CLAIMS

What is claimed is:

1. A method of a network element functioning as a controller for a split architecture network, the method for assigning flow monitoring tasks to a set of switches in the network to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, each of the switches to sample flows as directed by the controller, wherein each of the switches samples the flows by monitoring a subset of the flows and recording measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the method comprising the steps of:

receiving, by the controller, an input including an estimated traffic volume for each of the paths defined by each of the OD pairs;

calculating, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches;

receiving a request to set up a new flow to traverse one of the paths in the network; and

assigning the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint.

2. The method of claim 1, wherein d_{ij} represents a sampling fraction for one of the switches (S_j) ($j \in [1,K]$) and one of the paths (P_i) ($i \in [1,N]$), the step of calculating

further comprises the step of calculating the sampling fractions that maximize $\Sigma_i (f_i x T_i)$, wherein $f_i=\Sigma^K_{j=1} d_{ij}$ is the total fraction of flows on the path (P_i) to be sampled, T_i is the number of flows traversing on the path (P_i).

3. The method of claim 2, wherein the step of calculating further comprises the step of calculating the sampling fractions subject to the bandwidth constraint ($\forall j$, $c\Sigma_i (T_i x d_{ij}) < B_j$) and the memory constraint of ($\forall j$, $c\Sigma_i (T_i x d_{ij}) < M_j$), wherein c is a constant that represents a size of a flow record generated by sampling the flows.

4. The method of claim 1, wherein the step of assigning further comprises the steps of:

identifying whether the new flow can be assigned to a first switch on the one of the paths (P_i) within the sampling fraction for the first switch and the path (P_i), the bandwidth constraint and the memory constraint; and

if the first switch cannot accommodate the new flow, assigning the new flow to a second switch on the path (P_i) within the sampling fraction for the second switch and the path (P_i), the bandwidth constraint and the memory constraint.

5. The method of claim 1, wherein the network comprises a plurality of controllers that direct the switches to forward and sample the flows, the method further comprises the steps of:

hashing an identifier of the new flow to produce a hash value;

assigning, by the controller, the new flow to one of the switches that are located on the one of the paths if the hash value is within a predetermined hash range associated with the controller; and

updating routing information stored in the controller to keep track of assigned flows for each of the switches.

6. The method of claim 1, wherein receiving the estimated traffic volume further comprises the step of:

receiving a traffic matrix that contains a set of entries, wherein rows of the traffic matrix correspond to ingress ones of the switches, columns of the traffic matrix correspond to egress ones of the switches, and each of the

entries in the traffic matrix contains a number representing the number of flows that traverse on a path defined by a corresponding ingress switch in a same row and a corresponding egress switch in a same column.

7. A system of a split architecture network for assigning flow monitoring tasks to a set of switches in the network, the system to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, wherein each of the switches samples the flows by monitoring a subset of the flows and recording measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the system comprising:

a controller to direct forwarding and sampling of the flows; and

the set of switches coupled to the controller via the network,

wherein the controller comprises:

memory to store an estimated traffic volume for each of the paths defined by each of the OD pairs;

a processor coupled to the memory, the processor adapted to calculate, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches; and

a network interface to receive a request to set up a new flow to traverse one of the paths in the network, wherein the request causes the processor to assign the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction

for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint.

8. The system of claim 7, wherein d_{ij} represents a sampling fraction for one of the switches (S_j) ($j \in [1,K]$) and one of the paths (P_i) ($i \in [1,N]$), wherein the process is adapted to calculate the sampling fractions that maximize $\Sigma_i (f_i x T_i)$, wherein $f_i=\Sigma^K_{j-1} d_{ij}$ is the total fraction of flows on the path (P_i) to be sampled, T_i is the number of flows traversing on the path (P_i).

9. The system of claim 8, wherein the sampling fractions (d_{ij}) are calculated subject to the bandwidth constraint ($\forall j, c\Sigma_i (T_i x d_{ij}) < B_j$) and the memory constraint of ($\forall j, c\Sigma_i (T_i x d_{ij}) < M_j$), wherein c is a constant that represents a size of a flow record generated by sampling the flows.

10. The system of claim 7, wherein, when assigning the new flow, the processor is adapted to:

identify whether the new flow can be assigned to a first switch on the one of the paths (P_i) within the sampling fraction for the first switch and the path (P_i), the bandwidth constraint and the memory constraint; and

if the first switch cannot accommodate the new flow, assign the new flow to a second switch on the path (P_i) within the sampling fraction for the second switch and the path (P_i), the bandwidth constraint and the memory constraint.

11. The system of claim 7, further comprising a plurality of controllers that direct the switches to forward and sample the flows, wherein the processor of the controller is adapted to:

hash an identifier of the new flow to produce a hash value;

assign the new flow to one of the switches that are located on the one of the paths if the hash value is within a predetermined hash range associated with the controller; and

update routing information stored in the controller to keep track of assigned flows for each of the switches.

12. The system of claim 7, wherein the estimated traffic volume stored is a traffic matrix that contains a set of entries, wherein rows of the traffic matrix correspond to ingress ones of the switches, columns of the traffic matrix correspond to egress ones of the switches, and each of the entries in the traffic matrix contains a number representing the number of flows that traverse on a path defined by a corresponding ingress switch in a same row and a corresponding egress switch in a same column.

13. A network element functioning as a controller for a split architecture network, the network element for assigning flow monitoring tasks to a set of switches in the network to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, each of the switches to sample flows as directed by the network element, wherein each of the switches samples the flows by monitoring a subset of the flows and recording measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the network element comprising:

memory to store an estimated traffic volume for each of the paths defined by each of the OD pairs;

a processor coupled to the memory, the processor adapted to calculate, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches; and

a network interface to receive a request to set up a new flow to traverse one of the paths in the network, wherein the request causes the processor to assign the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint.

14. The network element of claim 13, wherein d_{ij} represents a sampling fraction for one of the switches (S_j) ($j \in [1,K]$) and one of the paths (P_i) ($i \in [1,N]$), wherein the process is adapted to calculate the sampling fractions that maximize $\Sigma_i (f_i x T_i)$, wherein $f_i = \Sigma^K_{j=1} d_{ij}$ is the total fraction of flows on the path (P_i) to be sampled, T_i is the number of flows traversing on the path (P_i).

15. The network element of claim 14, wherein the sampling fractions (d_{ij}) are calculated subject to the bandwidth constraint ($\forall j$, $c\Sigma_i (T_i x d_{ij}) < B_j$) and the memory constraint of ($\forall j$, $c\Sigma_i (T_i x d_{ij}) < M_j$), wherein c is a constant that represents a size of a flow record generated by sampling the flows.

16. The network element of claim 13, wherein, when assigning the new flow, the processor is adapted to:

identify whether the new flow can be assigned to a first switch on the one of the paths (P_i) within the sampling fraction for the first switch and the path (P_i), the bandwidth constraint and the memory constraint; and

if the first switch cannot accommodate the new flow, assign the new flow to a second switch on the path (P_i) within the sampling fraction for the second switch and the path (P_i), the bandwidth constraint and the memory constraint.

17. The network element of claim 13, wherein the controller is coupled via the network to a plurality of controllers that direct the switches to forward and sample the flows, wherein the processor of the controller is adapted to:

hash an identifier of the new flow to produce a hash value;

assign the new flow to one of the switches that are located on the one of the paths if the hash value is within a predetermined hash range associated with the controller; and

update routing information stored in the controller to keep track of assigned flows for each of the switches.

18. The network element of claim 13, wherein the estimated traffic volume stored is a traffic matrix that contains a set of entries, wherein rows of the traffic matrix correspond to ingress ones of the switches, columns of the traffic matrix correspond to

egress ones of the switches, and each of the entries in the traffic matrix contains a number representing the number of flows that traverse on a path defined by a corresponding ingress switch in a same row and a corresponding egress switch in a same column.

5

What is claimed is:

1. A method of a network element functioning as a controller for a split architecture network, the method for assigning flow monitoring tasks to a set of switches in the network to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, each of the switches to sample flows as directed by the controller, wherein each of the switches samples the flows by monitoring a subset of the flows and recording measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the method comprising the steps of:

receiving, by the controller, an input including an estimated traffic volume for each of the paths defined by each of the OD pairs;

calculating, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches;

receiving a request to set up a new flow to traverse one of the paths in the network; and

assigning the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint.

2. The method of claim 1, wherein the step of assigning further comprises the steps of:

identifying whether the new flow can be assigned to a first switch on the one of the paths (P_i) within the sampling fraction for the first switch and the path (P_i), the bandwidth constraint and the memory constraint; and

if the first switch cannot accommodate the new flow, assigning the new flow to a second switch on the path (P_i) within the sampling fraction for the second switch and the path (P_i), the bandwidth constraint and the memory constraint.

3. The method of claim 1, wherein the network comprises a plurality of controllers that direct the switches to forward and sample the flows, the method further comprises the steps of:

hashing an identifier of the new flow to produce a hash value;

assigning, by the controller, the new flow to one of the switches that are located on the one of the paths if the hash value is within a predetermined hash range associated with the controller; and

updating routing information stored in the controller to keep track of assigned flows for each of the switches.

4. The method of claim 1, wherein receiving the estimated traffic volume further comprises the step of:

receiving a traffic matrix that contains a set of entries, wherein rows of the traffic matrix correspond to ingress ones of the switches, columns of the traffic matrix correspond to egress ones of the switches, and each of the entries in the traffic matrix contains a number representing the number of flows that traverse on a path defined by a corresponding ingress switch in a same row and a corresponding egress switch in a same column.

5. A method of a network element functioning as a controller for a split architecture network, the method for assigning flow monitoring task to a set of switches in the network to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, each of the switches to sample flows as directed by the controller, wherein each of the switches samples the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the method comprising the steps of:

receiving, by the controller, an input including an estimated traffic volume for each of the paths defined by each of the OD pairs;

calculating, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches;

receiving a request to set up a new flow to traverse one of the paths in the network; and

assigning the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint, wherein d_{ij} represents a sampling fraction for one of the switches (S_j) ($j\epsilon[1,K]$) and one of the paths (P_i) ($i\epsilon[1,N]$), the step of calculating further comprises the step of calculating the sampling fractions that maximize $\sum_i (f_i \times T_i)$, wherein $f_i=\sum_{j=1}^{k} d_{ij}$ is the total fraction of flows on the path (P_i) to be sampled, T_i is the number of flows traversing on the path (P_i).

6. The method of claim 5, wherein the step of calculating further comprises the step of calculating the sampling fractions subject to the bandwidth constraint ($\forall j$, $c\sum_i (T_i \times d_{ij}) < B_j$) and the memory constraint of ($\forall j$, $c\sum_i (T_i \times d_{ij}) < M_j$), wherein c is a constant that represents a size of a flow record generated by sampling the flows.

7. A system of a split architecture network for assigning flow monitoring tasks to a set of switches in the network, the system to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, wherein each of the switches samples the flows by monitoring a subset of the flows and recording measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the system comprising:

a controller to direct forwarding and sampling of the flows; and

the set of switches coupled to the controller via the network,

wherein the controller comprises:

memory to store an estimated traffic volume for each of the paths defined by each of the OD pairs;

a processor coupled to the memory, the processor adapted to calculate, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches; and

a network interface to receive a request to set up a new flow to traverse one of the paths in the network, wherein the request causes the processor to assign the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint.

8. The system of claim 7, wherein, when assigning the new flow, the processor is adapted to:

identify whether the new flow can be assigned to a first switch on the one of the paths (P_i) within the sampling fraction for the first switch and the path (P_i), the bandwidth constraint and the memory constraint; and

if the first switch cannot accommodate the new flow, assign the new flow to a second switch on the path (P_i) within the sampling fraction for the second switch and the path (P_i), the bandwidth constraint and the memory constraint.

9. The system of claim 7, further comprising a plurality of controllers that direct the switches to forward and sample the flows, wherein the processor of the controller is adapted to:

hash an identifier of the new flow to produce a hash value;

assign the new flow to one of the switches that are located on the one of the paths if the hash value is within a predetermined hash range associated with the controller; and

update routing information stored in the controller to keep track of assigned flows for each of the switches.

10. The system of claim 7, wherein the estimated traffic volume stored is a traffic matrix that contains a set of entries, wherein rows of the traffic matrix correspond to ingress ones of the switches, columns of the traffic matrix correspond to egress ones of the switches, and each of the entries in the traffic matrix contains a number representing the number of flows that traverse on a path defined by a corresponding ingress switch in a same row and a corresponding egress switch in a same column.

11. A system of a split architecture network for assigning flow monitoring task to a set of switches in the network, the system to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, wherein each of the switches samples the flows by monitoring a subset of the flows and recording a measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the system comprising:

a controller to direct forwarding and sampling of the flows; and

the set of switches coupled to the controller via the network,

wherein the controller comprises:

memory to store an estimated traffic volume for each of the paths defined by each of the OD pairs;

a processor coupled to the memory, the processor adapted to calculate, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches; and

a network interface to receive a request to set up a new flow to traverse one of the paths in the network, wherein the request causes the processor to assign the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint, wherein d_{ij} represents a sampling fraction for one of the switches (S_j) ($j \epsilon [1,K]$) and one of the paths (P_i) ($i \epsilon [1,N]$), wherein the process is adapted to calculate the sampling fractions that maximize $\sum_i (f_i \times T_i)$, wherein $f_i = \sum_{j=1}^{K} d_{ij}$ is the total fraction of flows on the path (P_i) to be sampled, T_i is the number of flows traversing on the path (P_i).

12. The system of claim 11, wherein the sampling fractions (d_{ij}) are calculated subject to the bandwidth constraint ($\forall j, c\sum_i (T_i x\ d_{ij}) < B_j$) and the memory constraint of ($\forall j, c\sum_i (T_i x\ d_{ij}) < M_j$), wherein c is a constant that represents a size of a flow record generated by sampling the flows.

13. A network element functioning as a controller for a split architecture network, the network element for assigning flow monitoring tasks to a set of switches in the network to optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, each of the switches to sample flows as directed by the network element, wherein each of the switches samples the flows by monitoring a subset of the flows and recording measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the network element comprising:

memory to store an estimated traffic volume for each of the paths defined by each of the OD pairs;

a processor coupled to the memory, the processor adapted to calculate, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches; and

a network interface to receive a request to set up a new flow to traverse one of the paths in the network, wherein the request causes the processor to assign the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and the one of the paths, the bandwidth constraint and the memory constraint.

14. The network element of claim 13, wherein, when assigning the new flow, the processor is adapted to:

identify whether the new flow can be assigned to a first switch on the one of the paths (P_i) within the sampling fraction for the first switch and the path (P_i), the bandwidth constraint and the memory constraint; and

if the first switch cannot accommodate the new flow, assign the new flow to a second switch on the path (P_i) within the sampling fraction for the second switch and the path (P_i), the bandwidth constraint and the memory constraint.

15. The network element of claim 13, wherein the controller is coupled via the network to a plurality of controllers that direct the switches to forward and sample the flows, wherein the processor of the controller is adapted to:

hash an identifier of the new flow to produce a hash value;

assign the new flow to one of the switches that are located on the one of the paths if the hash value is within a predetermined hash range associated with the controller; and

update routing information stored in the controller to keep track of assigned flows for each of the switches.

16. The network element of claim 13, wherein the estimated traffic volume stored is a traffic matrix that contains a set of entries, wherein rows of the traffic matrix correspond to ingress ones of the switches, columns of the traffic matrix correspond to egress ones of the switches, and each of the entries in the traffic matrix contains a number representing the number of flows that traverse on a path defined by a corresponding ingress switch in a same row and a corresponding egress switch in a same column.

17. A network element functioning as a controller for a split architecture network, the network element for assigning flow monitoring tasks to a set of switches in the network of optimize network-wide flow monitoring by maximizing the number of monitored flows and reducing overhead of the flow monitoring, each of the switches to sample flows as directed by the network element, wherein each of the switches samples the flows by monitoring a subset of the flows and recording measurements of the subset of the flows, each of the flows to traverse the network via a path defined by one of a set of (Origin, Destination) ("OD") pairs, wherein the Origin is an ingress one of the switches from which the flow enters the network and the Destination is an egress one of the switches from which the flow exits the network, the network element comprising:

memory to store an estimated traffic volume for each of the paths defined by each of the OD pairs;

a processor coupled to the memory, the processor adapted to calculate, for all of the switches and all of the paths, sampling fractions that maximize the number of the flows sampled by the switches, wherein a sampling fraction for a given switch and a given path is defined as the number of flows that pass through the given switch and the given path and are sampled by the given switch divided by a total number of the flows that pass through the given path, the calculating being based on the estimated traffic volume and subject to a bandwidth constraint and a memory constraint of the switches; and

a network interface to receive a request to set up a new flow to traverse one of the paths in the network, wherein the request causes the processor to assign the new flow to one of the switches that are located on the one of the paths, based on the sampling fraction for the assigned switch and one of the paths, the bandwidth constraint and the memory constraint, wherein d_{ij} represents a sampling fraction for one of the switches (S_j) ($j\epsilon[1,K]$) and one of the paths (P_i) ($i\epsilon[1,N]$), wherein the process is adapted to calculate the sampling fractions that maximize $\sum_i (f_i \times T_i)$, wherein $f_i = \sum_{j=1}^{K} d_{ij}$ is the total fraction of flows on the path (P_i) to be sampled, T_i is the number of flows traversing on the path (P_i).

18. The network element of claim 17, wherein the sampling fractions (d_{ij}) are calculated subject to the bandwidth constraint ($\forall j$, $c\sum_i (T_i \times d_{ij}) < B_j$) and the memory constraint of ($\forall j$,

$c\sum_i (T_i \times d_{ij}) < M_j$), wherein c is a constant that represents a size of a flow record generated by sampling the flows.