Cloud Computing

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

Lecture – 37

Green Cloud

Hello, we will continue our discussion on Cloud Computing, different aspects of cloud computing. So, what we are discussing during this phase of this lecture is more looking at the different other aspects of cloud computing which make this computing more efficient, what are the different supportive technologies around it and how the overall cloud computing may be made more practical.

So, one of the aspects of this cloud is its energy consumption. We have seen in case of when we discussed about resource management that resource is in the cloud though theoretically that is infinite volume of resources, but the resources to be managed appropriately. So, that the service can be provided to the maximize maximum or to maximize its benefit right. So, it is a type of situation it what we looked at that where it is a win-win situation for both consumer and the provider and how this resource can be managed. One of the aspect of a resource management what we have discussed is to reduce the energy consumption right. What we see in that this overall major the dark side so called dark side of this cloud computing is huge volume of energy consume or huge amount of energy consumption.

If you have experience in any type of going through or if you have red any type of data centers or if you go through this internet different resources you will find that typically if you look at a data center the cost of the computing equipments or computing infrastructure is somewhere match with the other environmental infrastructure like the space environment ac power and so and so forth. So, it is something like x is spent on these is more or less x is also spent on that side right. Rather in some cases is more costly because of the cost of the particular space and power and type of things.

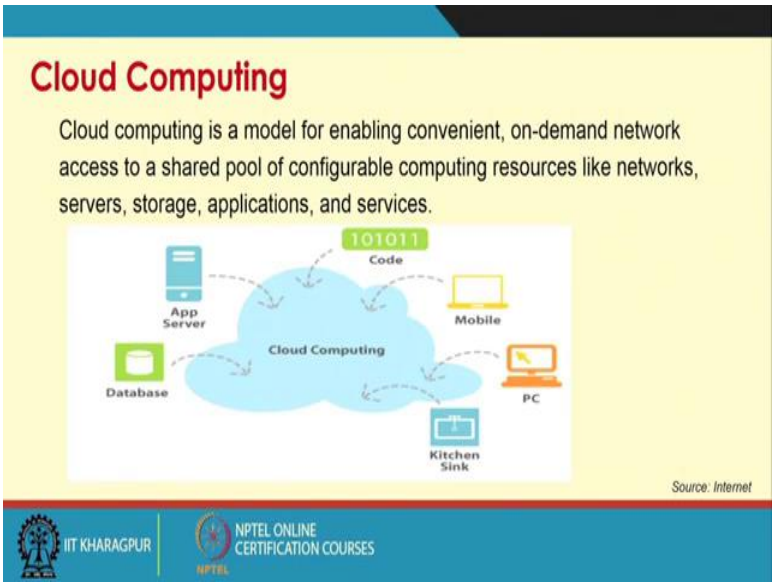
Secondly the amount of energy spent in maintaining this type of infrastructure is enormous right and what is been seen that unless this a overall consumption things consumption pattern can be reduced or the volume of consumption can be reduced in

some way or other it may not be feasible to scale up after extent right after a particular limit. So, though someone can have infinite or someone can have money to spend, but there may not be the supply is not available to the things.

So, in other and another aspect what we see that it is becoming sometimes it is a more hazardous or more negative effect on our environments specially the carbon footprint more you consume more energy being produced more may more is the carbon foot print. So, going to another unconventional energy sources may be other things etcetera. Nevertheless the overall these computing worlds need to look for some sort of a more quote unquote green computing. So, what we will discuss today briefly is that that different aspect of green cloud computing though it has a direct connection or direct relations with our resource management, but we will try to look at the different aspects of this green cloud computing and if we whenever we are setting up any infrastructure so this should keep we should keep in the mind not only the cost of the infrastructure, but also these consumption of energy how to reduce that consumption of energy and so and so forth type of things right.

So, that is what we discuss today is the green cloud.

(Refer Slide Time: 05:03)



If we quickly revisit our definition, so what it says that it is a cloud computing is a model for enabling convenient on demand network access to a shared pool of configurable resources like networks servers storage applications and services like this is the first line.

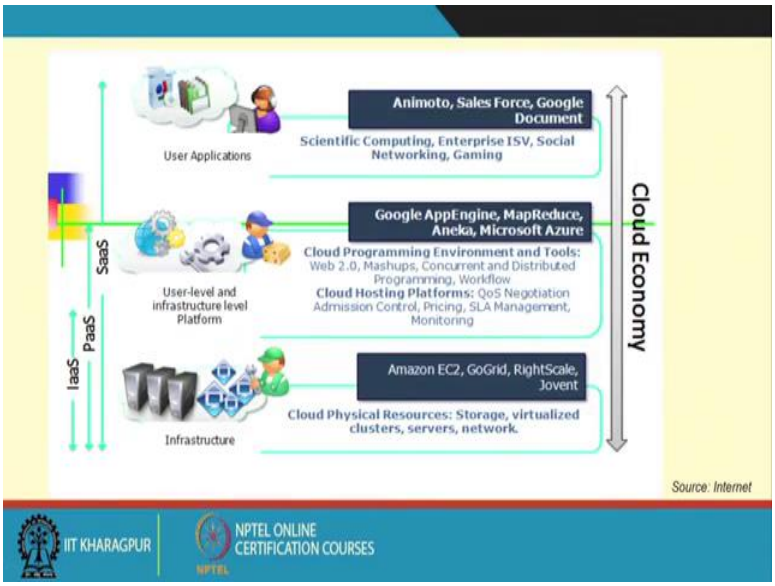
There are different characteristics like we have say that it can scale up or infinite scaling it has the ability of metered service broad network action and different characteristics are there omnipresent or ubiquitous access to these computing facilities and type of things.

Nevertheless these definitions what if we look at other in a little deeper way there is a lot of energy hungry resources are there right like if we may look at networks, servers storage and applications and services we run of the things they all take lot of energy if we can efficiently use those there may be there may be a chance that we can reduce the overall energy consumption otherwise the energy consumption may be considerably high. Like say if I say that if I have, if I my running VMs are say number of servers I am having 4 my number of running VMs are again 4.

So, every server can accommodate say running VM is 8 and every server can accommodate 4 VM. So, it may so happen that every server is distributed to VM per server it may look good that there is a load is distributed, but if you look at the energy point of view the energy consumption may be pretty high instead I could have packed them into two server of 4 4 each and so and so forth and other two server I can put on a off mode or a sleep mode or a passive mode and this could have saved energy in a bigger way.

So, these are the aspects which we need to try to look at and there is there is definitely a it is not like that that simple how we are trying to pose the problem how we are discussing, but there may be a lot of calculation of projects and type of things there are issue of SLAs, a key OS and type of things, but never the less taking all those in to consideration there is lot of opportunities may be there to go green right.

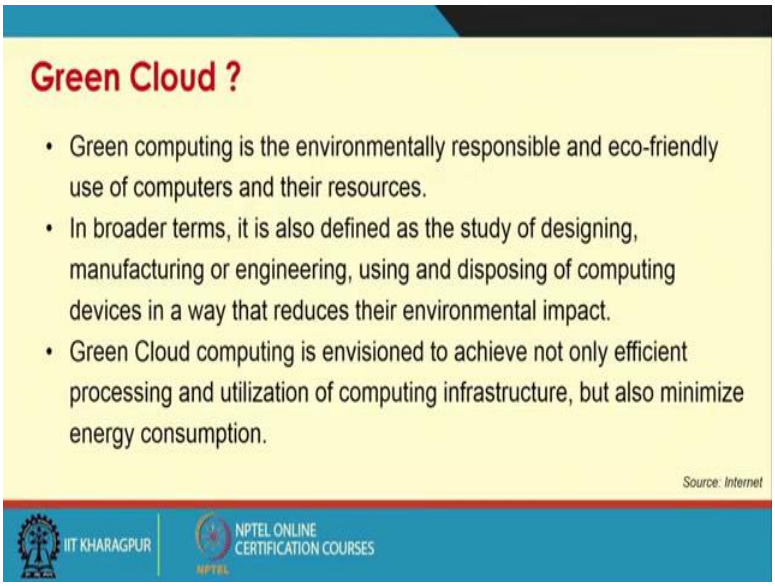
(Refer Slide Time: 07:36)



So, if you if we if we visit this means see the other slide. So, at the down we have a these infrastructure right where cloud physical resources storage virtualize clusters servers networks like Amazon EC2, go grid and different other solutions are there. At the middle we have a cloud programming environment or platform and cloud hosting platforms which uses this infrastructure and a there are Google app engine map reduce Microsoft Azure and Aneka type of things and at the top we have this SaaS or software as a service scientific computing etcetera.

Now these in turn every layer upper layer in turn uses the downwards layer and a more efficient is this computing at every more efficient is the implementation of each layer may help us in reducing the overall energy consumption right. So, it is not only that efficiency out here though we when we look about energy we mostly look at the IaaS type of things, but it is also that at a higher layers also contribute type some miss also contribute towards this proper energy management right. Like if I have an algorithm which takes unnecessary loops and takes more CPU time that may be energy inefficient then more efficient algorithm which is which is which where we can reduce the complexity and type of things right. So, it at much higher level I can do something which intern reduces my energy consumption or CPU usage time or network usage time and that intern reduces the overall energy consumption of the cloud infrastructure right.

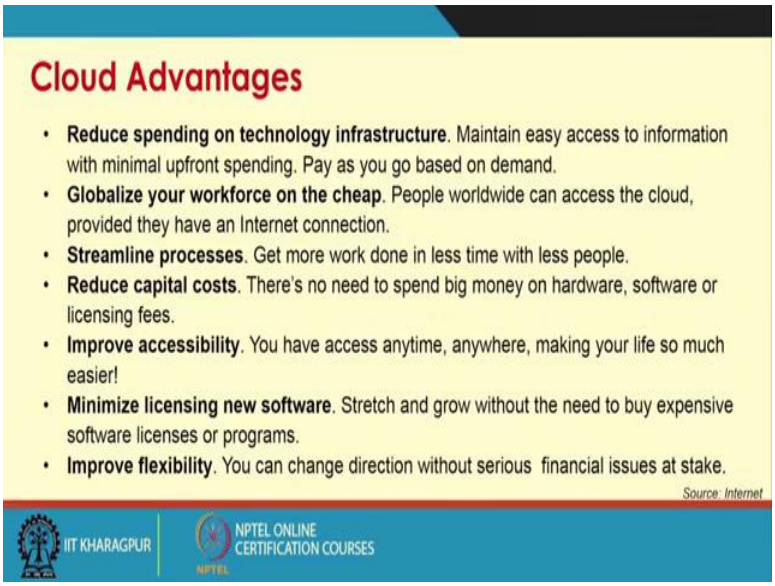
(Refer Slide Time: 09:37)



So, looking at these if we try to look at a green cloud, so green computing is a environmentally responsible and eco-friendly user computers and their resources right. So, its first start is that it should be environmentally responsible means minimum carbon foot print and so on and eco-friendly use of computers in broader terms it is also defined as a study of designing manufacturing or engineering using and disposing of computer devices in such a way to reduces the environmental impact right. So, if we look at that how the designing manufacturing engineering all those things go on green cloud computing is envisioned as a to achieve not only efficient processing and utilizes now computing infrastructure, but also minimize the energy consumption right.

So, it is not only how not only it will be efficient in processing and computing and giving maintaining key OS and SLA, but also it gives minimize the energy consumption, so what when we look about talk about green cloud computing we look both side of the thing.

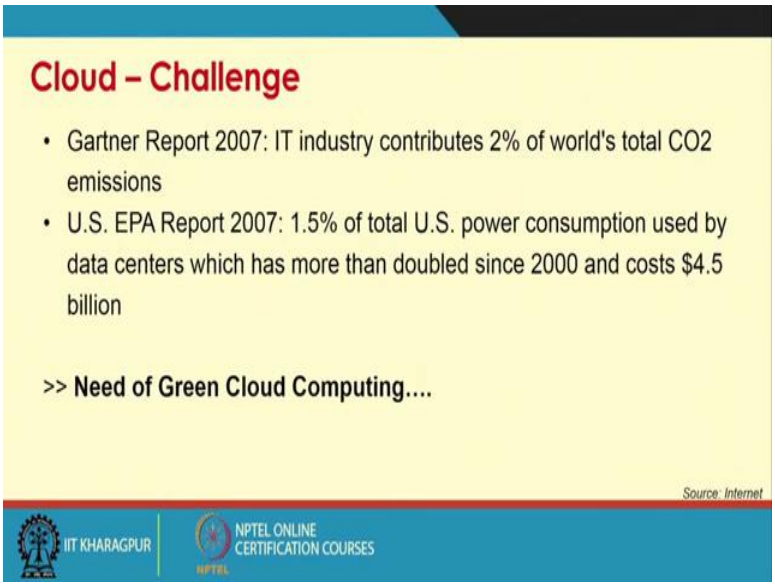
(Refer Slide Time: 10:50)



Cloud advantages are well known if we again look quickly reduce spending on technology infrastructure. So, what we feel that that instead of infrastructure I shift to the cloud infrastructure globalized your work force in a cheap. So, that it is omnipresent stream line processes reduce capital cost improve accessibility minimize licensing of new software improve flexibility.

So, there is a host number of advantages are there which may interns try to reduce the computing at different in a different environment and in a and going like in terms of setting up your infrastructure at your local things and going to the things. So, in sense it may appear that we are in a sense we are reducing a energy consumption, but at the other end the consumption things increases. So, what we are trying to look at is not that that what we say about here we are more at look at the how to make those service provider ends how things can be made more efficient.

(Refer Slide Time: 12:08)

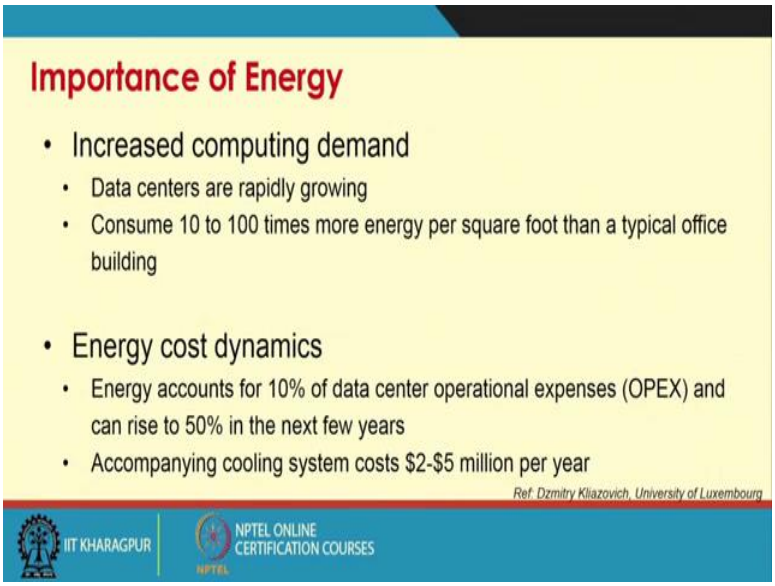


So, challenge specially in terms of energy consumption and carbon footprint like if you see the Gartner Report on 2007 it industry consume two percent of the world contributes to the 2 percent of the world's total carbon dioxide emission. So, it is a pretty high right it though its talks about it industry as a whole, but this major player or the cloud service provider are major contributor of the things right they are cooling they are overall energy consumption is pretty high.

Like big data centers installation people say that it is like a it is mini city like consuming power, so highly like. So, there is another report we say 1.5 percent of the us power consumption use by data center which has more than double since two thousand seven it is also around the around 2000 since 2000 and cause 4.5 billion dollar right it is also a report which is around 2007. So, in down the line 7 years it has more than doubled right.

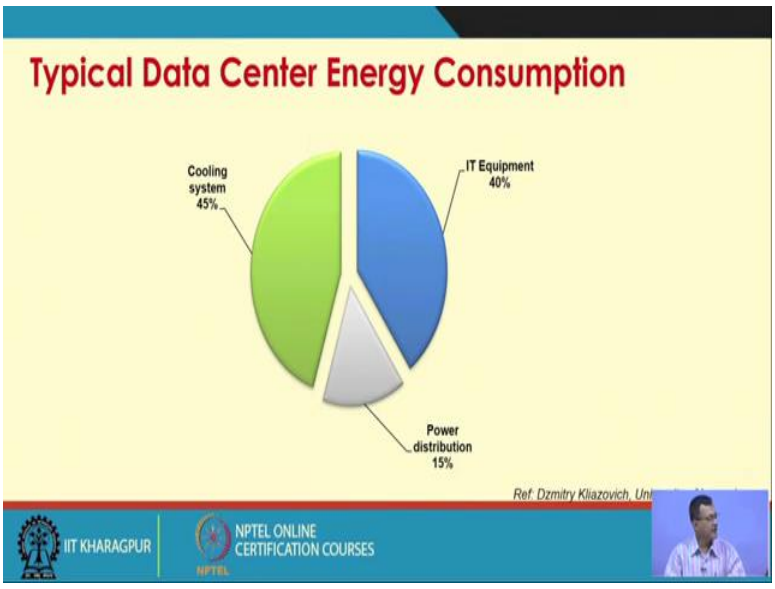
So, as it is ever increasing phenomena that this shifting towards cloud or cloud industry is growing at a much higher pretty higher rate. So, this figure of energy consumption is likely to increase much more than getting reduced. So, it is more demand for cloud more energy consumption and so and so forth right. So, there is a need of green cloud computing definitely there is a need of green cloud computing.

(Refer Slide Time: 13:56)



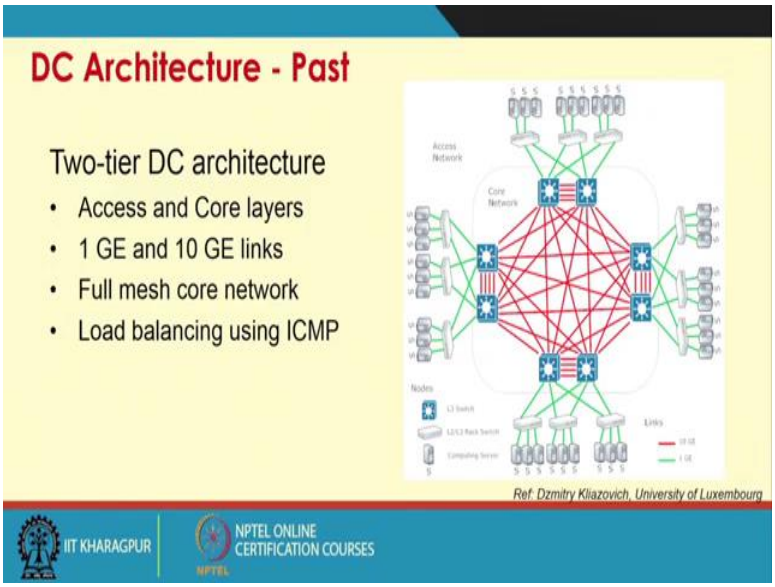
So, importance of energy increased computing demand data centers are rapidly growing consume 10 to 100 times more energy per square foot then typical of his building right this is a one rough cut thing energy is cost dynamics energy accounts for 10 percent of the data center operational cost right what other things might you say what we say OPEX and can rise to 50 percent of the next few years what they say that it is such a high rise going on it again one some report said that the cooling system cost around 2 to 5 million dollar per year this is also somewhere 2030 or so, that report which comes up.

(Refer Slide Time: 14:42)



So, again a rough cut if we see the energy consumption, so the equipment takes around 40 percent power distribution takes around 15 percent cooling system around 45 percent. So, if you say the if you look at the computing equipment is taking 40 percent over the total energy rest 60 percent is taking by this infrastructure is mesh thus end of a mesh cloud environment infrastructure itself right to set up this cloud that we need to set the environment that is sc power distribution appropriately power distribution etcetera that talking over the 60 percent of the energy the computing is around 40 percent. So, computing is still less.

(Refer Slide Time: 15:28)

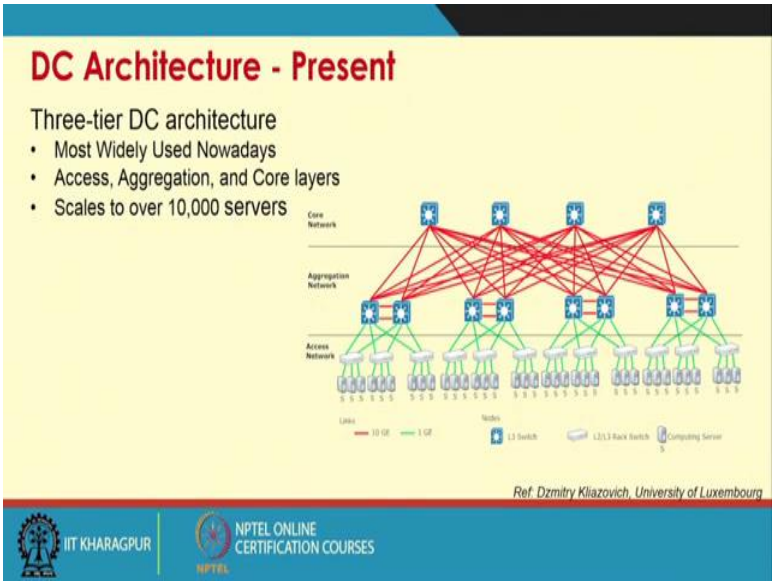


Now, if you look at the DC architecture. So, this has also evolving right it is also changing over time and more complex equipment are coming of course, more energy efficient equipment are also coming, so two-tier DC architecture access. So, initially it was access and cold layers. So, you have the centralized coal layer which is a highly mesh they are interconnected with a, vary concentrated mesh network and there are access layers from here the things are access these are the different access layers right this is the if you look at the DC architecture type of things computing architecture.

So, full mesh core network load balancing using ICMP protocol right, so this is typically there, so there are different type of switches then they are three switches they are two and they are three rack switches and computing servers. So, computing servers at the end mile or at the age and this layer three switches which routes traffic are at the core of the

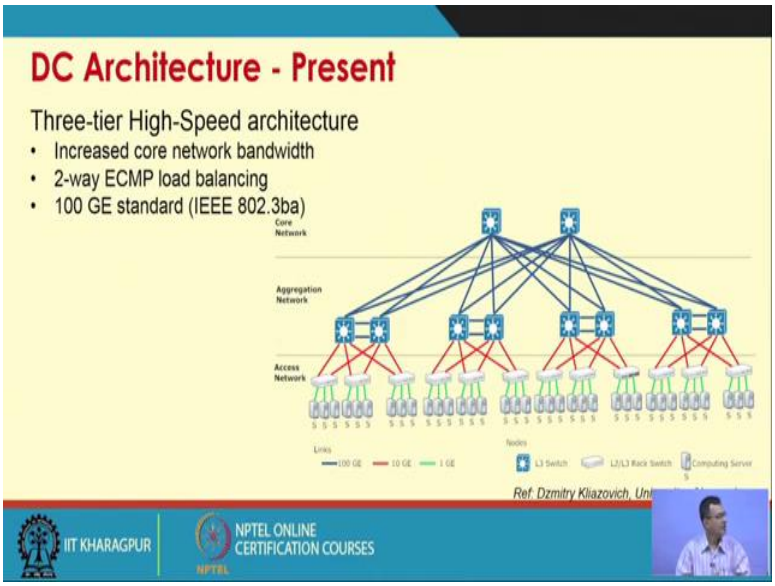
network. So, this is the typical architecture which were typically there in means old or past DCs though there are still some of things are still existing.

(Refer Slide Time: 16:57)



So, rather like all other large network this DC architecture present which see as a three tier DC architecture that is most widely nowadays access aggregation and core layers. So, there is a access layer, there is a aggregation layer. So, it is not only these layer three switches at the core, but also there are three switches at the aggregation layers and then we have that then this access layer. So, it is a hierarchical structure made to proper management and appropriate distribution which minimizes and of or better load balancing and type of things and so and so forth right. So, this is more for, it is more it is amicable for two scale up its scales up much better than our previous things like scales over ten thousand servers and so and so forth for a particular disease. So, it is this sort of architecture are there.

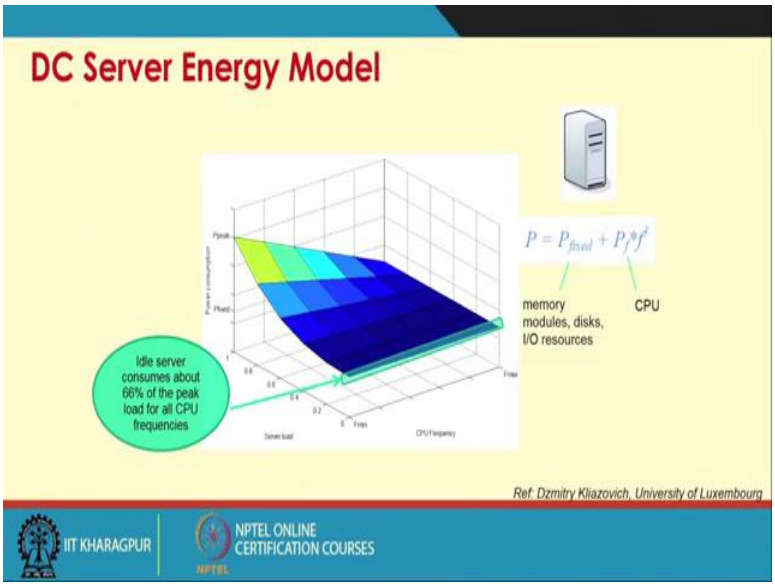
(Refer Slide Time: 18:03)



So, along with that it has high speed architecture three tier high speed architecture in increased core network bandwidth, 2-way ECMP load balancing. So, 100 giga bit ethernet standard connection over the thing. So, it is a much higher thing. So, what have thing previous thing was a standard connectivity with 1 to 10 giga bit connectivity whereas, here we have a jump of 100 giga bit connectivity. So, these links are very high speed links and then we have much lower links at the down the line right. So, it is better aggregations better scaling up and better management of that whole infrastructure is there. So, it is more better responsive type of architecture.

So, this present days architecture whenever we deploy all those things these are again we need to look at the energy consumption of the things. It is not only facilitates better computing or better accessibility of the computing things, but also it consume at times much more energy right. So, we need to have a tradeoff that whether there is a requirement of sub such things or that facilitating or providing services and increasing energy performance versus energy trade off should be three though we do not comprise on; do not want to comprise on performance, but making it efficient or energy efficient is one another goal.

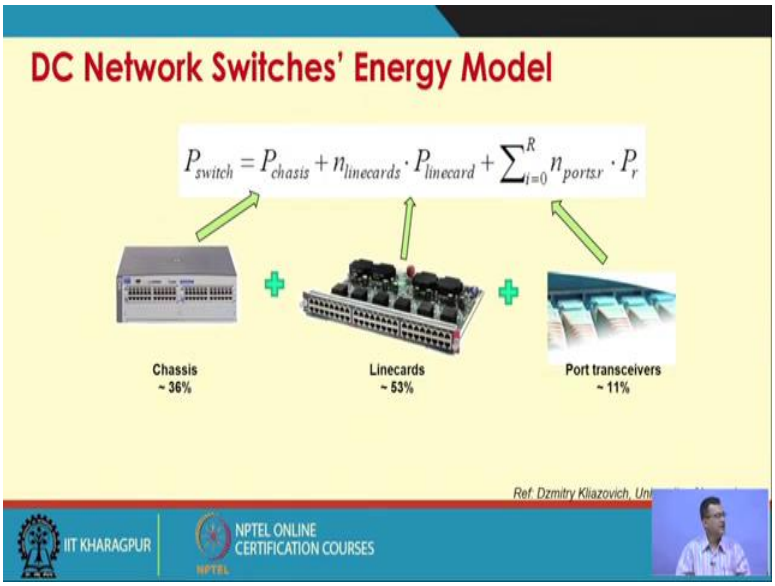
(Refer Slide Time: 19:55)



So, this is a typical energy model of DC server energy model what you see that even in the idle server consumes around 66 percent of the pig load of the CPU frequency. Even in this idle state it takes a considerable amount of energy so that it is always running energy. So, ideally if a particular service provider has no loads something it has to maintain the things. It is like if we consider that is in a if you in a particular shops say ice cream parlor or so, even there is no customer or not that particular customer it has to maintain as a particular level of cooling and type of things right. So, it is a some sort of energy level is there.

There are typical there are different model of energy models. In this case there is we have a fixed power model where the memory modules disk etcetera another CPU based on that frequency n number of fix CPU up etcetera we can have another other modules.

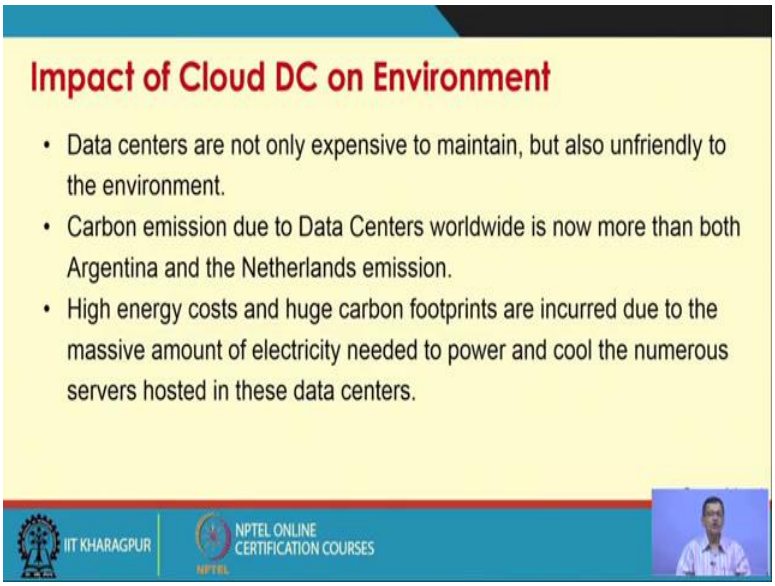
(Refer Slide Time: 21:10)



Similarly, if we look at the switches angry models so there are different category of things one is that what we see that the chassis, the chassis where if we have say a bunch of blade servers. So, it goes into a chassis right. So, typically say chassis contain the sixteen blades or of half height size so that chassis itself consume energy. So, if you look at that, the 36 around 36 percent of the a typical figure just to show you a rough cut figure that how things how the energy is important a chassis consume around 36 percent the line cards around 43 percent and these port trans receiver where the data being transmitted or transmission received is 11 percent right.

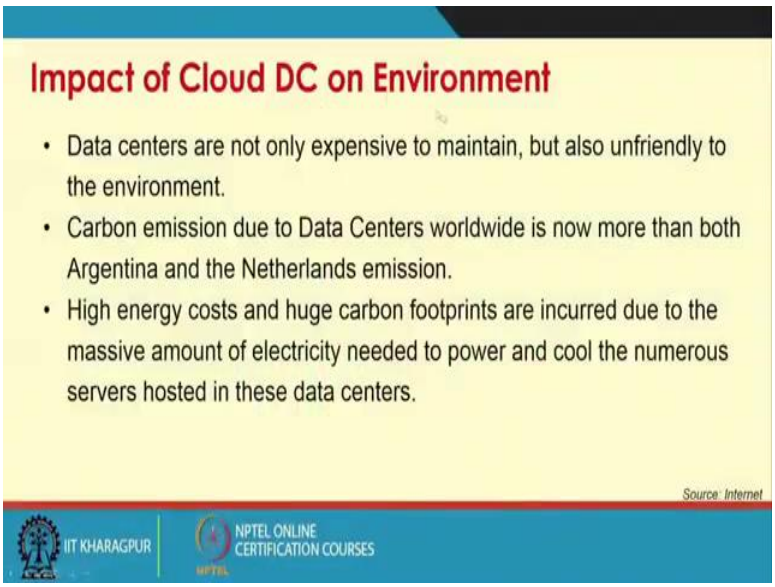
So, this model that p chassis plus number of line cards into p line cards and that number of aggregation of the number of ports and along with that the summation of things with that number of ports is that summation of the energy consumption will be the overall switching energy consumption. And that is also if we see that it can be considerable based on that if it is properly loaded and type of things like I can have traffic appropriately distributed I can have a better figure out here.

(Refer Slide Time: 22:46)



So, this all those things has a has a definite impact on the environment right or environment, this environment is not what we are talking about the cloud environment, but it is environment as a whole that overall environment like as we have talked about carbon footprint there may be effect of a heating there may be other different sort of a polluting effect right.

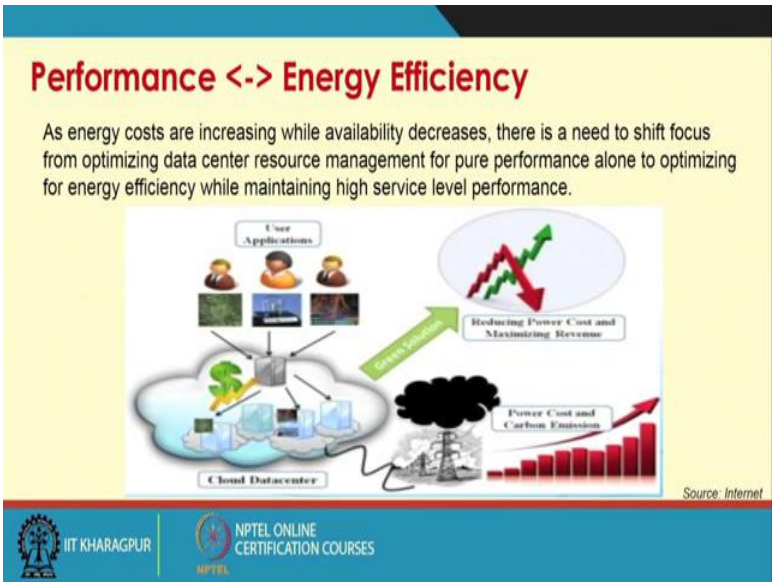
(Refer Slide Time: 23:11)



So, data centers are not only expensive to maintain, but also unfriendly to the environment it can be unfriendly carbon emission due to datacenter worldwide is now

more than more than both Argentina's and Netherland's emission. So, this is one figure it shows that the carbon emission is due to data centers worldwide is more than two counties overall emission high energy cost and huge carbon foot prints are incurred due to massive amount of electricity needed to power and cool numerous servers hosted in this data centers right. So, this is another major challenge of looking at it. So, it is a huge energy consumption.

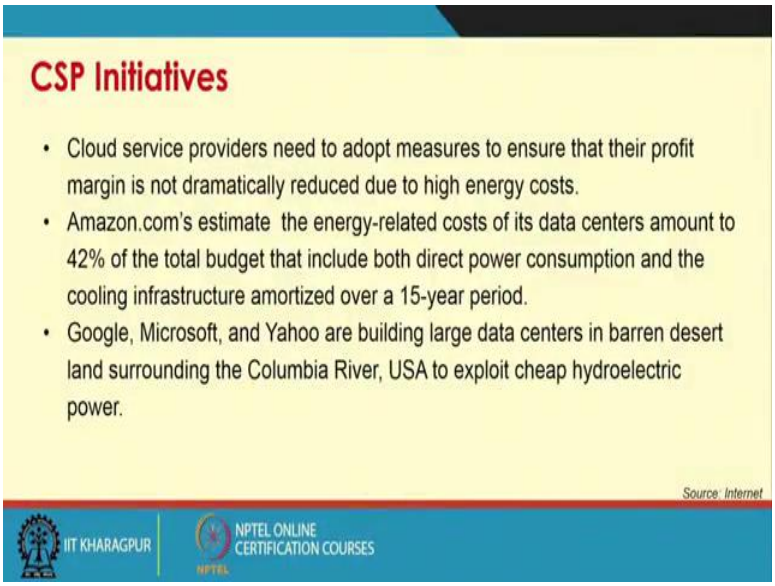
(Refer Slide Time: 23:53)



So, you need to balance between these performance versus energy efficiency, as the energy cost increasing while availability decreases there is a need to see it focus from optimizing data centers resource management for pure performance along to optimizing energy efficiency while maintaining high service level performance. So, as the like when we talk about resource management we may be looking primarily on the performance alone right. So, you need to look at that performance vis-a-vis these energy management. So, how overall this can be achieved by performance versus energy efficiency can be achieved is there, is like reducing power cost and maximizing revenue may be the thing like power cost and carbon emissions are ever increasing. So, you need to look at in a more practical way.

So, if you look at in the modelling terms. So, in our this model this energy components would come into play right that how that need to be managed how need to be controlled should come into play.

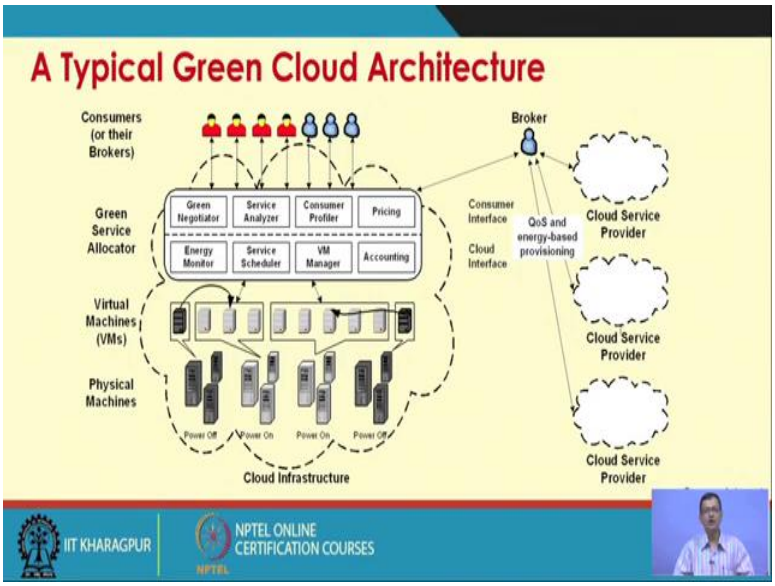
(Refer Slide Time: 25:05)



There are several initiative needed from the cloud service providers end like cloud service provider need be adapt measures to ensure that their profit margin is not dramatically reduced due to high energy cost right. Amazons estimate one figure shows that the energy related cost to the data center amount to 42 percent of the total budget and include both direct power consumption and cooling infrastructure amortizes to over 15 year period.

Google, Microsoft, Yahoo are building large datacenter in barren desert land surrounded by Columbia river to exploit the cheap hydroelectric power etcetera. So, that is a tendency of make the datacenter more near to the power to the power generation unit. So, that your transmission of power transmission loss etcetera are reduced to a drastic element.

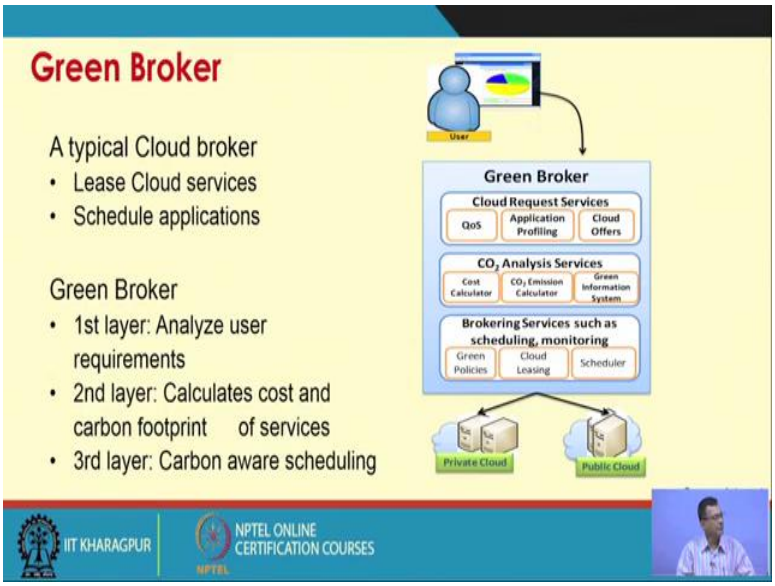
(Refer Slide Time: 26:04)



So, taking, so if this is typical green cloud architecture if we look at, so there are at the bottom end is physical machines there are several virtual machines and this green cloud allocator what we say now the green cloud brokering system, which brokers in favor of the consumers. So, this is a grid negotiator, service analyzer, consumer profiler, pricing, energy monitor, service scheduling, VM manager accounting this components where otherwise also they are in the architecture. What we see more is looking at that energy consumption energy consumption related parameter or matrix come into play right.

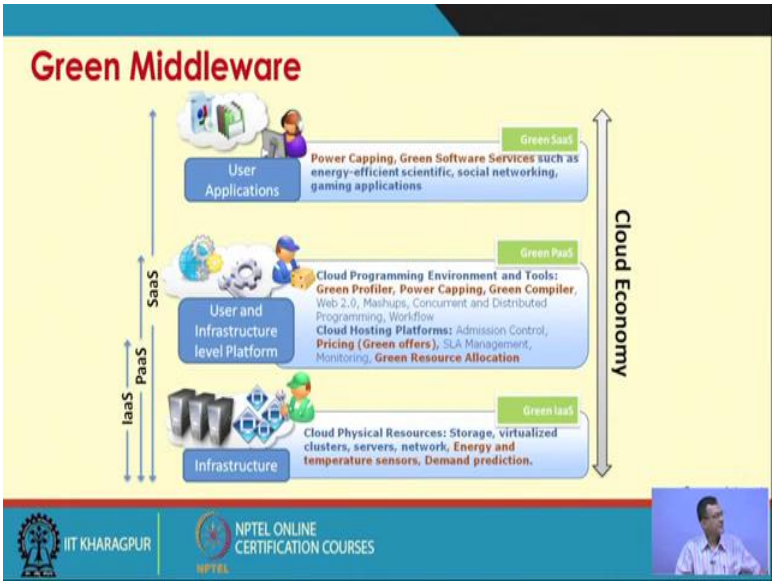
So, broker which look at QS and energy based provisioning of the different cloud service provider based on these they provision on these different cloud service provider. So, it is more energy aware provisioning of services.

(Refer Slide Time: 27:19)



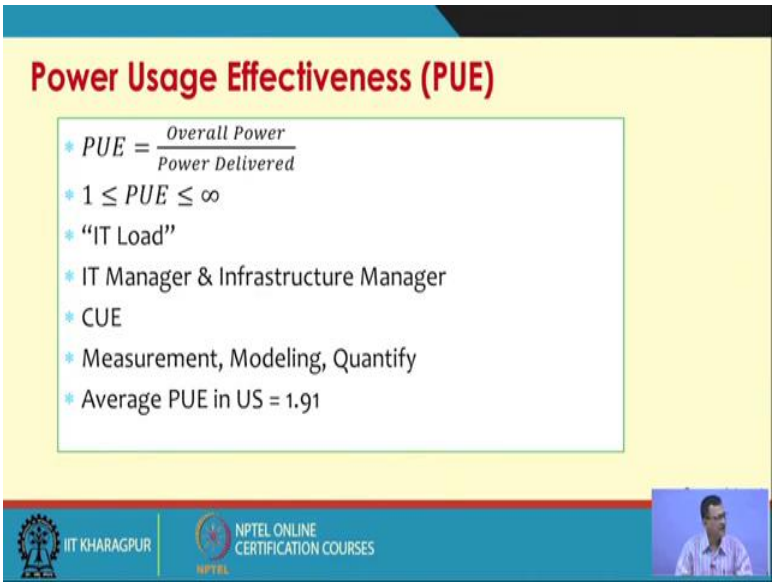
So, since that green if we look at the green broker. So, it has, so it a typical ground broker lease cloud services schedule application that is the measure duty of the things. So, when you look at the green broker. So, analyzing user requirement calculate cost and carbon footprint of the services carbon aware scheduling. So, in now the scheduling is carbon aware scheduling right brokering services scheduling monitoring carbon dioxide analyzing services cloud request services. So, it does a carbon aware scheduling.

(Refer Slide Time: 27:54)



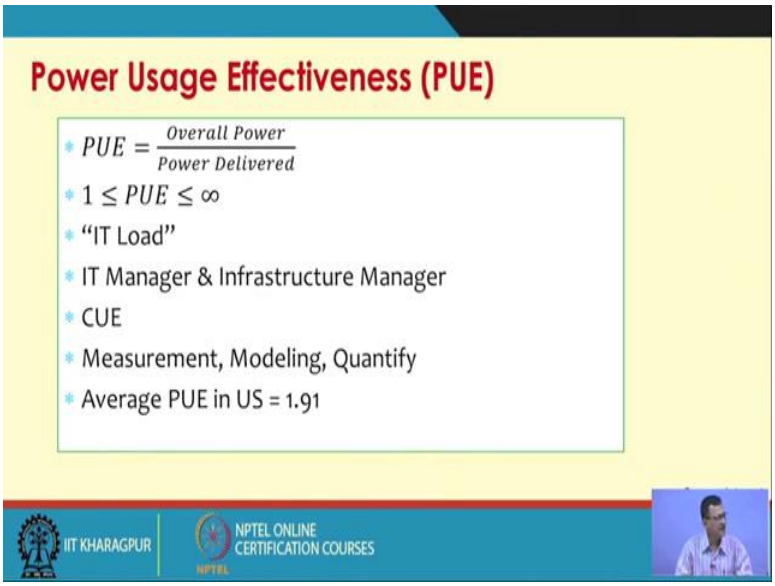
Similarly, we have a green middleware where what we say that green IaaS. So, what if you remember the initial figure, so we said storage virtualize service etcetera energy temperature sensor demand predictor is added. Here also we have at the PaaS label that green profiler power capping green compiler similarly green resource allocation system at the PaaS label, in the SaaS label power capping green software services and so and so forth. So, these are the different way we try to address this issues.

(Refer Slide Time: 28:36)



There is some effort of looking at the power usage effectiveness measure to find out that what is PUE and how to measure that how the how effectiveness is this power uses. So, there are different theories behind this. So, this there rough cart measures that how whether we can measure this power usage effectiveness of a typical infrastructure or a typical cloud service provider.

(Refer Slide Time: 29:05)



So, to summarize clouds are essentially data centers hosting application services offered on subscription basis; however, they consume high energy to maintain their operations, so high operational cost plus additionally environmental impact which we try to ignore. So, that is one of the major aspects. Presented, so what we look at like look at that carbon aware green computing frame work to look at. So, there are several open issues lot of research to be carried out to maximize energy efficiency as a cloud centers, developing regions or to benefits the most benefits more that wares would be situation and so and so forth.

So, what we see overall that a overall this sort of computing aspects has a major concern not only from the service provider or consumer point of view it is a concerned worldwide from the environment point of view that the huge amount of energy being consumed which has a carbon footprint and a better and there is a need for better energy management to for so that this type, this sort of cloud computing environment in with lot of benefits are able to exploit it by the consumer. So, we need to head for some sort of a green cloud computing environment.

Thank you.