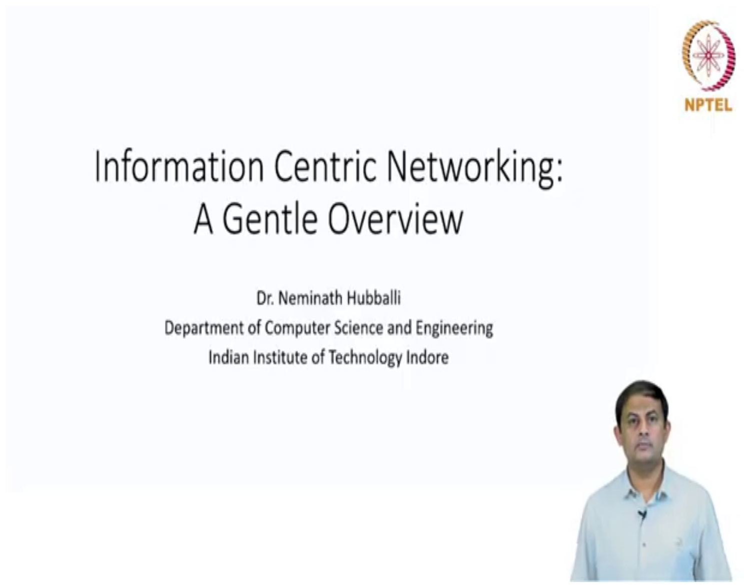
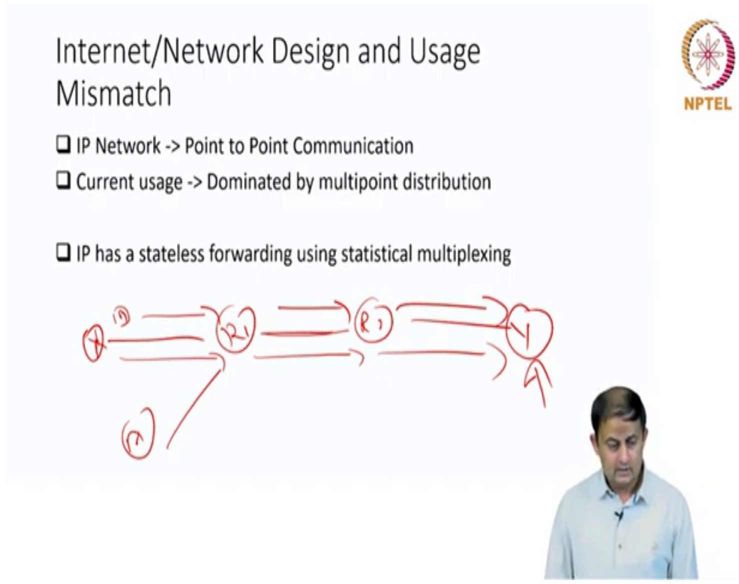
Advance Computer Network
Professor Doctor Neminath Hubballi
Department of Computer Science Engineering
Indian Institute of Technology, Indore
Lecture 63
Information Centric Networking Part 1

(Refer Slide Time: 00:16)



Welcome back. In the last lecture, we saw some of the techniques for Content Distribution in the IP network, so we said that the current IP network is designed for point-to-point communication, and then because the applications are there, they are doing the multi-point distribution, some of the things that were designed and retrofitted into the point-to-point communication, so particularly the CDN networks, the caching techniques, be it the proxy server level caching or the network service providers doing the caching and some other peer-to-peer networking mechanisms so all of them we saw.

(Refer Slide Time: 00:59)



So, the summary of the discussion that we had in the previous lecture goes something like this. So, the IP network is fundamentally a point-to-point communication so it means to deliver the content or transfer the content from point X to Y, and using that underlying communication mechanism, I am actually building an application which is doing the multi-point distribution.

So, although many of these techniques that we discussed, like the CDN networks, improved the response time and localized the content storage to a large extent bringing efficiency but still the underlying mechanism, the transmission that is going on is still using the point to point communication so, that is, the inherent efficiency of the point-to-point communication is still existing in the underlying network.

So, because the IP network is actually point-to-point communication and it has been engineered to work in a stateless fashion, what I mean by that is if X is the one which is transmitting the content and Y is the receiver, and there are a bunch of intermediate routers, so if X transmits let us say a packet P1 to the receiver Y, and if X desire to transmit the same packet same content 1000 times, the network blindly transmits that content 1000 times. In effect, what I am trying to say is that R1 does not remember whether packet 1 has been previously transmitted on this, from the same source to the same recipient. So, let us say P1 is carrying an image, and the same image is getting transmitted 1000 times to Y, that is, the inherent inefficiency.

And the inefficiency in the sense we are looking from a perspective which is the usage perspective in today's Networks. So back then, this was the best choice that one could have made, and that is how it has been engineered so it is stateless and works in a passion which is something called as statistical multiplexing.

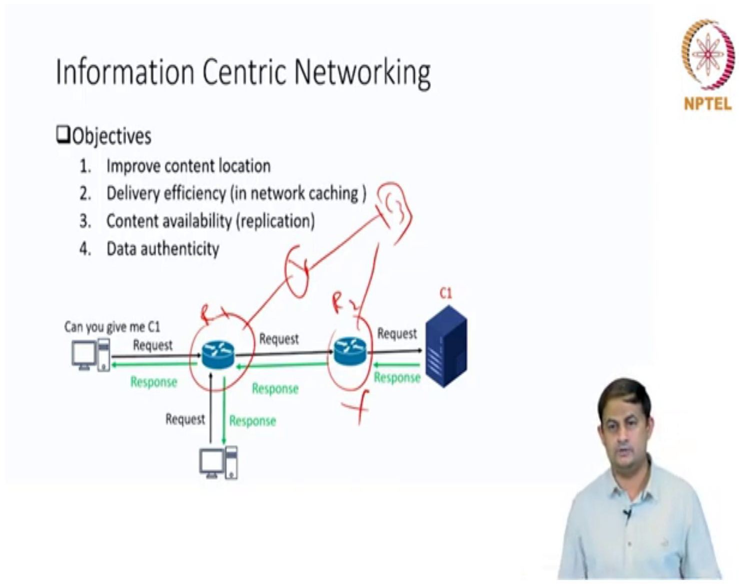
So, what it means is if there are multiple senders, X maybe and A is also sending, R1 looks at what is the available bandwidth, the bandwidth on which the packet has to be transmitted, and if there is bandwidth available, it will simply transmit that without differentiating whether X is transmitting or giving the packets, or A is giving the transmitted packets.

So, both of them, from the R's perspective are the same. So as long as there is capacity to transmit, you simply transmit, that is, the statistical multiplexing means that at some point, if only X is transmitting, then the whole bandwidth available for transmission is given to only X transmission. So, that is the best approach that one could have used, given the scarce resources. Resource in the sense here is both the ability of the R1 to process those packets and perform the lookup operation. And then also mean the bandwidth available at the R1's disposal to transmit the packets.

So this is still there; we are still using the same underlying mechanism to do the transmission, so that is a cause of worry for us because of the efficiency, so now this kind of communication, whatever is there in the IP network, is because it is stateless and it is doing the statistical multiplexing. It opens the doors for several cyber-attacks as well, for example, if X is having malicious intent and he wants to bombard Y by sending multiple number of the packets let us say 1 million packets per second he is transmitting without any purpose, nothing actually prevents X from doing that so you indicate the Y does not have any control over what it can receive, a simple thing that can do is maybe it can filter the traffic from coming from the X.

But if X is not using his own IP address, if he is randomizing the IP address and then sending those packets, then Y has nothing to do at its disposal with which it can safeguard itself so it becomes a hard choice for Y to mitigate this particular traffic. So, that is one another issue that you have in the current IP network.

(Refer Slide Time: 06:01)



So, with that background in mind, let us see what Information-Centric Networking is all about and what are the design goals and objectives for this Information Centric Networking. So if you could recollect the previous lecture discussion at the end, what we said is a user wants some content, so he wants an image, he wants a video, he wants some audio content to be delivered to itself, but then how do I you actually get that content. So, if a communication model which is centered around the information or the data is what we are looking for that is precisely what the Information Centric networking means. You keep the information at the center of the communication, and around that, you build the mechanism to do the transmission and the distribution in simple terms that is what it means.

So, let us see how that Information Centric networking is supposed to work, at least in the idealistic scenario, so what we have in this picture is a server that is where the content is actually produced, and somebody who is interested in receiving that content, so the guy on the left-hand side makes a request, I want content C1 assuming C1 is available at this server. He says that to the network, in essence, to the nearest router to which the client itself is connected or the user is connected.

So, you ask can give me content C1, and the network or the routers figures out where exactly C1 is, so there is nothing called IP address, there is nothing called a TCP connection between the client and the server all that the client asks is to the network, I want C1, it does not matter from where you bring it, can you give me C1? that is all it says.

And the router internally knows how to figure out where exactly C1 is or the host which actually has the content C1 and then it forwards the client's request to another intermediate router which is closer to the content, where the content is available, so here in this case the server is the one which is having the content, so it forwards that request to the server, not necessarily only two of them but through a bunch of intermediate such routers who can actually understand what is being requested and where to find out that content. So the request finally goes and reaches the content producer, and the producer, in response to that, sends the content back in the reverse direction, so first if you name this router as the R1 and then this router as R2, the response or the content C1 first comes to R2, and then it goes to R1, and finally it is delivered to the requester so the client which made the request, the content is delivered to it.

So, now the role of the R1 and R2 here is when the content is going in the reverse direction; they also make a very informed choice to keep a copy of that content C1 within its own store, so in effect, what it means is every router not only have an understanding of what content is being requested and where exactly the content is presently, but they also have an inbuilt storage mechanism where you can store the content.

Here, in this case, assuming for the time being, both R1 and R2 make a choice to store the content C1 when it is forwarding the content in the reverse direction, if subsequently, another client here, in this case, the client at the bottom of this figure makes a request to R1, so instead of going to the server on the right-hand side, what it does is, it responds back supplying the copy that is stored in the previous communication.

So, that is a very nice thing, assuming the content producer is some 20 hops away, so the client in the second case gets the content in one hop communication, that is a very nice thing to have, so by virtue of this ability to cache the content and then an ability to do and forwarding operation by looking at the content requests, you minimize lot of work that is happening in the network. So, this is what we call as the Information Centric Networking.

So, you keep everything happening in terms of the information in terms of the content, keeping that content at the center of the transmission you are operating. So, now this is what we exactly want, let us formalize this ability of this network in terms of some characteristics of what exactly we want to achieve and what are the design objectives of such a network.

So one of the primary objectives is to improve the content location in the sense that what I mean by that is, in an idealistic scenario, no matter who is requesting the content, whether it is client 1, client 2, client n, all of the requests made by all these clients in the entire global network are serviced by the routers which are closer to the requesters themselves in the sense that when client 1 requests, the first hop neighbour of the client 1 is able to service that request; when client 2 makes a request for some other content not necessarily the same content that time also I will be able to do the servicing or get back with the content to that client within one hop communication.

So, every request is serviced by not more than one hop communication. So, in effect, what I am trying to say is wherever the clients are located and the nearest routers have all the contents that these clients are going to make in the future, that is an idealistic situation we want to have, by arguing that we want to keep a copy of the content what we wanted to achieve is this one. So, every content request originating from every client would be serviced by the one hop, from the nearest router that is connected to the client itself. So we want to achieve this; whether we will be able to achieve this or not is a secondary question, but this is where we want to be, at least idealistically.

And second thing is, since we keep a copy closer to the client which is consuming that content, this brings the delivery efficiency. So in the previous scenarios, what we discussed many-many times, the same content is being transmitted over the network that brings the redundant data transmission, so that is avoided. I want to achieve a zero redundant transmission at least idealistically; that is the second objective of such Information-Centric Networking.

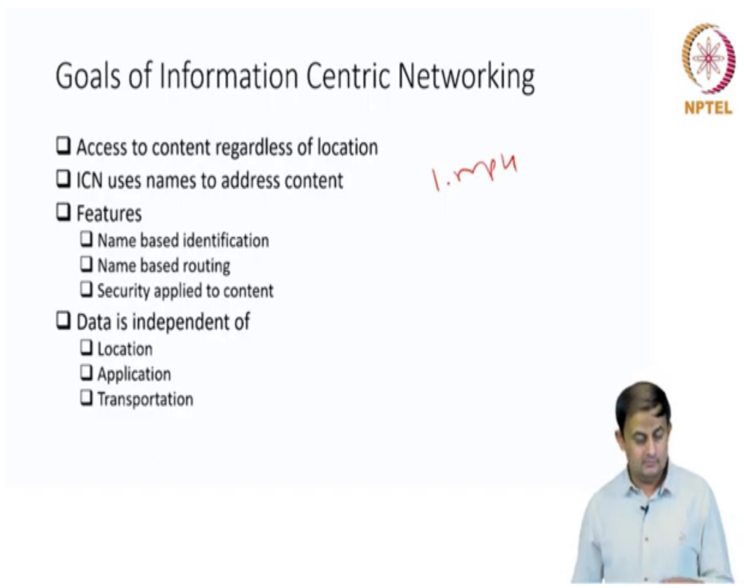
And then since as a by-product of this operation, I am keeping many-many copies in the network, many of the routers have a copy of the transmitted content, so there is inherent redundancy, and that replication improves the availability. If the original producer or some other router in the network having a copy is not available or down for some reason, then you can go elsewhere and find out the same content.

So, if let us say R2 is down for some reason, somebody who is connected to it, let us say, client 3 wants to access that content assuming he has got connected to it, or it can go and reach R1 through some other link you can still find out the content R1 within the network so that is what we want to achieve. By replicating the content or increasingly caching the content at different locations in the network, we want to improve the availability of the content itself.

And unlike the current scenario of data delivery, where the security notion is a tunnel that exists between the data consumer and the producer, we want to depart from that notion of security, and what we want to achieve, at least idealistically, is every piece of data that exists in the network and delivered as a response to any request originating from any client is guaranteed to be generated from the authentic source that is what we want to achieve.

So, it does not matter whether the router or R1 service provides that content or router R2 provides the content or the original producer, data producers, the server itself delivers the content, or every copy that exists in the network can be verified to be generated from one source which is the authentic source that is what the design objective is. So these are the idealistic objectives of such a network and then how do we achieve this is the secondary question.

(Refer Slide Time: 16:44)



So, let us go and see what are these objectives that translate into something called as the design goals, so what we said access the content regardless of the location, meaning that the client really does not care whether the content is available at a router that is directly connected to it or at a router which is 10 hops away or a producer which is 20 hops away.

So, all that he knows is what content he wants, so there is a movie xyz dot something file extension, so the client asked can you give me this movie xyz dot something and then the network figures out where is that xyz dot may be in mp3 or whatever the extension it is mp4 or something like that so client really does not know where the content is, and it should be able to access that content without actually establishing a connection to that host which is having the content, there is nothing like a TCP connection going from the client to the place where the content is located be the router or the one of the data producer.

So, that ability we want, so this goal is coming from the objective that we put forth in the last slide, and because the client is asking the network can you give me this content it is identifying, it is proposing a name and using that name the network is delivering the content. So, in effect, that means that the network knows what is the naming convention and it understands what a particular name means.

So, for example, if it is asking can you give me a movie that is 1.mp4 and then what is 1.mp4, the network knows that, router R1 understands what is 1.mp4 and router R2 also understands what is 1.mp4 and every router in such a network understand what is 1.mp4. So, everything is operating in terms of the name that is the second design code.

Now how do I make the routers aware that to operate using such names is the question you have. It also means that you also have a name-based identification of the content, and the routers know how to forward the request using the names and store; what is there inside my own cache because every router is equipped with a storage that is called cache.

So what is there in my cache, what is there in my neighbor's cache, and what is there in the router which is 10 hops away, everything every router, at least in an idealistic scenario, should be aware of. And third is even when I keep a copy of the content in one of the routers, the security of that content should be preserved and these are the end goals. How do I achieve this in practice?

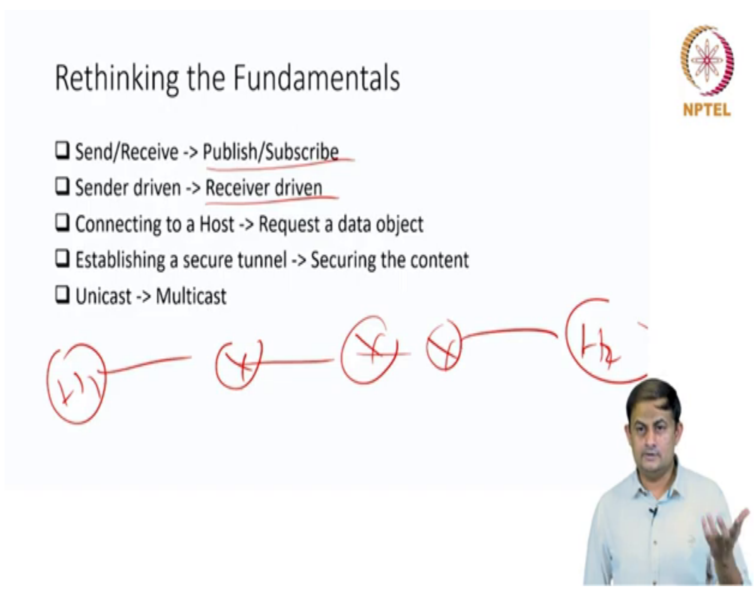
And then we want to separate out the data, that is, the naming convention is from the client's perspective. When you forward the request from the client to the nearest available content store, be it the router or the producer, so these are the abilities we require; you require naming, you require name-based routing and you require the content security, and from the other perspective when the data is moving and getting cached in the reverse direction you do not want to attach a location to that data it does not matter whether R1 caches it or R2 caches it or R10 caches it, any or all of them caches it.

So, I do not want to tag a particular piece of data at any location in the sense here you can roughly map it to the IP address, so any data content should not be mapped to one particular location by virtue of caching that content. You can goto any of the router and ask you have this content can you give me, that is a different story.

But I do not want to hard-code, tag a particular piece of information to one of the IP addresses or the location. And it does not matter which application is consuming that content and what kind of transport protocols or the mechanism that we use to carry that content C1 from the producer or the location where exactly it is currently residing to the intended recipient, so these are the end goals.

So, everything is operating in terms of the name, and then there is nothing called one particular host or or a particular router which is having the content there is no tagging connection between this name and this content storage, so that is what the end goal is. Now, these are the things that we want to achieve.

(Refer Slide Time: 22:18)



Now the question is how do we achieve this, Is it possible in the IP network? maybe or may not be. So the answer to this is how do I translate these design goals and objectives into reality? It boils down to revisiting some of the fundamental concepts that were used in designing the TCP/IP network. So the current TCP network is operating something on a semantics called send and receive mechanism, somebody wants to send a data, a piece of information to someone else he can send it.

So, if my computer wants to send a packet to someone else's computer, if I know the IP address of that computer, I can really send that particular piece of information. What he does with that information is a secondary thing, but as long as my intent is to transmit that packet or the data to that recipient, I can still do that.

And we want to depart from such a semantics where the sender has got an absolute control over what he can send to whom, to a semantics where the consumers can control what they can receive, so meaning if I am interested in receiving the packets or the data originating from a particular source or particular type of the content I should be able to do that; from a sender controlled approach to a receiver controlled approach that is what we want to achieve in practice. So, this is what we want achieve is the receiver driven operation. So, if I do not want to receive some sort of the content, I should not be receiving that content, that ability we want to bring in.

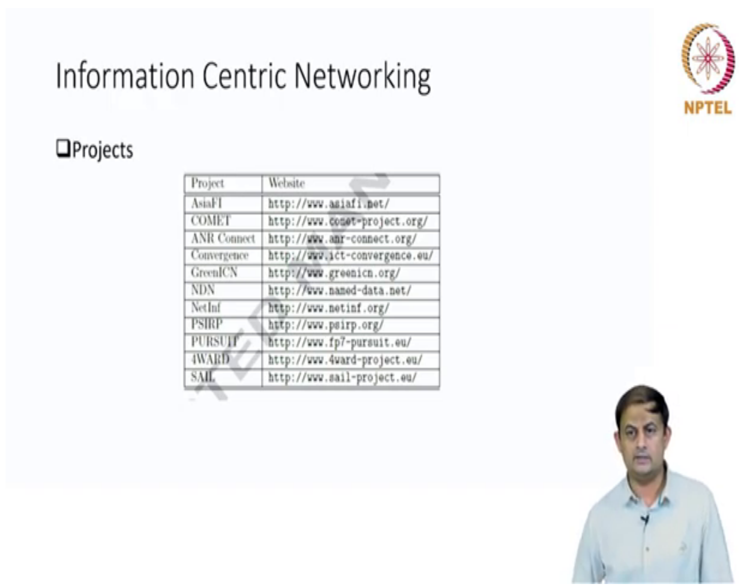
And if I am interested in only a particular type of content, I can subscribe to that particular piece type of content and get only that type of content. So, in the previous lecture, we saw how the published subscribe mechanism works where any data producer will tag the content with a particular name or the topic and the subscribers show interest in receiving such type of the content and whoever has shown the interest, the content will be delivered to them.

So, we want to move to that kind of semantics, so the publisher here does not have any control over to whom all he can push that content to. so the receiver has complete control over what it wants, from whom, which topic, content of the what type, he wants to receive. And then the third conceptual revisit that we want to do is establishing a connection, so between two hosts which are connected to the network through some intermediate routers.

What we want to do is just that some of the host makes a request to some of the content, and the network knows how to search that content and figure out what is the nearest place I can get the content, whether it is one hop, two hops or ten hops from there, it pulls out the data and then gives it to the intended recipient who has made the request and we do not want to establish a secure tunnel, instead of that, we want to secure the content itself. So, I want the authenticity of this content is originated from so and so source, from whom or whose content I as a consumer want to consume.

So, since this we are talking about content distribution, that is a multicast operation, a subset of all the recipients globally exist will be interested in content generated by some producers, so from a unicast model of communication, we want to go to a purely multicast transmissions semantics, so these are the fundamental revisits that are required are the differences that exist between the IP networks and the idealistically designed Information-Centric Networking.

(Refer Slide Time: 26:59)



So, these are things that we want to want to achieve, so now how do we achieve it. As an answer to whatever we discussed in the previous slide are all the design choices or the fundamental things that exist in the networking and then in order to translate those to the requirements into reality several research projects started at least in the last decade also.

So, these are some of the research projects that are exploring how exactly we can realize this Information-Centric Networking, how do we bring the ability of the storage, how do we bring the ability of the naming the content, how do we do design routers which are aware of analyzing or retrieving the content based on the names; so there are people who are exploring different mechanisms.

So, although we do not have at this point of time a particular method which is working in practice or at least having shown promise, but there are many parallel projects in use, exploring the different directions of achieving these research objectives, so many of them in fact have the websites where you can visit and then explore the details of what exactly they are bringing they are proposing the different architectures for such a network, which is operating on using the content as a central piece of information.