Cloud Computing

Prof. Soumya Kanti Ghosh
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 33
Fog Computing – I

Hello. So, we will be continuing our discussion on cloud computing. So, what we have seen that in case of cloud computing, what we are trying to do we are trying to offload our computing and computing processes and data to the cloud right. So, that it is maintained by a third party and the on the other end the customer or the consumer or the user more concentrate on the business processes process.

So, that is the basic objective of or model of the things right. There are a lot of technical technicalities at the backend, but nevertheless we are offloading the thing. So, what for that what we need a very strong backbone right or a strong backbone network which should be always up and able to transfer data on a large volume. As we see as along with the development and being most of the things are digitally enabled we are what we are getting a huge volume of data.

In other sense a huge volume of data maybe need to be transmitted from this customer end or consumer end to this cloud service provider being executed the results in some cases are transmitted back or transmitted in other places right. The major issue is this huge volume or transfer of data and what we have what we see in recent development with number of activities specially internet of things, coming up and more only and huge variety of sensors in place.

So, we have lot of multimedia data which need to be transmitted right. And that is one part of the story that we require a huge backbone and type of things as for as the cloud is concerned what we consider it is a huge computing power much higher than what we what the devices can do and it is some sort of infinite computing power is there.

On the other hand what we see that a huge volume of data are being generated and being transmitted over the network. Again what we see that the devices starting from as we discussed about mobile devices, smart mobile devices or other type of devices even

intermediate network devices, what they are becoming is more powerful in terms of computation and more resourceful things ok.

Or in other sense we are not all the times exploiting the resources available in the things sic say consider a particular sense sensor node or a local sync node of a sensor which are collecting the data and transmitting to the in the upward path maybe to the cloud. So, this this could have been done some processing at the end like I can say that suppose in this particular room or a particular lab, I have say 10 temperature sensors right.

So, what is my basically business model? This temperature should be varying may between say 18 to 22 degree centigrade that is the operating range of this temperature.

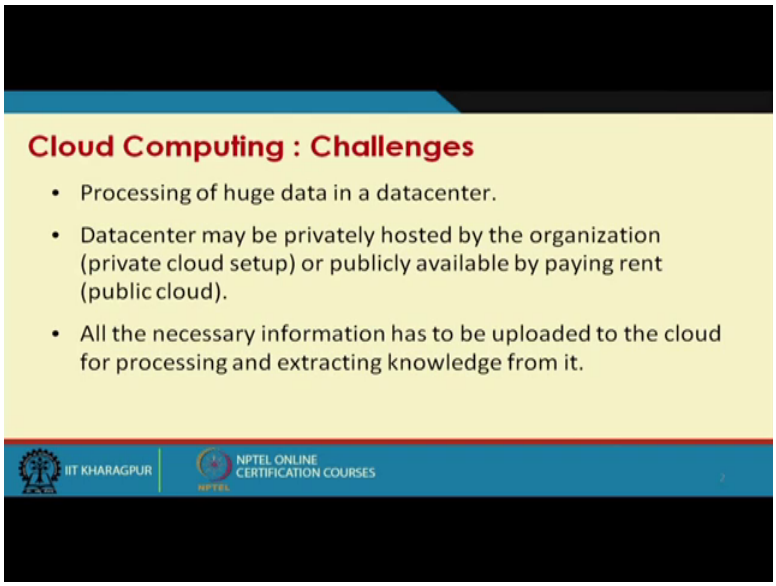
Now, what we are doing this all these 10 or sensors, are sending the data to the up in the server maybe in a cloud that is calculating, whether the within the limit and type of things and I can have the say 10 such labs. So, there are 100 such data are going on and if the temperature is varying somewhere it is sending error.

Now, if you consider this particular a enclosed lab or single things otherwise I have could taken, I taken a local ds and whether my temperature is in the lab by the sync node of the sensors which are collecting this data of this particular room and it takes a call that whether higher or upper, and then sense that say it is some statistical data or what we say some aggregated data to the sensor, it may be the average or it may be average with other standard deviation etcetera to the things.

In other sense it is not in other than sensing this transmitting this 10 sensors data I am sending one average data or and which has my purpose. Even you can say that the if the if my sync node is intelligent enough, it can take a call that whether the temperature is within this operating range yes or outside thing some 01 or yes no type of things and then transmit this.

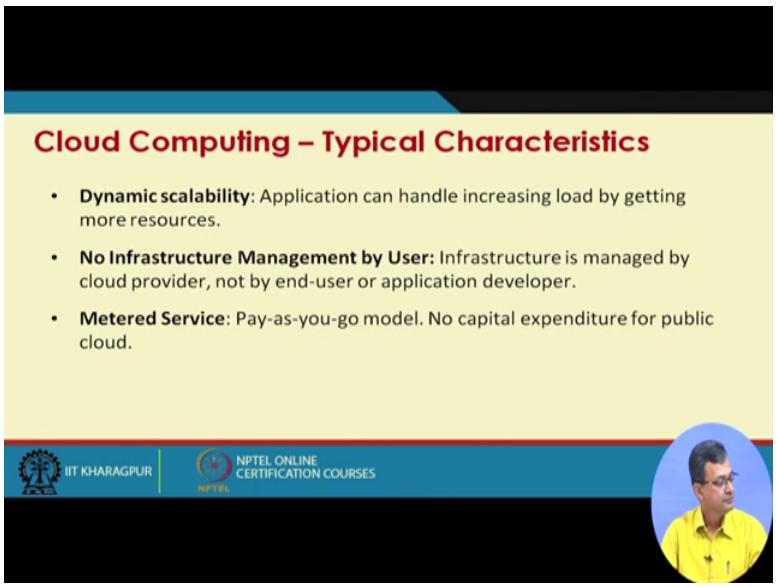
In other sense this is taking a some part of computing at the things. So, with the intermediate devices becoming more intelligent, whether there is a possibility of pushing the computing from logically centralized cloud to somewhere more down the line right towards the edge of the things right, that is exactly what we are trying to discuss today is what we say this sort of computing is fog or from cloud to fog right. So, cloud is the whole thing and fog computing.

(Refer Slide Time: 05:46)



So, as we see the challenges or the data what the cloud computing todays is doing the processing of huge data in the datacenters, may be privately hosted or publicly available by paying rent that is it can be a public cloud or a private cloud, all necessary information has to be uploaded or transmitted to the cloud for processing and extracting knowledge of it right.

(Refer Slide Time: 06:20)



So, the whole data as we are discussing need to be transmitted to the cloud.

Now, also we have seen the typical characteristics of cloud for which we are the today's world is inclined towards is that dynamic scalability, I can scale up or scale down based on my need. So, another is that no infrastructure management or practically minimal infrastructure management at the user end. So, if I offload everything that computing etcetera on the cloud. So, I require very less infrastructure management at my user end and secondly, and finally, what we have a metered service right pay as you go model.

So, these three things that dynamic scalability minimal management or all my infrastructure management pushing it to the cloud and metered service pay as you go model these are primary features of the cloud which makes is popular there are several other things which are which are there, but never the less these are the three things which are the driving force. So, whatever we do we do not want to lose out of the things if we compromise on those type of features then the very motivation to going towards cloud may be challenged.

(Refer Slide Time: 07:38)

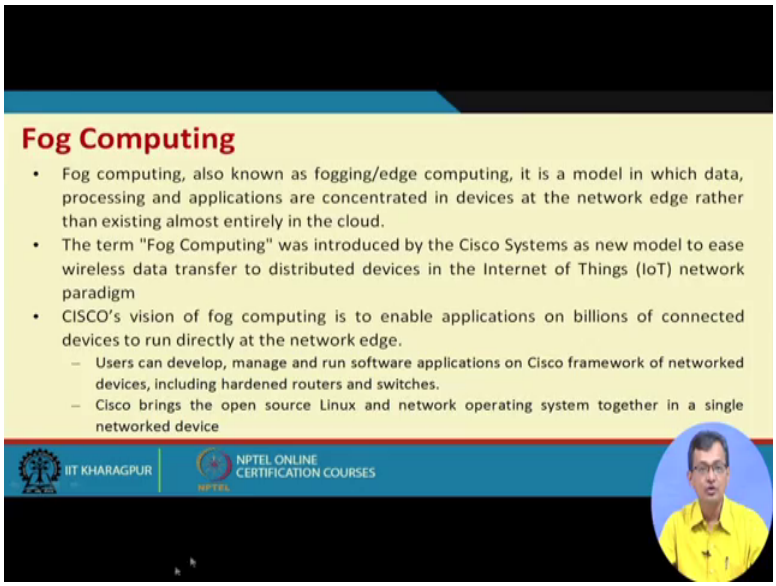
Now, there are issues with cloud only computing. So, what we say that only the cloud is computing register sitting duck or maybe the issues, especially in today's applications which variety of sensors, variety of real time operations and lot of redundant data right there are lot of data which are redundant like if I am sending temperature things it may not mean may be meaningless to sense all the sensors data, which are more or less same information unless there is a different in some sensor data I may not want to send the

data all are reporting between around 20 degree centigrade, it does not require a cloud to take a call it could have been done as a much lower level. So, that or in other sense I have a huge amount of digital data to be transmitted.

So, communication takes place takes a long time due to hum human smart for interaction and type of things, if the state still datacenters are centralized right datacenters and in some woodson. So, all the data from different region can cause congestion in the core right. So, being transmitted things especially in case of exigencies where a lot of volumes of data suddenly pushed into the thing right in case of say some disaster or some huge amount of in flux due to some event, this is a lot of volume of data suddenly in flux. So, there is a huge volume of data to be transmitted and there can be congestion and such a task requires very low response time to prevent further crashes (Refer Time: 09:15).

So, if I have this sort of things which has a some sort of accident, some accident prevention mechanisms can into should be activated. So, where we require a very low response time, immediately need to be act acted. So, waiting for that cloud to take a call revert back and all those things may take lot of time. So, that is another problem.

(Refer Slide Time: 09:40)



So, the emergence of a concept called fog computing. So, on the cloud we are to the we are talking about fog that is little bit bringing tao down to the ground or in other sense we

are pushing this computing thing from the centralized datacenter or the cloud datacenters to this edges right or intermediate or the edges of the network aj edge of the network.

So, fog computing also known as fogging and edge computing though some people have little other views of that edge computing, but nevertheless it is a fogging or edge computing it is a model in which data process applications are concentrated in devices at the network edge rather than existing almost entirely on the cloud. So, now, not only the cloud at the centralized things the data application and processes are distributed between the edge right which somes effect of some way of distributing this whole processing whether things. What it helps us? It helps us in reducing the data load in the communication, I can have a local decision and which is not needed for the global type of things they say smart traffic light management system.

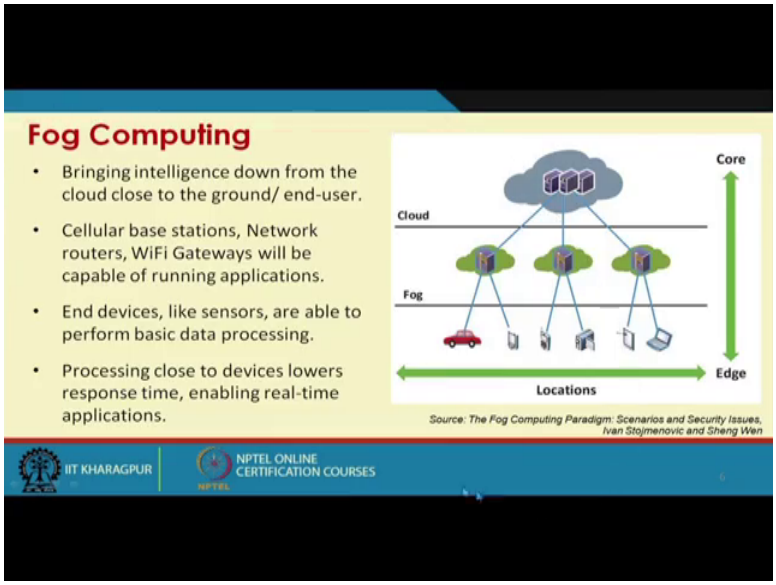
The traffic light management system in Kolkata is nothing to do with the traffic light management, this system in Delhi apparently right for day to day traffic management right. So, I could have done it locally or even I can say that a region of a particular city may have only aggregated data which need to be transmitted at the higher level for traffic management right. So, that basic intermediate management could be done locally.

So, those things could be done in a concept of what we say fogging or fog computing. The term fog computing was first introduced by CISCO as a new model to ease wireless data transfer to distributed devices in the internet of things network paradigm. So, as IoT is becoming omnipresent or IoT is becoming a everywhere it is there internet of things. So, it is huge volume of data devices which mass computing capability or resources much higher resources can do a bit of a job which could have been solved at a lower level.

So, CSICO’S vision if we look at that fog computing is to enable application, on billions of connected devices to run directly on the network edge; since CISCO is primarily a network driven organization. So, it has a huge number of devices across the world and those devices are somewhat managed etcetera managed by a some sort of a homogeneity is therefore, because upon the one make and there are resourceful devices which could have done some sort of computing things, and I can even run applications on the devices and doing so on and so forth right.

So, user can develop manage run software application of CISCO framework of network devices, including harden hardened routers switches etcetera. CISCO brings say open source Linux and network operating system together in single network devices. So, it helped to do not only computing, but it was if you want to do computing you need to give some sort of a platform to run the applications for the computing things right. So, those things are they are in the devices and this is possible because of there are resources available at different layer of the network towards the edge.

(Refer Slide Time: 13:39)



So, if we look at a view. So, this cloud are at the top it is still there and it should be there, are intermediate devices which we are now helping only a so far was only transmitting the data, now can they do some sort of a computing what we say fogs fog computing and there are end user devices which are spread over different locations, starting form say smart vehicles or which can communicate, devices servers smart cameras and anything which can do write any device which can capture detailed data compute and transmit right.

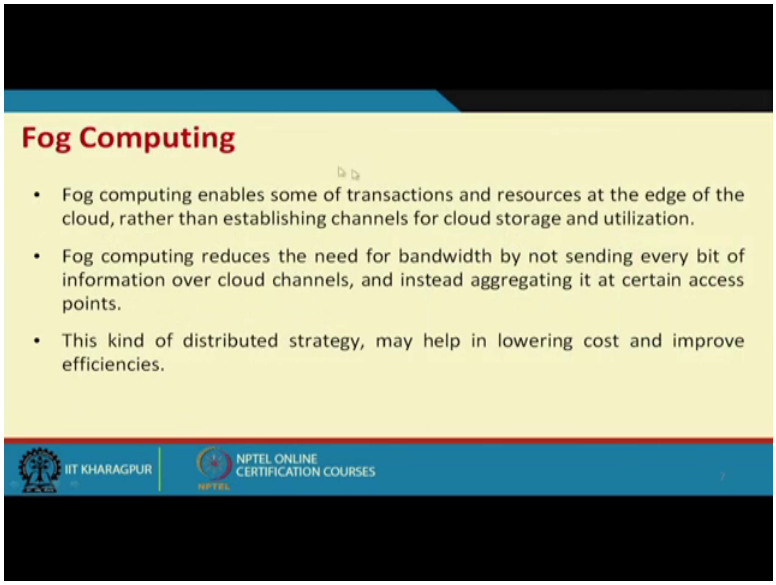
So, bringing intelligence down from cloud closer to the end user of the edge of the network that is one of the things. Cellular base station network routers Wi-Fi gateways will be capable of running these applications right. So, there are because whenever I communication we have cellular networks, Wi-Fi router into place and if those are having surplus resources and they are able to do that. So, that a my application can run

say I want to run a application for monitoring the environment of different labs, starting from temperature to humidity may be some sort of a what sort of air pollution or air content etcetera.

So, this sort of things can be done; end devices like sensors are able to perform basic data processing right. So, the sensors can do a basic data processing. Processing close to the devices lowers the response time enabling real time applications right. So, whenever we process close to the devices. So, the response time reduces that is obvious and I can do lot of real time processing of the things right. So, I can do a real time processing of a of a say of any applications like I do a application based on that what we say dynamic signalling mechanism of a traffic light based on the traffic on the road.

So, the cameras which are on the road capturing that how many what is the traffic flow based on that I the traffic (Refer Time: 16:05) signalling may change if that is the need of this traffic management. So, that is local right local to a particular portion, local to a region, local to a city right. So, that definition of locality may vary from application to application, but what we require that your devices like that traffic light device etcetera should be able to run this application which can take a call right. So, those are things nevertheless this is this is about the fog.

(Refer Slide Time: 16:33)



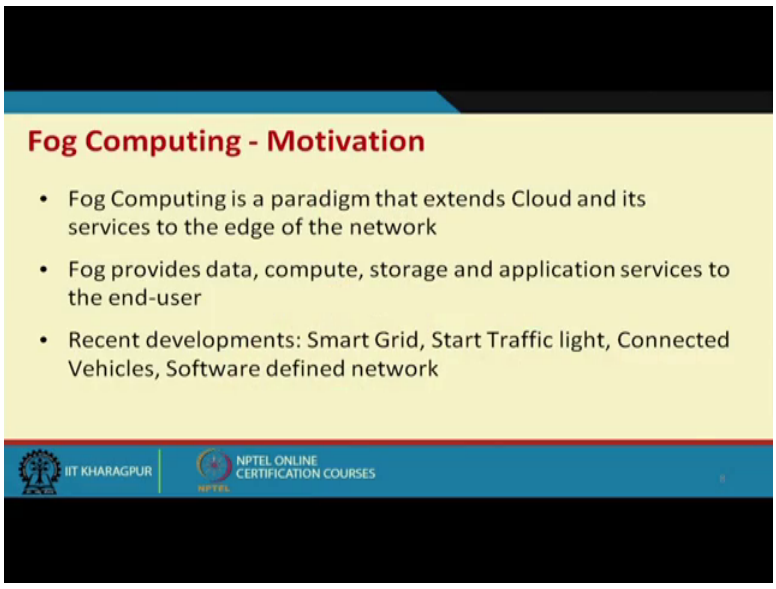
So, if we look at fog computing enable some transactions and resources at the edge of the cloud, rather than establishing channels for the cloud storage and you utilization. So,

rather than just transmitting, it do some sort of a transaction processing or application running on the things; fog computing reserves reduces the need of bandwidth by not sending every bit of information to the cloud channels over the cloud channel, instead aggregating at a certain axis point.

So, it aggregates and send the aggregate data this kind of distributed strategy may help in lowering cost and improve efficiency. So, this sort of it is a distributed strategy and this type of distributed phenomena may help us in lowering the overall cost not only in terms of monetary, if the cost of transmission in terms of time etcetera and I can we can do efficiency right I can run several applications which can be real time and type of things.

So, this motivation is obvious already whatever we have discussed the motivation, the fog computing a paradigm that extends cloud and it is services to the edge of the network. Fog provides data compute storage, application services to the end user if you see it says some sort of a small form of a instance of the cloud for that local type of things right. So, it is a doing some sort of a computing or giving some sort of a cloud service at that time at that portion of that region recent. And we and there is another side of the things because we have several series of developments one is the smart grid other is the smart traffic lighting in cities, specially cities connected vehicles or strong regular networks which is coming up and also the software defined network.

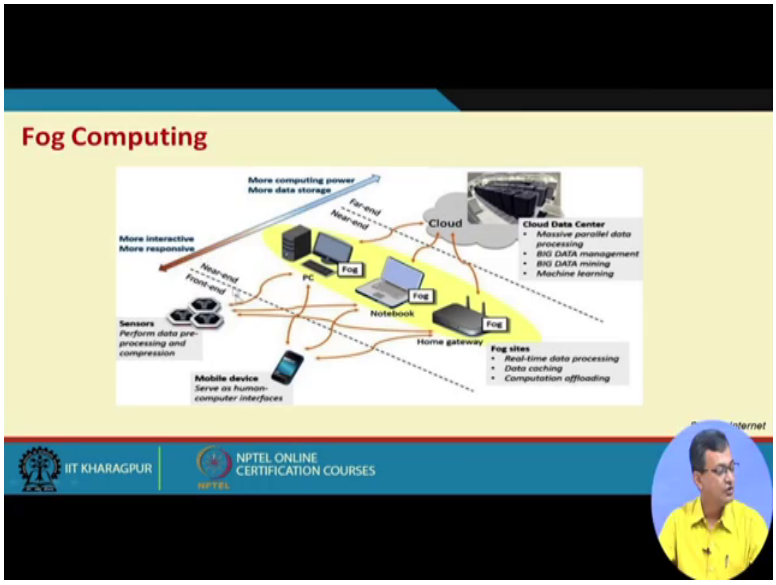
(Refer Slide Time: 18:29)



So, these are the different aspects which are itself is a topic to work at, but the smart grid smart traffic lighting, smart vehicles, as software defined network and so on and so forth they are becoming pretty popular, and in turn they generate huge volume of data right everyone is generating huge volume of data, which are being transmitted at the higher up in the layer for doing that.

So, all this different aspects has motivated or what a what it has pushed the push the processing towards doing it at the edges or intermediate layer rather than pushing everything to the cloud. So, this is what we look at the fog.

(Refer Slide Time: 19:32)



So, this is the same thing what we discussed so, we have one in this cloud. So, which has a datacenter with huge capability massive parallel data processing, big demand, big data mining machine learning algorithms etcetera, which is they are and should be their; intermediate layer which is more near to this edge or the devices. So, they are can act as a fog. So, they are they can be there can be fog sides with real time data processing data caching computation of offloading and those type of things.

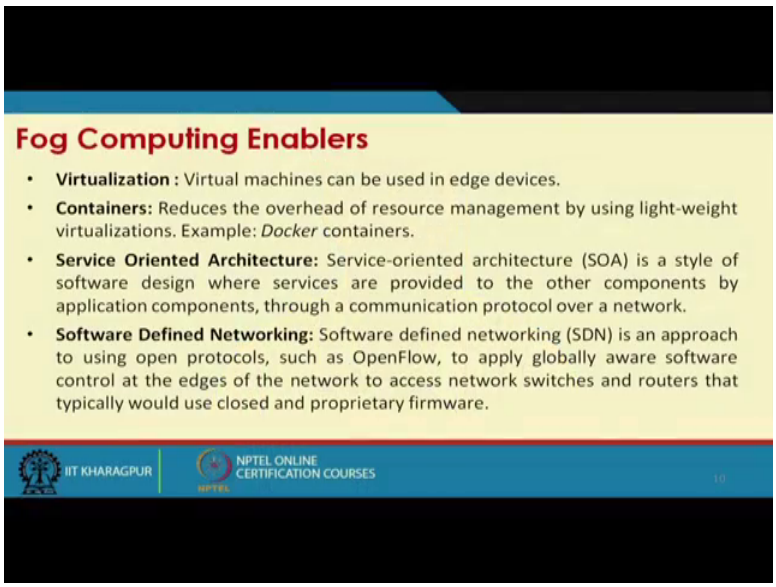
So, these are not. So, powerful at that, but as such they are intermediate devices we are which are used for transmitting data. So, that these are this can be used at the end or the at the front end or the edge or the last mind what we say, what we have the sensors we are connecting different type of data, perform data pre-processing and compression,

mobile device serve as a human computer interfaces like this these are the different type of things which are transmitting out here and in turn transmitting to the things.

So, these some sort of communication yes if we see that both way arrow, can be taken a call at this end itself right without transmitting the whole data at the things. It may be some sort of aggregated reporting and type of things or aggregating the data and taking putting it to the cloud for running some intelligent algorithm and machine learning based algorithm and type of things.

So, we have more interactive and more responsive end, to more computing power and more storage end at the other end.

(Refer Slide Time: 21:34)



So, instead of just putting a channel to transmit everything to the cloud and compute and come back we are doing some intermittent the provisioning of intermediate processing for to serve the application based. So, this is definitely a major motivation.

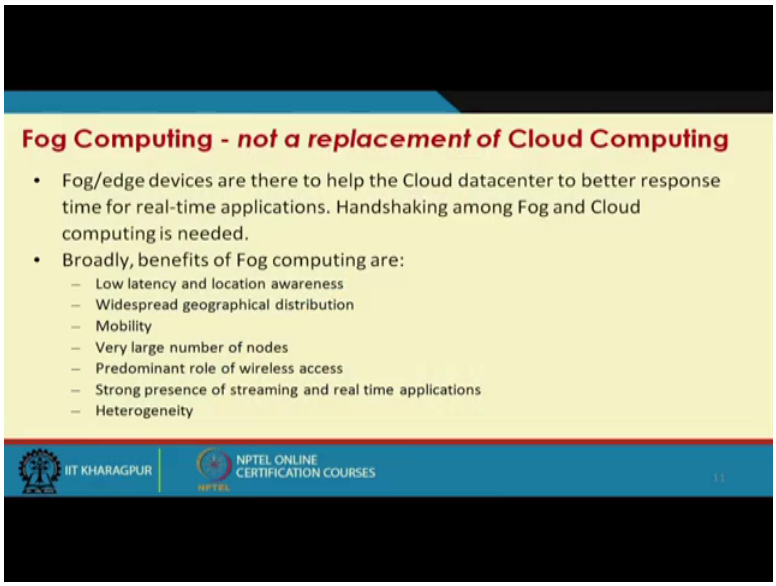
And what we try to look at that has as we have seen that the typical properties of cloud that having here infinite scalability, theoretically or quote unquote infinite scalability or off loading or having infrastructure no need of maintaining infrastructure at the client end or meter services those need to be preserved or respected right in case of a fog and those are definitely are still there as what we are discussed.

So, there are fog computing there are several enablers as those are true for our cloud computing also, one is that virtualization. So, virtual machines can be used as the edge devices right. So, there are there can be virtual machines containers or containers services are reduces the overhead of resource management by using lightweight virtualization or what we say container based application or services is. One of the popular container is Docker container right. So, it the idea is it docks into that particular things and run on the thing. So, you do not have to that dependencies is carries along with the thing right. So, it is a again a separate topic it possible we will discuss sometime, but that is this docking or container services are becoming very popular. So, that is another enabling technology out here.

Such this oriented architecture as we have discussed which is a enabling technology for cloud also is here also that SOA is a style of software design where services are provided to the other components by application component a component through a communication protocol over a thing. So, you have a service oriented architecture which three major component of service provider, service comma consumer and service registry so that heterogeneous loosely coupled parties can talk to each other right.

So, in SOA architecture is one of the driving enabling technology, and also what we are looking seeing at is the software defined network right SDN. So, SDN is an approach of using open protocols like for example, open flow to apply globally aware software control at the edges of the network to access network switches routers that are typically would use closed and proprietary from one. So, this is another technology which as becoming pretty popular or already popular in software defined network, and which is enabling technology for fog also.

(Refer Slide Time: 24:45)



So, with this several enabling technology fog is a becoming a reality, and being deployed and used in several cases.

So, in looking at we should not see that fog as a replacement of cloud, it is not a replacement of count not neither a competitor in that sense right. It is basically offloading some of these workload from the cloud to this edge devices because the resources are available because there are applications which are real time and needs more quick responses and overall process may be cost effective and efficient.

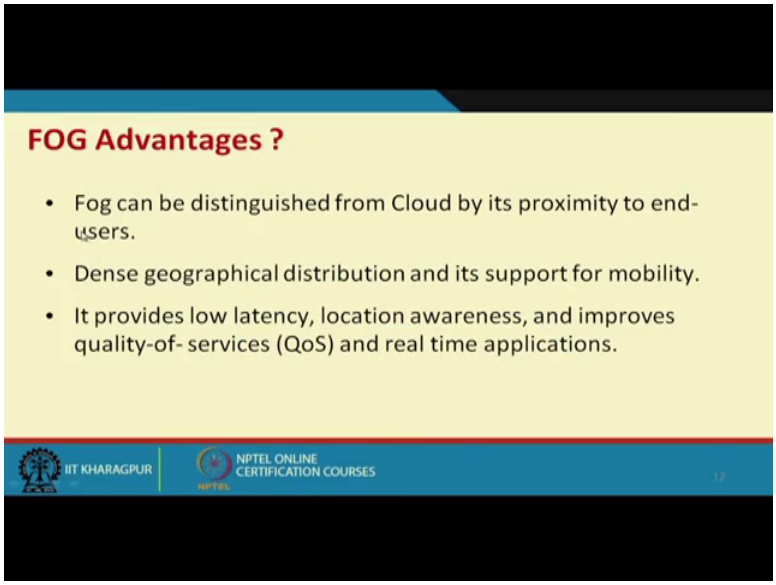
So, fog edge devices are there to help cloud datacenters, to better response time for real time applications right handshaking amount fog and cloud is needed; appropriate handshaking or synchronization between these fog and cloud is very much needed. Broadly benefits of flog computing can be that low latency and location awareness. So, it is aware that which location is operating, widespread geographical distribution especially with the sensors etcetera. So, it has thing mobility there is another important things like nowadays devices are we have lot of mobile devices right. So, the distance from the cloud or the intermediate devices which a div say end device passing through the intermediate devices will change based on the mobility of the things.

Now, this require a resynchronization reestablishment of the path; had it been in locally somewhere it may is the computing and response time. So, low latency and location awareness widespread these mobility, very large number of nodes with as we are

discussing with sensors and things predominant role of wireless access right huge volume of wireless accesses strong presence of streaming and real time applications. So, these days we are having a huge streaming and real time applications, and which requires quick response time huge volume of data and type of things need to be processed quickly and may not require all data to be transmitted right.

So, this huge volume of data can be locally processed and aggregated data can be transmitted so that the overall response time improves in a considerable way, strong presence of streaming and real time and heterogeneity, different sort of devices different type of the mix and inner and the heterogeneous. So, I can have it some sort of a fog sort of intermediate framework which basically talked to some devices which may be different from other devices. Like I have a group of sensors, I have a their sink node which talks to the sensor also do this aggregation which may be different in another other set of sensors which has a different sink note, but nevertheless when they do this aggregated data that is in a more standardized format. So, I can handle heterogeneous devices.

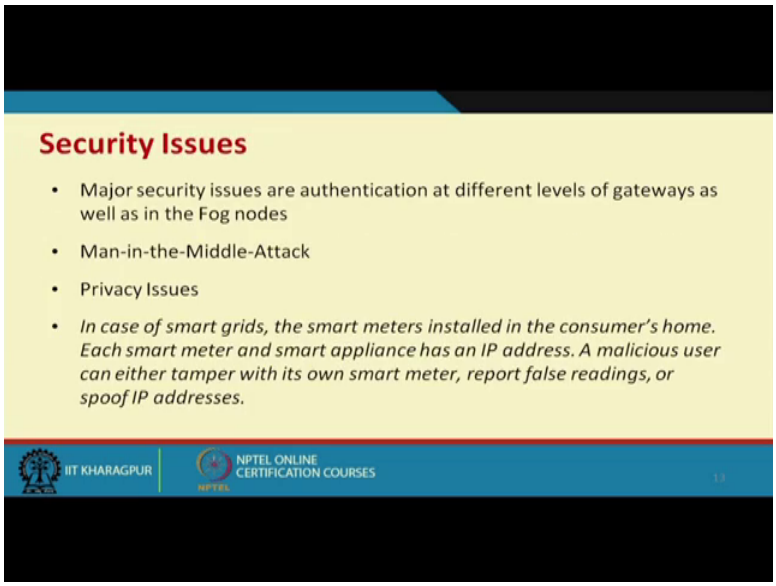
(Refer Slide Time: 27:53)



So, advantage is already we have already discussed. So, can be distinguished from cloud by proximity to the end user that is one of the advantage or over this free service cloud, dense geographical distribution and it is support for mobility right. So, we can have instead of scintillate I can have a lot of distribution. It provides low latency low

awareness and improves quality of service and real time applications right. So, there is a chance of better performing the things rather we try to look at it is not isolated fog, but fog plot cloud a as a whole can give a better service to these consumers right in terms of cost, in terms of scalability, in terms of your efficiency and type of thing specially applications where we have high quality of services and real time services streaming videos and type of things.

(Refer Slide Time: 28:52)



There are of course, some issues related to security. So, one is that as devices are dispersed right. So, maintainability of the security protocols at different fog devices is a serious challenge right. So, it is at different location now had it been cloud you have a provider at a particular centralized things, you can put lot of security mechanism in the place, but if you once you distribute over the form then you have to maintain so many things on the on different edge devices.

So, it is not only the data processing etcetera. So, there can be security issues. So, an man in the middle attack type of things can happen. So, as devices are disparts differ a as this computing data are being there in the in different edge devices. So, there is a issue of things of man in the middle attack can be there are issues of privacy issues as a as same that it is being processed at different edges and whether the data leakage is there.

Then whether you know about the things like if I consider smart grid or connected vehicles. So, if you do the intermediate ports processing whether you are basically

tracking the vehicle or looking at the processing of the consumption of a individual house or home, and type of utilization those can be there. Like in case of a smart grid smart meter installed at the consumer home, each smart meter and smart appliance as an IP address a malicious users can either tamper with it is own smart meter, report false reading or spoof IP addresses and so on and so forth.

So, whatever it comes with typical network security related issues may also come may also be problem out here. So, may be a challenge. So, there are definitely security issues there are security issues in the cloud, but this extend that to much more things as you have different devices are activated.

So, what we sees today that this fog is just not extension of the cloud, it is a necessity based on the different application, huge volume of data and the devices intermediate devices becoming more resourceful right and they are able to capable to do this type of calculations computation and. Secondly, in case in order to in doing so, I may not be doing all these high profile computation, but I can we can basically do some sort of a aggregation of the informations and sending only the aggregated informations, which lowers the basic bandwidth requirement intermediate bandwidth requirement also lowers the data load at the cloud end. So, what we see it is a technology which is a need of the hour and especially with IOTs and other things coming in a big way.

So, with this we will stop today.

Thank you.