek funds from their business unit managers, as well as from other business units. They can also seek corporate funding directly.
2. **New Venture Formation** – Product inventors can form their own team by recruiting through various networking forums within 3M. The recruits too can check the inventor's record and choose to work with them or not. If the product fails, everyone continues with his or her previous jobs.
3. **Dual-career Ladder** – Scientists don't necessarily have to take up managerial roles to move up the ladder. They are awarded the same prestige, compensation, and perks as managers.

Their reward system encourages success, tolerates mistakes, and risks failures. A culture to create stories of famous failures that eventually created breakthrough products inspires innovation among scientists.

Fig 2 – 3M's equation regarding innovation returns

3M believes that innovation goals can be achieved with the right technology used with the right customer insight.

The collective efforts are put in by the organization to cultivate innovative mindset as its company culture. This is not a short-term or one-time activity, but an on-going process, which turns out to be a game-changing asset for an innovating company.

Wipro

Wipro is one of India's largest IT outsourcing companies. They realized that the success in the outsourcing industry is heavily dependent on innovation. Clients are now looking for vendors who have innovation framework and have successfully delivered projects using these frameworks. This is judged on the basis of reduced costs, improved process efficiency, and increased profitability for both Wipro and its clients.

Studying the innovation methods in companies like 3M, Nike, and Home Depot, Wipro came up with a framework and guidelines that helped viable ideas move from the idea generation stage to commercialization of the idea.

Wipro Intranet Web enabled employees to share their ideas. Innovation teams conducted workshops and set up guidelines. **Innovation Council** lead by CTO and five members evaluated the ideas for commercial viability. The five "**gateways**" of Innovation Council are depicted in Figure 3.

TechForum, an annual event, is conducted for presenting papers and demos of ideas of the employees. Funds and infrastructure is granted to nurture ideas into products. By 2003, successful products like Flow-Brix, workflow solutions for publishing, i-Desk, collaboration tool for HR management, are results of these initiatives.

In 2006, Erehwon Innovation Consulting Pvt. Ltd designed **Quantum Innovation Program**. The aim was to generate new ideas, motivate people, and drive business growth. By 2007, more than 50 projects were designed.

Applied Innovation Framework too was formulated to come up with new ideas and ways of doing things. The aim was to improve business outcomes, without any major changes. It was a complete 360-degree business approach that covered process, delivery, business, and technology innovations. It helped Wipro work closely with clients to cut costs and improve speed-to-market. They also developed successful products like eCargo suite for Airline industry, TINA platform (SOA), etc.

Wipro promoted **Innovation Evangelism** to generate awareness of innovation and the need of the hour, motivate people that anyone can innovate, and give confidence of sufficient facilities to promote innovation. Innovation Evangelism is conducted through Wipro's Innovation Camp, Innovation Bazaar, Story Book on Innovation, Systematic Creativity Workshops, and Inflection Point newsletter, etc.

In 2010, Wipro Technologies and Knowledge@ Wharton started conducting **Global Innovation Tournament** to select the most innovating managerial tools that the clients can put to use. The innovations are judged on four parameters: Novelty, Feasibility, Track Record, and Overall Potential.

Wipro has successfully imparted innovation culture to expand its boundaries for the solutions it delivers. Wipro has been successful in some of the notable innovative products like India's most powerful supercomputer, a mobile switching router for emerging markets, Base Terminal Station - cost-effective for GSM networks in underserved markets, and Wipro Energy Manager using M2M communication.

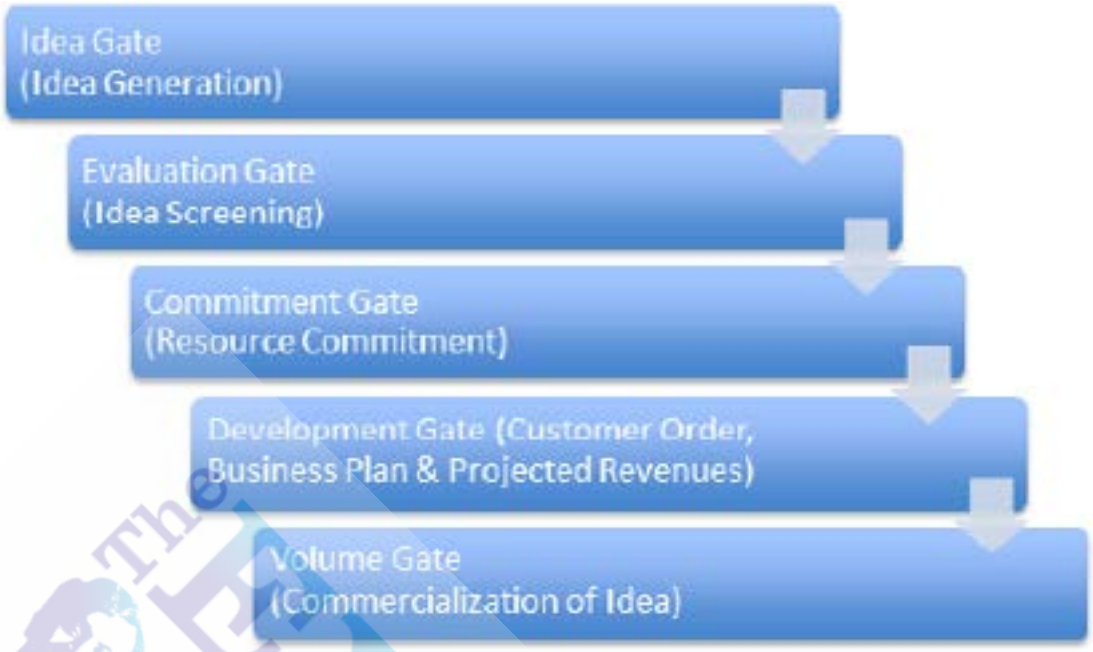
Rise of Chief Digital Officer (CDO)

We are presently in the era of "Digital-Native" companies. "Digital" does not merely mean information, technology, or data. These serve only as raw materials. Digital refers to a larger transformation. This "Digital Transformation" is focused on building a unique experience for the end customers.

The traditional businesses concentrated on internal and external efficiencies. Chief Information Officer (CIO) concentrated his or her efforts on improving internal systems and processes of the organization, while Chief Technology Officer (CTO) concentrated his or her efforts on building innovative products for their clients that created a cutting edge over the competitors in the industry. But both these roles fell short to work on the Digital Disruption brought in by the Digital-Native Companies.

Digital-Native Companies like Amazon, Google, Facebook, Uber, etc. have transformed industries. They have disrupted the traditional form of industries, rather than targeting only a few companies as competitors. Agile practices seem to be a forte of these digital-native companies, as they are able to create, design, deliver, and update their products at lightning speed.

Customers and technology are moving faster than the corporations. Technology is first introduced to the customers and then the corporation adopts it. "Digital" now refers to a superset of technology, information, data, and marketing to some extent.



To a company that is a digital-native, the Chief Executive Office (CEO) is essentially the Chief Digital Officer (CDO). In this case, every employee is working in the "digital" mode – optimizing the business digitally. Hence an explicit CDO may not add value. But for a company that is already in its digital transformation stages, the role of CDO may differ. In traditional business, a huge gap for applying customer-focused marketing skills to technology was left empty.

Chief Digital Officer (CDO)'s role intersects marketing with IT to deliver products and services to customers. CDO, a newly emerged leader, helps reshape business by using data and technology. CDO is primarily focused on making customer-focused products and services more digital.

References

1. http://www.gartner.com/it-glossary/cio-chief-information-officer/
2. http://www.webopedia.com/TERM/C/CIO.html
3. http://www.investopedia.com/terms/c/cio.asp
4. http://www.investopedia.com/terms/c/chief-technology-officer.asp
5. https://www.techopedia.com/definition/24134/chief-information-officer-cio
6. https://www.techopedia.com/definition/28874/chief-technology-officer-cto
7. http://searchcio.techtarget.com/definition/CIO
8. http://www.careerexplorer.net/executive/cto
9. http://www.fhinnovation.com/the-evolving-role-of-cto-innovation-leader/
10. http://www.techrepublic.com/blog/tech-sanity-check/sanity-check-whats-the-difference-between-cio-and-cto/
11. http://www.techrepublic.com/blog/career-management/the-cio-resume-vs-the-cto-resume/
12. http://www.brittenford.com/blog/cio-and-cto-who-gets-a-seat-at-the-table/

13. http://www.cisco.com/c/en/us/solutions/collateral/enterprise/cisco-on-cisco/cs-boit-05272014-supply-chain.html

14. http://www.3m.com/3M/en_US/company-us/about-3m

15. http://investors.3m.com/governance/corporate-officers/default.aspx

16. https://www.cebglobal.com/blogs/3ms-innovation-machine-2/

17. http://analysiscasestudy.blogspot.in/2011/11/innovation-wipro-limited-case-study.html

18. http://www.forbes.com/sites/danwoods/2014/04/14/dont-let-a-chief-digital-officer-steal-the-best-part-of-your-job

19. http://www.forbes.com/sites/danwoods/2014/08/30/why-the-rise-of-the-cdo-role-represents-a-power-grab/

Role of IT Solution Architect

- Saurabh Patwardhan

E-mail: saurabhpat@gmail.com

Definition of Solution Architect

Solution architecture is a practice of describing an architecture of a system delivered in context of a specific solution and as such it may include description of an entire system or only its specific section. Definition of a solution architecture is typically led by a **solutions architect**.

Following are few other available definitions.

1. The Open Group (2009) definition. A description of a discrete and focused business operation or activity and how IS/IT supports that operation. A Solution Architecture typically applies to a single project or project release, assisting in the translation of requirements into a solution vision, high-level business and/or IT system specifications, and a portfolio of implementation tasks.[1]

2. Definition provided by Gartner (2013) includes a hint of a relationship between the solution architecture and the enterprise architecture. They define it as follows. Solution Architecture (SA) is an architectural description of a specific solution. SAs combine guidance from different enterprise architecture viewpoints (business, information and technical), as well as from the enterprise solution architecture (ESA).

3. Greefhorst and Proper (2013) define solution architecture as follows. Architecture of a solution, where a solution is a system that offers a coherent set of functionalities to its environment. As such, it concerns those properties of a solution that are necessary and sufficient to meet its essential requirements

Understanding definitions

Almost all definitions agree that the characteristic of solution architecture is a specific solution or deliverable as opposed to an entire enterprise or group of an enterprise.

Talk of solution architecture is increasing; solution architecture is more about specific projects or initiatives than about ensuring interoperability between projects. The solutions architect leverages all the tools and design principles delivered by enterprise and SOA architects to make each solution a success within the larger framework. Solutions architects must also understand how company systems, applications and requirements work together. They must find the most appropriate industry patterns and standards for each solution. They need a background in the subject matter and a wealth of experience applicable to the project at hand. Most important, solutions architects must solve problems on deadline and on budget.

Solution architecture activities

Solution architecture is a set of key methods by which enterprise architecture delivers value to the organization. Solution architecture activities take place during solution ideation, solution design, and solution implementation.

- During ideation, solution architecture establishes the complete business context for the solution and defines the vision and requirements for the solution, it means it understands the business requirements and creates a solution to achieve and complete the requirement.
- During design, solution architecture elaborates potential options, which may include RFIs (request for information), RFPs (request for proposal) or prototype development. It selects the most optimal option and develops the roadmap for the selected solution.
- During implementation, solution architecture communicates the architecture to the stakeholders, and guides the implementation team. This implementation phase include translating the solution design into technical language and make it useful for the developers to build a system. The last phase stands at very critical in SDLC life cycle. This step stands as a pillar of entire software solutions or project.

Relationship with enterprise architecture

Relationship between enterprise architecture and solution architecture is generally well understood. According to the Federation of Enterprise Architecture Professional Organizations, solution architecture includes business architecture, information architecture, application architecture, and technology architecture operating at a tactical level and focusing on the scope and span of a selected business problem. In contrast, enterprise architecture, which also includes aforementioned four types of architecture, operates at the strategic level and its scope and span is the enterprise rather than a specific business problem. Consequently, enterprise architecture provides strategic direction and guidance to solution architecture.

Solution architects vs. enterprise architects

Solution architects in large organizations often act as the bridge between enterprise architecture and applications architecture.

An enterprise architect's deliverables are usually more abstract than a solution architect's, but not always. The main distinction between the two lies in their different goals. The enterprise architect is primarily employed to design, plan, and govern strategic and cross-organizational rationalization or optimization of an enterprise's services, processes, or components. The solution architect primarily helps programmers and project managers in the design, planning, and governance of projects of any kind.

A solution architect may have a reporting line to an enterprise architect, but the influence the enterprise architect team has on solution architects depends on an organization's policies and management structure. So, the extent to which a solution architect's work derives from an enterprise architect's road maps vary from 0 to 100 percent.

When the solution architect starts and stops depends on the funding model for the process of solution identification and delivery. For example, an enterprise may employ a solution architect on a feasibility study, or to prepare a solution vision or solution outline for an invitation to tender. A supplier may employ a solution architect at bid time, before any implementation project is costed or resourced. Both may employ a solution architect to govern an implementation project, or play a leading role within it.

Role of Solution Architect

The essence of the Solution Architect (SA) role is the conversion of the requirements into an architecture and design that will become the blueprint for the solution being created. This conversion is based largely upon the previous design patterns that the SA has been involved with in the past through reading

and staying abreast of the latest techniques, or through personal experience.

It is this conversion part of the role - the role of the SA -that most often is underestimated in its complexity. Just as the ability of the Functional Analyst to create a requirements document is one part science and wrote procedure so is the creation of the architecture. The rest, however, is an art form. Creating effective architectures to create a solution requires the careful balance of dozens of development concepts ranging from "Keep it Simple Stupid" to "Fail to Safe".

The final component to the role of solution architect is the motivation and guidance of the development leads. Development leaders need to buy into and accept the architecture, to know how the pieces will fit together at a high level

Overview of the role

An individual performing the role of a solution architect focuses on converting requirements into the architecture and design that ultimately constitute the blueprint for the solution. In that process, the solution architect usually relies on design patterns from their previous engagements, published reference architectures, and on guidance from enterprise architecture. In their efforts, solution architects balance architectural concerns of the projects with the concerns of the enterprise.

The solution architect is often the development team leader. As such, they are often expected to provide motivation and guidance to the entire development team during the systems development life cycle. The solution architect must ensure buy-in from the development team, so that the team is motivated to match the detailed design of the solution to the higher-level architecture.

Solution architects play an important role in ensuring that the solution architecture aligns with the roadmaps established by the enterprise architecture, and that it adheres to the enterprise architecture principles. Solution architects are both a consumer and contributor to enterprise architecture collateral. Often, the patterns and guidance solution architects develop becomes reusable in a broader enterprise architecture context.

Getting Started as a Solution Architect

For most people becoming the SA on a large project doesn't just happen. It's not like winning the lottery where one day your name is drawn out of the proverbial hat. It is, instead, a slow steady progression of learning and developing. A person may find their way to this coveted role within only a few years of professional experience but more frequently it takes a dozen or more years to consistently find themselves in this role.

The starting point is generally being the only person on a very small and sometimes insignificant project. The project may be small enough that a single person may fill every role - including the role of solution architect. These little projects can even be ones where the organization hasn't identified the project as something that needs to be done yet but are items that a member of the software development team realizes that would be helpful.

The distinction between a development lead and the SA are often subtle. Where the development lead focuses on detailed knowledge of a particular area the SA is very broad. This allows the SA to view the problem from a different perspective. Instead of getting mired down into the details of implementing one specific thing the SA focuses on integrating various parts of the solution into one cohesive network that solves the larger problem.

Where's the position heading anyway?

There is a great deal of pressure in the developed countries to move development to countries with cheaper labor. While the SA role could be outsourced, there is some insulation because of the need to work closely with the Functional Analysts in the gathering and organization of requirements. The distributed software of the global world requires more effort on the part of the SA and increases their need.

The overall need for SAs will continue to increase as the problems that the SMEs present are more

complex and thus they require more complex solutions. The more complex the solution, the more SAs will be required to create it.

The software development tools were supposed to reduce the effort of the SAs and therefore reduce their need for the role, however, that increase in efficiency has been far outstripped by new demand.

The Good, the Bad, and the Ugly side of solution architect

- Good: Key, High-Value Position - The SA is a key role and one which can provide immense value if done correctly. This generally means a healthy salary.
- Good: An SA is likely to get to interact with many of the key members of the development team as well as key members of the user community. This makes it a very visible position.
- Bad: Hard to keep up - Being a SA means keeping up to date on a wide variety of new techniques, patterns, and tools. The effort to keep up can be very draining at times.
- Bad: Difficult to get right - The role requires balancing so many factors that it's difficult to get right. In other words, it's easy to fail.
- Ugly: Requirements - Although a good Functional Analyst can provide great requirements a moderately skilled one may not. The difficulty is that most people, including seasoned SAs, have trouble spotting bad requirements documents before it's too late. The SA must always have to consider that the requirements may require the SA to do a lot more research and legwork into what the client really needs.
- Ugly: If a project fails, the SA is at the top of the list for people to blame.

Conclusion

The solution architect role may be the most sought after role in the software development process but it's not without its challenges. Learning the broad array of skills, shouldering the responsibility, and dealing with the consequences can be more than the average mortal may want to take on.

References

1. https://en.wikipedia.org/wiki/Solution_architecture

Computer Networks - Part 2

- Nandan Bhalwankar

Email : nandanbhalwankar123@gmail.com

Introduction

In the last part of this article we saw the basic meaning of 'Network', its types and topologies. Now we will see the most common example of term 'Network' which we use in our daily life, yes, as you would have guessed correctly, Internet. It is composed of hundreds of thousands of interconnected users. The question would have come in our minds how exactly does this Internet work? Let's have an overall view of the Internet.

Today's Internet

The internet is made up of many wide and local area networks joined by connecting devices and switching stations. Today most end users who want internet connection use the services of Internet Service Providers or ISP in short.

As we can see in Figure 1, the International ISPs connect Nations together. National ISPs are connected by complex switching stations called as NAP (Network Access Points) and they usually operate at a very high data speed (about 600Mbps). Regional ISPs are connected to many Local ISPs which in turn are connected to the actual internet users.

What is a Protocol?

In short, this 'Internet' works on the 'Connection' or 'Communication'. For this communication to occur, the entries must agree on a set of rules that governs data communication. These rules – what is communicated, how it is communicated and when it is communicated are commonly known as 'protocols'.

Let us have a look at the key elements of this 'protocol'.

1. Syntax – The structure or the format of the data, the order in which data is presented.
2. Semantics – The meaning of each section of bits, like how is particular pattern to be interpreted, what action is to be taken based on this interpretation?
3. Timing – When the data should be sent and how fast it can be sent on the public network.

OSI Model

OSI model (Open Systems Interconnection) is the layered model that dominated data communication and networking. 'Open Systems' is the set of protocols that allow any two different systems to communicate regardless of their underlying architecture.

As the name indicates, this model shows how to facilitate communication between different systems without requiring any changes to their underlying logic/software or hardware. Hence, OSI model understands and designs the network architecture that is flexible, robust and interoperable.

OSI model consists of seven separate, but related layers, each of which defines a part of the process of moving information across a network.

Passing of data and network information down through the layers of the sending device and back up through the layers of receiving device is made between each pair of adjacent layers. Each interface (the connection between each layer) defines what information and services a layer must provide for the layer above it.

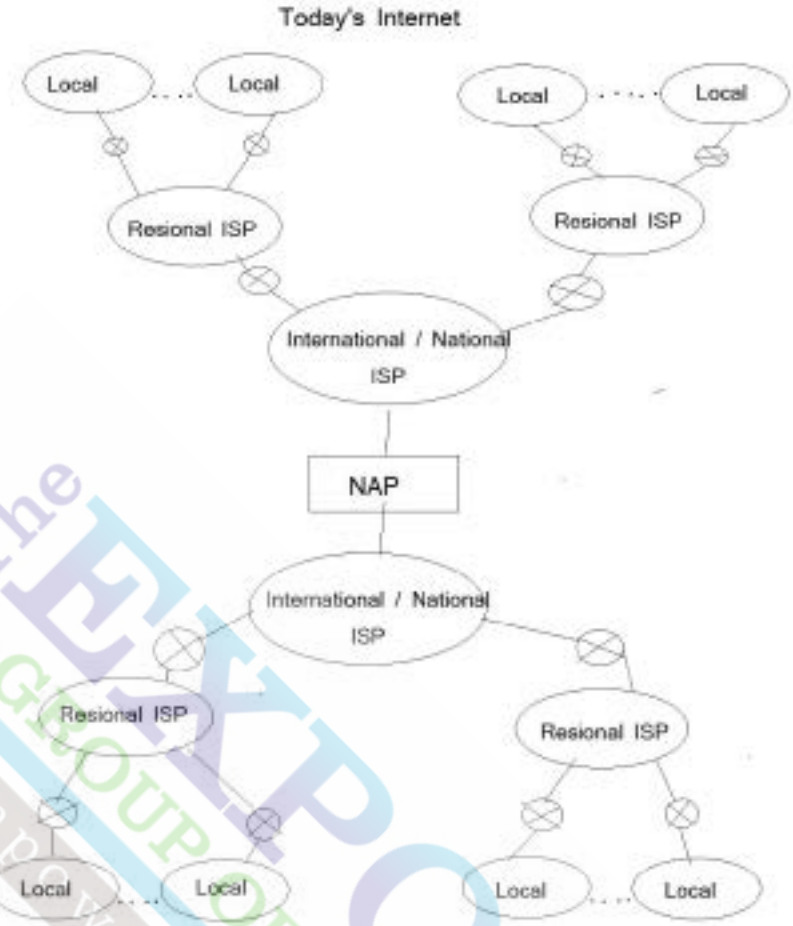
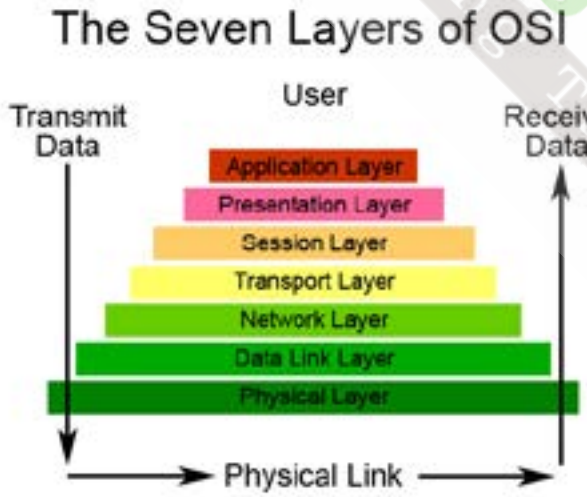


Figure 1. Today's Internet



As we can see in the diagram above, there are total 7 layers in the OSI model. The bottom 3 layers (Physical, Data Link and Network) deal with the physical aspects of moving data from one end to the other. Top 3 layers (Session, Presentation and Application) can be called as user-support layers. The layer lying between these 2 groups (Transport layer) is basically the linking between 2 groups and it ensures what below 3 layers have sent is in the form that upper layers can use.

This is just the overall view of the OSI model. Let us understand the functions of each and every layer of OSI model.

Layers in the OSI model

Physical Layer

Functions and concerns of Physical Layer are as follows.

1. Physical characteristics of Interfaces and Transmission Media. Physical layer.
    a. Coordinates the functions required to carry bit stream over a physical medium.
    b. Deals with electrical and mechanical specifications of interface and transmission media.
    c. Defines transmission media. Transmission media, as we know is the actual wire through which data / information is passed.

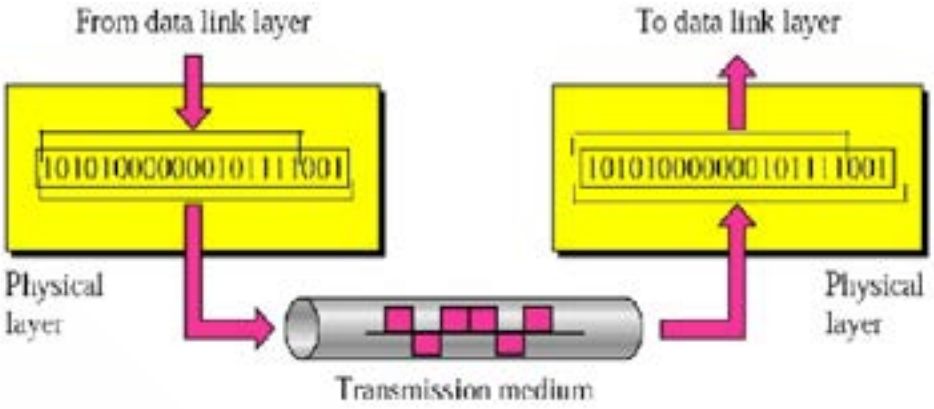


Figure 2. Physical Layer

2. Line Configuration. There are 2 types of Line Configuration.
    a. Point to Point: Two devices are connected together through a dedicated link.
    b. Multipoint: The link is shared between several devices.
3. Transmission Mode. There are 3 types of Transmission Mode.
    a. Simplex: Only one device can send and the other can only receive, or we can say it's a one-way communication.
    b. Half-Duplex: Two devices can send and receive, but not at the same time.
    c. Full Duplex: Two devices can send and receive at the same time.
4. Physical Network Topology. In first part of this article, we have seen the meaning of term – 'Topology' (the way in which devices are connected in the network).
5. Representation, synchronization of bits and their data rate. In Physical layer, bits are represented in 0s and 1s. Sender and receiver of the message must be in synchronization in order to send a message. In spite of this, physical layer also defines the transmission rate or data rate, which is the number of bits, sent per second.

Data Link Layer

Following are the functions and concerns of Data Link Layer.

1. Framing. The layer above Data Link Layer is the Network Layer, which sends stream of bits to the Data Link Layer. Data Link Layer then divides the stream of bits received into the manageable frames / data units.
2. Physical Addressing. Every node has some physical address so that the information / data can be sent / received. This layer adds header to the frame to mention the sender or the receiver's physical address. If the communication is for the outside network, then the header will contain the immediate receivers address and then the chain continues. Physical address can be called as MAC address or Hardware address
3. Flow Control. It may happen that the speed by which the data is sent and received is different. This will cause trouble for receiver to accept data at high speed than its current speed. By understanding this problem, Data Link Layer imposes flow control mechanism for the flow of information.
4. Error Control. It may happen in a network that frames get damaged or duplicated. Data Link Layer adds trailer to the frame to achieve this error control.
5. Access Control. If the same link / line is shared between two or more nodes / peers, and if two or more peers want to transfer frames at the same time, Data Link layer applies the access control mechanism to decide whom the access of the link is to be given.

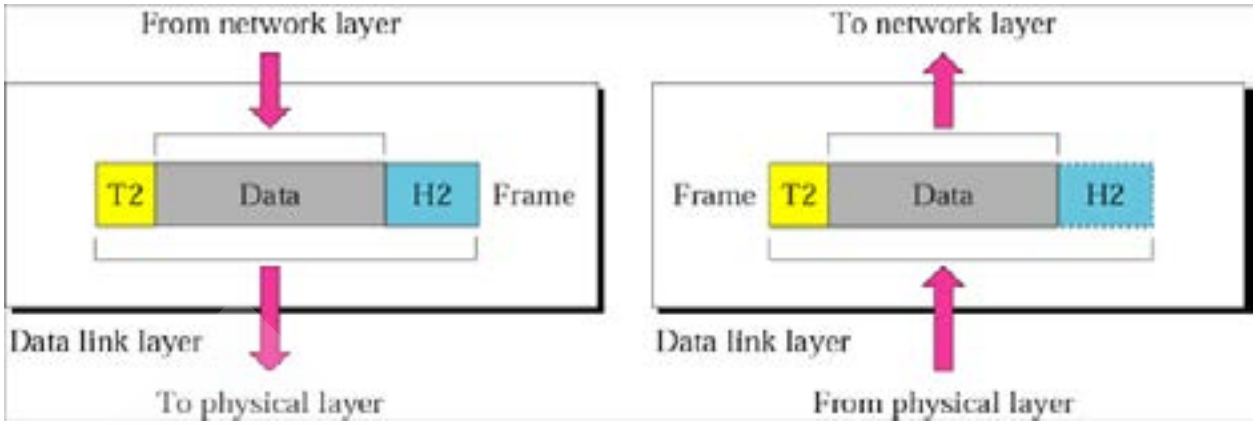


Figure 3. Data Link Layer

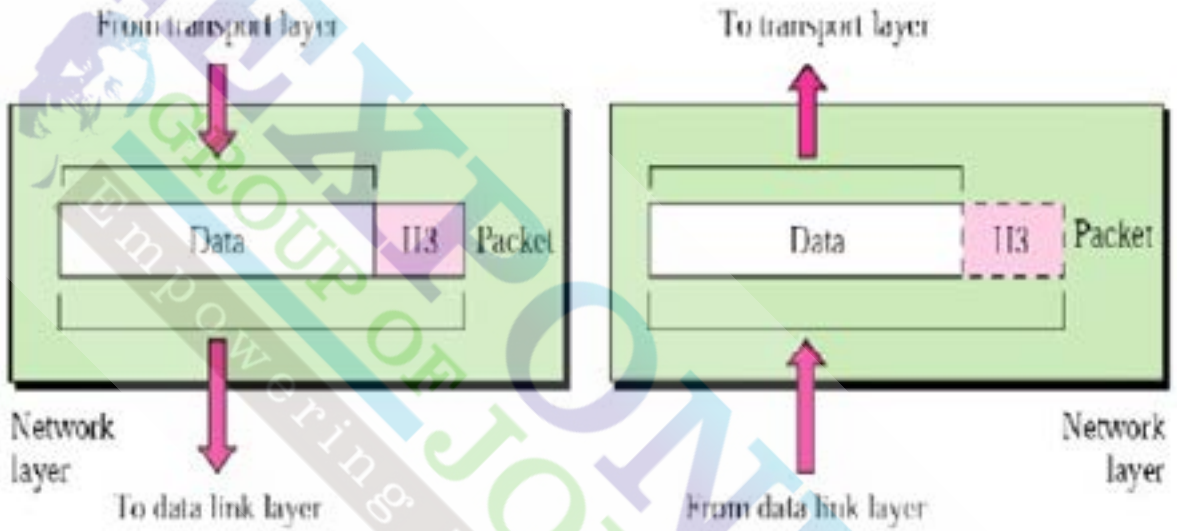


Figure 4. Network Layer

Network Layer

Following are the functions and concerns of Network Layer.

1. Source to destination delivery with the use of logical addressing. As we know, the concept of frames is used in Data Link Layer; Network Layer treats the bunch of information as a data unit known as 'packet'. On internet, each node or peer has some logical address / IP address by which it is identified. Network layer mentions this IP address in the header of the information attached by achieving end to end delivery.

2. Routing. Every small network while connecting to other networks needs some interconnecting device, known as Routers / Switches / Gateways in order to pass information. This passing or forwarding of information is known as 'routing'. Network layer ensures this proper routing by end to end delivery.

Transport Layer

Following are the functions and concerns of Transport Layer.

1. Process to process delivery. Computers run several programs / process at the same time. We have seen as the network layer transfers the packet of information from one host to the other using its IP address, transport layer on the above that forwards the incoming packet to the correct process. This correct process on that computer is identified by specific port number present on that computer. For

storing this port number, transport layer adds a header to the information.

2. Segmentation of the packet and its reassembly. While sending, the message from sender gets divided into smaller units called as segments. Each segment is uniquely identified by a sequence number. It may happen that in a network, due to a number of factors, many segments of message arrive at destination in haphazard way. Some segments may also get lost. Due to the presence of sequence number in a segment, transport layer at destination machine can identify the loss of any segment and verify the correctness of a complete message.

3. Flow and Error Control. We have already seen the meaning of flow control and error control in case of Data Link Layer. Transport Layer too works in the similar way. However, the fundamental of Transport Layer is process to process delivery. The transport layer at the sending side ensures that entire message arrives at the destination without damage, loss or duplication by retransmitting the faulty segments.

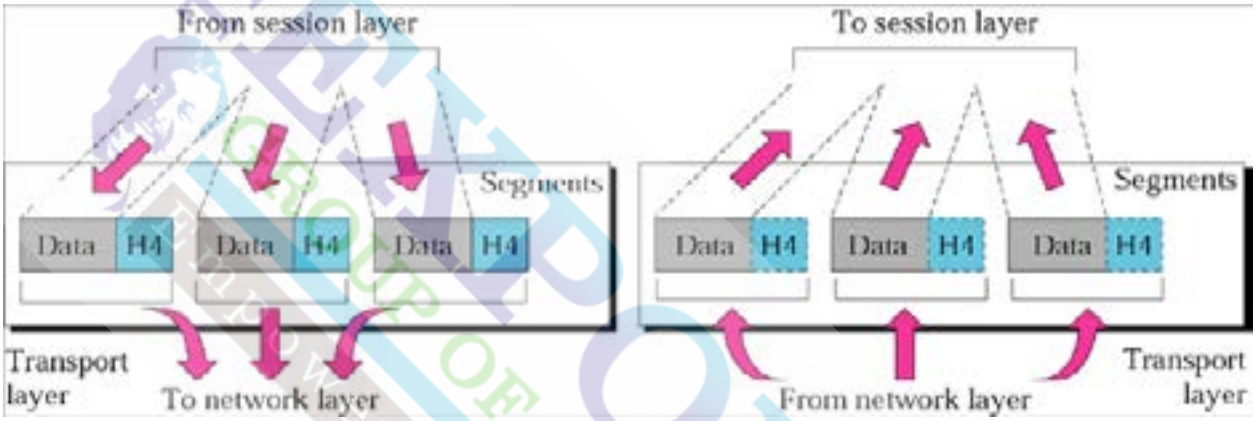


Figure 5. Transport Layer

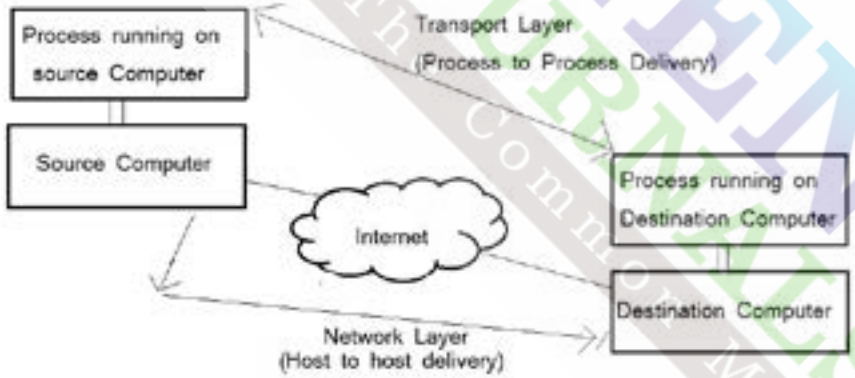


Figure 6. Process to Process /delivery using ports

Session Layer

Following are the functions and concerns of Session Layer.

1. Dialogue Control. The main objective of the session layer is to make the sending and receiving ends enter into a dialogue. With this dialogue control mechanism, the two ends can communicate with each other in 2 modes – half duplex and full duplex.

   In Half-Duplex communication, as we can figure out from the diagram above, at a time a sender can send information to the receiver. However with Full-Duplex mode, the sender can both transmit and

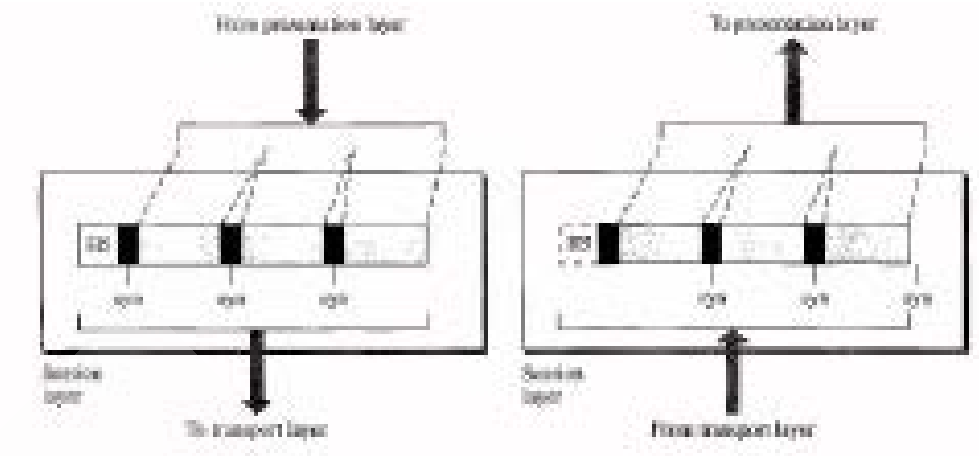
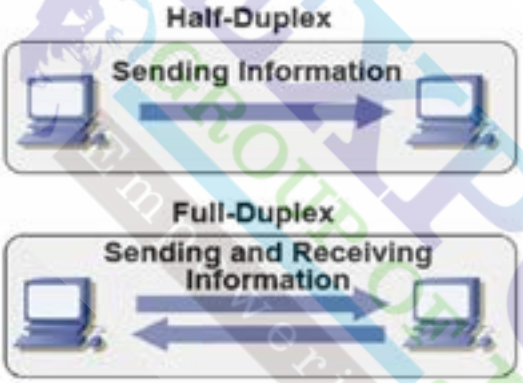


Figure 7. Session Layer

receive as well, both at a time. The best example of Half-Duplex mode is a text message that we send through mobile carriers and a phone call can be the example of Full-Duplex mode of communication.



2. Synchronization. Session layer adds checkpoints into a stream of data. Every checkpoint in the sending message is acknowledged so that data loss gets avoided. For an example, if a sender sends a pdf file of 1000 pages long. The session layer in this case adds checkpoints after every 50 pages to ensure that the set of 50 pages is received at the opposite end. In case of any network disturbance, a retransmission is needed. But because of these checkpoints, network does not need to send the complete file again. Rather only the pages from the missing page set onwards will be sent. The rest will automatically be recovered by the session layer.

Presentation Layer

The presentation layer is concerned with the syntax and semantics of the communication system.

Following are the functions and concerns of Presentation Layer.

1. Conversion and Translation. As we have seen in the first part of this article, information that is transferred from sender need not be in the format the receiver uses. Here, Presentation Layer comes into picture. This Layer makes sure that the information is delivered in such a form that the receiving system will understand and use it.

2. Encryption. Data that is passed over large network needs to be secured well in advance as there are many possibilities of network threats. Presentation Layer adds encryption to the message for this security. Encryption is the process of transforming original data into another form for passing over a network. Presentation Layer at sender uses encryption mechanism for passing information over a network and the receiver at the opposite end uses decryption is the process that transforms encrypted message back to its original form.

3. Compression. The data that is passed over network is not only textual, but it may contain images and

videos. Images and videos take a lot of time for travelling over network as they have large number of bits. Presentation Layer reduces the number of bits contained in the information.

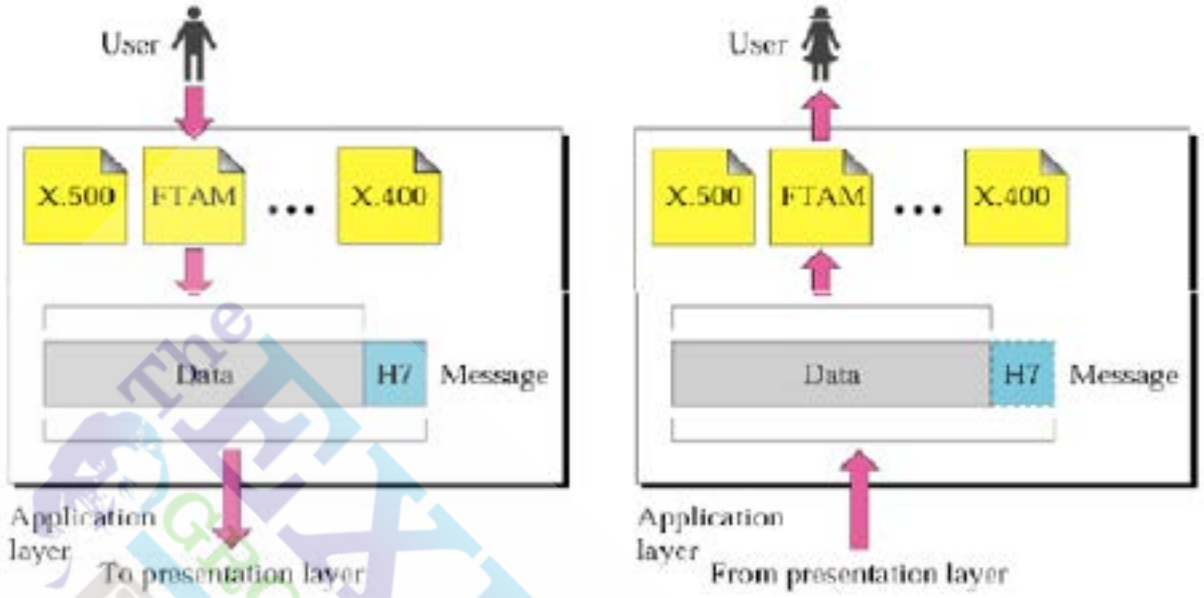


Figure 9. Application Layer

Application Layer

This is the layer where user interacts with the network. This Layer provides user interfaces and other services like email, database access and remote file access etc.

This was about the layers in OSI model.

Till now we have seen the basics of network, how these networks are connected and the uses of such connections. This part contains information of functions of every layer in the OSI model. In the subsequent parts of this article, we will see Transmission Media, Network Control Devices and Protocols.

References

1. Computer Networks by J. S. Katre – Reference Books
2. TCP / IP Protocol suit by Behrouz A. Forouzan
3. http://ecomputernotes.com/

IT Trivia

- Shyamkant Matondkar

Email : shyamkant@gmail.com

1. MPEG-4 file format is based on…
   A. MP3 format
   B. QuickTime format
   C. Wav format
   D. None of the Above

2. IBM developed a computer that defeated Gary Kasparov in 1997 is called as..
   A. Code Red
   B. Blue Sea
   C. Deep Red
   D. Deep Blue

3. Who invented Mouse?
   A. Douglas Engelbart
   B. Ada Lovelace
   C. Jeff Bezos
   D. Steve Jobs

4. Which computer system used mouse first?
   A. Microsoft
   B. DOS
   C. Xerox Alto computer
   D. Apple

5. Metric established by the IEC in 1998 to define 1024 bytes?
   A. Thousandbyte
   B. Kilobyte
   C. Megabyte
   D. Gigabyte

IT Trivia – Answers

1) B
2) D
3) A
4) C
5) B