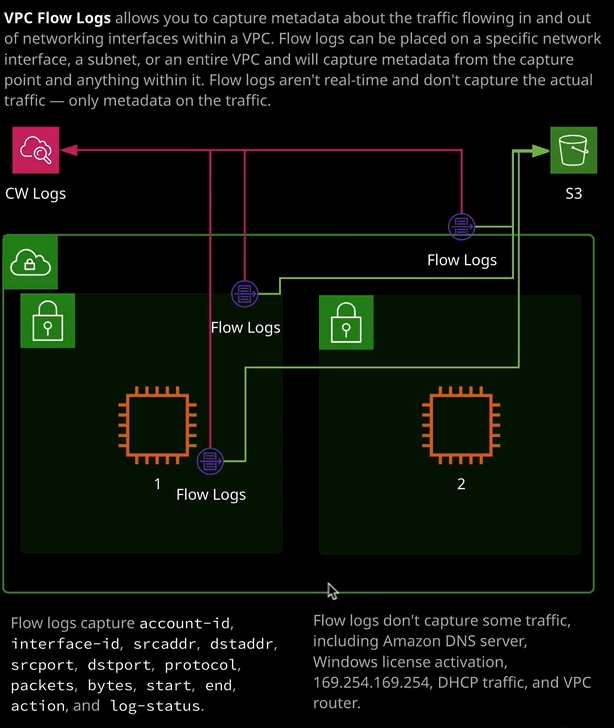
Welcome back and in this lesson I want to talk about VPC Flow Logs. Now, the exam won't expect you to understand how to implement VPC Flow Logs, but you might well get a question which tests your understanding of exactly what is monitored by VPC Flow Logs as well as the overall architecture.



So I think the easiest way to help you understand exactly how Flow Logs work is to actually demonstrate its implementation. So to enable the VPC Flow Logs is actually controlled from the VPC console. So I'm going to move across to VPC. Now in this account currently all I've got configured is the default VPC. So if you are testing this in your own environment and you don't have a VPC configured then the easiest way is to go to actions and then just go ahead and create a default VPC. Now, I won't be able to do this because I've already got one created. So this is my default VPC. To enable VPC Flow Logs you just go to actions and then create flow log. So VPC Flow Logs does not monitor any traffic within VPCs only monitors traffic metadata. So essentially for any traffic that's traversing any of these monitoring points, and I'll talk about what a monitoring point is in a second, it logs information such as account ID, interface ID, source IP address, and destination IP address, source and destination port, the protocol, number of packets, the bytes, the start and end time of the period, the action so whether the traffic was allowed or not, and then a log status value. So just to stress and this is really important for the exam. If you face any questions where you required to sniff traffic or monitor the contents of any traffic then Flow Logs will not work. It is not a product that can monitor the contents of IP traffic. Now when you're creating a VPC Flow Log. You get a number of options of where to attach the monitoring to. Now the options are you can attach it to a VPC which is what I'm doing here because I created it from the VPC. You can attach it to a specific subnet within a VPC, or you can attach it directly to a network interface. Now, if you attach it to a VPC what that means is the Flow Log monitors every network interface inside that VPC. If you attach it to a subnet, it monitors every network interface in that subnet. If you attach it directly to an interface, then logically it only monitors that single interface and it monitors IP data going in or out of any monitored interfaces. So if you monitor the entire VPC, then you'll see the traffic metadata going in or out being allowed or blocked from all the interfaces in that VPC. So you need to make sure that you place the VPC Flow Logs appropriately. Now, when you create a VPC Flow Log, you're able to specify a filter. So this determines what metadata you see. Do you only want to see metadata for traffic that's accepted? So that is allowed through NACLS and security groups? Do you only want to show traffic that rejected, or do you want to see all traffic? Now I'd recommend always selecting all because it's not the type of data that's going to consume any capacity, or is going to cost a huge amount of money, and you're always going to get access to any relevant data. You can always prune things later, but you can't get access to data after the fact. So if you've been exploited and you're using Flow Logs to determine what's happened, you'll never be in a position where you can get access to any data that historically you didn't monitor. So when in doubt, select all for this filter. Now, in terms of the destination, you can either store VPC Flow Logs, to an S3 bucket, or you can store it in CloudWatch logs. Now, in either of these two cases, you're going to need permissions. So Flow Logs uses IAM roles to give it the permissions it needs to write to this destination. So in this case, I'm going to configure it to store to CloudWatch logs and I'll need to make sure that the IAM role that I use has permissions to write into CloudWatch logs. If you do S3, you need to make sure it has the permissions for S3. So for destination log group, I'm going to go ahead and use flow. I'm going to need to create a new IAM role so I'm going to click on set up permissions. So this will take a few seconds. I'll need to create a new IAM role. It'll automatically name it. If I expand view policy document, you'll recognize these permissions. It's just being given create log group, create log stream, put log events, and then the ability to describe both log groups and log stream. So I'm going to go ahead and create this flow log role. Once I've done that, I can go back to this screen and hit refresh and then in this dropdown, I'll type flow, locate the flow logs role, select it, and hit create. So that's Flow Logs configured. It's been configured at the VPC level, and once it's set up, it'll start monitoring traffic and start storing all that data into CloudWatch logs. Now if I go to VPC, and I select the Flow Logs tab. I'll be shown any Flow Logs that are configured at the VPC level. So this is now active. It confirms the creation time. It shows me the IAM role that's being used for the permissions and it gives me the destination log group, and I can click on that to go to the CloudWatch logs group but I also want to demonstrate is if I go to subnets and select one of the subnets that's inside this VPC and then select the Flow Logs tab, I'll also see the Flow Logs. Now the reason for this is because if you create a Flow Log at a certain level so at the VPC or the subnet, it also captures anything further down the hierarchy. So if I create a Flow Log at the VPC level. It's automatically applied to every subnet and automatically applied again to every interface in that VPC. If I create it at the subnet level it's automatically applied to any interfaces. So it flows down this hierarchy, but it does not flow up. So if I create Flow Logs at the interface level, it will only monitor that interface. It won't monitor anything else in the subnet, and it won't monitor anything else in the VPC. Now to get some data inside this Flow Log what I'm going to do is I'm going to move across to the EC2 console. I'm just going to create a number of EC2 instances, So I want to go to running instances and I could either launch an instance manually. So set everything step by step but I also have is the launch template that I used in a previous lessons. I'm going to select that launch template. I'm going to select that launch template select version one. I'm going to set it to launch three instances. I'll accept the rest of the defaults, and then I'll launch those instances from the template. So this is going to launch three individual EC2 instances. So if I go to instances, I'll see they're all launching in availability zone 1a. Now, once they're loading, they'll start generating traffic naturally, so things will be trying to contact them. They'll all be set up with public IP addresses. They'll all be running the cat web website. So they'll be showing pictures of my cat Winky and ideally, because that publicly accessible, we'll find that they're getting scammed by various web scraping robots and potentially people looking to try and exploit them so automatically, without doing anything special, we should find that we're getting some accesses to them but just to ensure that we do I'm going to go to the DNS address of just one of these instances and open it and in a new tab and there we go. It's a picture of my cat Winky. So at least we've got some traffic that's gone into this particular instance. Now that we've got that I'm going to go to services and go to CloudWatch. Remember, I configured the VPC Flow Logs to inject the data into CloudWatch logs. So I go to logs. We've got the log group that I set up called Flow. So I'll select that, and so far it's got two log streams for two of the network interfaces for these EC2 instances. So I'm going to go to services, go to EC2, and just open that in a new tab, go to running instances and it was the BF 43 instance that I browsed to a second ago. So what I want to look for is what the network interface ID is for this instance. So I'm going to select eth0, and the network interface ends 3F32 so back to CloudWatch. I'll just refresh and I'm looking for 3F32 So this is the log stream that represents the network interface for this EC2 instance, that I just browsed to. So I'm going to open up this log stream and we've already got some metadata for the traffic that's being generated from this network interface. Now, just to reiterate Flow Logs only captures traffic metadata, and it only captures this information. So if I expand a particular entry, it captures the account ID. It captures the interface ID. So that's this listed. Then we've got the source IP and the destination IP, the source port, the destination port, the protocol, number of packets, the bytes, the start, the end, whether the traffic is allowed or rejected, and then just this status field that lets us know that Flow Logs is okay. I want to show you this not because you necessarily need to remember it but what you do need to remember is that Flow Logs is not monitoring actual traffic data. It is just metadata about the IP traffic that's inside the VPC. In this particular case it's the VPC because I created the Flow Logs at the VPC level and the way that Flow Log structured it's logging is that it's got one central log group. Inside the log group is a log stream for each network interface that's monitored. So in this case, it will be every interface inside the VPC. If I had selected the subnet it would be every interface in the subnet. If I'd have selected a specific interface, it would be only that one single interface and then inside the log stream, we've got these particular log events.

Now for the exam **Flow Logs are not real time**. So it takes some time to enable and once you've enabled it, there is a delay between the traffic itself occurring and when this metadata is logged inside the log group or delivered to S3. So you can't use this for any real time analytics. That's really important to understand for the exam and what's also important for the exam is that Flow Logs don't capture all types of traffic. So some examples of what Flow Logs don't capture is traffic to the Amazon DNS server, Windows license activation if you've got any Windows EC2 instances where the licenses are managed by AWS it does not capture that traffic. It doesn't capture any traffic to the 169.254.169.254 address, which is the instance metadata that occurs inside the instance. It does not go through the normal network interface of the instance. It doesn't capture any information about DHCP traffic or any information directly to the VPC router, so they're really important facts to remember for the exam. **It's not real time. It doesn't log actual traffic, the Flow Log collection points of VPC, subnet, or interface and then specifically the traffic that does not get logged.**